

Preface

This book came out of approximately ten years of continuing research at Yamagata University.

With the emergence of numerous algorithms for a variety of speech processing applications, such as coding, enhancement, and synthesis, a variety of distortion can now be observed. These disturbances degrade the speech quality in an unexpected manner. For example, some recent speech coding algorithms use the human auditory masking properties to “hide” the coding distortion to be unperceivable. Simple objective measures, such as the signal-to-noise ratio may be low in some samples, but may show surprisingly high subjective quality. This is why we need a subjective quality evaluation method that is stable, gives reproducible results, and is as inexpensive to administer as possible. The cost of administration would be, in this case, the length of the tests, the number of testers needed for stable results, and the stress on the testers.

Since we were initially working on speech enhancement algorithms, we needed such quality evaluation methods. The key to enhancement algorithms was to maintain the intelligibility of the enhanced speech, one of the aspects of speech quality. Accordingly, out of our own need, we attempted to define an efficient method for intelligibility measurement.

The Diagnostic Rhyme Test (DRT) seemed a logical choice for this task. Thus, we defined a new set of word-pair list for use as the Japanese DRT. The new Japanese DRT was tested, and proven to give stable, reproducible results even with naive users, which we only have available, and is thus a requirement. We applied this test to various speech processing applications we were working on, with the binaural localized speech application being the largest effort at the time. However, we realized that even with the efficient new testing method, it was not practical to do a full-scale subjective evaluation for each new condition that needs to be tested. Therefore, we investigated on the possibility of using objective measures, which we can compute on a computer, to estimate the subjective DRT intelligibility scores. We attempted two different approaches. We attempted to calculate objective measures, and map these measures to subjective scores using pre-trained mapping functions. We also attempted to use speech recognizers to mimic the

human recognition process of the DRT test. It turns out that both approaches give surprisingly good matches.

Many students contributed to this book. Ryo Izumi was involved in the initial formulation of the Japanese Diagnostic Rhyme Test (DRT). Masaya Fujimori and Eiichi Mikokai underwent initial psychoacoustic evaluations with the DRT. Rui Kaga was one of the first to attempt the estimation of DRT intelligibility using an objective measure, the PESQ-derived Mean Opinion Scores (MOS). Yuichiro Kitashima, Takahito Chiba, Naoki Saito, Naoya Anazawa and Yosuke Kobayashi applied DRT for the evaluation of localized speech intelligibility. Yusuke Takano attempted to estimate intelligibility using speech recognition systems. Futari Kano and Tomohiro Terada applied DRT to evaluate intelligibility of speech generated with parametric speakers. Tomohiro Komiyama applied DRT to the evaluation of masking efficiency of speech-privacy protection systems. Most of their work is described in this book.

The students and I received support from numerous foundations and organizations. The Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) provided vital financial support through their Grant-in-Aid program (number 20500151). The Yamagata University Faculty of Engineering Alumni Association, the NEC C&C Foundation, the Telecommunication Advancement Foundation, and the Ojima Foundation provided travel expenses to international conferences. The Research Foundation for the Electrotechnology of Chubu supported the publication of the results. Numerous other institutions provided financial support.

I would also like to express my gratitude to Professor Kiyoshi Nakagawa for his continuous support. I also thank the members of the Nakagawa Laboratory for their discussions, as well as long tiring psychoacoustic evaluation sessions. Without their support, the foundations of this book would not have been possible.

And last but not least, I would like to thank Yukari for her constant encouragement. Without her, I probably would not have survived the brutal world of academia to make this book a reality.

As a final note, the work this book describes is nowhere near completion, and is still ongoing. Our lab still has numerous projects running as of now. I also hope that many others will also follow in this interesting and important topic.

Kazuhiro Kondo

Subjective Quality Measurement of Speech
Its Evaluation, Estimation and Applications

Kondo, K.

2012, XIV, 154 p., Hardcover

ISBN: 978-3-642-27505-0