



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

Bishnu S. Atal

Speech Research Department, Room 2D-535, AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974, USA

Received 29 January 1996

The goal of this paper is to start a debate to encourage more creativity in ASR research. The paper makes a case for more tolerance in evaluating new approaches, because some of them might lead to higher error rates during the initial phase of research. They suggest that an increase in error rates might be acceptable, provided such research is based on sound methodology motivated by solid theoretical or empirical reasoning.

I agree with the premise stated in the paper that the narrow focus of research towards decreasing error rates can lead to local optimization and can miss promising directions. However, the authors fail to outline alternative strategies that ASR research community could follow to make scientific advances. Tracking recognition performance on speech databases with utterances recorded from many speakers in a variety of environments has been important in achieving the progress so far. On the negative side, this approach has sometimes led to solutions that are narrow and do not generalize to other situations.

The past two decades of ASR research has produced major advances leading to deployment of the ASR technology for increasingly difficult tasks and we are becoming more ambitious in seeking solu-

tions to bigger problems. Human communication by voice appears to be so simple that we often forget that the process our brain uses to decode speech could be very complex demanding scientific knowledge as well as processing capabilities that we do not have at present. As we move ahead to the next stage in the development of ASR technology, it is important to follow proper evaluation procedures that will help us to continue to advance vigorously. The paper provides very little guidance in this area. The paper however has raised a very important issue that merits continuing discussion in the ASR research community.

A major part of the paper has been devoted to presentation of research directions (new approaches) that the authors feel are promising but are a deviation from the mainstream research. These research directions are not new and have been discussed in publications. The paper presents outlines of many new approaches, but fails to include any results that could shed some light on how useful these new approaches could be. The paper would be much more useful if some results could be included with each of the new approaches pointing to its value in providing solutions to some of the difficult problems facing ASR.