

Abstract verbs

Richard Power

Department of Computing
Open University
Milton Keynes MK7 6AA, UK
r.power@open.ac.uk

Abstract

Existing generation systems use verbs almost exclusively to describe actions/events or to ascribe properties. In doing so, they achieve a direct concrete style of the kind often recommended in style manuals. However in many genres, including academic writing, it is common to find verbs expressing abstract relationships, with events etc. pushed down into nominalisations. This paper illustrates two important classes of abstract verb, one expressing discourse relations, the other expressing participant roles, and discusses some theoretical and practical reasons for studying such verbs and including them in generation systems.

1 Introduction

Many writing manuals enjoin us to freshen up our verbs. A bad writing style, they tell us, is characterised by overuse of abstract verbs like ‘involve’ and ‘characterise’ (oh dear), whereas good prose is enlivened by concrete verbs denoting actions and events. In his book *Style: Towards Clarity and Grace* (Williams, 1990) illustrates the two styles by comparing the following sentences:

(A) Because we knew nothing about local conditions, we could not determine how effectively the committee had allocated funds to areas that most needed assistance.

(B) Our lack of knowledge about local conditions precluded determination of committee action effectiveness in fund allocation to those areas in greatest need of assistance.

Sentence (A), most of us would agree, is clearly expressed, while its paraphrase (B) is turgid. The reason, according to Williams, is that the first sentence expresses actions through verbs (determine, allocate), while the second pushes them down into nominalisations.

Computational Linguists, like most academics, are often unable to resist the lure of the abstract verb. But how about our programs? Let us look at some output samples from Natural Language Generation

(NLG) systems, as surveyed by Pavia (1998).

Behrens’s principal activities were architecture and industrial design. He made electrical appliances and prototype flasks. He built the high tension plant and the turbine factory for AEG in 1908-1910. He built a housing for the workers of AEG in Henningsdorf.

Komet, (Bateman and Teich, 1995)

To schedule the appointment:

1. Choose the start time of the appointment.
2. Enter the description of the appointment.
3. Click on the Insert button.

DRAFTER-2 (Power and Scott, 1998)

This jewel is a necklace and is in the Art Deco style. It was made in 1920. It is made from moonstone and silver rock-crystal. In colour, it is coral. It differs from the previous item, in that whereas that was made by Arthur and Georgie Gaskin, this was made by H.G.Murphy.

ILEX, (Oberlander et al., 1998)

These typical products of current NLG present a paradox. On the one hand, they conform to the advice given by Joseph Williams and other style gurus. The verbs are concrete and usually denote actions (build, click, make). Noun phrases denote people and things rather than nominalised events or propositions. On the other hand, the texts are not really well-written; their short mechanical sentences suggest naivety and limited expressive resources. Concrete writing is not always good writing; turning this on its head, we might question whether the style manuals are really justified in dismissing an abstract style as necessarily bad. Perhaps we would do better to understand why abstract writing is common in some genres and whether it serves a purpose other than bamboozling the reader.

As Halliday among others has pointed out (Halliday and Martin, 1993), an abstract style is common in genres (like scientific writing) where **argument predominates over description**. In argument, it is often necessary to make comments upon propositions; therefore, we find many sentences in which the subject of the verb denotes a proposition (or a complex of propositions), either through a sentential

complement, or a nominalisation, or an anaphoric reference to a proposition mentioned earlier. Because sentences of this kind are often needed in argumentative writing, they have taken on the significance of a hallmark through which academics can certify their distinctive professional expertise — the literary counterpart of the doctor’s stethoscope or the lawyer’s wig. As a result, academics (and bureaucrats, etc.) may feel **obliged** to write in this way, even when the subject-matter is not argumentative, and could be expressed through straightforward narration or description.

If this analysis is right, we should not dismiss abstract writing as bad writing, either when writing papers or when designing NLG systems. We should aim for mastery of both styles, so that the system can choose whichever is most appropriate in any given case.

2 Discourse verbs

The concept of an abstract verb is initially perplexing: when events or propositions are pushed down into nominalisations or sentential complements, what is left for the main verb to do? One obvious answer is that verbs can take over the task of signalling relations among propositions: in other words, they can express discourse relations.

Take for example sentence (B) from the last section, ‘Our lack of knowledge about local conditions precluded determination of committee action effectiveness...’. The main verb here is ‘precluded’, and its meaning corresponds to the relation NON-VOLITIONAL CAUSE from Rhetorical Structure Theory (RST) (Mann and Thompson, 1987), with the extra twist that the nucleus is negated: in general, ‘X precluded Y’ means (roughly) ‘Because X happened, Y could not happen’.

How common are verbs expressing discourse relations? This is a complex question that would require investigation of corpora from varied genres, but merely by consulting thesauruses or works like the *Cambridge English Lexicon* (Hindmarsh, 1980) we can confirm that there are many such verbs, especially for causal and inferential relations. Here for example are some causal verbs roughly grouped:

CAUSING: bring about, cause, induce, lead to, make
 RESULTING: derive from, result from, stem from
 INITIATING: launch, originate, set in motion, start
 PROVOKING: arouse, elicit, evoke, provoke, trigger
 HELPING: contribute to, facilitate, help, stimulate

Most of these can be used as alternatives to discourse connectives like ‘because’ and ‘as a result’, except that their meanings are usually more precise. For instance, ‘Losing his job triggered John’s depression’ could be paraphrased roughly by ‘John was depressed because he lost his job’, but the discourse

verb carries the extra implication that the depression acquired a momentum beyond the original cause.

Considering the obvious prevalence of such discourse verbs, the NLG literature is curiously silent. Almost everybody follows the standard line that rhetorical relations are realised by discourse connectives. In work based on RST, discourse relations are assumed to hold between clauses, as in the manual for marking up the RST Discourse Treebank (Carlson and Marcu, 2001), where discourse verbs are not mentioned at all. The only references we can find in NLG are from Laurence Danlos, who was the first to point out that verbs like ‘cause’ and ‘precede’ can signal the relations expressed by discourse connectives like ‘because’ and ‘next’, and to label such verbs *discourse verbs* (Danlos, 2006). Among other things, Danlos observes that the subject of a discourse verb is often an anaphoric reference to an event (or proposition) mentioned in a previous sentence, or sometimes to the agent of this event:

Ted left. This preceded Sue’s arrival.

Ted didn’t stop joking. This (He) caused hilarity among his friends.

We have seen examples of verbs for CAUSE and SEQUENCE; how about the other RST relations? Danlos does not pursue this issue because she is interested mainly in issues of lexical representation in the framework of Segmented Discourse Representation Theory. To show that verbs can be used across the whole spectrum of discourse relations, rather than one or two exceptional cases, the table below lists the original RST relations from Mann and Thompson (1987) with examples (where possible) using discourse verbs.

ANTITHESIS: Taking part in sport beats watching it on television
 BACKGROUND: [No example]
 CONCESSION: John’s occasional mistakes do not mean he is not a good referee
 ENABLEMENT: Filling in the enclosed form enables you to apply for membership
 EVIDENCE: My learning the system in five minutes shows that it is easy to use
 JUSTIFY: [No example]
 CIRCUMSTANCE: The writing of his first novel coincided with his stay in Rome
 SOLUTIONHOOD: Adding a scroll bar solves the problem of text not fitting on the screen
 VOLITIONAL CAUSE: The dark clouds forming overhead motivated him to bring an umbrella
 VOLITIONAL RESULT: Our boat journey was motivated by the serious flooding
 NON-VOLITIONAL CAUSE: The abundant harvest brought about a surplus of corn
 NON-VOLITIONAL RESULT: The bomb exploding in the shopping centre resulted in 30 people being hurt

PURPOSE: His trip to Selfridges served to buy a present for his sister
 CONDITION: Claiming benefit requires that the mother is less than 18 years old
 OTHERWISE: Late submission of papers will preclude their consideration
 EVALUATION: Our offices being open 7 days a week adds up to a better service
 RESTATEMENT: Omitting needless words means keeping your writing concise
 SUMMARY: [No example]
 SEQUENCE: Ted's departure preceded Sue's arrival
 CONTRAST: John's liking for beetroot contrasts with Mary's disgust
 JOINT: [No example]

3 Theta verbs

We have noted one kind of abstract verb, the category that Danlos calls 'discourse verbs'. Are there others? A plausible further candidate can be found in verbs like 'performed' and 'underwent' which express *thematic roles* — that is, relations between eventualities and their participants. When an event is described directly, these thematic roles are expressed implicitly by syntactic relations like subject and object:

The police interrogated the suspect

However, to highlight a thematic role we can also express it by a verb:

The police performed an interrogation
 The suspect underwent an interrogation

As a label for this group I suggest *theta verbs*. Here is a list of examples for some common thematic roles; note that some circumstantial roles are equivalent to discourse relations so that we get some overlap with discourse verbs.

Participant roles

AGENT The doctor examined Mary ⇒ The doctor performed an examination (of Mary)
 PATIENT The doctor examined Mary ⇒ Mary underwent an examination (by the doctor)
 EXPERIENCER Mary feared her boss ⇒ Mary experienced a fear of her boss
 BENEFICIARY John paid Mary compensation ⇒ Mary benefitted from (John's) payment of compensation
 INSTRUMENT Mary opened the drawer with a hairpin ⇒ A hairpin was employed (by Mary) to open the drawer

Circumstantial roles

CAUSE John fell over by treading on a banana skin ⇒ Treading on a banana skin caused John's fall
 PURPOSE John sacked Mary to reduce his costs ⇒ Reducing his costs motivated John's dismissal of Mary
 MANNER John dismissed Mary quickly and insensitively ⇒ Speed and insensitivity characterised John's dismissal of Mary

PLACE The committee met in the city hall ⇒ The city hall housed the meeting of the committee
 TIME The minister arrived in the afternoon ⇒ The afternoon saw the arrival of the minister

Some roles can be expressed by several verbs, with different shades of meaning (e.g., for AGENT we also have achieve, accomplish, commit, engineer, and so forth). There is also at least one verb that can express *any* thematic role, thus achieving a vagueness that may assist the non-committal writer (but not the reader):

AGENT: The dismissal involved an American employer
 PATIENT: The dismissal involved an elderly secretary
 EXPERIENCER: The fear of dismissal involved the staff
 BENEFICIARY: The compensation payment involved the secretary
 INSTRUMENT: Opening the drawer involved a hairpin
 CAUSE: John's fall involved treading on a banana skin
 PURPOSE: The dismissal involved reducing John's costs
 MANNER: The dismissal of the secretary involved unseemly haste
 PLACE: The committee meeting involved the city hall
 TIME: The minister's arrival involved a Monday afternoon

4 Implications

One reason for studying abstract verbs is simply to enhance the quality and variety of generated texts. This would require (a) an expansion of grammars to cover formulations with abstract verbs, and (b) some kind of empirical investigation into when the use of an abstract verb is appropriate. However, discourse verbs (in particular) are also theoretically interesting because they challenge existing assumptions about rhetorical relations and their realization in discourse:

1. **Argument role:** In RST, argument roles are labelled *nucleus* and *satellite*; what happens to these roles when the relationship is expressed by a discourse verb? Using a verb introduces distinctions like given-new, and notions like agency, which are not relevant for discourse connectives.
2. **Syntactic features:** A discourse relation expressed by a verb can draw on the distinctions expressed by syntactic features like tense, aspect, voice, modality, and negation, most of which are irrelevant for discourse connectives. For instance, if the relation is EVIDENCE, the issue of when the evidential relationship was first noticed is made explicit by the tense of the verb.
3. **Span structure:** Discourse verbs provide additional evidence against RST's theory of span structure, as described by Knott et al. (2001).
4. **Repertoire of relations:** The classification of rhetorical relations might look very different

when based on evidence that includes discourse verbs as well as connectives.

Points 3 and 4 require some elaboration, which is given below.

4.1 Span structure

RST makes two central claims about span structure (Knott et al., 2001): first, that atomic spans are clauses; and second, that rhetorical relationships between spans (or their meanings) are expressed by combining adjacent spans through a schema application. The prototypical case is illustrated by the following text from the CLEF domain (Hallett et al., 2007) which realizes an EVIDENCE relation:

The patient looks anaemic, so a packed red cell transfusion is probably needed.

The atomic spans are adjacent clauses, linked by a discourse connective ('so') expressing the rhetorical relation between them. Now consider the same message expressed using a discourse verb:

The patient looks anaemic. This suggests the need for a packed red cell transfusion.

It is striking how an apparently minor change throws the RST analysis into confusion. First, the nucleus of the EVIDENCE relation is now expressed by a noun phrase; this means for example that it could not be marked up as a rhetorical argument in the RST Discourse Treebank (Carlson and Marcu, 2001). Second, this noun phrase is not linked directly to the span describing the anaemia, but rather to a pronoun ('this') coreferring anaphorically with this span. One could argue for some kind of presentational relation between the first sentence and the second (e.g., ELABORATION), but on such an analysis the rhetorical structure assigned to the text no longer addresses its main point.

4.2 Repertoire of relations

If we try to classify discourse relations with reference to discourse verbs rather than connectives, the most obvious change is that the verbs are more specific. We have illustrated this point already in listing causal verbs, which subclassify causal relations along several dimensions identified by Talmy (1988) (e.g., change/stasis, generation/enablement, weak/strong control, temporary/enduring effect). Consider for instance the following more specific alternatives to 'John was drowsy because he took antihistamines':

Taking antihistamines made John drowsy
Taking antihistamines kept John drowsy
Taking antihistamines enabled John to be drowsy
Taking antihistamines helped make John drowsy
Taking antihistamines triggered John's drowsiness

Such examples suggest that discourse verbs are far more effective than connectives in putting discourse

relations under the microscope, at least for the groups of relations concerned with causality and inference.

References

- J. A. Bateman and E. Teich. 1995. Selective information presentation in an integrated publication system: an application of genre-driven text generation. *Information Processing and Management*, 31(5):753–767.
- Lynn Carlson and Daniel Marcu. 2001. Discourse tagging manual. Technical report, ISI Tech Report ISI-TR-545.
- Laurence Danlos. 2006. Discourse verbs. In *Comparative study of connectives and discourse markers in the framework of a dynamic discourse semantics*, Toulouse, France.
- Catalina Hallett, Donia Scott, and Richard Power. 2007. Composing queries through conceptual authoring. *Computational Linguistics*, page (forthcoming).
- Michael Halliday and James Martin. 1993. *Writing Science: Literary and Discursive Power*. The Falmer Press, London, UK.
- R. Hindmarsh. 1980. *Cambridge English Lexicon*. CUP, Cambridge, UK.
- Alistair Knott, Jon Oberlander, Michael O'Donnell, and Chris Mellish. 2001. Beyond elaboration: The interaction of relations and focus in coherent text. In *Text representation: linguistic and psycholinguistic aspect*, pages 181–196, Amsterdam: Benjamins.
- W. Mann and S. Thompson. 1987. Rhetorical structure theory: a theory of text organization. In L. Polyani, editor, *The structure of discourse*, Norwood, NJ. Ablex.
- J. Oberlander, M. O'Donnell, A. Knott, and C. Mellish. 1998. Conversation in the museum: experiments in dynamic hypermedia with the intelligent labelling explorer. *New Review of Hypermedia and Multimedia*, 4:11–32.
- Daniel S. Pavia. 1998. A survey of applied natural language generation systems. Technical report, ITRI Technical Report, University of Brighton, UK.
- R. Power and D. Scott. 1998. Multilingual authoring using feedback texts. In *Proceedings of the 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics*, pages 1053–1059, Montreal, Canada.
- Leonard Talmy. 1988. Force dynamics in language and cognition. *Cognitive Science*, 12:49–100.
- Joseph M. Williams. 1990. *Style: Towards clarity and grace*. University of Chicago Press, Chicago, USA.