

LETTER TO THE EDITOR

IS MT LINGUISTICS?

(Reply to Letter to the Editor, Volume 12, number 1)

No! It is properly an *application* of linguistics, and therein lies the critical point that Kimmo Kettunen seems to miss. Other than that, I have only a few problems with what Mr. Kettunen says.

MT is an application of linguistics; as such, MT system developers are, first and foremost, required to produce an operational system. It is doubtful that anyone who has not produced a real-world application can appreciate the sacrifices of theoretical niceties that entails. More to the point, it is unlikely that any such individual can appreciate the degree to which linguistic theory fails to account for linguistic fact. Thus, while it is incontestably the case that theoretical advances are necessary, it is nevertheless true that MT system developers have less time to engage in theorizing than they might like. Theoretical linguists, for their part, seldom if ever engage in application development. MT system developers generally circumvent this impasse by applying the [parts of] theories that admit application, and either developing new theory on-the-fly or, more often, constructing necessarily ad hoc solutions to the problems they face. Such is the fate of developers everywhere, who applaud the fabrication of more nearly accurate, better specified, and more applicable theories.

My second point relates to the open-endedness of language. If languages are open-ended to the extent that they “cannot be easily – if at all – described with deterministic methods, i.e., using computers and algorithms”, then one can rest assured that linguistic theory will be stymied for exactly the same reasons. After all, a theory, properly speaking, is necessarily testable (by definition).

My third point relates to the way things are. The statement “So, to be cost-effective, an MT system has to produce output that is good enough to need little *or no*

human post-editing” is seriously flawed. As a matter of record, it has been demonstrated, for several MT systems, that cost-effectiveness has been achieved even though a significant amount of post-editing is performed. More astounding, the discussant casually assumes that one should expect MT systems to be usable with “. . . no human post-editing”, all the while agreeing that “everyone with some experience in translation knows” that “it is not unusual for [translation] products to be revised many times.” This is the standard trap that I worked so diligently to point out in my paper: where organizations employ post-editing [of human translations] now, they will continue to employ it [for machine translations] in the future. In other words, a second opinion *will be rendered*, in consonance with standard translation practice today. Why should things be otherwise, in principle or in practice?

In closing, I note the discussant’s twice-stated claim that “MT . . . should pay more attention to its linguistic premises.” He is apparently unfamiliar with the literature regarding recent MT systems, so I refer him to the other papers in issues 1 through 3 of Volume 11 of this journal. The point is, modern MT development projects most assuredly are attending to linguistic premises. It is sad, but true, that these are inadequate, as anyone attempting linguistic applications knows. We all look forward to testing newer and better theories. All too often, “theories” as stated are not testable – or, if testable, are falsified – so we all eagerly await the maturation of linguistics as a science. Perhaps the discussant will join us in contributing to this process, whether as a theoretician or an experimentalist.

Jonathan Slocum
MCC
Austin, TX 78759