HOW FAR DO SPEAKERS BACK UP IN REPAIRS? A QUANTITATIVE MODEL

Elizabeth Shriberg *Andreas Stolcke*

SRI International, Menlo Park, CA 94025
http://www.speech.sri.com

ABSTRACT

Speakers frequently retrace one or more words when continuing after a break in fluency. Syntactic principles constrain the points from which speakers retrace; however syntactic principles do not provide predictions about the relative usage of different allowable retrace points. Such predictions are useful for automatic processing of repairs in speech technology, particularly if they use information readily available to a speech recognizer. We propose a quantitative model that predicts the overall distribution of retrace lengths in a large corpus of spontaneous speech, based only on word position. The model has two components: (1) a constant, position-independent probability for extending a retrace by one more word; and (2) a position-dependent probability to "skip" to the beginning of the sentence. Results have implications for modeling repairs in speech applications and constrain explanatory models in psycholinguistics.

1. INTRODUCTION

When speakers resume after a disfluency, they often retrace back one or more words before continuing, producing simple repetitions as well as repeated words in repairs. A question important to modeling repairs in both psycholinguistics and in speech technology is: *when speakers retrace, what predicts how far back they go?* Previous accounts of retracing in linguistics and related fields have illuminated syntactic constraints on retracing—namely that speakers retrace to points that correspond to the onsets of syntactic phrase boundaries, and which can produce a well-formed syntactic coordination between the original utterance and the continuation [3]. The syntactic phrasing accounts match native speaker judgments about what constitutes a "bad" retrace point. However, they do not predict which of many possible remaining retrace points are chosen. In English as well as other right-branching languages, many locations in an utterance correspond to the onsets of constituents, and a large subset of these correspond to points that produce a well-formed coordination—including retracing back to the sentence onset. For example in the following case, all possible previous words constitute viable retrace points:

At the end of the road – ((((at) the) end) of) the block

Our goal in this study was to explore whether overall corpus statistics on the length of retracings could be predicted using information readily available to a speech recognizer. We focus on word position for this purpose, since retracing is inherently constrained by this factor and since information on words can be easily modeled in speech systems.

2. METHOD

2.1. Data

Data consisted of transcripts from the Switchboard corpus of human-human dialog over the telephone [2], distributed by the Linguistics Data Consortium (LDC). We used a subset of 1115 conversations (roughly 1.4M words, 350 different speakers) that had been marked for sentence boundaries and for disfluencies by the LDC, as described in [5]. Word correspondences within disfluencies were not marked, but retraced words could be detected automatically with high accuracy by aligning reparandum and repair regions via dynamic programming. We recorded all instances of simple repeats, as well as cases of retracing before changed words in other repairs. This resulted in a set of 30,524 disfluencies containing one or more retraced words.

2.2. Measures

We characterize retracing in terms of two measures, the number of retraced words (retracing length, k) and the position relative to the start of the utterance at which the retraced word sequence ends, (retracing position, m) as illustrated in Figure 1.