

Minimal Discourse Units: Can we define them, and why should we?

Liesbeth Degandⁱ & Anne Catherine Simonⁱⁱ

ⁱ FNRS/Université catholique de Louvain
Centre for Text and Discourse Studies
degand@lige.ucl.ac.be

ⁱⁱ Université catholique de Louvain
Valibel
simon@rom.ucl.ac.be

Abstract: In spite of its crucial role in discourse segmentation, there is no consensus in the literature on what a minimal discourse unit is and how it should be identified. We claim that the minimal discourse unit should be defined in terms of two linguistic criteria: syntax and prosody. We develop a heuristics for identifying minimal discourse units in (spoken) discourse and illustrate this with a piece of spontaneous conversation.

1. What is a Minimal Discourse Unit?

In discourse analysis it is commonly acknowledged that discourse is (hierarchically) structured (Grosz & Sidner, 1986; Hobbs, 1979; Mann & Thompson, 1988; Polanyi, 1988; Roulet et al., 2001). The assumption is that a piece of discourse is built up from smaller “building blocks” related to one another in a coherent way. What these building blocks actually look like differs according to the discourse model at stake. A number of authors avoid the problem of defining the minimal segments. Mann and Thompson (1988) content themselves with stating that the minimal units are “typically clauses”, and that the same relations can hold between larger spans of text too. Others limit themselves to a (sometimes naïve) working definition in view of (automated) discourse segmentation (for a detailed overview, see Passonneau & Litman 1997: 105-108 and den Ouden 2004: 13-22).

Most of the authors tackling explicitly the theoretical issue of defining MDUs acknowledge their bidimensional status with a linguistic face (in the form of syntactic clauses) and a textual face (in the form of contextualised information units), and that the maximal units at the micro-syntactic level (‘clauses’) do not necessarily correspond to the minimal discourse units at the macro-level (see also Hannay & Kroon 2005). In other words, there is no clear-cut boundary between the two levels (Roulet 2002). More recently, there has been a tendency to consider MDUs as being multidimensional, i.e. they are defined in terms of different criteria: the realisation of an illocutionary force, a conceptual content, syntactic dependency relations and a prosodic contour (Steen to appear). “Typical” basic discourse units would amalgamate these four dimensions, while “non typical” discourse units may lack one of these dimensions or may be otherwise non co-extensive (for a detailed account on this kind of mismatches, see e.g. Auchlin & Ferrari 1994, Simon 2001, 2004, Hannay & Kroon 2005). On the basis of Selting’s (2000) work on Turn-Constructional Units in the framework of Conversation Analysis, we share this multi-dimensional view of Minimal Discourse Units (MDU) and will consider them as being “the smallest interactionally relevant complete linguistic unit[s], in a given context, that is constructed with syntactic and prosodic resources within their semantic, pragmatic, activity-type-specific, and sequential (...) context” (477).

2. Overview of Problems in identifying MDUs

In discourse-based studies, be they linguistic, computational or psycho-linguistic, the idea that there is such a thing as a minimal discourse unit has become quite uncontroversial even if there is no real consensus on how to identify this MDU. Among the definitions used in the above mentioned models, we identify three major problems.

2.1. Problem 1: MDUs sit between coherence relations

Most models assume that minimal discourse segments are (hierarchically) linked together by coherence relations. Thus, identifying the coherence relations would also automatically bring up the minimal discourse units. However, it remains far from clear “what may or may not be admitted as a coherence relation” (Bateman & Rondhuis, 1997: 3). Furthermore, evaluating the adequacy of the minimal discourse segmentation in terms of the possibility for the MDUs to enter a hierarchical dependency structure (Rossari 1996) leads to circularity in defining segmentation criteria.

Another related problem arises from the tree-like representation of such discourse structures in which a position proper must be assigned to each of the MDUs, with a dependency relation to the other segments (one idea per unit). While productive in many cases, this constraint does not take into account that there may be different types of MDUs, and that they do not all fulfil the same role in the hierarchical discourse structure (see also example (1), below).

2.2. Problem 2: MDUs are mental representation updates

Understanding a piece of discourse is an incremental process in which new segments are integrated with the preceding ones to construct a coherent mental representation of the discourse content. Consequently, MDUs can be defined as devices for updating this representation (Van Dijk 1999, van Dijk & Kintsch 1983).

A number of authors take this updating function as an operational discourse segmentation principle. Roulet, for instance, defines the act (his minimal text unit) as « délimitée de part et d’autre par un passage en mémoire discursive » (Roulet 1999 : 210). This leads to a twofold problem. On the one hand, it is far from trivial to set up a list of linguistic cues that would index a given segment as having operated a mental model update (but see, e.g. Millis et al. 2000, on punctuation marks). On the other, it is possible that certain textual segments do not function as updates of the mental model under construction, but play a role only on the level of textual or interactional management.

- (1) (a) ouais bè c’est leur métier quoi
well, yeah, it’s their job, y’know
- (b) mais je veux dire ils /
but I mean they /
- (c) c’est une boîte
it’s a firm
- (d) euh je veux dire /
uhm I mean /
- (e) qu’on peut recommander quoi [BLAjv11, 71]
that one could recommend y’see

In (1), each of the listed segments is realized by an independent intonation unit, but segments (b) and (d) need not be related to the other segments by a coherence relation, they do not play a role in updating the current discourse model. Rather, they strictly concern the way in which the speaker evaluates and construes his discourse. The general problem of such models is then that they do not make explicit the link between the textual segments and the conceptual ideas they vehicle, since the mental representations are composed of concepts (ideas), not of “linguistic sentences”.

2.3. Problem 3: MDUs are static and clear-cut units

Most discourse models consider MDUs as relatively static and bounded units¹ offering non satisfactory solutions when having to define a (coherence) relation between segments (a) and

¹ Selting (1996, 1998) working in the framework of Conversation Analysis is an exception. Her work on conversational talk points at the dynamic, expandable and flexible character of MDUs.

(b) in example (2), for instance.

- (2) (a) et je me suis sentie dans un état heureux
and I felt in a happy state
(b) plus qu'heureux [Valibel, RADw11r]
more than happy

Segment (b) is a kind of “after thought expansion” of a unit initially presented as syntactically complete. These kinds of expansion are typical of spoken discourse and we would rather not consider them as truncated, thus discarding them from the analysis. The same goes for incidental clauses (Simon 2004) interrupting a unit under construction, or for “delayed self repairs” which should not be considered as misfires, but are more often “structuring devices for complex units” (Auer 2004).

3. Towards an operational definition of MDU

Considering the problems listed above a workable definition of MDUs should take into account the following points:

1. Only observable linguistic criteria can be taken into account (syntax and intonation/punctuation)². This precludes the use of the “representation update model” and the “one idea per unit” criterion to operationalize an MDU definition.
2. It cannot be assumed that coherence relations exist between all MDUs, nor that all MDUs contribute equally to the (conceptual) discourse structure. Metaphorically speaking, there are “leafs” that are not attached to the final discourse structure tree.
3. (Spoken) discourse is an incremental process, but the resulting structure is static. This means that some MDUs may be set up in the course of discourse, but recycled later on, or left unused, i.e. some leafs may fall on the ground... In addition, a number of segments may have no other function than regulating turn taking, an activity that can interfere with the conceptual one.
4. Discourse segmentation implies working on a completed product, i.e. our perspective will be one of discourse comprehension, rather than of production. MDUs, then, are the minimal segments the hearer uses when reconstructing what the speaker is saying. It is an open question whether the MDUs resulting from this segmentation procedure can also tell us something about the production process.

To cope with these issues, we will here adopt the distinction proposed by Hannay & Kroon (2005) between *strategic units* (‘acts’) and *conceptual units* (‘ideas’). The latter correspond to the conceptual content of the discourse, individualized in the form of *ideas*, i.e. “those items of information” that “can function as domains of activation cost” (Chafe 1994: 80). Thus, they are abstract entities involved in text processing (production and comprehension), but not as such materialised within the discourse. Strategic units, then, are entirely dependent on the situated discourse. They constitute the steps by which the “ideas” are planned, processed and expressed by the speakers/writers. They are linguistically realized in the form of intonation units or punctuation units, and potentially isolating major constituents of the clause, or grouping several clauses into a single intonation unit. Following Chafe (1994: 63) they come in three types: (i) *fragmentary units* are truncated, and will at this stage be disregarded from the analysis; (ii) *substantive units* convey ideas of events, states or referents; (iii) *regulatory units*, regulate the interaction of information flow; that is, the development of discourse, the interaction between participants, the expression of the speaker’s mental process, and/or the judgement of the validity of the information. According to us, only substantive units contribute to the mental representation of the ongoing discourse, whereas the

² In our view, this may include semantic notions as long as they have clear syntactic repercussions.

regulatory ones would only contribute to construing the image of the text itself. In terms of the tree metaphor, the two types of units are leaves of different trees. However, we shall not develop here how these two dimensions (mental representation of the discourse and image of the text itself) can interact.

3.1. Segmentation principles

A workable set of segmentation principles implies the use of explicit criteria, i.e. criteria based on observable, linguistic features. They should provide the discourse analyst with a two-steps procedure: (i) segmenting the flow of speech into MDUs, and (ii) distinguishing between the three types of units: fragmentary, substantive and regulatory.

Here, we focus on spoken discourse, but the procedure should apply to writing as well. As far as speech is concerned, the segmentation features result from two linguistic levels: syntax and intonation. Among others (Auchlin & Ferrari 1994, Auer 1996, Grobet 1997), Selting (1996, 1998) has indeed given evidence that, in interpreting units in conversational talk, hearers rely on the interplay of syntax and prosody for processing constructional units and managing continuation or completion of the turns for turntaking. Since segmentation cannot be realized with reference to syntax only or intonation only, the first step in our procedure consists in carrying out two independent analyses of the corpus at hand: one segmentation into syntactic units and one into intonation units.

3.1.1. Syntactic units

Syntactic units are determined making use of a dependency grammar developed for spoken French by Blanche-Benveniste and colleagues (1984, 1990; see van de Eynde & Mertens 2003 for a more recent account). Typical syntactic units are comprised of a “nucleus” (mostly verbs, but also nouns or adjectives) accompanied with all its “dependants”. Each “nucleus” may govern various kinds of dependants. Very briefly, (i) actants are specific dependants (usually, but not necessarily, compulsory) belonging to the restricted valency of the verb³, (ii) circumstants are also dependent on a verb, but out of its strict valency pattern, and (iii) adjuncts can be added to any construction, but in a “looser” way. Examples⁴ (3) - (5) illustrate the different types of dependants.

(3) tu sonnes à David Lombard_{ACT} par exemple_{ADJ} # [BLAjv11, 64]
you call David Lombard for instance

(4) de toute manière_{ADJ} j'ai pas le permis_{ACT} pour le moment_{CIRC} quoi / # [BLAjv11, 93]
in any case I do not have the licence [driver's] for the moment

(5) et Charleroi_{ACT} ça te tente pas du tout_{CIRC} # [BLAnd0, 91]
and Charleroi it does not tempt you at all

Of course, one could argue that this syntactic segmentation being based on the clearly semantic notion of valency, it is not strictly formal and thus could leave free way to annotation problems. However, according to the pronominal approach of Blanche-Benveniste and colleagues (1984), valency is defined in terms of pronominal types that can replace the complements. Thus, it is possible to operationalise the three types of dependants without having recourse to semantic criteria. The major distinction is that between elements governed by the verb (actants and circumstants) and those that are not (adjuncts or associated elements), a difference that can be described in syntactic terms (Blanche-Benveniste et al. 1984: 62-63).

1. Governed elements have an interrogative pronominal counterpart, while associated elements have not. In (4) above, the governed elements *le permis* (actant) and *pour le*

³ For sake of clarity, we restrict ourselves to verbs, but the same holds for nouns and adjectives which can also function as autonomous nuclei.

⁴ In what follows syntactic units are marked with a #, pauses with a /, intonation units are given on separate lines.

- moment* (circumstant) can be substituted by *quoi* and *quand*, respectively. Such a possibility does not exist for *de toute manière* (adjunct).
2. Governed elements can be subject to pseudo-clefting or extraction, associated elements cannot. Thus, (3a) is grammatically correct, while (3b) is not.
 3. Governed elements know no (contrastive) modality restrictions, while associated elements do (cf. 4a vs. 4b).

- (3a) C'est à David Lombard que tu sonnes par exemple.
 (3b) * C'est par exemple que tu sonnes à David Lombard.
 (4a) de toute manière j'ai pas le permis mais la licence provisoire
 (4b) *de toute manière mais pas de toute façon j'ai pas le permis

Typical oral discourse markers, such as *tu vois*, *ben*, *quoi*, *j'veux dire*, also belong to the category of adjuncts, since they are not governed by the verb.

- (6) elle est complètement vraie quoi tu vois / # je suis sûr tu vois / # [BLAjv11, 781]
 she's really it y'know you see / I'm sure you see /

Once completed, this syntactic analysis provides us with a segmentation of the verbal chain into nuclei and adjuncts. Each nucleus or adjunct is a potential candidate for constituting an MDU, but its effective status, as MDU, as fragment of MDU or as grouping of more than one MDU, depends on the intonation contour of the chunk.

3.1.2. Intonation units

Intonation units are identified in terms of major intonation boundaries (a sequence of syllables grouped by a final accent carrying a high tone projecting “more to come” or a low tone signalling finality), which are usually signalled by the presence of silent pause (Mertens 1993). “Minor continuity” is considered as an internal boundary in the intonation unit, as is the stretching of sounds (intra-MDU hesitation). We do not take “focus accents” into consideration for this segmentation purpose.

Once completed, this intonation analysis provides us with a second segmentation of the verbal chain into “intonation groups”. The interest now lies in comparing the two segmentations (which are far from being completely equivalent) and it is this interaction that makes it possible to decide which segment may be considered an MDU, and which may not.

3.1.3. Confronting syntactic units and intonation units

Although there may be one-to-one correspondences between the syntactic units and the intonation units, there are very often mismatches. In (7), one “simple” nucleus (verb + actant) is realised within one IU (co-extension), whilst the same structure is realised in two separate IUs in (8).

- (7) mais j'ai des heures gratuites_{ACT} euh/ # [BLAjv11, 146]
 but I have free hours uhm

- (8) elle a besoin /
 she needs
d'un homme # [BLAjv11, 64-65]
 a man

Our working hypothesis is that an MDU amalgamates one intonation unit with one syntactic unit (7). This corresponds to Chafe's (1994:65) observation that “[m]any substantive IU have the grammatical form of single clauses”. Hannay & Kroon (2005) take a similar position to define their *discourse acts*. We believe however, in line with Selting (2000), that the syntactic (dependency) units should also play their role when segmenting the discourse into minimal

discourse units. While it seems plausible that the ‘one IU – one MDU’ position is justified in a perspective of discourse production, where every intonation unit constitutes a step in the discourse planning process, spoken discourse understanding is however not strictly intonation based but recourse to syntactic patterning too (see e.g., Monschau, Kreyer, & Mukherjee, 2003; Cutler, Dahan, & Van Donselaar, 1997). Thus, the minimal discourse units cannot be determined with reference to intonation alone. “As long as syntax, in yet unfinished units, has projected a continuation that is continued and fulfilled in the following constructions, then syntax overrides prosody for the interpretation of the units.” (Selting, 2000: 203-504). Therefore, when one syntactic unit is broken down into two or more intonation units, the MDU will amalgamate several intonation units with one syntactic unit. Example (9) is composed of two intonation units, the first of which projects a syntactic continuation (relative clause) on the second one, forming together a single MDU. In (10) five intonation units are grouped within one syntactic unit. It is the entire (complex) syntactic clause which should be interpreted as a single MDU.

- (9) mais elle disait que
but she said that
tu es quelqu’un de très chouette # [BLAjv11, 20-21]
you are someone very nice
- (10) à cause de de / cons qui ont abusé
because of of / idiots who have abused
ben moi
well I
je ne peux plus euh (silence)
I can’t anymore uhm (silence)
fumer mon stick euh / pénard
smoke my stick uhm / cool
après une reunion de staff ou quoi / # [BLAjv11, 790-796]
after a staff meeting or else

The third possibility is one in which two or more syntactic units are comprised within one intonation unit. A distinction should be made between (11) and (12).

- (11) c’est un bon pilote Oli # c’est un fou /
he’s a good driver Oli # he’s crazy /
- (12) de temps en temps elle m’appelle # parce qu’elle a / elle a besoin / d’un
homme# [BLAjv11, 35]
once in a while she calls me because she needs a man

In (11), two syntactically complete units are grouped sequentially without any syntactic projection between them. The two syntactic nodes express two ideas that are here condensed into one MDU, that is, as if it were one single idea. In (12) the causal clause is prosodically integrated with the preceding main clause and presented as a syntactically dependent circumstance of the verb *appeler* (‘call’). We would therefore claim that we have here one single MDU. This should be contrasted with (13), where the causal clause is added later in a new intonation unit. Although there is also a syntactic link between main and causal clause, we would like to argue that they are presented in different MDUs because the causal clause has not been projected previously and because of the presence of the discourse marker *quoi* (‘you know’) that leaves place to the hearer to react if she wants to (see Selting, 2000:497-498, for a similar view). Thus, (12) and (13) show how the clauses of a complex sentence can be prosodically displayed in different ways.

- (13) ben ça m’a fait un choc quoi / #
well it was a choc you know

parce que c'est mon plus jeune frère quoi (silence) [BLAjv11, 330]
because it is my youngest brother you know

3.2. Different types of MDUs

The last step in the MDU-identifying procedure consists in determining which MDUs are substantive and which ones are regulatory. From a theoretical point of view, this places our endeavour in the line of Chafe (1994) and Hannay & Kroon (2005) in assuming that all MDUs do not contribute in the same way to the construal of the mental discourse representation.

According to Chafe an IU is substantive as soon as it contains at least a noun, verb, or adjective. Hannay & Kroon define them as “usually hav[ing] a predicational structure and [being] characterized by pragmatic autonomy in the form of a topic-focus structure” (p. 97). Regulatory IUs regulate the interaction of information flow: development of discourse, interaction between participants, expression of the speaker’s mental process, and judgement of the validity of the information. Chafe, as well as Hannay & Kroon, list a number of examples of substantive and regulatory IUs/acts, they do however not provide a list of linguistic surface cues that can be used to identify one or the other unit in actual discourse analysis. We hypothesize that substantive MDUs correspond to MDUs conveying single nuclei (actants with or without circumstants) and that regulatory MDUs correspond to MDUs containing adjuncts.

4. Analysis of a spontaneous conversation

To put the segmentation procedure to a test we started analysing a piece of spontaneous conversation in French involving two acquainted speakers. The discourse consisted of approx. 7800 words, and 360 turns. The syntactic analysis, performed in absence of sound according to the heuristics described above, resulted in 809 syntactic units and 754 adjuncts. Among these adjuncts we can distinguish three main types: (i) connectives making explicit the development of the discourse such as *mais* (‘but’), *parce que* (‘because’), and *donc* (‘so’) amounting together to 149 occurrences (nearly 20%), *mais* accounting for 114 of those; (ii) ‘conversational markers’, such as *quoi* (188) (‘y’know, I mean’), *je veux dire* (22) (‘I mean’), *enfin* (56) (‘well, I mean’), *tu vois* (46) (‘you see’), *ben* (51) (‘well’), etc. or combinations like *enfin voilà quoi* (‘well that’s it you know’). Here the primary function is to regulate the interaction between the participants and to express the speaker’s mental process. They account for the majority of adjuncts in the analysed piece of discourse with 590 cases (more than 78%). A special case is the use of *oui/ouais* (40/80) (‘yes/yeah’) and *non* (72) (‘no’). First of all, they very often come in combination. Therefore they should probably not be counted as individual adjuncts, as in (14) and *ouais ouais* in (15). Furthermore, they seem to fulfil rather distinct functions in the discourse. They can serve as a response to validate or invalidate the interlocutors prior intervention. In this case the adjunct appears mostly in prosodic isolation (15). According to Blanche-Benveniste et al. (1994: 104), *oui/non* should not be viewed as adjuncts in those cases but as non-verbal nodes. *Non* can also be used to call for validation from the hearer (15), in which case it always appears at the end of the clause, prosodically integrated within it. They also seem to function as a kind of trace of the speaker’s processing difficulties, a kind of filled pauses indicating idea planning difficulties (14) (Chafe, 1980). A further noteworthy use is the clause initial *non* (15), which is prosodically integrated, and seems to have a kind of discourse developmental role, a way of (re)introducing a topic into the ongoing discourse, functioning as a turn-holder (Stenström, 1994).

(14) mais oui non mais je voulais pas tu vois / # [BLAjv11, 736]
but yes no but I didn’t want to you see

- (15) BLAnd0 : et sur la route
 and on the road
 il fait peur non #
 he is frightening, isn't he
- BLAjv1 : ouais ouais #
 yeah yeah
non il fait peur #
 no he is frightening
 il est il a des fois euh /
 he is he has sometimes uhm /
non moi je l'engueule quoi #
 no me I shout at him you know

Finally, we have (iii) the attitude and validity markers such as *franchement* (3) ('frankly'), *en tout cas* (1) ('in any case'), *par exemple* (3) ('for example'), etc. whose main function is to evaluate the validity of the information. With only 15 occurrences, they only account for approximately 2% of the adjuncts.

The prosodic analysis of the full-length conversation sample was performed. Taking into account only major intonation boundaries, 1074 intonation units were identified. The confrontation of the intonation units with the syntactic units following the rules described above yielded 961 MDUs. Among those, approx. 74% (715) are substantive and 26% (246) are regulatory. Most of the substantive MDUs consist of a simple nucleus, i.e. a dependency relation without circumstant nor adjunct (64%) or a nucleus with a circumstant or adjunct (20%). It seems thus that in spontaneous conversation, 84% of the MDUs demonstrate a one-to-one correspondence between a syntactic unit and an intonation unit. Another 8% accounts for those MDUs in which a syntactic unit is broken down in two or more intonation units, and a final 7% corresponds to MDUs which condense more than one syntactic unit into one intonation unit.

With respect to our hypothesis that substantive MDUs convey single nuclei and that regulatory MDUs contain adjuncts, we must admit that the issue is far from straightforward. Adjuncts may be grouped within the construction, or they may constitute a separate intonation unit, depending on the way speakers "package" them. In (16) the adjunct is realised within the same IU as the nucleus, it is not in (17).

- (16) franchement_{ADI} je vois ma grand-mère elle
 frankly I see my grand mother she
- (17) mais il s'est quand même vachement calmé tu sai_{ADJ}
 but he's really calmed down you know
franchement_{ADJ} euh /
 frankly uhm /

Actually most of the adjuncts appear to not have their own intonation contour. Are they nevertheless regulatory or do they belong to the substantive MDU? In other words, are the only regulatory MDUs those who are prosodically independent? Another issue is whether sub-clausal acts (like 18a) constitute substantive or regulatory units.

- (18) a. et conduire
 b. tu aimes ça

The verb-adjunct (in 18) does not by itself add new content to the mental representation under construction: it rather functions as a reminder of a topic activated earlier in the conversation. On the other hand, (18a) is syntactically dependent on (18b) through left-dislocation, thus pleading for a single MDU interpretation. Such units are often labelled "subsidiary", but strict criteria are still needed to distinguish them from main acts – if necessary. Further analyses of additional material shall hopefully provide us with more conclusive answers.

Acknowledgements

The first author is a research associate at the Belgian Fund for Scientific Research (FNRS). This work was supported by grant F.R.F.C. n° 2.4535.02 and by grant n° ARC 03/08-301 from the *Communauté française de Belgique*.

References

- Auer, P. (1996). On the Prosody and Syntax of Turn-continuations. In: E. Couper-Kuhlen & M. Selting (eds), *Prosody in conversation : interactional studies*. Cambridge : Cambridge University Press, pp. 57-100.
- Auer, P. (2004). Delayed self-repairs as a structuring device for complex turns in conversation, *LiSt - Interaction and Linguistic Structures*, No. 40. <http://www.uni-potsdam.de/u/inlist/issues/40/index.htm>
- Bateman J. & Rondhuis K. (1997). Coherence relations: Towards a general specification. *Discourse Processes*, 27: 3-49.
- Blanche-Benveniste, C., Deulofeu, J., Stefanini, J. & van den Eynde, K. (1984). *Pronom et syntaxe. L'approche pronominale et son application au français*. Paris : SELAF.
- Blanche-Benveniste, C., Bilger, M., Rouget, Chr. & van den Eynde, K. (1990). *Le français parlé : études grammaticales*. Paris : Éditions du CNRS.
- Chafe, W. (1980). *The Pear Stories: Cognitive, cultural and linguistic aspects of narrative production*. Norwood, NJ: Ablex.
- Chafe, W. (1994). *Discourse, consciousness and time*. Chicago : The University of Chicago Press.
- Cutler, A., Dahan, D., & Van Donselaar, W.A. (1997). Prosody in the comprehension of spoken language: A literature review. *Language and Speech*, 40: 141-202.
- Den Ouden, H. (2004). *Prosodic realizations of text structure*. PhD dissertation. Tilburg: University of Tilburg.
- Grobet, A. (1997). La ponctuation prosodique dans les dimensions périodiques et informationnelles du discours. *Cahiers de linguistique française*, 19: 83-123.
- Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12: 175-204.
- Hannay, M. & Kroon, C. (2005). Acts and the relationship between discourse and grammar. *Functions of Language*, 12: 87-124.
- Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8: 243-281.
- Mertens, P. (1993). *Intonational grouping, boundaries, and syntactic structure in French*. In: D. House & P. Touati (ed). *Proceedings of the ESCA Workshop on Prosody*, September 27-29, 1993, Lund (S) *Working Papers 41* (Lund Univ., Dept. of Linguistics), pp. 156-159.
- Millis, K., King, A., & Hyung-Jeong, K. (2000). Updating Situation Models From Descriptive Texts: A Test of the Situational Operator Model. *Discourse Processes*, 30: 201-236.
- Monschau, J., Kreyer, R., & Mukherjee, J. (2003). Syntax and semantics at tone unit boundaries. *Anglia*, 121: 581-609.
- Passoneau, R. J. & Litman, D. J. (1997). Discourse Segmentation by Human and Automated Means. *Computational Linguistics*, 23: 103-139.
- Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, 12 : 601-638.
- Rossari, C. (1996). Identification d'unités discursives : les actes et les connecteurs. *Cahiers de linguistique française*, 18 : 157-177.
- Roulet, E. (1999). *La description de l'organisation du discours. Des dialogues oraux aux textes écrits*. Paris : Didier.

- Roulet, E. (2002). Le problème de la définition des unités à la frontière entre le syntaxique et le textuel. *Verbum*, XXIV/1-2 : 161-178.
- Roulet, E., Filliettaz, L. & Grobet, A. in collaboration with Burger, M. (2001). *Un modèle et un instrument d'analyse de l'organisation du discours*. Bern : Peter Lang.
- Selting, M. (1996). On the interplay of syntax and prosody in the construction of turn-constructional units and turns in conversation. *Pragmatics*, 6: 371-388.
- Selting, M. (2000). The Construction of Units in Conversational Talk. *Language in Society* 29 (4), 477-517.
- Simon, A-C. (2004). *La structuration prosodique du discours en français*. Bern : Peter Lang.
- Stenström, A-B. 1994. *An introduction to spoken language interaction*. London : Longman.
- Van den Eynde, K. & Mertens, P. (2003). La valence: l'approche pronominale et son application au lexique verbal. *Journal of French Language Studies*, 13: 63-104.
- Van Dijk, T. A. (1999). Context models in discourse processing . In: H. van Oostendorp & S.R. Goldman (Eds.) *The construction of mental representations during reading*. Mahwah, NJ. USA: Lawrence Erlbaum Associates. pp.123-148.
- Van Dijk, T.A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New-York: Academic Press.