

ASR Future: A Consumer's Wish List

Julia Hirschberg
Columbia University
8 October 2003

An ASR System....

- With acoustic and language models and even confidence measures that adapt on the fly based on context, e.g.
 - For voicemail: whose ext. did it come from?
 - For a call center: customer's records
 - Is this a non-native speaker?*
 - What is likely to be the topic?*
 - For a news show: what's in today's print news?

- That tells me who is talking
 - Speaker identification
 - This is Peter Jennings.*
 - Age, gender, speaking rate, native language
 - This is a child.*
- That gives me information about the speaker's current state
 - More than just negative/angry or not
 - U: Um Windows 98. (degree of confidence)*
- That segments speech into utterances, paratones, and speaker turns as well as words

- That identifies disfluencies explicitly rather than treating them as noise

I'd like to tr- Oh, so this is from my checking. I'd...

- That tells me which words are emphasized and which are not

U: I DON'T want to go to Boston. → S error?

(in a mtg) WE CANNOT DO THIS. → disagreement?

- That tells me which intonational contours people are using

S: Did you say you want to go to Boston?

U: Boston./Boston?