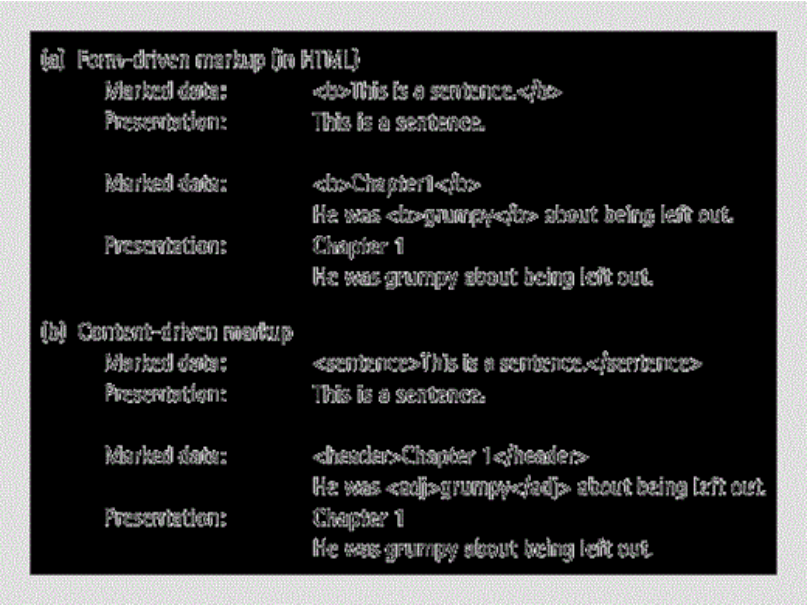
no way to quickly change the presentation form of just one kind of element (e.g. so that the adjective *grumpy* is italicized instead of bolded), nor is there any way to search for structurally different elements, e.g. headers only. Imagine having to comb through a 10,000-entry lexicon, entry by entry, just to make a minor format adjustment that could instead be accomplished with a single click or a couple of keystrokes.

p. 95 A more relevant example is that of a dictionary project. A single well-structured lexical database can give rise to a number of different presentation formats: on paper, on-line, with linked sound files and images, just as a simple wordlist, with reverse-language indices, and so forth. These many presentations of dictionaries are possible if contents of the elements (lexemes, glosses, definitions, example sentences, etc.) are described with, for example, backslash[9] field markers, rather than just with formatting information. Similarly, texts have logical structures: they usually include an orthographic representation of the spoken words, a morphemic parse, a morpheme gloss, and a free gloss as shown in the example of a Toolbox text fragment in Fig. 4.2, where each line is marked by a code (e.g. \tx = text; \mr = morphemes; \mg = morphemic glosses; \fg = free gloss). Describing or marking the content of these structures in our data allows various presentation formats to be generated for different audiences.

Figure 4.1.



Form-driven versus content-driven markup.

Structuring linguistic data by content allows it to be arranged and rearranged in new ways, providing visualization of relationships within the data. Such visualization allows new generalizations to be made, and also helps ensure consistency in the data. For example, being able to sort a lexicon on various fields puts
p. 96 similar types of information ↳ next to each other which can also be useful for creation of topical lists where the lexicon is sorted by semantic field rather than just alphabetically. Mistakes, such as duplicate entries or typos, are more easily located when such sorting is carried out.