GENERATING CONNECTIVES

Michael Elhadad
Kathleen R. McKeown
Department of Computer Science
450 Computer Science Building
Columbia University
New York, N.Y. 10027
ELHADAD@CS.COLUMBIA.EDU
MCKEOWN@CS.COLUMBIA.EDU

ABSTRACT

We present an implemented procedure to select an appropriate connective to link two propositions, which is part of a large text generation system. Each connective is defined as a set of constraints between features of the propositions it connects. Our focus has been to identify pragmatic features that can be produced by a deep generator to provide a simple representation of connectives. Using these features, we can account for a variety of connective usages, and we can distinguish between similar connectives. We describe how a surface generator can produce complex sentences when given these features in input. The selection procedure is implemented as part of a large functional unification grammar.

1. INTRODUCTION: MOTIVATION

A language generation system that produces complex sentences must be able to determine which connective (e.g., *"but," "although," "since," "because," "and,"* etc.) best links its embedded sentences. Previous text generation systems (McKeown, 1985, Mann, 1984, Davey, 1979, Hovy, 1987)[1] have generally used a notion similar to rhetorical relations to describe the connection between propositions. They make a one-to-one mapping from these relations to connectives for generation (for example, the relation "opposition" would be realized by the connective "but"). In this approach it is difficult to distinguish between similar connectives (e.g., because vs. since or but vs. although). These connectives can not be used interchangeably, however, and a generation system must be able to make the correct choice.

In this paper, we describe a model for connective choice which distinguishes between similar connectives. It is based on a representation of utterances - called "interpretative format" (IF) (Elhadad & McKeown, 1988) - which captures several dimensions of their usage. We present an implemented procedure to select an appropriate connective given IFs for two propositions. We demonstrate how our surface generator uses IFs to choose between the four connectives *but*, *although*, *since* and *because*.

Each connective is described as a set of constraints between the features of the propositions it connects. This allows for a simple representation of the connective but one that captures a wide variety of different uses. An IF contains four pragmatic features in addition to the propositional content and speech act of the proposition: *argumentative orientation* (Ducrot, 1983), the set of conclusions that the proposition supports; *functional status* (Sinclair & Coulthard, 1975, Roulet *et al*, 1985), its structural relationship to the remaining discourse segment; *polyphonic features* (Ducrot, 1983), indicating whether the speaker attributes the utterance to himself or to others; and a *thematization procedure*, which describes the connection between discourse entities in the propositions. Connective selection is implemented through constraint satisfaction using a functional unification grammar.

2. PREVIOUS WORK ON CONNECTIVE DESCRIPTION

The most basic constraint on connection is often referred to as homogeneousness condition: two propositions can be conjoined if "they have something in common." Which features of the conjuncts must be homogeneous is a difficult question: (Chomsky, 1957, p.36) stated a constraint on syntactic homogeneousness (conjuncts must be "of the same type"); a purely syntactic constraint is, however, largely insufficient to satisfy the needs of a text generation system, since the decision to conjoin must be made before the syntactic structure of the conjuncts is determined. (Lakoff, 1971) proposed a semantic approach to the problem of homogeneousness: conjuncts must have a "common topic" for conjunction to be possible (p. 118). Based on this definition of homogeneousness, she distinguished between a "semantic" meaning of "but" (to express a semantic opposition) and a pragmatic usage of "but" (to deny expectations), for cases which would not satisfy the homogeneousness constraint (*e.g.*, "John is rich but dumb"). Such a distinction between a semantic

---

and a pragmatic analysis of connectors is criticized in (Abraham, 1979, p.104) (Lang, 1984, pp172ff) and (Ducrot *et al*, 1980). Lang (1984) presents a general semantics for conjunction that does not distinguish between pragmatic (or contextual) and semantic levels. Lang attributes to conjunctions an operative semantics: conjunctions' meanings are sets of ''instructions'' for ''carrying out certain mental operations'' (p. 96)[2]. The meaning of connectors is a ''program'' that controls how a ''common integrator'' can be constructed from the meaning of each conjunct. In our work, we use a similar approach for the definition of connectives, but, since we work on generation (as opposed to interpretation), we describe the meaning of connectives as sets of constraints that must be satisfied between the conjuncts as opposed to ''instructions.'' We use the notion of *thematization procedure* to account for the homogeneousness condition (cf. Section 5). In this paper, we concentrate on the distinctions between similar connectives rather than on the general properties of the class of connectives.

Work on the structure of discourse (Cohen, 1984, Reichman, 1985, Grosz & Sidner, 1986) has identified the role of connectives in marking structural shifts. This work generally relies on the notion that hearers maintain a discourse model (which is often represented using stacks). Connectives give instructions to the hearer on how to update the discourse model. For example, ''now'' (Hirschberg & Litman, 1987) can indicate that the hearer needs to push or pop the current stack of the model. When used in this manner, connectives are called ''cue (or clue) words.'' This work indicates that the role of connectives is not only to indicate a logical or conceptual relation, but also to indicate the structural organization of discourse. The distinction between cue and non-cue usages is an important one, and we also attempt to capture cue usages, but the structural indication (which often has the form of just push or pop) under-constrains the choice of a cue word - it does not control how to choose among the many markers indicating a *pop*.

Halliday (Halliday, 1985) proposes that the connection between clauses can be described on three dimensions: *taxis*, *expansion* and *projection*. This model is implemented in the *Nigel* system (Mann & Matthiessen, 1983). It provides a fine-grained classification of a broad set of connectives. However, labels used to describe the type of relation between two propositions within the *expansion* system are similar to rhetorical relations and precise definitions of these relations, to date, have tended to be subjective.

Like Halliday, we also attempt to provide a fine-grained characterization of connectives and our model has features that are similar to Halliday's *taxis* and *projection* systems. However, the use of argumentative features and a thematization procedure allows us to avoid reliance on rhetorical relations.

Our work is influenced by work in pragmatics on implicature (Levinson, 1983, Karttunen & Peters, 1979) which proposed a two-level representation of utterances (propositional content and implicatures). It is also based on a ''multi-dimensional'' description of utterances and describes connectives as devices acting on each pragmatic dimension.

3. DISTINCTION BUT VS. ALTHOUGH: FUNCTIONAL STATUS

''But'' and ''although'' can be distinguished by their influence on the discourse structure in which they are embedded. We draw upon a theory of conversation organization common in conversation analysis (Sinclair & Coulthard, 1975, Taylor and Cameron, 1987, Roulet *et al*, 1985, Moeschler, 1986) to explain this distinction. The model describes conversation as a hierarchical structure and defines three levels of constituents: *speech acts*, *move* and *exchange*. A *move* corresponds to a turn of a speaker in a conversational *exchange* between two or more speakers. It is made up of several *speech acts*. In the structure of a move, one speech act is *directive*; all others are *subordinate* - they modify or elaborate the *directive* act (Roulet *et al*, 1985). Intuitively, the *directive* act is the reason why the speaker started speaking. It constrains what can follow the move in the discourse. While a move may consist of several subordinate speech acts in addition to the directive act, the directive controls the possibilities for successive utterances. Thus, it determines what is accessible in the structure of the preceding discourse.

To see how this characterization of discourse can explain the distinction between ''but'' and ''although,'' consider the following examples:

(1) * He failed the exam,
    although he is smart. Let's hire him.
(2) He failed the exam,
    but he is smart. Let's hire him.

In both (1) and (2), the first sentence expresses a contrastive relation between two propositions. But, the full sequence (2) is coherent, whereas the sequence (1) sounds peculiar in most situations. This can be explained by the fact that in ''P but Q'' Q has directive status while in ''P although Q,'' Q has subordinate status. In (2) then, ''he is smart'' has directive status, whereas in (1) it is subordinate. Therefore, the argumentative orientation of the complex sentence as a whole in (1) is the argumentative orientation of ''he failed the exam'' and it is the argumentative orientation of ''he is smart'' in (2). The conclusion (let's hire him) is only compatible with ''he is smart.''

This distinction is similar to Halliday's taxis system (the classic subordinate/coordinate distinction) but operates at a different level. Although ''but'' is a conjunction, meaning that P and Q have the same syntactic status, P and Q have a different influence on the following discourse. We therefore require the input to the surface generator to indicate the ''point'' of a move, but to leave the syntactic status of each proposition unspecified. This more delicate decision is made by the

---

surface generator.

4. DISTINCTION BECAUSE/SINCE: POLYPHONIC FEATURES

''Because'' and ''since''[3] have the same argumentative behavior and give the same functional status to the propositions they connect. Their different usages can be explained using Ducrot's theory of *polyphony* (Ducrot, 1983). Ducrot distinguishes between the *speaker* and the *utterers*: in an utterance, some segments present beliefs held by the speaker, and others present beliefs reported by the speaker, but attributed to others - the utterers.

Using this theory, the difference between ''because'' and ''since'' is as follows: in the complex ''P since Q,'' the segments P and Q can be attributed to different utterers (''since'' is polyphonic), whereas in ''P because Q,'' they must be attributed to the same utterer (''because'' is monophonic).

Others have described ''because'' and ''since'' by noting distributional differences such as:

1. To answer a ''why'' question, only ''because'' works:

   A: Why did Peter leave?
   B: <u>Because</u> he had to catch a train.
   B: *<u>Since</u> he had to catch a train.

2. ''Because'' has a tendency to follow the main clause while ''since'' has a tendency to precede it (Quirk *et al*, 1972, 11.37).

3. ''because''-clauses can be the focus of cleft sentences (Quirk *et al*, 1972):

   It is <u>because</u> he helped you
   that I'm prepared to help him.
   *It is <u>since</u> he helped you
   that I'm prepared to help him.

The given/new distinction gives one interpretation of these differences: ''because'' introduces new information, whereas ''since'' introduces given information (where given is defined as information that the listener already knows or has accessible to him (Halliday, 1985)). Halliday also indicates that, in the unmarked case, new information is placed towards the end of the clause. And indeed ''because'' appears towards the end, the unmarked position of new information, and ''since'' towards the beginning. ''Because'' can be the focus of an It-cleft sentence which is also characteristic of new information (cf (Prince, 1978) for example). ''Because'' can answer a why-question, thus providing new information to the asker. Presenting given information in response could not serve as a direct answer.

There are many different types of given information, however (Prince, 1981). Polyphony is one type of given information but it adds an additional parameter: each piece of given information is attributed to a particular

utterer. That utterer can be one of the speakers (this is similar to indirect speech), or it can be a mutually known previous discourse. The ability to distinguish how the ''since'' clause is given (i.e., which utterer contributed it) is crucial to correct use of sentences like (3).

From a father to his child:
(3) Since you are so tired, you must sleep.

In (3), the speaker presents the hearer as the source of ''you are tired,'' and uses the fact that the hearer has previously uttered this sentence as the argument for ''you must sleep.'' If the hearer is not the source of the sentence, this strategy cannot convince him to go to sleep. Given/new in this case is therefore a polyphonic distinction, and polyphony provides an added dimension to the distinction.

In summary, ''because'' and ''since'' have the same argumentative and functional status definitions, but they have different polyphonic definitions. ''Because'' requires P and Q to have the same utterers, while ''since'' does not.

5. THEMATIZATION PROCEDURE: CUE VS. NON-CUE USAGE

As mentioned in Section 2, the most basic constraint on the use of all connectives, is that the two related propositions say something about the same ''thing'' (Lakoff, 1971, p.118). It must be possible to find a discourse entity that is mentioned in both P and Q for a connection PcQ to be acceptable. We call the set of discourse entities mentioned in an utterance the *theme* of a proposition. The constraint is that the themes of P and Q intersect. For example, in (2) ''he failed the exam but he is smart,'' the entity in common is the person referred to by ''he'' in both P and Q. In simple cases, this common entity can be found among the participants in the process described by the proposition. In many cases, however, the common entity cannot be found in the propositional contents of P and Q, and yet the connection is coherent as shown in (4), (5), and (6).

(4) Are you going to the post office?
    - because I have some letters to send
    [i.e., I ask this because ...] (Quirk *et al*, 1972, p.752)

(5) He paid for the book, because I saw him
    [i.e., I claim that because...]
    (Quirk *et al*, 1972, p.559)

(6) A: where is she?
    B: She is sick,
    since you want to know everything.
    [i.e., I talk because you insist...] (Roulet *et al*, 1985)

We explain these connections by introducing the notion of *thematization procedure*. The elements of the theme are not limited to the entities mentioned in the propositional content of a proposition. They can also be derived from other aspects of an utterance. In (4) and (5), the theme contains the speech act realized in P: ''because'' justifies the fact that the locutor asked a question or asserted knowing something, and not the fact asserted or questioned. We say that ''because''

---

[3]we consider only the causal meaning of ''since'' here

```
;; Polyphonic mention of a known principle: use since
Since turning the switch to the left causes the power to decrease,
the transmission capacity decreases.

;; Explanation by a new fact: use because
The transmission capacity decreases because you turn the switch to the left.

;; Subordinate act is an imperative - use but
Replace the battery, but keep the old battery.

;; Subordinate act can be syntactically subordinate - use although
Although you replaced the battery, keep the old battery.
```

Figure 6-1: System generated complex clauses

links on the speech act rather than on the propositional content. The *SA thematization procedure* adds the feature `Speech-Act` to the *theme* of the proposition Q. In (6), ''since'' links on the Utterance Act: the fact that B utters ''she is sick'' is justified by A's insistence on knowing everything (note that ''since'' does not justify the assertion but the fact that B is speaking at all).

It is characteristic of certain connectives to allow linking on certain features or not - that is, to allow the use of a certain thematization procedure. (4) and (5) show that ''because'' allows the use of the Speech Act thematization procedure and (6) shows that ''since'' allows the use of the utterance act procedure. We currently use the following thematization procedures in our implementation: Propositional Content, Argumentative Derivation, Functional Status, Speech Act and Utterance Act.

In a complete text generation system, the ''deep component''[4] given certain information to convey, decides when it is possible to make some of it implicit by using a certain thematization procedure. The effect is to remove certain discourse entities from the propositional content to be generated. Using a non-PC thematization procedure therefore allows to implicitly discuss certain features of an utterance that may be difficult to address explicitly. The deep module we are currently developing (Elhadad, 1990a) will use politeness constraints (Brown & Levinson, 1987) to decide which thematization is most appropriate.

CUE VS. NON-CUE USAGE: Thematization procedures allow us to distinguish cue and non-cue usages of connectives. When a connective links on a feature that is not the propositional content, it does not affect the truth conditions of the propositions, at least in the traditional view. This suggests that non-content linking is in some ways similar to the cue/non-cue distinction discussed in section 2. Our approach does therefore capture this distinction, but with several differences. It describes the structural move performed by the connective (whether it is a push or a pop, for example) using features of the

''normal'' (*i.e.*, non-cue) interpretation: if C introduces a directive act, it would work as a ''pop,'' if it introduces a subordinate act, it would be a ''push.'' Thus, a cue interpretation of a connective differs from non-cue by the thematization procedure; cue usage would be indicated by linking on the functional status, and possibly speech act or utterance act.

It remains open whether cue connectives retain all other features of non-cue usage: does a connective loose its normal meaning when used as a cue? Some researchers (Grosz & Sidner, 1986, Hirschberg & Litman, 1987) seem to argue that it does: the cue and non-cue usages are actually two distinct words. If that is the case, it would be difficult for a generator to choose among the different cue words that can perform the same structural task. On the other hand, we have no evidence at this point that cue words are not interchangeable (*e.g.*, that ''but'' is used for one kind of pop and ''now'' another).

6. IMPLEMENTATION
The procedure for selecting connectives is part of FUF, a larger surface generator using the functional unification formalism (Elhadad, 1988, Elhadad, 1990b, McKeown&Elhadad, 1990). Each connective is represented as a functional description capturing the relations between the features of the segments it connects. Functional unification is well suited for our model because constraints on each pragmatic dimension can be described separately and the formalism handles interaction between these dimensions. The generated sentences in Figure 6-1 typify the kind of sentences our system currently produces.

7. CONCLUSIONS
We have presented a model that distinguishes between similar connectives. This work synthesizes theoretical work in argumentation (Anscombre & Ducrot, 1983), conversation analysis (Sinclair & Coulthard, 1975, Roulet *et al*, 1985, Moeschler, 1985), polyphony and given/new studies (Ducrot, 1983, Halliday, 1985, Prince, 1981) into a coherent computational framework. Connective choice is implemented using functional unification.

---

[4]Generation systems are generally divided into two modules: a deep module decides what to say and a surface module decides how to say it.

REFERENCES

Abraham, Werner. (1979). BUT. *Studia Linguistica*, *XXXIII*(II), 89-119.

Anscombre, J.C. & Ducrot, O. (1983). *Philosophie et langage. L'argumentation dans la langue*. Bruxelles: Pierre Mardaga.

Brown, P. & Levinson, S.C. (1987). *Studies in Interactional Sociolinguistics 4. Politeness : Some universals in language usage*. Cambridge: Cambridge University Press.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Cohen, R. (July 1984). A computational theory of the function of clue words in argument understanding. *Coling84*. Stanford, California: COLING, 251-258.

Davey, A. (1979). *Discourse Production*. Edinburgh: Edinburgh University Press.

Ducrot, O. (1983). *Le sens commun. Le dire et le dit*. Paris: Les editions de Minuit.

Ducrot, O. *et al*. (1980). *Le sens commun. Les mots du discours*. Paris: Les editions de Minuit.

Elhadad, M. (1988). The FUF Functional Unifier: User's manual. Technical Report CUCS-408-88, Columbia University.

Michael Elhadad. (1990). Constraint-based Text Generation: Using local constraints and argumentation to generate a turn in conversation. Technical Report CUCS-003-90, Columbia University.

Elhadad, M. (1990). Types in Functional Unification Grammars. *Proceedings of ACL'90*. Pittsburgh.

Elhadad, M. and McKeown, K.R. (1988). What do you need to produce a 'but'. Technical Report CUCS-334-88, Columbia University.

Grosz, B. and Sidner, C. (1986). Attentions, intentions, and the structure of discourse. *Computational Linguistics*, *12*(3), 175-204.

Halliday, M.A.K. (1985). *An Introduction to Functional Grammar*. London: Edward Arnold.

Hirschberg, J. and Litman, D. (1987). Now let's talk about Now: identifying clue phrases intonationally. *Proceedings of the 25th Conference of the ACL*. Association for Computational Linguistics, 163-171.

Eduard Hovy. (1987). *Generating natural language under pragmatic constraints*. Doctoral dissertation, Yale University.

Karttunen, L. & S. Peters. (1979). Conventional Implicature. In Oh & Dinneen (Ed.), *Syntax and Semantics*. Vol. 11: *Presupposition*. New York: Academic Press.

Lakoff, R. (1971). IFs, ANDs and BUTs: about conjunction. In Fillmore & Langendoen (Ed.), *Studies in Linguistic Semantics*. New York: Holt, Rinehart & Winston.

Lang, Ewald. (1984). *SLCS*. Vol. 9: *The Semantics of Coordination*. Amsterdam: John Benjamins B.V. Original edition: *Semantic der koordinativen Verknupfung*, Berlin, 1977.

Levinson, S.C. (1983). *Pragmatics*. Cambridge, England: Cambridge University Press.

Mann, W.C. (1984). Discourse Structure for Text Generation. Technical Report ISI/RR-84-127, Information Sciences Institute.

Mann, W.C. and Matthiessen, C. (1983). Nigel: a Systemic Grammar for Text Generation. Technical Report ISI/RR-83-105, USC/ISI.

McKeown, K.R. (1985). *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge, England: Cambridge University Press.

McKeown, K. and M. Elhadad. (1990). A Contrastive Evaluation of Functional Unification Grammar for Surface Language Generators: A Case Study in Choice of Connectives. In Cecile L. Paris, William R. Swartout and William C. Mann (Eds.), *Natural Language Generation in Artificial Intelligence and Computational Linguistics*. Kluwer Academic Publishers. (to appear).

Moeschler, J. (1985). *LAL. Argumentation et Conversation: Elements pour une analyse pragmatique du discours*. Paris: Hatier-Credif.

Moeschler, J. (1986). Connecteurs pragmatiques, lois de discours et strategies interpretatives: parce que et la justification enonciative. *Cahiers de Linguistique Francaise*, (7), pp. pages 149-168.

Prince, E.F. (December 1978). A Comparison of Wh-Clefts and It-clefts in Discourse. *Language*, *54*(4), 883-906.

Prince, E.F. (1981). Toward a Taxonomy of Given-New Information. In Cole, P. (Ed.), *Radical Pragmatics*. New York: Academic Press.

Quirk, R. *et al*. (1972). *A Grammar of Contemporary English*. Longman.

Reichman, R. (1985). *Getting computers to talk like you and me: discourse context, focus and semantics (an ATN model)*. Cambridge, Ma: MIT press.

Roulet, E. *et al*. (1985). *L'articulation du discours en francais contemporain*. Lang: Berne.

Sinclair and Coulthard. (1975). *Towards an Analysis of Discourse*. Oxford, England: Oxford University Press.

Taylor, T.J. and Cameron, D. (1987). *Language &*

Communication Library. Vol. 9: *Analysing Conversation: Rules and Units in the Structure of* *Talk*. Oxford: Pergamon Press.