Our comments derive from the experience of designing and implementing proposals for the translation of compounds within a multilingual MT system, which aims at minimising transfer, Eurotra. We concentrate here on noun-noun compounds.

In the ideal case, nothing would have to be added to the lexicon to handle compounds, since all relevant information would be there for the treatment of syntactic structure. We would claim that this situation is approached as far as argument structure is concerned, i.e. the system of argument structure for nouns which has been developed in Eurotra is also adequate when these nouns occur as heads of compounds. This helps when translation as a compound is, unusually, not permitted (e.g. German Menschenmenge — English crowd of people, not *people crowd). Crowd is one of a number of words denoting sets that take an arg1 (classifying argument) in Eurotra terminology, and which are not generally available to function as heads in English compounds. It also helps when translation to a non-compound is regularly impossible, as is often the case in Germanic-to-Romance translation, e.g. English integrated circuit manufacturers — Spanish los fabricantes de circuitos integrados.

Information about semantic features of arguments is also essential to disambiguation in analysis, as compounds provide fewer overt clues to grammatical relations than syntax does. For instance, comparing the English compounds project funding and government funding, in the absence of textual context there may be no way to decide on the grammatical role associated with the non-heads. However we can develop preferences for role assignment by assuming that humans are more likely to activate subject slots and inanimate entities object slots. The relevant semantic features must therefore be coded in the lexicon.

In addition, the distinction between predicate and function nominalizations of verbs (or process vs. result nominals) is also crucial in determining argument structure. For example, in the compound computer analysis a process interpretation of the head implies an object role for the non-head, whereas a result interpretation arguably implies a modifying non-head projecting some kind of instrumental relation. In general, little progress can be made in translating compound nouns without a fairly full theory of nominal argument structure.

Languages with very different types of compounding mechanism will have very similar lexica as far as argument structure is concerned, since (a) arguments can be realised either inside or outside compounds, and (b) argument structures are very similar (NOT identical) across languages.

One area where extra information is needed, however, is that of relational adjectives, i.e. denominal adjectives which have the same semantics as their underlying nouns. Languages differ, often idiosyncratically, as to whether a relational adjective is a possible, preferred or obligatory replacement for a noun as non-head in a compound. E.g. English car production — Spanish producción automovilística (the latter with a relational adjective). Clearly, when analysing a compound with a relational adjective, it is the semantics of the underlying noun which is crucial in establishing the relation to the head. E.g. in the case of the compound producción automovilística, it is the semantic features of automóvil we need in order to establish whether the subcategorisation requirements specified for the internal argument of producción are filled. Information about relational adjectives (e.g. that solar is the adjective from sun) is not given consistently in printed dictionaries that have been exploited for NLP purposes.

As far as modifying non-heads in compound nouns are concerned (e.g. software problems), we have made no attempt to classify the types of modifying relation, on the grounds that so much non-linguistic knowledge is involved in their interpretation. Semantic information in the lexicon may be exploited to some extent, but we do not see this as primarily a linguistic problem.