Towards a Detailed Annotation Scheme for Clustered Disfluencies

Peter A. Heeman[1] J. Scott Yaruss[2] Andy McMillin[3]
[1]*Center for Spoken Language Understanding, OGI School of Science & Engineering*
Oregon Health & Science University, 20000 NW Walker., Beaverton OR 97210
[2]*Stuttering Center of Western Pennsylvania, Dept. of Communication Sciences and Disorders*
University of Pittsburg, 4033 Forbes Tower, Pittsburgh PA 15260
[3]*Hearing and Speech Institute, 3515 SW U.S. Veterans Hospital Road, Portland OR 97239*
heeman@cslu.ogi.edu

In determining the disfluencies in a speech sample, the emphasis is on the number of disfluencies. Little attention has been paid to the structure of them. Even the approach of Systematic Disfluency Analysis (Campbell, Hill, and Driscoll, 1991), which is probably one of the most detailed analyses, does not precisely indicate which words are involved in the disfluency. Although it might be trivial to automatically determine for most disfluencies, this is not the case for *clustered* or *multi-component* disfluencies, which are very common in stuttered speech (Logan and LaSalle, 1999). Having a detailed annotation of disfluencies would be useful for a number of reasons, especially for building a computer system to automate disfluency counts.

In our earlier work in modeling disfluencies of non-stutterers, we proposed a way of annotating the clustered disfluencies that non-stutterers typically make (Heeman, 1997). We have applied this scheme to annotating clustered disfluencies in stuttered speech. A key aspect of the scheme is that it tracks where the speaker is in the utterance and when the speaker *backtracks* through a repetition or revision. The spot where a speaker backtracks is called an interruption point. After each interruption point, we start a new line indented so that the words from before the interruption and the words from after the interruption point that correspond to each other line up. This is illustrated in the first panel of Figure 1 for the utterance "we can go t- t- go to um can can go to the store." After the words are lined up, we can mark the *reparandum* for each repair by determining how far back we need to start the next line. This is illustrated in the second panel. Finally, we label the word above the line with the corresponding repair index, which is shown in the third panel. For the example, this gives a word-level annotation of "we can3 go2 t-1 t-2 go3 t-3 can4 can go to the store," where the indices indicate which backtracking point (or reparandum) the word belongs to. As this example illustrates, the backtracking points can incorporate both parts of a word and a sound repetition. However, from the annotations, we can automatically compute the sound, word, and phrase repetitions that make up the clustered disfluency.

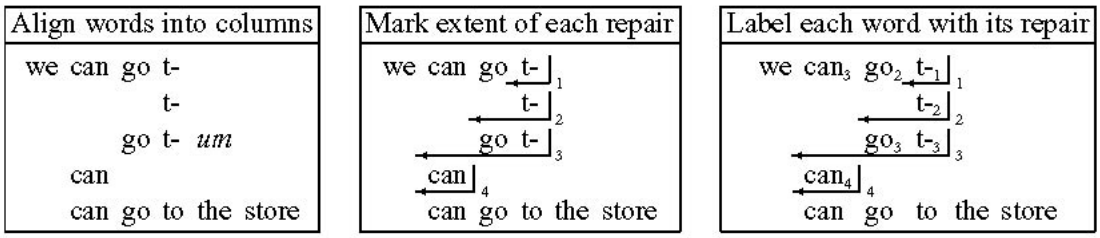

Figure 1: Determining the reparandum of clustered disfluencies

In the full paper, we will discuss how interjections, starter words, blocks, prolongations and revisions can be incorporated into the scheme.

References:

Campbell, J., D. Hill, and M. Driscoll. 1991. Systematic Disfluency Analysis: Using SDA to determine stuttering severity. In *Annual Convention of the American Speech-Language-Hearing Association*.

Heeman, P. 1997. Speech repairs, intonational boundaries and discourse markers: Modeling speakers' utterances in spoken dialog. Doctoral dissertation, Dept. of Computer Science, University of Rochester.

Logan, K. and L. LaSalle. 1999. Grammatical characteristics of children's conversational utterances that contain disfluency clusters. *Journal of Speech, Language, and Hearing Research*, 42:80–91, February.