Licensing by Information Structure

JENNIFER MACK
Yale University

0. Introduction
Information structure is the interface between the cognitive organization of discourse and the formal representation of sentences. The view of information structure as an interface level is implicit in the following definitions – “the tailoring of an utterance to meet the particular assumed needs of the intended receiver” (Prince 1981a:224) and “the formal expression of the pragmatic structuring of a proposition in a discourse” (Lambrecht 1994:5) – as well as in Chafe’s (1976) term “information packaging”. Current research continues to explore the scope of the information structure interface: the set of formal operations driven by the roles that information conveyed by a sentence play in a discourse.

Informational relations influence the formal expression of propositions on the syntactic, morphological and prosodic levels. In the domain of syntax, informational constraints are taken to motivate operations as diverse as displacement (Prince 1981b, Davison 1984, Birner and Ward 1998), dislocation (Lambrecht 1994, Grosz and Ziv 1998), the (non)-expression of syntactically optional arguments (Goldberg 2001), and the conditioning of lexically represented argument structure alternations (Goldberg 1995, Ruppenhofer 2004). This paper demonstrates that the reach of information structure extends to the domain of licensing. The evidence for this claim comes from English topical exclamative (TE) constructions such as the bold sentences in (1)-(3). The present paper is the first exploration to date of TE sentences. The name “TE” reflects the observation that such sentences are instances of the exclamative sentence type (Michaelis and Lambrecht 1996b) that obligatorily express the topic relation.

(1) A: Have you met Madonna or had much to do with her throughout your career to date?

I would like to thank Betty Birner, Yael Fuerst, Larry Horn, Knud Lambrecht, Laura Michaelis, David Pesetsky, Gregory Ward, and the audiences at BLS 33 and Yale University for helpful comments and suggestions on aspects of this research.
Jennifer Mack

B: No I haven’t met her but I would love to, she’s amazing the way she reinvents her image and is always ahead of the bunch.

(2) People are amazing when you give them the information and the tools they need. They’re amazing the responsibility they’ll accept.
(www.sbnonline.com/National/Article/124/9794/Squeezing_the_tube.aspx)

(3) A: Are you sure the windows are closed?
B: Yeah, of course!
A: No, but it’s hard to tell with them. Those windows are weird the way they close.  
(KM to JM in conversation, 1/20/2007)

In this paper, I demonstrate that the matrix subject is not selected by the main predicate in TE – in fact, the subject is not assigned a semantic role at all. Instead, the subject is licensed to express topicality: the informational relation of “about-ness” that holds between a referent and a proposition (Strawson 1964, Reinhart 1981, Lambrecht 1994). Although the licensing of non-argument topics in extra-clausal, “detached” positions is a well-known phenomenon (Li and Thompson 1976), intra-clausal licensing by information structure is, to the best of my knowledge, previously unattested. This analysis of TE carries implications for the extent to which informational constraints interact with other components of the grammar in determining sentence form. In particular, the phenomenon of licensing by information structure is a challenge to models of language that equate licensing with the projection of lexical argument structure representations.

1. The English Topical Exclamative Construction
The English TE construction resembles right dislocation (RD; Lambrecht 1994, Grosz and Ziv 1998) and nominal extraposition (NE; Michaelis and Lambrecht 1996a) but is distinct from both constructions with respect to form and informational relations. Examples (4)-(6) display the formal contrasts between the three sentence types, each of which contains an “extraposed” noun phrase (indicated by italics) which follows the main predicate. The term “extraposition” indicates that the post-predicate NP is assigned the semantic role canonically assigned to the subject position (cf. Michaelis and Lambrecht 1996a); it is not meant to imply a movement analysis of extraposition.

(4) It’s annoying, the way she always cracks her knuckles. [RD]
(5) It’s annoying the way she always cracks her knuckles. [NE]
(6) She’s annoying the way she always cracks her knuckles. [TE]

One key distinction is the semantic status of the subject NP, which is indicated by bold font. Across all three sentence types, the initial NP can be identified as the grammatical subject on the basis of morphosyntactic diagnostics: it appears in pre-verbal position, triggers verbal agreement, and receives nominative case (see Keenan 1976 for a discussion of diagnostics for subjecthood). Example (4)
illustrates RD, which is characterized by co-reference between the matrix subject and the post-predicate NP. In contrast, the main subject in NE (5) is non-referential (Michaelis and Lambrecht 1996a). Finally, TE (6) contains a referential subject that is not co-indexed with the post-predicate NP. With respect to prosody, TE broadly patterns with NE as both are pronounced under a single intonation contour. In contrast, RD is characterized by a clear prosodic break between the matrix predicate and the following NP.

TE is also distinct from NE and RD with respect to the informational relations that it expresses. As I will argue in Section 4, TE has a topic-focus structure: the predication between the matrix predicate and the following NP is in focus, and the asserted information is construed as being about the matrix subject. This type of topic-focus structure – in which the topical referent is not an argument of the in-focus predication – is pervasive in “topic-prominent” languages but relatively rare in “subject-prominent” languages such as English (Li and Thompson 1976). Neither RD nor NE has a topic-focus articulation of this sort. RD expresses a topic-focus structure that is more typical of English information structure: the post-predicate NP is a topic whose referent is an argument in the focal predication. In contrast, NE is an all-focus construction. Its main predication is in focus, but there is no element of the construction that is conventionalized to express topicality (Michaelis and Lambrecht 1996a).

TE is a common construction in contemporary American English, but appears to be restricted to relatively informal registers. Nearly all of the naturally-occurring tokens that I have collected come from spontaneous speech and informal written genres such as blogs and message boards. Prescriptively, TE is dispreferred. Although all of the American English speakers that I surveyed rate TE as at least “somewhat natural” (in at least some contexts), many perceive it as lazy, sloppy, or casual. One possible source of this perception is that TE appears to be a “young” construction, a relatively recent innovation. As recently as 1991, the year in which the Switchboard corpus of spoken American English (Godfrey et al. 1992) was recorded, TE was rare. However, a Google search performed in May 2006 revealed that 20% of the 45 sentences containing the phrase annoying the way are TE. The contrast between these results suggests that TE has become more common over the past 15 years. Of course, it is possible that sampling differences in dialect and register may have contributed to this effect.

2. Informational Constraints and Sentence Structure
Following Lambrecht (1994), I take “topic” and “focus” to express informational

---

2 I found no instances of the construction in a large sample of environments in which the construction could potentially occur. No tokens of TE were present within a randomly sampled set of 588 of the 1296 instances of the phrase “the way” found in the Switchboard corpus. (The post-predicate NP in TE very frequently takes the form “the way CP”.) I also found no instances of TE in 20 corpus searches of the form “ADJ the”, where ADJ is an exclamative adjective known to be compatible with TE (e.g. amazing, annoying, unbelievable).
relations that hold between a proposition and the mental organization of a discourse. They determine what information should be entered into the representation of an unfolding discourse and where this information should be filed. Topic and focus are intended to be informational constructs. They are not defined formally, nor are they necessarily reducible to any aspect of sentence form at the morphosyntactic level, including sentence-initial or subject position (Reinhart 1981, Lambrecht 1994, contra Halliday 1967 and the Prague School literature on functional sentence perspective, e.g. Firbas 1964). Though topic and focus are signaled by prosodic structure – topics tend to be unaccented and occur at the edge of a prosodic contour, while focal information is prosodically prominent – the relations are not defined at this level; instead, they are defined by the roles that they play in the organization of a discourse (e.g. Reinhart 1981). However, topics are still elements of sentence grammar; topical referents must be realized as sentence constituents even though they are defined on the informational level. This notion of “topic” is distinct from the construct of “discourse topic”, which is defined at the level of discourse coherence and may or may not be realized as a sentence-level constituent.

The relation of topic makes a pragmatic assertion of “aboutness”: topicality requires the discourse representation to indicate that a proposition is about a given referent (e.g. Strawson 1964, Reinhart 1981, Lambrecht 1994). Reinhart (1981), following Strawson (1964), connects the intuitive notion of aboutness with the storage and evaluation of information in a developing discourse. When an informative proposition enters a discourse representation, it may be stored under the organizational heading provided by a topic. If the topic relation holds, the truth of the proposition will be evaluated in terms of the available information about the topic. The evaluative component of the aboutness relation restricts the way in which the truth of a proposition is computed, but does not change the resulting truth value.

Informally, focal information contributes substantially to the body of knowledge that is active in a discourse representation. Lambrecht (1994:213) defines focus as “the semantic component of a pragmatically structured proposition whereby the assertion differs from the proposition”. Within a proposition, the focus relation may occur in isolation or embedded within the topic relation. Here I argue that the formal distinction between NE and TE encodes just this pragmatic distinction: an all-focus information structure in NE and a topic-focus articulation in TE. In other words, TE gives more specific directions about how the focal information should be represented. As parallel NE and TE sentences have identical truth conditions but differ with respect to information structure, they are “allosentences”, defined as “available … grammatical alternatives for expressing a given proposition” (Lambrecht 1994:6).

3. Argument Structure in TE
The primary goal of this section is to establish that the post-predicate NP in TE is an argument of the main predicate – and essentially, that the matrix subject is not.
Initially, this claim may seem far-fetched. For example, the most natural interpretation of (7) for some speakers is, informally, as follows: Amelia is annoying with respect to the way she always cracks her knuckles. In this interpretation, the matrix subject is apparently assigned a semantic role by the main predicate: the sentence seems to assert that Amelia is annoying.

(7) Amelia is annoying the way she always cracks her knuckles.

However, if this were the true interpretation, it would be problematic for most theories of argument structure. In such an analysis, a single semantic role (that of “being annoying”) would be assigned to two semantically and syntactically distinct constituents (the subject NP and the post-predicate NP) within a minimal domain of predication. On both empirical and theoretical grounds, projectionist theories of argument structure explicitly prohibit such an operation, e.g. through the θ-criterion (Chomsky 1981). Thus, there is pressure from the perspective of argument linking to devise an alternative account of the interpretation of TE. The rest of this section will demonstrate that distributional evidence also fails to support this analysis.

Main predicates in TE express judgment. They assert that a semantic constituent is subjectively assigned a value in a particular domain. As such, they select for an argument that undergoes evaluation, along with an evaluator. In TE, the “evaluator” role is typically unexpressed. As for the “evaluated” role, the rest of this section will demonstrate that it is linked to the post-predicate NP, and not to the subject NP. The truth conditions associated with TE provide support for this claim.

For a moment, let us return to the hypothesis that both the subject NP and the post-predicate NP in TE are assigned the “evaluated” role, in line with the apparently natural interpretation of the construction discussed at the beginning of this section. The predictions of this hypothesis are not borne out, as contradiction of the two hypothesized predications (between the predicate and each of the two NPs) results in very different truth conditions, as illustrated in (8)-(9).

(8) # She’s annoying the way she cracks her knuckles, but THE WAY SHE CRACKS HER KNUCKLES isn’t annoying.
(9) She’s annoying the way she cracks her knuckles, but SHE’S not annoying.

Example (8) is clearly contradictory, which is to be expected if the post-predicate NP is indeed assigned the “evaluated” role. All survey respondents indicated that (8) contains a “clear contradiction”. However, (9) – which tests whether an

---

3 I distributed a formal survey to six naïve native speakers of American English (age range: 20-54, diverse regional backgrounds); the judgments presented in Sections 3 and 4 are averaged from the responses to this survey.
argument-predicate relation holds with the subject NP – is entirely well-formed: five out of six respondents indicated that (9) contains “no contradiction”, while one responded that it contains a “possible contradiction”. These examples demonstrate that only the post-predicate NP is assigned the “evaluated” role in TE.

Examples (10)-(12) provide further evidence that the subject NP is not an argument in TE, as its presence does not contribute to the truth conditions of the sentence. Most speakers find (10) at least somewhat contradictory. However, the TE construction in (11) is not contradictory, demonstrating again that the “evaluated” role is not assigned to the subject NP. The source of this contrast cannot be the potential lexical ambiguity of funny (“strange” or “humorous”). For the survey respondents, only the “humorous” sense is available, as demonstrated by the perceived contradiction in (10); if the “strange” sense were available to these comprehenders, then (10) should be felicitous. In other words, the interpretation of (10) illustrates that funny is not lexically ambiguous for the survey respondents. As such, the well-formedness of (11) indicates that the subject NP is not assigned a semantic role by funny.

(10) ??She’s funny, but she has no sense of humor.
(11) She’s funny the way she always drops things, but she has no sense of humor.
(12) It’s funny the way she always drops things, but she has no sense of humor.

The comparison between (11) and the NE construction in (12) is striking: they are equally well-formed, suggesting that the presence of a referential subject in TE does not affect the sentence’s truth-conditions. (11) and (12) are thus allosentences. Section 4 will demonstrate that the differences in their distribution are motivated at the informational level.

These contrasts make it clear that the subject in TE is not assigned either of the semantic roles that are known to be lexically selected by the main predicate. However, there remains the possibility that the subject is assigned a different semantic role. As the claim that TE subjects are licensed by information structure is dependent on the subject not being a semantic argument, it is necessary to explore this possibility further.

Clearly, the hypothetical semantic role could not emerge from the predicate itself, as at least some TE predicates maximally select for “evaluator” and “evaluated” roles. It is possible, though, that the additional semantic role could be

---

4 Four out of six respondents rated (10) as “possibly contradictory”; one rated this example as “clearly contradictory,” while the final respondent saw no contradiction in (10). In contrast, five out of six respondents saw no contradiction in either (11) or (12). For all respondents, (11) and (12) were judged to be less contradictory than (10).

5 Pesetsky (1987) proposes a three-argument structure for experiencer verbs such as frighten and amaze: “experiencer”, “cause of emotion”, and “object of emotion”. Though many experiencer predicates appear in TE, this three-argument articulation could not be extended to all TE predi-
projected from the top down, i.e. from a constructional event representation. The essential question is what this semantic role would be. Lambrecht (p.c.) suggests that this role is assigned to an individual who possesses the property that is evaluated in TE; the individual role is assigned to the subject NP, while the “evaluated” role is assigned to the post-predicate NP.

Indeed, this semantic relationship between the referents of the two NPs is typical of TE. The referents of most TE subjects do play a role within the eventuality expressed by the post-predicate NP. (13) is a typical example: the referent of the matrix subject is the agent of the event expressed by the post-predicate NP.

(13) She’s annoying the way she always cracks her knuckles.
(14) Marilyn Monroe is amazing the way people still read books about her.

However, this semantic relationship is not required. Example (14) is a well-formed token of TE even though the referent of the matrix subject, Marilyn Monroe, does not participate in the embedded event (cf. Reinhart 1981). She is certainly relevant to the event, but this relation of relevance is not established through event participation. This suggests that the relationship between the subject and the post-predicate NP in TE is not defined at the semantic level. Section 4 will demonstrate that the relation is constrained at the informational level, specifically by the pragmatic requirements for topicality.

4. Information Structure in TE

As the subject in TE is topical, its relationship with the extraposed NP is constrained by a requirement of relevance: “A statement about a topic can count as informative only if it conveys information which is relevant with respect to this topic” (Lambrecht 1994:191). In TE, the requirement of relevance is often satisfied through the assignment of an event role – the referent of the subject NP is frequently a participant in the event expressed by the post-predicate NP. While event participation appears to be a sufficient condition for relevance, it is not a necessary one, as illustrated by example (14). A further exploration of the conditions for relevance within the topic relation is outside the goals of this paper. However, the acceptability of (14) demonstrates that the referents of the two NPs in TE are obligatorily linked at the pragmatic level, though not necessarily linked in semantic event representations.

The rest of this section will provide further evidence that TE obligatorily expresses a topic-focus information structure. Here I will emphasize the empirical basis for the claim that TE subjects are topical; I refer the reader to Michaelis and Lambrecht’s (1996a) account of NE – which as demonstrated above is the allosentence of TE – for arguments that the post-predicate NP is in focus. The
Jennifer Mack

evidence for topicality in TE is drawn from diverse sources: topic-selecting constructions, the activation status of TE subjects, and the effect of ratified discourse topics on the acceptability of TE in context.

First, TE is compatible with topic-selecting constructions such as “As for X, Y”, “Speaking of X, Y” and “What’s up with X?”, where X is required to be a topic (Reinhart 1981). When the topic-selecting construction subsumes TE within a discourse, the topical element X must co-refer with the TE subject, lending support to the claim that TE subjects are obligatorily topics. This is illustrated in (15)-(17).

(15) As for Mary, she worries me the way Bill keeps beating her at Scrabble.
(16) # As for Bill, Mary worries me the way he keeps beating her at Scrabble.
(17) A: What’s up with Mary?
   B1: Well, she worries me the way Bill keeps beating her at Scrabble.
   B2: # Well, Bill worries me the way he keeps beating her at Scrabble.

In (15), the topic Mary co-refers with the matrix subject of the TE construction, and the sentence is pragmatically well-formed. However, when the established topic and the TE subject do not co-refer, as in (16), the result is infelicitous. This is the case despite the fact that Bill is the agent of the embedded event in (16), and therefore is a potential topic by the criterion of relevance. In fact, in the absence of a discourse context Bill seems to be a better potential topic than Mary, as agentivity tends to correlate with topicality. The contrast between the B1 and B2 responses in (17) extends the same observation to a different topic-selecting construction. The infelicity of (16) and (17-B2) thus indicates that the subject position in TE conventionally expresses the topic relation.

Further evidence that subjects in TE are topical emerges from constraints imposed through the interaction of the topic relation with activation status, i.e. the extent to which a referent is “activated” in the mental representation of a discourse. Although information status and topicality are dissociable – inactive referents, often coded as indefinite NPs, may occasionally serve as topics – there is a positive correlation between topicality and activation or accessibility (Reinhart 1981, Lambrecht 1994). This correlation is a natural consequence of the informational role the topic relation plays. The referents of topical constituents tend to be active within a discourse because they specify the domains in which information is evaluated and stored. However, the correlation is necessarily imperfect; at some point, every potential topic must be introduced into the discourse and raised to active status.

The effects of this imperfect correlation are evident in TE. Subjects in TE are frequently unaccented pronouns, which in English is the preferred means of encoding highly active referents (Lambrecht 1994). In contrast, indefinite NPs – which tend to code inaccessible referents – are generally barred from the subject position in TE. This constraint is illustrated in (18), which is at least somewhat degraded for most speakers. However, there is not an absolute ban on inaccessible
Licensing by Information Structure

subjects in TE, as shown by (19). In this sentence, the matrix subject is both indefinite and inaccessible. However, the sentence is rescued by the fact that the form this (one) X serves to introduce new topics into a discourse (e.g. Prince 1981c). To summarize, subjects in TE are constrained by the same accessibility requirements that generally apply to the topic relation: inaccessible referents make poor topics unless their form explicitly signals the introduction of a new topic.

(18) (#) At the conference last weekend, a woman was annoying the way she kept cracking her knuckles.
(19) At the conference last weekend, this one woman was annoying the way she kept cracking her knuckles.

Finally, the acceptability of TE sentences improves when the subject co-refers with a ratified (i.e. strongly established) discourse topic (Lambrecht and Michaelis 1998). Though the constructs of “sentence topic” and “discourse topic” are distinct, they are interrelated: the referents of sentence topics tend also to be topical on the discourse level. Thus, the presence of a ratified discourse topic tends to make topic-selecting sentences more natural. Example (20), which is adapted from a naturally occurring discourse, has a prominent discourse topic: What I think of Bill explicitly selects Bill as the topic of conversation. The TE construction in Continuation (1), which appeared in the actual discourse, is fully felicitous in this environment. In fact, English speakers consistently find Continuation (1) more natural in the context provided than in isolation. This effect cannot be attributed to the presence of just any context, as the given context does not improve the felicity of the NE sentence in Continuation (2). In addition, some speakers prefer the TE construction to its NE allosentence in this environment, even though the NE construction is generally judged to be more natural in isolation.

(20) **Context:** What I think of Bill: He's a cool guy. People go on about him not speaking English much, but he's so good at French!
**Continuation (1):** He's great the way he sort of corrects my French when we're talking without making a big deal of it.
**Continuation (2):** It's great the way he sort of corrects my French when we're talking without making a big deal of it.

(adapted from http://www.coastkid.org/hd4a.html)

This double dissociation supports the analysis of TE and NE as informationally distinct allosentences. NE signals an all-focus information structure, while TE additionally expresses the relation topic.

5. **Concluding Remarks**
The analysis of the English TE construction presented here opens the door to a new exploration of the interaction between information structure, argument
structure, and morphosyntax – and in particular, between the relations of topic, argument, and subject. TE subjects, which are classified as such on morphosyntactic grounds, are licensed on the informational level but not on the semantic level. The fact that subjects are not arguments in TE goes against the generalization that a subject “always has a selectional relation with some predicate in the sentence” (Li & Thompson 1976:462). Even if this statement is amended to accommodate expletive subjects that are licensed at the syntactic level, the referential subjects of TE resist the generalization. They also point to broader implications: TE demonstrates that a purely projectionist model of argument structure – that is, a theory that requires that every referential expression within the clause be selected by some lexically represented predicate – cannot fully account for the distribution of referential NPs. Argument structure and information structure interact in the domain of licensing.

One important issue for future research is how general the phenomenon of licensing by information structure is: whether it is restricted to TE, or whether it is a productive operation at the information structure interface. Though to the best of my knowledge TE is the first attested case of licensing by information structure, the phenomenon is certainly consistent with the long tradition of research linking topicality to subject position (e.g., Firbas 1964, Halliday 1967, Reinhart 1981, Lambrecht 1994) as well as the “conspiracy of syntactic constructions” that front predictable or topical information (Prince 1981a:247, see also Horn 1986, Birner and Ward 1998). As such, the possibility that it is a productive process should be carefully investigated. One potential source of the apparent rarity of the phenomenon is that the “input conditions” for the licensing of a topical subject are rarely satisfied: an all-focus construction with an expletive subject that could potentially support the topic relation. Some initially attractive candidate constructions with expletive subjects are out on informational grounds: sentential extrapolation constructions (21) are not necessarily all-focus and it is hard to imagine how so-called “thetic” constructions (22) could support a topic.

(21) It’s amazing that Bill left.
(22) It’s raining.

The typological status of licensing by information structure thus remains an open question. The existence of the phenomenon, though, calls for a reexamination of the scope of the information structure interface, as well as the mechanisms that underlie licensing.

References

Chafe, Wallace. 1976. Givenness, Contrastiveness, Definiteness, Subjects,
Licensing by Information Structure


281


Jennifer Mack
Yale University
Department of Linguistics
370 Temple St.
New Haven, CT 06520

jennifer.mack@yale.edu