Identifying Assertions in Text and Discourse:
The Presentational Relative Clause Construction

Cecily Jill Duffield, Jena D. Hwang, and Laura A. Michaelis
Department of Linguistics, Institute of Cognitive Science
University of Colorado at Boulder
Boulder, CO 80309 USA
`{cecily.duffield, hwangd, laura.michaelis}@colorado.edu`

Abstract

In this paper we investigate the Presentational Relative Clause (PRC) construction. In both the linguistic and NLP literature, relative clauses have been considered to contain background information that is not directly relevant or highly useful in semantic analysis. In text summarization in particular, the information contained in the relative clauses is often removed, being viewed as non-central content to the topic or discourse. We discuss the importance of distinguishing the PRC construction from other relative clause types. We show that in the PRC, the relative clause, rather than the main clause, contains the assertion of the utterance. Based on linguistic analysis, we suggest informative features that may be used in automatic extraction of PRC constructions. We believe that identifying this construction will be useful in discriminating central information from peripheral.

1 Introduction

Identifying and extracting relevant information in a given text is an important task for human readers and natural language processing applications. To do this, proper identification and treatment of complex sentences containing relative clauses and other embedded structures such as appositive clauses (e.g., *My mother, **a patient at the center,** met him last year,*) and participial clauses (e.g., *Once he ate Werthers, **including the wrapper,***) is necessary.

Thus, the tasks of text simplification and text summarization in NLP have focused their efforts on finding effective ways of simplifying long and complex sentences into shorter and simpler ones. This has in turn proven useful in machine translation (Chandrasekar et. al., 1996), parsing and information extraction (Chandrasekar and Srinivas, 1997), as well as document simplification designed to make texts accessible to wider audiences. Such audiences include readers at low literacy levels (Siddharthan, 2003), second language learners (Petersen and Ostendorf, 2007) and aphasic readers (Devilin and Unthank, 2006).

The goal of text simplification and summarization is to reduce syntactic or structural complexities while preserving the central meaning or relevant information in the given text. Unfortunately, syntactic simplification algorithms often assume a uniform treatment of syntactic structures. This is especially true in the domain of relative clauses.

Relative clauses are often considered to contain parenthetical information. That is, their putative role in the sentence is to provide background information about the mentioned entity or entities. Consider (1): [1]

(1) You [get] <u>a guy down the street</u> **who comes up, uh, carrying a knife**.[2]

---

[1] Relative clauses are shown in boldface and the referent noun phrase is underlined. The matrix verb, the verb of the sentence in which the relative clause is embedded, is in brackets.

[2] Unless otherwise noted, all examples in this paper are taken from the Switchboard Treebank corpus (Godfrey et al. 1992, Marcus et al. 1993), a syntactically parsed version of the Switchboard corpus of American English telephone conversations.

According to the above definition of relative clauses, they key information in (1) is in the main clause *You get a guy down the street*. That is, (1) means something like: the entity *you* comes to possess *a guy down the street*. If we accordingly consider the relative clause as the background or incidental information and remove it from the semantic analysis, the assertion – the key piece of the information – would be lost.

In this paper, we discuss the Presentational Relative Clause (PRC) construction, as seen in (1). In the PRC, the relative clause, rather than the main clause, contains the assertion of the utterance (Duffield & Michaelis, 2009). Moreover, we analyze the construction in detail to assess its potential usefulness in NLP applications. Based on linguistic analysis, we suggest features that may be useful for implementation of automatic identification of PRC constructions. We believe that the identification of this construction will be useful in discriminating central information units from the peripheral ones.

2 Relative Clauses

Relative clauses are constructions in which a verbal clause modifies a nominal element, the "head," as shown in (2) and (3):

> (2) I [like] <u>cars</u> **that _____ are designed with human beings in mind**.
> (3) I [like] <u>those movies</u> **that you watch ____ time and time again**.

In (2), the relative clause *that are designed with human beings in mind* describes the head nominal *cars*, while in (3), the head nominal *those movies* is described by the relative clause *that you watch time and time again*. The gap in the relative clause indicating the position of the co-referential noun phrase is shown.

Relative clauses are typically embedded in main clauses, with the result that the nominal element satisfies a semantic-role requirement of two different verbs. For example, consider sentence (2), where the head nominal *cars* serves as an argument of the verb *like* while a gap that shares its referent with the head nominal marks the argument of the passive verb *designed*. Likewise, in (3) the referent shared by the head nominal *those movies*

in the main clause and the gap in the relative clause satisfies the requirements of two separate verbs, *like* and *watch*.

2.1 Types of Relative Clauses

In the linguistic tradition, relative clauses are typically classified into restrictive and non-restrictive types, as seen in (4-5), respectively.

> (4) And, you know, I [want] <u>a car</u> **that I can work on ____**, because I think it just costs too much even to get the oil changed anymore.
> (5) And once you've [reached] <u>the river walk area</u>, **which ____ is the tourist area**, it's usually pretty safe during the day.

In (4), the speaker has asserted that he would like a car; the restrictive relative clause specifies the type of car as one belonging to the set of cars that he could work on, as opposed to a type of car that he would be unable to repair. In (5), the relative clause does not identify the river walk area as one out of a set of areas, but simply provides additional information about it. In neither case does the relative clause assert information in the discourse; rather, it expresses a presupposed proposition (e.g., 'I can work on x,' 'x is the tourist area') that is assumed to be known by both the speaker and the addressee.

Accordingly, relative clauses are assumed to provide background information concerning the entities they modify. This background material serves either to distinguish the referent from others of its kind, as in a restrictive relative clause, or provide additional material, as in a non-restrictive relative clause, rather than asserting something new about the referent.

2.2 Relative Clauses in NLP

In line with the linguistic consensus, work in NLP has also viewed relative clauses as expressing background information about a referent. In syntactic simplification, the structural complexity is resolved by splitting a sentence into multiple ones (Siddharthan, 2003; Chandrasekar, 1996). In effect, the relative clause is pulled out of the main clause into an independent sentence. For example, (5) would be simplified into (6).

(6) And once you've reached the river walk area, it's usually pretty safe during the day. **The river walk area is the tourist area**.

In text summarization, where background information is considered parenthetical and identified with non-key content, relative clauses are simply disregarded (Siddharthan et al., 2004). That is, if we consider the relative clause to contain parenthetical information, it is reasonable to simply remove the non-key content from the text prior to any semantic analysis.

3 Presentational Relative Clauses

As the analysis of (1) above suggests, however, not all relative clauses contain parenthetical information. In fact, many linguistic studies have argued that subordinate clauses can make assertions (Goldberg, 2006; Menn, 1974; McCawley, 1981; Fox & Thompson, 1990), as a counterpoint to the studies that view them as expressing exclusively backgrounded information (Shibatani, 2009).

Here, in line with Michaelis and Lambrecht (1996), Kay and Fillmore (1999) and other work in the Construction Grammar tradition, we analyze a particular construction, the Presentational Relative Clause construction (PRC), as a productive idiomatic pattern. The PRC is a construction in which the material presented in the relative clause is not backgrounded, either in the sense of being unnecessary parenthetical material, or in the sense of being already known to both speaker and addressee (Duffield & Michaelis, 2009). In the PRC, information is asserted in the relative clause that modifies the nominal element, which is introduced by a semantically bleached main clause. Examples of the PRC include (1) and the following:

(7) They [had] <u>some guy</u> **that ___ was defending himself**.

(8) And I [know] <u>people</u> **who ___ have been drug tested and who have not, you know, been hired by a corporation**.

(9) And they've [got] <u>a fifteen year old</u> **that ___'s their boss that ___ is carrying a gun**

Each of the examples above were examined in their original contexts to determining that an uninformative main clause introduces the head nominal, while an assertion is contained within the relative clause. In (7), the main clause that introduces the referent *some guy* is semantically uninformative. By this we mean that it does NOT assert that the entity *they* possessed *some guy*. Rather, the asserted proposition in the utterance is in the relative clause, and (7) can be paraphrased as *A guy was defending himself*. In (8) the important information is not that the speaker knows a certain set of people. Instead, the key assertion here is: "Some people have been drug tested and have not been hired by a corporation." Likewise, (9), illustrating what might be described as a double-PRC, could be rephrased as, "A fifteen-year-old is their boss and is carrying a gun."

3.1 Anatomy of the PRC construction

The PRC construction is typically characterized by three main properties: a semantically empty main clause, a head nominal in the object position of the main clause that is newly introduced into the discourse, and a subject-gap relative clause that modifies the head nominal.

First property: The semantically bleached main clause serves to convey the restriction on the range of the existential quantifier rather than an assertion. In (7), for example, the main clause conveys the restriction 'x is a guy'. Consequently, a PRC, unlike a restrictive relative, is not optional. It is a required part of the clause in which it appears, exemplified by the fact that (7) cannot reasonably be construed as asserting 'They had some guy'.

Being uninformative, the main verbs of PRC tend to have low semantic weight, as in (10-12).

(10) I've [seen] <u>some statistics</u> **that ___ say it's more expensive to kill somebody than to keep them in prison for life**.

(11) You [get] <u>a guy down the street</u> **who ___ comes up, uh, carrying a knife.**

(12) When our kids were small we [had] <u>a couple of, uh, good women</u> **who ___ would often come to the house**.

The bracketed main verbs in (10-12), which otherwise denote relations of perception, obtaining, and possession, respectively, here appear simply to 'set the stage' for their object referents. In other words, (10) does not assert that the speaker sees something, (11) does not assert that the addressee

obtains something and (12) does not assert that some people possessed someone. Rather than predicating a property or action of the main clause subject, the main clause predications in (10-12) provide an explicit or inferred center of perspective from which to view the entity denoted by the head nominal (Koenig and Lambrecht, 1999).

Second property: The discourse-new head nominal is in the object position of the main clause. Thus, the PRC enables the speaker to avoid violating a hearer-based information-packaging constraint that Lambrecht (1994) refers to as the Principle of Separation of Reference and Role (PSRR): "Do not introduce a referent and talk about it in the same clause" (ibid). In other words, to aid the hearer in anchoring the new referent in discourse, the speaker introduces it in the object position of the main clause, and then predicates upon it in the relative clause, as in example (13):

(13)   *Speaker A:* We have more options now then (sic) we did when my kids were born, with being able to take off full-time longer, you can phase your schedule in so that it 's not full-time for up to six months.
       Speaker B: Oh boy, that's great.
       Speaker A: It 's really neat. I've [had] <u>a couple of assistants</u> **that ___ came back just three days a week** or they've, you know, whatever schedule they want from a pay standpoint.

Consider Speaker A's second turn, restructured as a declarative clause rather than as a PRC, and thus violating the PSRR. This time the assertion is conveyed in the main clause, but with the new entity in the in the subject position the result is pragmatically awkward:

(14)   *Speaker A:* It 's really neat. ***A couple of assistants came back just three days a week*** or they've, you know, whatever schedule they want from a pay standpoint.

An additional example is provided in (15):

(15)   *Speaker B*: I've never liked D.C. a whole lot and a really hate the Redskins. And a lot of it's because, you know, I

[got] <u>a lot of people, you know, at work with and everything</u> **that __ are big Redskin fans**. (*??A lot of people I work with and everything are big Redskin fans.*)

Thus, a crucial identifying characteristic of the PRC is that it always modifies head nominals that are main-clause objects.

Third property: Third identifying property of the PRC is the presence of a subject-gap relative clause. That is, the relative clause modifying the head nominal contains a gap in the subject position of the relative clause that is co-referential with the head nominal, as in (7) repeated here as (16):

(16)   They [had] <u>some guy</u> **that ___ was defending himself**.

There are, however, cases in which the head nominal is modified by an object-gap relative clause, which conveys an assertion, as in (17):

(17)   Everybody [gets] <u>five pounds of garbage</u> **that they can throw away____** you know uh but more than that every week uh you've got to pay by the pound.

In the example above, as with the more prototypical subject-gap PRC, the main clause does not make an assertion (in this case, the main clause does not assert that everybody receives five pounds of garbage). Rather, the assertion in the relative clause is demonstrated by the appropriate paraphrase, "Everybody can throw away five pounds of garbage." While speakers do produce object-gap sentences to convey assertions, subject-gap PRC tokens account for the majority of assertoric relative clauses in spoken discourse (Duffield and Michaelis, 2009). This results in the subject-gap structure being a useful property for identifying prototypical instances of the PRC.

4 Why identify the PRC construction?

As argued above, identifying the PRC is important because, unlike restrictive and non-restrictive relative clauses, the PRC does not present backgrounded or parenthetical information. Rather, the loss of information asserted in a PRC results in the loss of inferences crucial for the discourse.

4.1 "My son is an animal lover."

So far we have seen sentences or utterances that would inarguably be interpreted as sentences containing a PRC. However, there are PRCs, which, while equipped with every relevant PRC characteristic, initially appear to contain relative clauses expressing parenthetical information. Consider the following sentence:

(18)   I [had] <u>a son, he's now gone from the home,</u> **that ___ was an animal lover**.

In isolation, (18) could be interpreted as asserting that the speaker has a son, who now happens to have left home. The relative clause *that was an animal lover* would be treated as background information about the son. Yet an examination of the context of the conversation reveals that the relative clause contains crucial information with regard to the discourse as a whole:

(19)   *Speaker A*: Do you want to hear about my other animals I've had?
       Speaker B: Sure, sure.
       Speaker A: I've had a skunk
       Speaker B: Yeah.
       Speaker A: I've had a Burmese python, I've had rats, I've had mice.
       Speaker B: Wow.
       Speaker A: Uh, let's see, I've had gerbils, I have, I [had] <u>a son, he's now gone from the home</u>, **that ___ was an animal lover**.
       Speaker B: Uh-huh.
       Speaker A: So at one point I had a snake, skunk, dog and a cat running loose in the house.

In this case, we see even a more compelling reason to identify this sentence as a PRC. Disregarding the relative clause in (18) and treating the main clause as containing an asserted proposition, results in a radically different reading: the speaker's son is among the animals that the speaker claims to have owned (i.e. "I've had a skunk, I've had a Burmese python, I've had rats, I've had mice, I've had gerbils, I had a son…"). By classifying this sentence as a PRC, we reach the intended assertion, "My son was an animal lover," which in turn explains why the speaker has been the proud owner of a menagerie of animals.

4.2 Other examples

Much like the example in (18), the discourse context of the other PRCs presented in this paper substantiates the claim that they present information central to the discourse. The PRC in (7) "They [had] <u>some guy</u> **that was defending himself,**" used in a conversation describing a trial, signals that the situation departs from the prototypical courtroom schema in a crucial respect (the defendant is without a lawyer).

Other instances of the PRC, such as (8) "And I [know] <u>people</u> **who ___ have been drug tested and who have not, you know, been hired by a corporation**," like (18), provide explanatory information: the reason for the speaker's negative view of drug-testing. Finally, in (15) "I [got] <u>a lot of people, you know, at work with and everything</u> **that ___ are big Redskin fans**," the PRC utterance explains why the speaker dislikes a particular football team. Treating these clauses as background information, restricting categories of entities, or removing them from semantic analysis results in the loss of information about causal connections in the text.

5 Identifying the PRC construction

Thus far we have presented the identifying linguistic properties of the PRC constructions. We will now demonstrate how these properties (see Section 3.1) lend themselves to features that could be useful for automatic identification and classification of PRCs. For the purposes of this section we make the assumption that we will only retain instances that can be parsed by an automatic parser (Collins 1999, Charniak, 1997). The features we suggest are based on the results of a corpus study carried out by Duffield and Michaelis (in prep) examining the role of the PRC in the distribution of relative clause types in spoken discourse.

5.1 The distribution of the PRC in discourse

In the study by Duffield and Michaelis (2009, in prep), 1000 sentences (500 each of subject-gap and object-gap relative clauses) from the Switchboard Corpus (Godfrey, 1996) were manually examined for the first two of the identifying properties of PRC tokens as described in Section 3.1. In addition, each of the 1000 sentences was examined

within a context of 50 lines of previous discourse to determine whether or not the relative clause conveyed an assertion.

Their results showed that three properties, namely, a semantically bleached verb, discourse-new head nominals, and an assertion in the relative clause, were found to significantly predict relative clauses of the subject-gap type, suggesting that PRCs account for the prevalence of subject-gap relative clauses in discourse. In fact, 22.4% of subject-gap relative clauses were PRCs, while only 6.8% of object-gap relative clauses displayed features of the PRC. The manner in which Duffield and Michaelis manually annotated their data, although based on linguistic analysis as discussed above, easily lends itself to a list of properties that could be automatically used to identify PRCs in larger corpora.

5.2 Verb in the main clause

Corresponding to the first property (Section 3.1) of a semantically empty main clause, Duffield & Michaelis have observed that PRC tokens have the tendency to co-occur with verbs of existence, perception and discovery. Table 1 lists these verbs.

Be	Get	See	Hear	Tell
Have	Find	Know	Look	Wonder

Table 1: Main-clause verbs likely to appear in PRCs.

This suggests that encoding the lemmatized verbs as features may help in automatic classification of PRCs.

5.3 Position of head nominal

The head nominal of the relative clause was found to occur in two positions relative to the main-clause verb. It was either the second argument of the main-clause verb (20) or the complement in the prepositional phrase (21), which in turn was the second argument of the main-clause verb.

(20) They [had] {some guy **that ___ was defending himself**.}-NP

(21) I have a friend who was [telling] me {about her brother **who ___ gets high all the time**.}-PP

Duffield and Michaelis also observed that there was a correlation between the main-clause verb and the position in which the head nominal was found. That is, the verbs such as *look, tell,* and *wonder* were regularly found when the head nominal was the complement in the PP, while other verbs in Table 1 more frequently occurred with the head nominal in the direct argument position of the main-clause verb.

Furthermore, Duffield and Michaelis found that in cases where the head nominal was the complement of the PP, the head of the PP was found to be either *of* or *about.*

This suggests that in conjunction with the features derived from the lemmatized verb, the position of the head nominal in relation to the main-clause verb could be encoded as a feature. That is, for each of the sentences examined, a feature can be coded for either a NP complement or PP complement, given which type of complement the relative clause sits in.

In addition to the position of the head nominal, for those relative clauses that are found in the PP complement a feature can also be coded for the preposition heading the phrase.

5.4 Head nominal: noun and modifiers

Corresponding to the second property (Section 3.1) of PRCs, discourse-new status of modified head nominals, is indefinite form. Although Duffield & Michaelis recognize that the distinction between the 'given' and 'new' discourse statuses is not the same thing as definite versus indefinite form, discourse-active entities tend to be formally marked as definite, while discourse-new entities tend to be marked as indefinite (Prince 1992).

Head nominals considered as indefinite include bare plural nouns (e.g., *engineers*), determinerless nominals modified by adjectives or cardinal numbers (e.g., *about forty kindergarteners*), bare mass nouns (e.g., *material*), nominals with weak quantifiers (e.g., *some companies*), indefinite pronouns (e.g., *somebody, anybody*) and nominals containing the indefinite article *a* (e.g., *a fish*).

Definite head nominals include those containing the definite article *the* (e.g., *the thing, the resources*), demonstrative determiners (e.g., *this recording, that attitude*), possessive determiners (e.g., *my bass*), strong quantifiers (e.g., *every story, all these people*), demonstrative pronouns (e.g., *that, those*) and proper nouns (e.g., *Rockport, Albany*). Partitive nominal expressions with

indefinite heads (e.g., *one of those things*, *some of my friends*) are also considered as indefinite.

To turn the above into linguistic features that are characteristic of definite and indefinite head nominals, we suggest a number of possible features for classification. Several of these relating to the head noun phrase may contribute to the identification of PRCs:

Head nominal features:
- the phrasal categories of the sisters to the noun in the head noun phrase. These features will encode the presence of any adjectival or prepositional phrases within the head noun phrase. The inclusion of these features will account for the existence of any adjectival modification on the head noun phrase and/or partitive nominals.
- the existence of named entities in the head noun phrase to ascertain the existence of any proper nouns in the head nominal.

Head nominal features encoding:
- whether or not the nominal is a pronoun. This will serve to introduce the indefinite and demonstrative pronouns into the classification of PRCs.
- singularity/plurarity of the head nominal.

Modifier features encoding:
- articles and determiners,
- quantifiers, and
- possessive pronouns that modify the head noun or noun phrase.

5.5   Gap in the relative clause

The third and final property shared by PRCs concerns the gap in the relative clause. The gap occurs in subject position and is co-referential with the head nominal. This can be identified in the syntactic parses by the presence of a trace in the syntactic position co-indexed[3] with the relative pronouns *that, who,* or *which*.

---

3 Note that coindexation is distinct from coreference. In Treebank, coindexation involves the creation of a syntactic link between the trace and the constituent that was moved out of the position trace now occupies. Coreference is the relationship between the gap and the referent. Most parsers, however, do not supply co-indexation.

The syntactic position of the gap can be coded as a feature. These would also include a feature for cases where the gap is entirely missing from the relative clause. This is to account for cases of relative clauses containing a pronoun in the position where the gap should be (e.g. The gap in this example is filled with the pronoun *it*: "Here['s] a journal **that I'm in the board of it**.").

5.6   Subject position of matrix clause

In addition to the above features, based on Duffield & Michaelis' characterization of the PRC, there is one other syntactic characteristic worth investigating—the subject of the matrix clause. Consider the following PRCs:

(22)   *They* [had] some guy **that ___ was defending himself**.

(23)   *There*['s] a lot of people **that fall into that category**

(24)   *It* [was] a moving man ____ **pulled right up to her house, broke in and stole everything she owned[4].**

General observation of PRCs is that they seem to display a tendency to have either a pronoun (22), or an expletive *there* (23) or *it* (24) in the subject position of the matrix clause. This suggests that the lexical content of the subject position may be a useful predictor for PRC classification.

6 Discussion and Conclusion

In this paper, we have presented identifying properties of the PRC construction. We recognize that individual properties as presented here contribute to but do not determine the final meaning of the PRC construction as a whole, but in combination, they are likely predictors. Not all syntactic forms can be treated in the same way. By not privileging the syntactic level, but rather treating lexical, morphological, and syntactic features equally, we are able to identify key indicators that could be used to identify the function of a relative clause in discourse as conveying an assertion as opposed to backgrounded information.

---

4 Certain types of PRCs, such as in this example, are produced without the relative pronoun. Such PRCs are referred to as amalgams (see Lambrecht, 1988 for discussion)

For the purposes of NLP, we must work within the framework of phrasal structures, constrained by the resources currently available. Yet as we have suggested here, those resources, although not constructionally based, can be used to identify constructions for the purpose of extracting relevant information from naturally occurring data. We have further investigated the applicability of a construction-based approach to identifying relative clause types when the individual components, such as lexical items themselves are not themselves effective predictors. This clause-level information allows for richer representations of textual meaning.

Our future plans include experiments with implementing automatic classifiers of relative clause type based on these features. Such empirical study will give us a better understanding of the degree of usefulness of these features in identifying PRCs in text data. We anticipate that additional features will be discovered during the implementation process.

Acknowledgements

References

Chandrasekar, Raman, and Bangalore Srinivas. 1997. Automatic Induction of Rules for Text Simplification. *Knowledge-Based Systems*, 10(3): 183-190.

Chandrasekar, Raman, Christine Doran, and Bangalore Srinivas. 1996. Motivations and Methods for Text Simplification. In *Proceedings of the 16th International Conference on Computational Linguistics (COLING 1996)*, pages 1041-1044, Copenhagen, Denmark.

Charniak, Eugene. 2000. A Maximum-Entropy-Inspired Parser. In *Proceedings of NAACL00*, pages 132-139, Seattle, WA, USA.

Collins, Michael John. 1999. Head-Driven Statistical Models for Natural Language Parsing. Ph.D. thesis, University of Pennsylvania, Philadelphia, PA.

Devlin, Siobhan, and Gary Unthank. 2006. Helping Aphasic People Process Online Information. In *Proceedings of the Eighth International ACM SIGACCESS Conference on Computers and Accessibility*, pages 225-226, Portland, OR, USA.

Duffield, Cecily Jill and Laura A. Michaelis. 2009. Why Subject Relatives Prevail: Constraints versus Constructional Licensing. Presented at the *2009 Annual Meeting of the Linguistics Society of America. San Francisco*. CA, USA

Duffield, Cecily Jill and Laura A. Michaelis. in prep. Why Subject Relatives Prevail: Constraints versus Constructional Licensing.

Kay, Paul and Charles J. Fillmore. 1999. Grammatical Constructions and Linguistic Generalizations: the What's X doing Y? Construction. *Language,* 75(1):1-33.

Fox, Barbara, and Sandra Thompson. 1990. A Discourse Explanation of the Grammar of Relative Clauses in English Conversation. *Language* 66:51-64

Goldberg, Adele E. 2006. *Constructions at work. The nature of generalization in language*. Oxford University Press, Oxford.

Menn, Lise. 1974. Assertions not made by the main clause of a sentence. *Studies in the Linguistic Sciences (University of Illinois)* 4(1):132-143.

Koenig, Jean-Pierre and Knud Lambrecht. 1999. French Relative Clauses as Secondary Predicates. In Francis Corbin, Carmen Dobrovie-Sorin, and Jean-Marie Marandin, editors, *Empirical issues in Formal Syntax and Semantics 2*. Thesus, pages 191-214, The Hague, The Netherlands.

Lambrecht, Knud. 1994. Information structure and sentence form: Topic, focus, and the mental representation of discourse referents. In *Cambridge Studies in Linguistics* 71. Cambridge, Cambridge University Press.

Lambrecht, Knud. 1988. There was a Farmer had a Dog: Syntactic Amalgams Revisited. In S. Axmaker, A. Jaisser and H. Singmaster, editors, *The Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*. BLS, Inc., pages 319-339 Berkeley, CA.

McCawley, James D. 1981. The Syntax and Semantics of English Relative Clauses. *Lingua,* 53: 99-149.

Michaelis, Laura A. and Knud Lambrecht. 1996. The Exclamative Sentence Type in English. In Adele Goldberg, editor, *Conceptual Structure, Discourse and Language*. Center for the Study of Language and Information, pages 375-389, Stanford, CA.

Petersen, Sarah E. and Mari Ostendorf. 2007. Text simplification for language learners: a corpus analysis. In *SLaTE-2007*, pages 69-72, Farmington, PA.

Siddharthan, Advaith. 2003. Syntactic simplification and Text Cohesion. Ph.D. thesis, University of Cambridge, UK.

Siddharthan, Advaith, Ani Nenkova and Kathleen McKeown. 2004. Syntactic Simplification for Improving Content Selection in Multi-Document Summarization. In *Proceedings of the 20th International Conference on Computational Linguistics (COLING 2004)*, Geneva, Switzerland.