Invited position paper in *Computational Intelligence* 7(4), December 1991, a special issue on natural language generation, pp. 361-362.

"One Tree – One Unit"
a hypothesis for the conceptual sources underlying generation

David D. McDonald
August 1991

14 Brantwood Road, Arlington MA 02174-8004, USA
(617) 646-4124 MCDONALD @ BRANDEIS.EDU

The work-a-day business of generation researchers is how to account for the enormous variation in expression that is available to human speakers, and how to bring that diversity under control for deliberate, contextually appropriate use by machine speakers. A necessary first step in any such account is to establish what the generator starts from—different assumptions can lead to dramatically different accounts. Under contract to a funding agency, we take whatever representation the application we are interfacing to uses and make do. But in doing basic research we would like a firmer basis for our accounts than just engineering adequacy.

To this end we are free to posit what our ideally designed program should supply as a source, though we need to justify these assumptions carefully. Psychology can only help us indirectly, since the representations in the mind are not susceptible to direct experimental control, and their results will always require interpretation. Instead we need to turn to general design principles that appeal to us aesthetically or which make for stronger (i.e. more readily disprovable) hypotheses. If we can assemble a coherent and internally consistent model for mental, ideal-program -representations using these criteria then our work will have a firm basis—even if we cannot yet say that it is "true". Here are three such criteria. I will take them up in reverse order.

- Established explanatory accounts of well understood surface phenomena should be projected as hypotheses about unknown deep phenomena.
- Prefer more efficient mechanisms over less efficient.
- Prefer representations that can provide explanations.

The key to an explanatory model is a reduction in the expressive power of the notation and mechanisms of the theory underlying the model. If the notation makes it literally impossible to state rules or carry out processes that are in fact not observed (e.g. the presence of some grammatical relations and not of others), then we have an explanation of the observations by pointing to the use of that notation.

One example of this is how Tree Adjoining Grammar can provide an explanation for the problem of why texts are grammatical. As an engineering matter we will of course arrange for the texts our generators produce to be grammatical, but the steps we take in doing this do not necessarily constitute an explanation. But since a TAG predefines all the possible minimal phrases (trees) that can go into the surface structure representation of a text and all the ways they can combine, it removes grammaticality from being a contingent property of how a grammarian happens to

write individual rules to a property that follows inescapably from the use of the formalism (assuming that the individual trees are correctly defined).

If we accept the value of TAGs as an account of surface structure—the domain of phenomena that we understand very well—then we can project this back as a candidate hypothesis about the nature and organization of structures at earlier stages. Meteer (this volume) has done this to great benefit in her "Text Structure" level of representation, where adopting basic notions such as matrix phrases (initial trees) and adjunction, along with a carefully selected set of detail-encapsulating semantic types, has led her to an account of expressibility.

If we project the principles of TAGs still further back as a hypothesis about the granularity of the conceptual source for the generator, and add in efficiency criteria (below), I conjecture that, as seen by the generator, information in the speaker's representation of his/her/its mental state is reified into units that correspond to the granularity in a TAG: *one tree–one unit* (see McDonald & Pustejovsky 1985; compare Jackendoff 1990, Shieber & Schabes this volume).

The argument from efficiency considerations runs as follows. It is more efficient to be told what to do than to search through a state space to classify it and from that classification assemble an expression to be realized. Even ignoring the effort of classification, checking for compatible interactions among the alternative realizations of the different terms in a source expression can require polynomial time to sort out, whereas if the information in that expression has been already reified into a unit, the unit can project directly to a single, precomputed realization in one step. (In other words, efficiency can be gained over time by remembering earlier classifications and their composite realizations.)

Since we have hypothesized that surface realizations are organized as trees in a TAG, this means that the natural reifications in the mind of the speaker should chunk information into tree-sized units (i.e. units corresponding to syntactic maximal projections: clause, NP, PP, etc.). To see what this comes to, consider the three texts below—texts which arguably contain much the same information, but which differ markedly, if subtly, in the situations they are suited to.

(1) "*I can only stay until 4:00.*"

(2) "*I have to leave by 4:00.*"

(3) "*I can stay, but only until 4:00.*"

The first two are single clauses; the third is two clauses. This means that even though we might imagine that the third is a reworking of the information in the first, we have to take it as the conjunction of two mental units, otherwise we will have violated the hypothesized design constraint of one unit one surface tree (e.g. one clause). This is an example of how the criteria act to reduce the potential freedom of action we could imagine in a generator (e.g. it can not be permitted to take the source for example one and produce example three), thereby simplifying any account of surface variation by reducing its available degrees of freedom).

Within the source for each of these clauses there must obviously be smaller conceptual units: the internal representation of the speaker ("I"), of the actions stay/leave, of necessity, of specific times and capped intervals, etc. Under the "one tree – one unit" hypothesis these need to be taken as arguments to be bound to positions (lambda variables) in pre-defined relational types, corresponding to grammatical arguments/positions at the surface. Thus the compositional patterns of such smaller units within the maximal units have to be tightly regulated in order to sustain the efficiency of the hypothesized direct mapping from conceptual sources to surface linguistic resources. Note that the hypothesis does not say that the

smaller units do not exist, rather that when the speaker presents units to the generator for realization those units must correspond to full elementary trees, the smaller units will not be presented except in this larger context. (If the ultimate utterance is just, say, a small unit, e.g. “70%”, then the full context must be available in the discourse.)

Having adapted this hypothesis about the organization of conceptual structure, we have to consider whether we are comfortable with all its implications. For example if we take a narrow reading of what should be included in clause trees in a TAG, e.g. restricting them to just obligatory subcategorized arguments, the two single clause examples should be taken as composites: “I can stay” + “only until 4”, “I have to leave” + “by 4”. Are we comfortable with an open formula like “only until 4” as a first-class (i.e. composable) unit in the ontology?

A comparable issue, this time within the generator, is what mechanism we are to posit that will handle the information “suppression” that must have occurred in sentence three. Under this hypothesis, the second, subordinate clause in that sentence must have had as its source the same units as underlies sentence one. The matrix of that unit must be suppressed as redundant, given the context of the main clause. This is a new kind of operation for a generator to perform, as it does not correspond to any of the conjunction reductions of surface grammar, and I would argue that it should be done at a level where the generator is manipulating conceptual units rather than linguistic structure.

A final matter is a promissory note. It is one thing to establish that a given organization of conceptual structure is advantageous to the generator, it is another entirely to show that that organization is actually advantageous to the reasoning system whose conceptual structure this is. I believe that it will be, and I suspect that the benefits will accrue when the reasoner has to navigate multi-million item databases, where the complex organization now required of the reasoner’s knowledge can be exploited to focus its attention according to what is salient at the moment. Natural language texts excel as vehicles for communicating salience, and this competence should be projectable onto the reasoner.

References:

Jackendoff, Ray (1990) *Semantic Structures*, MIT Press.

McDonald, David & James Pustejovsky (1985) "TAG's as a Grammatical Formalism for Generation", proc. Association of Computational Linguistics, Univ. of Chicago, July 8-12 1985, pp. 94-103.

Meteer, Marie (this volume) “Bridging the `Generation Gap’ between Text Planning and Linguistic Realization”.

Sheiber & Schabes (this volume) “Generation and synchronous tree-adjoining grammars”.