Daniel Dejica

Polytechnic University of Timisoara

Probing for Relevance: Information Theories and Translation

Summary

Recent studies claim that the more translators know about the structure and the dynamics of discourse, the more readily and accurately they can translate both the content and the spirit of a text. Similarly, international research projects highlight directions of research which aim at helping translators make reasonable and consistent decisions as to the relevance and reliability of source text features in the target text. Other recent studies stress the importance of developing information structure theories for translation. In line with such current research desiderata, the aim of this article is to test the relevance of information theories for translation. In the first part, information theories are presented from different linguistic perspectives. In the second part, their relevance for translation is tested on a series of texts by examining how they have been or can be applied to translation. The last part presents the conclusions of the analysis.

Key words: information structure, source text analysis, translation studies

V iskanju relevantnosti: informacijske teorije in prevajanje

Povzetek

Zadnje raziskave kažejo, da bolj ko so prevajalci seznanjeni s strukturo in dinamiko diskurza, bolj vestno in bolj natančno so sposobni prevesti tako vsebino kot duh besedila. Podobno mednarodni raziskovalni projekti izpostavljajo smernice za sklepanje razumnih in doslednih prevajalskih odločitev glede relevantnosti in zanesljivosti izhodiščnih lastnosti v ciljnem besedilu. Druge sodobne študije poudarjajo pomen razvijanja teorij informacijske strukture za prevajanje. V skladu s takšnimi znanstvenimi prizadevanji je cilj pričujočega članka preveriti relevantnost informacijskih teorij za prevajanje. Prvi del predstavi različne jezikoslovne vidike informacijskih teorij. Jedro drugega dela predstavlja preizkus njihove relevantnosti na vrsti besedil v smislu njihove dosedanje in potencialne prevajalske rabe. Prispevek sklenejo ugotovitve analize.

Ključne besede: informacijska struktura, analiza izhodiščnega besedila, prevodoslovje

Probing for Relevance: Information Theories and Translation

1. Introduction

In the past decades there has been a veritable surge of interest in the way information is structured in communication. The terminology used to describe this field of research is varied and under-formalized: Functional Sentence Perspective, used by the scholars of the Prague School of linguists (Firbas 1964, 1992), information structure (Halliday 1967), information packaging (Chafe 1976), or informatics (Vallduví 1992), to name just a few.

Under these headings, linguists usually speak of two information-structural parts of the sentence or the clause known as "Theme" and "Rheme" or "topic" and "comment". Another set of terms include, on the one hand, "given," "old" or "known" information, and, on the other hand, "new" or "unknown" information.

Some of these notions and other new ones have been applied to higher levels of discourse than sentence or clause level: macro- and hyper-Themes (Daneš 1974) or grounding (Hopper 1979). In general, all these approaches imply that certain parts of discourse are more important, more salient or more foregrounded than others, and illustrate how connections between these parts account for discourse structure which, in its turn, accounts for the information flow in discourse and discourse coherence.

Translation studies could not remain unresponsive to the input that such theories might bring to its practical or theoretical development. In fact, recent studies claim that the more translators know about the structure and the dynamics of discourse, the more readily and accurately they can translate both the content and the spirit of a text (Nida 1997, 42). Similarly, international research projects highlight directions of research which aim at helping translators make reasonable and consistent decisions as to the relevance and reliability of source text features in the target text (Gerzymisch-Arbogast 2005, 7). Other recent studies stress the importance of developing information structure theories for translation (Gerzymisch-Arbogast 2005, Gerzymisch-Arbogast/Kunold/Rothfuß-Bastian 2006).

In line with such current research desiderata, the aim of this article is to test the relevance of major information theories for translation. In the first part, information theories are presented from different linguistic perspectives. In the second part, their relevance for translation is tested on a series of texts by examining how they have been or can be applied to translation. The last part presents the conclusions of the analysis.

2. Information theories – an overview

Information structure, information packaging or informatics: under these concepts, linguists show how information is structured in sentences or texts with the aim of providing a better text understanding.

In describing the *information structure* of a clause, Halliday (1994, 274–5) uses the concepts *given* and *new.* Instead of the mathematical concept of 'information', which has been developed

in cybernetics, Halliday sees it as the interplay between given and new in the information unit, which can coincide with the clause. In more pragmatic terms, Lambrecht (1994, 5–7) states that *information structure* is concerned with the relationship between linguistic form and the mental states of speakers and hearers. According to him, the linguist dealing with information structure must deal simultaneously with formal and communicative aspects of language. In his perspective, information structure is viewed as related to sentence grammar (Lambrecht 1994, 5).

Information packaging is a structuring of sentences by syntactic, prosodic or morphological means that arises from the need to meet the communicative demands of a particular context or discourse (Vallduví and Engdhal, 1996). In particular, information packaging indicates how information conveyed by linguistic means fits into a hearer's mental model of the context or discourse. When communicating a proposition, a speaker may realize it by means of different sentential structures according to his or her beliefs about the hearer's knowledge and attentional state with respect to it. The concept of information packaging is due to Chafe (1976):

> "I have been using the term packaging to refer to the kind of phenomena at issue here, with the idea that they have to do primarily with how the message is sent and only secondarily with the message itself, just as the packaging of toothpaste can affect sales in partial independence of the quality of the toothpaste inside." (Chafe 1976, 28)

In his theory of *informatics*, Vallduví (1990) proposes a hierarchical articulation of information. Sentences are divided into the *focus*, which represents that portion of information that is hearer-new, and the *ground*, which specifies how that information is situated within the hearer's knowledge-store. The ground is further divided into the *link*, which denotes an address in the hearer's knowledge-store under which he is instructed to enter the information, and the *tail*, which provides further directions on how the information must be entered under a given address.

What these theories have in common is that, under different headings and using sometimes different approaches, they all involve the processes of information identification, analysis and structuring. Based on this shared-ground principle and wishing to avoid entering a terminological debate in this direction, the general term *information theories* is used in this article to refer to the previously mentioned theories. From this perspective, the term is to be distinguished from Information theory as a branch of applied mathematics and electrical engineering involving the quantification of information.

2.1 Putting information theories into practice – various approaches and perspectives

The information theories presented in the previous section display various approaches to analysing information both at sentence and text levels. Most of these approaches see information organized as a two-part system in a sentence, one being always more salient, foregrounded, focused, etc. than the other; just like the theories which they stand for, they use differing terminologies. Among the most often distinguished notions in the area of information identification and analysis are Theme-Rheme, topic, focus, given-new, and foregrounding-backgrounding. Combinations of these may occur: Theme-contrast, Theme-tail, background-focus, etc.

2.1.1 Theme and Rheme

The semantic organization of the sentence or clause into two parts, the "Theme" and the "Rheme" has long been dealt with in linguistics. There are several detailed reviews of ideas and issues of clause level Theme (Dik 1981, 1989; Tomlin 1986; Firbas 1992; Halliday 1994). The standard approach – Theme as starting point of the sentence – considers the Theme to be whatever comes in first position in the sentence; whatever follows it is the Rheme. In this respect, every clause has the structure of a message: it says something (the Rheme) about something (the Theme). Other approaches regard Theme as "aboutness," that is to say the Theme is what the sentence is about. In this respect, it acts as a point of orientation by connecting back to previous stretches of discourse and thereby maintaining coherence.

Mathesius and Daneš were among the first linguists to develop the theory of Theme as "starting point of the utterance". In this conception, the Rheme is everything else that follows in the sentence and consists of "what the speaker states about, or in regard to, the starting point of the utterance" (Mathesius 1942). The main functions of the Theme in this approach are to connect back and link to the previous discourse, maintaining a coherent point of the text, and to serve as a point of departure for the further development of the discourse (Daneš 1974). In the past decades, this approach has been explored in depth by a series of functional linguists, including, but not limited to Siewierska (1991), Firbas, Fortesque, Harder and Kristoffersen (1992), and Halliday (1994).

Gundel (1974) and Lambrecht (1994), on the other hand, do not approach Theme as the starting point of the sentence, but relate it to the definition of "subject" in traditional grammars. In their view, the Theme of a sentence is what the proposition expressed by the subject is about.

A similar approach of Theme as aboutness has been developed by Mudersbach (2008, 3–6) who proposed a *Theme-Rheme model* which uses pragmatic parameters, such as *the scope of shared attention, joint stock of information, background knowledge,* etc., for the for the identification of Themes and Rhemes in oral communicative situations. The model has been adapted for written communicative situations and it has been used for the analysis of coherence and cohesion of a text by illustrating various textual relations between Themes and Rhemes (Dejica 2006, 103–10).

It has been rightfully claimed that "the selection of an individual theme of a given clause in a given text is not in itself particularly significant. But the overall choice and ordering of themes, particularly those of independent clauses, plays an important part in organizing a text and in providing a point of orientation for a given stretch of language" (Baker 1992, 126).This internal organization of Themes in units larger than the sentence has been labelled *thematisation* by Brown and Yule (1983), who see it as "a discoursal rather than sentence process" (1983, 133).

Before than Brown and Yule, and in more technical terms, Daneš (1974) used the phrase *thematic progression* to refer to the way subsequent discourse re-uses previous Themes or Rhemes according to an overall textual plan. Thematic progression relates the way Themes and Rhemes concatenate within a text to the hierarchic organisation of the text and ultimately to rhetorical purposes. Daneš observed different patterns of matching sentence arrangements described in

terms of linear progression, progression with constant Theme, and progression with derived Theme. These types of progression and new ones, i.e. progression with derived Rheme, have been presented in detail and exemplified in Dejica and Superceanu (2004).

Thematic progression at paragraph (episodic) level can be represented hierarchically; since sentences consist of more than one clause, they will have several layers of thematic structure. Each clause will have its own Theme-Rheme structure which may be subordinated to a larger Theme-Rheme structure. This hierarchical nature of Theme-Rheme structures on a text is illustrated in Example 2.1; the identification of Themes and Rhemes at sentence and clause levels is represented in Fig. 2.1, and the comparative analysis of the relations between them is presented in Fig. 2.2:

(2.1)   The City of Los Angeles comprises 1,215 sq km (469 sq mi) and had a population of about 3.7 million people at the 2000 census. It is the largest municipality (in terms of size and population) among all the cities in Los Angeles County. It is irregular in shape because it has grown over the years through the annexation of surrounding territory and cities. The city proper is shaped like a lighted torch: its narrow handle extends north from the Port of Los Angeles to downtown Los Angeles, and its flames flicker irregularly to the north, west, and northwest. Several separate cities—such as Beverly Hills, Santa Monica, and Culver City—are partly or completely surrounded by the City of Los Angeles. The city is bisected by the Santa Monica Mountains, which run east to west.

(*Los Angeles*, in Microsoft® Encarta® Encyclopedia 2003. © 1993-2002 Microsoft Corporation)

The City of Los Angeles comprises 1,215 sq km (469 sq mi) and had a population of about 3.7 million

T1	R1	T2 (the city)	R2
T1	R1		

people at the 2000 census. It is the largest municipality (in terms of size and population) among all

...R2	T3	R3
...R1	T2	R2

the cities in Los Angeles County. It is irregular in shape because it has grown over the years through

...R3	T4	R4	T5	R5
...R2	T3	R3		

the annexation of surrounding territory and cities. The city proper is shaped like a lighted torch: its

...R4	T6	R6	T7
...R3	T4	R4	

narrow handle extends north from the Port of Los Angeles to downtown Los Angeles, and its flames

...T7	R7	T8
...R3		

flicker irregularly to the north, west, and northwest. Several separate cities—such as Beverly Hills,

R8	T9
...R4	

Santa Monica, and Culver City—are partly or completely surrounded by the City of Los Angeles.

...T9	R8
...R4	

The city is bisected by the Santa Monica Mountains, which run east to west.

T10	R10	T11 (the SMM)	R11
T5	R5		

Fig. 2.1   *Hierarchical thematic analysis of a text at clause and sentence levels*
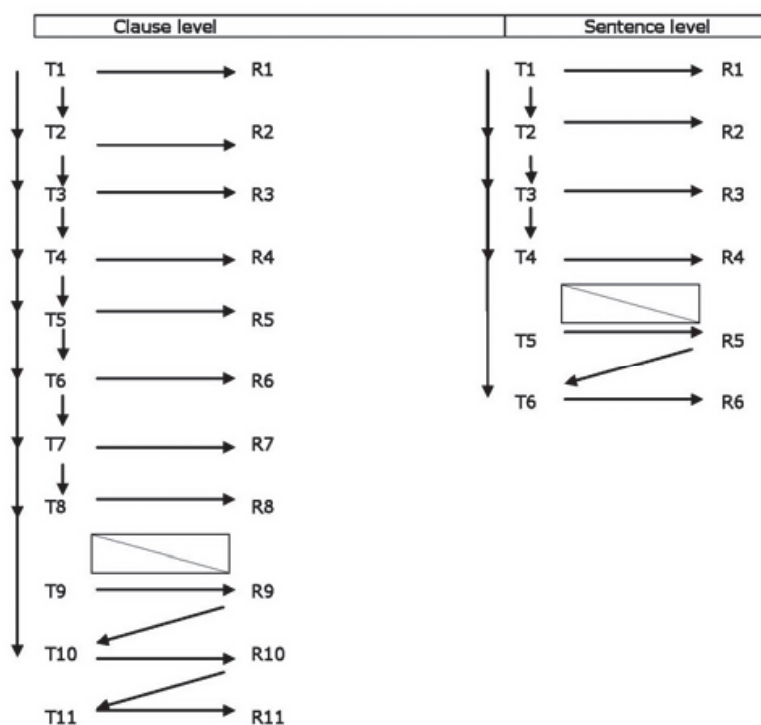
Fig. 2.2   Hierarchical thematic representation of a text at clause and sentence levels

The thematic representation (Fig. 2.2) of the thematic analysis (Fig. 2.1) shows that there are two kinds of progression which characterize the paragraph at both sentence and clause levels: progression with derived Theme, and linear progression. In non-technical terms, the emerging pattern for the text may be described as "parallel," as the same Theme is repeated at the beginning of almost all the clauses in the first part of the text, and "zig-zag," as the Rhemes of the last sentences become the Themes of the immediately succeeding sentences.

The above paragraph (example 2.1) is a relevant example for how successive Themes are related to a single preceding Theme (or *hyper-Theme* as Daneš terms it). According to Daneš (1974), the hyper-Theme functions as what would be termed a Topic Sentence in rhetoric – in other words, as the Theme of the paragraph, rather than the Theme of a clause. Martin (1992) provides his own definition of the term, which is related to Daneš's and which he extends to Themes which may occur at different levels of discourse. For Martin,

> "a *hyper-Theme* is an introductory sentence or group of sentences which is established to predict a particular pattern of interactions among strings, chains and Theme selection in following sentences. . . . On the basis of this definition of hyper-Theme, 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 in rhetoric. . . . Macro-Themes may be themselves predicted as *super-Themes*, super-Themes as *ultra-Themes* and so on, depending on the number of layers of structure in a text." (Martin 1992, 439–91)

According to Martin's classification, in longer texts, this pattern of macro-Themes predicting hyper-Themes can be extended, with hyper-Themes themselves functioning as macro-Themes in their own right. Once texts develop to this level of internal complexity, titles, subtitles, headings

and subheadings are commonly deployed to keep track of the composition structure being erected. In other words, in much the same way as clause-level Themes are tied between them, so, too, are the units connected to yet higher-level structures. In writing, the use of macro-Themes to predict hyper-Themes, which in turn predict a sequence of clause Themes, is an important aspect of coherence, and texts which do not make use of predicted patterns may be considered as less coherent or incoherent.

2.2.2 Topic

There are mainly two distinct uses of the term *topic* at sentence level. On the one hand, *topic* is used widely to capture ideas similar to clause level Theme. On the other hand, *topic* is seen as part of the predication, different from Theme.

In some studies, a clause level *topic* is identical to clause level Theme, the two terms being used virtually synonymously, especially by the academic inheritors and developers of the Prague School theories. This can be found in Sgall (1987), who uses the terms topic-comment to denote the Prague approach in which the sentence is divided into its two functioning parts, Theme (topic) and Rheme (comment). Thus, the topic, just as the Theme as starting point, links each sentence with what has gone before, while the comment cumulatively moves on the meaning which the discourse is to communicate. The topic can also function as Theme as aboutness, where the topic of the sentence is what the proposition expressed by the sentence is about (Lambrecht 1994).

The notion of topic has also been applied to structures higher than sentence structures. For instance, Keenan and Schieffelin, discussed in Brown and Yule (1983), use the term *discourse topic* in an attempt to distinguish the notion of topic from sentence topic. For them, discourse topic is not a simple noun phrase, but a proposition which represents the whole of the fragment. Van Dijk (1982) also presents an explicit formal account for discourse topic or, as he terms it, *topic of discourse* (1982, 132). According to him, topic is basically as a macrostructure drawn from all the information contained in a text. To arrive at the topic, the addressee or the interpreter assembles all the factual statements of the text and summarizes them in a statement called a *macroproposition*. It should be stressed that van Dick's approach is based on an underlying semantic representation of the text rather than on the sequence of sentences which constitute the text.

This latter aspect of topic continuity in a text is an element which was developed by Givon and which accounts for the discourse structure (1983). He analyses how clauses are combined into larger units of discourse which he calls *thematic paragraphs* (1983, 7). Thematic paragraphs may further combine into larger discourse units such as sections, chapters, parts, or stories. Givon discusses the process of *continuity of discourse* at paragraph level in terms of *thematic continuity, action continuity,* and *topic/participant continuity* (1983, 8). It is to be noticed that Givon associates topic continuity with the higher-level Theme of the paragraph which stands for what the paragraph is about. This concept differs from the notion of hyper-Theme which was discussed in the previous section and which, in turn, is the introductory sentence of the paragraph, or the Topic Sentence in rhetoric.

Topic continuity accounts for the coherence of a paragraph, and ultimately, for the internal structuring of a text, since a text consists of a series of paragraphs. If a different idea which has no connection with the main one occurs in the paragraph, one may speak of topic shift or topic discontinuity, which leads to paragraph incoherence; in order to avoid this, the new idea should be developed, exemplified, or illustrated in a different paragraph. Since a text consists of a series of paragraphs which express different ideas, it is inevitable for its internal building and structuring to make use of both continuity and discontinuity.

2.1.3 Given vs. new information

As shown before, the sender may choose from different techniques to make certain parts of the message the *topic* of what he or she is saying, and others simply *comments* about that topic or Theme. These, however, are not the only interpretations of the information structure of discourse. Another interpretation which accounts for the two-part structure of each clause is the *given-new* approach. Despite the difference of terminology, this approach also agrees that the clause has a bi-partite structure but the ordering of information within that structure is determined by the sender's hypotheses about what the receiver does and does not know. With this interpretation, information can be divided into two types (that which the sender thinks the receiver already knows, and that which the sender thinks the receiver does not yet know) and these two types can be labelled *given* information and *new* information, respectively. Similarly to Theme and topic, there are two basic ideas about given information: (1) given information represents a referent shared in some way by speaker and listener, or (2) given information is a cognitively activated referent.

Mathesius (1942) analyses how information is packaged in discourse and what resources are available to speakers and writers for indicating to their addressees the status of information which is introduced in discourse. He suggests that one portion of the utterance represents information that is assumed to be possessed by the listener from the preceding context or may be inferred by him or her from the context. Such information is *known (old, given)* information. It is contrasted with the portion of the utterance which the speaker presents as *new (unknown)* information and which is the content of the utterance. Mathesius examines how this status of information is signalled via strategies such as word order, intonation, and other constructions.

Many of the insights developed by Mathesius and other Prague School scholars were first brought to the attention of Western scholars by Halliday (1967), who elaborated and developed those aspects of the Prague school which related to the structure of texts. Particularly, he was concerned with relating each unit of information in a given sentence to the preceding discourse. He also drew a distinction between *given* information and *new* information. New information represents information the speaker treats as not known to the listener. Given information represents information the speaker treats as known to the listener. Thus, according to Halliday, in example 2.3,

> (2.3)   A sedan is *a car* which usually has *a fully enclosed passenger compartment, a permanent roof, two or four doors, front and rear seats, and a separate trunk.*
>
> (Encarta World English Dictionary)

car is given or old information because it refers to a concept previously introduced – in example 2.3 *sedan* – whereas the rest of the sentence, not know to the reader, *a fully enclosed passenger compartment, a permanent roof, two or four doors, front and rear seats, and a separate trunk* is new information. The given-new principle is best at work in longer units of discourse, where old or given information repeats a concept previously mentioned in discourse ensuring thus the coherence of that unit of discourse. Using the given-new principle, the coherence of a discourse at episodic and global levels is exemplified and analysed in example 2.4,

(2.4) *1* *American football* is **a body-contact, 11-man team sport**
    played with *an oval ball on a rectangular field with goalposts at each end.*
    Each team tries to score *points*
    by carrying or passing (to a **team-mate**) **the ball**
  5 over **the opponent's goal line** for *a touchdown*
    or by kicking **the ball**
    between **the goalposts** for *a field goal.*
    A team must advance **the ball** *10 yards* (**9.1 metres**)
    in *four attempts*, called **downs**,
  10 or turn *possession* over to **the opponent**.
    As **the possession of the ball** changes from **side to side**,
    defensive and offensive teams alternate **positions on the field**.
    This gridiron sport, as **it** is called because of the **field's markings**,
    is indigenous to **the United States**
  15 and has not been taken up in *the rest of the world*
    to the degree that *other American sports*,
    such as **basketball and baseball**, have.
    It has spread to some **other countries**, however,
    and has achieved a degree of *international popularity*
  20 through *television viewing*.
      (*American* Football, in Encyclopaedia Britannica, 2004 CD-ROM edition)

where the bold-faced NPs are generally taken to be given information and the italicized NPs new information. Overall, this paragraph describes American football. Several of the clauses exhibit straightforward cases of given information as well as straightforward cases of new information. The bold-faced NPs in 1, 3, 4, 6, 8, 9, 10, 11, 12, 13, 14, 17, and 18 represent given information because they have been mentioned before in the text. The italicized NPs in 1, 2, 3, 5, 7, 8, 9, 10, 15, 16, 19, and 20 represent new information, for they have just been introduced. Other cases are a bit less clear. The NPs in 5, 7, and 11(2) are marked bold, and their putative status as given must be related to knowledge shared by writer and reader about American football fields.

One important problem in accounting for reference management in discourse is understanding how speaker and listener keep track of referents during discourse production and comprehension. According to Tomlin (1997, 80), keeping track of referents involves three related problems: (1) introducing referents to the discourse, (2) sustaining reference once a referent has been introduced, and (3) reintroducing referents after a long hiatus. In example 2.4, the newly-introduced referent *American football* is sustained through anaphoric forms such as noun phrases (*team sport*) or pronouns (*it*), or is reintroduced (elliptically) after a long pause in the paragraph. Therefore, a

unit of new information (*American football* in the previous example), may change its status as the discourse proceeds, and what was new in one sentence becomes given in the next, precisely because it has just been said. This given-new dichotomy is essential in the process structuring information to achieve unity and coherence in discourse. Cook (1993) elaborates on the topic, and states that

> communication might be defined as the conversion of new information into given information; and a successful communicator as a person who correctly assesses the state of knowledge of his or her interlocutor. If we misjudge, and treat what is given as new, we will be boring; in the reverse case when we assume the new to be given, we will be incomprehensible. (Cook 1993, 46)

2.1.4 Focus

Like Theme or Topic, the term *focus* has also been used in recent linguistics in differing, albeit to some extent related. Focus is used to refer to the newness of particular pieces of information in the utterance, a position traditionally linked with the Prague School's notion of Rheme. Focus is also related to the cognitive notion of prominence or salience being referred to as "object of attention." Finally, focus may be regarded as the addressor's attempt to get the addressee to replace some incorrect information with the correct information, aspect which is usually referred to in terms of contrast.

In a similar way to that in which Halliday (1967) uses the given-new dichotomy, he states that in communication, "the speaker marks out a part (which may be the whole) of a message block as that which he wishes to be interpreted as informative" (1967, 204). In his approach, it is suggested that "the focus of a message ... is that which is represented by the speaker as being new, textually (and situationally) non-derivable information" (Halliday 1967, 204).

The second aspect of Focus is related to the cognitive notion of prominence or salience. In this case, focus is not represented by the Rheme of a sentence but rather by the information that "stands out" from other information and which can be placed anywhere in a sentence in order to achieve a particular effect or to draw the attention of the addressee. Erteschik-Shir (1997), one of the exponents of this theory, states that "the focus of a sentence S = the (intention of a) constituent c of S which the speaker intends to direct the attention of his/her hearer(s), by uttering S." (Erteschik-Shir 1997, 11). For Erteschik-Shir, a sentence in discourse has only one main focus which is assigned to a syntactic constituent which may be a noun phrase, a verb phrase, or the whole sentence. Dik (1981) was also concerned with what the focus of a sentence is, but, unlike Erteschik-Shir, he states that focus may be assigned to more than one constituent. Thus, the answer to the question "Who ate what in the restaurant?" must have two referents which act as an object of attention.

Last but not least, Dik (1981) defines focus as the pragmatic function which characterises constituents which present – relatively speaking – the most important information with respect to the pragmatic information of the speaker and the addressee. Dik offers a very useful and detailed discussion of focus as contrast and identifies five different sub-types of focus phenomena

found in language, each sub-type having a special function. All five sub-types of focus identified by Dik are contrastive since, according to him, "the speaker tries to put some piece of information in opposition to other information, either explicitly or implicitly" (1981, 149).

2.1.5 Foregrounding vs. backgrounding

Like the notions discussed so far, the idea of *foreground* versus *background* information in discourse arises from attempts to explain structural alternations in language for which no obvious semantic explanations are apparent. The major contributions made to the theory of foregrounding find their roots in Russian Formalism; they were elaborated in Prague Structuralism and later on developed by Western linguists (Grimes 1975; Levinson 1983; van Peer 1986; Tomlin 1986; Vallduví 1992) in extended studies. This theory, like focus theory, implies that information in a sentence is structured in two main parts: the foreground – the focused information, the most important information – and the background – the rest of the sentence.

One way of foregrounding information is to alter the usual word order in a sentence. The closer to the beginning or to the end of a sentence the information is placed, the more foregrounded it is because it is in those positions that information is most likely to be noticed or to be more prominent to the readers. This strategy of moving information to the first or last position in a sentence is not the only one through which information can be foregrounded. The same effect can be obtained by preserving the normal ordering and by adding emphasizers to particular bits of information in the sentence.

The foreground/background theory can also be used to account for the structure of larger units of discourse. Hopper (1979) uses the term *foreground* to refer to the parts of the narrative which relate events belonging to the skeletal structure of the discourse, and *background* to refer to the supportive material which does not itself narrate the main event. To support his claim, Hopper (1979, 214) uses a translated text from Swahili of a nineteenth century traveller's tale:

> (2.5)   We returned to the camp, and ran away during the night, and we travelled for several days, we passed through several villages, and in all of them we did not have to pay tribute.

The distinction between foregrounded and backgrounded sentences becomes clear if they are arranged in a chart, with the chronology of events running from left to right, and events not on the main route indicated by a subroutine to the side:
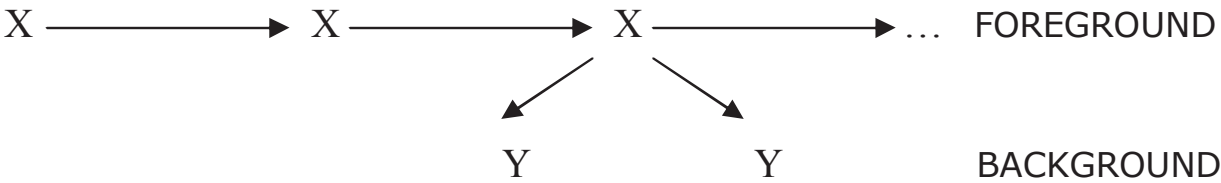

Fig. 2.4   Chronologic representation of events

The notions of foregrounding and backgrounding have led many researchers to consider them as crucial factors in structuring information because, according to them, the way a discourse is structured has a prominent effect on the process of discourse interpretation.

3. Relevance of information theories for translation

The notions discussed so far, even if varied in terminology, assume that information is usually structured at sentence level into two parts, one part which contains more prominent or salient information usually foregrounded, focused or staged, and another part which is backgrounded or which contains information about the first part. In the same way, at higher levels of discourse, certain information seems more central or important to the development of discourse than other.

One of the characteristics of coherent discourse is that entities, once introduced at a given point in the text, are often referred to again at a later point (Tomlin 1997, 77). As shown before, the problem of how reference is managed in discourse production and comprehension / interpretation has been the focus of considerable research on discourse because it is fundamental in accounting for its structure, development, and coherence; since translators are concerned with both discourse interpretation and production, they should be the main beneficiaries of such research.

To probe the relevance of information theories for translation, only the theories which analyse information at levels higher than sentence level are tested here: Thematic progression – Theme as starting point (Daneš 1974, Halliday 1994), topic continuity and topic structuring (Givon 1983), given information as shared information, (Halliday 1967), foregrounding and backgrounding (Hopper 1979). This decision is justified by the general premise that the translator can take the most appropriate decisions only when faced with a holistic view of the text to be translated or analysed. It is therefore assumed here that the other approaches, even if useful for the translator at sentence level, since they are not developed into approaches which look at information at higher text levels, cannot help the translator take overall decisions regarding aspects of information structure, focus, coherence, etc.

To exemplify Daneš's (1974) thematic progression the following text and its translation into English has been selected from a bilingual magazine:

(3.1) *Clubul Investitorilor*	*The Investors' Club*
La finele anului trecut, la inițiativa conducerii Comisiei Economice a Senatului--respectiv a președintelui Dan Mircea Popescu și a secretarului general Ion Zara, s-a înființat Clubul investitorilor. Asociație nonguvernamentală, cu personalitate juridică și apolitică, clubul se bucură de toate drepturile conferite de lege pentru promovarea unui climat favorabil afacerilor și investițiilor în România.	The Investors Club was founded at the end of last year, at the initiative of Dan Mircea Popescu and Ion Zara, president and general secretary respectively of the Senate Economic Commission. A non-governmental, non-party political association, the club enjoys all the rights accorded by the law for the promotion of a climate favourable to business people and investors in Romania.

(*Timisoara--What, Where, When*, 2002: 7)

The following Themes and Rhemes can be identified using Halliday's (1994) Theme-Rheme principle, i.e., Theme as the starting point of the utterance:

Clubul Investitorilor (ST)		The Investors' Club (TT)	
T1 -- *La finele anului trecut, la inițiativa conducerii Comisiei Economice a Senatului - respectiv a președintelui Dan Mircea Popescu și a secretarului general Ion Zara*	**R1** -- s-a înființat Clubul investitorilor	**T1** -- *The Investors Club*	**R1** -- was founded at the end of last year, at the initiative of Dan Mircea Popescu and Ion Zara, president and general secretary respectively of the Senate Economic Commission.
T2 -- *Asociație nonguvernamentală, cu personalitate juridică și apolitică, clubul*	**R2 --** se bucură de toate drepturile conferite de lege pentru promovarea unui climat favorabil afacerilor și investițiilor în România.	**T2** -- *A non-governmental, non-party political association the club*	**R2** -- enjoys all the rights accorded by the law for the promotion of a climate favourable to business people and investors in Romania.

Fig. 3.1   Exemplification of Theme-Rheme identification – Theme as starting point (Halliday 1994)

Using Daneš' (1974) approach to Theme-Rheme analysis at the level of the whole text, the following thematic representations can be obtained:
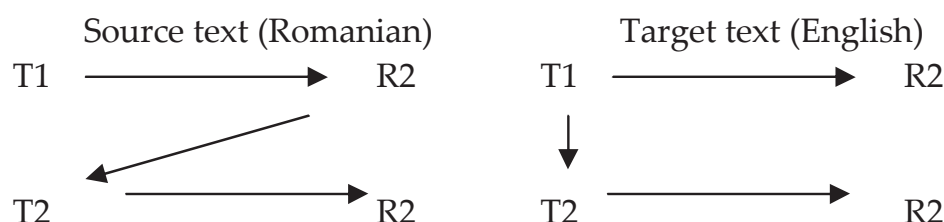

Fig. 3.2   Exemplification of a thematic progression in a source text and of its re-creation in a target text

Even if in both texts the macro-Themes (which in this case are *The Investors' Club*) predict what the texts will be about, the difference between the thematic organizations of each text is obvious. The "zig-zag" pattern in the ST, typical for linear progression, places the importance on *when* and *who* set up the club (a fact which might be of importance in a wider context), while the "parallel" pattern in the TT displays a progression with constant Theme and places the importance on the club itself. This shift in point of view gives a subjective flavour to the TT, and it can even be regarded as a way of manipulation.

Such analyses are useful for translators in that they can (a) identify the thematic structure of the ST and decide whether to preserve it or change it in the TT according to the requirements of the client, of the target language preferences for particular structures, etc., (b) label a source text as coherent or incoherent and if necessary, make the appropriate changes in the target text, etc. Translation studies researchers can also apply this approach to information structure to draw comparative structural analyses between parallel texts and establish source or target language

preferences for particular thematic structures. Last but not least, genre analysts can use it to establish genres' preferences for particular thematic structures.

The following example has been selected to see how and if *thematic progression* can work at higher text levels:

(3.2)   Rob Roy
Rob Roy, real name Robert MacGregor (1671-1734), Scottish brigand, [*was*] sometimes called the Scottish Robin Hood. Known as Rob Roy, or Robert the Red, because of his red hair, he was a member of the outlawed Scottish clan Gregor. After his lands were confiscated in 1712 by James Montrose, 1st duke of Montrose, to whom he was in debt, Rob Roy became a leader of uprisings and a freebooter. He later lived in peace for a time under the protection of John Campbell, 2nd duke of Argyll, taking the name Robert Campbell, but he was imprisoned by the English in the 1720s. The Scottish novelist Sir Walter Scott turned the brigand into a romantic hero in his popular novel *Rob Roy* (1818). In 1995 a movie about his struggle was released called *Rob Roy.*

(Rob Roy, *Encarta® World English Dictionary* © & (P) 1999, 2000 Microsoft Corporation)

The underlined Themes in example 3.2 have been identified using the same Theme as starting point principle. The Themes are presented in the order of their appearance in Fig. 3.3. In the same figure, the thematic progression containing the links between the Themes and Rhemes is illustrated:
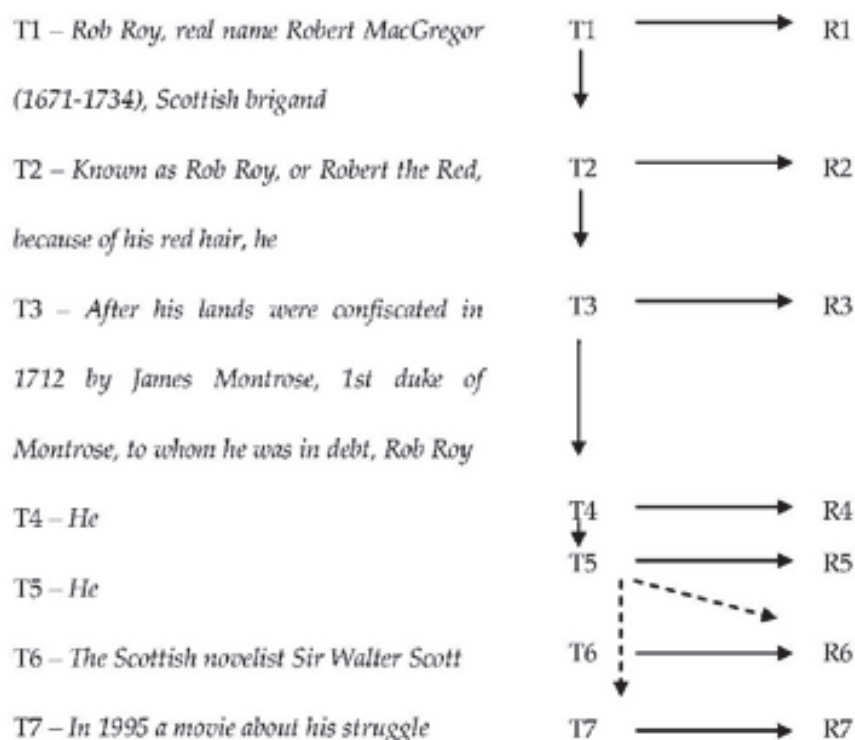

Fig. 3.3   *Exemplification of Theme identification in a text and of its thematic progression (Theme as starting point)*

At a first sight, the paragraph reads well and does not seem to cause any translation problems.

When trying to identify the thematic progression of the text (Fig. 3.3) using Theme as starting point, the translator might encounter a problem when establishing the Theme in sentence 3.

The problem is generated by the fact that it is difficult to establish where the Theme ends and where the Rheme begins: (1) if one considers only the beginning of the sentence, i.e., "after his lands", as Theme, it would be difficult to establish the progression and the links with the remaining text; (2) if one takes T3 as Theme, this problem would be solved, but one would immediately notice that the Theme contains rhematic information in itself, i.e., information about Rob Roy, e.g., that his lands were confiscated by in 1712 by James Montrose, 1st duke of Montrose, and that he was in debt to the duke, information which, according to this approach, follows the Theme.

Another problem is that, if using the same approach to Theme-Rheme analysis, it is difficult to establish links between (T5) and the remaining text, represented with dotted arrows in Fig. 3.3, respectively (T6), (R6), (T7), (R7), which can be considered as new information, indirectly related to the hyper-Theme of the text, i.e. Rob Roy. An indirect implication of this problem is that, from this analysis, the paragraph might be considered incoherent.

The above discussion and examples prove that not all the information that comes at the beginning of a sentence is necessarily thematic, and that the identification of Themes and Rhemes of a text from a syntactic perspective, which takes into account the leftmost information unit in a sentence as being thematic, can be inaccurate. Also, from the same perspective, it is not always easy to establish links between Themes and Rhemes and account for text coherence.

To conclude, if the syntactic identification and analysis of Themes and Rhemes can sometimes work on smaller or simpler texts in terms of sentence structure and organization, this is not the case with longer units displaying complex sentences, information shifts, voice changes, etc.

As shown in section 2.1.2, Givon (1983) discusses *topic continuity* in discourse in terms of *thematic continuity, action continuity,* and *topic/participant continuity*. Givon uses the same parameters for Theme identification as the proponents of the approach who see Theme as starting point of the sentence: this means that when applied to example (3.2), the thematic continuity would display the same results and problems as presented above. What topic continuity brings as new is the analysis of *action continuity* of events which facilitates text understanding and proves the chronological coherence of a text. In example (3.2), the action continuity can be represented as follows:

1671 – 1734	(life of) Robert MacGregor
↓	
1712	Lands confiscated by James Montrose
↓	
1720	Imprisoned by the English
↓	
1818	Novel about his life
↓	
1995	Movie about his struggle

Fig. 3.4   Representation of action continuity in discourse

The analysis of *participant continuity* in the same example displays the following results:

Rob Roy, real name Robert MacGregor (1671-1734), Scottish brigand, [was] sometimes called the Scottish Robin Hood. Known as Rob Roy, or Robert the Red, because of his red hair, he was a member of the outlawed Scottish clan Gregor. After his lands were confiscated in 1712 by James Montrose, 1st duke of Montrose, to whom he was in debt, Rob Roy became a leader of uprisings and a freebooter. He later lived in peace for a time under the protection of John Campbell, 2nd duke of Argyll, taking the name Robert Campbell, but he was imprisoned by the English in the 1720s. The Scottish novelist Sir Walter Scott turned the brigand into a romantic hero in his popular novel *Rob Roy* (1818). In 1995 a movie about his struggle was released called *Rob Roy*.
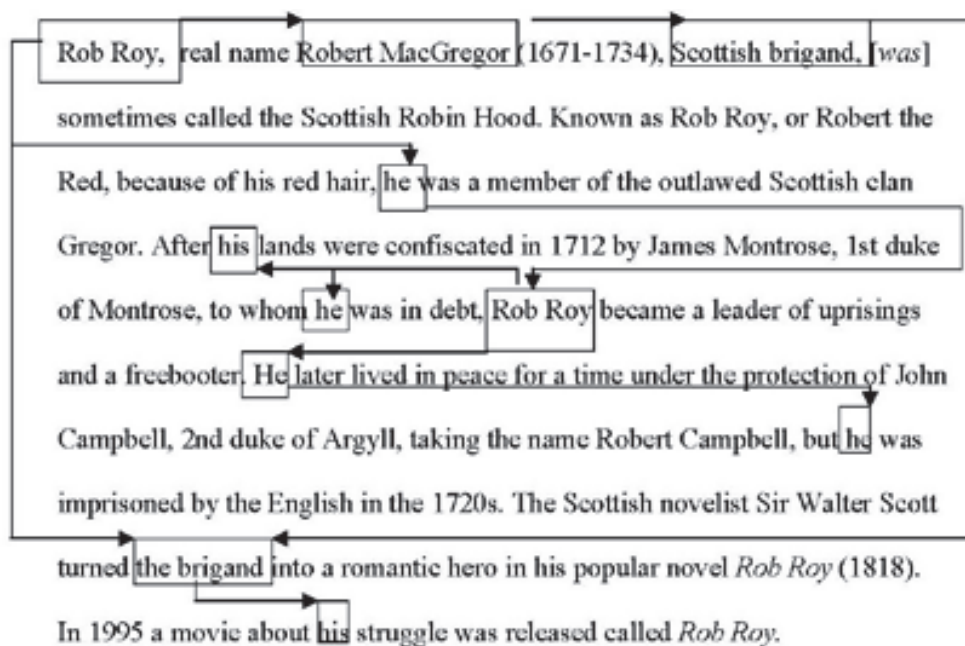
Fig. 3.5 Representation of participant continuity in discourse

As figures 3.4 and 3.5 show, the text is coherent from the viewpoint of chronology of events and identity of referents. Analyses of action continuity and participant continuity may prove to be extremely helpful for the translator when accounting for and preserving source text coherence in the target text.

The advantage of using the *given-new* analysis, at whole text level is that it can account for text coherence; in the following representation of example (3.1), the elements in *italic* are New, and the elements in **bold** are Given:
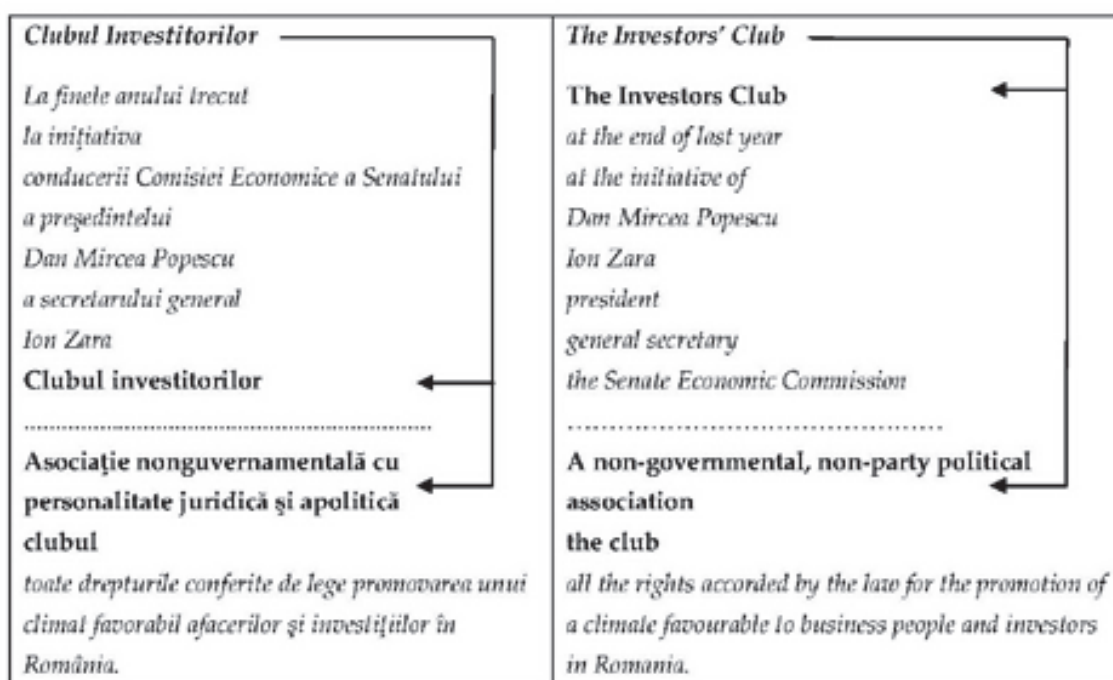

Clubul Investitorilor	The Investors' Club
La finele anului trecut	**The Investors Club**
la inițiativa	*at the end of last year*
conducerii Comisiei Economice a Senatului	*at the initiative of*
a președintelui	*Dan Mircea Popescu*
Dan Mircea Popescu	*Ion Zara*
a secretarului general	*president*
Ion Zara	*general secretary*
Clubul investitorilor	*the Senate Economic Commission*
....	....
Asociație nonguvernamentală cu personalitate juridică și apolitică	**A non-governmental, non-party political association**
clubul	**the club**
toate drepturile conferite de lege promovarea unui climat favorabil afacerilor și investițiilor în România.	*all the rights accorded by the law for the promotion of a climate favourable to business people and investors in Romania.*

*Fig. 3.6   Identification and analysis of **given-new** elements in a text*

Fig. 3.6 shows that the given and the new elements are the same at all levels in both texts, a fact which analytically proves that the target text is as coherent as the source text. However, what the given-new approach brings as new, as compared to the previous approaches, is that, in identifying the given and new elements, it offers a perspective on the word order in the two texts: preserving in the target text the same given or new information as in the source text, but ordered differently, may also constitute a subject of study for translation.

The last approach analysed here at whole text level is foregrounding-backgrounding. The analysis of sequentiality in narrative discourse proposed by Hopper (1979) (Fig. 2.4) is very much identical with the concept of action discourse proposed by Givon (1983) (Fig. 3.4). What differs in Hopper's perspective is that actions can also be presented in the background, usually subordinated to those in the foreground. The relevance for translation is that both approaches account for chronology of events, and implicitly logical coherence of discourse.

4. Conclusion

The aim of this paper was to present an overview of existing linguistic theories and approaches to information identification and analysis in texts and to probe their relevance for translation. Several theories and approaches which use different parameters for the identification and analysis of information in texts, i.e. syntactic, semantic, pragmatic, were covered, and eventually, their relevance to translation was tested.

As to their relevance for translation, several conclusions can be drawn. On the one hand, knowledge of information theories may prove extremely helpful for professional translators in that by applying them, they can

- identify the information structure of the source text and decide whether to preserve it or change it in the target text according to the requirements of the client, of the target language preferences for particular structures, etc;
- label a source text as coherent or incoherent and if necessary, make the appropriate changes in the target text, etc;
- identify the chronology of events and identity of referents which leads to understanding the source text better;

The information theories presented here can also be of interest to translation scholars, in that they can

- apply various approaches to information structure to draw comparative structural analyses between parallel texts and establish source or target language preferences for particular structures;
- establish genres' preferences for particular information structures.

On the other hand, the analyses performed here showed that

- not all of the theories focusing on the identification of information can be applied to the analysis of information on larger text segments; they work only on isolated excerpts of texts,

e.g. at local or sentence level, and due to this limitation, they cannot offer the translator a holistic view needed for text understanding in terms of information structure and relations;

- most of the parameters used by their proponents are semantic or syntactic in nature and cannot provide consistent identification of salient information;
- they cannot always account for text coherence on all levels (local, episodic, global);
- none of these theories, as far as we know, have been integrated into a translation method. Most of the studies dealing with information theories in translation were comparative in nature and focused on differences between language structure, ordering of information, etc. in the source text and in the target text. This can be a serious limitation especially to translation trainees.

Clearly, information theories can play a decisive role in the translation process, especially in the stages of source text understanding and target text re-creation. However, further research is necessary to polish these theories and eventually validate their relevance for translation.

Bibliography

Baker, M. 1992. 1994. *In other Words.* London: Routledge.

Brown, G., and G. Yule. 1993. *Discourse Analysis.* Cambridge: Cambridge University Press.

Chafe, W. 1976. Givenness, Contrastiveness, Definiteness, Topics and Point of View. In *Subject and Topic*, ed. C. Li, 25–55. New York: Academic Press.

Cook, G. 1993. *Discourse.* Oxford: Oxford University Press.

Daneš, F. 1974. Functional sentence perspective and the organization of the text. In *Papers on Functional sentence Perspective*, ed. F. Daneš, 106–28. The Hague: Mouton.

Dejica, D. 2006. Towards a New Model for the Identification of Thematic Information in Discourse. In *Comunicare Profesională şi Traductologie, Lucrările Conferinţei Internaţionale. Timişoara. 29–30 septembrie 2005*, eds. R. Superceanu and D. Dejica, 103–10. Timişoara: Editura Politehnica.

– – –. 2008a. Identification and Analysis of Theme-Rheme Relations for Discourse Production and Translation. *Scientific Bulletin of Politehnica University of Timisoara. Transactions on Modern Languages* 7(1/2): 35–41.

– – –. 2008b. Using hol-atomistic and holistic lexical and grammatical relations in translation. In *Professional Communication and Translation Studies* 1: 143–7.

Dejica, D., and R. Superceanu. 2004. Thematic-Structure Analysis of the Section Statement of Problem in Project Proposals. In *Scientific Bulletin of Politehnica University of Timisoara. Transactions on Modern Languages* 3(1/2): 45–52.

Dik, S. 1981. 1989. *Functional Grammar.* Dordrecht: Foris Publications.

Erteschik-Shir, N. 1997. *The Dynamics of Focus Structure*. Cambridge: Cambridge University Press.

Firbas, J. 1964. On Defining the Theme in Functional Sentence Perspective. In *Travaux linguistiques de Prague* 1: 267–80.

– – –. 1992. *Functional Sentence Perspective in Written and Spoken Communication*. Cambridge: Cambridge University Press.

Fortesque, M., P. Harder and L. Kristoffersen, eds. 1992. *Layered Structure and reference in a functional perspective.* Amsterdam: John Benjamins.

Gerzymisch-Arbogast, H. 2005. Introducing Multidimensional Translation. In *MuTra: Challenges of Multidimensional Translation.* Proceedings of the Marie Curie Euroconferences – Saarbrücken 2–6 May 2005, eds. H. Gerzymisch-Arbogast, and S. Nauert. *http://www.euroconferences.info/ proceedings/2005_Proceedings /2005_proceedings.html (*accessed 1 March 2008).

Gerzymisch-Arbogast, H., J. Kunold and D. Rothfuß-Bastian. 2006. Coherence, Theme/Rheme, Isotopy: Complementary concepts in text and translation. In *Text and translation. Theory and methodology of translation.* Jahrbuch 6, 2005/2006 Übersetzen und Dolmetschen, eds. C.K. Heine, C. Schubert and H. Gerzymisch-Arbogast, 349–70. Tübingen: Narr.

Givon, T. 1983. *Topic Continuity in Discourse: A Quantitative Cross-Language Study ¸* Amsterdam: John Benjamins.

Grimes, J. 1975. *The Thread of Discourse*. The Hague: Mouton.

Gundel, J. 1974. The Role of Topic and Comment in Linguistic Theory. Ph.D. diss., University of Texas, Austin.

Halliday, M.A.K. 1967. Notes on transitivity and Theme in English. Part II. In *Journal of Linguistics* 3: 199–244.

– – –. 1994. *An Introduction to Functional Grammar.* London: Edward Arnold.

Hopper, P.J. 1979. Aspect and Foregrounding in Discourse. In *Discourse and Syntax,* ed. T. Givon, 213–42. New York: Academic Press.

Lambrecht, K. 1994. *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*. Cambridge: Cambridge University Press.

Levinson, S.C. 1983. 1992. *Pragmatics.* Cambridge: Cambridge University Press.

Martin, J. R. 1992. *English Text: System and Structure*. Amsterdam: John Benjamins.

Mathesius, V. 1942. From co-operative word studies. In *CMS,* 28.

Mudersbach, K. 2008. A New Theme in the Given Theme-Rheme Debate. Translated by H. Gerzymisch-Arbogast. Saarbrücken. Hildesheim: Olms Verlag.

Nida, E.A. 1997. The Principles of Discourse Structure and Content in Relation to Translating. In *Transferre Necesse Est*, eds. K. Klaudy and J. Kohn, 37-42. Budapest: Scholastica.

Sgall, P. 1987. Prague functionalism and topic vs. Focus. In *Functionalism in Linguistics*, eds. R. Dirven and V. Fried, 169–90. Amsterdam: John Benjamins.

Siewierska, A. 1991. *Functional Grammar.* London: Routledge.

Tomlin, R. 1986. *Basic Word Order: Functional Principles.* London: Croom Helm.

Tomlin, R., L. Forrest, M.M. Pu and M.H. Kim. 1997. Discourse Semantics. In *Discourse as Structure and Process*, ed. T.A. van Dijk, 63–111. London: SAGE.

Vallduví, E. 1992. *The Informational Component.* New York: Garland.

Vallduví, E., and E. Engdhal. 1996. The Linguistic Realization of Information Packaging. In *Linguistics* 34: 459–519.

van Dijk, T.A. 1982. *Text and Context.* London: Longman.

van Peer, W. 1986*. Stylistics and Psychology: Investigations of Foregrounding*. London: Croom Helm.