

# Information Structure in Subordinate and Subordinate-like Clauses

Nobo Komagata

*Department of Computer Science, The College of New Jersey*

*PO Box 7718, Ewing, NJ 08628, USA*

komagata@tcnj.edu

**Abstract.** While information structure has traditionally been viewed as a single partition of information within an utterance, there are opposing views that identify multiple such partitions in an utterance. The existence of alternative proposals raises questions about the notion of information structure and also its relation to discourse structure. Exploring various linguistic aspects, this paper supports the traditional view by arguing that there is no information structure partition within a subordinate clause.

**Keywords:** Information structure, Theme/Rheme, Topic/Focus, Subordinate clause, Discourse structure, Utterance

## 1. Introduction

Traditionally, information structure (IS) has been viewed as a single partition of information within an utterance (e.g., Mathesius, 1975; more recently, Vallduví and Zacharski, 1993 and Komagata, 2001). According to this approach, even a complex utterance has only one IS partition, as can be seen below (adapted from Lambrecht, 1994).

(1) Q: Why did you hit him?

A: [I hit him]<sub>Theme</sub> [because he insulted me]<sub>Rheme</sub>.

Here, and throughout this paper, the IS labels ‘theme’ and ‘rheme’ (*T* and *R*, respectively, in later examples) are used, most closely following Steedman (2000).<sup>1</sup> Although various IS-related terms are used in different ways in certain respects, we try to limit our discussion to the essential properties of theme and rheme involving binary informational contrast between them (cf. Communicative Dynamism of Firbas, 1964).

In contrast to the traditional view, more complex IS analyses have also been proposed. For example, the following analysis has been put forward by Kruijff-Korbayová and Webber (2001a).<sup>2</sup>

<sup>1</sup> ‘Topic’ and ‘focus’ are other common labels for ‘theme’ and ‘rheme’, respectively. However, ‘topic’ and ‘focus’ are considered more overloaded with a variety of other notions.

<sup>2</sup> SMALL CAPITALS are used to indicate pitch accents.

- (2) Although [Clyde married]<sub>T</sub> [BERTHA]<sub>R</sub>, [he]<sub>T</sub> [did not inherit a PENNY]<sub>R</sub>.

Furthermore, there also is a recursive analysis of IS (Partee, 1996).<sup>3</sup>

- (3) [What convinced Susan that [our arrest]<sub>T</sub> [was caused by HARRY]<sub>R</sub>]<sub>T</sub>  
[was a rumor that [someone]<sub>T</sub> [had witnessed Harry's confession]<sub>R</sub>]<sub>R</sub>.

Another possibility that may undermine the traditional IS analysis is the presence of IS-marking syntax within subordinate clauses. The following examples from the British National Corpus (BNC) demonstrate such a case.<sup>4</sup>

- (4) a. The shape seemed to be looking through a book, although [what the book was] [Henry could not tell]. [ASS 676]  
b. His mother was always telling him that it was important for teachers to give, although [what they were supposed to give] [she did not say]. [HR 831]  
c. Although [what that could possibly be], [I have no idea], Melissa thought to herself as she put down the phone. [GVP 1963]

In these examples, the possible IS partitions are indicated by bracketing without reference to theme or rheme. The theme/rheme distinction in these types of constructions will be discussed in Section 2.

Although analysis of IS has been an active area of study for a long time, investigation of complex structures is still one of the most neglected areas.<sup>5</sup> Occasionally, we notice a remark such as, “it is quite plausible that clauses have IS.” However, the fact is that such *plausibility* has never been justified. The existence of different IS analyses is an obstacle to understanding the nature of IS, including its definition and application.

This paper supports the traditional view of IS, i.e., that there is exactly one IS partition within an utterance. In doing so, the paper defends its main hypothesis: there is no IS partition within a subordinate clause. Although it will mainly discuss *although* and *because*-clauses, the paper argues that the discussion applies to subordinate clauses in general. The arguments in this paper are partly based on Komagata (2001) and are closely related to the idea that the linguistic marking of information structure is a matrix-level phenomenon (Komagata, 1999, p. 37).

<sup>3</sup> There is another recursive approach by Hoffman (1995).

<sup>4</sup> Thanks to Bonnie Webber, p.c.

<sup>5</sup> For example, Sgall et al. (1986) and Hajičová et al. (1995) focus on simple sentence structures.

The issue at hand is also relevant to the analysis of the relation between IS and discourse structure (DS), the main theme of this special issue. To see this point, let us assume that DS is the organization of discourse units corresponding to clauses (e.g., Grosz and Sidner, 1986). Then, DS boundaries (based on clauses) and IS boundaries can be interlaced, as can be seen in the following example (taken from Kruijff-Korbayová and Webber, 2001b).

(5) [If it's Sunday] $_{T_1}$ , [we buy wine] $_{T_2}$  [over the state line] $_R$ .

This situation is referred to as “entangled IS” in Komagata (2001). Such an entanglement might appear to introduce complications to discourse analysis. With a certain nontraditional view of IS that accepts multiple IS partitions within an utterance, e.g., example (2), it would be possible to eliminate this type of entanglement. However, as we will see in Section 4, the traditional IS analysis can offer a simpler way of viewing the IS-DS connection in general. In addition, this paper also explores a few implications of the present position. The idea that every utterance has exactly one IS could strengthen the elusive notion of ‘utterance’.

Before closing this section, a few notes are in order. One might consider the question addressed in this paper as a black-and-white problem. The final verdict may well be that way. However, it is important to carefully evaluate hypotheses, especially at this stage of mixed views, so that we develop a better understanding of IS.

Next, the constructions we focus on in this paper are subordinate and subordinate-like structures. As for sentential coordinate structures, we assume that they are a sequence of utterances, each of which may contain its own IS. Furthermore, if a multiple-clause structure is considered as a coordinate structure, e.g., a nonrestrictive sentential relative clause, we may consider a separate IS for each component within the current analysis.

Another qualification is that we do not discuss a special case of IS partition within an embedded clause, such as the following.

(6) *Q*: What did you think Marcel proved?

*A*: [I thought Marcel proved] $_T$  [completeness] $_R$ .

While these types of nontraditional constituents are fairly common (Steedman, 2000), they do not appear in the subordinate structures discussed in this paper.

The rest of this paper is organized as follows. Section 2 discusses evidence for the argument that there is no IS partition within a subordinate clause. Section 3 explains the most challenging case of po-

tential counterexamples, i.e., example (4), as involving independent subordinate-like clauses. Section 4 investigates the connection between IS and DS, especially with reference to the notion of ‘utterance’.

## 2. Unavailability of IS Partition within Subordinate Clauses

In this section, we argue for the main hypothesis that there is no IS partition in a subordinate clause. We will examine the applicability of the so-called question test, as well as the roles of prosody, morphology, syntax, and semantics.

### 2.1. QUESTION TEST

One of the most common ways of identifying an IS partition is to use the question test. In fact, we have already used the method implicitly in example (1). Roughly speaking, the component in the response corresponding to the *wh*-word in the question can be considered the rheme of the response. This method is useful in many cases and is applicable to any language. However, there are several limitations. In this section, we identify potential problems with the test and clarify the applicability of this test.

First, it is always possible to respond to a question indirectly. An indirect response would immediately introduce inferences. Since the effect of inference on IS is still an uncharted area, we cannot reliably use the test for indirect responses. Second, the use of a yes-no question can be problematic. In a sense, except for the answers “yes” and “no,” the response to a yes-no question must be seen as an indirect response, which would introduce inferences. Third, the use of multiple questions is also problematic. How these questions affect the IS of later utterances is an open question. Fourth, it is obvious that the question test cannot be used to analyze IS in a text in its original form.

As a result, the only case to which we can safely apply the question test is a *direct* response to a single question. For example, the example (1) satisfies this condition. With this limitation in mind, we can still analyze certain cases of IS involving a subordinate clause. For example, the following IS pattern is possible for an utterance involving an *although*-clause.

(7) Q: I know Clyde married one of those rich women. But what happened to him after the woman died?

A: [Although Clyde married BERTHA]<sub>T1</sub>, [he]<sub>T2</sub> [did not inherit a PENNY]<sub>R</sub>.

While the main clause has an IS partition, the *although*-clause does not. Note that it is not crucial that there are two discontinuous themes above. There is a way to derive a single partition of IS from discontinuous themes/rhemes through the use of structured meaning (Komagata, 1999).

On the contrary, potentially problematic cases tend to exhibit inconsistency with linguistic marking, as we will observe in later subsections. In the following example, we attempt to analyze the IS of the response based on two questions (note that the first one is a yes-no question, which is a problem by itself).<sup>6</sup>

(8) Q: Did Clyde propose to Ada or to Bertha? And what about the dowry?

A: I don't know. But if [he proposed]<sub>T<sub>1</sub></sub> [to BERTHA]<sub>R<sub>1</sub></sub>, [he won't get]<sub>T<sub>2</sub></sub> [a PENNY]<sub>R<sub>2</sub></sub> (, so I guess he'll want to marry Ada).

Without going into the details of this particular example, we must be able to notice the limitations of the question test. Assuming an IS such as the above stretches the use of the question test beyond the applicable range.

As for examples like (2) and (3), it appears quite difficult to apply the question test to directly support the proposed IS partitions. Since the question test is incomplete, the inability to apply the question test to these examples does not immediately reject these analyses. However, this requires that the proponents of these analyses provide some other evidence.

In summary, we conclude that the question test is applicable only to a direct response to a single question. The correctness of IS analyses for the direct response to a single question can be confirmed by linguistic marking of IS such as by prosody, morphology, and syntax, as will be discussed in the following subsections.

## 2.2. PROSODY

This subsection explores the role of prosody in English as an IS marker. According to Steedman (2000), certain pitch accents signal the presence of a theme or a rheme as shown below.

(9) Q: I know which result Marcel PREDICTED. But which result did Marcel PROVE?

A: [Marcel  $\underbrace{\text{PROVED}}_{L+H^*}$ ]<sub>T</sub> [ $\underbrace{\text{COMPLETENESS}}_{H^*}$ ]<sub>R</sub>.

<sup>6</sup> Thanks to an anonymous reviewer of JoLLI.

Although Steedman (2000) discusses a broader range of pitch accents, this paper focuses on the pitch accent L+H\* as the main element of a theme marker and H\* as the main element of a rheme marker. Note that these tones do not necessarily appear in every utterance. In particular, themes are marked with L+H\* only when they are ‘contrastive’. The discussion of this section depends on the intuition of a native speaker. However, it would also be possible to evaluate analyses by using synthesized speech/intonation, e.g., by applying the approach of Prevost (1995).

Let us now turn to the examples introduced earlier. First, the most natural prosody for example (1A) would have an H\* pitch accent on *insulted*. This is consistent with the IS analysis based on the question test. As for example (7), the most natural prosody would have L+H\* on *Bertha* and H\* on *penny*. Again, it is consistent with the question test.

Next, we discuss the examples that cannot rely on the question test, i.e., examples (2), (8), and (3), in that order. In order to satisfy the IS analysis of example (2), the prosodic pattern must be as follows.

$$(10) \text{ Although [Clyde married]}_T \underbrace{[\text{BERTHA}]}_{H^*}_R, \\ [\text{he}]_T \text{ [did not inherit a } \underbrace{\text{PENNY}}_{H^*}_R.$$

According to a native speaker of English who is familiar with this type of analysis, this prosody is infelicitous in any context. That is, the above prosodic marking is incompatible with both readings of *although*: denial of expectation and concessive opposition.<sup>7</sup> This observation, coupled with the lack of question test, strongly suggests that the IS analysis in example (2) is not correct.

Next, example (8) is analogous to example (2). H\* on both *Bertha* and *penny* would yield an infelicitous prosody, regardless of the context. This point must also be seen in connection to the limitation of the question test.

Finally, in example (3), the prominent word *Harry* could be assigned L+H\* or H\*, depending on the context. Thus, it can be a part of the theme or the rheme. However, once the pitch accent signals either a theme or a rheme, the same pitch accent cannot signal the other. Thus, it would be impossible to embed a theme within a rheme or a rheme within a theme within the current analysis of pitch accent.

<sup>7</sup> As for denial of expectation, the *although*-clause delivers a proposition that would contradict the main clause. As for concessive opposition, the *although*-clause and the main clause contrastively support/reject an implied conclusion.

In summary, we were able to confirm only the IS analyses of the examples that are consistent with the question test. All other cases were either rejected or inconclusive, even with the prosodic analysis.

### 2.3. MORPHOLOGY

In Japanese as well as some other languages, morphology plays a role in marking IS. This subsection examines the use of this property for analyzing IS.

Japanese has an explicit theme marker, i.e., particle *wa* (Kuno, 1973), which is useful for analyzing IS. A caveat is that the same morpheme also serves as the contrastiveness marker. However, the contrastive case requires that there be a phonological prominence within the noun phrase that is suffixed with *wa* (summarized in Komagata, 1999). Thus, the instances of *wa* suffixed to a non-prominent noun phrase can be considered thematic.

One of the properties of the thematic *wa* observed by Japanese linguists, including Kuno (1973), is that thematic *wa* cannot occur in an embedded environment. This is in accordance with the main position of this paper. But there are also some potential counterexamples. For example, Noda (1996) discusses classifying subordinate clauses into the following categories, which include potential counterexamples.

(11)

Type	Examples	<i>ga</i>	<i>wa</i>
Strong	<i>(re)ba</i> “if”, <i>toki</i> “when”, <i>koto</i> “that” <i>node</i> (focal) “because”	yes	no
Weak	<i>node</i> “since”, <i>ga</i> “but”	yes	yes

Although Noda states that the thematic *wa* cannot be used in strong subordinate clauses, he also lists several examples that he considers exceptional. But the *wa*-marked subjects in these examples are either the subject of the main clause or contrastive. Thus, they are not exceptions at all.

In addition, Noda also discusses examples of ‘weak’ subordinate clauses that involve a thematic *wa*. One case involves the conjunctive particle *ga* “but” as the sentence connector. However, consisting of two utterances in series, this case must be classified as a sentential coordinate structure. The other case involves a particle that normally serves as a ‘subordinator’ but lacks the main clause as shown below.

- (12) *memorii-wa* ... *mottomo anzen-na basyo-dakara-da.*  
 memory-TOP      most      safe      place-because-COP

“Because the memory is the safest place (for viruses to stay).”

The subject NP suffixed with *wa* is considered as the theme of this utterance. Even though this sentence contains the subordinator-like particle, due to the lack of the main clause, it must be analyzed as an independent utterance. Then, it is no surprise that there is an IS partition within it. In this case, the subordinator-like particle *dakara* “because” must be considered a discourse connector. We will come back to these types of constructions in Section 3, especially in connection to example (4).

The above observation confirms the hypothesis that thematic *wa* cannot occur in subordinate clauses. Although this result is limited to Japanese (and other languages where the same analysis is applicable), it is also possible to make a connection to the question test in general and the prosodic analysis in English. Examples (1) and (7) would have a translation in Japanese with an appropriate morphological marking. However, examples (2), (8), and (3) cannot be translated into Japanese with the corresponding IS-marking particles, because the thematic *wa* cannot be placed on the subject within the subordinate clause, even though these analyses demand theme marking on these subjects.

To summarize, the fact that the Japanese theme marker *wa* cannot appear in a subordinate clause is consistent with our analyses involving the question test and prosody in English.

#### 2.4. SYNTAX

Syntactic IS marking is used in various languages in various ways. In this subsection, we briefly discuss the use of syntax in analyzing IS in English.

The strongest syntactic IS marker in English is topicalization/focus movement (Prince, 1984). Syntactically, these two are identical. Thus, in a written text, it is entirely contextual to identify the preposed element as a part of the theme or rheme. In a spoken corpus, prosodic information may be used to distinguish the two. In either case, the presence of topicalization/focus movement is one of a few syntactic IS markings in English. Since examples (2), (3), and (8) do not involve these constructions, we cannot positively identify an IS partition within the subordinate clauses. On the other hand, example (4) does include topicalization/focus movement within the subordinate clauses. As a result, these are stronger counterexamples to the main hypothesis of



this paper. For this reason, we will discuss these examples in detail in the next section.

There are other constructions that call for some clarification. In particular, *it*-clefts and pseudoclefts are occasionally misunderstood as IS markers. However, especially in embedded environments, clefts do not necessarily mark an IS partition. According to the findings of Prince (1978), Collins (1991), and Delin (1995), *it*-clefts serve heterogeneous functions of marking IS, contrastiveness, and referential status. Also due to Prince (1978) and Collins (1991), the free relative part of a pseudocleft is either ‘evoked’ or ‘inferrable’, which is analogous to the referential status of the definite expression. As definite expressions can appear in themes and rhemes not necessarily marking information structure, a pseudocleft cannot be a direct IS marker either.

To summarize, we note that topicalization/focus movement is the strongest syntactic IS marker in English. The presence of this construction must be analyzed carefully.

## 2.5. SEMANTICS

This subsection examines the semantic motivation behind the analysis in examples (2) and (3), in that order.

As for example (2), Kruijff-Korbayová and Webber (2001a) argue that the semantics of *although* can be explained by considering an IS partition for each clause, as in example (2). Their idea is that the conventional implicatures for *although* can be specified in terms of the alternative sets associated with the themes and the rhemes for the two readings of *although*, i.e., denial of expectation and concessive opposition. Kruijff-Korbayová and Webber adopt the framework of Steedman (2000), which is based on alternative semantics (Rooth, 1985). Their analysis does clarify the interpretation of *although*. However, we question whether the semantic effect of *although* is indeed due to IS.

More specifically, the proposed idea is that their themes and rhemes in example (2) would update the context so that each of these components can affect the context in a specific way. However, an analogous contextual update must be dealt with for contrastive case within a noun phrase as shown below.

- (13) a. The marriage of Clyde to BERTHA did not let him inherit a PENNY.
- b. The marriage of Clyde to BERTHA and his inheritance of a PENNY would not arrive at the same conclusion in terms of his happiness.



Negation: This time John's cousin didn't cause our VICTORY.  
 ≧ We won.

The proposition “we won” is a presupposition of (a) but only an ‘allegation’ of (b), because its negation cannot entail the proposition. Hajičová's argument is that when the presupposition-triggering material is in the theme (rheme), it results in a presupposition (allegation). Partee (1996) attempts to connect this analysis to Heim's analysis of presupposition along the recursive tripartite structure. According to Partee, the presupposition/allegation distinction can be observed recursively, just as in the case of tripartite structure.

However, we note that the presupposition/allegation distinction can also be observed within a noun phrase, as shown below.

- (16) a. the RECORD of our arrest ≧ We were arrested.  
       Negation: no RECORD of our arrest ≧ We were arrested.  
       b. the record of our ARREST ≧ We were arrested.  
       Negation: no record of our ARREST ≧ We were arrested.

Thus, the distinction must be analyzed independently of IS as argued in conjunction with the analysis of *although*. In fact, this distinction can be made in terms of the notions CB (contextually-bound) vs. NB (non-bound) (e.g., Sgall et al., 1986). This situation seems to correspond to the point that the analysis of Kruijff-Korbayová and Webber (2001a) could be cast within the focus-background distinction of Steedman (2000).

Based on the above discussion, we introduce the following two conjectures. First, the semantic/pragmatic effect that can be observed entirely within a noun phrase cannot be an IS effect. Second, IS only provides a bound on the domain of tripartite structure, a much weaker view of the relation between IS and tripartite structure.

We are now in a position to summarize this section. We analyzed the possibility of an IS partition within a subordinate clause, mainly by examining examples (2), (3), and (8). First, we are unable to rely on the question test for these examples, because we cannot construct a single question that would give these examples as direct responses. Second, observations made on linguistic marking, including prosody, morphology, and syntax, point to the same conclusion. None of these are in direct support of the proposed IS partitions within subordinate clauses. Finally, the semantic motivations for examples (2) and (3) are insufficient to justify the proposed IS partition either. All of these results are consistent with the present position: there is no IS partition in a subordinate clause.

### 3. Independent Subordinate-like Clauses

In this section, we turn to example (4), potentially the strongest counterexample to the paper's main hypothesis. As discussed in Subsection 2.4, the subordinate clauses in these examples contain a syntactic IS marking, i.e., topicalization/focus movement. This suggests that there is an IS partition within these subordinate clauses. However, we also noticed in Section 2.3 (for Japanese) that a subordinate-like clause can appear as an independent utterance with its own IS partition. Along this line, we will argue that example (4) actually involves independent subordinate-like clauses with their own IS partitions under a certain condition.

#### 3.1. INDEPENDENT SUBORDINATE-LIKE CLAUSES IN ENGLISH

As in the example in Japanese, i.e., example (12), subordinate(-like) clauses can appear as an independent utterance in English as well (Quirk et al., 1985).

- (17) Q: Why did he do it?  
A: Because he was angry.

The word *because* can be analyzed as a discourse connector because it connects the proposition “he did it” with the response as the reason for the proposition.

We can also find a naturally occurring example of independent *although*-clauses, e.g., in Biber et al. (1999).

- (18) A: They wouldn't shore that up because there are only about  
eight houses up there, so it doesn't pay them to <unclear>  
B: Although they own that bit

The word *although* can be considered as a concessive discourse connector. For this type of utterance, it is natural to consider an IS that is felicitous to the context. While IS marking is not necessarily clear in (written) English, we have already observed an example in Japanese with an IS partition, i.e., example (12). This observation counts as evidence for the existence of independent subordinate-like clauses.

#### 3.2. SYNTACTIC MARKING IN GERMAN

In English, the distinction between matrix and subordinate structures is not necessarily clear (e.g., Quirk et al., 1985, p. 927). Thus, it is not possible to verify that a subordinate-like clause is indeed at the matrix

level. However, the matrix-subordinate distinction is clearly marked in German, due to the Verb-second (V2) phenomenon. In the following, we observe the *obwohl*-clause (*although*-clause) in German with respect to the syntactic matrix-subordinate distinction.

While *obwohl*-clauses typically have the verb-final pattern (i.e., subordinate clause), Günthner (1996) observes the growing tendency of the V2 configuration in *obwohl*-clauses, especially, in spoken colloquial German as shown below.

- (19) A: DU ich brauch en kleinen STIFT  
       “hey I need a small pencil”  
       B: moment mal  
       “just a second”  
       A: obwohl NE eigentlich weiß ichs auch AUSwendig glaub ich  
       “although no actually I do know it by heart”

Günthner argues that these types of *obwohl*-clauses form a coordinate structure rather than a subordinate structure. We do not necessarily view this construction as a coordinate structure. However, we must analyze the *obwohl*-clause as a matrix clause. Then, this instance of *obwohl* can be considered as a discourse connector. Günthner also analyzes the condition for using *obwohl*-clauses at the matrix level as follows: the relation between the main clause and the *obwohl*-clause is very loose (or an independent illocutionary force for the *obwohl*-clause). According to Günthner, the availability of the two patterns and the condition for the *weil*-clause (*because*-clause) are analogous to the *obwohl*-clause.

In summary, we see that subordinate-like clauses can exist as independent clauses. This is confirmed by syntactic marking in German. Such independent subordinate-like clauses can also have their own IS marking as observed in the Japanese example (12).

### 3.3. WEAKLY LINKED SUBORDINATE-LIKE CLAUSES

So far, we have been looking at completely independent subordinate-like clauses. In this subsection, we investigate the possibility of analyzing some forms of subordinate structures in complex sentences as independent utterances. We will do so by examining example (4) and other related examples.

Except for example (4c), all the examples that involve IS-marking (i.e., topicalization/focus movement) within a subordinate-like clause have the *although*-clause after the main clause. We first observe that example (4c) is a special case.

With a closer look at the relation between the *although*-clause and the following clause, we can see that it is neither denial of expectation nor concessive opposition. The key to analyzing this example lies in the examination of this sentence as it appears in the original context (in a novel, *Over the Edge* by Betty Rowlands, p. 137-138).

- (20) “I’ll leave in a few minutes and be with you in about half an hour.  
Tell Madame Gebrec I will do whatever I can.”

Although what that could possibly be, I have no idea, Melissa thought to herself as she put down the phone.

Since the *although*-clause involves topicalization/focus movement, the subordinate clause must have its own IS partition. However, the clause that follows the *although*-clause is not the main clause connected to the *although*-clause. The main clause can be found in the preceding direct speech of the speaker (Melissa): “I will do whatever I can.” The *although*-clause is only realized as a thought of Melissa, which is an independent subordinate-like clause reported indirectly by Melissa. Therefore, this example turns out to be analogous to the completely independent subordinate-like clauses seen in the earlier subsections.

We now focus on example (4a, b), where *although*-clauses with an IS partition follow the main clause after a comma. The main-subordinate ordering is the only pattern we still have as a potential counterexample. According to the literature, this is a ‘marked’ pattern. For example, the majority of *although*-clauses precede the main clause (57% of all the instances involving an *although*-clause in the ACLDCI corpus from LDC). Furthermore, according to Quirk et al. (1985, p. 919), one of the semantic characterizations of subordination is that the subordinate clause presents information as if it is presupposed as given rather than asserted as new. As a result, many subordinate clauses are part of the theme. Halliday (1967) also points out that the first components in an utterance tend to be thematic. More specific analyses, also in connection to the position of adverbial phrase in general, are available (Ramsay, 1987; de Swart, 1999). These observations are consistent with our analysis of the IS partitions in example (2).

Let us now turn to the potential counterexamples involving the main-subordinate ordering. We first observe an example (Quirk et al., 1985, p. 1077) where placement of phonological prominence affects the interpretation.

- (21) a. Raven didn’t leave the party early because CAROL was there.  
b. Raven didn’t leave the party EARLY, because CAROL was there.

While the scope of the negation includes the subordinate clause in example (a), it is not the case in example (b). This suggests that there is a difference between examples (a) and (b) with respect to the strength of the connection. We can analyze the entire sentence in example (a) as a single utterance. However, the sentence in example (b) can be analyzed as a series of two utterances, analogous to the observation of Günthner (1996) in German. Then, the two utterances in (b) must have their own IS partitions. The subordinate clause in such an example would actually be a subordinate-like independent clause where the subordinator functions as a discourse connector. This is our explanation for example (4a, b).

Finally, we can also make a prediction based on the IS-related distinction between *since* (only thematic) and *because* (no IS restrictions) (Quirk et al., 1985; Lambrecht, 1994). That is, because a *since*-clause cannot have a rhematic component, it would not give rise to an independent utterance as the *because*-clause does.

To summarize this section, the presence of independent IS in the subordinate-like clauses in the potential counterexamples is not at all inconsistent with the current position, and thus these are not considered counterexamples. We also note that subordinate-like clauses with their own IS partitions only appear after the main clause or independently. The analysis of subordinate-like clauses proposed here is not specific to the *although* or *because*-clauses but applicable to subordinate (and subordinate-like) clauses headed by various subordinators (or discourse connectors).

#### 4. Utterance as Information Structure

If there is exactly one IS in an utterance as we argue for in this paper, we can also characterize ‘utterance’ in terms of IS, i.e., a pair of theme and rheme. In this section, we explore some implications of this idea.

Although we use the term ‘utterance’ quite freely, it is not so straightforward to give a complete definition. For example, Levinson (1983) defines ‘utterance’ as a sentence (or its analogue) in a context. However, fragments can be full-fledged utterances in certain contexts (e.g., Vallduví, 2001). On the other hand, Laver (1994) defines ‘utterance’ as a stretch of speech without pause. But it is difficult to identify exactly what kind of pause would signal the end of an utterance.

By no means do I propose to replace these characterizations of ‘utterance’ with the formula “utterance = IS.” However, it seems beneficial to see IS as the major property of utterance. Naturally, the alternative

views of IS could not take advantage of this position because, in their view, an utterance may have multiple or recursive ISs.

If every utterance has an IS and if every IS has a theme (we will come back to this point shortly), the utterance must have some connection to the context through the theme. Then, it would be possible to analyze thematic progression (Daneš, 1974), which in turn can be used to establish the DS (based on the thematic connection).<sup>9</sup> This situation can be represented as hypothesis (22) below, where (b) is an abbreviation for (a).

- (22) a. The DS of a discourse can be determined by the DS prior to the current (last) utterance and the IS of the current (last) utterance.  
 b.  $DS + IS = DS'$

Again, this picture will not apply to the alternative views of IS because of the more complicated relation between utterance and IS.

One consequence of the above scheme is the possibility of incremental DS development. Some discourse theories, e.g., Mann and Thompson (1988), are static in the sense that the theory provides an account of the snapshot for a complete discourse. Their theory does not address incremental processing. Grosz and Sidner (1986) are more specific about the processing mechanism associated with a discourse structure. In a sense, the idea in hypothesis (22), explicitly relating DS and IS via utterance, can be seen as a variant of the ‘linguistic structure’ of the DS proposed by Grosz and Sidner.

Let us now return to the point of whether every IS has a theme. Since fragments without a theme are possible (Vallduví, 2001), this statement is not true as it stands. On the other hand, we could analyze that an utterance is always associated with a proposition and thus a rheme-only utterance still has a semantic component that would correspond to the theme if the proposition is fully expressed. Then, we may consider that rheme-only utterances have ‘deleted’ themes. In the spirit of Erteschik-Shir (1998), we now consider the following hypothesis.

- (23) Every contextually appropriate utterance has an IS, i.e., a pair of a theme (or a deleted theme) and a rheme.

Vallduví’s work on fragments can be seen as an analysis of various types of realization (or deletion) of theme depending on appropriateness.

Let us now return to the potential counterexample (4c). The *although*-clause has its own IS signaled by the topicalization/focus movement.

<sup>9</sup> The idea is also related to the work of Webber (1991), where the right frontier has a special status.



Thus, the *although*-clause is an utterance on its own. Then, the main clause must be in the context. This is consistent with the observation made in Subsection 3.3, i.e., the relation between the *although*-clause and the subsequent clause is not that of subordinate and main clauses.

Finally, here is a preliminary answer to the question of whether there is a principled reason for the lack of an IS within the subordinate clause that precedes a main clause. We first suppose that such a pattern is possible, as in example (2), where each clause is an utterance. If we assume hypothesis (22), immediately after the subordinate clause is uttered, the DS up to (and including) the subordinate clause can be identified. This last utterance has a connection to the context through the theme. At the same time, it also has a connection to the main clause through the subordinator. Whereas there are discourse markers that cataphorically signal the structure of the forthcoming discourse segments, subordinators do not seem to have this ability. To see this point, we turn to the following parallel example from Webber and Joshi (1998).

- (24) *i.* On the one hand, John is very generous.  
*ii.* For example, suppose you needed some money.  
*iii.* You would just have to ask for it.  
*iv.* On the other hand, he's a bugger to find.
- (25) *i.* Although John is very generous,  
*ii.* giving money to whoever asks,  
*iii.* when you actually need it,  
*iv.* you'll see that he's a bugger to find.

While the two examples above may appear equivalent, we can also see some difference if we insert the long parenthetical shown below between utterances (*iii*) and (*iv*) in examples (24) and (25).

- (26) “by the way, he is the head of one of the most prestigious linguistics departments in the United States.”

My speculation is that subordinate clauses cannot support an arbitrary connection to a later discourse segment, unlike, say, *on the one hand*. Then, the use of a subordinate clause as an independent utterance would be limited to the case where the main clause (explicit or implicit) is already in the context. If the above analysis is correct, it is possible to say that a subordinate clause cannot have its own IS except when

it is used as a completely independent utterance or an independent utterance that follows the main clause with a weak link.

In summary, the idea of characterizing utterance in terms of IS seems quite natural. In all of the cases discussed above, if we take one of the alternative views of IS, the story would be substantially different.

## 5. Conclusion

In support of the traditional view of IS as a non-recursive, utterance-level partition, this paper argues that IS partitions do not occur within subordinate clauses. Equivalently, the paper also examines the contrapositive of the main point, i.e., if there is an IS division in a subordinate-like clause, it is not a subordinate clause and is actually an independent utterance. Examining aspects in prosody, morphology, syntax, semantics, and pragmatics in multiple languages, we observe evidence in support of the main point but no evidence against it.

With the traditional view of IS, we can characterize the notion of ‘utterance’ in terms of IS. This approach allows a cleaner account of the relation between IS and DS than the alternative views. It is also compatible with an incremental analysis of building DS based on the previous DS and the IS of the current utterance. Thus, the proposed idea seems to provide an informal, yet interesting, starting point for further discussion in response to the main theme of this special issue, the IS-DS connection.

Despite its usefulness in accounting for various phenomena, IS has also been a source of confusion, partly because of the lack of a precise definition. This paper attempts to eliminate a particular type of confusion associated with IS partitions in subordinate and subordinate-like clauses. For further advancement concerning IS, we will need to address other aspects as well. The author is currently working on a characterization of IS based on information theory, where the essence of IS can be reduced to comparison of the entropies of the theme and the rheme.

## Acknowledgements

I would like to thank Claire Gardent, Aravind Joshi, Geert-Jan Kruijff, Ivana Kruijff-Korbayová, Martha Palmer, Mark Steedman, Enric Vallduví, Bonnie Webber, the reviewers/participants of the ESSLLI 2001 workshop, and the reviewers/editors of JoLLI for comments and discussion.

## References

- Biber, D., S. Johansson, G. Leech, S. Conrad, and E. Finegan: 1999, *Longman Grammar of Spoken and Written English*. Longman.
- Borschev, V. and B. H. Partee: 2001, 'The Russian genitive of negation in existential sentences: The role of Theme-Rheme structure reconsidered'. In: *Travaux de Cercle Linguistique de Prague*, Vol. 4.
- Collins, P. C.: 1991, *Cleft and pseudo-cleft constructions in English*. Routledge.
- Daneš, F.: 1974, 'Functional Sentence Perspective and the Organization of the Text'. In: F. Daneš (ed.): *Papers on functional sentence perspective*. Prague: Academia and The Hague: Mouton, pp. 106–128.
- de Swart, H.: 1999, 'Position and Meaning: Time Adverbials in Context'. In: P. Bosch and R. A. van der Sandt (eds.): *Focus: Linguistic, Cognitive, and Computational Perspectives*. Cambridge University Press, pp. 336–361.
- Delin, J.: 1995, 'Presupposition and Shared Knowledge in *It*-Clefts'. *Language and Cognitive Process* **10**(2), 97–120.
- Erteschik-Shir, N.: 1998, 'The Syntax-Focus Structure Interface'. In: P. W. Culicover and L. McNally (eds.): *Syntax and Semantics, Vol. 29: The limits of syntax*. Academic Press, pp. 211–240.
- Firbas, J.: 1964, 'On Defining the Theme in Functional Sentence Analysis'. *Travaux Linguistiques de Prague* **1**, 267–280.
- Grosz, B. J. and C. L. Sidner: 1986, 'Attention, Intentions, and the Structure of Discourse'. *Computational Linguistics* **12**(3), 175–204.
- Günthner, S.: 1996, 'From Subordination to Coordination? Verb-Second Position in German Causal and Concessive Constructions'. *Pragmatics* **6**(3), 323–356.
- Hajičová, E.: 1984, 'Presupposition and allegation revisited'. *Journal of Pragmatics* **8**, 155–167.
- Hajičová, E., B. H. Partee, and P. Sgall: 1998, *Topic-Focus Articulation, Tripart Structures, and Semantic Content*. Kluwer.
- Hajičová, E., H. Skoumalová, and P. Sgall: 1995, 'An Automatic Procedure for Topic-Focus Identification'. *Computational Linguistics* **21**(1), 81–94.
- Halliday, M. A. K.: 1967, 'Notes on Transitivity and Theme in English (Part II)'. *Journal of Linguistics* **3**, 199–244.
- Heim, I.: 1982, 'The Semantics of Definite and Indefinite Noun Phrases'. Ph.D. thesis, University of Massachusetts at Amherst.
- Hoffman, B.: 1995, 'The Computational Analysis of the Syntax and Interpretation of "Free" Word Order in Turkish'. Ph.D. thesis, University of Pennsylvania.
- Komagata, N.: 2001, 'Entangled Information Structure: Analysis of Complex Sentence Structures'. In: I. Kruijff-Korbayová and M. Steedman (eds.): *Proceedings of ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*. pp. 53–66.
- Komagata, N. N.: 1999, 'A Computational Analysis of Information Structure Using Parallel Expository Texts in English and Japanese'. Ph.D. thesis, University of Pennsylvania.
- Kruijff-Korbayová, I. and B. L. Webber: 2001a, 'Concession, Implicature and Alternative Sets'. In: *Proceedings of the International Workshop on Computational Semantics (IWCS-4), Tilburg, Jan., 2001*. pp. 227–248.
- Kruijff-Korbayová, I. and B. L. Webber: 2001b, 'Information Structure and the Semantics of "otherwise"'. In: I. Kruijff-Korbayová and M. Steedman (eds.): *Proceedings of ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*. pp. 67–83.

- Kuno, S.: 1973, *The Structure of the Japanese Language*. MIT Press.
- Lambrecht, K.: 1994, *Information Structure and Sentence Form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press.
- Laver, J.: 1994, *Principles of Phonetics*. Cambridge University Press.
- Levinson, S. C.: 1983, *Pragmatics*. Cambridge University Press.
- Mann, W. and S. Thompson: 1988, 'Rhetorical structure theory: Toward a functional theory of text organization'. *Text* 8(3), 243–281.
- Mathesius, V.: 1975, *A Functional Analysis of Present Day English on a General Linguistic Basis*, edited by Josef Vachek. The Hague: Mouton.
- Noda, H.: 1996, *WA-to GA (WA and GA)*. Tokyo: Kurocio Publishers.
- Partee, B. H.: 1996, 'Allegation and local accommodation'. In: B. H. Partee and P. Sgall (eds.): *Discourse and meaning: papers in honor of Eva Hajičová*. John Benjamins, pp. 65–86.
- Prevost, S.: 1995, 'A Semantics of Contrast and Information Structure for Specifying Intonation in Spoken Language Generation'. Ph.D. thesis, University of Pennsylvania.
- Prince, E. F.: 1978, 'A Comparison of Wh-clefts and It-clefts in Discourse'. *Language* 54(4), 883–906.
- Prince, E. F.: 1984, 'Topicalization and Left-Dislocation: A Functional Analysis'. In: S. J. White and V. Teller (eds.): *Annals of New York Academy of Sciences, Vol. 433: Discourses in Reading and Linguistics*. New York: The New York Academy of Sciences, pp. 213–225.
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik: 1985, *A Comprehensive Grammar of the English Language*. Longman.
- Ramsay, V.: 1987, 'The Functional Distribution of Preposed and Postposed "if" and "when" Clauses in Written Discourse'. In: R. S. Tomlin (ed.): *Coherence and Grounding in Discourse*. John Benjamins.
- Rooth, M. E.: 1985, 'Association with Focus'. Ph.D. thesis, University of Massachusetts at Amherst.
- Sgall, P., E. Hajičová, and J. Panevova: 1986, *The meaning of the sentence in its semantic and pragmatic aspects*. D. Reidel.
- Steedman, M.: 2000, 'Information Structure and the Syntax-Phonology Interface'. *Linguistic Inquiry* 31(4), 649–689.
- Vallduví, E.: 1990, 'The informational component'. Ph.D. thesis, University of Pennsylvania.
- Vallduví, E.: 2001, 'Fragments in Information Packaging'. In: I. Kruijff-Korbayová and M. Steedman (eds.): *Proceedings of ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics*. pp. 15–16.
- Vallduví, E. and R. Zacharski: 1993, 'Accenting phenomena, association with focus, and the recursiveness of focus-ground'. In: *Proceedings of the Ninth Amsterdam Colloquium*. pp. 683–702.
- Webber, B.: 1991, 'Structure and ostension in the interpretation of discourse deixis'. *Natural Language and Cognitive Process* 6(2), 107–135.
- Webber, B. L. and A. K. Joshi: 1998, 'Anchoring a Lexicalized Tree-Adjoining Grammar for Discourse'. In: *Proceedings of the Workshop on Discourse Relations and Discourse Markers (COLING-ACL 98)*.