



Comments on “Towards increasing speech recognition error rates” by H. Boullard, H. Hermansky, and N. Morgan

Sadaoki Furui^{a,b}

^a *NTT Human Interface Laboratories, 3-9-11, Midori-cho, Musashino-shi, Tokyo, 180 Japan*

^b *Tokyo Institute of Technology, 2-12-1, Ookayama, Meguro-ku, Tokyo, 152 Japan*

Received 26 January 1996

As the authors state, the main goal of this paper is “to promote the notion of ‘deviant’ research paradigms”. I agree with the authors that “new approaches will most likely be necessary” for solving the remaining difficult problems in speech recognition, and I am sure that it is necessary to “consider more radical deviations from current wisdom in order to improve our overall understanding” of speech. It is true that “there is much more unknown than known”, and that “ASR is still a field in which many of basic problems are still barely touched, and in which new ideas from newcomers will be critical”.

However, the way this paper is written is somewhat misleading. It is quite obvious that “increasing speech recognition error rates” is NOT a necessary NOR a sufficient condition for a new innovation. I have not seen any technical papers on major speech recognition technologies that have reported increasing error rates.

On the contrary, it is crucial to verify the signifi-

cance of a new idea by reducing the error rates, or at least to indicate some promise for ultimately improving performance in the end-applications. Even if the error rate can not be significantly reduced for the time being, it should not be increased. Every technical paper should at least satisfy the following three conditions indicated by the authors:

1. solid theoretical or empirical motivations,
2. sound methodology (so that something can be learned from the results), and
3. deep understanding of state-of-the-art systems and of the specificity of the new approach.

I hope the readers will correctly understand what the authors intended to say and that this paper will stimulate ambitious undertakings in new speech recognition and understanding technology. Misunderstanding of the paper’s message could lead to an acceleration in the publication of irresponsible, half-baked papers. I trust this will not happen.