

FOCUS CONSTRAINTS ON LANGUAGE GENERATION

Kathleen R. McKeown

Department of Computer Science
Columbia University
New York, NY 10027

This paper also appears in *Proceedings of the Eight International Joint Conference on Artificial Intelligence*, Karlsruhe, Germany, 1983

Focus Constraints on Language Generation¹

Kathleen R. McKeown
Department of Computer Science
Columbia University
New York, N. Y. 10027

ABSTRACT Computer generation of natural language requires the ability to make reasoned choices from a large number of possible things to say as well as from a large number of expressive possibilities. This paper examines in detail how one influence on a generated text, focus of attention, can be used to constrain the many possibilities that a generation system must consider. A computational treatment of focus of attention is presented that can be used to constrain what the system needs to consider when deciding what to say next. In this process, information is produced that provides constraints on which words and syntactic structures best express the system's intent, thus ensuring that its resulting text is coherent. This analysis has been used in the fully implemented TEXT system which generates paragraph length responses to questions about database structure.

1 Introduction

Computer generation of natural language requires the ability to make reasoned choices from a large number of possibilities and from a variety of knowledge sources. A system that communicates with its users must be able to decide *what* information to communicate, *when* to say what, and *which* words and syntactic structures among many possibilities best express its intent. Previous papers (e.g., [MCKEOWN 80], [MCKEOWN 82A], [MCKEOWN 82B]) illustrate how the final text is influenced by a variety of factors, including commonly used rhetorical strategies, semantic information, focus of attention, and the discourse goal. This paper examines in detail how one of those sources, focus of attention, can be used

¹This work was partially supported by National Science Foundation grant #MCS81-07290, awarded to the Computer and Information Science Department of the University of Pennsylvania and by ONR contract N00014-82-K-0156, awarded to the Department of Computer Science, Columbia University.

to constrain the many possibilities that a generation system must consider. A computational treatment of focus of attention is presented that constrains what information to communicate and its order, and in the process produces information that constrains which words and syntactic structures best express its intent. Examples are given of how this analysis has been used in the fully implemented TEXT system, which generates paragraph length responses to questions about database structure.

Focus constraints are only part of the TEXT system's mechanism for responding to a question. TEXT also uses discourse plans and a mechanism for determining relevancy [MCKEOWN 82A]. To answer a question, TEXT first circumscribes a subset of the knowledge base containing information relevant to the given question. A discourse plan (called a schema [MCKEOWN 82A]) is then used to guide the construction of an answer. The focusing mechanism aids in this process by constraining the selection of information to talk about next to that which ties in most appropriately with the previous discourse. Thus, focus information doesn't primarily determine the content of the response, but provides constraints on the many possibilities that must be considered and aids in shaping a coherent response. TEXT was implemented using an ONR database containing information about military vehicles and weapons. Examples are taken from this domain. In the following sections, various choices that a generation system must make at different phases of processing are first described. How focus information can be used to influence these decisions is then discussed.

2 Choices

One of the first steps in speaking or writing is the narrowing of attention to knowledge relevant to the purpose at hand. Speakers and writers are capable of ignoring information in their large body of knowledge about the world which is not specific to the current discourse purpose. This process, called *global focusing* [GROSZ 77] is modeled in TEXT by restricting the information that needs to be considered when constructing an answer to a subset of the knowledge base which contains information that could potentially be included as part of the answer. Although this process is not discussed further here, the fact that it does occur is critical for the success of later processes (see [MCKEOWN 82A] for further details)

Once a system has determined what information is likely to be relevant to its current discourse goal, it also must be able to determine what to say first, what next, and how to close the discourse. Order of information can be crucial to a reader's understanding of a text. Textual sequence alone can cause a reader to draw inferences about the relation between two propositions, including temporal sequence, causality, and exemplification, among others. While textual sequence need not always correspond to, for example, temporal sequence, the absence of textual connectives specifying otherwise (e.g., "when", "after", "while") may indicate that it does. It is important, therefore, that careful attention be given to how propositions are ordered.

At the surface level, a generator must be able to make reasoned decisions about the best lexical items to use, when to use pronominal reference, and about the syntactic construction that should be used. Examples illustrating these choices are shown in 1-3 below:

- 1 Lexical choice (bought vs sold)
 - A) Jane bought \$3 00 worth of bobby socks from Michael
 - B) Michael sold \$3 00 worth of bobby socks to Jane

- 2 Pronominal reference (Linda vs she)
 - A) Linda flew to Washington.
 - B) She flew to Washington.

- 3 Syntactic structure (active vs passive)
 - A) John gave the book to Mary.
 - B) Mary was given the book by John.

Textual order and surface choice are both influenced by a speaker's focus of attention. In the next two sections this influence is characterized in such a way that it can be used by a language generator to resolve decisions in textual order and surface choice.

3 Immediate focus and generation

On producing a single utterance (controlled by a schema), TEXT narrows its focus of attention to a single object (or set of objects) in its pool of relevant information. Having made a decision about what to talk about first, it must support that decision in succeeding utterances if it wants its text to be easily understood. That is, having decided to focus on a particular object(s), its utterances constrain the set of possibilities for what can be said next if the system is to avoid jumping around from one topic to another. These are termed *immediate focus* constraints since they apply locally between utterances.

TEXT uses constraints developed by Sidner [SIDNER 79] on how focus of attention can shift or be maintained from one sentence to the next. Sidner showed that a speaker can either maintain his/her current focus, shift to focus on an item just introduced, return to a previous focus, or focus on an item implicitly related to the current focus. The TEXT system uses these constraints to limit the number of possibilities it must consider when deciding what to say next. If its discourse plan allows for several next utterances, the system only considers propositions that have an element that can be focused in one of these ways.

While Sidner's constraints are sufficient for interpreting natural language, for generation a speaker may have to decide which of the constraints is better than any other at any point. An ordering on Sidner's constraints was developed for generation which dictates which of these is preferable (see Figure 1 below). The preference ordering suggests that a speaker should shift to focus on an item just introduced into conversation if s/he has something to say about it. If the speaker chooses not to do so, that item will have to be re-introduced into conversation at a later point before the additional information can be conveyed. If, on the other hand, the speaker does shift to the item just mentioned, there will be no trouble in continuing with the old conversation. In that case, the speaker is returning to a topic of previous discussion, a legal focus move.

Several consecutive moves to items just introduced are not a problem. In fact, consecutive focus shifts over a sequence of sentences occurs frequently in written text. If this rule were applied indefinitely though, it would result in never-

ending side-tracking onto different topics of conversation. However, the model of generation assumes that information is being presented in order to achieve a particular goal (e.g., answer a question). Only a limited amount of information is within the speaker's scope of attention because of its relevance to that goal (as defined by global focus). Hence only a limited amount of side-tracking can occur.

The second preference indicates that a speaker should continue talking about the same thing rather than returning to an earlier topic of conversation where possible. By returning to a previous discussion, a speaker closes the current topic. Therefore, having introduced a topic (which may entail the introduction of other topics) one should say all that needs to be said before returning to an earlier topic. The second preference guarantees that a speaker will avoid implying that s/he is finished talking about the current subject when in fact there is more to be said. If neither of the first two preferences apply then the speaker must return to an earlier topic of discussion (preference 3).

In cases where a speaker must choose between two propositions with the same focus, the preferences described so far proscribe no course of action. Rather than making an arbitrary choice, a speaker tends to group together in discourse those properties that are in some way related to each other. When the system has a choice between two propositions with the same focus, it chooses that proposition with the most mentions to previously mentioned items (preference 4)

This ordering doesn't dictate *absolute* constraints on the system. Just as a speaker may choose to suddenly switch topics, the system may choose to do so also. The ordered focus constraints are preferences which indicate the system's best move when faced with a choice. If the system's discourse plan indicates that no next choice meets these constraints, it will follow its plan making note of the abrupt switch in focus. This switch can then be syntactically marked to ease the transition for the user.

1. shift focus to item mentioned in previous proposition
2. maintain focus
3. return to topic of previous discussion
4. select proposition with greatest number of implicit links to previous proposition

FIGURE 1. Ordered Focus Constraints

4 Choosing Surface expressions

There are many different ways in which a proposition can be expressed in English. If the system makes an arbitrary decision about which to select in a given situation, an inappropriate decision could easily be made. For example, if the propositions shown in 1-3 above are to be expressed as parts of discourse sequences, then one of the choices in each pair is clearly inappropriate (4-6 below).

4. Jane was in a hurry to finish her shopping. It was a chore she particularly despised. First,
 Jane bought \$3.00 worth of bobby socks from Michael.
 *Michael sold \$3.00 worth of bobby socks to Jane.
5. We knew that Mary took the train to New York with Linda, but didn't realize that
 Linda flew to Washington from there.
 *she flew to Washington from there.
6. John bought that great new book on data structures. He read the first three chapters and then
 he gave the book to Mary.
 *Mary was given the book by John.

In these discourse sequences, the inappropriateness of the starred choice in each case can be explained by the speaker's focus over the discourse. A number of linguists (e.g., [HALLIDAY 67], [FIRBAS 66]) have discussed how thematic (or focus) information can affect the ordering of sentence constituents, suggesting that new (or unfocused) information usually occurs towards the end of a sentence. In order to place this information in its proper position in the sentence, structures other than the unmarked active sentence may be required (for example, the passive). Structures such as it-extraposition, there-insertion, topicalization, and left-dislocation can be used to introduce new information into discourse. Pronominalization is another linguistic device associated with focused information [SIDNER 79]; it is often used to signal the speaker's focus of attention.

In the example discourse sequences, focus accounts for the choices made in the following ways. In sequence (4), the focus is on Jane who is doing the action. The verb "bought" is selected over "sell" in order to allow Jane to be described as the focused participant in the action. In (5), Mary is the focused element of the factive "to know". If the second reference to Linda is pronominalized, we are likely to interpret the referent of the pronoun as the focused element, or Mary. In (6) the first sentence focuses on John. The active sentence is more appropriate since it places the focused information in surface subject position.

5 Surface Choice in the TEXT System

Since focus information has been used to constrain the selection of propositions in the TEXT system, a record containing each proposition's focus and its potential focus list is available for the system to use when determining the specific syntactic structures that should be used in the answer. It can examine this information to determine how a proposition is related to previous discourse and therefore which of various surface-level choices should be made.

In TEXT, focus information is used in some limited situations to test whether pronominalization can be used. Part of the answer to the question "What is a ship?" where pronominalization was selected is shown in (7) below. In the first sentence of the answer, the ship is being focused on and reference to it in the following sentence can therefore be pronominalized.

7) (definition SHIP)

What is a ship?

A *ship* is a water-going vehicle that travels on the surface. *Its* surface-going capabilities are provided by the DB attributes DISPLACEMENT and DRAFT

Focus information is also used to discriminate between use of the passive and active construction. The passive can be used to place the focused constituent in surface subject position when the logical subject of the sentence is not in focus. In TEXT, this construction may be necessary in describing binary relations in the ONR database. In the ONR database, weapons are associated through the relation "carry" with different vehicles. When answering a question about missiles, a weapon, the passive construction is used since the missile is in focus (see (8) below). When answering a question about the ECHO II, a type of submarine, the active construction is used in order to attribute information to the "ECHO II" (see (9) below)

8) (difference MISSILE TORPEDO)

What is the difference between a missile and a torpedo?

The missile's target location is indicated by the DB attribute DESCRIPTION and its flight capabilities are provided by the DB attribute ALTITUDE. Other DB attributes of the missile include PROBABILITY_OF_KILL, SPEED, ALTITUDE, LETHAL_RADIUS_&_UNITS and TIME_TO_TARGET_&_UNITS. *Missiles are carried by water-going vehicles and aircraft.*

9) (information ECHO-II-SUBMARINE)

What do you know about Echo II's?

There are no sub-classes of echo II in the ONR database. *Echo IIs carry 16 torpedoes, between 16 and 99 missiles and 0 guns.*

The use of there-insertion by the TEXT system is shown below in (10) in a portion of the answer generated to the question "What is a guided projectile?". Use of there-insertion in this situation is one way to introduce the set of sub-classes of the guided projectile as focus into the discourse.

10) (definition GUIDED)

What is a guided projectile?

There are 2 types of guided projectiles in the ONR database: torpedoes and missiles. The missile has a target location in the air or on the earth's surface. The torpedo has an underwater target location.

6 Focus Constraint Implementation

TEXT constructs an answer to a given question by repeatedly selecting² propositions to include in the answer, each of which is eventually translated to a sentence. A proposition consists of a predicate and its arguments which are realized in the final English version as the verb of the sentence and its case roles. TEXT uses its schema, or discourse plan, as a guide in determining what to include next and is constrained in its choice in the two ways discussed so far: 1) it is constrained in what it considers for inclusion by global focus and 2) the text constructed so far constrains what it can say next. As a proposition is added to the answer, a focus record is filled out and added along with it. The focus record

²Actually by constructing the proposition since propositions don't exist as a whole in the knowledge base

associated with each proposition includes the focused argument and the potential focus list (other arguments which are candidates for a shift in focus). This record is used both to constrain what proposition can follow as well as for the basis of surface choice.

Immediate focus constraints (the preference ordering shown in Section 3.0) determine how the focus record is filled out. If any possible next proposition (of those that have been selected by the schema) has an argument that can be focused such that the first preference is met (i.e., the argument was a member of the previous proposition's potential focus list), that proposition is selected and its argument is recorded as the current focus of the proposition. All other arguments are included in the proposition's potential focus list as they are candidates for a shift in focus. If the first preference cannot be met, the same procedure is repeated for each of the remaining preferences until a proposition is selected.

Surface choice is made on the basis of the focus record associated with the proposition. It is used to select the sentence voice (active, passive, or there-insertion) and to determine whether pronominalization can be used. On selection of a verb for the sentence, sentence voice is also selected. When a verb is selected to translate a predicate, the predicate's arguments are mapped onto the case roles of the verb (e.g. protagonist, goal). If the protagonist is the focused argument, the active voice is selected, if the goal is focused, the passive voice is selected. The selection of voice there-insertion is slightly more complicated as it is based on the introduction of a set whose members are focused in succeeding utterances (see MCKEOWN 82A] for more details).

As an example, consider the proposition, consisting of a predicate, its arguments, and its focus record, shown in (11A) below. (11B) shows the proposition at an intermediate stage of generation. The verb "to carry" has been selected to translate the analogy-relation ON and the arguments of the relation have been mapped onto the case roles of the verb. The carrier has been mapped to the protagonist and the weapon to the goal. Since the goal is in focus, the passive voice is selected and the final sentence shown in (11C) is generated.

11A) predicate = analogy-relation
 relation = ON
 carrier = (AIR-VEHICLE
 WATER-VEHICLE)
 weapon = MISSILE
 current focus = MISSILE
 potential focus list =
 ((AIR-VEHICLE WATER-VEHICLE)
 ON
 analogy-relation)

B) verb ==== carry
 protagonist =
 conj ==== and
 head1 ==== aircraft
 head2 ==== water-going vehicle
 goal ==== missile
 voice = passive

C) Missiles are carried by
 water-going vehicles and aircraft.

Pronominalization is determined on selecting a lexical item to translate a predicate argument. When choosing pronominalization, the focus record of the last proposition is checked. If the argument was in focus, pronominalization is selected in place of the full reference for the argument.

7 Limitations and Unimplemented Effects

The current formulation and implementation clearly show how focus information can be successfully used as the basis for surface choice. Further improvements can be made, however, by encoding the tests for surface choice as part of the grammar. The grammar used in TEXT is based on Kay's functional grammar formalism [KAY 79] and in fact allows for the explicit encoding of focus information. Enough information is available at that time to make the tests for an entire category (e.g., verb or noun phrase) instead of a lexical item, thus allowing for more generality in determination of choice.

The implementation can also be extended by including tests for additional

types of surface choice. The influence of focus information on lexical choice, noted earlier, is not currently implemented and would be one place to start. The use of other surface-level structures can also be signaled through focus information. Some of these include parallel sentence structure, subordinate sentence structure, and textual connectives. Parallel sentence structure can be used to increase the cohesiveness of text when focus remains the same from one sentence to the next. When focus shifts to an item just introduced into conversation, subordinate sentence structure can be used to combine the two adjacent propositions into a single complex sentence. When there has been an abrupt shift in focus, textual connectives can be used to ease the transition for the hearer. The implementation of these uses of focus information for surface-level choices remains a topic for future work.

8 Conclusions

The process of generating natural language has been shown to involve a system of choices across a wide spectrum of knowledge sources. A method has been presented that provides a theoretical basis which constrains generation decisions. Furthermore, it illustrates how information arising from decisions about what to say can be used to constrain choices in the surface level expression. While these choices can be arbitrarily determined, an inappropriate decision could easily be made. As systems become more sophisticated, it is imperative that they produce appropriate utterances in order that they communicate effectively with their users.

ACKNOWLEDGMENTS

I would like to thank Aravind K. Joshi, Bonnie Webber, Michael Lebowitz, and Kathy McCoy for their comments and suggestions on various drafts of this paper.

References

- [1] [FIRBA 66]. Firbas, J., "On defining the theme in functional sentence analysis." *Travaux Linguistiques de Prague* 1, Univ. of Alabama Press, 1966.
- [2] [GROSZ 77]. Grosz, B. J., The representation and use of focus in dialogue understanding. Technical note 151, Stanford Research Institute, Menlo Park, Ca (1977)

- [3] [HALLIDAY 67] Halliday, M. A. K., "Notes on transitivity and theme in English." *Journal of Linguistics* 3, 1967.
- [4] [MCKEOWN 80] McKeown, K.R., "Generating relevant explanations natural language responses to questions about database structure." in *Proceedings of AAAI*. Stanford Univ., Stanford, Ca. (1980). pp 306-9.
- [5] [MCKEOWN 82A] McKeown, K.R., Generating natural language text in response to questions about database structure. Technical Report MS-CIS-82-5, Univ. of Pennsylvania, Philadelphia, Pa. (1982)
- [6] [MCKEOWN 82B] McKeown, K.R., "The TEXT system for natural language generation: an overview" in *Proceedings of the 20th Annual Meeting of the ACL*, University of Toronto, Toronto, Canada (1982).
- [7] [SIDNER 79] Sidner, C.L., Towards a computation theory of definite anaphora comprehension in English discourse. Ph.D. Dissertation, MIT, Cambridge, Mass (1979)