On Theme, Topic and Givenness: The state of the art*

MARÍA ÁNGELES GÓMEZ-GONZÁLEZ
Universidade de Santiago de Compostela

1. INTRODUCTION

Whereas it seems that a consensus has been reached that Theme/Topic should be described within some kind of functionalist framework, very different positions have been taken on the appropriate criteria for the definition of the notions and, as a result, with respect to their linguistic manifestation. Indeed, these notions have been associated with a myriad of terms in the relevant literature, such as (De-)Topic(alization), Theme Proper, Diatheme, Functional Sentence Perspective, Communicative Dynamism, Topic-Focus Articulation, Psychological Subject, Basis, etc., as well as with different constructions, such as zero anaphora, clitic pronouns, dislocations, definite NPs, the passive voice, preposings, cleft constructions, etc. (cf. Güblig/Raible 1977; Allerton 1978; Brômser 1982). As I see it, two factors are re-

* A preliminary version of this paper appeared in Gómez-González (1996). I am especially grateful to M.A.K. Halliday, Christopher S. Butler, Margaret Berry, Angela Downing, Peter Fries, T. Fanego Lema, T. Jiménez Juliá and Bari Samita for questions and comments that I endeavoured to bear in mind as I was writing up the paper. None of them, of course, is accountable for the way I have interpreted their recommendations. This research was funded in part by the Xunta de Galicia and by the Spanish Ministry of Education and Science DGICYT, grant numbers PB90-0370 and PB94-0619, whose support is gratefully acknowledged.

Halliday’s Systemic-Functional Grammar exemplifies the functionalist approach, which: (i) is paradigmatically-based; (ii) is influenced by ethnography and rhetoric, and posits an iconic, or non-arbitrary, relation between form and meaning; (iii) describes language on purely functional terms, so that the terms functional and function may be used in two senses: (a) as a functional label, that is, indicating a proportional, or paradigmatic, relation (or a term in such a relation or some means whereby such a relation is expressed) alongside classes or categories), and (b) as an indication of the meaning or purpose with which language is used (see Halliday 1994: xxviii); and (iv) implies a historical passage from a probabilistic preference for a particular structure, based on universal tendencies, to the development of a convention that is fully determinate. By contrast, Chomsky’s Transformational-Generative model embodies the formalist approach, which: (i) is syntagmatic in orientation; (ii) posits an arbitrary relation between form and meaning; (iii) rooted in logic and philosophy, focuses on universals and uses intuitions to elicit them; and (iv) stratifies linguistic description in terms of syntax, semantics, pragmatics and so on.

Moenia, 3 (1997), 135-155
sponsible for this state of the art: (1) different elements of communication have been taken as criterial to Theme/Topic, that is, either the message or the interactants; and (2) the multiplex nature of the notion(s) at issue, which have been related to virtually all of the established components of linguistic description. In view of the resulting variety of approaches, the author has considered it necessary to propose a rough categorisation into the three following groups, according to their underlying rationale that:

1. Theme/Topic is realized by clause initial position, which is suggested by syntactic interpretations;

2. Theme/Topic conveys Given/Known information, as hypothesised by informational approaches;

3. Theme/Topic establishes a relationship of ‘aboutness’ in a message, which is supported by semantic analyses.

These three tenets were all implicit in Mathesius’s (1929) characterisation of základ as ‘that which is known or at least obvious in the given situation and from which the speaker proceeds’, and have been subject of continuous controversy as they have gone through repeated stages of modification and reinterpretation. In the following, I will critically review the three above approaches, pointing out what are felt to be their major weaknesses. In the course of our discussion, I will have the opportunity to examine different aspects of the notions of ‘Givenness’ and ‘aboutness’, and I will also touch on some related issues as: (1) discourse staging (Hyper-/Macro-Theme/Rheme, Thematic Progression), (2) presupposition, (3) discourse saliency/relevance, and (4) thetic-categorical judgements. As a conclusion I will suggest treating Theme, Topic and Givenness as three different categories, which may, but need not coincide in one wording. In general, the use of the first term will be restricted to the notion of ‘syntactic Theme’, or clause initial position; whereas Topic and Givenness will be related to ‘aboutness’ and to the coding of different types of ‘given/known information’, respectively.

2. SYNTACTIC INTERPRETATIONS OF THEME/TOPIC

Echoing Aristotle’s notion of double judgement, syntactic interpretations conceive the informational structure of the clause as a bipartite construct consisting of: a point of departure, or syntactic Theme/Topic, and the main issue, or Rheme/Comment. Further, relying solely on morpho-syntactic structure and linking the concept ‘point of departure’ with the linear quality of language (i.e. the constraint that words must be ordered into sentences and these into texts according to some organising principle), syntactic analyses equate Theme/Topic with the leftmost, or initial, constituent of the clause as a message (prefield (Vorfeld) position). But, for some advocates of the syntactic approach, not only the leftmost, but also the rightmost NP constituent of the sentence qualify for the category of syntactic Theme/Topic (e.g. Die Frau da, sie kommt
ON THEME, TOPIC AND GIVENNESS: THE STATE OF THE ART


According to the supporters of the syntactic approach, syntactic Themes/Topics play a frame-establishing and/or perspective-marking function. Thus, Chafe (1976: 53) emblematically explains that Topic ‘limit[s] the domain of applicability of the main predication to a certain restricted domain (...) set[ting] the spatial, temporal or individual framework within which the predication holds’, while Lowe (1987: 6) suggests that this domain may expand over the ensuing discourse span, i.e. ‘any unit, usually larger than the sentence, contributing to the topic continuity or discontinuity of texts’ (cf. Chafe 1976; Magretta 1977: 132; Davison 1984, Lowe 1987; Halliday 1967, 1994; Fries 1987, 1995; Downing 1991; Vasconcellos 1992).

As I see it, syntactic approaches pose five debatable issues (cf. Lyons 1977; Pasch 1982; Hudson 1986; Huddleston 1988). The first is that syntactic analyses embody a wide range of structurally and communicatively heterogeneous constructions across different languages, which, despite their common highlighting of the clause initial and final slots, makes a unitary treatment of such structures hard to justify. Witness to this (1) below (from Gundel 1988: 224 ff. [my emphasis], except for (viii) and (ix)) (cf. e.g. Altmann 1981: 46; Davison 1984; Gundel 1985, 1988):

(1) (a) Die Frau da, sic kommt aus Berlin. (Left dislocation)
‘The woman there she come from Berlin’.

(b) My work, I’m going crazy. (Double-Subject construction, Chao 1968, Bland 1980)

(c) al sratim ka ele rina mamaš meta. (Topicalization, Berman 1980: or fronting, Y-movement)
‘Movies like that, Rina is really crazy about.’

(d) (To) Sto menja udivljaet, eto eë mudrost. (Cleft construction)
‘What surprises me, it’s her wisdom.’

(e) i. Your battery seems to be dead vs. It seems that your battery is
dead. (Subject to Subject raising)
ii. George is difficult to talk to vs. It is difficult to talk to George.
(Object to Subject raising)
iii. My soup has a fly in it vs. There is a fly in my soup. (Have-constructions)

(f) He’s shrewd, that one. (right dislocation)

(g) C’est pour ça qu’il a gagné le prix. (it-cleft)
‘It’s for that that he won the prize.’
(h) i. I saw the man. (Subject in an active declarative clauses)
ii. The man was seen. (Subject in a passive declarative clauses)
(i) Paul, you can’t do that. (Vocative).

Although numerous studies attest to the diverse communicative status of each or of some of the syntactic Theme/Topic constructions in (1), it would appear that no investigation has yet substantiated the contrast among the different thematic options within/across languages with enough quantitative and qualitative natural discourse evidence.

The second drawback of syntactic accounts lies in the fact that what constitutes the initial constituent of a clause as a message is far from being a well-settled issue. For one thing, in unplanned discourse, for example, such phenomena as repetition, hesitation, elliptical structures and so on, make it hard to identify syntactic units and hence the borderlines between their thematic (and rhematic) constituents. Further, the criteria used to identify syntactic Theme/Topics are heterogeneous and not necessarily concurrent. Indeed, scholars have invoked as markers of thematic status: (1) phonological criteria (initial tone group), (2) syntactic markers (initial syntactic constituent, either nuclear and peripheral), (3) semantic factors (first semantic role), and/or (4) metasemantic considerations (first constituent fulfilling an experiential, or referential, function). As a result, depending on the criteria cited, different scholars identify different types of constituents as Theme/Topic (and Rheme/Comment). By way of example, Halliday (1994: 56 [my emphasis]) claims that (Multiple) Theme ‘extends up to (and includes) the Topical Theme’, that is to say, the first experiential element, preceded or not by other Textual and/or Interpersonal Themes, while Lautamatti (1978) and Downing (1991) dissociate Halliday’s (Multiple) Theme from Topic, i.e. clausal topical participants (viz. Subject/Object), and Dik (1978, 1989) or Foley and Van Valin (1984, 1985) distinguish between the first peripheral slot (viz. a clause initial constituent that is not bound to the predicate) and the first nuclear slot (i.e. the initial constituent that is linked to the predicate). (cf. Enkvist 1974; Berry 1989; Vasconcellos 1992).

A third area for consideration is that while clause (internal or external) initial and final position always qualify for thematic status from a syntactic point of view, that is, they always fulfil a syntactic function, the constituents in such positions need not always have a semantic, or referential, correlate and therefore the communicative relevance of such thematic constituents is harder to justify (e.g. it in English impersonal structures such as It is raining, there in existential constructions like There is a fallacy in your argument, etc.).

A fourth shortcoming of the syntactic approach is that it seems to coalesce two qualitatively different approaches. For some of its advocates treat the Theme/Topic-Rheme/Comment pattern as a structural, or clausal, relationship, i.e. as a form of organisation that gives the clause the status of a communicative event, whereas others
take it to represent a non-structural, or textual, strategy of organisation. According to
the former, Theme/Topic is realized by clause initial position, whereas the latter prin-
ciple refers to the fact that information is presented, or staged, throughout discourse
into background, or initial position, and foreground, or final position, surpassing any
clear-cut structural, or grammatical, borderlines (cf. Grimes 1975; Halliday 1995). As
a result, Themes/Topic-Rhemes/Comment patterns are also regarded to extend across
clausal boundaries, over paragraphs or even over whole texts. Following this trend,
Anderson et al. (1977) coin the term Hypertheme (in opposition to that of Hyper-
Rheme) to refer to the introductory clause or group of clauses (headline(s), title(s),
paragraph(s), etc.), which are established to predict a pattern of Theme/Topic selec-
tion over a whole text; whereas Martin (1992) uses the label Macro-Theme (in con-
trast with the Macro-Rheme) to embrace a clause or a combination of clauses predict-
ing one or more Hyper-Theme(s) (e.g. the introductory paragraph of a text, its Table
of Contents, etc.) (cf. the discussion on Thematic Progression by Daneš 1964; Ada-

Embracing the previous four, the fifth and last debatable issue posed by the
syntactic analyses is, as remarked by Taglicht (1984), whether or not the category of
Rheme/Comment is profitable at all, as attention is centred almost exclusively on the
syntactic coding of Theme/Topic.

3. INFORMATIONAL INTERPRETATIONS OF THEME/TOPIC

Influenced by the belief that our ideas flow from given information towards
new information, informational approaches assess the category of Theme/Topic as
Given/Known information, that is, as the part(s) of a message known to the listener, or
that the speaker/writer expects the listener to know, or that can be deduced from the
c(o)n)text. In this connection, it should be borne in mind that, while the acolytes of the
informational interpretation take Theme/Topic and Given information as one and the
same category, the supporters of the syntactic approach regard Theme/Topic and
Given information as two independent, though orthogonal, choices, which tend to
coalesce, but may not do so. In Halliday’s (1967:212) words:

while ‘given’ means ‘what you were talking about’ (or ‘what I was talking about
before’), ‘theme’ means ‘what I am talking about’ (‘or what I am talking about
now’) and as any student of rhetoric knows the two do not necessarily coincide.

Leaving aside Halliday’s argumentation, it seems to me that combining inter-
pretations pose two major questions: (1) how reliably can message/clause elements be
identified as Theme/Topic if the notion of ‘Givenness’ has been approached from
many different perspectives and associated with several different notions that have of-
ten been indistinctly used in the literature?; (2) how is the chosen formulation of
Theme/Topic linguistically relevant? Regarding the question in (1), I agree with
Chafe's (1976: 3) and Allerton's (1978: 133, 151 ff.) observation that *Givenness* should firstly be distinguished from *presupposition*. In my view, Givenness is a discourse notion referring to the informational status of the constituents of a message, which is determined by the speaker/reader's view of the situational and linguistic context (including the addressee's communicative needs), and which is indicated by attenuated morpho-syntactic and phonological forms (e.g. pronominalisation, definiteness, weaker stress, etc.). Conversely, presupposition is a logico-semantic notion realised by sentence form which involves a *proposition* (i.e. a potential sentence having the capacity of being true or false), whose assumability is required for the success of the message (see Jackendoff 1972: 276-8). It follows that a proposition may convey Given information, but need not be presupposed (e.g. *It can't be true*, as an answer to *I saw the man*, where *it* stands as Given information, but the proposition is not presupposed). Or, vice versa, informationally new items may occur within a presupposition (e.g. *What the duke gave to my aunt was that teapot*, where it is presupposed that the duke gave something to someone, but the identity of that someone is presented as news).

Now, focusing on Givenness, it seems to me that this term has been used in at least three distinct senses that render different analyses of Theme/Topic: (1) the *relational*, or reciprocal, interpretation; and (2) two types of *referential* accounts, i.e. (a) contextual and (b) activated (cf. Gundel 1988: 211-2).

**Table 1: Trends in informational interpretations of Theme/Topic**

<table>
<thead>
<tr>
<th>clause level of analysis</th>
<th>discourse level of analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>relational interpretation</td>
<td>referential interpretations</td>
</tr>
<tr>
<td>Given information in relation to New information</td>
<td>Recoverability (Giv_R)</td>
</tr>
<tr>
<td></td>
<td>Predictability (Giv_P)</td>
</tr>
<tr>
<td></td>
<td>Shared Knowledge (Giv_K)</td>
</tr>
<tr>
<td></td>
<td>Assumed Familiarity (Giv_F)</td>
</tr>
<tr>
<td></td>
<td>entities participants</td>
</tr>
<tr>
<td></td>
<td>are attending to</td>
</tr>
<tr>
<td></td>
<td>at the time of</td>
</tr>
<tr>
<td></td>
<td>the utterance</td>
</tr>
</tbody>
</table>

In what follows it will be shown that these three informational accounts of Theme/Topic rest upon different theoretical primitives that operate at two different levels of analysis, and therefore pursue different goals and yield different results.

To begin with, the relational, or reciprocal, approach works at individual clause/utterance/single speech act level, whereas contextual and activated approaches operate at the level of discourse. Indeed, in relational interpretations Theme/Topic is regarded as Given in relation to Rheme/Comment within the domain of *individual clauses/utterances* (cf. Gundel 1988; Firbas 1964). Alternatively, contextual referential approaches conceive Theme/Topic as the element that relates the clause/utterance to the discourse context in which it occurs, so that shared familiarity by the interlocutors becomes the necessary condition for felicitous thematicity/topichood. This pic-
tecture is further complicated for, in my view, contextual accounts invoke four non-constituent readings of Givenness: Recoverability (GivR), Predictability (GivP), Shared Knowledge (GivK), and Assumed Familiarity (GivF).

Recoverability (GivR) is upheld, for example, by Halliday (1967, 1974, 1994) and his advocates in Systemic Functional Grammar (SFG), who believe Given ‘express[es] what the speaker is presenting as information that is recoverable from some source or other in the environment – the situation or the preceding text’ (Halliday and Hasan 1976: 326). In contrast, supporting the notion of Predictability (GivP), de Beaugrande and Dressler (1981) and Kuno (1972, 1978, 1980) argue that Given represents what the addressee expects the speaker/writer is likely to say/write, while New relates to expressions not following (or not predicted by) these expectations. In turn, Clark and Haviland (1977: 4) describe Shared Knowledge (GivK) as what the speaker ‘believes the listener already knows and accepts as true’ (but is not necessarily thinking about it) (cf. Haviland and Clark 1974); whereas Allerton (1978: 147) and Prince (1981: 233-37) argue for Assumed Familiarity (GivF) as a scalar notion, the latter scholar distinguishing seven degrees of familiarity which are broken down into three groups:

1. **New**, or entities firstly introduced in discourse by the user, are classified into three types (viz. Brand-New Unanchored, Brand New Anchored, Unused);

2. **Inferable**, or entities that the speaker/writer assumes her/his addressee can infer through reasoning;

3. **Evoked**, or entities that are already present in discourse (viz. situationally Evoked or textually Evoked).

In their turn, in the last version of informational analyses, that is, in activated referential interpretations, Givenness (GivR) refers to entity (-ies) which the interac-
tants are not only familiar with, but are actually attending to (thinking of) at the time of utterance. A case in point is Chafe’s (1974, 1976: 30, 1987: 30 ff.) definition of the Given as ‘that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance’, for which its referent must have been explicitly introduced in the discourse, i.e. be present in the physical context or be categorised in the same way as a referent previously introduced or physically present. By contrast, Givón (1988, 1992: 10) proposes an alternative description of GivR as a discrete process of attention activation within a cognitively based framework. Put succinctly, Givón establishes a correlation between two textual dimensions of topicality, referential accessibility (i.e. Givenness as derived from the cultural knowledge, the speech situation and/or the preceding discourse) and thematic importance (i.e. the text frequency of referents in subsequent discourse), on the one hand, and two cognitive domains, on the other, that is, mental storage (i.e. where a currently inactive node that is already in storage is identified) and attentional activation (whereby a referent Topic,
or file, is open, or activated, to receive incoming information) (cf. Copeland and Davis 1983; Lambrecht 1988, 1994).

Discrepancies arise within informational approaches because, on closer inspection, relational, referential and activated interpretations of Theme/Topic are not independent of each other, and because they use syntactic, contextual and activated criteria in a loose ambiguous manner. It seems, for example, that all types of Givenness qualify for $\text{Giv}_K$, although the reverse is not necessarily the case. Indeed, in order for information to be presented as recoverable, salient or predictable the speaker/writer must assume that the hearer ‘knows’ or can infer a particular thing. Moreover, if the speaker/writer assumes some information to be in the consciousness of the addressee, it is also likely to be recoverable and/or predictable from the linguistic or situational context. Likewise, it seems to me $\text{Giv}_G$ definitions assume a knowledge of the workings of memory and attention/consciousness which is simply not available at the moment, so that there is no way of empirically validating whether an element is Given or not. Some vexing questions are: for how long can mental representations be active and eligible for thematic/topical status, if they quickly fade from consciousness unless refreshed by continual mention? How can the ‘silent’ mental shift from Given to Accessible be measured and regarded in terms of thematicity and/or topicality?

Behind all these questions lies the second problem posited by informational approaches, that is, whether or not they provide operational or empirically verifiable analyses of Theme/Topic. It is my impression that most of these accounts are somewhat elusive, because they do not identify the Theme/Topic of a message directly. Instead, this category is indirectly described:

(1) in relation to the Comment/Rheme (New) at clause/utterance level (in relational accounts);

(2) in relation to such slippery concepts as ‘recoverability’, ‘predictability’, ‘shared knowledge’ and ‘saliency’;

(3) as a discrete value along a scale of topicality;

(4) as a process of attention activation.

Neither is it clear to what extent the Given-New status is determined by: (1) syntactic form (viz. word order, specific Givenness/newness markers, etc.), (2) the context or (3) the speaker. The latter hypothesis is supported by Halliday (1967) and his followers in SFG, but, for Dik and the supporters of Functional Grammar (FG), as well as for many Prague School advocates, it is the co(n)text, while CD acolytes rely on intonational Givenness.

In addition, most informational accounts tend to restrict their explanatory power to NPs: thematic/topical items are identified with zero anaphora and with
(modified) definite NPs, whereas indefinite NPs receive rhematic/topical status. This identification leads to three additional debatable issues. One is, if it is true, as noted by Givón (1984; 1992: 42), Du Bois (1987) and Chafe (1987), among others, that clauses/utterances may contain more than one Given NP but only one is perceived as Theme/Topic, then it follows that additional criteria apart from that of Givenness have to be invoked to select among the potential thematic/topical items. Secondly, the assumption that only (Given) NPs qualify for thematic/topical status is not totally accepted, but has been criticised by some scholars such a Allerton (1978: 157), Reinhart (1982: 72), Jordan (1985) and Lütscher (1985). The latter, in (1985: 207-208), states that ‘topicalized elements can be either thematic [referential] [...] or rhematic [non-referential]’ in languages such as German, Russian, Japanese, Chinese, French or English, although these languages seem to vary in the ease with which non-referential Topics are expressible.

My last concern about informational approaches regards the treatment of all-New messages (i.e. all New clauses, newspaper headlines, titles, etc.). For the equation of Givenness with Theme/Topic renders all such messages as devoid of Theme/Topic functions and therefore as being ‘groundless’ or informationally incomplete. To preclude this possibility, Chafe (1987) affirms that all constituents would be good candidates for thematic status, were we to calculate their ‘activation state’. However, it seems to me that, if Chafe’s procedure were to be adopted (difficult though it may appear to calculate the degree of activation of some word classes such as conjunctions, adverbs, prepositions etc.), then the category of Theme/Topic would embody such a vast and heterogeneous territory that the concept would be virtually useless for explaining the structure of individual messages/utterances/clauses and/or connected discourse.

4. SEMANTIC INTERPRETATIONS OF THEME/TOPIC

According to semantic interpretations Topic (less frequently, Theme) expresses a relationship of ‘aboutness’: it indicates ‘what the message is about’. This description echoes Grice’s (1975) maxim of relevance: ‘Make your contribution relevant in terms of the existing topic framework’, that is, make your contribution fit closely to the most recent elements incorporated in discourse. However, this section will show that despite addressing the notion of ‘aboutness’ as criterial for topical/thematic status, many semantic approaches have ended up using positional or informational cues to identify the Theme/Topic of a message, a practice which, as shown, is the protocol of syntactic and informational analyses, respectively. Furthermore, it will be argued that ‘aboutness’ has been invoked in at least three non-coterminous readings, namely: (1) two message-centred interpretations, i.e. (a) relational and (b) referential, and (2) one context-centred interpretation, i.e. the interactive, or framework, interpretation, as illustrated in Table 2 below:
Table 2: Trends in semantic interpretations of Theme/Topic

<table>
<thead>
<tr>
<th>clause level of analysis</th>
<th>discourse level of analysis</th>
<th>context-centred interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>message centred interpretations</td>
<td>referential interpretation</td>
<td>interactive interpretation</td>
</tr>
<tr>
<td>relational interpretation</td>
<td>contextual interpretation</td>
<td>activated</td>
</tr>
<tr>
<td>proposition</td>
<td>discourse entity</td>
<td>propositional/problem</td>
</tr>
<tr>
<td>in a complex clause</td>
<td>or proposition</td>
<td>framework</td>
</tr>
<tr>
<td>clause/utterance entity</td>
<td>discourse entity</td>
<td>saliency/relevance</td>
</tr>
<tr>
<td>or proposition</td>
<td>or proposition</td>
<td></td>
</tr>
</tbody>
</table>

The relational, or reciprocal, semantic interpretation conceives the informational structure of the clause as a bipartite construct entailing a relation of ‘aboutness’ between a Topic/Theme, or an entity (i.e. referent / participant / constituent / proposition) and a clausal predication, or Comment/Rheme (cf. Li and Thompson 1976: 463; Kieras 1981: 2; Reinhart 1982: 48; Lambrecht 1994: 127). From this perspective, Topic/Theme has been related to such psycholinguistic notions as frames of reference, cognitive hooks, pegs, links, background information and perspective-taking devices by, for example, Bühler (1934), Halliday (1967, 1994), Lutz (1981), Adamec (1981: 226), Kieras (1978), Fillmore (1968), Kuno (1976), Dik (1978), or Duranti and Ochs (1979).

Alternatively, the referential semantic interpretation describes Topic/Theme as the referent/participant/constituent/idea that establishes a relationship of ‘aboutness’ between individual clauses/utterances and the overall discourse:

1. either as determined by the co(n)text, in contextual semantic interpretations;
2. and/or as processed by the (decoder’s/receiver’s) mind, as proposed by activated semantic accounts;


As I see it, semantic accounts raise two debatable issues: (1) whether or not human judgement/statements are double, that is, whether or not they distinguish a Topic/Theme from a Comment/Rheme (either implicitly or explicitly); and (2) whether or not there exist systematic markers of ‘aboutness’. The relational premise in (1) was challenged in the 19th century by Franz Brentano and Anton Marty, who ar-
gued for the existence of *thetic judgements*² (thrẹtikόs ‘positive’, or einfaches Urdteil, ‘simple judgement’), as opposed to the double, or *categorical ones* (cf. Marty 1918; Paul 1909; Bally 1932) Thetic messages are characterised as single unstructured kind of judgements involving only the recognition or rejection of some judgement material independently of a recognised Subject, as in:

(1) impersonal expressions (e.g. events, states, facts, etc. such as *It is raining*);

(2) presentative constructions (e.g. *Byl jednou jeden král* ‘Once upon a time there was a king’, Mathesius 1961: 83);

(3) any state of affairs presented as a compact whole representing nothing but new information (e.g. *A girl broke a vase, The sun is shining*).

The thetic-categorical distinction has been approached in linguistics from two different perspectives: the logico-semantic approach and the informative-pragmatic one. In the former view, the contrast represents two different points of view, or communicative perspectives, from which a given state of affairs can be shaped into a clausal predication: categorical clauses are logically complex, whereas thetic ones are logically simple and thereupon constitute a category of their own that cannot be captured with principles of information structure such as Theme/Topic-Rheme/Comment patterns (cf. Kuroda 1972, 1984, 1985; Dahl 1974, 1976; Sasse 1984, 1987; Ulrich 1985). On the other hand, the supporters of the informative-pragmatic approach share the premise that the contrast between the thetic and categorical judgements has to do with the activation and identifiability state of the Subject referent (cf. Mathesius 1929; Bolinger 1954; Firbas 1966; Halliday 1967; Chafe 1974; Schmerling 1976; Faber 1987; Kuno 1972; Horn 1989). Thus, Lambrecht (1987: 366; 1994: 137 ff.), for example, explains that thetic structures, his *sentence focus* (SF) constructions, introduce a referentially non-salient or non-individuated, and therefore non-topical (i.e. relatively inactivated or unidentifiable) NP or event into the universe of discourse, differing from categorical messages (his *predicate Focus* (PF) constructions) in the way the Subject NP is grammatically marked:

(1) accented vs. non-accented Subject NP in English or German;

(2) postverbal vs. preverbal Subject in Romance, Slavonic, German and Chinese;

(3) clefted vs. detached NP in French, Welsh and Arabic;

---

² There is no established terminology concerning thetic structures. Among the labels applied to these statements stand: Bolinger’s (1954) *presentational sentences*, Kuno’s (1972) *neutral descriptions*, Schmerling’s (1976) *news sentences*, Lambrecht’s (1986) *event reporting sentences*, or Lambrecht’s (1987, 1994) *sentence focus structures*. 

---
(4) special morphological marking, such as ga-marked vs. wa-marked NP in Japanese, or similar examples in Bantu.

However, Lambrecht emphasises that cognitively thetic clauses may be either structurally coded as categorical, that is, endowed with a syntactic Topic/Theme-Comment/Rheme (Subject-Predicate) structure (i.e. clauses with pronominal or null Subjects, such as Es regnet, Pluit, or It is raining). Or they may not involve any grammatical Subject at all (e.g. V zádech me boli vs. Boli me v zádech in back-LOC me-ACC hurt-3 sg. ‘My back HURTS’ vs. ‘MY BACK hurts’). This latter point connects directly with the second issue raised by semantic interpretations of Theme/Topic, that is to say, whether or not there exist systematic markers of ‘aboutness’.

It is my impression that neither text linguists nor conversational analysts endorsing semantic interpretations of Theme/Topic as a discourse entity or proposition have yet provided an operational description of how to derive, represent, and relate the most important ideas in a text. To begin with, the postulated scales for entity Topic/Theme show three deficiencies:

(1) they limit Theme/Topic to a particular syntactic/semantic function, fundamentally Subject/Agent;

(2) they leave questions of referent resolution unaddressed since syntactic and semantic constraints on reference are ignored;

(3) their validity is seriously flawed because the scales are:

(a) data dependent (they depict at best the coding of protagonists in narratives), and so skew expectations for other data types; and

(b) by and large based on quantitative analyses.

By the same token, propositional semantic interpretations do not offer a means of identifying ‘the Theme/Topic’ of a piece of discourse, but propose an alternative method of producing a single sentence summary for the text, which can only to a certain extent be objective. This seems to be, as pointed out by Brown and Yule (1983: 110), ‘a formula for determining, not the Topic of a discourse, but the possible topics of a discourse [my emphasis].’

Finally, it appears that most interactive approaches avoid, instead of providing answers to, the difficulties inherent in the notion of Theme/Topic. In the first place, they allow for so many types and treatments of Theme/Topic as to render this type of analysis both subjective and virtually useless in accounting for linguistic structure. For one thing, if Themes/Topics are viewed as problem frames to be treated in context-changing utterances, theoretically fewer restrictions have to be imposed on the possible candidates for Topics/Themes, to the point that, within a given proposition, any constituent (or the proposition itself) may appear as creating a frame of cogni-
tive/communicative discrepancies calling for treatment. For this reason, Schank (1977: 424) and Maynard (1980) explain that most discourse analysts, aware of the profusion of topical/thematic candidates, do not try to elicit Topics/Themes themselves, but find it more feasible to identify the formal markers of a Theme/Topic shift in discourse, such as:


2. **indentations** and **paragraphs**, the latter being either orthographic or semantic, that is, demarcated by the unity of participants and setting adverbials (cf. Hinds 1977: 83; Grimes 1975: 109; Longacre 1989: 116),

3. **genre-specific Topic shift markers** (e.g., changes of place and time in narratives, of the thread of argumentation, etc.) (cf. Grimes 1975: 102; van Oosten 1986),

4. in spoken discourse, **paratones**, or speech paragraphs, (cf. Brown 1977: 86), and **kinesics** (e.g., gaze, body movements, etc.) (cf. Kendon 1967).

Nevertheless, the problem remains that the enumerated Topic/Theme -shift markers represent optional tendencies only. In other words, the question needs to be addressed as to how the analyst can determine the occurrence of a shift in Topic/Theme systematically, as Theme/Topic shift markers may be absent from discourse, or they may not be used as expected.

5. CONCLUSIONS

In the light of the discussion presented so far it seems that a ‘definition’ of Topic/Theme that would cover all instances of the three aspects of the terms, syntactic, informational and semantic, is unobtainable. And, by the same token, it appears that none of the posited Topic markers in the so-called **Topic-prominent languages** (cf. Li and Thompson 1976), e.g., Japanese (-wa), Tagalog (-ang), Korean (-nun), etc., can be simply equated with any existing definition of Topic/Theme. As an illustration, behold (2) below (from Tsutsui (1981:164) and Kitagawa (1982:176)), where the Japanese suffix -wa, allegedly a Topic marker, is used to mark: contextually given information (2i), a constituent detached from the clause (2ii), and a focus of contrast (2iii), while Gundel (1988: 218) illustrates the existence of topical items lacking this particle in (2iv):

---

3 Schachter (1976, 1977) and Schachter and Oaines (1972: 81) raise similar points with respect to the -ang particle.
What solution is left then? In this paper I advocate the independent treatment of all linguistically relevant categories (at all levels) that have been factored out into the Topic/Theme equation, i.e. initial position, ‘aboutness’, and Givenness, which nonetheless may interact with each other (cf. Gómez-González 1994, 1996). Hence, Theme is here identified with the category pursued by syntactic approaches, that is, with syntactic Theme, or the starting point of the clause. This represents the source of the expressed event and is cognitively related to ‘schemes of imagery’, which imply semantic networks of the type ‘source’, ‘goal’, ‘path’, etc. (cf. Langacker 1987: 39 ff.; Lutzeier 1191). However, as noted in section 2 above, such syntactic analyses need to:

1. produce an operational criterion that identifies syntactic Themes;
2. demonstrate the discourse-functional relevance of clause initial (final) position, by embodying within the same paradigm the wide range of constructions that highlight thematic/topical position and by exploring such constructions in real texts, within and across languages;
3. determine whether the Theme-Rheme pattern is either a structural grammatical relationship or a non-structural principle of organisation of discourse.

On the other hand, Topic is regarded to entail a relationship of ‘aboutness’ with the rest of the message, as supported by semantic approaches. Nonetheless, in section
4. It has been shown that uncovering what is topical in a ‘semantic’ sense requires firstly a clarification of what type of ‘aboutness’ is at issue:

1. relational, established by an entity or a proposition with respect to a clausal predication;
2. referential, entailed by a referent with respect to the overall discourse;
3. interactive, evoking salient/relevant information in discourse.

In addition, relational semantic approaches need to demonstrate that individual messages/utterances/clauses are necessarily duplex, which demands the elicitation of Topic markers across/within languages. Similarly, in order to be more consistent, the scales for entity Topic postulated by the referential semantic interpretations need to substantiate with sufficient empirical evidence why the category of Topic should be limited to a particular syntactic/semantic function (e.g. Subject/Agent); whereas interactive interpretations must provide the means to objectively identify speakers’ and/or discourse Themes/Topics.

Finally, though closely related to the other two, Givenness is here treated as a category of its own. For, I believe that Givenness is neither a sufficient nor the most elegant criterion to explain the phenomena of thematicity or topicality, because (1) it does not define the category of Topic or Theme independently of other categories, and (2) it restricts its explanatory power mainly to NPs. Furthermore, section 3 has demonstrated that Givenness evokes an elusive notion that can be approached from three different perspectives:

1. relational, as Given information in relation to what is presented as New in individual clauses/messages in isolation;
2. contextual, as information rendered as Given by the co(n)text, in terms of:
   a. recoverability (GivR), or information which is recoverable from the co(n)text;
   b. predictability (GivP), or information that is predictable from the co(n)text;
   c. shared knowledge (GivK), i.e. the knowledge shared by the interactants at the moment of the exchange;
   d. assumed familiarity (GivF), or a scalar notion of information which the speaker assumes her/his addressee can retrieve or infer from the co(n)text;
3. activated Givenness (GivS), referring to what interactants have in the mind.

REFERENCES


ALLERTON, David J. (1978): “The notion of ‘Givenness’ and its relations to presupposition and to Theme”, Lingua 44, 133-68.


——— (1988): “The pragmatics of word order: predictability, importance and attention”, In M. Hammond et al. (eds.), 243-84.


