

Csapó Tamás Gábor
V. Jahrgang

Steigerung der Natürlichkeit synthetisierter Sprache

TDK Vorlesung
16. November 2007

Sprachlicher Konsulent: Clemens Prinz
Fachliche Konsulenten: Dr. Németh Géza, Dr. Fék Márk



Inhalt

- Ziel
- Theoretischer Hintergrund
- Experiment
- Test
- Weiterentwicklung

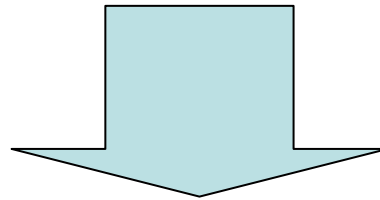
Ziel der Arbeit

Menschliche Sprache:

- derselbe Satz mehrmals ausgesprochen immer anders
- Variabilität in Melodie, Betonung, Rhythmus

Vorlesautomat:

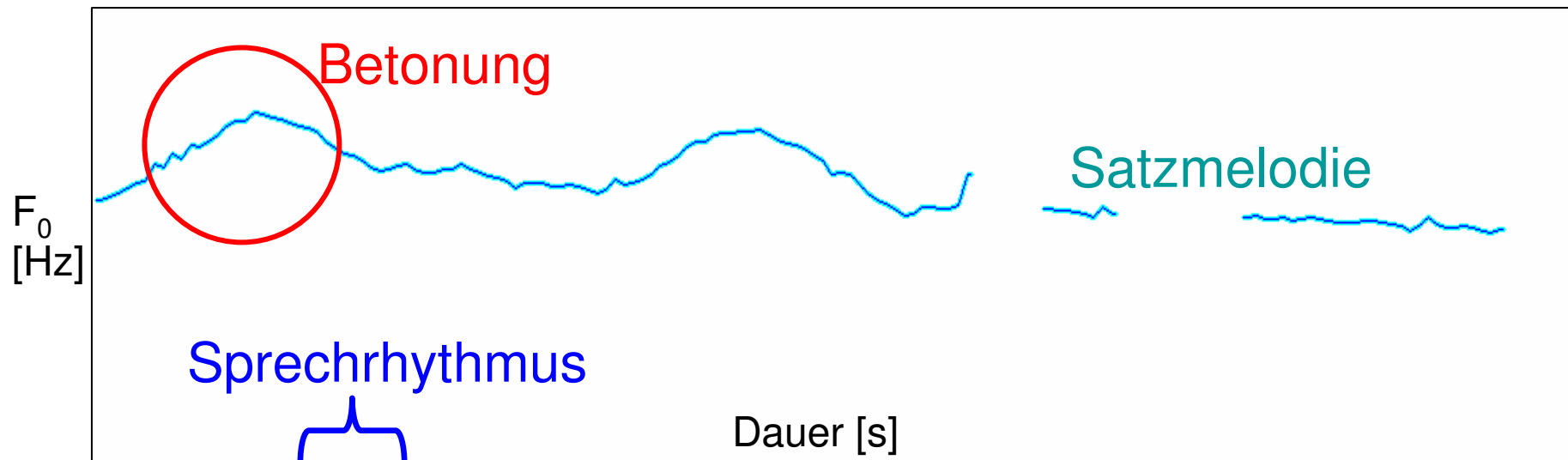
- derselbe Satz klingt immer gleich
- keine Variabilität



Ziel: abwechslungsreiche Sprache bei Vorlesautomaten

Theoretischer Hintergrund (1): Prosodie

- 3 Komponenten:

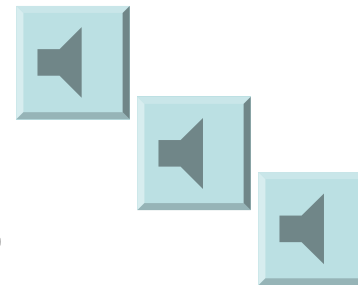


wichtigste Komponente: Satzmelodie
=> ich beschäftigte mich nur mit Melodie



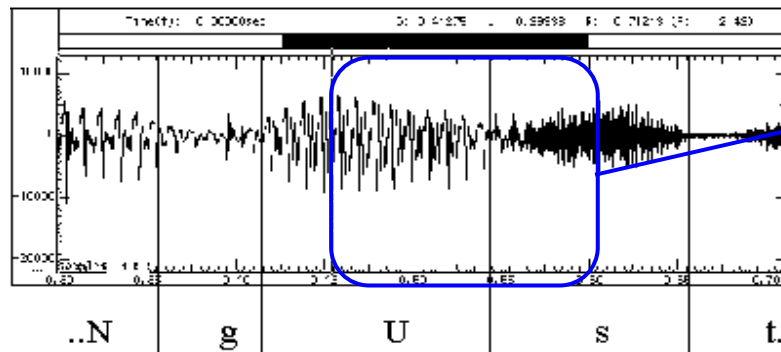
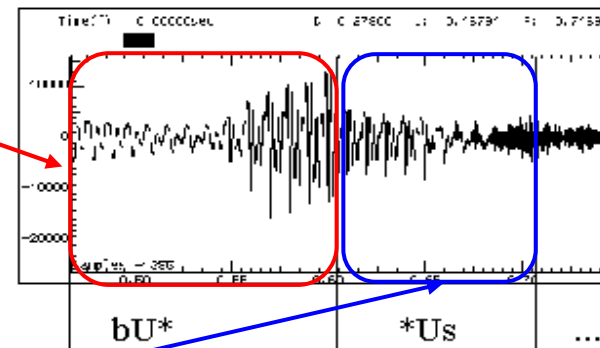
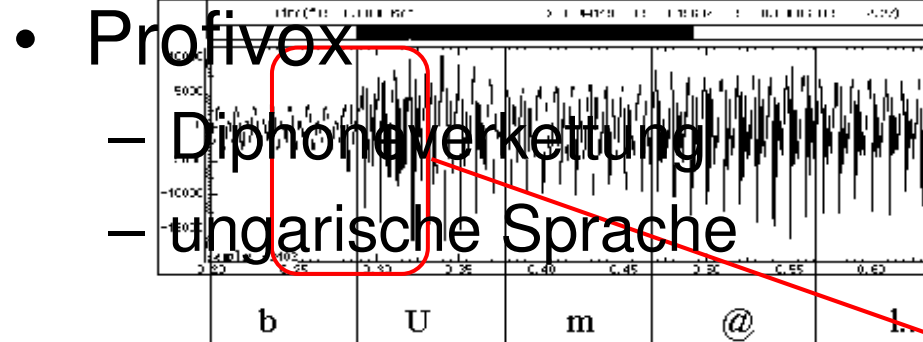
Theoretischer Hintergrund (2): Sprachsynthese

- die Erzeugung von gesprochener Sprache durch einen Computer
- aus natürlicher Sprache ausgeschnittene Teile verbunden
- Signalverarbeitung (z.B. Melodie erhöhen/senken)
- 3 Generationen:
 - Formantsynthese
 - Elementverkettung
 - Unit-Selection (Elementauswahl)



Theoretischer Hintergrund (3): Sprachsynthese: Elementverkettung

- Diphone = Lautübergang (b->U, U->s)

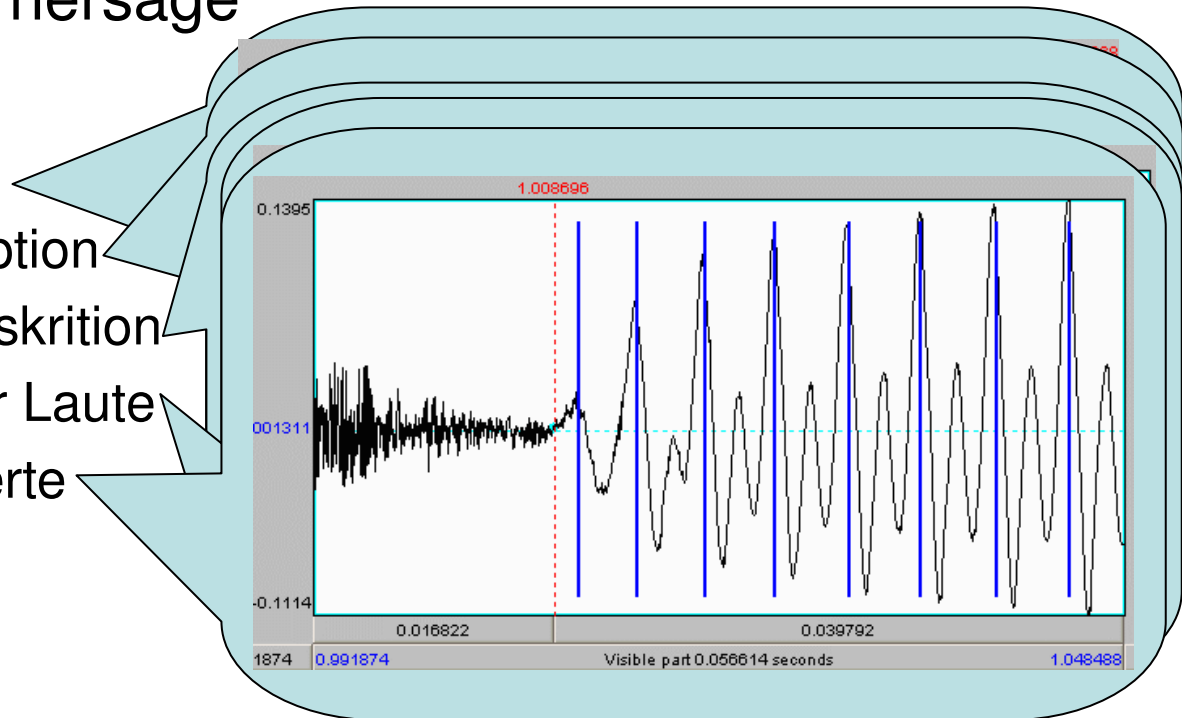


Theoretischer Hintergrund (4): Prosodiemodelle

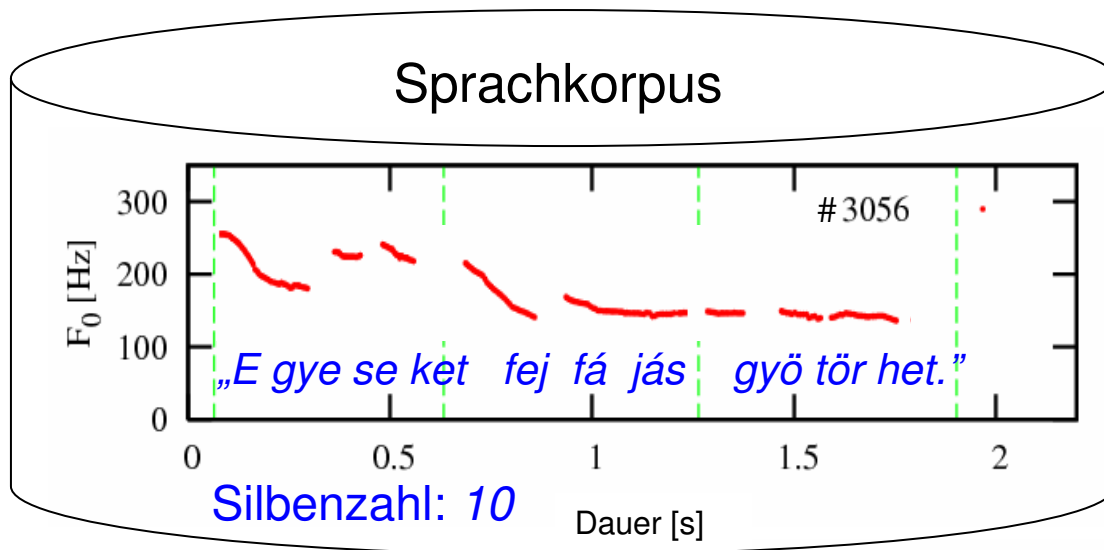
- natürliche Prosodie durch Computer modellieren
 - regelbasierte Modelle
 - Regeln, die vom Menschen definiert sind
 - Vorteil: immer die gleiche Qualität
 - Superpositionsmodelle
 - verschiedene Teile addieren (=superponieren):
 - Satz
 - Wort
 - Laut
 - Profivox
 - regelbasiertes Superpositionsmodell

Experiment (1): Sprachkorpus

- 200 ungarische Sätze
- Thema: Wettervorhersage
- Inhalt:
 - Tonfile
 - textliche Transkription
 - phonetische Transkription
 - Beginn / Ende der Laute
 - Grundfrequenzwerte



Experiment (2): Meine Methode: Melodiekopierung



Referenz
(#3056)



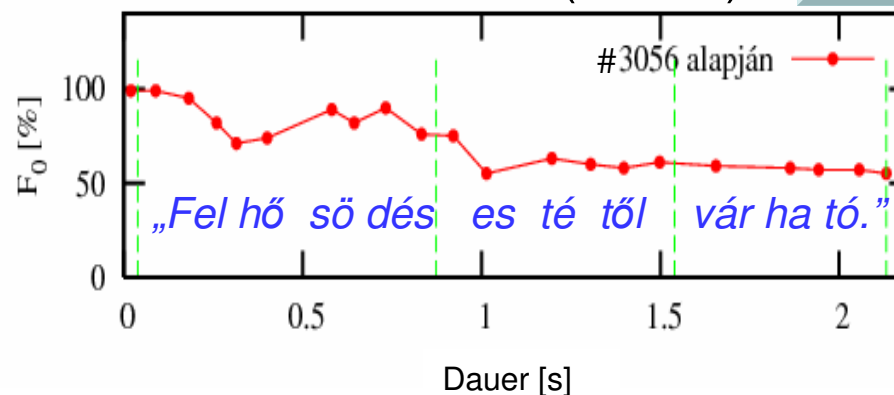
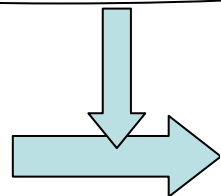
verändert
(#3373)



Silbenzahl: 10

„Felhősödés estétől várható.“

Profivox regelbasiert
(#3373)

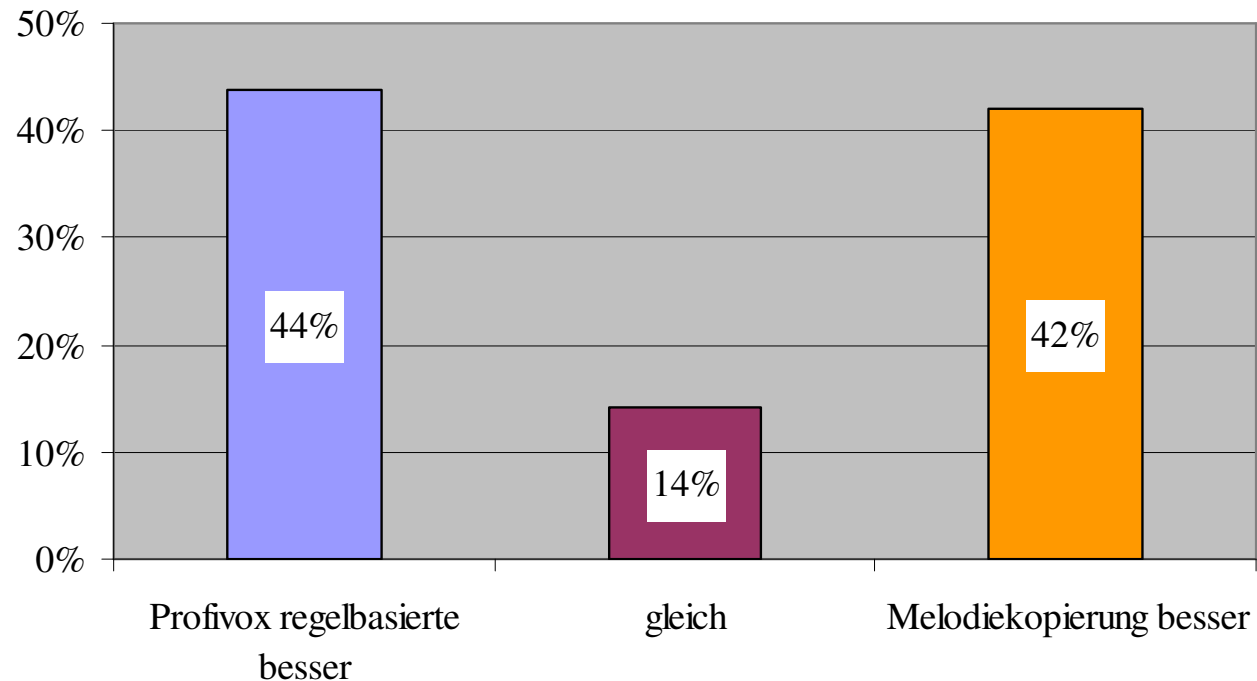


Experiment (3): Testbedingungen

- 28 Satzpaare: in jenem 1 Satz in 2 Melodie-Varianten (alte vs. neue Methode)
- 1. besser / 2. besser / kein Unterschied
- per Internet
- 23 Testpersonen
- Ziel des Tests
 - ob die verschiedenen Melodievarianten akzeptabel sind
 - ob der Unterschied zwischen ihnen bemerkbar ist

Experiment (4): Ergebnisse des Tests

- bei fast der Hälfte aller Sätze wurde meine Methode (Melodiekopierung) präferiert



Weiterentwicklung

- die Eigenschaften der Lautdauer zu untersuchen (hier: nur Satzmelodie)
- andere Sprachen (hier: nur Ungarisch)
- größerer Sprachkorpus (hier: nur 200 Sätze)
- Verwendung:
 - SMS-, E-Mail-, Buchvorlesautomat
 - längerer Text (ohne Monotonie)

Zusammenfassung

- Ziel der Arbeit
- Komponenten der Prosodie
- Sprachsynthese durch den Computer
- Prosodiemodelle
- Neue Methode: Melodiekopierung
- Test, Ergebnisse
- Weiterentwicklung, Anwendung