One-Anaphora and the Case for Discourse-Driven Referring Expression Generation

Robert Dale

Centre for Language Technology
Macquarie University
Sydney NSW 2109
`Robert.Dale@mq.edu.au`

Abstract

Conventional approaches to the generation of referring expressions place the task within a pipelined architecture, typically somewhere between text planning and linguistic realisation. In this paper, we look at the issues that arise in generating *one*-anaphoric referring expressions; examination of this task causes us to reflect on the current predominant architectural models for natural language generation, and leads us to suggest an alternative architecture where decisions that influence forms of reference happen much earlier in the process of natural language generation.

1 Introduction

Referring expression generation is a much-explored task within natural language generation: given an internal symbol that corresponds to an entity in some real or imagined world, we need to work out what properties of that entity should be used to describe the entity so that our hearer will be able to identify it as the intended referent. Many different algorithms have been developed to address this task, which is generally conceived of as mapping from a symbol—effectively, a referent—to a set of properties—a sense. The computation of the appropriate set of properties to use takes account of the other potential referents in the context, selecting properties which rule these distractors out of consideration.

Conventional approaches to the process of generating referring expressions place the task within a pipelined architecture, where it is assumed that questions of what content should be conveyed in a text are resolved before questions of surface form are considered; this is the well-known STRATEGY vs TACTICS distinction first discussed in the context of natural language generation in the mid-1970s. However, it is not clear exactly where in the pipeline the process of generating referring expressions should belong. In Reiter and Dale [2002], reflecting current practice in the field, we positioned it in the microplanning stage, where microplanning is an intermediate stage lying between text planning and surface realisation. Even there, however, we noted that there are interactions between the three microplanning tasks of sentence planning, lexical selection and referring expression generation that argue for a more interleaved constraint-based approach to the problem.

The principal focus of existing work has been the generation of definite noun phrase references; relatively little has been written on generating other kinds of referring expressions. In particular, there is virtually no work on the generation of *one*-anaphora. Taking up some ideas first explored in [Dale 1992, 1995], this paper looks at how a consideration of where *one*-anaphora fits into the gen-

eration process might cause us to review the kinds of architecture that are required for natural language generation.

Section 2 first summarises the conventional approach to referring expression generation, and reviews how this fits into the standard architectural models for natural language generation. Section 3 introduces the phenomenon of *one*-anaphora, before going on to explore how the generation of *one*-anaphora might be integrated into existing approaches of referring expression generation. Section 4 then suggests an alternative approach, where the decision to use a *one*-anaphor is made much earlier in the generation process. Section 5 concludes by discussing how this alternate approach might impact both on other aspects of referring expression generation, and on natural language interpretation.

2 Conventional Approaches to Referring Expression Generation

Anaphoric reference to an entity previously mentioned in a discourse can be carried out using any of a number of different strategies: in particular, pronominal anaphora, definite noun phrase anaphora and *one*-anaphora may each be used in appropriate discourse contexts, as demonstrated in examples (1)–(3) respectively.

(1) a. John has a red jumper.
     b. He wears *it* on Sundays.
(2) a. John has a red jumper and a blue cardigan.
     b. He wears *the jumper* on Sundays.
(3) a. John has a red jumper and a blue one.
     b. He wears *the red one* on Sundays.

There is now a well-established body of work in natural language generation that focusses on the problem of generating definite noun phrase anaphora; see Chapter 5 in Reiter and Dale [2000] for a review. Work on the generation of pronominal anaphora is somewhat less developed, with researchers often falling back on some notion of focus as the prime determinant of whether pronominalisation is possible; the major problem here is coming up with an independently motivated notion of what it means to be 'in focus'. The generation of *one*-anaphoric expressions, however, has been virtually ignored, apart from some initial explorations in Davey [1979], Jameson and Wahlster [1982], and Dale [1992, 1995].

A high-level characterisation of the algorithm that underlies much work in referring expression generation is shown in Figure 1. This is deficient in a number of regards: pronouns may be used even if the intended referent is not in focus—see, for example, the centering algorithm of Grosz *et al* [1983]—and a definite noun phrase may be used even if the referent has not been mentioned before, or alternatively its form may be further constrained in some way by the structure of the discourse. However, these complications are not important for our present purposes. The question this paper addresses is as follows: how does the decision to use a *one*-anaphoric expression fit into this kind of algorithm?

3 *One*-Anaphora

3.1 *One*-Anaphora as Syntactic Substitution

The phenomenon of *one*-anaphora is reasonably well discussed in the linguistics literature: in terms of X-bar theory, for example, the pro-form *one* is generally characterised as a substitute for an $\bar{\text{N}}$ constituent (see, for example, Radford [1981:94–95], McCawley [1988:185–186]); and the systemic literature provides some discussion of the nature of *one* as a substitute (see, for example, Halliday and Hasan [1976:89–98]). Although these treatments differ in a number of respects, both characterise effectively the same syntactic constraints on when *one*-anaphora

Given an intended referent r:
begin
    if r is in focus **then** use a pronoun
    elseif r has been mentioned in the discourse already
    then build a definite noun phrase
    else build an initial indefinite reference
end

Figure 1: A Skeletal Referring Expression Generation Algorithm

is possible: the *one* form is seen to substitute for a head noun and some number of modifiers of that noun.

For the purposes of natural language generation, we could take this notion of substitution literally: each time we generate a noun phrase structure, we could then compare this against noun phrases in some locally specified discourse context, and then replace any replicated substructure by the form *one*. Assume, for the moment, that a *one*-anaphor always has its antecedent in the previous clause.[1] The generation of *one*-anaphora can then be characterised as follows. Suppose P is a set consisting of the noun phrase structures that appear in the previous clause:

- Given an intended referent r, determine the semantic content needed to identify this referent to the hearer.

- Work out the syntactic structure that realizes this semantic content; call this s.

- Compare s against each $p \in P$, and look for common substructure starting at the head noun and working outwards; replace the largest common substructure found in s by the form *one*.

So, given an antecedent noun phrase as in (4a) and a subsequent noun phrase as in (4b), we can substitute the *one* form to produce (4c), with the *one*-anaphor substitut-

ing for the $\bar{\text{N}}$ constituent *mouldy Germanic manuscript*. [2]

(4)   a.   [a [large [mouldy [Germanic [manuscript $_{\text{N}}]_{\bar{\text{N}}}]_{\bar{\text{N}}}]_{\bar{\text{N}}}]_{\text{NP}}$]
       b.   [a [small [mouldy [Germanic [manuscript $_{\text{N}}]_{\bar{\text{N}}}]_{\bar{\text{N}}}]_{\bar{\text{N}}}]_{\text{NP}}$]
       c.   [a [small [one $_{\bar{\text{N}}}]_{\bar{\text{N}}}]_{\text{NP}}$]

There are a number of problems with this approach. First, it sanctions the use of *one*-anaphora where we would want to rule it out on semantic grounds, as in the following constructed example:

(5)   a.   Do you have any wine bottles?
       b.   No, but I have a red one.

Second, it rules out *one*-anaphora in cases where the syntactic structures are more distinct, yet we would still want to allow the use of *one*-anaphora, as in the following example:

(6)   a.   Mary chained her bicycle to a steel fence.
       b.   Fred chained his to one made of wood.

But quite apart from these concerns (see [Dale 1992:215-230] for a discussion), it also

---

[1]This is not always true, but the algorithm described here can be trivially extended to deal with other cases.

[2]We will fairly randomly switch between consideration of definite and indefinite *one*-anaphoric forms: for the purposes of the present discussion, any complications introduced by this aspect of discourse status appear to be orthogonal to the issues we are concerned with.

seems a rather wasteful approach. Since the commonality between the antecedent and the anaphor has something to do with shared *semantic* content, why should we go as far as working out the *syntactic* structure required to realise the second NP in order to determine if *one*-anaphora can be used? Syntactic substitution may be an appropriate way to characterise the behaviour of the *one* form when discussing it as a linguistic phenomenon, but that does not mean it should serve as the basis of a generation algorithm.

3.2 *One*-Anaphora as Semantic Substitution

The above objection to the syntactic substitution approach suggests a better solution: look for shared structure at the semantic level. Suppose we have the semantic structure that corresponds to the noun phrase *the red jumper*, and suppose we have gone as far as to generate the semantic content that could be ultimately realised as the noun phrase *the blue jumper*. These semantic structures could be represented as follows:

(7) $\text{type}(x1, \text{jumper}) \wedge \text{colour}(x1, \text{red})$

(8) $\text{type}(x2, \text{jumper}) \wedge \text{colour}(x2, \text{blue})$

By identifying what it is that the antecedent NP and the anaphoric NP have in common at the level of semantics, we both avoid unnecessary work in building syntactic structure, and at the same time constrain more correctly the use of *one*-anaphoric forms. This method is elaborated further in [Dale 1992:220–226], and is based on observations made in the work of Webber [1979].

This approach provides us with a way of generating *one*-anaphoric expressions that fits into the general algorithmic structure we sketched in Figure 1; all that is required is that the algorithm maintain a distinction between determining the semantic content of a referring expression and the linguistic realisation of that content, a fairly

standard separation useful for other purposes in any case.[3] We then complicate the algorithm to check for the possibility of using a *one*-anaphoric construction once the semantic content has been determined, simply by checking whether there is a replication of semantic content that includes at least the content that would be realised by the head noun. A revised version of the skeletal algorithm is shown in Figure 2.

In suggesting this approach, we have effectively shifted the decision to use *one*-anaphora further back in the generation process, replacing a process of syntactic substitution by one of semantic substitution. In the next section, we argue that we can shift the decision further back still: if we take the stance that *one*-anaphora is typically used to achieve a specific range of discourse functions, that it makes sense to have the discourse planning stage of a generation system impose a requirement that *one*-anaphora should be used when those discourse functions are being realised.

4   Discourse-Driven Generation of *One*-Anaphoric Expressions

4.1   The Functions of *One*-Anaphora in Discourse

Observation suggests that *one*-anaphoric forms are used to achieve particular discourse functions; a common such function, for example, is when a speaker contrasts two entities.[4] It seems reasonable to suppose that, at the discourse planning stage, a generator will already know that it is contrasting two entities; but if the system knows that it is performing a contrast, then at that stage it should already be able to suggest that a *one*-

---

[3]For example, it allows us to generate *a red jumper* and *a jumper which is red* as variants of the same basic semantic content.

[4]Clearly, an appropriate corpus analysis would determine which particular discourse functions are characteristic of the use of *one*. Just such an analysis is currently underway by Gardiner (forthcoming).

```
Given an intended referent r:
begin
    if r is in focus then use a pronoun
    elseif r has been mentioned in the discourse already
    then begin
        build the semantics for a definite noun phrase
        if there is shared structure with a previous noun phrase then elide it
    end
    else begin
        build the semantics for an initial indefinite reference
        if there is shared structure with a previous noun phrase then elide it
    end
end
```

Figure 2: A Revised Skeletal Referring Expression Generation Algorithm

anaphor may be used. In other words: why construct an elaborate mechanism to determine a semantic structure that can be subsequently elided if this means rediscovering something the generator already knew?

The idea that *one*-anaphora is used in the context of particular discourse functions has been noted in the literature before: Dahl [1985] and LuperFoy [1991:114–159] both discuss this aspect of *one*-anaphora at some length. LuperFoy's observations are closest to those that lie behind the view taken here. She suggests that uses of *one*-anaphoric forms correspond to three particular discourse functions: to contrast two sets of individuals, to denote a representative sample of a set introduced by the antecedent, and to refer to a new specimen of a type that is salient in the discourse; examples of each of these categories are provided in (9)–(11) respectively:

(9)   a.   John has a magenta Capri.
       b.   Robert has *a reef-green one.*

(10)   a.   John has several cars.
       b.   *The smallest one* is a Capri.

(11)   a.   John has several old cars.
       b.   Mary wants to buy him *a new one.*

In the terms of Rhetorical Structure Theory [Mann and Thompson 1987], the discourse function in (9) is one of CONTRAST, and those in (10) and (11) are instances of the ELABORATION relation.

4.2   How We Might Integrate *One*-Anaphora in Text Planning

We are concerned in the first instance with the monologic case, where both the sentence containing the *one*-anaphora and the sentence containing its antecedent are spoken by the same conversational participant; as will become clear, a quite separate explanation is likley to be required for dialogic uses of *one*.

In the sample discourses just presented, it seems plausible to suggest that the two sentences are 'spoken as pairs'. In (9), the speaker utters the two sentences precisely in order to draw a contrast; in (10) and (11), the second sentence is only a coherent contribution to a discourse given the background provided by the first sentence.

In a natural language generation system which performs text planning, we take the view that the contrast or elaboration that is being performed is the most important issue; the particular linguistic expressions con-

structed are subsidiary to these aims. Viewed in this way, it makes sense for the text planner to PRESELECT some of the linguistic features of the utterances to be produced when the discourse relation has been decided upon.[5] Clearly there are other forms of contrast than those realised by means of *one*-anaphors, and a fully-fleshed out model of how this preselection mechanism might work will require finding an appropriate level of abstraction for expressing 'linguistically-realised contrast'; however, for present purposes we can focus on instances of contrast where it is similar entities that are being contrasted, and assume for simplicity that *one*-anaphora is the only appropriate contrastive device available.

Any text planning component has to decide when it wants to use rhetorical devices such as contrast. The proposal here is that, when such a goal has been selected for whatever reason, then, provided some additional constraints are met, the text planner can already at that point determine that specific linguistic forms should be used. In effect, the choice of a specific discourse relation brings with it linguistic consequences. This is entirely plausible where discourse connectives are concerned: a decision to use, for example, a relationship of CAUSE might lead automatically to the decision to use the discourse connective *because*. Here, we are extending this idea to cover also elements within the expression of the propositions to be conveyed. The present case under discussion is shown schematically in Figure 3, where a desire to use a CONTRAST relation, combined with a particular configuration of knowledge in the knowledge base, results in the use of a specific rhetorical structure with some prespecified lexical content. In this case, the constraints on the configuration of knowledge are that the two entities x_1 and x_2 share the same
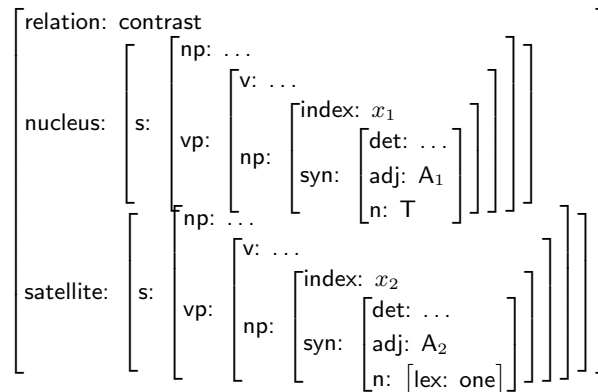

Figure 3: A schematic discourse structure that preselects lexical material

semantic type but have differing values for other attributes, here expressed by the adjectives A_1 and A_2; precisely the circumstances under which *one*-anaphora is possible.

Clearly the picture is considerably more complicated than this simple sketch implies, but the general idea should be clear. As suggested above, what this view does is to push the decision to use a *one*-anaphoric expression further back still in the generation process. This sites the decision in a far more appropriate place: deciding when a contrast should be made is a much larger question that must be faced by any text planning system. Ultimately, the view taken here is that contrast is just one device that we use to produce coherent discourse: one way of characterising the general problem for a text planner is as the decision of what to say next, and here notions like topic maintenance and topic chaining are crucially important. Contrasting two clusters of information stored in a knowledge base is just another of these associative devices that can be used to build a coherent text on the basis of relations that reside in the underlying knowledge base.

---

[5]The use of the term 'preselect', a term from work in systemic approaches to generation, is deliberate. What we are arguing for here amounts to an interstratal preselection of lexicogrammatical features in the sense of Matthiessen and Bateman [1991:62–65].

5   Conclusions and Future Work

We have argued that *one*-anaphora is best viewed as a linguistic phenomenon that is a natural consequence of the speaker's choice to use specific subject-matter discourse relations, and that a consequence of this is that the decision to use *one*-anaphora should, at least in part, be determined at the level of discourse planning. We have sketched how this might work in the case of contrastive uses of *one*, but a similar story can be told for the set-elaborative function.

At the outset, we asked how the generation of *one*-anaphora could be integrated into existing referring expression generation algorithms. These algorithms assume that they are given some symbol that corresponds to the intended referent, and then attempt to determine what content should be used to identify this intended referent. This model is incompatible with the approach proposed here, since the approach we have argued for lacks a distinct stage in the processing where the intended referent is only indicated by some internal symbol. In order to integrate the generation of *one*-anaphora into conventional generation algorithms, the assumption that the referring expression generator is given nothing more to work with than the symbol that corresponds to the intended referent has to be abandoned, and the bandwidth of communication between the discourse planner and the referring expression generator increased: ideally, the referring expression generator is told not only what the intended referent is, but also what its function in the discourse is.

There is some precedent for this idea. McDonald's [1980] work on referring expression generation within MUMBLE includes a facility whereby the expert system driving the generator can specify that a message element (i.e., an internal symbol corresponding to the intended referent) is 'ontologically of a sort that cannot be pronominalized' [1980:217]:

this allows the expert system to specify that some information has to be expressed for descriptive, rather than purely referential, purposes. A similar broadening of the bandwidth is visible in McKeown's TEXT [McKeown 1985], where the text planning component can indicate to the linguistic realisation component that a particular entity is the focus of the utterance, resulting in pronominalisation; and the same idea finds expression in the use of the CENTRE attribute in Dale's EPICURE [1992:170–171]. The present work suggests that these devices can be seen as instances of a more general mechanism where the discourse purpose of a referring expression plays a role in how that referring expression is best realised. Above, we have discussed one specific discourse function, which we might characterise more precisely as CONTRAST-TWO-SIMILAR-ENTITIES; other instances of the use of *one* would be characterised by the discourse function SELECT-ELEMENT-FROM-MENTIONED-SET. The same idea, however, can be used to provide a new way of thinking about existing well-explored reference tasks: so, for example, in appropriate discourse contexts, pronominalisation may be an automatic consequence of the discourse functions MAINTAIN-AS-FOCUS and SHIFT-INTO-FOCUS; initial reference might be best thought of as a consequence of the discourse function INTRODUCE-ENTITY; and different instances of subsequent reference might be cases of either DISTINGUISH-ENTITY or ATTRIBUTE-ADDITIONAL-INFORMATION, or even combinations of both.

Further work is required in order to determine how best to rearrange generation architectures to integrate these observations. By abandoning traditional architectural divisions into pipelined components, systems based on systemic functional grammar (see, for example, [Matthiessen and Bateman 1991]) already allow sufficient flexibility to incorporate the mechanisms discussed here. However, the absence of distinct processing

modules with well-defined interfaces between them is generally considered to make it more difficult to build practical systems which can be easily re-used and maintained. A question for further research is whether, taking the observations of this paper into account, we can characterise the required interactions between referring expression generation and other aspects of the generation task in such a way that modular systems can be built.

An additional interesting direction that is opened up by this view is that of how we might revise our models of natural language analysis to take account of the interstratal relationships between discourse planning elements and surface forms. If, for example, we can characterise the generation of a pronominal form as discourse-planning construct that has as its base a discourse function of MAINTAIN-AS-FOCUS, then we may also be able to use such a multi-level construct when interpreting a pronoun: the idea here would be that, instead of using a more traditional level-by-level analysis (syntactic analyis, then semantic analysis, followed by interpretation in context), we can hypothesis information at all of these higher levels simultaneously on the basis of the presence of the surface form. Of course, this is only a sketch, and there is significant work to be done in fleshing this out; however, the basic idea offers a novel way of thinking about both language analysis and language generation.

References

D Dahl [1985] *The Structure and Function of One-Anaphora in English.* Indiana University Linguistics Club.

R Dale [1992] *Generating Referring Expressions.* MIT Press.

A Davey [1978] *Discourse Production.* Edinburgh University Press.

M Gardiner [forthcoming] *Identifying and resolving one-anaphora.* Honours Thesis, Centre for Language Technology, Macquarie University.

B Grosz, A Joshi and S Weinstein [1983] *Providing a Unified Account of Definite Noun Phrases in Discourse.* In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, Massachusetts, 15th–17th June 1983.

M Halliday and R Hasan [1976] *Cohesion in English.* Longman.

E Hovy [1991] *Approaches to the Planning of Coherent Text.* Pages 83–102 in C Paris, W Swartout and W Mann (eds), *Natural Language Generation in Artificial Intelligence and Computational Linguistics.* Kluwer Academic Publishers.

S LuperFoy [1991] Discourse Pegs: A Computational Analysis of Context-Dependent Referring Expressions. PhD Thesis, University of Texas at Austin.

J McCawley [1988] *The Syntactic Phenomena of English*, Volume 1. The University of Chicago Press.

D McDonald [1980] Natural Language Generation as a Process of Decision-Making Under Constraints. PhD Thesis, MIT.

K McKeown [1985] *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text.* Cambridge University Press.

W Mann and S Thompson [1987] Rhetorical Structure Theory: A Theory of Text Organisation. USC/ Information Sciences Institute Technical Report RS–87–190.

C Matthiessen and J Bateman [1991] *Text-Generation and Systemic Functional Linguistics.* Pinter.

A Radford [1981] *Transformational Syntax.* Cambridge University Press.

E Reiter [1990] Generating Appropriate Natural Language Object Descriptions. PhD Thesis, Harvard University.

E Reiter and R Dale [1992] A Fast Algorithm for the Generation of Referring Expressions. In *Proceedings of Coling–92*, Nantes, France.

E Reiter and R Dale [2000] *Building Natural Language Generation Systems.* Cambridge University Press.

P Sibun [1991] Locally Organized Text Generation. COINS Technical Report 91–73, Department of Computer and Information Science, University of Massachusetts.

H Thompson [1977] Strategy and Tactics in Language Production. In *Papers from the Thirteenth Regional Meeting of the Chicago Linguistics Society*, Chicago.

B Webber [1979] *A Formal Approach to Discourse Anaphora.* Garland Press.