On English Topicalization and Left-Dislocation from an Information-Structural Perspective

Péter Szűcs

The aim of this paper is to investigate the Information-Structural (IS) properties of two English constructions featuring constituents in a non-canonical, left-peripheral position: Topicalization (TOP) and Left-Dislocation (LD). Pulling several research threads together from generative and functional linguistics, it will be argued that seeing these as simple topic-marking devices is a too simplistic approach: in reality, LD marks a subtype of (non-contrastive) topics, Thematic Shifters, while TOP is used for contrastive IS categories: Contrastive Topics (C-TOPIC) and Contrastive Focus (C-FOCUS).

Keywords: *topicalization, left-dislocation, information-structure, English*

1 Introduction

Languages commonly use a variety of methods to express the information-structural (IS) features of a sentence. Besides intonation and certain morphemes (like the Japanese topic marker *wa*), word order variation is one of the prime tools for such procedures. This is even true for English, a so-called "configurational language", which is commonly assumed to have a relatively fixed word-order.

In this paper, I will investigate the information-structural properties of two English structures which utilize word-order variation for such purposes. (1a) and (1b) provide examples for the constructions.

(1) a. Tom, I like.
b. Tom, I like him.

The common name in the literature for the configuration in (1a) is "Topicalization" (abbreviated as TOP henceforth), while (1b) is most commonly called "Left-Dislocation" (abbreviated as LD). Both feature an argument in a non-canonical, left-peripheral position.[1] The obvious difference between the two is that while in TOP, the canonical position of the fronted constituent is empty (or, from a transformationalist viewpoint, occupied by a trace), in LD, it is filled with a coreferential resumptive pronoun.

Both of these constructions are commonly regarded as topic-marking devices (e.g. in Lambrecht 1994 and Dalrymple 2001: 391).[2] Despite the intuitive appeal of this

[1] Following Birner & Ward (1998), the discussion of TOP and LD should be limited to lexically subcategorized elements. Adjuncts can also occur in the left-peripheral position, but their function is more like frame-setting and they occur much more freely than one would expect from topicalized or left-dislocated elements. E.g. (ia) can be discourse initial, unlike (ib) with a genuine TOP.

(i) a. In New York, there's always something to do. (felicitous discourse-initially)
b. #In a basket, I put your clothes. (infelicitous discourse-initially)

[2] The name of TOP, given by Ross (1967), is an unmistakable sign for this.

characterization, not everybody has shared these ideas. There are several functionalist researchers who have called these assumptions into question. For instance, Prince (1999) writes the following about TOP:

> A glance at the literature over the past thirty years shows that this assumption has been maintained by syntacticians as well as by functionalists, although it has never been proven or even, to my knowledge, seriously investigated.

Prince argued in several papers (Prince 1981, 1998, 1999) that rather than being simple topic-marking devices, both TOP and LD may actually have several functions, and marking a topic is crucially *not* one of them. In this paper, I argue that Price's claims are partially correct. The claim that TOP and LD simply mark topics cannot be maintained (especially for TOP), but a more fine-grained view of IS-notions can capture the generalizations about these constructions. Such a view has been emerging in the generative research tradition, so a possible convergence between generativists and functionalists is possible.

Before discussing the details about TOP and LD, it is essential to clarify the basic concepts of information structure, topic and focus. I will do so in the next section.

2 Basic concepts of Information Structure

2.1 Topic

According to the widely accepted generalization "the topic of a sentence is the thing the proposition expressed by the sentence is about" (Lambrecht 1994: 118).[3] From this basic tenet, various constraints follow which delimit what can serve as a topic expression. It is commonly accepted that topics must be at least referential, otherwise they could not serve as targets for a proposition. According to Gundel & Fretheim (2004), this has to do with the definiteness or presupposition effect that topics have. They cite a Japanese example. In (2a) the subject is followed by the nominative marker *ga* and it can be interpreted as either definite or indefinite. In (2b), by contrast, the subject is followed by the topic marker *wa* and it can be interpreted only as definite (and it can also have a generic meaning).

(2) a. *Neko* ***ga*** *kingyo* *o* *ijit-te*… (Japanese)
cat NOM goldfish OBJ play-and
"The cat/A cat is playing with the/a goldfish and…"

b. *Neko* ***wa*** *kingyo* *o* *ijit-te*…
cat TOP goldfish OBJ play-and
"The cat/*A cat is playing with the/a goldfish and…"

An English example for the presupposition effects of topics can be illustrated with the "lie-test" (Erteschik-Shir & Lappin 1979, cited by Lambrecht 1994:52). Consider (3):

[3] Note that this definition does not include "discourse topics", as its scope is limited to the sentence-level.

(3) a. John is my friend.
b. My friend is John.

(3a) is assumed to be about *John*, so *John* is the topic. Imagine someone challenges the claim in (3a) by saying "That's not true!" This would be understood as claiming "John is NOT your friend," but the existence of John would still be taken for granted. Since it is presupposed, it is outside of the scope of sentence negation. In fact, the denial could be felicitously complemented with "you don't have any friends," which indicates that only the existence of topic (*John*) is presupposed, the content of the comment is not. Conversely, uttering "That's not true!" in response to (3b) where the topic is *my friend*, would still presuppose that I have a friend (just not *John*).[4]

It is evident that the claim that a sentence is "about" a topic is vague in itself. In principle, one can argue that sentence (3a) is not just about John, but also about the speaker, or the speaker's friends. To remedy this problem, various tests have been proposed in the literature to identify topics. Prince (1999), citing Gundel (1974/1985) and Reinhart (1981), lists 3 tests that could be used for this purpose:

(i) *The "as for X"-test*:
Can the sentence be plausibly paraphrased with an initial "as for X"-phrase, where X is the supposed topic expression?

(ii) *The "what about X"-test:*
Can the sentence plausibly answer a "what about X"-question, where X is the supposed topic expression?

(iii) *The "say about X that…"-test:*
Could the sentence be plausibly reported about using an initial "Y said about X that…"-phrase, where X is the supposed topic expression?

If we use these tests on (3a), we can verify that *John* is indeed the topic of the sentence (and not e.g. the speaker's friends). Note that the test-results would be the exact opposite in the case if (3b).

(4) a. As for John, he is not my friend.
a′. #As for my friend, John is not him.
b. A: "What about John?"
B: "John is not my friend."
b′. #A: "What about your friend?"
B: "John is not my friend."
c. He said about John that John is not his friend.
c′. #He said about his friend John is not him.

It has also been suggested that the entity that is denoted by the topic expression should be accessible in the discourse universe. Gundel (1985) calls this the "familiarity condition" on topics. Lambrecht (1994: 159) offers a striking example. (5) is a beginning of a telephone conversation, where someone had dialed the wrong number.

[4] Note that the test even works if we replace *John* with a definite expression like *the king of France*. Of course one could say to (3a) "That's not true, because the king of France doesn't even exist!", but that would be an explicit modification of the presupposition.

(5) A: Is Alice there?
B: a. #Alice isn't here.
b. There is no Alice here.

Even though *Alice* is a referential, definite expression, and is clearly discourse old by the time B replies, the reply in (5a) is undoubtedly strange. The problem in B's first reply is that *Alice* is not properly established in the universe of the discourse, since B doesn't know what *Alice* could A refer to. The way to circumvent this problem is to remove *Alice* from the position where she is interpreted as a topic, as in B's second reply.

There are also other tendencies which have been noted in connection with topics. One of these is that topics tend to be animate entities. Many languages are like English in this respect in that there's no direct syntactic correlate for this, but the interested reader is referred to Dalrymple & Nikolaeva (2011, chapter 1 and references therein) for an outline of the correlation between animacy and topichood in the context of object-marking in a number of languages.

Another tendency is related to the intuition that topics are somehow centers of attention in a discourse. This, combined with the accessibility constraint mentioned earlier, is formalized by Centering Theory (Walker, Joshi and Prince 1998), which says that if anything is referred to with a pronoun in the subsequent discourse, it should be the backward looking center of a sentence. This is, informally speaking, the element that links the sentence to the previous discourse, which roughly corresponds to the notion of topic in Centering Theory.[5] The idea is that since topics are established in the discourse, active in the interlocutors' minds, they can easily be referred to with pronouns.

It has been recognized that a single term "topic" is inadequate to cover all the uses of topics. According to Frascarelli (2007), at least three subtypes should be distinguished: Continuing Topics, Shifting Topics, and Contrastive Topics. According to Frascarelli, each of these has specific characteristics (intonational pattern and syntactic behavior) in Italian.

Continuing Topics, which are always linked to the discourse, refer back to some already established entities. Shifting Topics can be "newly introduced, newly changed or newly returned to" (Frascarelli 2007). They either introduce completely new topics to the discourse, or they introduce a subtopic. These kinds of topics are also referred to as "Thematic Shifters" by Erteschik-Shir (2007). Gazdik (2012) mentions that in Hungarian, continuing topics contrast with thematic shifters in that only the latter can be overtly realized (Hungarian is a pro-drop language). So because the topic is unchanged from the previous sentence, the subject pronouns are anomalous in (6a). By contrast in (6b), the subject is subtopic of a previously introduced discourse topic.

(6) a. *Tamás szeret olvasni. (#Ő) intelligens, szorgalmas és sokra fogja vinnni.*
Thomas likes read.INF he intelligent hard-working and much aux.3SG reach.INF
'John likes reading. He is intelligent, hard-working and he will achieve a lot.'

b. *Mesélek a barátaimról, Tamásról, Péterről*
tell.1SG the my friends.POSS1PL.DEL Thomas.DEL Peter.DEL

[5] Centering Theory offers another candidate for topichood, the Preferred Center, the most prominent newly introduced entity. For details, see Prince (1999).

és Katiról. Tamás egy régi barátom, Pétert az
and Kate.DEL Thomas an old friend.POSS.1SG Péter.ACC the
egyetemről ismerem, Marival pedig együtt dolgozunk.
university.DEL know.1sg Mary.INSTR and together work.1PL
'I'll tell you about my friends, John, Paul and Mary. John is an old friend of mine from school, Paul, I know him from college, and Mary, I work together with her.'

Example (6b) also shows that topics are not always discourse-old, since the proper names themselves had not been mentioned before. As subtopics, they are accommodated from the discourse, from a general "question under discussion".[6]

Contrastive Topics contrast the topic entity to other entities in the discourse, like in (7):

(7) $\text{Tom}_{\text{C-TOPIC}}$ ate the beans.

Contrastive Topics are associated with a specific intonational contour and an interpretation that evokes that there is more to say, only a partial information has been given. It means that the sentence implies that there are additional people who ate other food items (e.g. Mary ate the meat, Joe ate the cake, etc.), or other people did something else to the beans (e.g. Mary saved them for further use). For more on Contrastive Topics, see Büring (1999, 2003).

2.2 Focus

Focus is usually regarded as the part of the sentence that contains new information. According to Gazdik (2011: 152), this is "related to the assumption that that focused constituents are the ones that answer constituent questions." She rejects this on the basis of an example like (8), where the focused constituent in B's reply is clearly not a new discourse entity, as it was already mentioned in the question.[7]

(8) A: Who did you invite, Tom or Mary?
B: I invited $\text{TOM}_{\text{FOCUS}}$

However, Gundel & Fretheim (2004) point out that it is important to distinguish between two kinds of given-new dichotomies: referential and relational. Although the

[6] An anonymous reviewer notes that in a framework like Prince (1998), the names in (6b) could be considered Contrastive Topics, as they are members of a set, introduced by the previous sentence. I think set-membership itself is not a sufficient condition to be recognized as contrastive. I follow Titov's (2013) definition of contrastiveness which says that that the contrastive entity itself activates other entities in the discourse (see section 2.2) In (6b), it is the context, not the topic-entities themselves indicates other discourse referents. A genuine CT like (6) evokes alternatives in the hearer without any explicit context.

[7] Although the overall intonational phonology of sentence (8) might not differ from that of a netural, broad focus sentence, it is fairly uncontroversial in the literature that the object phrase *Tom* is an information structural focus. In Lambrecht's (1994) system, it would be focus, since it is the part of the proposition where the assertion (the one that I invited = Tom) differs from the presupposition (I invited someone). Krifka (2008) also mentions that one of the basic functions of focus is to serve as a congruent answer to a question.

denotation of *Tom* is referentially given, in the sense that it is already present in the discourse, its relation to the predicate is new. In other words, what is new is that *Tom* can instantiate the variable in the evoked proposition *I invited X*.[8]

There are also subtypes of focus. The two main ones are Contrastive Focus and Information Focus. Both represent new information, but a Contrastive Focus also indicates that there are alternative candidates for the focus value. According to Titov (2013), for a focus to qualify as contrastive, "the set of alternatives must become active in the discourse at the point the sentence containing the contrastive element is uttered. No sooner and no later." In this sense, B's answer in (8) does not contain a contrastive focus, since the alternatives are already evoked in the question. The following exchange possibly contains a Contrastive Focus in B's response.

(9) A: Who did you invite?
B: It was $\text{TOM}_{\text{C-FOCUS}}$ who I invited.

Under the most natural interpretation, B's response entails that there were several other people whom B could have invited, but B chose *Tom*.

Focus is always highlighted in linguistic expressions in some way. English normally relies on prosody, so the focused element is accented. Hungarian, beside the prosodic means, also uses syntactic highlighting, placing the focused element into preverbal position, which may result in the well-known focus-induced inversion of particles, see (10).

(10) a. *Tomi*$_{TOPIC}$ *el-ment a koncertre.*
Tom away-went the concert.TO
"Tom went to the concert."
b. *Tomi*$_{FOCUS}$ *ment el a koncertre.*
Tom went away the concert.TO
"TOM went to concert."

The general problem with defining topic and focus is that both of them are multifaceted phenomena, with syntactic, semantic and pragmatic repercussions. It has been noted by several researchers (Gundel & Fretheim 2004, Prince 1999) that consistency concerning them is seriously lacking in the linguistic literature. It is a significant task for linguistic research to reconcile the different views and approaches. What I have provided in this section is far from satisfactory in general, but it should be satisfactory enough for our current purposes, namely the closer examination of the Information-Structural properties of TOP and LD. I shall carry this out in the next sections.

3 Topicalization

Below I repeat sentence (1a) as (11), which serves as our example for TOP. An object is fronted from its canonical position, leaving a "gap" behind.

[8] Even though Tom was mentioned as a possible candidate for the value of X, in the answer it is made an actual value of X. So its relation has changed, making it a new piece of information.

(11) Tom, I like.

The very first thing that we should note about sentence (11) is that the fronted constituent may actually have two distinct functions: it can be interpreted as a "topic-like" entity (we will return to this soon), or as some kind of focus. This was first noted by Prince (1981). The obvious question is this: what kind of focus is involved here?

Choi (1997), referring to Ward (1988), asserts that the fronted phrase actually refers to two discourse elements: one, a set or a scale, and two, a specification of a value or an element in that set or scale. In this example this would mean that the sentence evokes a set of people that I may like and picks *Tom* as a member of that set.

If this is correct, then the sentence meets the criteria for contractiveness defined by Titov (2013), mentioned earlier: the set of alternatives becomes active in the discourse at the point the sentence containing the contrastive element is uttered. When TOP is used this way, the sentence has only one pitch accent, an H* tone (which Jackendoff 1972 calls A-accent) on the fronted constituent.

In the other use of TOP, the sentence has two accents. On the initial expression, it has an L+H* tone. This is called B-accent by Jackendoff (1972), and there's also an accent on the verb or the subject. This alone indicates that what we are dealing with in this use is not a simple topic either.

Further doubt on the topichood of the initial element is cast by Prince's (1999) observations. She cites the following naturally occurring data, containing a topicalized phrase:

(12) Thanks to all who answered my note about asking about gloves. I didn't look at this bb for several days and was astounded that there were 11 answers. *Some I missed, darn.*

Prince (1999) points out that the topicalized phrase fails on all three topichood-tests we have mentioned earlier:

(13) a. Thanks to all who answered my note about asking about gloves. I didn't look at this bb for several days and was astounded that there were 11 answers. #*As for some, I missed them, darn.*

b. A: Thanks to all who answered my note about asking about gloves. I didn't look at this bb for several days and was astounded that there were 11 answers.
B: #*What about some?*
A: Some I missed, darn.

c. She thanked everyone who answered her note about gloves. She said she didn't look at this bb for several days and was astounded that there were 11 answers. #*She said about some that she missed them.*

The problem that underlies the intuition that these sentences fail the tests is that the noun phrase *some* is not definite. As Gundel & Fretheim (2004) note that indefinites are not generally used to refer to familiar entities, thus they fail the familiarity condition, discussed in section 2.1. To put it differently, one may assert that in (13a-c), the word *some* fails to provide an adequate referent about which the sentence could predicate something. The fact that in (12), the "topicalization" is felicitous nevertheless strongly suggests that the fronted constituent is not a topic.

Moreover, it was established that topics should at least be referential. Considering this, it is striking that there are several grammatical elements that may be topicalized, but would not count as referential under any basic understanding of the concept: verbs ((14a) and (14b)), adjectives (14c) and propositions (14d). If topicalization was about (referential) topics, all these examples would be predicted to be unacceptable.

(14) a. Surrender, we never will.
b. To win, we at least tried.
c. Happy, Tom will never be.
d. That Tom was a movie star, we would never have guessed.

Finally, we should mention that Prince (1999) notes that in the corpus of Gregory Ward, which is a collection of naturally occurring Object-Subject-Verb structures, not one case can be found where the topicalized phrase is a 3rd person pronoun (she picks 3rd person pronouns because that is where one may choose between a lexical NP and a pronoun). As was discussed, according to Centering Theory, if anything is referred to with a pronoun in a sentence, it should be the backward looking center, which is one of Centering Theory's candidates for the notion of topic. The fact that such elements are seldom topicalized raises further skepticism about the fact that TOP is a topic-marking device.

Of course none of these arguments is a clincher. One may debate that the topichood-tests are reliable enough (Gundel & Fretheim 2004 note that pragmatic tests are not deterministic, so they cannot be used as foolproof methods for identifying topics), or one does not have to subscribe to Centering Theory. However, the arguments enumerated in this section all point to the same direction: Topicalization is not about topics. What then is it about?

We have already established that TOP may mark a Contrastive Focus in one of its uses. I would like to argue that the other use of TOP marks Contrastive Topics (C-TOPIC), giving us the generalization that TOP is a marker of contrastive Information Structural categories. For the example in (14) it means that the topicalized phrase implies that there are other answers which the speaker did not miss, so contrast is evoked. This indeed seems to be a plausible interpretation.

The claim that TOPs mark C-TOPICs also sheds some light on the question of why it can be used with nonreferential expressions, demonstrated in (14). For reasons that are not clear to me at this point, the restrictions on what can qualify as a Contrastive Topic are lighter than on regular topics. The reasons for this should be subject to further investigation.[9] Nevertheless, the fact remains. For instance, Gécseg (2001) notes that in Hungarian (similarly to 14b-c), infinitives and adjectives can serve as C-TOPICs, unlike regular topics (the same fact holds for focus as well):

(15) a. *Úszni*$_{C\text{-}TOPIC}$ *tudok.* (Hungarian)

[9] An anonymous reviewer raises the possibility that this is "because fronting is not associated with the notion of 'contrastive topic' but with contrastiveness in general and so restrictions on topics don't always hold of fronted elements as they are not always topics." I think this view would be inadequate for several reasons. First, fronting itself is not necessarily contrastive, as we will see in the case of Left-Dislocation. Second, contrastive topics occur elsewhere as well and they are subject to the same looser categorical restrictions there (e.g. *Tom will never be happy*$_{C\text{-}TOPIC}$, *but at least will have money*).

swim.INF can.1SG
"To swim, I'm able to." (as opposed to e.g., to ski)

b. *Szépnek*$_{C\text{-}TOPIC}$ *szép a húgod.*
pretty.DAT pretty the sister.POSS.2SG
"Pretty, your sister in fact is." (but she may not be clever)

Having a constituent in a non-canonical position that might have two discourse functions is not unique to English. The Hungarian construction exemplified in (16) shows similar Information-Structural behavior.

(16) *Tomit mondtam, hogy láttam.* (Hungarian)
Tomi.acc said.1SG that saw.1PL
'Tom, I said that I saw.'

Whatever syntactic analysis of these structures we subscribe to (for two different views, see Gervain 2002 and Szűcs 2014), from an information-structural perspective it is clear that the initial constituent *Tomi* could be either C-FOCUS or C-TOPIC. Just like in English, the two interpretations correspond to different intonational patterns. However, the parallel is not perfect, as in Hungarian, the two interpretations would also lead to syntactic differences. Since in Hungarian, only the focus has to be adjacent to the verb, in the case of the C-TOPIC-interpretation, a verbal modifier can intervene between the fronted constituent and the verb, see (17).

(17) *Tomit meg-mondtam, hogy láttam.* (Hungarian)
Tomi.acc PERF-said.1SG that saw.1PL
'Tom, I did say that I saw.'

4 Left-Dislocation

Superficially, LD differs from TOP only in that it contains a resumptive pronoun in the canonical position of the initial phrase. Our example for it was (1b), which is repeated here as (18).

(18) *Tom, I like him.*

Prince (1998) claims that there are 3 basic functions for LD:

(i) island-amnesty,
(ii) simplifying discourse processing,
(iii) signaling a "poset-inference."

In the first use, it is actually applied as covert topicalization. The speaker would like to use TOP, but faces a syntactic obstacle, e.g. an island, and thus is forced to put a resumptive pronoun in the canonical position of the initial element. One such example is shown in (19).[10]

[10] While in some languages, the distribution of gaps and resumptive pronouns is more complex, it is fairly uncontroversial in the literature that English uses resumptive pronouns for a very

(19) Tom, the story about *(him) was funny.

As such uses are clearly forced by core syntax and have nothing to do with Information Structure, I exclude them from the scope of this paper.

The second function of LD is "simplifying discourse processing." According to Prince (1998) this means that by using LD, people remove discourse-new entities from positions that are dispreferred for them. Prince's (1998) example for this is the following segment:

(20) My sister got stabbed. She died. Two of my sisters were living together on 18th Street. They had gone to bed, and this man, their girlfriend's husband, came in. He started fussing with my sister and she started to scream. *The landlady, she went up,* and he laid her out. So sister went to get a wash cloth to put on her, he stabbed her in the back.

According to Prince (1998), *the landlady* in its original position would be a subject and subjects are generally dispreferred as discourse-new entities.[11] One can also approach this from the perspective of Lambrecht's (1994: 185) "Principle of the separation of reference and role": do not introduce a referent and talk about it in the same clause. This militates against viewing the left-dislocated element as simple topics, as it was discussed in section 4.2, that topics are preferably discourse-established entities.

The third use of LD according to Prince (1998) is to trigger an inference on the part of the hearer that the entity represented by the initial NP stands in a salient partially-ordered set relation to some entity or entities already evoked in the discourse-model. Partially ordered sets, "posets" are "defined by a partial ordering R on some set of entities, e, such that, for all e-1, e-2, and e-3 that are elements of e, R is either reflexive, transitive, and antisymmetric or, alternatively, irreflexive, transitive, and asymmetric" (Prince 1998). In essence, this means that the left-dislocated entity has some set relation with other elements.

Prince (1998) sees these functions as separate entities. However, subsequent research suggests that there may be a way to have a unified view of functions 2 and 3 (as was stated, the first function is set aside in this paper).

Gregory & Michaelis (2001) have conducted a corpus study on TOP and LD. They suggest that the overarching function of LD is that of "topic promotion", that is, to bring entities into the discourse. They have compared all the LD tokens with all the TOP tokens and have found 3 factors that back this claim up.

First, they examined the givenness of LDs, compared to TOPs. They used Gundel, Hedberg & Zacharski's (1993) cognitive statuses to determine the referential givenness of an element in the discourse. These are (from the lowest to the highest givenness): type

---

restricted set of purposes. Their main function is to neutralize island-violations like the one in (19), and possibly they can be inserted in some sentences for parsing purposes, for instance see (iib) from Falk (2002).

(ii) a. This is the girl that John likes (*her).
b. This is the girl that Peter said that John thinks that yesterday his mother had given some cakes to ?(her).

For more discussion on resumption see Aoun (2001), Falk (2002) and Gervain (2004).

[11] There is a traditionally assumed connection between subjecthood and topichood, see Lambrecht 1994, chapter 4.2.

identifiable, referentially uniquely identifiable, familiar, activated, in focus. In (18) there is an example for each status (examples 21a to 21e are from Gundel, Hedberg & Zacharski 1993).

(21) a. *Type identifiable*:
I couldn't sleep last night. *A dog* (next door) kept me awake.
b. *Referential*:
I couldn't sleep last night. *This dog* (next door) kept me awake.
c. *Uniquely identifiable*:
I couldn't sleep last night. *The dog* (next door) kept me awake.
d. *Familiar*:
I couldn't sleep last night. *That dog* (next door) kept me awake.
e. *Activated*:
I couldn't sleep last night. *That* kept me awake.
f. *In focus*:
I couldn't sleep last night because of your dog. *It* kept barking.

Th authors found that LD has relatively low givenness in the discourse, the most typical givenness status being uniquely identifiable. According to Gregory and Michaelis (2001), this is expected if LD is a topic-promotion device, since "uniquely identifiable status alone represents the intersection of discourse-new and hearer-old statuses," entities that can be identified by the hearer (a condition for topics, see example (5)), but are not in the current discourse yet. TOPs on the other hand had higher activation status, which is expected if they are contrasted to some discourse elements, as was established in the previous section.

Gregory & Michaelis's (2001) second target for investigation was the anaphoricity of left-dislocated and topicalized entities. They categorized tokens according to the type of the anaphoric link that the fronted element had to the discourse (from highest to lowest): directly mentioned, the entity is member of a set that has been mentioned, none. They found that LDs tended to have low anaphoricity, which is expected if their role is topic promotion.

Gregory & Michaelis (2001)'s final factor was topic persistence. They measured to what extent the fronted elements in LD and TOP tend to remain topics of the subsequent discourse. They found that LD has a high topic persistence, as opposed to TOP. This is in line with what we have discussed in connection with these structures: LD is a topic promoter, so one expects that the entity introduced by it is going to be talked about. We do not have such expectations for contrasted elements introduced by TOP.

What do these results of functionalist research mean from a generative perspective? I think considering all these it stands to reason to say that left-dislocated elements correspond to the discourse function Frascarelli (2007) and Erteschik-Shir (2007) refer to as "shifting topic" or "thematic shifter". This means that LD could be regarded as a topic-marking device, but what we have here is a subtype of topics. It either introduces a completely new topic (Prince's 1999 first function), or a subtopic of an existing discourse topic (Prince's 1999 second function). That LD can be used to introduce a brand-new topic is evident from the fact that it is conceivable that someone, looking for a particular Tom, enters a room and utters the following sentence, containing an LD:

(22) Tom, where is he?

The same could hardly be conceivable with TOP (as C-TOPIC and C-FOCUS are always related to the discourse and cannot be uttered out of the blue), though syntactic factors may also interfere in this particular example.

That LD is related to topics gets further support from two facts. First, recall that we discussed in section 3 that the fact TOP is grammatical with nonreferential entities supports the claim that it's not a topic-marking device. On the other hand, such entities make LD seriously degraded:

(23) a. ???Surrender, we will never do so.
b. ???Happy, Tom will never be like that.
c. ???That Tom was a movie star, we would have never guessed that.

Second, a corpus study by Snider & Zaenen (2006) found that there is a positive correlation between LD and animacy. This is expected if LD is a device for marking a kind of topics, Thematic Shifters, as it was noted in section 2, there is a general tendency for topics to denote animate entities.

5 Conclusions

I have argued in this paper that the general idea that Topicalization and Left-Dislocation are topic marking devices is too simplistic. I aimed to look into functionalist linguistic research regarding these constructions and I tried to create a synthesis with the concepts of generative research. In my view, TOP is a marker of contrastive IS categories, Contrastive Topic and Contrastive Focus, whereas LD corresponds to thematic shifters. The following passage, from Prince (1998) is a nice illustration of these statements:

(24) She had an idea for a project. She's going to use three groups of mice. *One, she'll feed them mouse chow*, just the regular stuff they make for mice. *Another she'll feed them veggies. And the third she'll feed junk food.*

In this passage, two LDs are followed by a TOP. The LDs represent subtopics of the groups of mice, introduced in the second sentence. They are part of a set, but not contrasted. What is contrasted is the third group of mice. This is quite obvious from the setting, the experiment, the aim of which most probably is to evaluate the effects of junk food. So what I claim is that although both TOP and LD may indicate that the denotatum of the fronted constituent is an element of a set, it is only TOP that actually evokes a contrast with other members of the set.

Chafe (1976) already suggested that TOP is necessarily contrastive. He defined "contrast" as assertion on the part of the speaker that one of "a limited number of candidates" is "correct". Birner & Ward (1998) criticizes this view on the basis of examples like (25):

(25) The only time the guy isn't considered a failure is when he resigns and announces his new job. That's the tipoff, "John Smith resigned, future plans unknown" means he was fired. "John Smith resigned to accept the position of president of X company" – then you know he resigned. *This little nuance you recognize immediately when you're in corporate life.*

They claim that "it seems unlikely that the speaker is asserting that one little nuance is the 'correct selection' from some set of little nuances." I think Chafe (1976) was right in his claim that TOP is contrastive, but he gave a wrong definition of contrastiveness. What counts is not whether the number of candidates is limited or not, or whether there is a "correct" selection, but the fact that the topicalized constituent does evoke the presence of a set of salient alternative members, as in Titov's (2013) definition of contrast. I think that this is intuitively true even for (25). It does evoke the inference that there are other nuances in corporate life that one could talk about. Left-dislocated entities may be set-members, but the other members of the set are not salient in relation to the left-dislocated element. This is particularly clear in example (22) (*Tom, where is he?*), which has zero implication suggesting that the speaker might also look for other people. What is important is the newly introduced topic (*Tom*), set-membership is non-existent in this case.

These notions (Contrastive Topic, Contrastive Focus, Thematic shifter) are recognized categories in generative research, so this attempt is a favorable move in the goal of bringing different research traditions closer to each other.

The formalization of these suggestions is subject to further research. This is not an easy task. One direction one could take is to regard "contrast" as a primitive IS notion that characterizes both C-FOCUS and C-TOPIC, but not Thematic Shifters or Information Foci (see e.g. Vermeulen 2009). Another intriguing possibility is Titov's (2013) recent suggestion that C-FOCUS and C-TOPIC are not separate IS categories, but one category in different configurations. A third option could be to try to define IS notions in terms of discourse-linkedness or prominence (e.g. Gazdik 2011).

Because of the several linguistic levels involved, I think models with multiple levels of representation like Lexical Functional Grammar would fare the best in formalizing these phenomena. In this model of grammar, there have already been some advances in the formal representation of Information Structure (see King & Zaenen 2004 and Gazdik 2011), phonology (Mycock and Lowe 2013) and discourse structure (Gazdik 2011).

References

Aoun, Joseph, Lina Chouieri and Norbert Hornstein. 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry* 32. 371–404.

Birner, Betty & Gregory Ward. 1998. *Information status and noncanonical word order in English*. Amsterdam/Philadelphia: John Benjamins Company.

Büring, Daniel. 1999. Topic. In Peter Bosch & Rob van der Sandt (eds.), *Focus – Linguistic, cognitive, and computational perspectives*. Cambridge: Cambridge University Press.

Büring, Daniel. 2003. On d-trees, beans, and b-accents. *Linguistics & Philosophy*, 26.5. 511–545.

Choi, Hye-Won. 1997. Information structure, phrase structure, and their interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG97 Conference*. Stanford: CSLI Publications.

Chafe, Wallace. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles Li (ed.), *Subject and Topic*, 25–55. New York: Academic Press.

Dalrymple, Mary 2001. *Lexical functional grammar*. San Diego: Academic Press.

Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.

Falk, Yehuda. 2002. Resumptive pronouns in LFG. In Miriam Butt, Tracy Holloway King (eds.), *Proceedings of the LFG 02 conference*, 154–173. Stanford: CSLI Publications.

Frascarelli, Mara. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form*, 87–116. Amsterdam/Philadelphia: John Benjamins.

Gazdik, Anna. 2011. Multiple questions in French and Hungarian. Université Paris Diderot & Eötvös Loránd Tudományegyetem, PhD dissertation.

Gervain, Judit, 2004. The syntax and semantics of the resumptive dependency in Hungarian focus-raising constructions. In Benjamin Shaer, Werner Frey & Claudia Maienborn (eds.), *Proceedings of the Dislocated Elements Workshop (ZAS Papers in Linguistics*, 35.1), 111–131. Berlin: ZAS.

Gundel, Jeanette K. 1974. The role of topic and comment in linguistic theory. University of Texas, Ph.D. dissertation.

Gundel, Jeanette K. 1985. 'Shared knowledge' and topicality. *Journal of Pragmatics* 9.1.83-107.

Gundel, Jeanette K., Nancy Hedberg and Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69.2. 74–307.

Gundel, Jeanette K. & Thorstein Fretheim. 2004. Topic and focus. In Ward, Gregory & Horn, Laurence R. (eds.), *The Handbook of Pragmatics*, 175–196. Blackwell Publishing.

Gécseg, Zsuzsa. 2001. A kontrasztív topik szintaxisáról és szemantikájáról. *Magyar Nyelv*, 2001. szeptember, december.

Gregory, Michelle L. & Laura A. Michaelis. 2001. Topicalization and left-dislocation: A functional opposition revisited. *Journal of Pragmatics* 33. 1665–1706.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge: MIT Press

King, Tracy Holloway & Annie Zaenen. 2004. F-structures, information structure, and discourse Structure. Conference talk presented at the LFG04 Conference, University of Canterbury.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55. 243–276.

Lambrecht, Knud. 1994. *Information structure and sentence form*. Cambridge: Cambridge University Press.

Mycock, Louise & John Lowe. 2013. The Prosodic marking of discourse functions. In Miriam Butt &, Tracy Halloway King (eds.), *Proceedings of the LFG13 conference*, 440–460. Stanford: CSLI Publications.

Prince, Ellen. 1998. On the limits of syntax, with reference to left-dislocation and topicalization. In Peter Cullicover, Louise McNally (eds.), *Syntax and Semantics 29 – The limits of syntax*. New York: Academic Press.

Prince, Ellen. 1999. How not to mark topics: 'Topicalization' in English and Yiddish. In *Texas Linguistics Forum*, Chapter 8. Austin: University of Texas.

Reinhart, Tanya. 1981. Pragmatics and linguistics: an analysis of sentence topics. *Philosophica* 27. 53–94.

Snider, Neal & Annie Zaenen. 2006. Animacy and syntactic structure: Fronted NPs in English. In Miriam Butt, Mary Dalrymple. & Tracy Holloway King (eds.), *Intelligent linguistic architectures: Variations on themes by Ronald M. Kaplan*. Stanford: CSLI Publications.

Szűcs, Péter. 2014. A magyar "operátoremelés", mint prolepszis. In Gécseg Zsuzsanna (ed.), *LingDok 13: Nyelvészdoktoranduszok dolgozatai*, 185–204. Szeged: Szegedi Tudományegyetem Nyelvtudományi Doktori Iskola. Electronic publication: http://nydi.bibl.u-szeged.hu/SZTE_NYDI/LingDok_kotetek_files/lingdok13.pdf

Titov, Elena. 2013. Do contrastive topics exist? *Journal of Linguistics* 49.2. 413-454.

Vermeulen, Reiko. 2009. Topics, contrast and contrastive topics in Japanese. In Reiko Vermeulen & Ryosuke Shibagaki (eds.), *MIT Working papers in linguistics 58: Proceedings of the 5th workshop on Altaic formal linguistics (WAFL 5)*, 361–372. Cambridge: MIT Press.

Ward, Gregory. 1988. *The semantics and pragmatics of preposing*. NY/London: Garland Publishing.