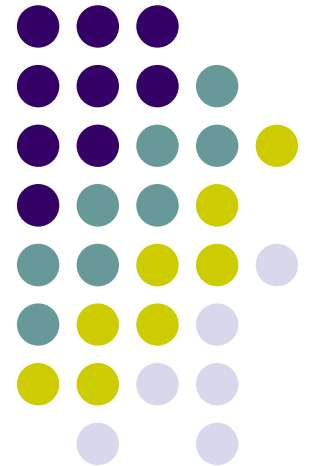


Genderspezifische Spracherkennung

Ein Vortrag von Meral Akyol im Rahmen des Seminars
Sprecherunterschiede bei Prof. Harrington
VIU 18.10.2007



Genderspezifische Spracherkennung

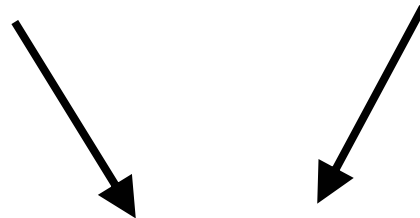
Grundlagen der Spracherkennung



Mensch <-> Maschine

Analog <-> Digital

Autonom <-> Abhängig



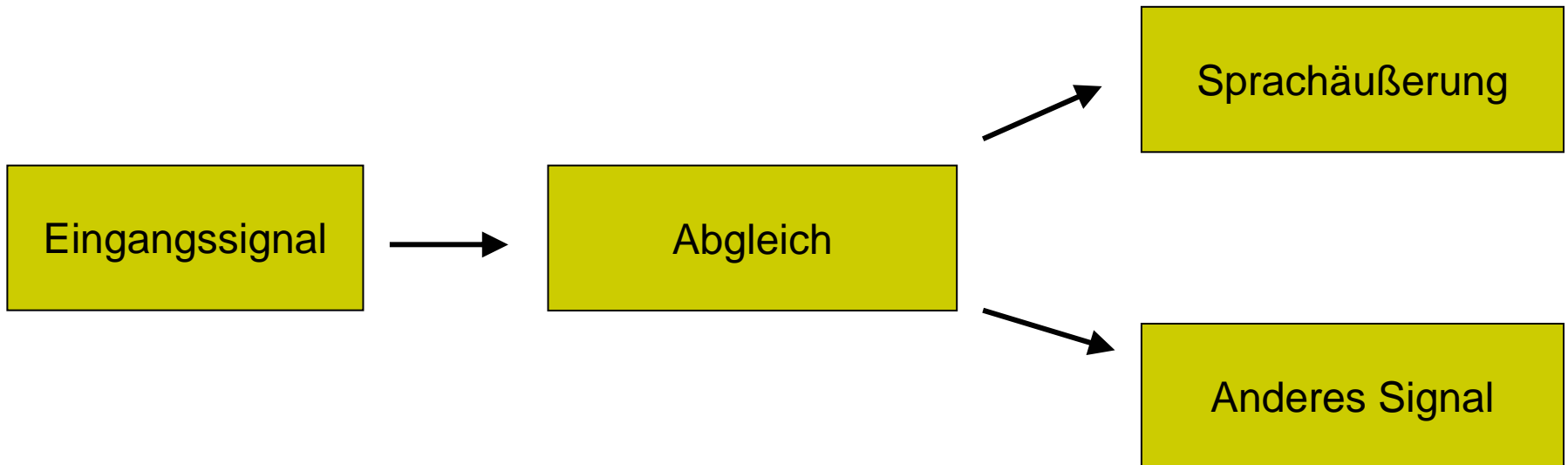
Gemeinsamkeiten?

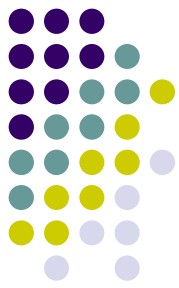
Genderspezifische Spracherkennung

Grundlagen der Spracherkennung



Wie erkennt die Maschine Sprache?

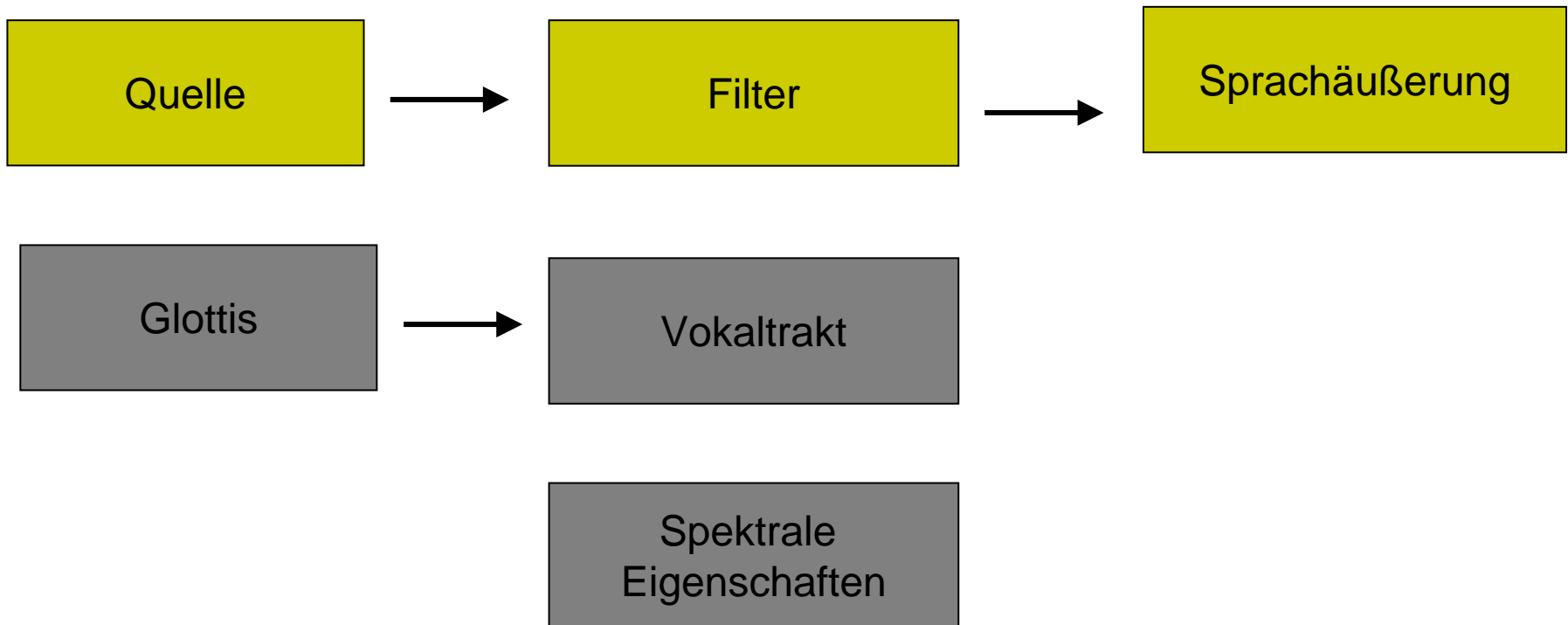




Genderspezifische Spracherkennung

Grundlagen der Spracherkennung

Wie erkennt die Maschine Sprache?





Genderspezifische Spracherkennung

Grundlagen der Spracherkennung

Wie erkennt die Maschine Sprache?

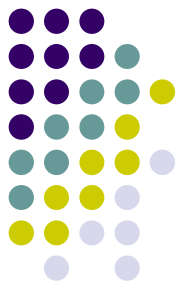
Formanten



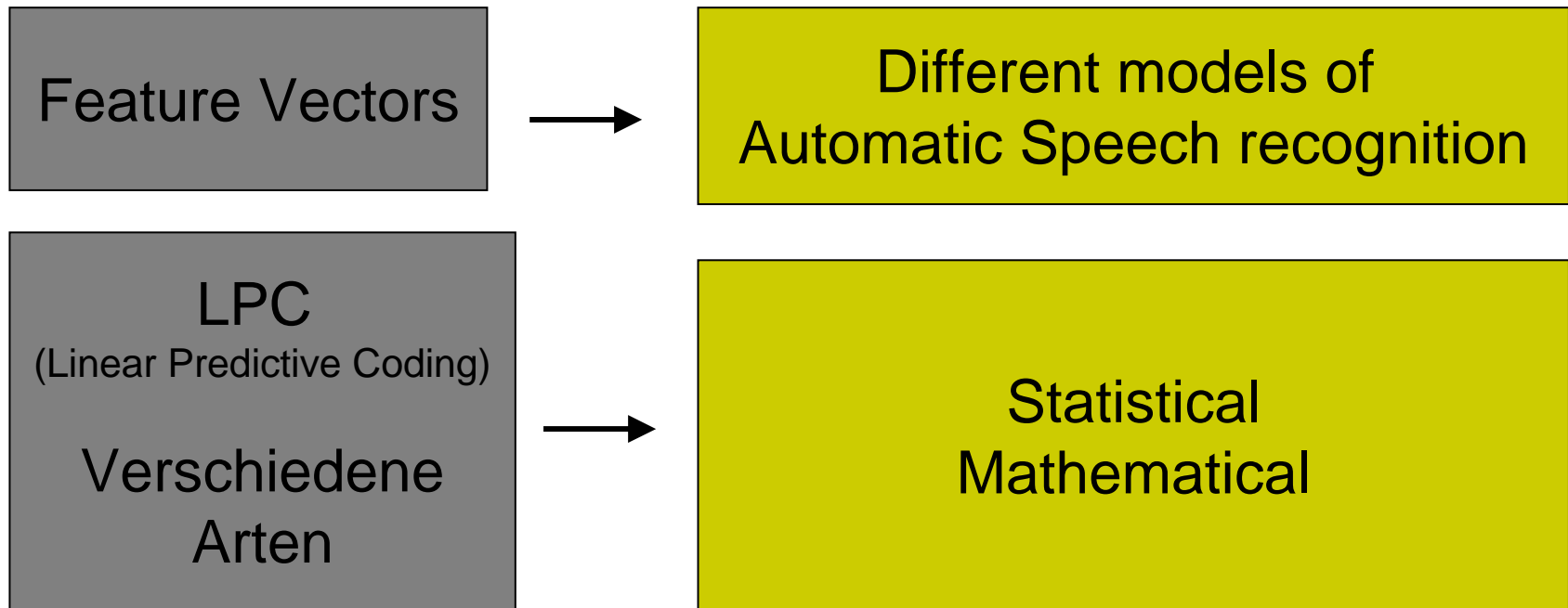
Merkmalsextraktion für Automatische Spracherkennung
Feature Vectors oder **Merkmalsvektoren**

Genderspezifische Spracherkennung

Grundlagen der Spracherkennung



Wie erkennt die Maschine Sprache?





Welche **Template Muster** eignen sich am besten zur Erkennung des Geschlechts?

TABLE I. Four recognition schemes.

	Test template from	Reference template from
Scheme TS	Token level	Speaker level
Scheme TG	Token level	Gender level
Scheme SG	Speaker level	Gender level
Scheme SS	Speaker level	Speaker level

Aus: Childers et al. (1991). Gender recognition from speech: Part I. Coarse analysis.



Welche Feature Vectors eignen sich am besten zur Erkennung des Geschlechts?

TABLE IV. Results for the Euclidean distance measure for recognition scheme SG for various filter orders and the four acoustic feature vectors.

		Correct recognition rate in %			
		Order = 8	Order = 12	Order = 16	Order = 20
Sustained vowels	ARC	78.8	78.8	78.8	82.7
	LPC	73.1	78.8	80.8	80.8
	FFF	N/A	98.1	N/A	N/A
	<u>RC</u>	<u>88.5</u>	<u>100.0</u>	<u>100.0</u>	<u>100.0</u>
	CC	82.7	92.3	90.4	90.4
Unvoiced fricatives	ARC	75.0	75.0	75.0	75.0
	LPC	80.8	69.2	71.2	71.2
	RC	80.8	80.8	80.8	80.8
	CC	71.2	75.0	84.6	82.7
Voiced fricatives	ARC	86.5	88.5	86.5	88.5
	LPC	92.3	92.3	92.3	90.4
	RC	94.2	96.2	96.2	96.2
	CC	94.2	98.1	98.1	96.2

Aus: Childers et al. (1991). Gender recognition from speech: Part I. Coarse analysis.



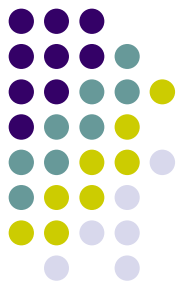
Welche Methode eignet sich am besten zur Erkennung des Geschlechts?

TABLE VI. Summary of the most effective feature vector and distance measure combinations for scheme SG (except as noted).

		Correct recognition rate in %			
		Filter order			
		8	12	16	20
Sustained vowels	LPC (pdf)		98.1		
	RC (pdf)		98.1		
	FFF (EUC)		98.1		
	<u>RC (EUC)</u>		<u>100.0</u>	<u>100.0</u>	<u>100.0</u>
Unvoiced fricatives	LPC (pdf)		86.5		
	CC (EUC) ^a		88.5	90.4	88.5
Voiced fricatives	LPC (LLD)			96.2	98.1
	LPC (LLD) ^a	98.1	96.2	96.2	
	CC (EUC)		98.1	98.1	96.2
	RC (EUC)		96.2	96.2	96.2

^a By using recognition scheme SS.

Aus: Childers et al. (1991). Gender recognition from speech: Part I. Coarse analysis.

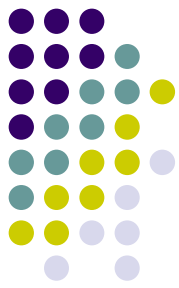


Welche **Kombination** eignet sich am besten zur Erkennung des Geschlechts?

RC-Merkmalvektoren

Euklidischer Abstand

Speaker vs. Gender



Welche **Charakteristika** eignen sich am besten zur Erkennung des Geschlechts?

Spektrale Charakteristika
sind ausschlaggebend
zur Erkennung des Geschlechts

Genderspezifische Spracherkennung



Welche **Formanten** eignen sich am besten zur Erkennung des Geschlechts?

	Childers II	Perry et. al
F ₀	96.2%	N. Signifikant
F ₁	90.4%	Signifikant
F ₂	98.1%	Hoch Sig.
F ₃	94.2 %	Signifikant
F ₄	96.2	Signifikant
F _{0+Spek.}	98.1%	-

M.Akyol

Genderspezifische Spracherkennung



Welche Erklärung lässt sich für diese Ergebnisse in Bezug auf F0 finden?

Perry et al.
VPN: Kinder
haben annähernd den selben F0
-> Trotzdem konnten die Probanden
Geschlechter unterscheiden

	4	8	12	16
m	67%	74%	82%	99.7%
f	62%	56%	56%	95%

Childers et al.
VPN: Erwachsene
klar unterscheidbare F0 -> Erkennungsrate bei 96.2 % (m: 100% - f: 92%)

Genderspezifische Spracherkennung



Ergebnisse

- Frauen haben eine etwas größere Variabilität in den Spektralen Eigenschaften
- Männer werden insgesamt etwas leichter erkannt
- F0 alleine scheint ausreichend zu sein um das Geschlecht erkennen zu können (ab 16 Jahren)
- Kombination von Formanten mit anderen Spektralen Charakteristika (Formant + Amplitude, Bandbreite) erhöht die Erkennungsrate Nicht.
- F2 scheint der stärkste Schlüssel zur Erkennung des Sprechergeschlechts zu sein

Genderspezifische Spracherkennung



Folgerungen

- Frauen und Männer unterscheiden sich in ihren Spektralen Merkmalen
- Sowohl Maschinelle (ASR) als auch Humane (Perzeption) Erkennung des Geschlechts scheint von F0 & F2 abhängig zu sein
- Für die Automatische Sprechererkennung scheinen andere Spektrale Merkmale wie Amplitude und Bandbreite keine Signifikante Rolle zu spielen

Genderspezifische Spracherkennung



Ausblick

- Man könnte nun mittels Perzeptionstests ermitteln ob für Menschen Amplitude und Bandbreite Auswirkungen auf das pezipierte Geschlecht haben
- Man könnte die Experimente von Childers mit Jugendlichen Stimmen nachstellen, um den Schwellwert (Alter) der Erkennungsmöglichkeit des Geschlechts herauszufinden

Genderspezifische Spracherkennung



Quellen

Wu, K. and D.G. Childers (1991). Gender recognition from speech: Part I. Coarse analysis. *Journal of the Acoustical Society of America*, 90, 1828–40.

Childers, D.G. and K. Wu (1991). Gender recognition from speech: Part II. Fine analysis *Journal of the Acoustical Society of America*, 90, 1841–56.

Perry, T., Ohde, R. & Ashmead, D. (2001). 'The acoustic bases for gender identification from children's voices.' *Journal of the Acoustical Society of America*, 109, 2988–2998.