There are three major difficulties with this paper. First, the case is not really made that this specific application area (e-learning of algebra) should *a priori* benefit from a natural language interaction. Second, this is too early in the lifecycle of the project to present a compelling or interesting story. Third, the system itself does not seem to reflect any breakthroughs or even substantial incremental improvements; rather, it is a cobbling together of existing resources. Despite issues 1 & 3, it might be possible to write an interesting and useful paper after the system is actually deployed and tested.

The paper begins with a review of adaptation in e-learning. The review is okay but not sufficiently comprehensive to publish as a review paper. Throughout this section, there is an untested assumption that a learner's preferred style is the one that should be adopted. Preference may be important, but it is not necessarily best in all circumstances. Another presumption is that the *least* cognitive load is the *best*. Clearly, too great a cognitive load can impede learning. But it is not obvious that minimal cognitive load insures maximum learning. In fact, this would seem to be at odds with "levels of processing" notions of Bloom and others. "Instructor-based" and "Learner-based" may be one important dimension of classification of systems, but it is not the only one.

The review of dialogue systems is similarly okay but neither extensive nor particularly insightful or original. Again, there are some questionable assumptions here that are not really discussed. Example:

Learner: "I ... I think ... understand improper fractions now, but I ...show me again how to convert them to decimal."

[According to the authors] "This contains several ungrammatical constructions as well as words superfluous to meaning."

Perhaps these words and bad constructions are not at all superfluous to what is going on. "I ... I think..." and the break-offs and switches may be important pedagogical cues to what is going on in the learner's head. To simply ignore this and focus on showing the learner how to convert improper fractions to decimal may be exactly the wrong thing to do.

It is very good that the authors intend to use iterative development. However, in terms of a journal publication, I believe that what should be published is the end result, not the beginning speculations along with a hub and spoke "architecture diagram." This is especially true because there is really nothing new here in terms of the technology. After several iterations and details are worked out and incorporated into the system and some results are forthcoming; e.g., showing that a spoken interaction version of this system produces superior learning compared with a typed I/O version (which I doubt but therefore would be all the more interested in reading about). Even more interesting, perhaps the authors can show some qualitative differences in the *kinds* of learning that takes place in a typed versus auditory version. Also of interest would be specific examples of difficulties in learning that trace back to specific technology issues; e.g., in this case, a misrecognized word resulted in this problem, while in this case, a failure to deal appropriately with context resulted in this problem, etc. Real examples of what happens when students actually do interact with the system will be a hundred times more interesting than speculations about what could happen.