Reasons Why We Use Dependency Grammar
Eva Hajičová
Charles University
Prague, Czechoslovakia

1. First, I would like to say why I do care grammar formalisms. The point is not only that I was trained as a student of linguistics and that I always have been interested in theoretical linguistics, but in the present context also, and mainly, that natural language processing systems mostly are too complex to be built, modified, complemented, enriched,..., without a solid theoretical background. As Prof.Nagao puts it, theory is important and valuable for the explanation and understanding; a language processing model should be understandable on the background of a powerful linguistic theory. On the other hand, I would like to stress that if linguistics wants to be useful and to make safe its own perspectives, then it has to be useful for linguistic engineering. This means for me that the theory has to be not only adequate, but also economical and modular. The task of the theory is to offer a relatively complete framework, which never captures all the details in their specific and often exceptional character, but which, as Karen Jensen notes in her point (7), offers a maximal coverage, i.e. which contains means necessary and sufficient for handling all such details as far as they are relevant for the given application field. In this respect, the theoretical framework can be compared to a fisherman's net, which need not be used whole, if this is not necessary for the given pool; some of the meshes may be left unused in the boat or ashore, but in a larger pool they may be useful. The most important point is that the meshes are there, and we know where they are and for what purpose they might be useful.
2. The formalism is not the only important ingredient of an NLP system, and it is not interesting here for its own sake. It is true that the bottleneck of an NLP system is in handling the "dirty" exceptional cases, rather than the cases directly fitting into the main body of this or that theory. As a matter of fact, using any theory, we have to face such intricate but common examples as Kirschner's cases of target language ambiguity (or vagueness) corresponding e.g. to that of English ing-forms, or the long but lexically bound sequences of nouns in terminological noun groups, or a procedure translating lexical items by modifying the productive affixes of international terms of Greek and Latin origin and other "emergency rules" ensuring that at least an approximate (at least partially readable) output will be achieved.
3. All this heavily supports the argument that the theoretical framework should be relatively economical. If two theories, which in the given conditions cannot be systematically compared with respect to their coverage, seem to exhibit a more or less equal degree of adequacy, then the simpler one will be preferred. With this framework it is more probable that there would be space enough for capturing all the necessary "dirty" details illustrated in Section 2.

Dependency grammar meets this condition, since its trees (or even the more complex representations required for a treatment of coordinated structures) are much simpler than any kind of phrase-structure based representations: it is possible to use here complex (although strictly limited) node labels and thus distinguish between the syntactically free lexical occurrences and function morphemes (the latter need not have nodes of their own); moreover, the representations can be conceived of without nonterminal symbols (if the kinds of dependency relation, similar e.g. to theta roles, cases, and valency, are denoted by the labels of the edges).
4. Besides the advantage of an economical description, a dependency based procedure may derive from the lexical data all the information necessary for the assignment of theta--roles (valency slots) and other complementations by a head. This property is extremely valuable for the formulation of a parser; once the verb is identified and looked-up for in the dictionary, it 'predicts' many pieces of information necessary for the identification of the functions of nominal (including prepositional) complexes (as complementations of the given verb). As far as our own experience with the build-up of NLP systems goes, this holds true not only when working with highly inflectional languages (the word order of which is, consequently, not grammatically determined), but also e.g. for a parser of English.

The information on the 'theta-roles' belongs to one dimension of the tree, namely the vertical one. The second dimension, horizontal, is left free to denote the topic--focus articulation and the deep word order, which contributes to the determination of operator scopes in semantic interpretation.

Coordination (as well as apposition) constitutes a third dimension, similarly as with other approaches; rules have been formulated (by Plátek, Sgall, Petkevič) which handle hierarchical as well as sequential coordinated structures in the linearized sentence representations.
5. The arguments brought forward against dependency as the basis of an integrated linguistic description are not so weighty as they may seem at first blush: anaphoric relations and the relative closeness of adjuncts or other modifications can best be solved on the basis of an account of topic and focus, and instead of "double dependency" with predicative complements we use a broader notion of manner adverbial. Our framework differs from Gaifman's in imposing no specific limits on the sequences of non-terminals used in the derivations of representations (Borota).

The framework has been used with advantage as the basis of the natural language processing systems built in our group, be it parsers for machine translation systems (English-to-Czech, Czech-to-Russian) or for a system modelling natural language comprehension (TIBAQ).

451