For the *Handbook of Semantics*
Craige Roberts, OSU
January, 2009; revised August, 2010

Topics

1. <u>Two notions of topic in discourse</u>

Quite often in making an utterance, a speaker in some way brings our attention to an entity that is relevant at that point in the discussion, in order to tell us something about it. The relevant entity may be an individual or it may be a situation or event. In any such case, we say that the entity to which our attention is drawn is the **Topic** of the utterance. Let us say that the constituent of the utterance which denotes or otherwise indicates the Topic is thereby **Topical**. Depending on the language in use, a Topical constituent may be an N(oun) P(hrase) or an adverbial element of some sort; a Yiddish Topic may even be indicated by preposing a verb stem, per Davis & Prince (1986). The adverbial PP *in any such case* in the third sentence of this paragraph is an interesting case in point. If the PP were removed from the sentence altogether, the passage would be a bit choppy: The first sentence is about circumstances where speakers bring their addressees' attention to such an entity as a prelude to telling us something about it. The second sentence enumerates a class of entities in the world—individuals, situations, events, etc.—which are typically attended to in those circumstances, while the third sentence is again about the circumstances themselves, discussing an arbitrary instance. The PP facilitates a shift in Topic, from the class of entities to the circumstances in which one of them is attended to. The PP could remain *in situ* in the matrix: *We say that the entity to which our attention is first drawn in any such case is the Topic*, but placing it utterance-initially arguably facilitates the smoothest transition from one utterance to the next: Recognizing what the entities are that the speaker is going to tell us about makes it clear how the new utterance is relevant to what comes before, and hence it is easier to readily grasp how the proposition uttered fits into the structure of the speaker's argument. If we take such smooth transitions to be a feature of a maximally cohesive text, then they are motivated by how they foster optimal comprehension.

A typical dictionary definition for *topic* cites two rather different notions: that of the subject-matter of a discussion, and that of the subject of a text (e.g., of a sentence uttered). This is reflected in two uses of the term in the linguistic literature, which, albeit closely related, are different in important respects, leading to some confusion and conflation of the two notions. As Erdmann (1990) puts it "The literature on *topic* (/ *comment*) suffers from the confusion of grammatical ("the first constituent in the sentence") with narrative ("what a text is about") features." I will say something brief about each of these notions, and about how they might be related, then focus on the utterance/"sentence"-specific notion, the one illustrated and discussed in the first paragraph of this essay. In what follows, I take it that an utterance is the ordered pair of the sentence or other linguistic expression uttered (under a linguistic analysis) and the context in which it is uttered (Bar-Hillel 1971). The linguistic phenomena motivating theories of dynamic interpretation might be taken to argue that this characterization is not

adequate, since context can change utterance-internal, but it will suffice for present purposes, especially as Topical constituents in the sense to be considered here typically are root phenomena, i.e. observed only in matrix sentences, at least in English.

With respect to what a discourse ("text", "narrative") is about, much recent work on formal pragmatics, drawing on older work by members of the Prague School and by Halliday and his colleagues, is based on the hypothesis that a central organizing factor in discourse is the *discourse topic*, also called the *issue* or *question under discussion*, or the *Theme* (Halliday 1985). (Note that this is *not* what the Prague School calls *Theme*, discussed below.) *Inter alia*, the topic in this sense is said to play an important role in determining the *prosodic focus* of an utterance (that portion of the utterance which is prosodically most prominent, in English marked by tonic stress) and in constraining the resolution of presuppositions, including those involved in discourse anaphora, domain restriction and ellipsis; in generating conversational implicatures; and in determining the speech act expressed in a given context. For more about the subject-matter of a discussion, see Ginzburg (1996. 2010), Roberts (1996, 2004), and the entry on **Information Structure** in this volume. In order to prevent confusion with the other sense of *topic* in use in linguistics, I will refer to the discourse topic as the **Question Under Discussion (QUD)**.

The term *theme* is similarly ambiguous, referring both to the discourse topic/QUD (whatever the discourse is about) and to that portion of a sentence which is congruent with the QUD in the context of utterance. I'll use the term **Theme** in the latter sense: the constituent or constituents of the uttered sentence which are congruent with the QUD in the context of utterance. In Halliday's (1985) terminology, the Theme is complementary to the **Rheme**, the portion of the sentence which correlates intuitively with the answer to the question being addressed. Those who use the term *topic* for what I'm calling *Theme* generally use the term *comment* instead of *Rheme*, as in the Erdmann quote above. But because of the ambiguity in the term *topic*, there is also marked confusion about the meaning of *comment*, so that it's sometimes taken to be the *Rheme*, in the sense defined just above, and others to be the complement of the Topical constituent. I'll avoid the term here, in the interest of clarity, but one could use it in a way consistent with the terminology I'm using to mean the complement of the Topic of an utterance, if any.

The Theme/Rheme distinction is illustrated by the following:

(1)   Who saw the tornado?
        [Mary]$_{Rheme}$ [saw the tornado]$_{Theme}$

Here *Mary* corresponds to *who* in the preceding question, the QUD in this little discourse. The Theme in the answer is congruent to the QUD in that abstracting on the *wh*-word in the question yields the property 'seeing the tornado', which is the denotation of the Theme, as well. Note that a Theme need not be Topical; arguably, *saw the tornedo* in (1) is not a Topical constituent in the intuitive sense discussed above—the speaker does not draw the addressee's attention to the denotation of the VP, then tell her something about it. Note also that (in English at least) a Rheme is prosodically focused: In (1) *Mary*

would typically bear the nuclear accent in the utterance, with no accent on any words in the Theme. Hence, the Theme in the sense of interest here is often called the Ground of Focus, taken to be complementary to the Focus of the utterance. But this terminology is misleading because sometimes constituents in the Theme can be prosodically focused as well—Halliday's *rhematic themes*, as will be illustrated below. Theme/Rheme here is purely a functional characterization of a partitioning over the sentence uttered, relative to the QUD. E.g., it is not necessarily reflected in (any level of) the syntactic structure of the sentence uttered.

The other notion of *topic* is embodied by a sentential constituent that plays the special pragmatic role in the discourse context sketched informally in the first paragraph, in some sense indicating what the sentence uttered is about. A Topical constituent seems to be more or less what Prague School authors mean by the *theme* of a sentence. And this is quite close to Vallduví's (1992,1993) notion of a *Link*, which he takes to be a constituent that indicates where the information conveyed by the sentence should be entered in the hearer's knowledge store. I'll use the adjective *Topical* to characterize a constituent with something like this special pragmatic role, thereby differentiating a Topical constituent from the Topic it introduces, and use **Topicality** to refer to the discourse functional role(s) associated with a Topic. Note that, as in the example discussed in the first paragraph, a Topical constituent sometimes does not directly denote the Topic: *in any such case* restricts the domain of circumstances to which the generic verb phrase *say that…* applies to those which are topical; it involves quantification over those cases and a quasi-locative preposition *in*; thus it is the domain of the quantifier *any* which is the Topic here, not the whole Topical PP (see Portner & Yabushita 1998 for examples in Japanese).

In what follows, the terms *Topic* and *Topical(ity)* will consistently be used in these senses, and it should be understood that what I am attributing to other authors is not this terminology (since that is used quite differently from author to author), but ideas about the underlying notions.

Ward (1985) speculated that Topics correspond to the Backward Looking Centers (CB) of Centering Theory (Joshi & Weinstein 1981; Grosz, Joshi & Weinstein 1983, 1995; Walker, Joshi & Prince 1998), which attempts to explain what constitutes a felicitous transition between utterances, focusing on structurally or grammatically based prominence rankings of the NPs in the utterances in question. Vallduví (1992) hypothesized a similar correspondence for those Links expressed by NPs, and the Topic/CB correspondence has subsequently been assumed by Gundel (1998) and Beaver (2004), among others. But Poesio, Stevenson, diEugenio & Hitzeman (2004), considering a variety of realizations of Centering Theory, argue that the identification of Topics with CBs is not strongly supported by the corpus data. And Gordon, Grosz & Gillion (1993) do not find experimental evidence for the claims of Centering Theory about preferences for certain types of transitions between utterances in discourse. This is not surprising if we take the core function of Topicality to be indicating what an utterance is intuitively about, since the most prominent Centering algorithms do not attempt to assess aboutness, tending to use features of the structure of the sentence uttered which

can be automatically retrieved, without reflection on meaning. Hence, with CB as usually defined, the correspondence between Topic and CB seems questionable.

Topic is often confused with the grammatical notion of Theme illustrated in (1), the notion sometimes called the *Ground of Focus*. This conflation is fairly natural, since the Theme, corresponding to the QUD being addressed, is old information in one sense, the Topic is often old information in some sense (see below), and both are under discussion. In fact, von Fintel (1996) argued that the two notions of *topic*, what I'm calling *QUD* and *Topic*, are one and the same, and captured the sense in which they are both (typically) old information by making them anaphoric. But Vallduví (1993) and Portner & Yabushita (1998) have argued against the conflation of the Topical constituent with the Theme, for a wide range of languages. Vallduví gives examples like the following, with the Topical constituent here (and below) marked with boldface:

(2)   a.   What about Mary? What did she give to Harry?
                    Mary gave [a shirt]$_{Rheme}$ to Harry.
        b.   What about Harry? What did Mary give to him?
                    To Harry Mary gave [a shirt]$_{Rheme}$.

In (2a) and (b) we have the same Rheme, expressed by *a shirt*, and hence the same Theme, the remainder of the sentence: *Mary gave…to Harry* (in some order). (Note that Vallduví did not use the term *Rheme* in his original examples (2) and (6) below, but marked the NP *a shirt* (in (2)) or the VP (in 6b) as Focus.) But there are distinct Topics, the denotations of *Mary* in (2a) and *Harry* in (2b). In both, the Rhematic portion of the answer is denoted by the constituent with the same grammatical role as that of *what* in the immediately preceding question (the QUD) *What did Mary/she give to Harry?*. Hence, the Topical constituents are only a proper sub-part of the Theme. This argues against the conflation of QUD and Topic proposed by von Fintel; see Portner & Yabushita (1998) for other arguments, especially as applied to the Japanese case.

Here is another thing Topics and Themes have in common: Both may be entirely absent from a given utterance. Poesio et al. (2004) found that the majority of the utterances in their corpus lacked a Backward Looking Center, providing empirical support for this contention. When the utterance has maximally broad focus, for example in response to a question like *What's happening?*, then it is all-Rheme. Similarly, when the rhetorical point of an utterance is not to tell us about a particular entity but to simply note a fact or noteworthy event, there is no Topical constituent. Kuroda (1972, 1992) distinguishes between cases where there is a Topic (and a Topical constituent) and those where there is none via the distinction between **Categorical** and **Thetic** judgments, drawing on the earlier work of Brentano (1973/1874). Ladusaw (1994) illustrates Kuroda's distinction with Japanese examples like the following (my proposed translations). Note that the Japanese nominal particles *ga* and *wa*, to be discussed further below, play a crucial role in signaling the two kinds of judgments in these examples:

(3)   neko ga   asoko de nemutte iru   (Thetic)
                a/the cat there   sleeping is
                'a/the cat is sleeping'

(4)   **neko wa** asoko de nemutte iru   (Categorical)
                the cat   there   sleeping is
                'as for the cat, it is sleeping'

(5)   **neko wa** inu ga   oikakete iru   (Categorical)
                the cat   dog   is chasing
                'as for the cat, a/the dog is chasing it'

Ladusaw continues:

> [The thetic judgment (3)]. . .simply affirms the existence of an eventuality of a certain type. [The categorical judgment (4)] might be used to describe the same situation, but in an essentially different way. . .: [I]t draws attention first to the cat, and then says of the cat that it is sleeping there. Correlated with this difference is the fact that in [(4)], the bare noun marked by *wa* cannot be taken as a nonspecific indefinite cat; it must be a particular cat. This follows from the presuppositional nature of the subject of a categorical judgment. . .

Similarly, in the categorical (5) the addressee's attention is first drawn to the cat, here denoted by the direct object, and then the rest of the utterance tells us about one of its properties.

Here is something else to note about the Topic in examples like (2a) and (2b): The first, *what about X?* question implies a contrast between the mentioned entity *X* and the other members of some implicit set of relevant entities. The second question is, then, about that individual, implying that the comparison is to be made via the property queried in this question. The answer given by the indicative then continues to be about the same individual, contrasted there with the other members of the implicit set with respect to the answer to the second question, the QUD. The denotation of the topical constituent in such a contrastive context (and often, by extension, the constituent itself) is called a **Contrastive Topic**. (I take it this is also what Kuno & Takami (1993:112) call a *Sorting Key*: "In a multiple *wh* question, the leftmost *wh*-word represents the key for sorting relevant pieces of information in the answer." It is also what Jackendoff (1972) calls an *independent focus*.)

Contrastive Topics are generally realized by prosodically focused constituents (in English and German, at least; Jackendoff 1972, Roberts 1996, Büring 2003, Rooth 2005). They typically carry a special type of prosodic contour, the so-called *B-accent* of Jackendoff (1972); in ToBI transcription (Beckman & Ayers 1994), this is the contour L+H* LH%. This illustrates why it would not be desirable in general to conflate the notion of Topic(al constituent) with that of the Focal Ground: It is not generally the case that a Topical constituent is in that portion of the uttered sentence which contains no prosodic focus. To underline this, Vallduví also gives minimal pairs with the same Contrastive Topical NP but two different Rhemes, both Focused constituents:

5

(6)   a.   What about Mary? What did she give to Harry?
              Mary gave [a shirt]$_{Rheme}$ to Harry
       b.   What about Mary? What did she do?
              Mary [gave a shirt to Harry]$_{Rheme}$

In fact, even non-Contrastive Topics are quite often prosodically focused (in English, at least), particularly those which are displaced to occur utterance-initial; for example, a realization in speech of the adverbial *in any such case* in the third sentence of this article bears focal prosody, constituting an entire phonological phrase in the sense of Beckman (1996). Since each phonological phrase includes a pitch accent, hence a nuclear accent, each such phrase has a Focus (Selkirk 1996, the article on **Focus**). The fact that the Rheme in examples like those in (2) and (6) may be a proper part of the complement of the Topical constituent reflects one of Prince's (1998) observations about English examples like (2b) (see also Ward 1985): that the non-Topical tonically stressed constituent in the utterance should be replaced with a variable, the result representing an "open proposition" which is "saliently and appropriately on the hearer's mind at that point in the discourse, the tonically stressed constituent representing the instantiation of the variable and the new information in the discourse". E.g.., in (2a,b) or (6a) we replace the tonically stressed *a shirt* with *x* to yield the open proposition *Mary gave x to Harry*; this is congruent with the QUD *What did Mary give Harry?*, which, as QUD, is on the mind of any attentive participant in the discourse.

Hence, examples with Contrastive Topics, like those in (6), illustrate a relationship between the two notions of topic in certain contexts, one which has been argued for in detail by Büring (2003). In such contexts, an entire section of a discourse reflects a strategy of inquiry (Roberts 1996) wherein the speaker singles out first one, then another member of some relevant set of entities, considering in turn which relevant property each of these entities has. Consider (7):

(7)   [after a trip to the zoo] What about the African animals? Who saw which animal?
       a.   **The zebra**, [Mary]$_{Rheme}$ saw
       b.   **The elephant** was seen by [Harry]$_{Rheme}$
       c.   and [Zach]$_{Rheme}$ saw **the giraffe**.

In (7a), *the zebra* is Contrastive Topic, marked both by the B-accent typical of Contrastive Topics and by Topicalization—the English construction wherein the Topical NP occurs sentence-initial, serving as the filler for a gap in the matrix clause—here in the direct object position following the verb. In (7b) the Contrastive Topic is *the elephant*, made utterance-initial by passivization and bearing B-accent. And in (7c) *the giraffe* is marked as Contrastive Topic by bearing B accent, even though it remains *in situ*. The speaker is effectively answering one sub-question of the explicit QUD at a time: *who saw the zebra?, who saw the elephant?, who saw the giraffe?*, one question for each (relevant) African animal. This illustrates how the overt question being replied to (*Who saw which animal?*) needn't be the QUD implicitly assumed by the speaker of an utterance, as

reflected in the utterance's Theme. The possibility of an implicit QUD is even clearer in the following:

(8)   (No prior discourse, at least on a related subject)
        A: [When are you going to China]$_F$?
        B: Well, I'm going to **China** in [April]$_{Rheme}$.   (Roberts 1996)

Here, B answers A's question, with A-accent on the Rheme *April*, but also uses B-accent on *China* to mark it as a Contrastive Topic, presupposing that there is a larger set of relevant entities (countries) for which one might pose the question of when B is going to visit them, and implicitly inviting A to inquire about those as well.

As illustrated by (7c) and (8), though Topical constituents are often utterance-initial, they need not be, at least not in English. So the notion of a Topical constituent should not be confused with that of a Topicalized constituent: Topical constituents needn't be Topicalized, and, as we will see below, Topicalized constituents are not always Topical. This is cautionary: We must be careful not to take *one way topicality may be encoded* to be the *unique* way it is encoded or even to always encode topicality. In studying the relationship between Topicality and Topic-marking in a given language, we must carefully control for function, in order to determine Topicality, and then determine what expressive options exist, and conversely, for any given construction, determine what functional roles for a particular sub-constituent are consistent with use of the construction across a range of possible contexts of utterance.

Table 1 summarizes the terminology and distinctions proposed here:

Notions of Topic:	Discourse Topic: **QUD**	Utterance **Topic**
Linguistic correlates:	Interrogative sentence; in indicative sentences reflected by the **Theme** (portion of sentence congruent with QUD). Cf. the complement of the Theme: the **Rheme**(s): the answer portion, correlated with the *wh*-element(s) in a preceding interrogative	**Topical constituent** (NP, AdjP, PP, adverbial phrase or clause, etc., sometimes *in situ*)
comparison:	Topical constituents, including those which are Topicalized or Left Dislocated, may be proper sub-parts of the Theme of an utterance. The traditional Topic/Comment distinction consists of the Topical constituent and its complement in the sentence uttered; this does not coincide with the distinction between Theme and Rheme. Every utterance has a Rheme, but Themes and Topical constituents are optional. When there is a Topical constituent, the utterance is **Categorical**; otherwise, it is **Thetic**. I.e., **Theme ≠ Topical constituent**	
Possible conventional indication(s) in an utterance:	Theme is generally unmarked, while Rheme is marked, e.g. by prosodic status and/or word order	Topicalized constituent (English) Left Dislocated constituent (English) Passive subject (English) *wa*-marked constituent (Japanese) *nun*-marked constituent (Korean)
comparison:	Thematic-status is the *unmarked case*, whereas Topicality tends to be *marked*, so the indications noted for Themes tend to be indirect, and may be over-ridden by other factors. E.g., English Rhemes are always prosodically focused, with Themes often unaccented; but some sub-constituents of Themes may be focused as well, e.g. Contrastive Topics. And while the Theme of an utterance tends to stay *in situ*, Topical sub-constituents of the Theme quite often do not.	
Pragmatic function	establishing or confirming the current discourse goal determines what's relevant	**topicality:** directing addressee attention to some relevant discourse referent
comparison:	Questions have a special status in the organization of discourse (Information Structure), and hence Theme/Rheme structure is generally associated with the Focus structure of an utterance. But we may wish to direct our addressees' attention to other relevant entities as well—individuals, events, situations, etc.—and these entities are thereby Topical. When the purpose is to contrast the Topical entity with others in a relevant set, we have a **Contrastive Topic**.	

Table 1: Notions of Topic in discourse

2. **Proposed Tests for Topicality**

The *what about* questions in (2) and (6) illustrate one of the tests proposed in the literature for identifying Topical constituents in English, the *what about* test of Gundel (1974,1985). A similar test is due to Reinhart (1981): the *say about X that S* test. Two others due to Gundel are the *as for* and *speaking of* tests. These tests and the differences between them are illustrated in the following examples (see Ward 1985, Chapter 2 for consideration of other proposed tests in the literature):

(9)   Then Tom Cruise went to work for Francis Ford Coppola, on this spring's semi-successful film version of "The Outsiders". <u>Coppola</u> he found to be "just like one of the guys. And he totally trusted me. He let me go anywhere I wanted to go with the character…" [*Philadelphia Inquirer*, p.8-C, 9/1/83, "His star is rising, but his feet remain firmly on the ground.", cited in Ward 1985:73]
       Tests:
            (i)   <u>About Coppola</u>, he said that he found him to be…..
            (ii)   #<u>What about Coppola?</u> He found him to be…
            (iii)   #<u>As for Coppola</u>, he found him to be…..
            (iv)   #<u>Speaking of Coppola</u>, he found him to be…..

The *about* test succeeds in (9), i.e. yields a felicitous, sense-preserving substitute for the underlined target constituent. Arguably this is because the utterance including the target *is* about Coppola, i.e. Coppola is the Topic of the utterance. But the *what about* and *as for* tests seem to implicate a contrast between Coppola and some other salient entity or entities, a contrast which isn't drawn, explicitly or implicitly, in this discourse. That is, the denotation of *Coppola* in the original text is not a Contrastive Topic, nor is there some salient contrast set to which he is being compared, so these tests strike one as infelicitous in this context (as indicated by '#'). *Speaking of X* does not necessarily contrast *X* with some other relevant entity. It does seem to presuppose that Coppola has been recently mentioned, but also that he was not under discussion at that point: Perhaps he was only mentioned in passing, in some other connection. Then the adverbial serves to signal a shift to talking about him. But in the context given, one might argue that at the end of the first sentence Coppola, or at least Coppola's relationship with Cruise, is as plausibly under discussion as is Cruise himself, so the implication of a shift seems unnecessary and infelicitous. In terms of Centering Theory, *speaking of X* seems to indicate might be called the establishment of a Backward Looking Center (CB) which was not in the set of Forward Looking Centers (CF) of the previous utterance. See Poesio et al. (2004). But Coppola *is* in the CF for the first sentence here.

Consider how the tests work with another example. (10) uses the four test frames as ways that the indicative reply to the QUD in (6a) or (6b) might be continued, on the assumption that Sue and Mary are both (merely) implicitly known by the interlocutors to both be members of some salient and relevant set of individuals:

(10)   a.   What about Mary? What did she give to Harry?
           Mary gave [a shirt]$_{Rheme}$ to Harry

b.   What about Mary? What did she do?
                  Mary [gave a shirt to Harry]$_{Rheme}$
       Possible continuations of the indicative responses in (a) and (b):
                  (i)   #<u>About Sue</u>, (I would say that) she gave Harry a scarf.
                  (ii)   <u>What about Sue?</u> (What did she do?)
                  (iii)   <u>As for Sue</u>, she gave Harry a scarf.
                  (iv)   #<u>Speaking of Sue</u>, she gave Harry a scarf.

About Sue seems to presuppose that Sue was already mentioned (as we saw in (9)); since this is not the case in the contexts given, it is infelicitous in these examples. This also leads to infelicity of *speaking of Sue*. But *what about Sue?* (with or without repetition of the main question from (10a/b)) simply seems to extend the contrastive strategy begun with *What about Mary?*, as does *as for Sue*, so that both are felicitous so long as Sue and Mary are plausibly both members of the same relevant contrast set. (Ward (1985) has claimed that this should be a partially ordered set, but that doesn't seem to be necessary in general. Here, for example, a simple set of close friends would suffice.)

The following (constructed) discourse illustrates appropriate use of *speaking of X*:

(11)   A: I was at the mall yesterday and I ran into Louise Clark, who was here visiting
           Sue Topping.
      B: Interesting. [interlude of talk about Clark, followed by:]
           (i)   #<u>About Sue</u>, Louise said that…
           (ii)   <u>What about Sue?</u> {What's she up to?/I heard she was moving.}
           (iii)   #(But) <u>as for Sue</u>, did you know…
           (iv)   But <u>speaking of Sue</u>, did you know she's engaged?

With *speaking of*, Sue is a non-contrastive Topic, because she isn't being compared with any other salient entities in some relevant respect—the speaker merely switches to talking about Sue and offers some interesting news about her. After *what about Sue?* Sue is also the Topic, but she may also be Contrastive Topic, indicated by the fact that reference to Sue in the following utterance involves B-accent. This is the case in the follow-up question *What's she up to?,* where Sue is being contrasted with Louise in respect to what each is up to and *she* receives B-accent. In the alternative follow-up *I heard she was moving*, if *she* receives B-accent, Sue is contrasted with Louise Clark (who presumably has moved away), while if *she* is unaccented Sue is merely Topic, without implied contrast with respect to the predicate in question, *moving*. The example also shows that felicity conditions on use of *what about X* differ from those for *as for X*. *What about Sue?* is felicitous in this context without prior indication that Sue and Louise were to be contrasted in this connection, even if Sue is the Contrastive Topic in the subsequent utterance. But *as for Sue* seems to presuppose that the speaker is working her way through a salient contrast set, indicating a turn of attention to someone already understood to be a member of that set; the contrast in this example might be facilitated by the use of *but*. Since there is no such salient set in the context given, (iii) is infelicitous.

As we have seen, these four tests proposed for Topicality in English actually differ in the felicity conditions involved, though they all do seem to insure Topicality for the entity mentioned (Copolla or Sue in these examples). But given these subtle differences, we can readily imagine that any of these tests might not be available in direct translation in another language, or that that language might offer ways of testing for Topicality other than direct translations of these particular utterance frames. Keeping this in mind, we turn to consider how Topicality may be marked cross-linguistically.

3. <u>Topic-marking across languages</u>

Although most of the examples considered above are English, Topicality is, of course, marked in some way or other in most, if not all human languages. A first pass through the catalog of ways in which various languages encode Topicality leads one to observe that they frequently involve syntactic constructions which place the Topical constituent utterance-initial, or near-initial, whether these involve a filler-gap relationship (e.g. Topicalization) or not (e.g., so-called Left Dislocation, where the utterance-initial dislocated constituent is associated with a coreferential resumptive pronoun *in situ* in the root clause). As we will see, in some languages, Topicality is at least sometimes morphologically marked, this marking sometimes combined with placement of the Topical constituent at the left-periphery of the root clause. However, when we look carefully at the specifics of the relevant constructions and morphological markings, we find that they vary a great deal from language to language, in syntactic detail and/or in corresponding pragmatic function(s) and felicity conditions. So though there may be many common factors, in the end it appears that what these various conventions realize is a family of closely related notions, rather than a syntactic or pragmatic universal. Following are some examples very briefly illustrating this diversity across a range of language families. Topical constituents are in boldface throughout.

Saaed (1984) claims that in **Somali** (Chadic) preposed Topical NPs are typically Left Dislocated, not corresponding to a syntactic gap in the root clause, as in his examples (12) and (13). (In Saaed's glosses *FOC* is a focus particle, which follows a focused NP. As he notes, the Topical NP in (13), *suuga*, is not grammatically locative.)

(12) **shandadaha** kuwa birta   ah baa   ka culus kuwa santa ah
      suitcases$_{def}$ those metal$_{def}$ are FOC more heavy those leather are
      'Suitcases, those which are metal are heavier than those which are leather.'
(13) **suuga**   hilib geelku aad buu   qaalisan   yahay
      market$_{def}$ meat camel very FOC   expensive is
      'The market, camel's meat is very expensive.'

This differs from **Hungarian** (Kiss 1998), a Finno-Ugric VSO language which has distinguished preverbal positions for Topical constituents, quantificational NPs (or *DP*s), and Focused NPs/PPs, in that order. Left dislocated constituents occur before Topical NPs at the left periphery of the root clause. Topical NPs bind a gap in the root clause, as in her examples (14) and (15):

(14) **Zsuzsának** [[János]$_{Foc}$ mutatta   be   Imrét]
      Susan.DAT John   introduced prev Imre.ACC
      'Susan was introduced to Imre by JOHN.'
(15) **Imrét** [[Zsuzsának]$_{Foc}$ mutatta   be   János]
      Imre.ACC Susan.DAT   introduced prev John
      'Imre was introduced by John to SUSAN.'

Note that neither (14) nor (15) is a passive structure; the Case-marking of the NPs is the same in both, regardless of their position and role as (non)-Topic. See also Roberts (1998) for more discussion of the Hungarian left periphery and its relationship to Focus.

According to Sturgeon (2006) the Slavic language **Czech** has two kinds of Left Dislocation constructions which can mark Topicality, Contrastive Left Dislocation (CLD) and Hanging Topic Left Dislocation (HTLD). But these differ both syntactically—for example, only CLD displays connectivity effects between the left dislocated constituent and the resumptive pronoun in the main clause—and pragmatically: While HTLD "promotes the discourse referent of the left dislocate to topic status", CLD marks the left-dislocate as a Contrastive Topic.

All three of these languages display the tendency to place Topical constituents on the left-periphery of the root clause, as in the English examples (2) – (7) above. But we also saw that in **English** there are a variety of constructions which achieve this: Topical NPs may be preposed via Topicalization, as in (2c) or (7a) above, but may also be promoted to subject by Passivization, as in (7b), or even be left *in situ* in subject position, as in (2a), (6a) or (6b). Prince (1998) argues that Left Dislocation can also mark Topicality in English, as in (16), which involves the mixed use of Left Dislocation (*one* and *another*) and Topicalization (*the third*):

(16) She had an idea for a project. She's going to use three groups of mice$_{i,j,k}$. **One$_i$**, she'll feed them$_i$ mouse chow, just the regular stuff they make for mice. **Another$_j$** she'll feed them$_j$ veggies. And **the third$_k$** she'll feed e$_k$ junk food. (SH, 11/7/81, reported in Prince '98:287)

But occurrence on the left-periphery is neither necessary nor sufficient to mark a constituent as Topical in English. English Topical constituents may also remain *in situ* post-verbally, as in (7c) or (8), where the Contrastive Topics are prosodically marked with the B-accent. And Prince offers evidence that Left Dislocated NPs may play another role than that of (Contrastive) Topic. This non-Topical role is illustrated by her (17) (with the target constituent underlined for ease of identification):

(17) 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. <u>The landlady$_k$, she$_k$ went up</u>, and he laid her$_k$ out. So sister went to get a wash cloth to put on her$_k$, he stabbed her in the back. But she saw her death. She went and told my mother when my brother was buried, "Mother," she said, "your trouble is not over yet.

You're going to have another death in the family. And it's going to be me." And sure enough, it was. *Welcomat*, 12/2/81, p.15.

In (17), the Left Dislocated NP *the landlady* has not been mentioned previously, i.e.in Prince's terms is Discourse New (although weakly familiar in the sense of Roberts 2004). Prince points out that the discourse is not about the landlady at this point, and in keeping with this, the example fails a number of the tests for Topicality considered above. None of the topicality tests in (18i) – (iii) is felicitous in the context given:

(18)   [Context] A: He started fussing with my sister and she started to scream.
        i. #She said he started fussing with her sister and she started to scream. She said about the landlady that she went up and he laid her out. So she said her sister went to get a washcloth to put on her, he stabbed her in the back.
       ii. B: #What about the landlady?
           A: The landlady, she went up and he laid her out**.** So sister went to get a washcloth to put on her, he stabbed her in the back.
      iii. #He started fussing with my sister and she started to scream. As for/Speaking of the landlady, she went up and he laid her out.

The *about* test fails in (18i) because the utterance in the original text is *not* about the landlady, but (still) about the occasion of the sister's murder. Nor is the landlady being contrasted in some relevant respect with other participants in the event, so that she isn't a Contrastive Topic. So switching to talk about her makes no sense in this context, yielding a discourse which is not optimally coherent. Similarly, because a *what about X* question makes the denotation of *X* the Contrastive Topic in any felicitous reply, introducing this question in (18ii) also leads to discourse incoherence, and in any case to a discourse which seems to switch Topic in a way not displayed in (17). The *as for* and *speaking of* tests also lead to infelicity in (18iii). Not only do they trigger the infelicitous Topic-switch, but both seem to have an additional presupposition, that the entity inquired about is already salient in the context. But the landlady in (18iii) is Discourse-new (Prince 1992) in this context, and so could not reasonably be taken to be relevant or salient in the discourse before her mention. Finally, the *as for* test seems to presuppose that the Topic is being contrasted in some respect with another entity recently under discussion, hence infelicitously making it a Contrastive Topic.

Prince argues that Left Dislocation is used in (17) because subject positions are dispreferred for Discourse-new entities (Halliday 1967; Geluykens 1992), so *the landlady* is Left Dislocated to introduce that person into the discourse, thereafter referring to her pronominally. Prince calls this use an instance of "simplifying Left Dislocation", about which she says that it "serves to simplify the discourse processing of Discourse-new entities by removing the NPs evoking them from a syntactic position disfavored for NPs evoking Discourse-new entities and creating a separate processing unit for them. Once that unit is processed and they have become Discourse-old, they (or, rather, the pronouns which represent them) may comfortably occur in their canonical positions within the clause." This kind of careful consideration of the discourse function(s) of occurrence on

the left periphery is lacking (so far as I can tell) in the discussions of topicality in other languages cited above.

In **Japanese**, Topical NPs and sometimes other Topical constituent types as well, are said to be morphologically marked with the enclitic *–wa* (see Kuroda 1965, Kuno 1973, and a wealth of literature; see the useful overview in Heycock 2007). Something quite similar occurs in **Korean**, which marks Topicality with *–(n)un* (Choi 1997, 1999; Lee 2003). We get a feeling for their distribution from the following story fragment in the two languages, due to Kubota & Lee (2007). (In these examples and those that follow, *NOM* is nominal case, *ACC* accusative case, *DECL* declarative; *HON* is honorific.)

(19)   **Japanese**:
Gakusei-ga/ #-**wa**   kyoozyu-to   menkaisi-te i-ta.
student-NOM/ TOP professor-with meet PROG-PAST
'a student was meeting with a professor'
　　Gakusei-wa   syukudai-nituite   situmon-o   si,
　　student-TOP   homework-about   ask.questions-ACC do
　　　　kyoozyu-wa   sore-ni   teineini   kotae-ta.
　　　　professor-TOP   it-DAT   carefully   answer-PAST
　　　　'the student asked questions about homework and the professor answered them carefully'
　　　　　　Hutari-wa   gakusei-no   taamu-purojekuto-nituite-mo hanasi-at-ta.
　　　　　　two-TOP   student-GEN   term-project-about-also   discuss-PAST
　　　　　　'they also talked about the student's term project'
　　　　　　　　Totuzen   kyoozyu-ga/ ?-**wa**   yuka-ni   taore-ta.
　　　　　　　　suddenly   professor-NOM/ TOP floor-DAT fall-PAST
　　　　　　　　'suddenly, the professor fell on the floor'

 (20)   **Korean**:
Han   haksayng-i/ #-**nun**   han   kyoswu-nim-kwa   manna-koiss-ess-ta.
one   student-NOM/ TOP one   professor-HON-with   meet-PROG-PAST-DECL
'a student was meeting with a professor'
　　Haksayng-un   swukcey-ey.tayhaye   mwul-ess-ko,
　　student-TOP   homework-about   ask-PAST-and
　　　　kyoswu-nim-un   chincelhakey   taytaphaycwu-si-ess-ta.
　　　　professor-HON-TOP   kindly   answer-HON-PAST-DECL
　　　　'the student asked him about the homework, and the professor answered the question kindly'
　　　　　　Twu   salam-un   kimal-puloceyktu-ey.tayhaye-to iyakiha-yess-ta.
　　　　　　Two people-TOP term-project-about-also   talk-PAST-DECL
　　　　　　'they also talked about the term project'
　　　　　　　　Kapcaki   kyoswu-nim-i/ ?-**un**   patak-ey ssuleci-ess-ta.
　　　　　　　　suddenly professor-NOM/ TOP   floor-on   fall-PAST-DECL
　　　　　　　　'suddenly, the professor fell on the floor'

In each, the first sentence introduces a pair of (discourse-new) participants, and use of the TOP-marker *–wa* or *–nun* is infelicitous. In the second and third clause in each language TOP-marker *-wa* or *–nun* is acceptable on the subject, denoting one of those participants; the same is the case for the subject of the fourth clause, denoting the pair. In these three

14

central clauses, the NOM-marker *-i* is also acceptable on the subject in Korean and *-ga* is marginally acceptable in Japanese, though there seems to be a clear preference for a uniform choice of marking throughout the sequence: all TOP (preferably) or all NOM. Note that although there arguably are uses of *–wa/-nun* that are Contrastive Topics, the TOP-marked subjects in these examples would not be treated as Contrastive Topics if translated into English, e.g. associated with the B-accent or paraphrased with *as for the student/as for the two of them*; they would not be the contrastive topics of Lee (2006a, 2006b). Nor is it clear that the *about X* test gives positive results here. The final utterance is of particular interest: In each language, although the subject is one of the same familiar entities TOP-marked previously, native speakers are clear that the nominative case-marker is preferable to *–wa* or *-nun*. Kubota & Lee hypothesize that the sharp shift marked by 'suddenly' is what makes the TOP-marking infelicitous. One might hypothesize, following Kuroda (1972), that unlike the preceding utterances reporting categorical judgments about the participants, the adverbial signals that what's of interest in the final utterance is the unexpectedness of the event, reported by a thetic judgment, and not what is said about the subject.

There are a number of other parallels between the uses of these particles in the two languages. For example, each language permits the use of the TOP-marker for a so-called "double subject" construction where the *-wa* or *–nun*-marked constituent does not correspond to a gap in the root clause, as in the following Japanese example (Kiss 1998) and its Korean counterpart (Jungmee Lee, p.c., omitted for reasons of space):

(21)   Sakana wa   tai ga   oisii.
         Fish   TOP red.snapper   delicious
         'as for fish, red snapper is delicious'

In this construction, as reflected in the gloss, the *wa*-marked NP does appear to function as a Topic, in the sense of what the utterance is about.

Also, each language admits of a second use of *–wa/-nun*, a non-Topical function indicating exclusivity (often called a "contrastive use" in the literature; see Kuno 1973; and more recent work on contrastive *–wa* by Noda 1991, Hara 2004, Oshima 2008 and Sawada to appear)—note that this is not always a Contrastive Topic in the sense defined above (and see Lee 2006a, 2006b). The non-topical use is illustrated by this Korean example (Jungmee Lee, p.c.), where the second *–(n)un*-marked phrase is non-Topical:

(22)   Mary-nun   John-un   chohaha-n-ta.
         Mary-TOP John-TOP like-PRES-DECL
         "as for Mary, she likes John (but not others)"

(22) can serve as an answer to *What about Mary? Who does she like?* (or *Who likes her?*), but not to *What about John? Who does he like?/Who likes him?*, arguing that *Mary-nun* here is a Contrastive Topic, while *John-un* is not Topical but instead serves as the focused Rheme.

Besides NPs and predicates (Heycock 2007), Korean and Japanese can mark with *–nun/-wa* a variety of adverbial constituents, as with the temporal adverbial clause in the Korean (23) (Jungmee Lee, p.c.):

(23) [nay-ka ttena-l-ttay   -nun]   John-i   ca-koiss-ess-ta
        I-NOM leave-REL-time-TOP   John-NOM sleep-PROG-PAST-DECL
        'as for the time when I left, John was sleeping'

(23) also has a non-Topical exclusive interpretation: 'John was sleeping when I left (but not at other times)'. The temporal adverbial may instead follow the subject *John-i*, in which case it has the same two readings.

But for all these parallels, Japanese *–wa* and Korean *–nun* differ in at least one important respect: sensitivity to the familiarity, or Givenness of the constituent they mark. Portner & Yabushita (1998; 2001) provide evidence that non-contrastive *–wa* NPs usually presuppose definiteness, technically realized as a kind of familiarity. Kubota (2007) argues that this should be Weak familiarity, in the sense of Roberts (2003). But Portner & Yabushita also observe that Japanese quantificational NPs can occur with *–wa* just in case they are presuppositional, which "indicates that [the] quantifier's domain is contextually given in some way or another." This is illustrated in their (2001:286) examples following:

(24) Heya   wa subete sansetto biichi ni   menshite-imasu.
        Room(s) TOP every Sunset   Beach LOC facing-be
        'every room faces Sunset Beach'
(25) Sansetto biichi wa   subete no heya   kara miemasu.
        Sunset   Beach TOP every of room(s) from seeable.
        'Sunset Beach can be seen from every room'
(26) Taitee no kan'kookyku ga   ∅   kin'itte-imasu.
        most of   tourists   NOM pro pleased-be
        'most tourists like it/them/that'

(24) and (25) have "essentially the same propositional content" but establish distinct topics, which lead to distinct possibilities for resolving subsequent anaphora. Following (24), the null object *pro* in (26) is understood to refer to either the rooms (the Topic and domain of *subete*) or the idea that every room faces the beach; while following (25), the *pro* can be taken to refer to either Sunset Beach (the Topic, denotation of the NP *Sansetto biichi*) or to the fact that the beach is seeable from every room.

In this vein, Kubota & Lee (2007) provide data arguing that *–wa* is only felicitous with those quantificational NPs which admit of a partitive interpretation, presupposing a familiar domain of quantification, and so presuppositional in Portner & Yabushita's sense. Hence the contrast between Kubota & Lee's (27), where the partitive interpretation is not available, and (28), where it is, as reflected in the gloss:

(27)　#{san-nin/go-nin-izyoo/takusan/oozei/kazoekire-nai-hodo/nan-nin-mo}-no
　　　　three-CL/five-CL-more/many/many/countless/numerous　　　-GEN
　　　　　gakusei-wa　siken-o　　uke-ta
　　　　　student-TOP　exam-ACC take-PAST
　　　'three/more-than-five/many/countless/numerous students took the exam'
(28)　Ooku-no　　gakusei-wa/ga　　　siken-o　　uke-ta
　　　many-GEN　student-TOP/NOM　exam-ACC take-PAST
　　　'Many of the students took the exam'

(With –*ga* (28) can also mean non-partitive 'there were many students who took the exam'.)

But Korean does allow non-partitive quantificational NPs with –*nun*, so that the Korean counterpart of (27) in (29) (as well as that of (28)) is grammatical:

(29)　{Sey-myeng-uy/ tases-myeng-isang-uy/ manhun}　haksayng-tul-un
　　　three-CL-GEN/five-Classifier-more_than-GEN/many　student-PL-TOP
　　　　sihem-ul　　chi-ess-ta
　　　　exam-ACC　take-PAST-DECL
　　　'three/more-than-five/many students took the exam'

The acceptability of (29) reflects that fact that, more generally, Korean contrastive –*nun* is acceptable with (non-specific) indefinites like *etten saram* 'certain man', in contexts where Japanese –*wa* would result in infelicity (Kubota & Lee 2007).

Italian presents yet another range of options for indicating Topicality. According to Rizzi (1997), in Italian there is a sharp distinction between the Clitic Left Dislocation (CLLD) structure, involving an IP-internal resumptive clitic coreferential to an initial Topical NP, as in his (30), and contrastive focus preposing, involving a preposed NP linked to a matrix-internal gap, as in his (31):

(30)　　**Il　tuo　libro**, lo　　ho　letto
　　　　def 3sg. book　clitic perf read
　　　　'Your book, I have read it.'

(31)　[Il　　tuo　　libro]_{Foc} ho letto (, non　il suo).
　　　def 2sg.poss book　　perf read　not def 3sg.poss
　　　'Your book I read (, not his)'

Rizzi isn't clear about the contexts in which these utterances would be felicitous. About (30), Rizzi notes that "the English gloss…is somewhat misleading" (286). And about (31) he says that "It could not be felicitously uttered as conveying non-contrastive new information, i.e. as an answer to the question 'What did you read?'." But since the gloss (with Rizzi's labeling of the preposed constituent as *Foc*) makes it clear that the preposed direct object is contrastive, it may be a Contrastive Topic: Recall that in English topicalized NPs in general, and topicalized Contrastive Topics in particular are Foci.

Arregi (2003) argues that in Spanish CLLD the left dislocated phrase is a Contrastive Topic, in the sense defined above. But there are differences from the Italian case, where CLLD obligatorily involves a clitic: Spanish left dislocated phrases display an interesting semantic pattern wherein those with clitics must be reconstructed at LF (to capture the correct scope potential of Left Dislocated indefinites), while indefinites may undergo Left Dislocation without clitics, and in the latter case they do not undergo reconstruction.

In French, the sharp distinction Rizzi draws between the pragmatic functions of CLLD and constructions like (31) does not seem to obtain. Abeillé, Godard & Sabio (2008) argue that in French, as we saw earlier for English, there is more than one way to mark a constituent as Topical. French CLLD is one of these. But they consider two other types of French construction where an NP occurs on the left periphery of the clause without an overt clitic—one in which the filler-gap relationship is sensitive to islands, as in English Topicalization and Rizzi's (31), and the other in which it is not, as in CLLD. They give evidence that in each of these French constructions the initial NP may be Topical, though it need not be. Moreover, in both constructions the preposed NP may alternatively serve as the (narrowly Focused) Rheme, in marked contrast with Rizzi's claim about the Italian constructions.

There is another striking difference between the Italian and English constructions. In the Italian, CLLD is acceptable in embedded clauses, as in the relative clauses in Rizzi's (32) and (33) (1997:306), whereas English Left Dislocation is a root (main clause) phenomenon, as we see in (34) and (35):

(32)   Un uomo a cui, il tuo libro$_i$, lo$_i$ potremmo dare.
        'A man to whom, your book, we could give it'
(33)   Un uomo che, il tue libro$_i$, lo$_i$ potrebbe comprare
        'A man who, your book, could buy it'

(34)   *A man to whom your book$_i$ we could give e$_i$
(35)   *A man who your book$_i$ could buy e$_i$.

The tests we considered in the preceding section all seem to treat Topicality as a function of the utterance, hence appropriate only at the root level. This prediction is borne out in English, but not in Italian, calling the pragmatic status of the Italian construction even more into question: Does *your book* in (32) or (33) stand in contrast to some other entity? Or is it what the whole utterance is about in some sense? Note, though, that Italian is not alone in permitting such embedded Topic-like constituents. For example, this is also observed in **Tz'utujil** (Mayan) according to Aissen (1992), and in Korean (Jungmee Lee, p.c.). But embedded non-contrastive *–wa* marked NPs are not acceptable in Japanese (Yusuke Kubota, p.c.).

The above illustrate just a few of the ways that languages differ in expressing something taken by theorists to be Topic. Though space precludes elaborating further, we might mention briefly a couple of other interesting differences gleaned from the literature:

a) Many languages, including Catalan (Valduví 1992:48), Korean (Jungmee Lee, p.c.), Somali (Saeed 1984), Hungarian (Kiss 1998), and Yucatec Mayan (Judith Tonhauser, p.c.), permit multiple preposed NPs which are said to be Topical (in some sense), whereas Reinhart (1981) convincingly argues that English generally only permits one Topical NP, whether Topicalized or Left Dislocated, at least according to her tests.

b) Some languages treat *wh*-elements as non-Topical, while others treat (some of) them as Topical. For example, Hungarian *wh*-NPs are preposed between the distinguished pre-verbal position for Topical constituents and that for Foci. English Topical constituents and preposed *wh*-elements may co-occur, in that order (*As for Tom, who likes him?*), and given that the language does not permit multiple Topical constituents, this argues that they are in distinct syntactic positions, as well as functionally distinct. But Bulgarian (Jaeger 2001) allows Clitic Doubling of *wh*-phrases, in which case they behave like Topical constituents, occurring before other *wh*-expressions and often interpreted as an echo-question or Contrastive Topic; see also Grohmann (2006). Kubota & Lee (2007) show that in Japanese only D-linked (Pesetsky 1987)—and hence familiar—*wh*-expressions may be marked with *–wa*, and the same seems to hold in Korean.

All this underlines the importance of trying to develop clear pragmatic tests and/or criteria (contextual felicity conditions) for determining whether, in a given language, a particular construction or other (e.g. morphological) means is consistently used to mark a constituent as Topical, and to determine whether that notion of Topicality corresponds with the notion as reflected in other languages, like English. Until we do this, we may well be comparing apples and oranges. And it raises a question that needs to be addressed in a rigorous way language-by-language: Just because a particular construction in a particular language serves as the best-translation for another construction in a distinct language, does that mean that the two have the same functional load from a semantico-pragmatic point of view?

4. The Universality of Topicality

What can we say about Topicality in human language generally? Is it reflected in a universal at some level, e.g. syntactic or pragmatic? A number of authors offer extended discussions of these questions: Reinhart (1981), Ward (1985), Gundel (1988), von Fintel (1988), McNally (1995), Prince (1998), Portner & Yabushita (1998), Jacobs (2001), Portner (2002), and Ward & Birner (2004), among many others. There is not the space here to go through their proposals in detail. But we can mention two features of Topics that many of these authors agree on, although, again, the terminology they use often differs:

A Topic is familiar.
> A Topic is must be something that is either familiar (or given) itself, or is an identifiable member of some familiar set of entities.
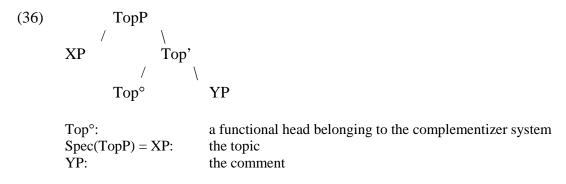
A Topic is what the utterance in which it occurs is about.

Of course, it has been noted by many that a Topical NP, in English for instance, need not be definite. But at least in Japanese, if a Topical NP is indefinite, it must be specific. Portner & Yabushita (1998, 2001) argue for a view of specificity in which the indefinite NP in question is taken to have a familiar (singleton) domain, unifying this case with that of the definites. Note that if the familiarity of Topics is taken to be the Weak Familiarity of Roberts (2003), which seems to be appropriate for both Japanese and English, it subsumes Portner & Yabushita's specificity.

Portner & Yabushita formally realize the *aboutness* criterion by taking the information conveyed by the utterance to update the discourse referent corresponding to its Topic, if it has one. This update procedure is an implementation of the notion of Linkhood of Vallduví (1992,1993), quite similar to Reinhart's (1981) notion of *aboutness* or Jacob's (2001) *addressation*. In all its instantiations in the literature, the notion of *aboutness* remains relatively vague, and as we saw, the tests proposed to check for aboutness give slightly different results for different examples, even in the same language. Across languages, it is even less clear that the exact same notion is operative. For example, it may be that categorical judgment is what is relevant for Japanese or Korean, as we saw in (19) and (20); clearly many examples with *–wa/-(n)un* would not be Topical according to the tests proposed for English. But English certainly would not support Topicalization across the full range of categorical utterances. I strongly suspect that Topicality is really a family of closely-related notions, rather than one notion which can be defined with a single set of necessary and sufficient conditions.

It does seem fairly certain that all languages have some means of expressing the combination of familiarity and aboutness which I take to characterize Topicality in the general case. But we must exercise caution in drawing this conclusion, in view of the evidence that Korean *–nun*—so closely related to Japanese *–wa* in other respects—does not always presuppose familiarity (or specificity).

What does all this tell us about language universals? The prevalence cross-linguistically of dedicated structural positions and/or morphological or prosodic indications of something like Topicality has led many to argue, or simply to assume that Universal Grammar (in Chomsky's sense) contains a functional category Topic, heading a Topic Phrase in the phrase structural characterization of any given language. For example, Rizzi (1997) argues for the existence of a number of phrases between the root node of a syntactic tree, the CP ("Complementizer Phrase") and the IP (the constituent in which are located tense, aspect, etc.); one of these is a Topic Phrase, headed by a functional head Topic, with the Topical constituent located in the Spec(ifier) of this functional head, and the remainder of the sentence (its **Comment**) as its complement:

(36)   TopP
             /   \
       XP   Top'
                   /   \
             Top°   YP

    Top°:   a functional head belonging to the complementizer system
    Spec(TopP) = XP: the topic
    YP: the comment

But the diversity observed in the languages briefly reviewed above argues that while this phrase structure may be appropriate to Italian, and perhaps for some other closely related Romance languages as well (though possibly not French), it is unlikely as a syntactic universal. Both structurally and in terms of function, those elements of the surveyed languages which are utilized to reflect Topicality include morphological and prosodic markers, as well as functionally distinguished syntactic positions (both with and without long distance dependencies) and scrambling, thus constituting a set of very different structural devices. Moreover, we often find multiple means utilized within a single language, even within a single discourse, as we saw in English and French.

I am reminded of a theme from Chomsky (1982:7ff,120f). Talking about notions like *passive* and *relativization*, he points out that in early work in generative grammar, there were said to be universal syntactic rules, realized in all languages, that corresponded to those notions. Some of us can remember when people debated about the character of the presumed universal Passive transformation. But Chomsky argues:

> The notions "passive," "relativization", etc., can be reconstructed as processes of a more general nature, with a functional role in grammar, but they are not "rules of grammar".
> 　　　We need not expect, in general, to find a close correlation between the functional role of such general processes and their formal properties, though there will naturally be some correlation. Languages may select from among the devices [available to them] to provide for such general processes as those that were considered to be specific rules in earlier work. At the same time, phenomena that appear to be related may prove to arise from the interaction of several components, some shared, accounting for the similarity. The full range of properties of some construction may often result from interaction of several components, its apparent complexity reducible to simple principles of separate subsystems.

Similarly, I would argue that Topic is not a structural universal that we expect to find in the *grammar* of all human languages. Instead, what we have is a loose functional universal, Topicality, so useful in human discourse that we tend to find specialized means of indicating it across a broad variety of languages. This may be the kind of thing that Jacobs (2001) has in mind in talking about *prototypical Topics*. It is useful because it helps lend coherence to discourse to talk about a single entity, often over an extended set of utterances, and to indicate when we have switched what we're talking about. But even

in languages as similar in many respects as Japanese and Korean, the realization of Topicality differs in subtle but interesting ways, depending in the brief data-set considered above on a difference in the presuppositions associated with the enclitics used *inter alia* to mark Topicality: whether they conventionally presuppose familiarity.

To resolve these questions requires careful, detailed work on a broader variety of languages, adopting a carefully defined terminology in order to facilitate comparison of results across those languages. It requires the development and refinement of tests which permit us to ferret out distinctions like those sketched so briefly above for English, so that for a given construction we have evidence of the functional role of any purported marker of Topicality. And in making claims about any essentially contextual function like this one, we must examine sentences not in isolation, but embedded in enough context to permit us to grasp the discourse dynamic in which the Topicality plays a part: The Topicality of a given constituent can only be assessed in context.

Bibliography:

Abeillé, Anne, Danièle Godard & Frédéric Sabio (2008) Two types of NP preposing in French. In Stefan Müller (ed.) *Proceedings of the 15th HPSG Conference*, CSLI Publications, Stanford, CA.

Aissen, Judith (1992) Topic and focus in Mayan. *Language* 68.1:43-80.

Arregi, Karlos (2003) Clitic Left Dislocation is Contrastive Topicalization. In *Proceedings of the 26th Annual Penn Linguistics Colloquium*, *Penn Working Papers in Linguistics* 9.1:31-44. Philadelphia: Penn Linguistics Club.

Bar-Hillel, Yehoshua. (1971) *Pragmatics of Natural Language*. Dordrecht: Reidel.

Beaver, David (2004) The optimization of discourse anaphora. *Linguistics and Philosophy* 27.1:3-56.

Beckman, Mary E. (1996) The parsing of prosody. *Language and Cognitive Processes* II (1/2): 17-67.

Beckman, Mary E. and Gayle M. Ayers (1994) *Guidelines for ToBI labeling guide, ver. 2.0*, on-line at The Ohio State University: www.ling.ohio-state.edu/~tobi/.

Brentano, Franz (1973) *Psychology from an empirical point of view*. Translated by A.C. Rancurello, D.B. Terrell, & L. McAl;ister from *Psychologie vom empirischen Standpunkt* (1874,1924). Sections V-IX.

Büring, Daniel (1997) *The Meaning of Topic and Focus: The 59th Street Bridge accent*. Routledge, London.

Büring, Daniel (2003) On D-Trees, Beans, and B-Accents. *Linguistics and Philosophy* 26(5):511-545.

Choi, H-W. (1997) Topic and focus in korean: The information partition by phrase structure and morphology. In Ho-min Sohn and John Haig (eds.), *Japanese and Korean Linguistics*, Vol. 6, 545-561. CSLI Publications, Stanford, CA.

Choi, H-W. (1999) *Optimizing structure in context: scrambling and information structure*. Stanford: CSLI.

Chomsky, Noam (1982) *Lectures on Government and Binding*, Foris, Dordrecht.

Davis, Lori J. & Prince, Ellen F. (1986) Yiddish verb-topicalization and the notion "Lexical Integrity". In A. Farley, P. PRLEY & K.-E. McCullough (eds.) *Papers from the 22nd Regional Meeting of the Chicago Linguistic Society*.

Erdmann, Peter (1990) *Discourse and Grammar: Focussing and defocusing in English*. Niemeyer Verlag, Berlin.

Fintel, Kai von (1996) *Restrictions on Quantifier Domains*. Ph.D. dissertation, UMass/Amherst.

Geluykens, R. (1992) *From Discourse Process to Grammatical Construction: On left-dislocation in English*. Benjamins, Amsterdam/Philadelphia.

Ginzburg, Jonathan (1996) Dynamics and the Semantics of Dialogue. In Jerry Seligman & Dag Westerstahl (eds) *Language, Logic and Computation*. Stanford, CA: CSLI.

Ginzburg, Jonathan (2010) *The Interactive Stance: Meaning for Conversation* (under review with Oxford University Press),

Gordon, P.C., B. J. Grosz & L.A. Gillion (1993) Pronouns, names and the centering of attention in discourse. *Cognitive Science* 17:311-348.

Grohmann, Kleanthes K. (2006) Top issues in questions: Topics – Topicalization – Topicalizability. In Lisa Lai-Shen Cheng & Norbert Corver (eds.) *Wh-movement: moving on*. MIT Press, Cambridge, Mass.

Grosz, B., Joshi, A., and Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In *Proc. ACL-83*, pages 44–50.

Grosz, B. J., Joshi, A. K., and Weinstein, S. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):202–225.

Gundel, Jeannette K. (1974) *The role of Topic and Comment in linguistic theory*. Ph.D. dissertation, University of Texas, Austin.

Gundel, Jeannette K. (1985) "Shared knowledge" and topicality. *Journal of Pragmatics* 9:83-107.

Gundel, Jeannette K. (1988) Universals of topic-comment structure. In M. Hammond, E. Moravcsik & J. Wirth (eds.) *Studies in Syntactic Typology* 209-39. Amsterdam: Benjamins.

Gundel, Jeannette K. (1998) Centering Theory and the Givenness hierarchy: Towards a synthesis. In Marilyn A. Walker, Aravind K. Joshi & Ellen F. Prince (eds.) *Centering Theory in Discourse*. Oxford University Press, 183-198.

Halliday, M.A.K. (1967) Notes on transitivity and theme in English: Parts 1 and 2. *Journal of Linguistics* 3:37-81 and 199-244.

Halliday, M.A.K. (1985) *An introduction to functional grammar*. Baltimore, Md.: Edward Arnold.

Hara, Yurie (2004) Implicature unsuspendable: Japanese contrastive *wa*. In *Proceedings of Texas Linguistics Society* vol. 8. Cascadilla Press.

Heycock, Caroline (2007) Japanese *wa*, *ga*, and Information Structure. In Shigeru Miyagawa & Mamoru Saito (eds) *Handbook of Japanese Linguistics*. Oxford University Press.

Jackendoff, Ray (1972) *Semantic Interpretation in Generative Grammar*. MIT Press. Chapter 6: "Focus and presupposition."

Jacobs, Joachim (2001) The dimensions of topic-comment. *Linguistics* 39.4:641-681.

Jäger, Gerhard (2001) Topic-Comment structure and the contrast between stage level and individual level predicates. *Journal of Semantics* 18:83-126.

Jaeger, T. Florian (2004) Topicality and Superiority in Bulgarian *wh*-questions. In O. Arnaudova, W. Browne, M.L. Rivero & D. Stojanovic (eds.) *Proceedings of FASL-12: The Ottawa Meeting 2003*, 207-228. Michigan Slavic Publications.

Joshi, A. K. and Weinstein, S. (1981). Control of inference: Role of some aspects of discourse structure–centering. In *Proc. International Joint Conference on Artificial Intelligence*, pages 435-439.

Kiss, Katalin É. (1998) Discourse-configurationality in the languages of Europe. In Siewierska (eds.) *Constituent Order in the Languages of Europe*. de Gruyter, Berlin. 681-727.

Kubota, Yusuke (2007) The Japanese topic marker *wa*. Ms., The Ohio State University.

Kubota, Yusuke & Jungmee Lee (2007) Topic marking in Japanese and Korean. Ms., The Ohio State University.

Kuno, Susumu (1973) *The Structure of the Japanese Language*. Cambridge, Massachusetts: MIT Press.

Kuno, Susumu & Ken-Ichi Takami (1993) *Grammar and Discourse Principles: Functional Syntax and GB Theory*. University of Chicago Press.

Kuroda, S.-Y. (1972) The categorical and the thetic judgment. *Foundations of Language* 9:153-185.

Kuroda, S.-Y. (1992) *Japanese Syntax and Semantics*. Dordrecht: Kluwer.

Ladusaw, William (1994) Thetic and categorical, stage and individual, weak and strong. *Proceedings of the Conference on Semantics and Linguistic Theory 4*. Ithaca, NY: Cornell University Department of Modern Languages and Linguistics, 220-229.

Lambrecht, Knud (1994) *Information Structure and Sentence Form: Topic, focus and the mental representations of discourse referents*. Cambridge University Press. See especially Chapter 4: Pragmatic relations: Topic.

Lee, Chungmin (2003) Contrastive topic and/or contrastive focus. *Japanese and Korean linguistics*. In William McClure (eds.) *Japanese/Korean Linguistics*, Vol. 12, CSLI Publications, Stanford, CA.

Lee, Chungmin (2006a) Contrastive Topic/Focus and polarity in discourse. In K. von Heusinger & K. Turner (eds.) *Where Semantics Meets Pragmatics*. CRiSPI 16, Elsevier. 381-420.

Lee, Chungmin (2006b) Contrastive (predicate) Topic, intonation, and scalar meanings. In C. Lee, M. Gordon & D. Büring (eds.) *Topic and Focus: Crosslinguistic perspectives on meaning and intonation*. Springer. 151-175.

McNally, Louise (1995) On recent formal analyses of 'topic'. In Ginzburg (ed.) *The Tbilisi Symposium on language, logic and computation*. CSLI Press, 140-162.

McNally, Louise (1998) Encoding of Information Packaging Instructions. In Peter Culicover & Louise McNally (eds.) *The Limits of Syntax*, Academic Press, 161-183.

Noda, Hisashi (1991) *Wa to ga (Wa and ga)*. Tokyo: Kurosio Publishers.

Oshima, David Y. (2008) Morphological vs. phonological contrastive topic marking. In Rodney L. Edwards, Patrick J. Midtlying, Colin L. Sprague & Kjerti G. Stensrud (eds.) *Proceedings of Chicago Linguistic Society (CLS)*, vol.41-1. Chicago: Chicago Linguistic Society. pp.371-384.

Pesetsky (1987) *Wh*-in situ: Movement and unselective binding. In E. Reiland & A. ter Meulen (eds.) *The Representation of (In)definiteness*. MIT Press, Cambridge, 98-129.

Poesio, Massimo, Rosemary Stevenson, Barbara di Eugenio & J. Hitzeman (2004) Centering: A parametric theory and its instantiations. *NLE Technical Note TN-02-01, CS Technical Report CSM-369*, available at: http://cswww.essex.ac.uk/staff/poesio/publications/CL04TN.pdf. Appears in abridged form in *Computational Linguistics* 30(3):309-363, 2004.

Portner, Paul (2002) Topicality and (non-)specificity in Mandarin. *Journal of Semantics* 19(3):275-287.

Portner, Paul & Katsuhiko Yabushita (1998) The semantics and pragmatics of topic phrases. *Linguistics and Philosophy* 21:117-157.

Portner, Paul & Katsuhiko Yabushita (2001) Specific indefinites and the information structure theory of topics. *Journal of Semantics* 18:271-297.

Prince, Ellen F. (1981) Topicalization, Focus-Movement, and Yiddish-Movement: a pragmatic differentiation. *Proceedings of the Seventh Annual Meeting*, BLS.

Prince, Ellen F. (1992) The ZPG Letter: Subjects, Definiteness and Information Status. In S. Thompson & W. Mann (eds.) *Discourse Description: Diverse analyses of a fund raising text*, John Benjamins, Philadelphia, 295-325.

Prince, Ellen (1998) On the limits of syntax, with reference to left-dislocation and topicalization. In Culicover & McNally (eds.) *The Limits of Syntax*, Academic Presss: 281-302.

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

Rizzi, Luigi (1997) The fine structure of the left periphery. In L. Haegeman (ed.) *Elements of Grammar: Handbook of generative syntax*. Kluwer, Dordrecht, 281-337.

Roberts, Craige (1996) Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In Jae Hak Yoon and Andreas Kathol (eds.) *Ohio State University Working Papers in Linguistics* Volume 49.

Roberts, Craige (1998) Focus, The Flow of Information, and Universal Grammar. In Peter Culicover & Louise McNally (eds.) *The Limits of Syntax*, Academic Press.

Roberts, Craige (2003) Uniqueness in Definite Noun Phrases. *Linguistics and Philosophy* 26:287-350.

Roberts, Craige (2004) Context in Dynamic Interpretation. In Horn & Ward (eds.) *Handbook of Pragmatics*, Blackwell, 197-220.

Mats Rooth (2005) Topic Accents on Quantifiers. In Greg Carlson and Jeffrey Pelletier (eds.) *Reference and Quantification: the Partee Effect*. Stanford: CSLI Publications.

Saaed, John (1984) *The Syntax of Focus and Topic in Somali*. Hamburg: H. Buske.

Sawada, Osamu (to appear) The Japanese contrastive wa: A mirror image of EVEN. In *Berkeley Linguistic Society*, vol. 33.

Selkirk, Elisabeth O. (1996) Sentence Prosody: Intonation, stress and phrasing. In J. Goldsmith (ed.) *The Handbook of Phonological Theory*, Blackwell, London.

Sturgeon, Anne (2006) *The Syntax and Pragmatics of Contrastive Topic in Czech*. Ph.D. dissertation, University of California at Santa Cruz.

Vallduví, Enric (1992) *The Informational Component*. Outstanding Dissertations in Linguistics Series, Garland, NY. Publication of the author's 1990 Ph.D. dissertation, University of Pennsylvania.

Vallduví, Enric (1993) *Information packaging: A survey*. Report of the Word Order, Prosody, and Information Structure Initiative, University of Edinburgh.

Walker, M. A., Joshi, A. K., and Prince, E. F. (eds) (1998) *Centering Theory in Discourse*. Clarendon Press / Oxford.

Ward, Gregory L. (1985) *The Semantics and Pragmatics of Preposing*. Ph.D. dissertation, University of Pennsylvania. Reprinted (1988) Garland Publishing Co., NY.

Ward, Gregory, & Betty J. Birner (2004) Information structure and noncanonical syntax. In L. R. Horn & G. Ward (eds.) *The Handbook of Pragmatics*. Oxford: Blackwell. 153-174.