

# Entwicklung einer silbenprominenzbasierten Prosodievorhersage für ein Unit-Selection-Sprachsynthesesystem

Andreas Windmann

26.05.2010

Das Prinzip der Unit-Selection-Sprachsynthese, bei dem neue Äußerungen durch die Rekombination von Sprachaufnahmen aus großen Korpora synthetisiert werden, ermöglicht die Erzeugung von qualitativ hochwertiger synthetischer Sprache. Die Modellierung einer angemessenen Prosodie stellt jedoch nach wie vor ein nichttriviales Problem dar. Bei der Unit-Selection-Synthese liegt die spezielle Herausforderung darin, dass nach Möglichkeit die inhärente Variation der Sprachdaten im Synthesekorpus ausgenutzt und nicht auf Signalmanipulationstechniken zur expliziten Modellierung einzelner prosodischer Parameter zurückgegriffen werden soll. Im Vortrag wird ein Ansatz vorgestellt, bei dem Prosodie systemintern in Form von abstrakten Prominenzwerten auf Silbenebene modelliert wird. Prominenz wird dabei als die graduell wahrgenommene Stärke einer Silbe relativ zu ihrer Umgebung verstanden. In verschiedenen Arbeiten wurde für eine Reihe von Sprachen ein stabiler Zusammenhang zwischen der wahrgenommenen Prominenz einer Silbe und bestimmten akustischen Eigenschaften des Sprachsignals nachgewiesen. Ausgehend von diesen Erkenntnissen wurde in der aktuellen Arbeit ein automatisches Annotationswerkzeug verwendet, das auf der Grundlage entsprechender akustischer Analysen Prominenzwerte für alle Silben im Synthesekorpus errechnet. Ergebnisse und Evaluierung der automatischen Prominenzannotation werden im ersten Teil des Vortrags vorgestellt. Im zweiten Teil wird ein Ausblick auf die weiteren Implementierungsschritte und die abschließende Evaluation des Systems gegeben.