Two Kinds of Alternative Sets and a Marking Principle – When to Say *also*

Kristina Striegnitz

Computational Linguistics, Saarland University
kris@coli.uni-sb.de

Abstract. Discourse anaphora other than definite NPs have not received much attention in generation. This paper proposes a strategy for generating the additive particle *also*, a non-nominal discourse anaphor. The strategy is based on a treatment of additive particles as markers rather than presupposition triggers (following Zeevat (2003)) and two sorts of contextually given alternative sets.

1 Introduction

Discourse anaphora are important devices for achieving cohesion in discourse (Halliday and Hasan 1976). In generation, work on anaphoric expressions has mainly concentrated on definite descriptions and pronouns. Other kinds of discourse anaphora, and in particular anaphoric discourse adverbials (Webber et al. 2003), have not received a systematic treatment as anaphoric expressions. Based on the assumption that being able to generate a wider variety of anaphora will allow us to build generation systems that produce more cohesive output, this paper presents a first step toward an approach to the generation of anaphoric adverbials. It proposes a strategy for deciding when to use additive particles, such as *also* or *too*, which establish an anaphoric link to an event previously mentioned in the discourse. While the paper discusses the special case of additive particles, I think that the concepts and mechanisms proposed will also be useful for other anaphoric expressions.

Section 2 reviews the basic properties of additive particles. Section 3 then characterizes the conditions under which additive particles are used. Following work in theoretical semantics, this characterization is based on sets of alternative entities. I further distinguish two kinds of alternative sets that play a role: sets which group entities due to their ontological category and sets which are induced by discourse structure. In Section 4, I introduce the general generation strategy which is based on a treatment of additive particles as context markers as proposed by Zeevat (2003). I argue that this view is better suited from a generation point of view than the traditional view of additive particles as presupposition triggers. Section 5 describes an implementation of the suggested strategy in an application that in particular requires the generation of object descriptions. Section 6 concludes the paper.

2 Additive Particles

Additive particles, such as *also, too, either*, *associate* with a constituent of the sentence they are used in (the *associated constituent*) and express that whatever is said about the referent of this constituent also holds of some *alternative* entity (see, e.g., Karttunen and Peters 1979; Krifka 1992; Rooth 1985; Reis and Rosengren 1997). In (1a), e.g., *also* associates with *Antonia* and expresses that besides Antonia somebody else gave a book to Norbert. In (1b) and (1c), *also* associates with *a book* and *Norbert* respectively, and the alternative entities are some chocolate and Charlie.

(1) a. *Charlie gave a book to Norbert. Antonia* ***also*** *gave a book to Norbert.*
b. *Antonia gave some chocolate to Norbert. She* ***also*** *gave a book to Norbert.*
c. *Antonia gave a book to Charlie. She* ***also*** *gave a book to Norbert.*

This paper aims at determining when to use an additive particle. It will not discuss issues pertaining to the realization, such as the choice between different additive particles or the position of them in a sentence. Therefore, all of my examples use the additive particle *also*, which is the most frequently used and the most flexible.

In order to devise a generation strategy two questions have to be answered: First, what characterizes the situations in which the use of an additive particle is licensed. In particular, we have to know when two entities count as alternatives. Second, how is the use of an additive particle triggered? The next two sections address each of these questions in turn.

3 Two Kinds of Alternative Sets

Above I said that *also* expresses that there is an entity which a) is an alternative of the entity referred to by the associated constituent and of which b) the property attributed to the associated constituent holds as well. So, in a generation system we will have to be able to decide whether an entity has alternatives, i.e., whether two entities belong to the same *alternative set*. Alternative sets are formed according to ontological and contextual restrictions (Krifka 1992; Vallduví and Vilkuna 1998), but what exactly that means is usually not specified any further. I now show that there are two different types of alternative sets that play a role. Consider the following example.

(2) (a) *Antonia invited friends for a Japanese dinner.* (b) *She prepared miso soup, sushi, and tempura.* (c) *She* ***also*** *made some green tea ice cream.*

The associated constituent of *also* is *green tea ice cream*. So, the relevant alternative set has to contain green tea ice cream as well as miso soup, sushi and tempura. There are two possible explanations. First, the alternative set in question could be the ontologically defined set of (Japanese) dishes. Second, the alternative set could be evoked by discourse structure. The *question under discussion* (Roberts 1996; Ginzburg 1996; van Kuppevelt 1995) for sentences (a) and

(b) is what Antonia prepared for dinner. Green tea ice cream, miso soup etc. all belong to the alternative set of all those things that Antonia prepared.

Example (2) does not let us decide which of the two possible alternative sets triggered the use of *also*, but here is an example which can only be explained with the help of structurally evoked sets.

(3) (a) *Antonia hit Norbert* (b) *because he had stolen her bicycle.* (c) *He had* ***also*** *called her a monkey.*

The associated constituent of *also* in the last sentence is *has called her a monkey.* The use of *also* marks that the event of Norbert calling Antonia a monkey (let's call it e_1) as well as the event of Norbert stealing Antonia's bicycle (let's call it e_2) are reasons for Antonia's hitting Norbert. So, the relevant alternative set has to be a set of events including both e_1 and e_2. It is not plausible to assume an ontological category that subsumes both of these events. Discourse structure, on the other hand, does allow us to group e_1 and e_2 into one set. The question under discussion for sentences (b) and (c) is *Why did Antonia hit Norbert?*, i.e., $\lambda x[reason(x, e)]$ with e being the event of Antonia hitting Norbert. The events e_1 and e_2 both fit this description and are therefore alternatives.

Example (4) shows that ontologically defined alternative sets are also important. In this example, *also* expresses that Charlie, besides Norbert, likes cake. So, the relevant alternative set contains at least Charlie and Norbert. The set of all human beings would be a possibility. Discourse structure, on the other hand, does not provide an appropriate alternative set. The issue addressed by sentence (b) is why Norbert ate all the cake, while the question under discussion that sentence (e) is answering is why Antonia scolded Norbert.

(4) (a) *Antonia baked a cake.* (b) *Norbert really likes cake* (c) *and ate it all.* (d) *Antonia scolded him* (e) *because Charlie* ***also*** *likes cake.*

To sum up, I call two entities *alternatives* if they both belong to the same alternative set $\{x|P(x)\}$ where P is either an ontological category or the question under discussion. I will furthermore say that a sentence describing the event e and attributing property $\varphi(e)$ to an entity a is *also-parallel* wrt. the context if there is an alternative b of a such that the discourse context entails that $\varphi(e')$ also holds of b for some discourse old event e'. Additive particles are used to express that a sentence is also-parallel wrt. the context.

The distinction between ontology based and discourse structure based alternative sets is based on an exploratory corpus study of 167 cases of *also* found in the Wall Street Journal and the Brown Corpus. 60% of the 62 cases examined in the Brown Corpus involved discourse structure based alternatives and 32% ontologically defined alternatives. In the remaining cases it was either not possible to decide between the two options (3%) or to analyze the example at all (6%).

4 Additive Particles as Context Markers

Traditionally, *also* is taken to trigger a presupposition. According to Karttunen and Peters (1979), e.g., *also* triggers the following presupposition: "There are

other x under consideration besides the entity e described by the associated constituent, such that what is said about e also holds of x." Zeevat (2003), however, points out that many particles (among them additive particles) have a number of characteristics which are untypical for presupposition triggers. He argues that the behavior of these particles can be explained better if a *marking principle* is assumed and they are treated as *context markers*, i.e., expressions that mark that the content of the current sentence is in a certain relation to the discourse context. The *additive marking principle* relevant here would look as follows: If a sentence is also-parallel wrt. the context, then this has to be marked by an appropriate marker.

Zeevat's (2003) approach is useful from a generation point of view because it predicts that an also-parallel sentence not containing an appropriate marker is infelicitous. Purely presuppositional analyses do not make such predictions. Assume that a generation system has to describe the eventuality *like(antonia,cake)* in a context in which *like(antonia,ice_cream)* holds. The sentence *Antonia likes cake* and the sentence *Antonia also likes cake* both express this content. The second version additionally carries a presupposition which is satisfied in the given context. This presupposition, however, is coupled to the use of the lexical item *also*. It imposes a restriction on the contexts the second sentence can be used in; it does not have any impact on the appropriateness of the version without *also*.

Marking principles impose well-formedness constraints on linguistic expressions. But for generation purposes they can also be viewed as rules licensing the introduction of markers. The next section describes an example of how to integrate a treatment of marking principles into a sentence planner.

5 An Implementation

A strategy for generating additive particles based on the notions described so far is currently being implemented in the generation module of a text based computer game. One of the main generation tasks in this application is the description of rooms and objects.

The system has access to the following knowledge bases: the state of the game world, what the user knows about the game world, and information about the discourse history, such as salience of entities. In the first step schemata are used to plan the text structure. The output of this text is a tree with communicative goals for individual sentences at the leaves. A SPUD like module (Stone et al. 2003) is then used to plan and realize these sentences.

5.1 Detecting Alternative Sets

For detecting ontologically defined alternatives I follow previous approaches, such as Prevost (1995), who uses ontology based alternative sets for generating contrastive intonation, and assume that all entities with the same parent class (the most specific class they belong to) are alternatives. For applications with small, shallow ontologies this seems to work quite well. For bigger ontologies,

user: *Look at the green frog.*
The frog is ugly and slimy. It is wearing a small crown
user: *Look at the brown frog.*
The brown frog also is ugly. It is carrying a tiny sword.

user: *Look at the sword and the crown.*
The sword is old and rusty. It has an inscription. The crown is golden. It also has an inscription.

user: *Look at the frog.*
The frog is green. It is wearing a small crown. It is also carrying a tiny sword.

Fig. 1. Example interactions involving also-parallelism.

however, it might be problematic that the design of the ontology directly influences what counts as alternatives. I furthermore consider entities alternatives if the user explicitly introduced them as a group by using a coordinated or plural NP. The examples at the top of Figure 1 involve ontologically defined alternatives. The green frog and the brown frog are alternatives because they both have the parent class *frog*. The sword and the crown are alternatives because the user introduced them as a group.

Discourse structure based alternative sets are determined during document planning. The document planning module is based on schemata providing templates for how to structure the information in the description. Schemata can be viewed as specifying which questions have to be raised (implicitly) to fulfill a particular communicative goal. Under this view, the steps of a schema (calls of sub-schemata or instructions to retrieve information from the knowledge base) correspond to questions under discussion. So, whenever a query to the knowledge base (corresponding to a question under discussion) yields more than one answer, these answers form an alternative set. The crown and the sword mentioned at the bottom of Figure 1 are alternatives evoked by discourse structure. The schema used for generating this description first presents the physical appearance of the object and then its accessory. The knowledge base query for retrieving the accessory returns several answers (the crown and the sword) which are taken to constitute an alternative set.

5.2 The Sentence Planner and Surface Realizer

SPUD (Stone et al. 2003) starts from communicative goals of the form $\langle Cat, a, \Gamma \rangle$ where Cat is the syntactic category of the linguistic structure that has to be built, a is the main discourse entity described by that structure (e.g., sentences describe events or states), and Γ is a set of facts that the output should convey. SPUD then simultaneously assembles the semantic content and syntactic structure of an utterance achieving this goal. This is done as follows: Starting from a tree consisting of one node (with label Cat), TAG (Tree Adjoining Grammar) elementary trees are added incrementally until a) the tree is syntactically complete, b) all references to entities known to the hearer are unambiguous, and c) all facts specified in Γ are conveyed. In each search step, SPUD computes

all possible ways of extending the current tree, the possibilities are ranked according to heuristics evaluating the progress toward the goal, and the best one is retained. In SPUD's grammar, elementary trees are associated with a representation of their assertions and presuppositions. The addition of an elementary tree is licensed if the asserted part of its semantics is supported by the system's knowledge and the presupposed part is supported by the shared knowledge.

Marking principles constitute additional constraints on the tree being built. The final tree therefore has to satisfy the following additional constraint: if a marking principle applies, the structure has to contain an appropriate marker, and vice versa. In each search step, all elementary trees for markers contributing information that is satisfying a requirement issued by a marking principle are licensed to be added. They are competing with all other possible extensions.

Elementary trees for markers are licensed to be added if their contribution satisfies a requirement issued by a marking principle. That means marking principles are checked in each search step.

The additive marking principle is defined to only apply to syntactically complete sentences. It checks whether the sentence refers to an entity a for which an alternative b is available. If this is the case, it checks whether replacing the reference to a with one to b yields a formula that follows from the shared knowledge. This is best illustrated by means of an example. Assume that the expression built so far is *the frog has a crown*. The grammar associates the presupposition $frog(X)$ and the assertion $have(Z, X, Y) \wedge crown(Y)$ with this sentence. The variables Z, X, and Y are bound to domain entities when choosing the corresponding elementary trees. Let's assume that the intended binding is $\sigma = \{Z \leftarrow e,\ X \leftarrow f_1,\ Y \leftarrow c_1\}$, and let's furthermore assume that the hearer can correctly resolve the presupposition, so that the information conveyed to the hearer is $have(Z, f_1, Y) \wedge crown(Y)$. Z and Y are two entities which are new to the hearer.

Let's say entity f_2 is an alternative of f_1. If there is a variable binding σ' such that $\sigma'(have(Z, f_2, Y) \wedge crown(Y))$ follows from the shared knowledge, then the sentence is also-parallel. Markers with the contribution $\textit{also-parallel}(\sigma(Z), \sigma'(Z))$ are licensed to be added. The pragmatic constraints associated with additive markers furthermore require that $\sigma'(Z)$ be discourse old.

If instead of entity f_1, entity c_1 has an alternative, e.g., c_2, then we have to check for a variable binding σ' such that $\sigma'(have(Z, F_1, c_2))$ follows from the shared knowledge.

6 Conclusions

I have proposed a strategy for generating sentences containing the additive particle additive particles. I assume that the use of additive particles is triggered by a marking principle along the lines of (Zeevat 2003) requiring that a sentence be marked if it is also-parallel wrt. the context. I believe that this view is useful for generating other kinds of anaphoric expression as well, such as other discourse particles, but also, e.g., NPs with the determiner *another*. Also-parallelism is de-

fined in terms of alternative sets. To make the notion of alternative sets useful for generation, we have to make more precise according to which properties entities are grouped in alternative sets. I argue that there are two kinds of alternative sets which play a role in the analysis of *also*: sets which group entities due to their ontological category and sets which are induced by the implicit questions structuring discourse (van Kuppevelt 1995).

This paper has concentrated on characterizing the conditions under which additive markers are used. Further investigation addressing the realization and the influence of salience is needed. However, I believe that the proposed strategy is a starting point for investigating the generation of other focus particles and, more generally, other expressions involving alternative sets, such as the alternative markers studied by Bierner (2001).

References

Bierner, G. (2001). *Alternative Phrases – Theoretical Analysis and Practical Application*. Ph. D. thesis, University of Edinburgh.

Ginzburg, J. (1996). Dynamics and the Semantics of Dialogue. In J. Seligman and D. Westerstahl (Eds.), *Logic, Language and Computation*. CSLI Lecture Notes.

Halliday, M. and R. Hasan (1976). *Cohesion in English*. London: Longman.

Karttunen, L. and S. Peters (1979). Conventional Implicature. In C.-K. Oh and D. A. Dinneen (Eds.), *Presupposition*. Academic Press.

Krifka, M. (1992). Compositional Semantics for Multiple Focus Constructions. In J. Jacobs (Ed.), *Informationsstruktur und Grammatik*. Westdeutscher Verlag.

Prevost, S. (1995). *A Semantics of Contrast and Information Structure for Specifying Intonation in Spoken Language Generation*. Ph. D. thesis, University of Pennsylvania.

Reis, M. and I. Rosengren (1997). A Modular Approach to the Grammar of Additive Particles: the Case of German *Auch*. *Journal of Semantics 14*.

Roberts, C. (1996). Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In J.-H. Yoon and A. Kathol (Eds.), *Papers in Semantics*, Volume 49 of *OSU Working Papers in Linguistics*. Ohio State University.

Rooth, M. (1985). *Association with Focus*. Ph. D. thesis, University of Massachusetts, Amherst.

Stone, M., C. Doran, B. Webber, T. Bleam, and M. Palmer (2003). Microplanning with Communicative Intentions: The SPUD System. *Computational Intelligence 19*, 311–381.

Vallduví, E. and M. Vilkuna (1998). On Rheme and Kontrast. In *The Limits of Syntax*, Volume 29 of *Syntax and Semantics*. Academic Press.

van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. *Journal of Linguistics 31*.

Webber, B., M. Stone, A. Joshi, and A. Knott (2003). Anaphora and Discourse Structure. *Computational Linguistics 29*.

Zeevat, H. (2003). Particles: Presupposition Triggers, Context Markers or Speech Act Markers. In R. Blutner and H. Zeevat (Eds.), *Optimality Theory and Pragmatics*. Palgrave-McMillan.