Syntactic form and discourse function in NLG

Cassandre Creswell
Department of Linguistics
University of Pennsylvania
creswell@ling.upenn.edu

Abstract

Previous research has shown that certain discourse conditions are necessary for the felicitous use of non-canonical syntactic forms like topicalization, left-dislocation, and clefts. However, the distribution of these forms does not correlate one-to-one with the presence of these conditions, and a system that generates these statistically-rare forms based only on these conditions will overgenerate. Instead, a generation algorithm must be based on additional communicative goals that can be achieved through the use of these forms. Based on a corpus study, I present three types of communicative goals that speakers achieve through the use of non-canonical syntax.

1 Introduction

1.1 Word order variation within the clause

Users of natural languages have many ways to encode the same propositional content within a single clause. In English, besides the “canonical” word order, options include topicalization, (2), left-dislocation, (3), it-clefts, (4), and wh-clefts, (5).

(1) Ed grilled the steak.

(2) The steak, Ed grilled.

(3) a. The steak, Ed grilled it.
b. Ed, he grilled the steak.

(4) a. It was Ed who grilled the steak.
b. It was the steak that Ed grilled.

(5) a. The one who grilled the steak was Ed.
b. What Ed grilled was the steak.
c. What Ed did was grill the steak.
d. What happened was Ed grilled the steak.

Corpus-based research has shown that these forms are appropriate only under certain discourse conditions (Prince, 1978); (Birner and Ward, 1998); among others. These include the membership of referents in a salient set of discourse entities (left-dislocations and topicalizations), or the salience of particular propositions within the discourse model (topicalizations and clefts).

The discourse conditions posited in the literature that allow the felicitous use of these and other non-canonical syntactic forms are necessary conditions. When they do not hold, native speakers judge the use of a non-canonical form infelicitous. They are not, however, sufficient conditions for their use (Vallduvi, 1990). Even when they hold, speakers may choose the option of not using a special form, or alternatively, the use of a special form may still be judged as odd. Moreover, although the use of these forms appears to be optional, in some contexts, substitution of a canonical sentence for the non-canonical yields a difficult-to-interpret discourse. We will illustrate each of these situations below in Section 2.

1.2 Relevance for NLG

Even when considering only a single main clause, there are usually many ways to encode propositional content. The purpose of an NLG is not simply to encode propositions as grammatical strings that occur with high frequency in a corpus, but rather to encode them as both grammatical *and* contextuallyappropriate strings from which users can derive the system's communicative intent and update their own knowledge store accordingly.

Previous NLG work on English clausal word order variation has attempted to integrate contextual information into the process of choosing a intra-clausal word order (Stone et al., 2001); (Geldof, 2000); (Klabunde and Jansche, 1998); (Humphreys, 1995). However, this work for the most part has not been grounded in corpus-based research on the discourse function of these forms. In addition, even work based on sound pragmatic research (Stone et al., 2001) cannot account for the pattern of usage described here because the model of choosing a form is too simple. Whenever the necessary conditions hold, a special form is generated. Given how rare non-canonical word order is, this model will result in overgeneration[1].

The purpose of this project is to study when human speakers generate different syntactic encodings of propositions in order to better characterize their conditions of use for utilization in an NLG system.

2 Patterns to be explained

This section first presents the previously posited factors conditioning the use of non-canonical syntax. Then, three patterns of use left unexplained by these factors will be discussed.

Topicalization and left-dislocation both involve an NP "displaced" to the left-periphery of the clause. In topicalizations this NP is coreferential with a gap/trace somewhere in the clause. In a left-dislocation, it is coreferential with a pronoun within the clause. As shown in (Birner and Ward, 1998), topicalizations are felicitous when two conditions hold: 1) the referent of the topicalized NP is a member of a salient partially-ordered set (poset) and 2) the open proposition expressed by the main clause, constructed by replacing the constituent receiving tonic stress by a variable, is salient to the hearer. The corresponding conditions for the topicalization in (6) are shown in (7).

(6) When mother was pregnant he said, "Nobody will believe it, but I hope it's a girl, **because a girl you can spoil**."(SSA)

(7) Poset P = { BOYS, GIRLS}; Open Proposition = YOU CAN DO X WITH y, SUCH THAT $y \in P$ AND X = SPOIL

Left-dislocations, in contrast, only require a single condition for felicitous usage. Here the clause-initial NP must stand in a salient poset relation with some previously evoked entity or entities in the discourse model, as illustrated in (8).

(8) I can see obvious disabilities in some individuals, **others you can't see a thing wrong with them**.(SSA)

(9) Poset P = {INDIVIDUALS EXAMINED}; OTHERS $\in P$

Both wh- and it-clefts have a two-part syntax with a *focus* constituent, post-copular in wh-clefts, and post-copular, pre-*that* clause in it-clefts, and a *presupposition*, an open proposition expressed by the complement clause.

(10) a. ...but they were vague in their minds then about what they meant by old age pensions. **Usually what they meant was a pension paid out of general revenues with some kind of an income test.**

b. Open Proposition = THEY MEANT X, X = PENSION PAID OUT OF GENERAL REVENUES

(11) a. You know that he never wanted to be President, **it was his wife that wanted him to be President.**(SSA)

b. Open Proposition = X WANTED HIM TO BE PRESIDENT, X = HIS WIFE

The discourse status of each varies by type of cleft. In a wh-cleft, the information conveyed by the presupposition must be material that (the speaker can assume) is in the hearer's consciousness at the time of utterance, either discourse-old or inferable from something else presented in the discourse (Prince, 1978). In an it-cleft, in contrast, the existential closure of the open proposition should be a belief of the speaker (Dryer, 1996).

2.1 Non-canonical entirely optional

In some cases, when the discourse conditions hold for the use of a non-canonical form, either a canonical or non-canonical is acceptable with little or no change in meaning. For example, in (12), the discourse conditions that permit topicalization hold. There is a salient open proposition, ADMINISTRATION FELT X ABOUT BALANCED BILLING LIMITS, and balanced billing is a member of a salient set, POLICIES THAT MIGHT BE ADOPTED. However, the speaker chose not to use a topicalized sentence. In

[1] Based on a tgrep search of the Penn Treebank Wall Street Journal and Switchboard corpora, these four forms appear with a frequency of about 200 and 850 tokens per million words, respectively. In the corpus used for this project, 58 transcribed oral histories from the online Social Security Administration Oral History Archives (SSA), these forms occurred with a frequency of 850 per 750,000 words.

contrast, in (13) the speaker uses a topicalized sentence even though canonical order does not seem different in this context.

(12) The AMA supported the fee schedule, opposed the expenditure targets and opposed the balanced billing limits,[...]The administration said they could live with the fee schedule if there were expenditure targets, **and they had no problems with balanced billing limits.** (SSA)

a. and balanced billing limits they had no problems with.

(13) I think we were fortunate in the kind of leadership we had, generally. **Some of them, as you know, I'm not enthused about,** but generally speaking, the quality of our leadership was quite high. (SSA)

a. I'm not enthused about some of them, as you know.

2.2 Non-canonical odd when conditions hold

In some contexts, the conditions licensing a non-canonical appear to hold, but the use of such a form would be odd. For example, (14) is an excerpt from an oral history of a soldier's experience in WWII. The implicit question the text answers is *What did the speaker do then?*. However, substituting a wh-cleft for a canonical sentence into an arbitrary point in the text is odd. In (15), the writer is replying to a message about choosing a laptop on a newsgroup about laptops. The writer can assume the salience of the poset LAPTOPS and the open proposition WRITER WOULD DO X. However, neither a topicalization nor a left-dislocation is felicitous here, as shown in (15a).

(14) And when I landed, they assigned me to a very, very bad transit camp on the other side of the river. And I couldn't stand it. It was muddy, difficult. I said "I'm not going to stay here." I walked out. I was lucky, because I was wearing bars on my shoulders, so I could get away with it. And I asked around and found out that there were a number of officers and other people sleeping at the Grand Hotel, right opposite the race course, right in the center of Calcutta. So I went over there. **And I found a bed.** And that's where I stayed in Calcutta as long as I was there. (http://fas-history.rutgers.edu/oralhistory/addison.htm)

a. ?? And what I did was found a bed.

(15) **I would recommend a Toshiba.** I just bought the 5105-S607 model and am quite pleased with it. (comp.sys.laptops, May 2, 2002)

a. ?? A Toshiba I would recommend (it).

2.3 Non-canonical form "obligatory"

In some cases, the non-canonical form is not only felicitous, but allows additional inferences. Using a canonical form instead would result in a different interpretation. For example, in (16), without the it-cleft the hearer would conclude the speaker was uncertain about whether the president was at the conference. With the it-cleft, however, the uncertainty can only be about the cause of the president's absence because the remainder of the clause is marked as presupposed. In (17), without the left-dislocation one would infer that the meeting of the second guy took place at the same time as the event in the previous clause.

(16) The conference was to take place in November. [...] We managed to bring it off in November–just when the President had his gall bladder surgery. **I think it was his gall bladder surgery that kept him from being there**, but the thing came off OK. (SSA)

a. I think his gall bladder surgery kept him from being there.

(17) "The first time was 1968, just to get out of my dad's house," she says. "**Second guy, I just met him and didn't have anything else to do.** Didn't work out...Third and fourth times were business partners. We got married for business reasons." (*Philadelphia Inquirer, p. 4-J, 7/3/88*)

a. I just met the second guy and didn't have anything else to do.

3 Choosing intraclausal word order

The previous section demonstrated that the distribution of non-canonical forms does not correlate one-to-one with the presence of the necessary conditions posited in the literature, and in some cases these optional forms play a crucial role in the meaning contributed by an utterance.

In this section we will outline a preliminary model for characterizing these choices as an augmentation of the SPUD system (Stone et al., 2001). Because SPUD explicitly connects communicative goals and the discourse context through patterns of linguistic form it is well-suited as a basis for characterizing a model of how to choose clause syntax.

By using non-canonical forms, speakers make explicit their assumptions about the discourse model, including which entities are in poset relations and which open propositions are currently salient or presupposed. Making these assumptions explicit can trigger further inferences (as shown in Section 2.3). Therefore, an algorithm for syntactic choice must incorporate goals characterized by when speakers want to trigger these inferences.

3.1 Sentence Planning Using Description

SPUD (Stone et al., 2001) is an NLG system that combines sentence planning and surface realization by choosing lexical items and their associated syntactic and semantic representations simultaneously. Any utterance generated by SPUD can be characterized by its COMMUNICATIVE INTENT, the set of inferences to which the speaker is committed to in uttering a sentence and that they expect the hearer to recover when interpreting the utterance. The source of SPUD's inferences are the conversational record, the system's beliefs, and the user's beliefs. SPUD's knowledge base keeps track of information to be asserted and the information status (discourse- and hearer-oldness) and salience of discourse entities (entities, poset relations, and open propositions). Currently, communicative intent is divided into three records:

- ASSERTIONS, the update to the conversational record that the utterance is intended to achieve
- PRESUPPOSITIONS, how particular elements in the utterance link to individuals already present in the conversational record
- PRAGMATICS, requirements on the status of individuals in the discourse

In Stone, et al. (2001), only the assertions of an utterance affect the conversational record. The choice of main clause syntactic form is related only to the pragmatics. For example, a transitive verb will be associated with multiple trees; a tree with canonical order can be chosen in any context. A tree with a topicalized order will be associated with the pragmatic requirements discussed in section 2 and will be selected if they are fulfilled Any tree, canonical or non-canonical, associated with the verb will achieve the same update to the conversational record.

3.2 Communicative goals of non-canonical syntax

Based on the current corpus study, the update of the conversational record that an utterance can achieve should be modified. In particular, trees with non-canonical syntax will be associated with not just the assertions of their canonical counterpart and some necessary pragmatic conditions but will also be associated with a richer set of potential assertions that they achieve by virtue of the fact that they can be used to fulfill some additional communicative goals. In this section I present three additional types of goals to be included in the system: attention marking, discourse relation, and focus disambiguation.

3.2.1 Attentional goals

The attentional structure of a discourse can be modeled as a stack of focus spaces that contains the individuals salient at each point in a discourse (Grosz and Sidner, 1986). Although the pragmatic constraints on the use of non-canonical forms in SPUD currently require certain entities (posets and open propositions) to be salient, in fact the use of the form is often better characterized as licensing an inference that this entity is relevant at a particular point in the discourse. Speakers can use a non-canonical form to efficiently indicate which discourse entities are currently relevant in order to have the hearers' model of the discourse match their own more closely. For example, in (18), the topicalization licenses the inference that the poset {ASPECT OF PRESS BEING DISCUSSED} is relevant here; i.e. the speaker is only making a statement about a single member of the poset (i.e. *press* = 'news stories') not any others.

(18) Q: Would you discuss your relations with the press and its attitude toward Social Security over the years?
Altmeyer: I don't know what you mean by the press. **The press, insofar as news stories are concerned, I don't think had much influence one way or another.**(SSA)

As such, uttering a topicalization Q in this case will fulfill both the goal COMMUNICATE($P \wedge$ IN-POSET(a,S)), where P is the semantics of Q, and a is the topicalized referent in poset S. As such it seems that IN-POSET(a) need not be explicitly part of the current attentional state of the conversational record, as long as it is inferable from the conversational record. In addition, SPUD will need to be altered so that the form fulfilling the most specific pragmatic requirements will not automatically be chosen unless those conditions contribute to achieving a communicative goal.[2]

[2] The oddness of (15a) can now be explained as a use of a topicalization when achieving this additional communicative goal is unnecessary. Given the context of the utterance, the membership of *a Toshiba* in the set LAPTOPS is salient and assumed; a cooperative speaker should not have a goal of communicating this information.

3.2.2 Discourse relation goals

In any text made up of more than a single utterance, the semantic relations that hold between utterances are an additional part of the meaning of the text supplementing the meaning that a single utterance contributes. These relations, referred to as *coherence, subject matter,* or *semantic relations* (Kehler, 2002; Hobbs, 1990; Halliday, 1985; Mann and Thompson, 1988), hold between two utterances and include, for example, the temporal relation holding between events or a contrast relation holding between the propositions. Because the linguistic material comprising a text, its clauses and phrases, can be combined into larger discourse segments, these relations may hold between sets of utterances. These groupings of utterances (or the intentions underlying them) are often modeled as a hierarchical tree structure (Grosz and Sidner, 1986).

Speakers can use non-canonical forms to communicate information about both coherence relations and discourse segmentation, as illustrated above in (17) and here in (19). In (17), the use of a left-dislocation changes the time interpretation of the event in the second sentence. The left-dislocation instructs the hearer that the relation between this clause and the previous is not NARRATIVE, but PARALLEL. The second clause is not a continuation of the event described by the first, but a separate event. In (19), the use of the it-cleft occurs after some intervening discussion of a separate topic marked by the hearer as an aside; it allows the speaker to mark his question as related to previous discussion because it marks the OP, YOU GOT TO MICHIGAN STATE AT TIME T, as presupposed. In a tree structure of this discourse, the cleft will correspond to an instruction to “pop” back to a higher level in the tree when attaching the utterance.

(19) G: I decided to go to academia after that and taught at Michigan State in economics and community medicine. One thing I should mention is that for my last three months in government, I had been detailed to work on the Price Commission which was a component of the Economic Stabilization program.[...]
B: **In what year was it that you got to Michigan State?**

The augmented goals fulfilled by these forms would be respectively COMMUNICATE($P \wedge$ PARALLEL(P,Q)), and COMMUNICATE($P \wedge$ ATTACH(N,P)) where N is some non-terminal node on the right frontier of the discourse tree.[3]

3.2.3 Focus disambiguation goals

In English, focus-ground structure correlates significantly with the prosodic effects of duration and amplitude (Hockey, 1998). The focus marks the part of an utterance which would correspond to the instantiation of the missing constituent in a wh-question with that utterance as the answer (Gussenhoven, 1984). In other words, focus-ground partition is relative to an implicit question being answered (Kuppevelt, 1995).

Although speakers must prosodically mark focus-ground structure on *every* utterance, this prosodic focus marking is often ambiguous. A single sentence final pitch accent may potentially correspond to multiple focus structures (Ladd, 1996). In addition, depending on its heaviness, even a single constituent may be realized with multiple pitch accents. Whether these pitch accents necessarily correspond to focus-ground partitioning is still a matter of debate (Ladd, 1996; Steedman, 2000).

In contrast, the syntactic forms here can mark focus-ground partitioning unambiguously and independently of their prosody. For example, a wh-cleft can disambiguate the focus-ground partitioning of an utterance, as in (20). Here the focused object NP can be realized with multiple prosodic phrases each with its own primary accent; its canonical counterpart would at the least be ambiguous with respect to whether the object or the entire VP were in focus (Ladd, 1996).

(20) There are those that would argue that **what we need is a quick and dirty decision at the state level based upon whatever information that was to come in the door**...(SSA)

This goal will be the most difficult to simply append to the SPUD generation system. The need to disambiguate the focus structure of an utterance is conditioned not only by a speaker’s goal COMMUNICATE($P \wedge$ FOCUS-PARTITION$(P,G(f))$), where G is the ground and f is the focus, but by the formal options and requirements the speaker has when realizing this goal

[3] Because SPUD’s grammar is a Lexicalized TAG, a discourse structure component in SPUD could be implemented by utilizing the tree structures of a Discourse LTAG (Webber et al., To appear). Part of creating a description of discourse entity would then correspond to selecting a discourse tree to adjoin to the preceding discourse structure.

prosodically. An implementation would require SPUD to choose among not only alternate syntactic trees but also prosodic realizations of those trees.

4 Conclusions and Future Work

A speaker's choice of forms is a complex piece of discourse and sentence planning. Rather than a simple function from a discourse condition to a form, it depends on the speaker's intention to communicate information about the status of entities in the discourse model, to relate the meaning of one utterance to another, and to disambiguate focus structure. The multiple intentions that a single utterance can achieve when realized with non-canonical syntax make syntactic choice a useful communicative tool.

However, there remain several problems with making it a practical tool for NLG systems:

- Any simple implementation is likely to overgenerate.
- Even given a set of goals that forms can achieve, it is not clear when a system should intend to achieve such goals.
- Multiple means of achieving these goals besides syntactic form are possible (e.g. use of discourse connectives or prosody).
- If multiple goals can be achieved with a single form (e.g. discourse segmentation and focus-ground partitioning), how will hearer's know how to update their discourse model?

In order to resolve these problems, the next step in this project will be to annotate texts with the set of above goals and apply a learning algorithm in order to determine which, if any, aspects of context and form correspond to ways—syntactic, prosodic, and lexical—of achieving these goals. Although it may seem a more general problem of AI planning to determine when speakers have particular goals, in the case of such "low level" linguistic goals, prior linguistic context may be enough to motivate these goals. A probabilistic model may be the most useful characterization of the interaction between the multiple goals and the multiple methods of achieving them in a particular context; I will test this claim through training and testing a learning algorithm.

References

B. Birner and G. Ward. 1998. *Information status and noncanonical word order in English*. John Benjamins, Amsterdam.

M. Dryer. 1996. Focus, pragmatic presupposition and activated propositions. *J. of Pragmatics*, 26:475–523.

S. Geldof. 2000. From context to sentence form. In *1st INLG Conference*, Mitzpe Ramon, Israel.

B. J. Grosz and C. L. Sidner. 1986. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

C. Gussenhoven. 1984. *On the grammar and semantics of sentence accents*. Foris, Dordrecht.

M. A. K. Halliday. 1985. *An introduction to functional grammar*. Edward Arnold Press, Baltimore.

J. R. Hobbs. 1990. *Literature and cognition*. CSLI Lecture Notes no. 21.

B. A. Hockey. 1998. *The Interpretation and Realization of Focus*. Ph.D. thesis, Univ. of Pennsylvania.

K. Humphreys. 1995. *Formalising Pragmatic Information for Natural Language Generation*. Ph.D. thesis, Univ. of Edinburgh.

A. Kehler. 2002. *Coherence, reference, and the theory of grammar*. CSLI.

R. Klabunde and M. Jansche. 1998. Abductive reasoning for syntactic realization. In *9th Int'l Workshop on NLG*, Niagara-on-the-Lake, Ontario, Canada.

J. Van Kuppevelt. 1995. Discourse structure, topicality, and questioning. *J. of Linguistics*, 31:109–147.

D. R. Ladd. 1996. *Intonational phonology*. CUP.

W. C. Mann and S. A. Thompson. 1988. Rhetorical structure theory: Towards a functional theory of text organization. *Text*, 8(3):243–281.

E. F. Prince. 1978. A comparison of wh-clefts and it-clefts in discourse. *Language*, 54(4):883–906.

M. Steedman. 2000. Information structure and the syntax-phonology interface. *Linguistic Inquiry*, 31(4).

M. Stone, C. Doran, B. Webber, T. Bleam, and M. Palmer. 2001. Communicative-intent-based microplanning: the spud system. Rutgers University.

E. Vallduvi. 1990. *The Informational Component*. Ph.D. thesis, Univ. of Pennsylvania, Philadelphia, PA.

B. L. Webber, M. Stone, A. K. Joshi, and A. Knott. To appear. Anaphora and discourse semantics. *Computational Linguistics*.