

# Altersbedingte Veränderungen der Stimme und Sprechweise von Frauen

Eine perzeptive und akustische Analyse

# Übersicht

- Stand der Forschung
- Ziele
- Daten und Methoden
- Ergebnisse der perzeptiven Analyse
- Ergebnisse der akustischen Analyse
- Zusammenfassung

# Stand der Forschung

physische Veränderungen im Sprechapparat:

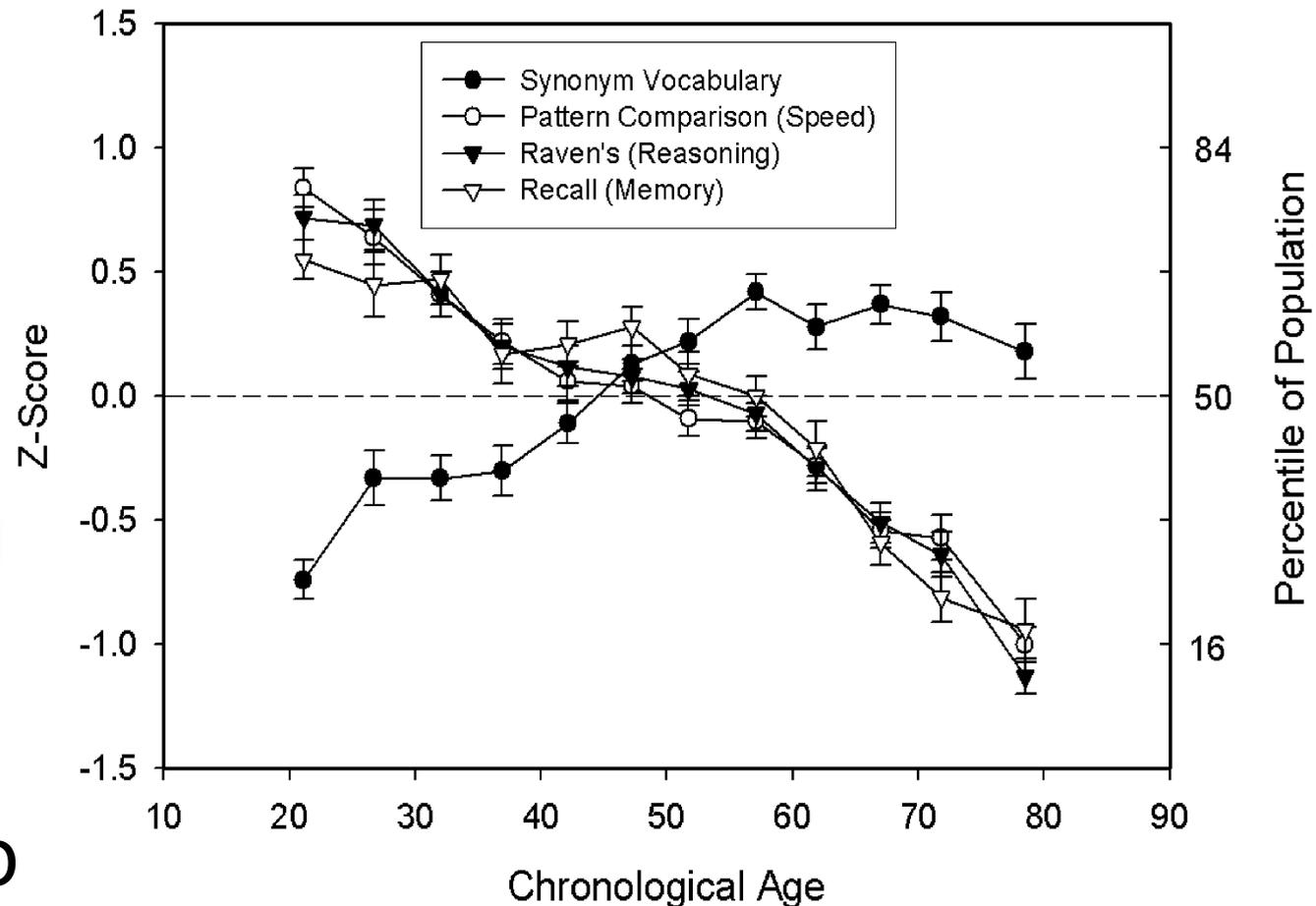
- Lunge wird kleiner, leichter, steifer, die Muskulatur wird weniger und schwächer
- Kehlkopf senkt sich ab; Kehlkopfknorpel verknöchern, Muskulatur versteift
- Stimmlippen werden länger, steifer, muskulärer Anteil nimmt ab
- Schädel wächst weiter, Gesichtsmuskulatur und Zunge werden schwächer und steifer
- Nervensystem verliert an Masse (Myelin aber auch Neuronen); Verringerung von Neurotransmittern

# Stand der Forschung

kognitive

Veränderungen:

- angeeignetes Wissen nimmt (bis zum 60. Lebensjahr) zu
- Effektivität der Verarbeitung nimmt stetig ab



# Stand der Forschung

Ist das Alter eines Sprechers über auditive Wahrnehmung zu ermitteln?

- Ptacek & Sander (1966): 99% richtige Zuordnungen gelesener Äußerungen zu jung vs. alt
- Linville & Fisher (1985): 51% richtige Zuordnungen gehaltener Vokale zu 3 Kategorien
- Shipp & Hollien (1969):  $r=0,88$  zwischen (gemittelten) Schätzungen des Alters (PA) in Jahren anhand gelesener Rede und dem chronologischen Alter (CA)
- Hartman (1979):  $r=0,93$

# Stand der Forschung

von (Laien-) Hörern genannte Charakteristika höheren Sprecheralters:

- tiefere Sprechtonlage
- vermehrte Heiserkeit, Rauigkeit der Stimme
- vermehrte Behauchung
- vermehrte Zitterigkeit
- vermehrte Stimm(-an-)spannung (engl.: strain)
- reduzierte Lautstärke
- langsamere Sprechweise
- weniger genaue Artikulation
- vermehrtes Verzögern des Redeflusses (hesitancy)
- längere Sprechpausen

# Stand der Forschung

akustische Veränderungen des Sprachsignals:

- mittlere Grundfrequenz:
  - ist konstant (Baken, 1987)
  - steigt (Hummert, 1999)
  - sinkt (Linville, 2001)
- Dispersion der Grundfrequenz:
  - max. Stimmumfang sinkt (Böhme & Hecker, 1970)
  - Spannbreite und SD sinkt in Rede (Decoster & Debruyne, 1997 sowie Morgan & Rastatter, 1986)
  - SD steigt in Rede (Morris & Brown, 1994 sowie Hummert, 1999)
  - SD steigt in kontrollierter Phonation (Linville, 2001)

# Stand der Forschung

- Perturbationsmaße:
  - steigen, aber eher durch schlechteren gesundheitlichen Zustand (Orlikoff, 1990 sowie Ramig & Ringel, 1983)
  - Jitter aber mehr als Shimmer (Brückl, 2003 sowie Winkler, 2008)
  - verändern sich nicht (Decoster & Debruyne, 1997)
- Modulationsmaße:
  - Perturbationsmaße, die zeitlich größere Abschnitte integrieren sind besser → Modulation? (Linville, 2001)
  - Frequenztremor-Intensitäts-Index steigt (Brückl, 2003)

# Stand der Forschung

- Maße der spektralen Energieverteilung
  - visuelle Begutachtung von Sonagrammen von Männern: spektrales Rauschen nimmt zu (Ramig, 1983)
  - LTAS-Bänder um 6300 Hz gelesener Rede zeigen erhöhte Energie (Linville, 2002)
  - H1-H2 gehaltener Vokale sinkt im Querschnitt aber nicht im Längsschnitt (Decoster & Debruyne, 1997)

# Stand der Forschung

- Maße spektralen Rauschens
  - MDVP-Maße steigen im Gruppenvergleich (Xue & Deliyski, 2001)
  - MDVP-Maße nicht verändert (Brückl & Sendlmeier 2003)
  - Praat-HNR (spezifischer Bänder) sinkt, aber abhängig vom Äußerungstyp und Art der Altersskala (Winkler, 2008):
    - gar nicht bei spontaner Rede
    - bei Wörtern mit CA
    - bei Vokalen mit PA

# Stand der Forschung

- Formanten
  - Point of Formant Concentration sinkt (Endres et al., 1971)
  - zunehmende Zentralisation (Benjamin, 1982, 1997 sowie Rastatter & Jacques, 1990)
  - nur im /u/ steigt F1 und sinkt F2 → eher konstante Formantlagen (Rastatter et al., 1997)
- Sprechgeschwindigkeit
  - AR sinkt (Oyer & Deal, 1985 sowie Smith et al., 1987)
  - AR bleibt konstant (Hoit et al., 1989)
  - AR sinkt nur beim Lesen (Brückl & Sendlmeier, 2003)

# Ziele

Beantwortung der Fragen:

- Wie genau lässt sich das an einer Äußerung perzeptiv ermittelbare Alter (PA) bestimmen?
- Wie genau lässt sich das chronologische Alter (CA) einer Sprecherin aus dem PA bestimmen?
- Welche akustischen Merkmale erzeugen die Alterswirkung?
- Welche akustischen Merkmale indizieren das CA?

# Daten: Sprecherinnen

- Ziele bei der Auswahl:
  - möglichst große Variabilität hinsichtlich des Alters
  - Verallgemeinerbarkeit der Ergebnisse auf möglichst große Grundgesamtheit
- ad-hoc-Stichproben
- Querschnitt: 88 Sprecherinnen, 20 bis 87 Jahre alt
- Längsschnitt: 9 Sprecherinnen, bei der 1. Aufnahme 20 bis 87 Jahre alt, 5 Jahre älter bei der 2.

# Daten: Äußerungen

- gehaltene Vokale
  - drei Qualitäten: /a/, /i/ und /u/
  - drei Abschnitte: Anfang, quasi-stationärer Mittelteil, Schluss
  - [Beispiel: stat. /a/ jung, stat. /a/ alt]
- gelesene Rede [Bsp.: jung, alt]
- spontane Rede [Bsp.: jung, alt]

# Methodischer Dreiklang

zur Analyse mündlicher Kommunikation:

- perzeptive Analyse
  - auditive
    - Generierung von Hypothesen
    - Bestimmung vergleichbarer Einheiten
  - perzeptive
- akustische Analyse
- statistische Integration



# Methoden: perzeptive Analyse

## Verfahren

1. Konstruktion von Wahrnehmungsverfahren
2. Bestimmung der Urteilerreliabilität
3. Bestimmung der Validität der Urteile (Korrelation)
4. Regression

## Zweck

1. Beurteilung des Sprecheralters durch Hörer
2. Prüfung der Übereinstimmung der Hörer; Beurteilung der Genauigkeit des PA; Zusammenfassung → (beste) Schätzer des PA
3. Prüfung des Zusammenhangs von CA und PA
4. Bestimmung von Modellen zur Vorhersage von CA aus PA

# Methoden: akustische Analyse

## Verfahren

1. Berechnung akustischer Maße
2. Korrelation mit CA und PA
3. Regression mit CA und PA

## Zweck

1. Abbildung von Konstrukten der Stimme/Sprechweise
2. Bestimmung des Zusammenhangs mit CA und PA
3. Bestimmung von Modellen zur Vorhersage von CA und PA und Auswahl unabhängiger Prädiktoren

# Wahrnehmungserhebungen

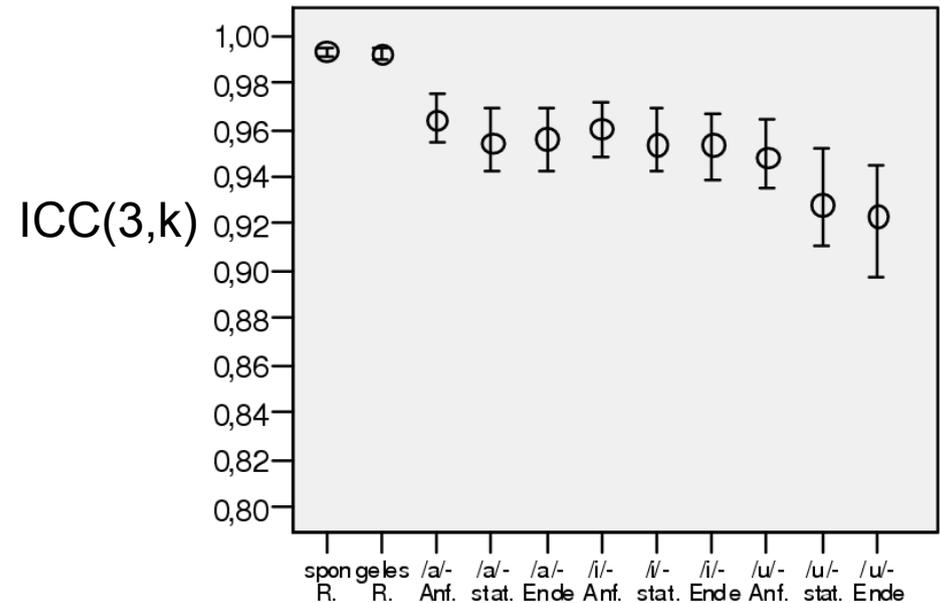
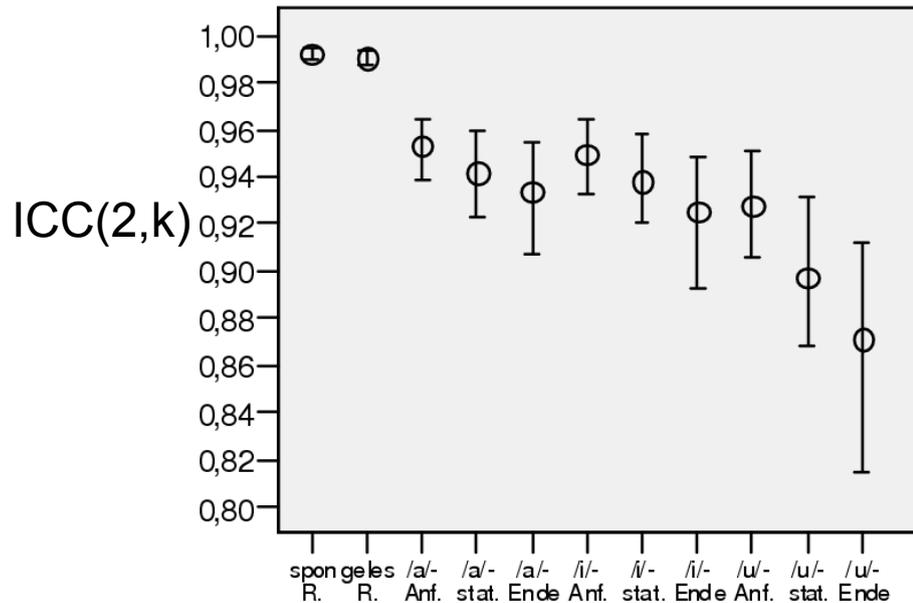
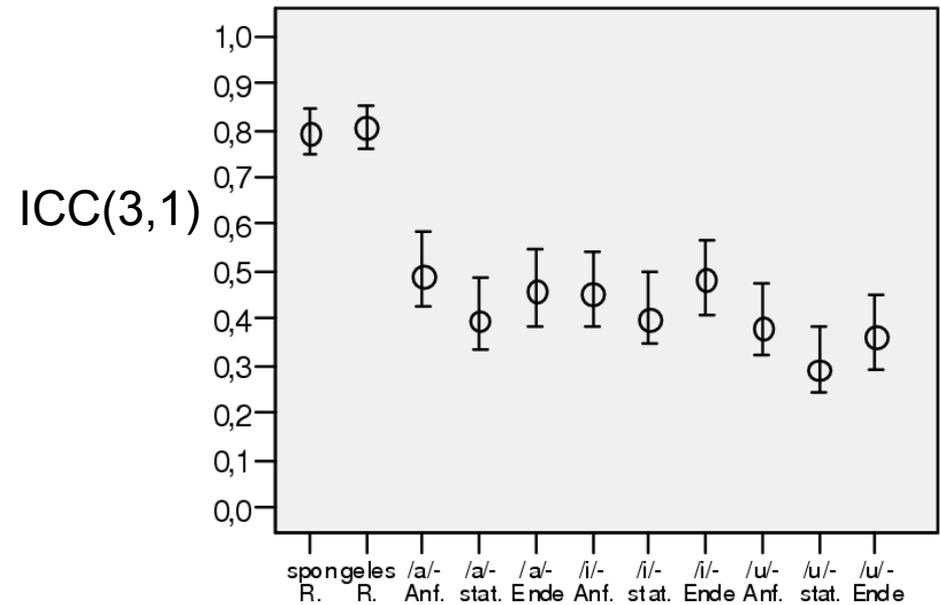
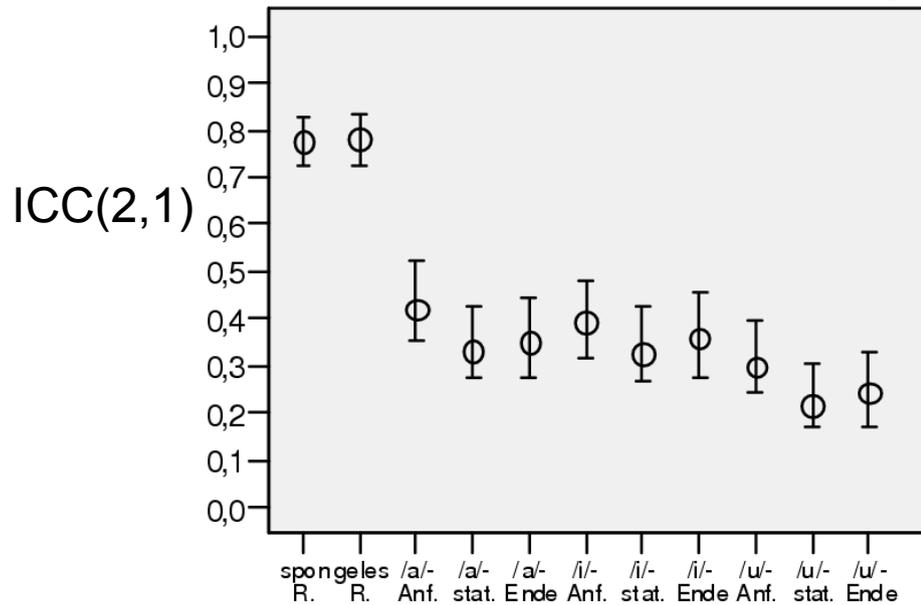
- 5 verschiedene Wahrnehmungserhebungen: 4 an den Querschnittsdaten, 1 an den Längsschnittsdaten
- Gemeinsamkeiten der Querschnittserhebungen: das (chronologische) Alter wird durch Hörer abgeschätzt
- Unterschiede der Querschnittserhebungen:
  - verschiedene Hörer
  - unterschiedliche Teilmengen und Abfolgen der Äußerungen
  - Papier und Bleistift vs. computergestützt

# Wahrnehmungserhebungen

- 115 Hörer in der (zusammengefassten) Querschnittserhebung
- im Mittel (pro Äußerungstyp) zwischen 37,92 (spontane Rede) und 21 (/u/-Enden)
- 34 Hörer in der Längsschnittserhebung, 32 davon beurteilen Vokaläußerungen
- Aufgabe der Hörer in der Längsschnittserhebung: Paarvergleich von gleichen Äußerungen gleicher Sprecherinnen

# Ergebnisse: Perzeptive Analyse

## Urteilerübereinstimmung im Querschnitt



# Ergebnisse: Perzeptive Analyse

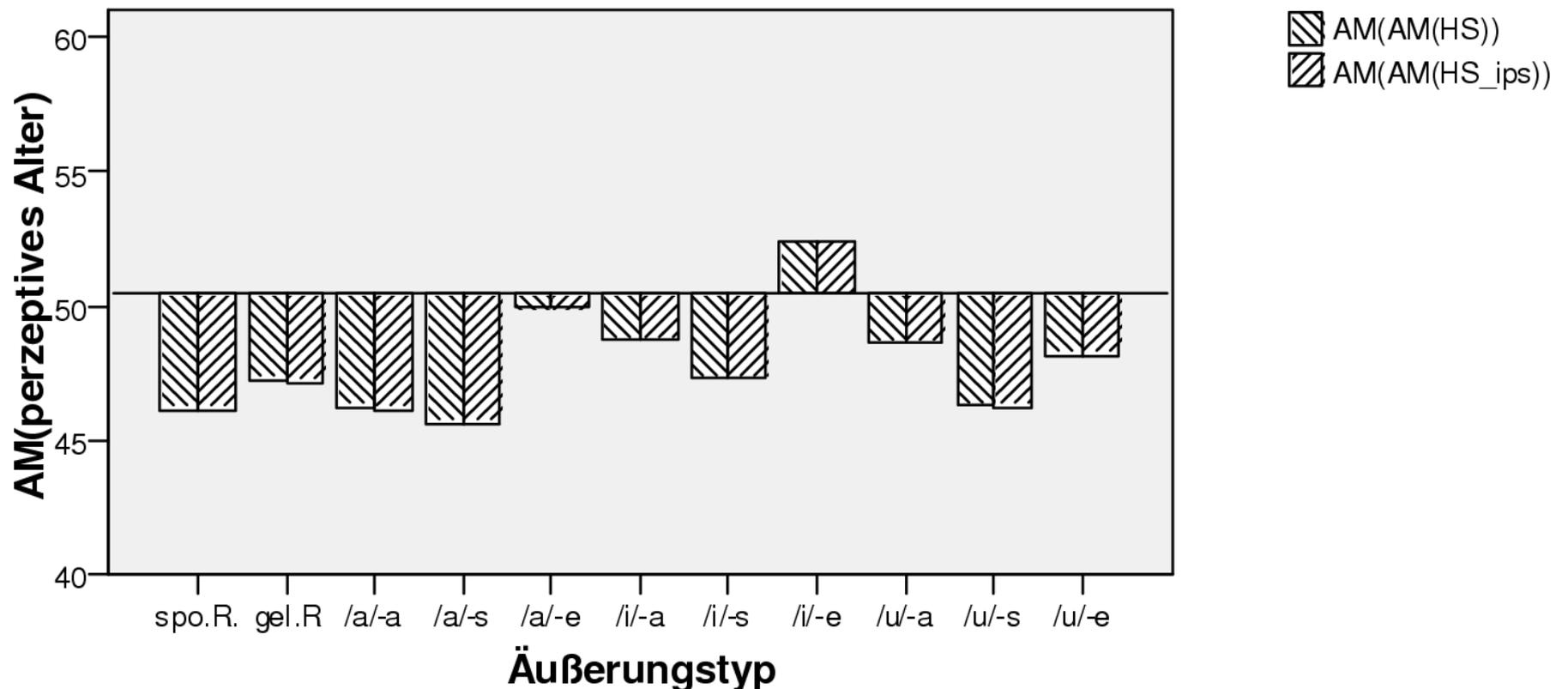
Messgenauigkeit der Hörer (Querschnitt):

aus den Reliabilitäten (ICCs) lassen sich Konfidenzintervalle für (neue) Hörschätzungen ableiten:

- $KI(95\%;2,1)$ :
  - spontane Rede: 25,43 Jahre
  - /i/-Enden: 37,39 Jahre
- $KI(95\%;2,k)$ 
  - spontane Rede: 4,68 Jahre ( $k=37,92$ )
  - /i/-Enden: 12,78 Jahre ( $k=22$ )

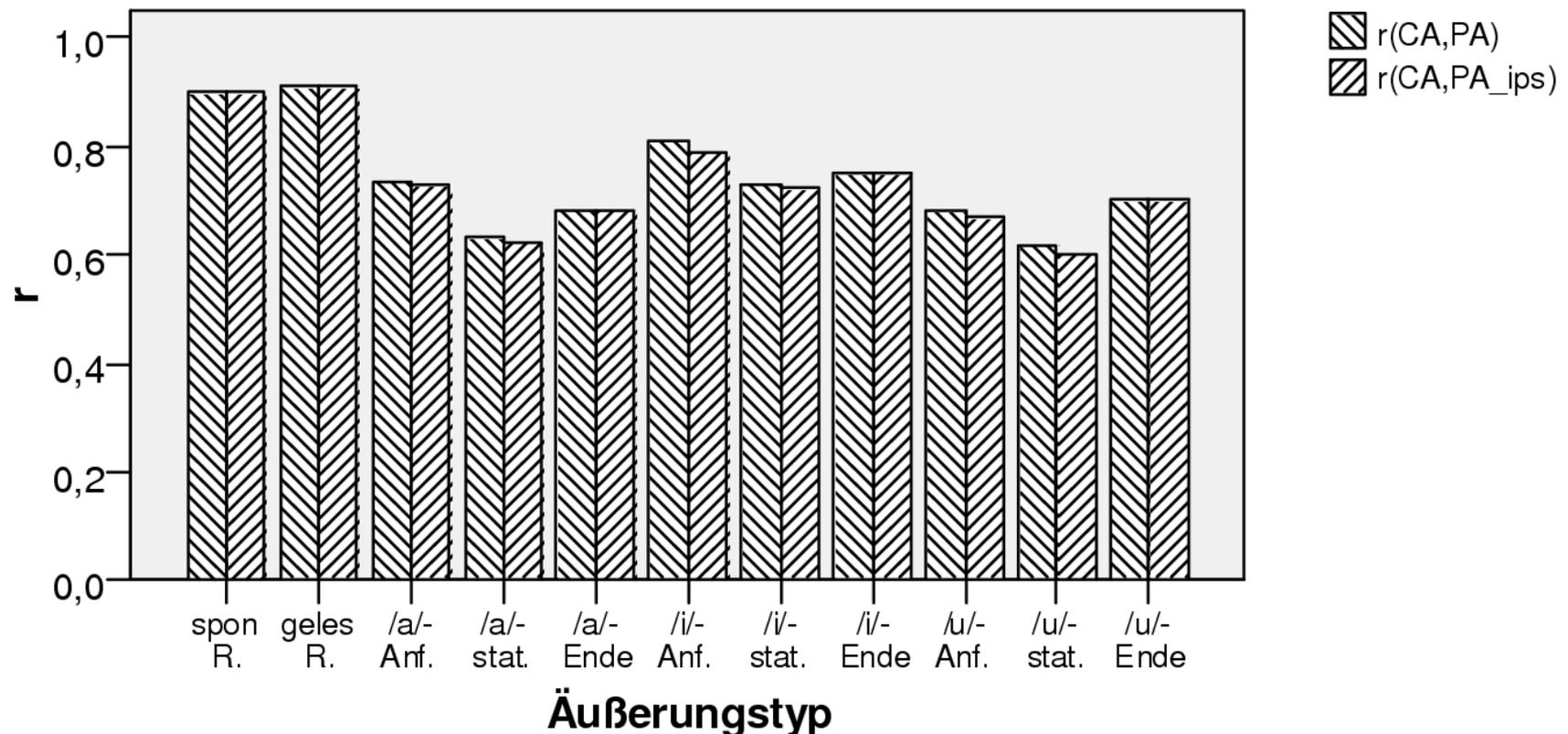
# Ergebnisse: Perzeptive Analyse

Die Mittel pro Äußerungstyp der (Schätzungen der) perzeptiven Alterswerte (Querschnitt):



# Ergebnisse: Perzeptive Analyse

Korrelation zwischen PA und CA pro Äußerungstyp (Querschnitt):

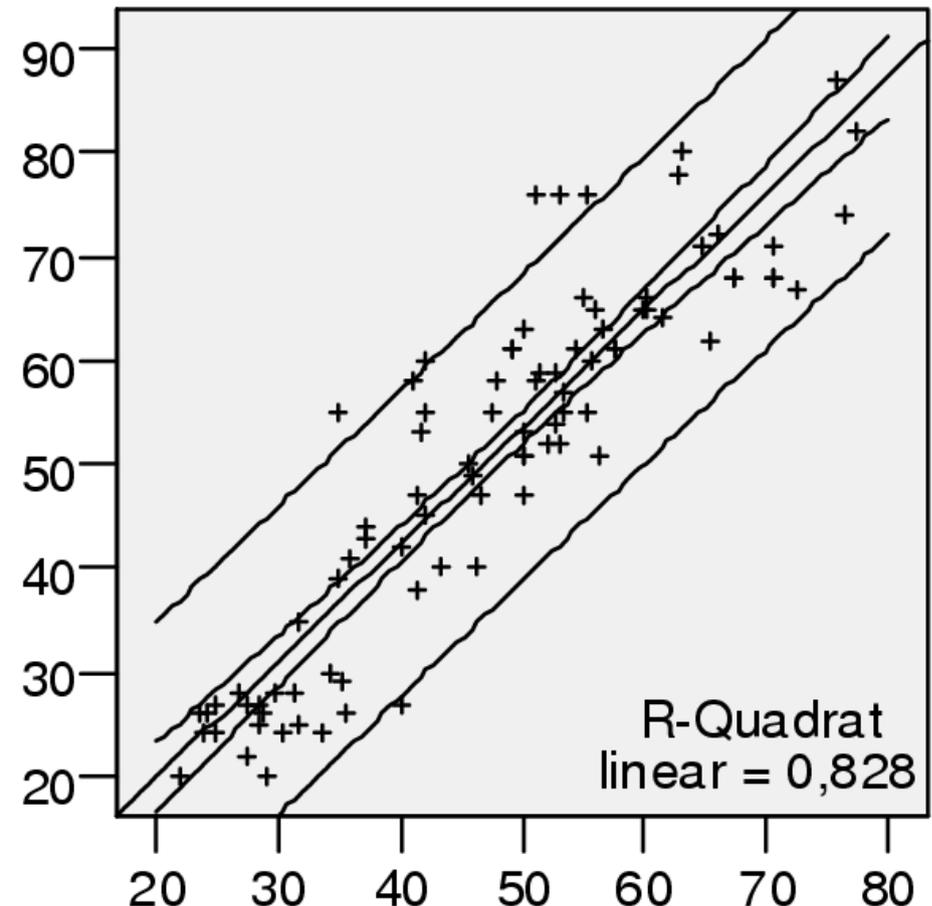


# Ergebnisse: Perzeptive Analyse

Regression zur  
Erklärung von CA aus PA  
(Querschnitt):

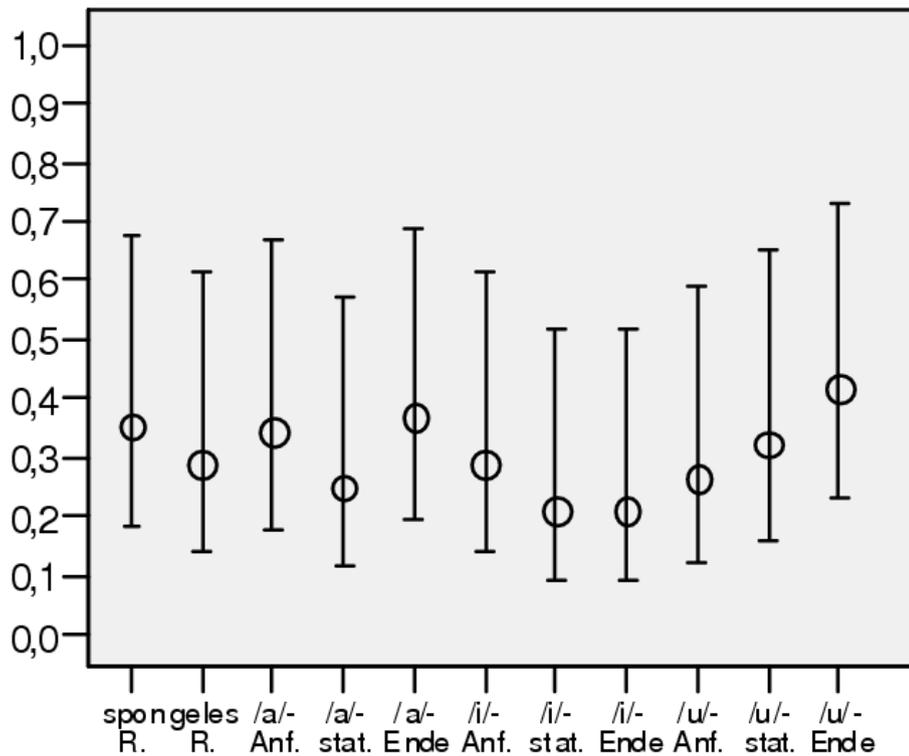
$$\begin{aligned} \text{CA}(\text{gelR.50}) &= \\ 1,13 * 50 - 2,8 &= \\ 53,7 \end{aligned}$$

$$\text{KI}(95\%, 50) = [38,9; 68,3]$$

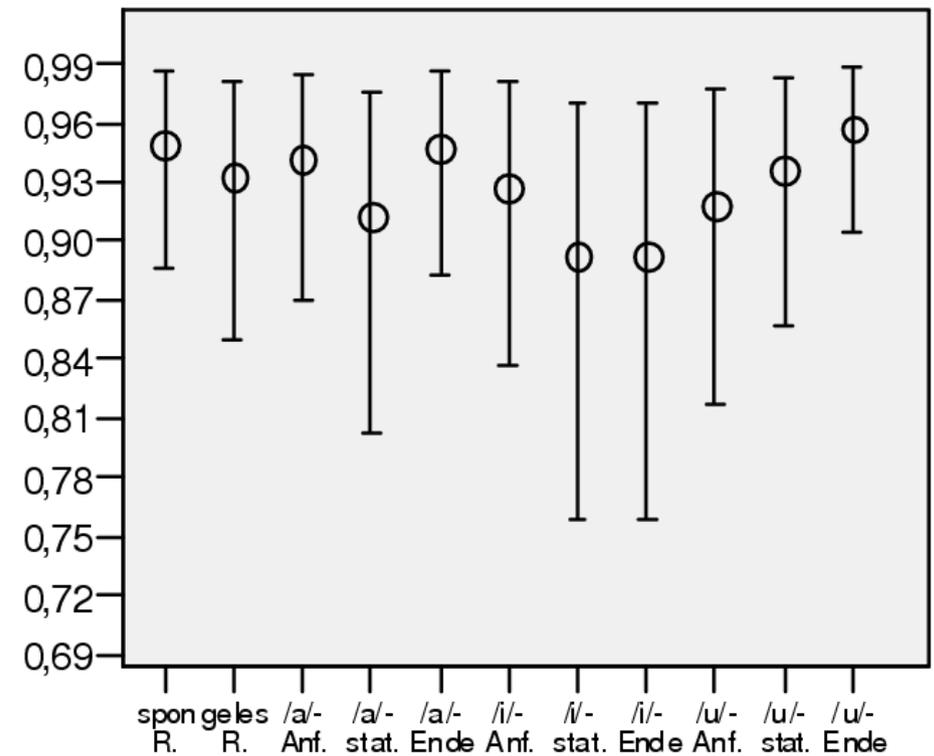


# Ergebnisse: Perzeptive Analyse

## Urteilerübereinstimmung im Längsschnitt



ICC(3,1)



ICC(3,k)

# Ergebnisse: Perzeptive Analyse

Korrelation zwischen PA und CA pro Äußerungstyp (Längsschnitt):

- auf 5%-tigem  $\alpha$ -Fehler-Niveau signifikant sind nur die Korrelationen anhand der Rede-äußerungen:
- $r(\text{spo.}) = 0,731$  ( $p = 0,008$ )
- $r(\text{gel.}) = 0,653$  ( $p = 0,020$ )

→ Hörschätzungen anhand der Vokale sind zwar reliabel, aber nicht valide

# Konstrukte und Parameter

## Konstrukte

Stimmtonlage

Stimmumfang

Stimmvariabilität

Rauhigkeit/Heiserkeit

Zittrigkeit der Stimme

## Parameter

Grundfrequenz (F0)

Minimum, Maximum und  
Range von F0

Standardabweichung, Varia-  
tionskoeffizient von F0

Grundfrequenzperturbation  
(Jitter)

Amplitudenperturbation  
(Shimmer)

Grundfrequenzmodulation  
(Frequenztremer)

Amplitudenmodulation (Am-  
plitudentremor)

# Konstrukte und Parameter

## **Konstrukte**

Behauchung, Harmonizität,  
Stimmhärte

Energieverteilung

Sprechintensität

Sprechpausen

## **Parameter**

Quotienten aperiodischer und  
periodischer Energie in be-  
stimmten Frequenzbändern)

Intensitäten in LTAS-Bän-  
dern

Intensität von Harmonischen  
Formantfrequenzen

vocal effort

Pausendauern

Pausenanzahlen

# Konstrukte und Parameter

## **Konstrukte**

Artikulationsgenauigkeit

Sprechgeschwindigkeit

Artikulationsmonotonie

## **Parameter**

Anzahl der Elaborationen und Reduktionen

Lautzahlminderungsquotient

Formanttransitionen in Diphthongen

Formantdifferenzen in Vokal-Nasal-Diphonem

Artikulationsrate

um Koartikulationseffekte

korrigierte Artikulationsrate

Phon-Jitter

Phon-Shimmer

Silben-Jitter

# Konstrukte und Parameter

## **Konstrukte**

Artikulationsmonotonie

Betonung

## **Parameter**

Silben-Shimmer

Veränderungen der Perturbationsmaße von Silbe zu Silbe

Dauern (kanonisch) betonter und unbetonter Vokale

Dispersion von F1 und F2 von unbetonten Vokalen

# Ergebnisse: Akustische Analyse

- kein einziger Parameter erzielt an allen Äußerungstypen höchstsignifikante Korrelationen
- nur die Standardabweichung der Periodendauern ist bei beiden Redetypen und auch bei gehaltenen Vokalen ein passables Korrelat beider Altersskalen:
  - $r(\text{spoR}, \text{CA}) = 0,431$
  - $r(/a/e, \text{PA}) = 0,610$

# Ergebnisse: Akustische Analyse

Beste Korrelate:

Typ	spoR	geIR	/a/a	/a/s	/a/e	/i/a	/i/s	/i/e	/u/a	/u/s	/u/e
<b>CA</b>	Pf0md _vow	Pf0- min_a	ATrl_ Ln	ATrl_ Ln	Pf0- min	Pf3	MFTR	FTrl_ Ln		Pt0sd	FTrl_ Ln
	-0,636	-0,583	0,585	0,489	-0,458	-0,350	0,529	0,383		0,363	0,405
<b>PA</b>	Pf0md _vow	Pf0- min_a	FTrl_ Ln	FTrl_ Ln	Pf0- min	Pf3	FTrl_ Ln	FTrl_ Ln	Pf3sd	Pt0sd	FTrl_ Ln
	-0,621	-0,648	0,521	0,635	0,669	-0,439	0,689	0,466	0,408	0,644	0,559

- Korrelate von PA und CA unterscheiden sich nicht wesentlich
- Parameter korrelieren tendenziell besser mit PA

# Ergebnisse: Akustische Analyse

Ergebnisse der Verknüpfung mehrerer Parameter zur Erklärung des Alters:

- multiple Regression erreicht 75% Varianzaufklärung von CA, Prädiktoren sind:
  - Minimum der Grundfrequenz (/a/ in gel. Rede)
  - Standardabweichung der Periodendauern (Nasale gel. Rede)
  - Geräuschintensität 70-4500Hz (spon. Rede)
  - Frequenztreemorintensität (gehaltene /i/-s)
  - korr. Artikulationsrate (gel. Rede)
  - Ampl.-Perturbation, Glättung 11 (Vokale in spon. Rede)
  - Intensität im LTAS-Band um 6000Hz (gel. Rede)

# Ergebnisse: Akustische Analyse

Ergebnisse der Verknüpfung mehrerer Parameter zur Erklärung des Alters:

- multiple Regression erreicht 68% Varianzaufklärung von PA (gel. Rede), Prädiktoren sind:
  - Minimum der Grundfrequenz von /a/-Segmenten
  - korr. Artikulationsrate
  - Standardabweichung des 4. Formanten vokalischer Segmente

# Ergebnisse: Akustische Analyse

## akustische Prädiktoren des Alters:

- spontane Rede:
  - Grundfrequenzparameter
  - (Amplituden-) Perturbationsparameter
  - Maße des spektralen Rauschens
- gelesene Rede:
  - Grundfrequenzparameter
  - Dauer- bzw. Sprechgeschwindigkeitsmaße
- gehaltene Vokale:
  - Tremormaße
  - Grundfrequenzparameter
  - (Amplituden-) Perturbationsparameter
  - Maße des spektralen Rauschens

# Ergebnisse: Akustische Analyse

Ergebnisse bzgl. der replizierten Maße:

- Formanten: sind untauglich
- Intensitäten der Harmonischen: untauglich
- LTAS-Intensitäten: um 6000Hz erhöht, unklar
- Sprechgeschwindigkeit: sinkt (beim Lesen)
- Grundfrequenz: sinkt, v.a. das Minimum
- Spektrales Rauschen: steigt
- Perturbationen: steigen, Amp.P geeigneter

# Ergebnisse: Akustische Analyse

Ergebnisse bzgl. der neu und weiter entwickelten Maße:

- korrigierte Sprechrate: besser als nicht korrigierte... Aufwand?
- Maße der prosodischen Veränderung von Silben und Phonemen: ok, aber verbesserungswürdig
- segmentweise Extraktion stimmqualitativer Maße: lohnend
- Maße spektralen Rauschens: besser als bestehende
- Tremormäße: liefern beste Korrelate (an Vokalen)

# Zusammenfassung

- verschiedene Äußerungstypen vermitteln unterschiedliches Ausmaß an Information über das Alter
- in verschiedenen Äußerungstypen wird Alter durch unterschiedliche akustische Merkmale vermittelt

# Zusammenfassung

- die präziseste Schätzung des Sprecheralters erhält man durch die Zusammenfassung mehrerer Hörerurteile
- Verfahren zur Bestimmung der Urteiler-Reliabilität wurden auf die Anwendbarkeit an unvollständigen Urteilmatrizen erweitert
- generell unterschätzen Hörer das Alter von Sprecherinnen
- EMU liegt bei ca. 5 Jahren

# Zusammenfassung

- Genauigkeit der Altersschätzung über akustische Algorithmen erreicht (erstmal) diejenige einer Hörergruppe
- besonderen Beitrag hierzu leisten:
  - Tremormaße
  - Maße spektraler Energieverteilungen
  - Messung stimmqualitativer Parameter an (phonetischen) Rede(-segmenten)
  - um Koartikulationseffekte korrigierte Sprechrate
  - Maße der Sprechtonlage

# Zusammenfassung

- Die Stimme älterer Sprecherinnen ist
  - zittriger
  - tiefer
  - rauher
  - behauchter
- die Sprechweise (beim Lesen) ist
  - langsamer und/oder weniger genau artikuliert
  - durch vermehrte/längere Sprechpausen unterbrochen