Approaches to the Textual Theme
Renata Pipalová
Charles University, Czech Republic

Abstract
Many linguists maintain that the entity unifying macrotextual units as well as all texts and inducing their coherence is their Theme (Textual Theme, Hypertheme, Discourse Topic, Macrotheme and related concepts). This function, to most treatments superimposed to utterance themes, may be broadly described as the pragmatic aboutness of texts. Surveying diverse approaches to the Textual Theme, the paper discusses its various aspects, functions, layers and dimensions.

Keywords
Corpus linguistics, English for Academic Purposes, Hedge expressions, Stance, Vagueness in language, Academic writing

Many linguists maintain that the entity unifying macrotextual units (such as paragraphs) as well as all texts, ensuring their integrity or coherence, and, among others, facilitating interaction, is their Theme (also referred to as Textual or Discourse Theme, Textual or Discourse Topic / Macrotheme / Hypertheme / and related concepts), see, e.g., Mathesius (1942/1982), Miko (1973), Hubacek (1987), Uhlířová (1987), Grepl et al. (1995), Daněš (1994; 1995), Cmejloková et al. (1999), van Dijk (1977; 1981; 1985), Brown and Yule (1983), Giora (1985; 1998), Sperber and Wilson (1986), Martin (1992), Mislíková (2003), Tomlin, Forrest, Pu, Kim (1997), Hearst (1997), Martin and Rose (2003), and Dubois (1997). For instance, on its significance Brown and Yule (1983: 73-4) remark: "The notion is attractive because it seems to be the central organizing principle for a lot of discourse. It may enable the analyst to explain why several sentences or utterances should be considered together as a set of some kind, separate from another set. It might also provide a means of distinguishing fragments of discourse which are felt to be good, coherent, examples from those that are, intuitively, incoherent concatenations of sentences." Or, as Giora (1998: 83) maintains, "in a well-organized text, each utterance within a given paragraph is interpretable in relation to the paragraph topic."

There is, however, much disagreement as to what to understand by the concept. Etymologically, it denotes something laid down (see, e.g., Webster's 1993). To our knowledge, the Theme has been defined as pragmatic aboutness, as a single referent, as an FSP function, as a proposition, as a topic sentence, as a main idea, as an abstract, as a single referent, as an FSP construction) "reduce the complex, detailed meaning structure of a text into a simpler, more general and abstract (higher level) meaning of a text." (ibid)

Moreover, they are recursive and "may apply again at each level of abstraction to produce even shorter abstracts. The result is a hierarchical macrostructure, consisting of several levels, each level consisting of a sequence of (macro-)propositions that 'summarize' a sequence of lower level (macro)propositions." (1985: 76) van Dijk takes the Topics for cognitive units. The propositional view has been adopted by other linguists as well (e.g., Giora 1985; Tomlin, Forrest, Pu, Kim 1997). For example, according to Giora (1985: 21) the Discourse Topic should be formulated "in terms of propositions or argument-predicate nominalizations, and not in terms of NPs alone."

Another approach suggested by van Dijk (1977; 1981; 1985), the concept of Theme is defined in propositional terms. Its concise formulation is provided in van Dijk (1985: 21): "a theme in this case is not simply a word or a single concept, but a (macro-) proposition". To achieve it, van Dijk proposes various "macro-rules, which map sequences of propositions onto sequences of (macro-) propositions." (1977: 188) The macrorules (i.e., deletion, generalization and (re-) construction) "reduce the complex, detailed meaning structure of a text into a simpler, more general and abstract (higher level) meaning of a text." (ibid)
“Those aspects of the context which are directly reflected in the text, and which need to be called upon to interpret the text, we shall refer to as activated features of context and suggest that they constitute the contextual framework within which the topic is constituted; that is the topic framework.” (ibid: 75). For them, the Topic framework represents “the area of overlap in the knowledge which has been activated and is shared by the participants at a particular point in a discourse.” (ibid: 83). Furthermore, they maintain that the Topic framework is consonant with the presupposition pool (by Venneman), in that it includes a number of discourse subjects. “Within the presupposition pool for any discourse, there is a set of discourse subjects and each discourse is, in a sense, about its discourse subjects.” (ibid: 80) By Topic Entity, on the other hand, they refer to a prominent discourse subject, a part of Discourse Topic, such as the “main character” in a novel. They discuss a number of studies which have shown that such foregrounding has a clear effect on interpretation as well as on subsequent recall. Focussing mainly on conversation, the authors further introduce the concept of the Speaker’s Topic constituted within the Topic framework, maintaining that “each participant expresses a personal topic within the general topic framework as a whole.” They note that in most conversations, “topics are not fixed beforehand, but are negotiated in the process of communication.” (ibid: 88-9).

Also exploring conversation, Downing, (2003: 120) views the Global Topic as a macro-speech act, whereas Episodic Topics are in her view “built up in the expansions which cluster round the core utterances and their responses.”

The authors surveying mainly monological texts, Čmejrková, Daněš and Švtála (1999: 105) nevertheless stress that the Theme is a textual function deliberately selected by the author as relevant: “Theme is thus understood as a specific textual function which the author assigns to elements of content selected by him/her, which s/he takes to be significant, relevant from the point of view of the construction of the sense of the text. This way the author organizes the whole thought content of his/her message.” (translated by R.P.) Other authors see it as a content and strategic starting point of communication (e.g., Kofensky et al. 1987).

In treatments presumably inspired by the rhetorical tradition, the Theme is occasionally co-extensive with the topic sentence (e.g., Martin 1992: Martin and Rose 2003). For instance, Martin (1992: 437) proposes a hierarchy of Themes, viz. macro-Theme, hyper-Theme and Theme. The first is correlated with the whole text, the second with the paragraph, and the last one with a clause: “A hyper-Theme is an introductory sentence or group of sentences which is established to predict a particular pattern of interaction among strings, clauses and Theme selection in following sentences...the term macro-Theme can be defined as a sentence or group of sentences (possibly a paragraph) which predicts a set of hyper-Themes; this is the introductory paragraph of school rhetoric.”

In Martin and Rose (2003: 182) this concept is further elaborated to form a dichotomy: “While hyperThemes predict what each phase of discourse will be about, new information accumulates in each clause as the phase unfolds. In written texts in particular, this accumulation of new information is often distilled in a final sentence that thus functions as a hyperNew to the phase. HyperThemes tell us where we’re going in a phase; hyperNews tell us where we’ve been...As a general rule, writing looks forward more often than it looks back. So hyperThemes are more common than hyperNews; there’s more ‘prospect’ than ‘retrospect’.” (ibid: 182) Further on, the co-authors note: “The Theme gives us orientation as to what is to come: our frame of reference as it were. Beyond this, the hyperTheme is predictive; it establishes expectations about how the text will unfold.” (ibid: 181). It should be noted that a related approach is adopted by Dubois (1997).

Some authors consider it a salient abstract idea (e.g., Baldick 2001; Cuddon 1999). For instance, Baldick (1990/2001) maintains that the Topic is “a salient abstract idea that emerges from a literary work’s treatment of its subject matter or a topic recurring in a number of literary works.” According to Wales (1994: 462), in literary criticism the Theme “is the ‘point’ of a literary work, its central idea, which we INFER from our INTERPRETATION of the PLOT, IMAGERY and SYMBOLISM, etc.” In a similar vein, Cuddon (1999: 913) defines the Theme as follows: “Properly speaking, the theme of a work is not its subject but rather its central idea, which may be stated directly or indirectly. For example, the theme of Othello is jealousy.”

In some treatments, especially the lower-rank Themes are related to more than one (see, e.g., Cervenka 1992; Quinn 2000). For example, for Quinn (2000: 323) the Themes represent “a significant idea in a literary text, sometimes used interchangeably with MOTIF. Theme is also used to describe a recurring idea in a number of texts...One problem with the varied uses of the term is the tendency to employ it as the equivalent of MORAL...”

Some authors treat the Theme with respect to foregrounding and backgrounding (Hausenblas 1969, 1971; Jones 1971; Tomlin et al. 1997). For example, Jones, L.K. (1971: 63) argues: “The essence of theme is to make something prominent at the expense of other things, which are therefore backgrounded.” Other researchers identify the Theme (Hypertheme) as a FSP function (e.g., Firbas 1995; Svoboda 1981; Adam 2002). Firbas (1995: 63) maintains that all the FSP thematic elements in a text form the thematic layer, which is in turn composed of further layers according to the individual thematic functions, viz. “those of hypertheme, theme proper and diatheme.” (ibid: 62) He distinguishes between “such constituents as appear in the thematic layer for the second time and such as have become more firmly established in it after than two occurrences. The former are referred to as themes proper, the latter as hyperthemes.” (ibid: 63) Similarly, Svoboda (1981: 129) claims that “both themes proper and diathemes may participate in constituting hyperthemes, i.e., thematic elements that several clauses have in common. Hyperthemes form strings of various lengths; composed of various thematic elements...Even a string of two elements represents in fact a hypertheme with regard to the two clauses concerned.”

Some treatments have identified the Theme with a cognitive structure. For instance Tomlin, Forrest, Pu, Kim (1997: 90) maintain that “a well defined global theme facilitates text comprehension; it functions as an advance organizer (Frase, 1975), scaffolding (Anderson et al., 1978), or anchor point (Pichert and Anderson, 1977) by evoking a mental model (representation) in the comprehender. Such a representation might be called schema (Rumelhart, 1980), frame (Minsky, 1975), script (Schlank and Abelson, 1977) or scenario (Sanford and Garrod, 1980).” According to Downing and Locke (1992: 224), Superoordinate Topics are in cognitive schemata. Discussing Thematic means as part of the structure of communication, Hausenblas (1993: 53) argues that they are “carried” by linguistic means, constitute a higher level of the semantic structure of a communication, but do not necessarily depend on
The present approach to textual theme

Taking particularly Brown and Yule (1983), Hausenblas (1969, 1971) and Daněš (1995; 1999) as our starting point, in Pipalová (2005) we suggested a three-layered approach to the Textual Theme (or Hyperthemes of various ranks, such as the Global Theme, the Chapter Theme, the Paragraph Group Theme, etc.). In what follows we shall briefly outline this treatment, modifying it somewhat to suit our present purposes. In the broadest sense, the Textual Theme involves all the elements inherently taken for granted in the particular speech event. If we adopt the framework proposed by Koňenský et al. (1987), then the broadest layer of Theme would correspond to a whole array of gradually established constituents derived from, and reflecting, the comprehensive structure of the communicative event. The latter involves the socio-psychological (sub)structure (i.e., the social, psycho-physiological and communicative features of the participants, their mutual relationships, their shared knowledge and experience, etc.), the communicative competence structure (the participants’ knowledge of the social and communicative shared norms, their shared experiential and cognitive pool, and their use of the verbal and non-verbal codes), the pragmatic structure (communicative intentions, strategies, goals, etc.), object structure (participants, present personal and non-personal objects, the communicative medium and channel, records of previous communications, etc.) and, the arguably most decisive Theme-and-content structure (i.e., the discussed personal and non-personal objects, and other content items, including metacommunicative ones).

The somewhat narrower layer (and simultaneously central layer) of the Textual Theme may be conceived of as a complex and hierarchized semantic (cognitive) structure, in monological texts selected by the author. Naturally, like the broader layer, it may, but need not, be expressed explicitly. In the latter case it stays in the background and may only be inferred. Moreover, even when it is encoded explicitly, it may never be expressed in its entirety. Indeed, from this structure the author deliberately selects elements to be encoded as utterance themes (U-themes). It is usually some of its most conspicuous, central elements that suggest it. Conversely, many of these entities may be solely presupposed and throughout the discourse will not be manifested by explicit exponents. In other words, the representation of this layer is always intentional and selective. It is never exhaustive, given the openness of texts (see van Peer 1989: 277).

We have seen that the author always selects to encode as U-themes various elements both from the broadest layer of Theme (communicative framework) as well as from the central layer (Theme-and-content structure). There are cases when he or she remains rather focused in his/her choices, and as a result, this consistency in choices assigns the item selected (and enacted as the main Thematic discourse subject) extra prominence. Therefore, in the narrowest sense, however, the content aspect of the Theme may be identified with some of the the most salient elements of the Theme-and-content structure, or, right away with its dominant, e.g., the subject of scrutiny in a scientific monograph; or a protagonist in an autobiographical novel (though itself a second-order sign). It should be noted, however, that even if a particular discourse subject remains the centre of attention throughout the discourse (through rather principled choices), it is always foregrounded against the respective background (i.e., whole semantic (cognitive) structure, the dominant of which it is taken to represent), whatever the degree of such foregrounding. Even if the background remains only implied,
cooperative participants in the communication act will activate the portions of world knowledge structures (frames, schemata, scenarios, etc.), pertaining to the dominant discourse subject and relevant for it. Presumably, the aforementioned tiers, among other things, suggest which elements constituting the complex Theme are typically foregrounded and which are not. It seems that each tier as such is incorporated in the immediately succeeding broader counterpart as its somewhat foregrounded constituent. Despite these tiers in the delimitation of the Theme, we tend to think of Theme as a complex cognitive entity which unites rather than separates, has an integrative force, lends sense to the selection and arrangement of hierarchically lower Themes, or even subsidiary Themes, motivating them.

It seems that all texts apart from athematic ones exhibit at least the first two layers of the Theme. Athemetic ones, presumably, display only the broadest layer. The centrality of the middle Theme-frame layer, presumably leads Downing and Locke (1992: 224) to the delimitation of what they call “Superordinate Topics” as cognitive schemata. Martin and Rose (2003: 181) identify them as “frames of reference.” The representation of this layer, however, may be backgrounded, whenever the choices from among its constituents are principled to such an extent that they lead to the unequivocal foregrounding of some of the Theme’s conspicuous discourse subjects, i.e., the Theme’s dominant.

Conclusion

From the foregoing survey of approaches, which by no means claims to be exhaustive, it is apparent that there is much disagreement among authors as regards Theme and that much more research will presumably be needed before the notion is fully understood. Despite the differences in point of view, most linguists contend that the notion is more abstract (e.g., Duszak 1994) or much vaguer than its utterance counterpart (e.g., Sperber and Wilson 1986).

Furthermore, whatever the approach adopted, most authors agree on its being superposed to the themes of individual utterances. Thus, we may presuppose, with Hausenblas, Daneš, and others, that as the U-themes the author selects elements which are at least to some degree relevant to the Textual Theme (or Hyperthemes of various ranks). For instance, as Uhlířová (1987: 108) argues, “the very existence of the hypertheme, against which all the sentence themes are in a particular, even if implied relation, is one of the necessary conditions for a number of consecutive sentences in a text to form a continuous, coherent text.”

Indeed, the Theme stabilizes and “grounds” the discourse and is relevant for the perception of its coherence. That is presumably why it is shown to decay from memory more slowly than other processing levels (see, e.g., Kintsch et al., 1990 cited in Brown 2005). Furthermore, unnegotiated changes in Theme tend to be identified as disturbance in coherence (see Bublitz and Lenk 1999: 166-172). Uistic markers of subjectivity can be found on any level of language description. For a quantifiable analysis, however, we need to limit the scope of our investigation to clearly identifiable markers. This means, the focus of the following study will be largely on morphosyntactic markers. Quantification comes at a cost: Valid results can only be obtained by looking at large amounts of texts as provided by corpora. Corpora provide a perspective that extends the experiential boundaries of an author or learner, thus enabling prospective text producers to learn on the example or to test via search engines whether a collocation is appropriate. A key issue here is representativeness of the corpus (e.g. Mukherjee 2005:5) as a mere study of other research papers may introduce and perpetuate in-group language which can inhibit readability and coherence considerably. As there can be no “general” corpus for an author in the academia (the sub-disciplines even within e.g. astrophysics are enormously diverse) the corpus compiler can only hope to capture some important strategies and successful conventions of producing coherent yet complex but always accessible texts.

Bibliography and references


Daneš, F., 1994. Odstavec jako centrální jednotka tematicko-kompoziční výstavby textu (na materiále textů výkladových) (Parágraf as a Central Unit of the Thematic and Compositional Build-up of the Text (On the Material of Expository Texts)). Slovo a slovesnost, 55, 1-17


Symbols and Abbreviations:
FSP – Functional Sentence Perspective
R.P. – Renata Pípalová
U-theme – utterance theme, i.e., theme on the FSP level
Theme – Theme on Textual Level