SENTENCE-INITIAL INFORMATION STRUCTURE IN ENGLISH AND ROMANIAN: A PRAGMATIC APPROACH

Silvia Florea, "Lucian Blaga" University of Sibiu

Abstract: My paper sets out to explore various aspects pertaining to how information structure may be constrained, determined and correlates with differences in meaning in English and Romanian. It argues that from a pragmatic point of view, accounts of sentential information structure correspond to truth-conditional effects and that topic-comment representation in both languages is reflective of discourse unbundling that serves specific communicative needs.

Keywords: information, sentence structure, pragmatic, topic, comment.

The last decades have witnessed an increasing interest in the unifying approaches to pragmatic phenomena, particularly in the development of overarching logical theoretic frameworks of information structure. Newly arising research platforms on the interaction of discourse coherence, implicatures and information structure have underscored the architectural interface of several fields, in particular of linguistics and philosophy: information structure, in terms of the ways in which the informational architecture (rheme vs theme; topic vs focus, given vs new, etc.) is reflected in the sentence structure; the semantics of discourse with regard to the way meanings of sentences contribute to text cohesion and dialogue; and the pragmatic inferences (implicatures) generated by the speaker's rationality and cooperativeness that contribute to the richness of sentence meaning. More specifically, the study of implicatures offers significant functional explanations of linguistic facts, produces simplifications in the structure (and content) of semantic descriptions and last but not least, provides explicit explanation and evidence of how it is possible to mean (and be understood accordingly) more than is said.

Several attempts have been thus made at subsuming conversational implicatures to semantics (Chierchia, 2004), at uniformly explaining information structure, presuppositions, and conventional implicatures (Simons et al., 2011), at exploring discourse theories in terms of segmented discourse representation theory (Asher and Lascarides, 2003), and game theoretic pragmatics (Benz et al., 2006), including the Optimality Theory recent research proposed by Benz and Mattausch (2011).

Along such theoretical framework approaches, a considerable number of experimental studies have pointed out that constituents such as the information structure and implicatures are closely interdependent. Starting with the Prague School theorists, Halliday (1967), and Vallduví (1992;1993) the theory of information structure has been described at sentence-level as a variation of sentential structure operating along certain parameters so as to regulate the presentation of the information in a way that relates that information to the prior context structure. The relationship is thus characterized by the primitive functional roles of theme/rheme, focus/ (back) ground, topic/link, old/new, etc. These primitives encapsulate the correlated sentence information structure and help explain the roles of particular syntactic focus constructions, of intonational focus, ellipsis (Winkler 2005), of topicalization and/or other

BDD-V604 © 2014 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:55:23 UTC)

displacement transformations, including of a wide range of other phenomena such as word order, the functionality and representations of distinguished structural positions, definiteness, specificity, use of affixes and negation, etc.

However, information structure is typically taken to be a pragmatic phenomenon, with primitives that represent pragmatic categories of the proposition and that can be defined on the basis of pragmatic information interpretation established in the unifying management sense of the Common Ground (Lambrecht 1994; Roberts 1996). Hence, the concept of 'topic' may be approached from several viewpoints: of *aboutness* (Gundel 1988; Lambrecht 1994); *addressation* (Reinhart 1981); *presupposition* (Roberts 1996) and/or *givenness* (Katz and Selkirk (2011). Similarly, the notion of 'focus' is generally interpreted across various property of information theories that differentiate between proposition and presupposition (Lambrecht 1994) or definiteness and givenness (Erteschik-Shir 2007). These approaches rest on the pragmatic properties and functions that the chunk of information has in a context, and although, most often than not, they seem to reduce the meaning component to a grammatical concept of processing instructions they nonetheless rely on pragmatic features. In other words, language specific studies explore how primitive pragmatic categories are expressed across languages, by explaining the aspect of form representing the assumed form–meaning relation.

According to Winkler (2005), this is reflected even in generative grammar where information structure involves two distinct cycles: a functional and a grammatical one. The first cycle constitutes in fact the default cycle that automatically derives the information focus (IF) of the phase that is handed on to LF (the syntax-semantics interface) which is responsible for the interpretation of information structural and discourse notions and which relates to pragmatics. The smallest phase represents the domain of IF. Cycle 1 determines whether a phase needs to be sent on to the second cycle on the basis of the presence of absence of formal features while the second cycle checks the grammatical status of the phase constituents. In so doing, it identifies the non-interpretable features that lead to syntactic displacement and word order variation. In English, cycle 2 generally interprets displaced constituents to a sentence initial position as either Contrastive Focus (CF) or Contrastive Topic (CT).Therefore the first cycle typically derives unmarked sentence structures, whereas the second cycle derives marked sentence structures. Such a derivational history (that is whether a phase passes the first cycle without being sent on to the second cycle) effects a direct impact on SSI (Surface Semantic Interpretation) a subcomponent of LF that is responsible for the interpretation of syntactic displacement.

Against this background, the present paper argues that both topic and focus, often termed as given (topic) versus new (focus), are delineated in natural languages across the whole spectrum of syntax, phonology, semantics, and pragmatics (still mostly controversial of them all), and that there is a high tendency, particularly in unmarked sentences, to map the contextually bound stretch of the sentence on to the subject and the contextually unbound stretch on to the predicate. With further exemplifications from English and Romanian we will point to various emerging issues of linearity arising mostly from the fact that the majority of world languages are either SVO or SOV.

By general acceptance, the sentence-level message is built along prototypical lines as a balanced representation of the given and the new, with a climax in the form of a focal point of

BDD-V604 © 2014 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:55:23 UTC)

information that usually comes in sentence-final position. However, there is no reason why any contextually bound (given/less dynamic/presupposed) part of the sentence should come first. Different approaches have examined the topic property of a constituent in terms of its occurrence in sentence-initial position, a generalization that can be nonetheless questioned. For linear considerations, only constituents that convey grammatical functions can have a topic function. Thus, languages may have topics that do not always occur in initial sentence position, such as circumstantial elements:

"Yesterday Jane was in a bad mood". (1)

In this sentence, the topic is not *yesterday*, rather it is the constituent that has the subject function. However, the first position may differ from topic position, although the first position is generally the topic position. Cheng (2012) exemplifies the differences encoded in the information structure provided by the pattern "Today Jean is playing" as compared to "Jean is playing today":

Question: When is Jean playing? (2)
Answer: Jean is playing // today. (3)
PRESUPPOSED FOCUS
Answer: ? Today // Jean is playing. (4)
FOCUS PRESUPPOSED
Question: Is Jean studying today? (5)
Answer: No, today Jean // is playing. (6)
PRESUPPOSED FOCUS

Although the natural position for focused information in English is sentence final, occasional deviations from this rule are allowed and, as in this case, may be signaled by a high falling pitch emphatic stress on the constituent.

As well, English cleft-sentences are commonly considered to be patterns of a sub-category of information structure characterized by exhaustivity (Kiss 1998, Büring and Križ 2013); however, its encoding variation of information structure is characterized by discourse features that determine the distribution of clefts (Delin and Oberlander 1995). The relative clauses in it-cleft constructions introduce information that is already known, but is not yet embedded in the present discourse context, which is why they typically come as focused information at the end of the sentence. For example:

"Jack wrote an essay" has encoding variation of information structure:
It was Jack who wrote an essay. (7)
FOCUS PRESUPPOSED
It was an essay that Jack wrote. (8)
FOCUS PRESUPPOSED

The relative clauses of wh-clefts (presupposed) appear in different discourse contexts than it-clefts, as they include foregrounded information that comes naturally in the topic position.

The one who wrote an essay was Jack. (9)
PRESUPPOSED FOCUS
What Jack wrote was an essay. (10)
PRESUPPOSED FOCUS

There is a very strong interdependence between the concepts of topic and subject. The topic-comment encoded information structure in English appears to be a bit more complicated by

BDD-V604 © 2014 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:55:23 UTC)

the common sentence initial position for subject nominal. A good way to alternate topic choice is by selecting different subjects.

Water streamed down the alley. (11)
TOPIC COMMENT
The alley streamed with water.
Dirt flowed in the pipes. (12)
TOPIC COMMENT
The pipes flowed with dirt.

Across natural languages, subjects tend to be nominal phrases, however sentential subjects (subjects that would ordinarily be analyzed as sentences) may function as subjects. An interesting point of view is provided by Lohndal (2013) who examines whether sentential subjects are structurally subjects, or sentential subjects are structurally topics.

[That Mary is late] annoys John. (13)

Active and passive voice representations constitute one of the most common information-structure encoding devices in English. Within topic-comment information structure, extra-positioning is a topic-altering device:

That Mary will come is sure. (14)
TOPIC COMMENT
Mary is sure to come. (15)
TOPIC COMMENT

On the other hand, in Romanian, which is a pro-drop language (16), sentence-initial information follows regular topic-comment information structure. However, a characteristic feature is that an overt pronoun in a sentence-initial subordinate clause is considered as disjoint from the main clause subject (17a,b) (Diaconescu and Goodluck 2003; Dobrovie-Sorin 1994), but for a sentence-final subordinate clause co-reference with the main clause subject is possible although typically dispreferred (18a,b).

pro mănâncă (16)
sing-PRESENT-3sg
'he eats'
Când *pro* stă, Nicolae nu mănâncă. (17a)
Când el stă, Nicolae nu mănâncă. (17b)
When he rests, Nicolae does not eat.
Nicolae nu mănâncă când *pro* stă. (18a)
Nicolae nu mănâncă când el stă. (18b)
When he rests, Nicolae does not eat.

Another interesting issue relating to topic–comment information structure is posed in Romanian by the syntax and status of the clitic (see Tigău 2014), traditionally considered as weak forms of their stressed pronominal counterparts that they may replace, as in (19):

L-am auzit pe el. (19a)
Him.cl-have.I heard *pe* him.
I have heard him.
L-am văzut. (19b)
Him.cl-have.I heard.
I have heard him.

BDD-V604 © 2014 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:55:23 UTC)

Pe cine ai auzit? (20a)
Pe who have.you heard
'Whom did you hear?'
L-am auzit pe el. (20b)
Him.cl-have.I heard *pe* him.
'I have heard him.
L-am auzit. (20c)
Him.cl-have.I heard.
'I have heard him'

For the purpose of our discussion, in (20b) the stressed form of the pronoun *pe el* represents the information focus, in fact the very new information sought after in (20a). The clitic constitutes known information in the topic partition and explains the discourse-linked character of the *wh*-element *pe cine.* As in (20c) some basic information is apparently missing, therefore (20c) does not represent an answer to (20a), hence it is considered ungrammatical. (20c) makes use of the clitic to convey this information, but this is not possible as the clitic may operate only as a topic making a connection to a presupposed situation.

All the above contexts have excluded phonological discussion contexts where pitch and duration may be manipulated in meaningful and well-defined ways so as to create a contrast between the word(s) under contrastive focus and the structures in pre-and post focal contrasts (Vigario et Al 2009). Typically, topic and subject are very closely connected, however they are simply not isomorphic; there are cases of topics that are not subjects, particularly when they occur in initial sentence position and precede the subjects as the following sentences:

TOPIC COMMENT
Last year// I heard 10 new songs in premiere. (21a)
As for Mary// she is a terrific bore. (21b)
Anul trecut //am auzit 10 piese noi in premieră. (22a)
Cât despre Mary// ea e foarte plictisitoare. (22b)

(21a) and (22a) represent clear cases of ordinary topicalization, while (21b) and (22b) operate as left-dislocations, in which the pronouns in both languages mark the otherwise expected topic nominal position. In English, as much as in Romanian, left-dislocations may be used to express a change in the topic of the discourse, typically to override previous topic (as for). Ordinary topicalization, on the other hand, may prevent as topic any constituent which is presupposed. For example:

Question: What's your favourite singer? (23)
TOPIC COMMENT
Answer: Pink// her voice is wonderful.

By way of conclusion, information structure variation may be critically architectured only by studies at the interface of syntax, pragmatics and computational linguistics. The range of variation across natural languages generally points to information that is (re)construed from the complementary point of view. The message is built along prototypical models in the form of a balance between the given and the new, upsurging in a focal point of information, that ultimately represents the packaging of a particular expression of a language.

BDD-V604 © 2014 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:55:23 UTC)

Bibliography

Asher, N. and A. Lascarides, (2003), *Logics of Conversation*, Cambridge University Press.
Benz, A., G. Jäger, and R. van Rooij, (2006), An Introduction to Game Theory for Linguists, in *Game Theory and Pragmatics* Benz et al., pages 1-82.
Benz, Anton, and Jason Mattausch, (2011), *Bidirectional Optimality Theory, in Linguistics Today 180*, John Benjamins.
Büring, Daniel, and Manuel Kriz, (2013, "It's That, and That's It! Exhaustivity and Homogeneity Presuppositions in Clefts (and Definites)", in *Semantics and Pragmatics* 6: 6:1–29.
Cheng, Shufang, (2012), "A Short Analysis of Information Structures of English Sentences", in *Theory and Practice in Language Studies*, Vol. 2, No. 3, Academy Publisher, pp. 492-497, March.
Chierchia, Gennaro, (2004), 'Scalar Implicatures, Polarity Phenomena, and the Syntax/Pragmatics Interface', in Adriana Belletti (ed.), Structures and Beyond: The Cartography of Syntactic Structures, vol.3, pp.39–103, Oxford: Oxford University Press.
Delin, Judy, and Jon Oberlander, (1995), "Syntactic Constraints on Discourse Structure: The Case of It-Clefts", in *Linguistics* 33 (3).
Diaconescu, Constanţa Rodica and Helen Goodluck (2003), "Structural and Discourse Factors in Romanian L2 English Learners' Interpretation of Pronouns", in Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002), ed. Juana M. Liceras et al., 71-75. Somerville, MA: Cascadilla Proceedings Project.
Dobrovie-Sorin, C., (1994), *The Syntax of Romanian. Comparative Studies in Romanian*, Mouton de Gruyter.
É. Kiss, Katalin, (1998), "Identificational Focus versus Information Focus", in *Language* 74 (2): 245–73.
Erteschik-Shir, Nomi, (2007), *Information Structure: The Syntax-Discourse Interface*, OUP Oxford.
Gundel, Jeanette K. (1988),"Universals of Topic-Comment Structure", in *Studies in Syntactic Typology*, edited by Michael Hammond, Edith A Moravcsik, and Jessica R Wirth, 209–39. Amsterdam/Philadelphia: John Benjamins.
Halliday, M.A.K., (1967), "Notes on Transitivity and Theme in English" (Part 2), in *Journal of Linguistics* 3:199-244.
Kahnemuyipour, Arsalan, (2009), *The Syntax of Sentential Stress*, Oxford Scholarship Online: September 2009.
Katz, Jonah, and Elizabeth Selkirk, (2011), "Contrastive Focus vs. Discourse-New: Evidence from Prosodic Prominence in English", in *Language* 87 (4): 771–816.
Lambrecht, Knud, (1994), *Information Structure and Sentence Form: Topic, Focus, and the Mental Representations of Discourse Referents*, Cambridge: Cambridge University Press.
Lohndal, Terje, (2013), Sentential Subjects in English and Norwegian. *Syntaxe et Sémantique*.
Reinhart, Tanya (1981), "Pragmatics and Linguistics: An Analysis of Sentence Topics in Pragmatics and Philosophy" I, in *Philosophica* 27 (1): 53–94.

BDD-V604 © 2014 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:55:23 UTC)

Roberts, Craige, (1996), "Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics", in *OSUWPL* Volume 49: Papers in Semantics, by Jae Hak Yoon and Andreas Kathol (eds.), The Ohio State University Department of Linguistics.
Simons, Mandy, Judith Tonhauser, David Beaver, and Craige Roberts, (2011), "What Projects and Why", in *Semantics and Linguistic Theory* (SALT) 20, in Nan Li and David Lutz (eds.), 3:309-327, *Language*.
Tigău, A., (), **"**Clitic Doubling and *Pe* Marking with Romanian Indefinite Direct Objects", available at:http://www.unige.ch/lettres/linguistique/enseignantschercheurs/equipesyntaxe/alina-tigau/. Retrieved Dec.01, 2014.
Vallduví, Enric, (1992), *The Informational Component*, Garland Press, New York.
*** (1993), *Information packaging: A survey*, Report of the Word Order, Prosody, and Information Structure Initiative, University of Edinburgh.
Vigário, Marina, Sónia Frota, M. João Freitas, (2009), *Phonetics and Phonology: Interactions and Interrelations*, John Benjamins Publishing.
Winkler, Susanne (eds), (2005), *Ellipsis and Focus in Generative Grammar*, Walter de Gruyter.

BDD-V604 © 2014 Arhipelag XXI Press
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 17:55:23 UTC)