Aging Female Voices:

an Acoustic and Perceptive Analysis

Technical University of Berlin, Germany

Institute for Language and Communication

Markus Brückl, Walter Sendlmeier

Sue Ellen Linville:

- “Firm conclusions as to the effect of aging on jitter and shimmer levels are not now possible.”
- “Amplitude SD in female speakers with aging has yet to be investigated.”
- “Research is necessary to examine spectral noise as a correlate of perceived age estimates from women’s voices.”
- “Studies have not been conducted correlating age estimates to speech rate in female speakers.”

Aims of this Study

Investigate

- Amp SD (and other perturbation measures)
- Articulation rate
- Spectral noise,

as a function of chronological age and perceived age

- Further acoustic parameters:
  - tremor measures
  - F0
- Relevance of vowel onset for age perception and age measurement

- 56 speakers, aged from 20 to 87 (AM=49.77, SD=16.01)

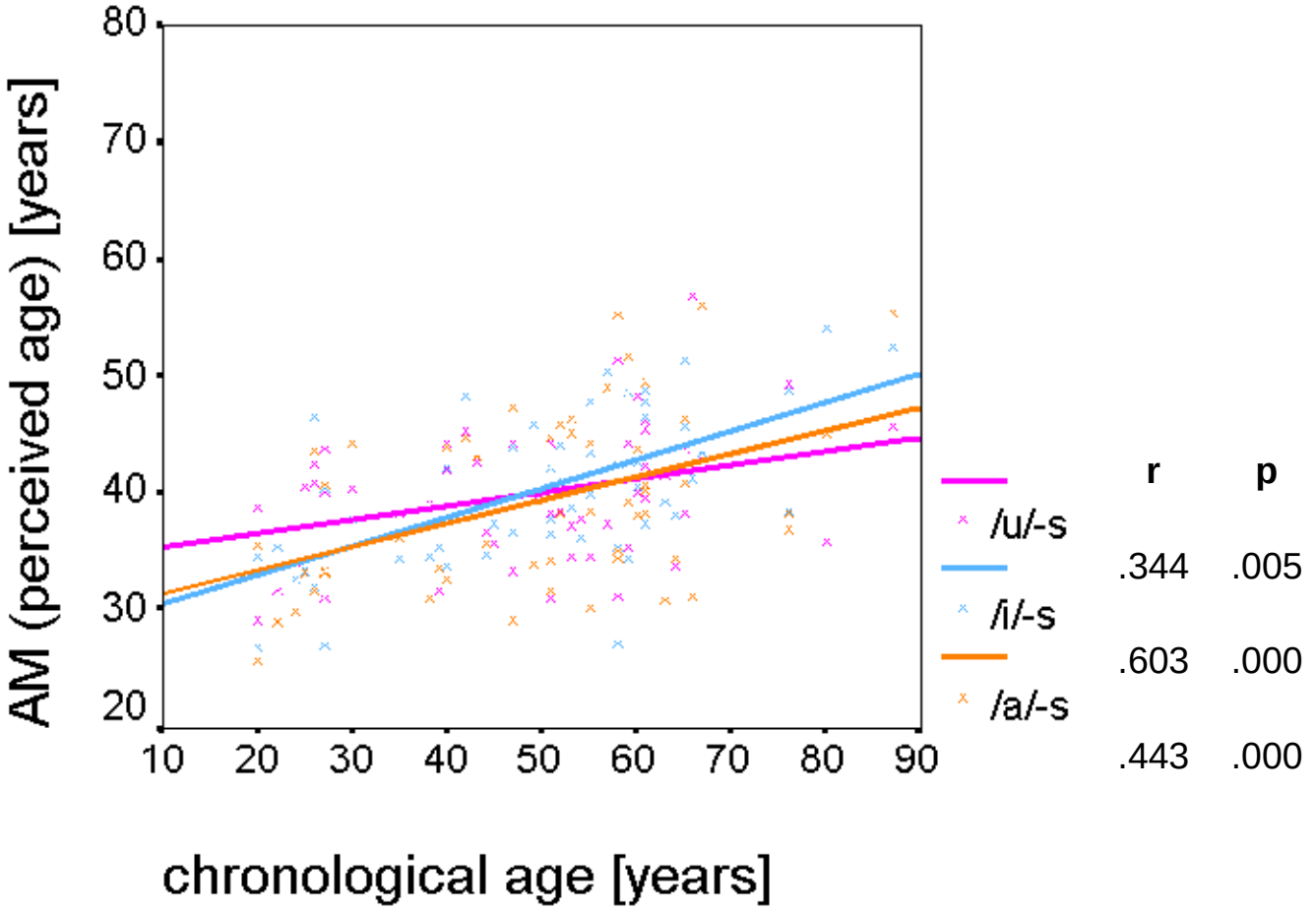
- 8 types of voice samples, assumed to differ in amount and type of age-related information:
  - Spontaneous speech (s-sp)
  - Read speech (r-sp)
  - Sustained vowels /a/, /i/ and /u/
    - Onset sample (e.g. /a/-o)
    - Quasi-stationary sample (e.g. /a/-s)

- 15 young adult listeners rated perceived age of each sample
- estimations are significantly concordant

  ⇒ listeners’ age perceptions are averaged

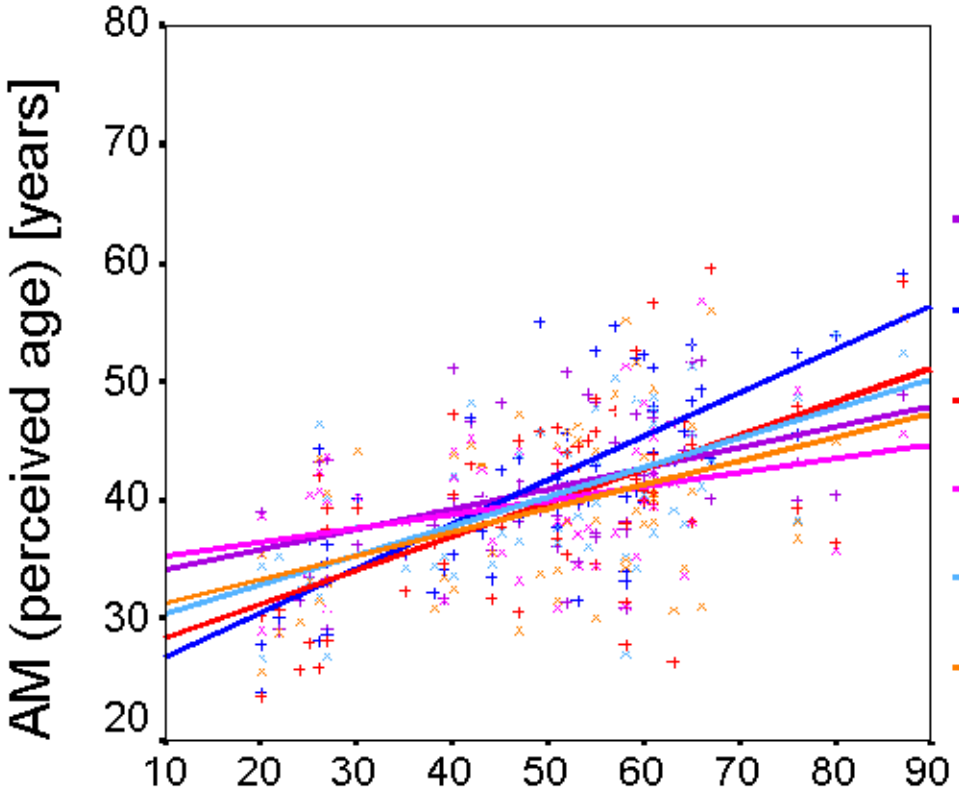
  ⇒ perceived age for each voice sample
- 22 Acoustic parameters are extracted separately for each voice sample
- Correlation of:
  - acoustic parameters and chronological age
  - acoustic parameters and perceived age
  - perceived age and chronological age

Accuracy of Perception



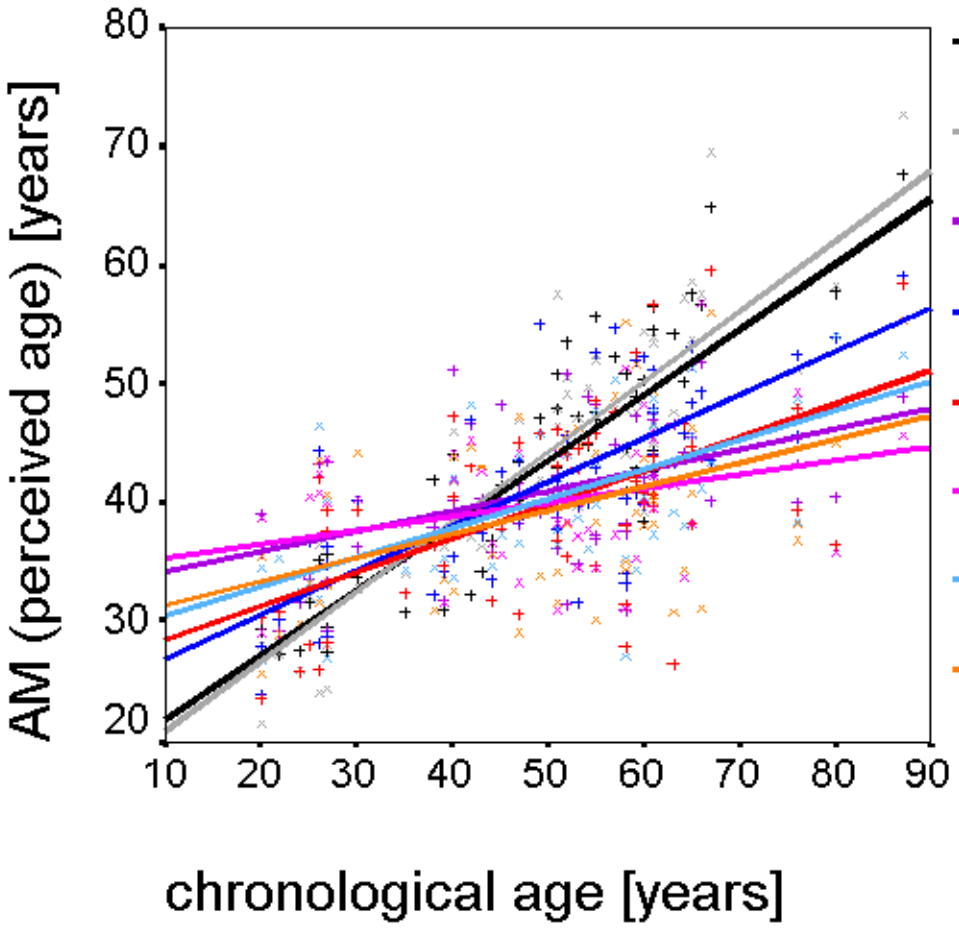
	r	p
/u/-s	.344	.005
/i/-s	.603	.000
/a/-s	.443	.000

Accuracy of Perception



	r	p
/u/-o	.460	.000
/i/-o	.738	.000
/a/-o	.559	.000
/u/-s	.344	.005
/i/-s	.603	.000
/a/-s	.443	.000

Accuracy of Perception



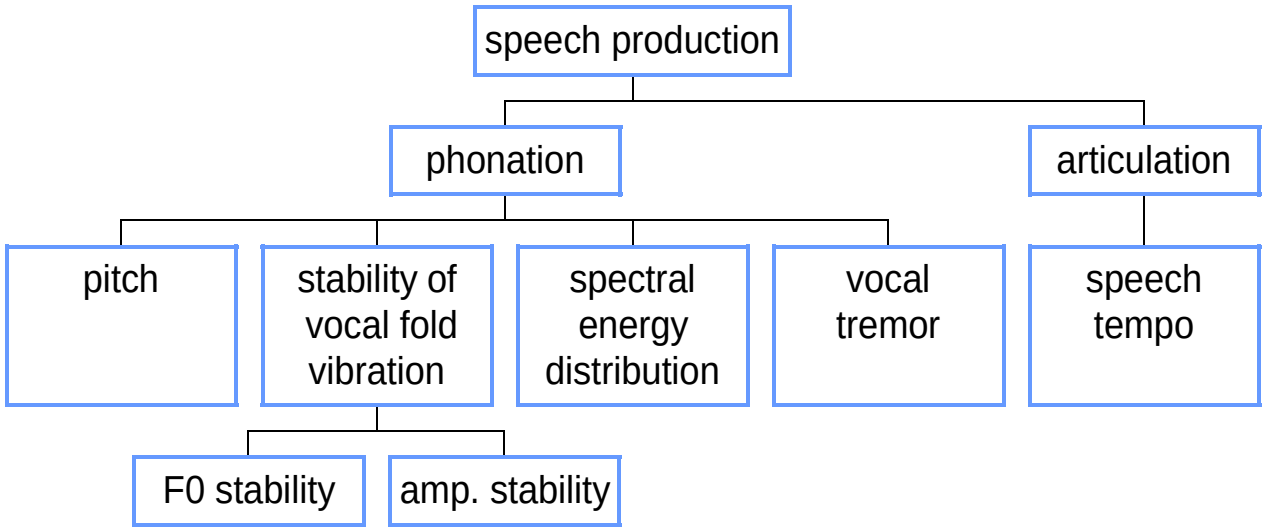
	r	p
s-sp	.864	.000
r-sp	.862	.000
/u/-o	.460	.000
/i/-o	.738	.000
/a/-o	.559	.000
/u/-s	.344	.005
/i/-s	.603	.000
/a/-s	.443	.000

Accuracy of Perception:

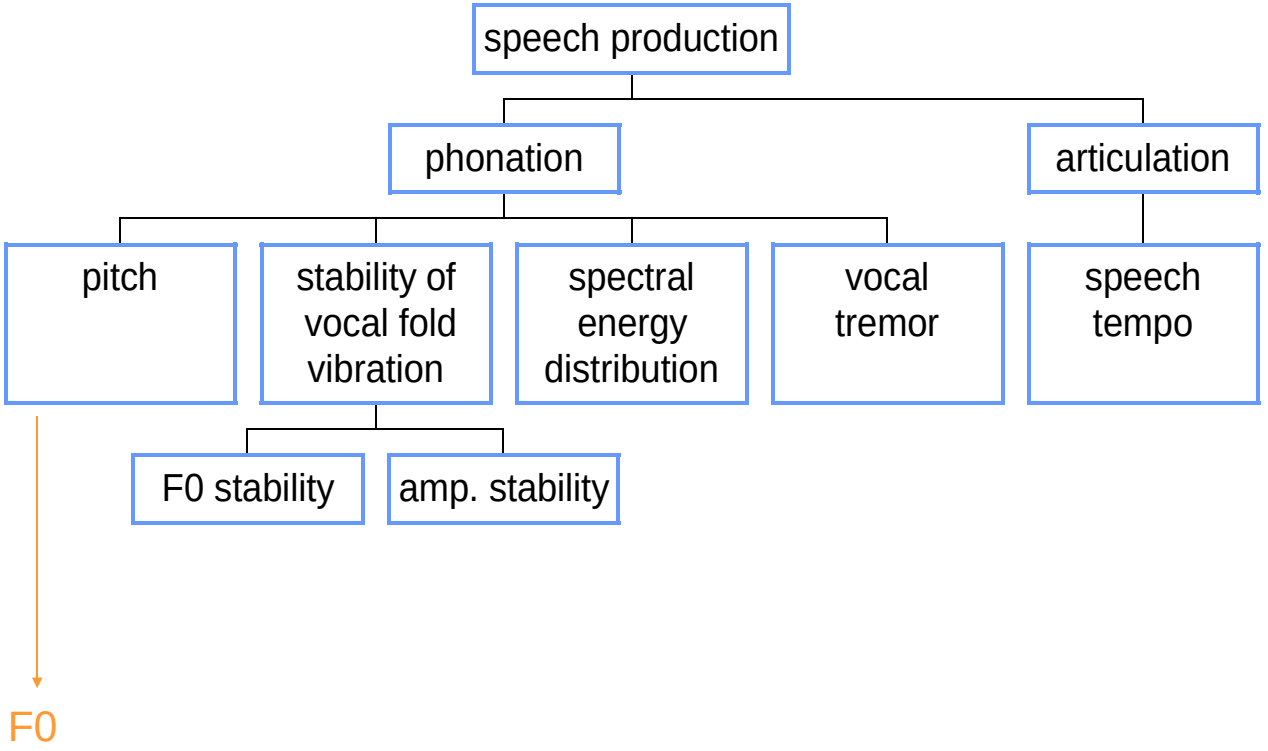
Summary:

- the most accurate age estimations result from spontaneous speech and read speech
- accuracy of age estimations on vowels differs according to
  - vowel type
  - onset criterion: vowels containing onset are rated more accurately

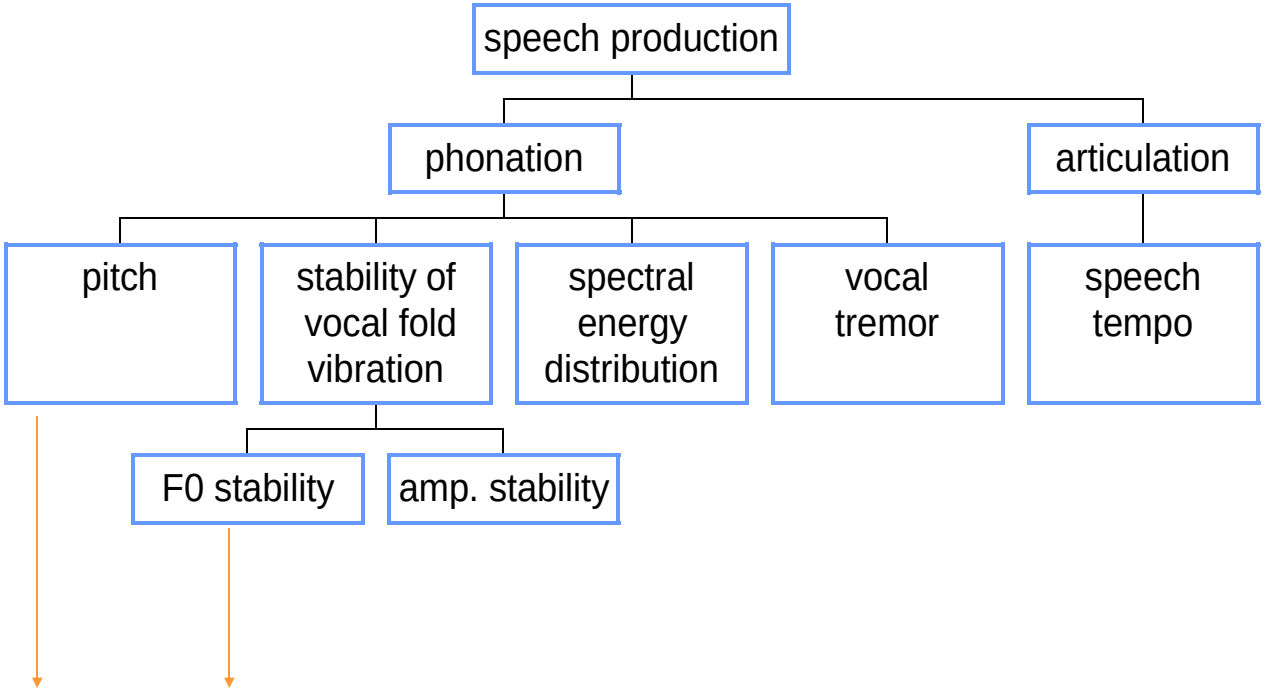
Acoustic Parameters



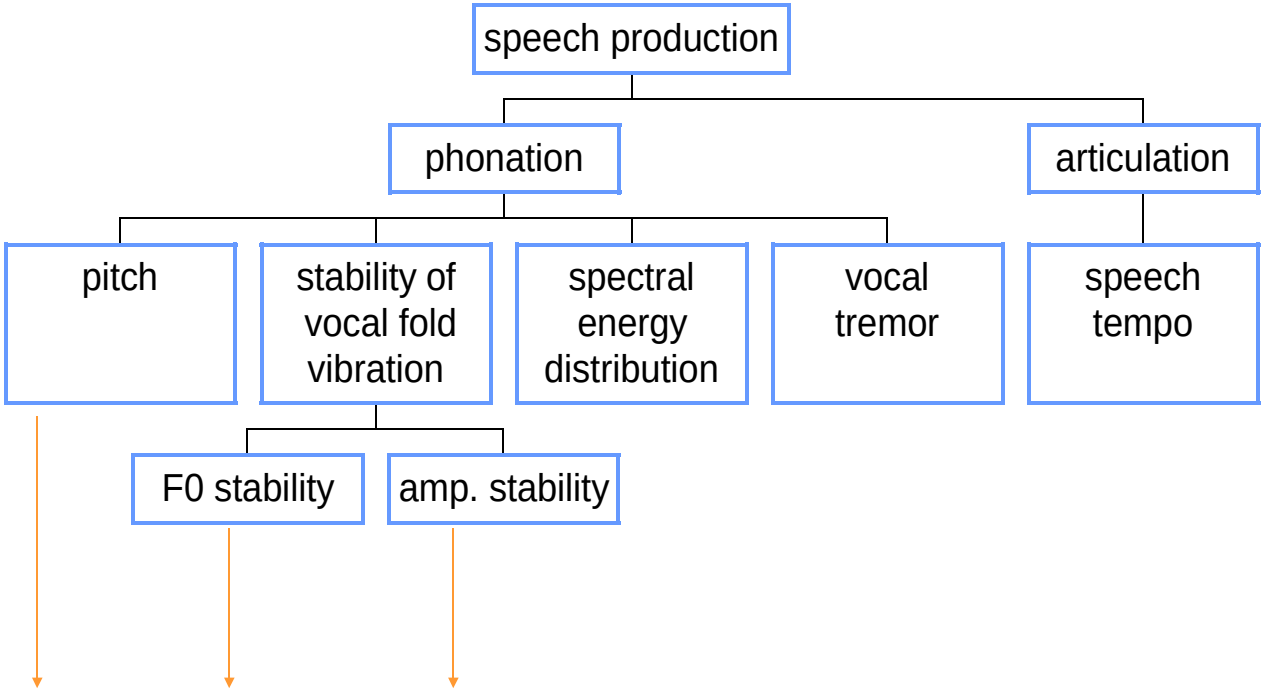
Acoustic Parameters



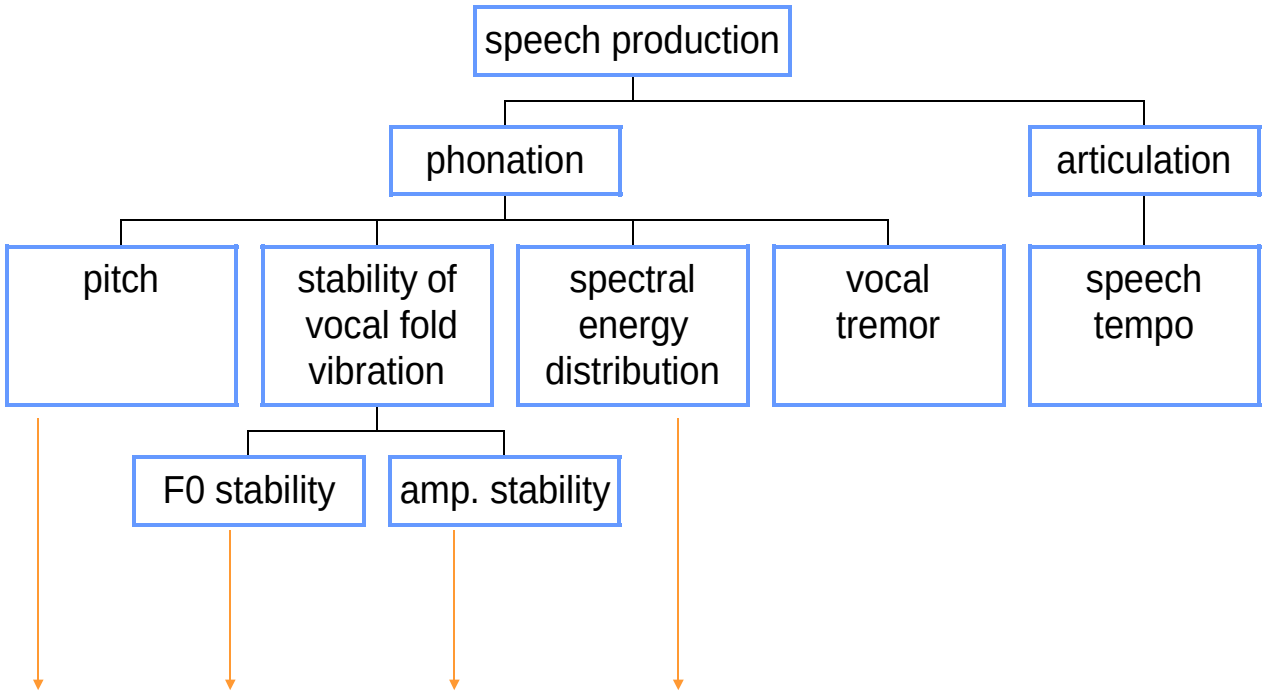
Acoustic Parameters



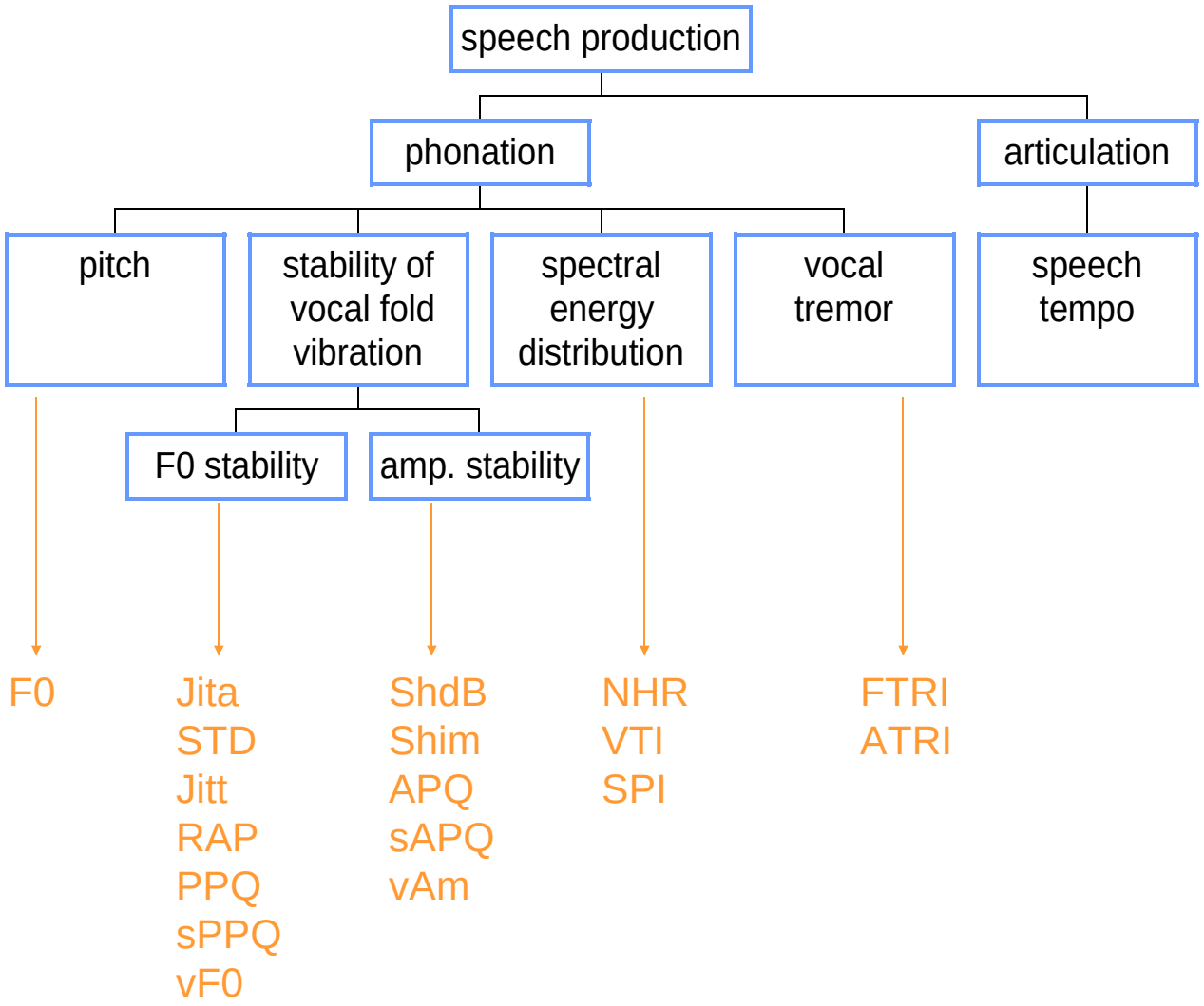
Acoustic Parameters



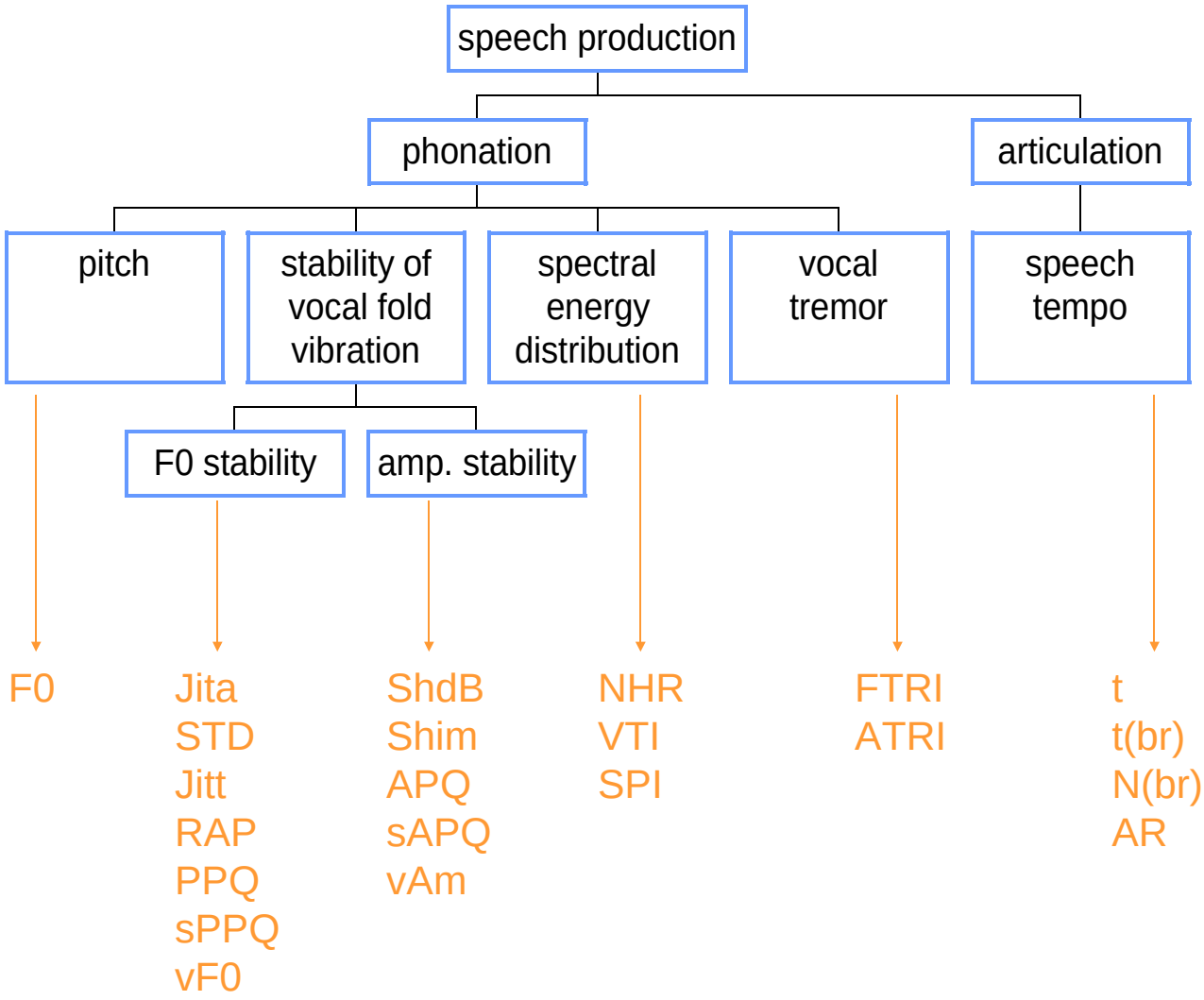
Acoustic Parameters



Acoustic Parameters



Acoustic Parameters



The estimated age values are generally stronger correlated with the acoustic measures than the chronological age

F0-measurements confirm former findings:

- F0 is steadily decreasing with increasing age in women’s voices
- not correlated to age in /i/ and /u/

  ⇒ intrinsic pitch

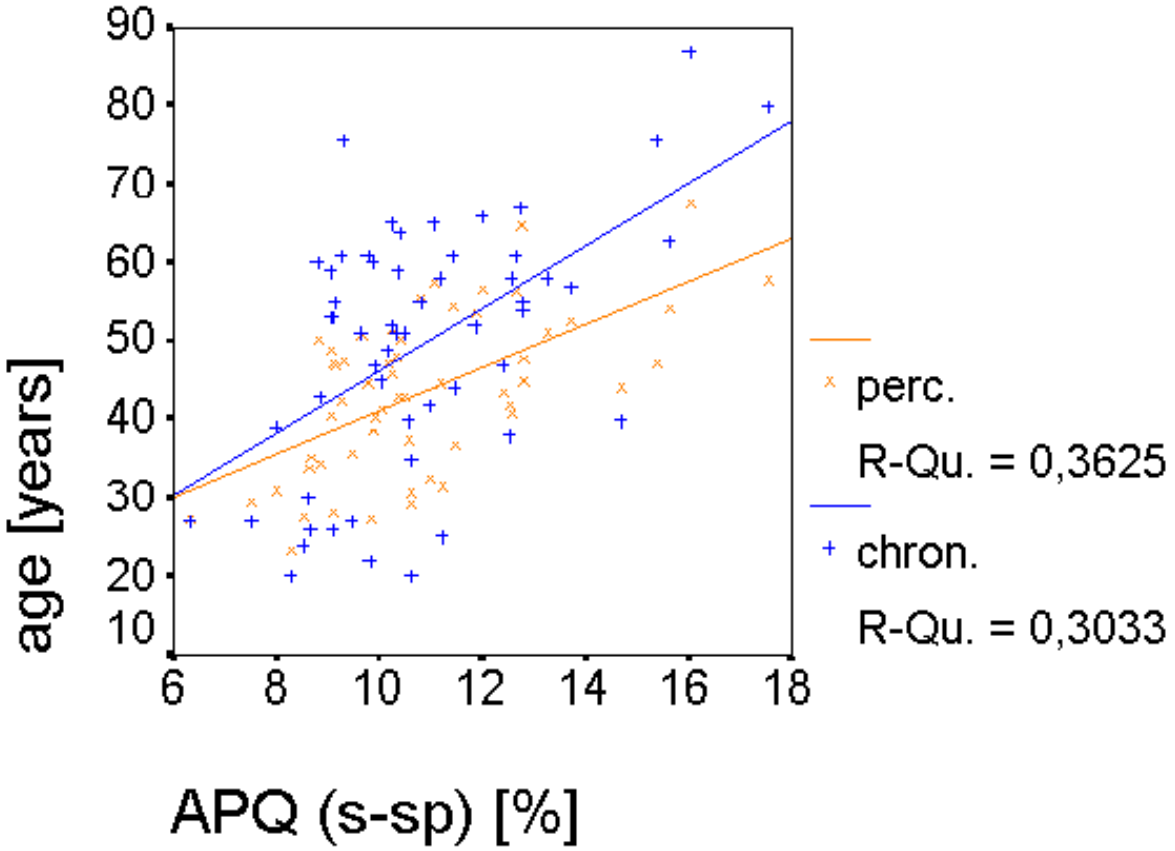
F0 perturbation measures:

- minor respectively sporadic correlations with age (compared to amplitude perturbation)

⇒ F0 perturbations are rather related to physical fitness

Amplitude perturbation measures:

APQ of spontaneous speech is the best acoustic measure of age in this study



Amplitude perturbation measures:

- The strongest relations of amplitude perturbation measures and age are achieved with a smoothing factor of 5 and 55 cycles (APQ and sAPQ)

  ⇒ AMP SD and shimmer less correlated

- relation of amplitude perturbation and age can not be found in read speech and in /i/ and /u/ vowels

Spectral energy distribution:

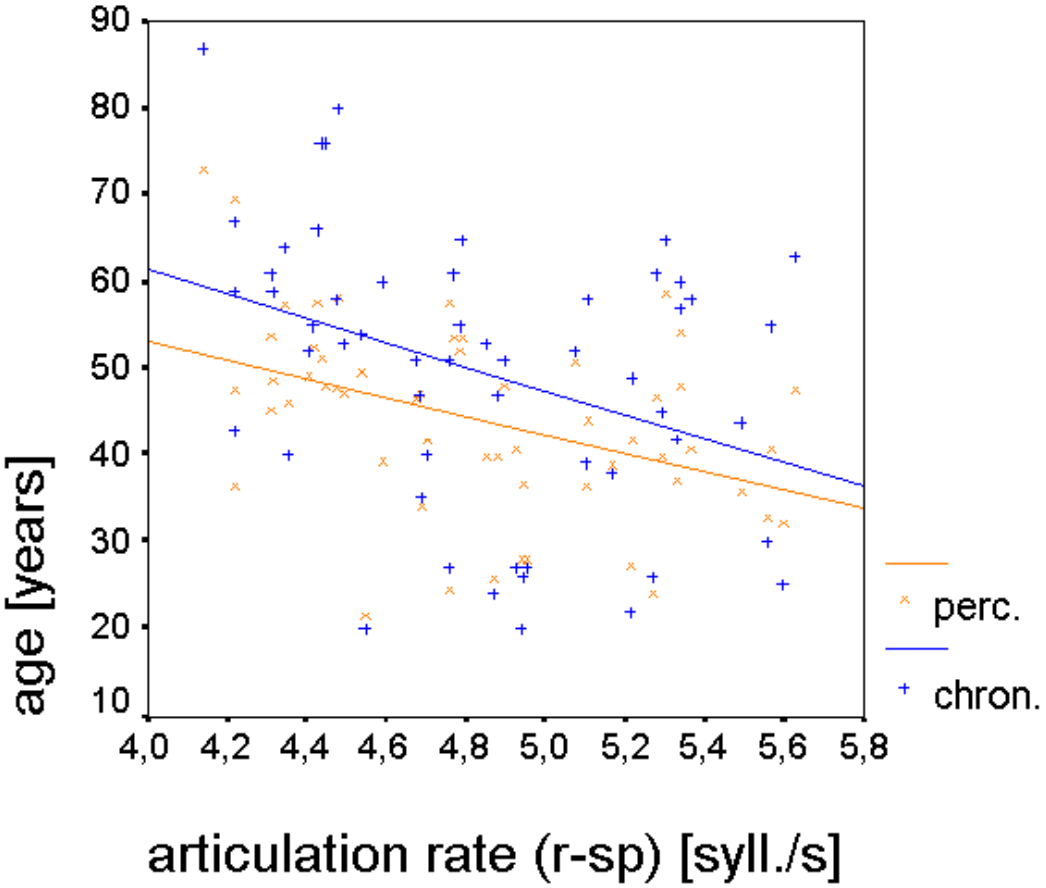
- SPI (soft phonation) correlates with perceived age in /a/ vowels
- NHR (spectral noise) shows only faint correlations with perceived age in /a/ vowels
- VTI (breathiness) is not correlated to age

Vocal Tremor:

- FTRI is increasing with age
  - more reliably than other measures
  - in all sustained vowels
  - but not in read and spontaneous speech
- ATRI is not correlated to age

Speech Tempo:

AR of read speech is correlated with age



Acoustic correlates of age:

- Amplitude perturbation quotient, best from spontaneous speech samples
- Frequency tremor intensity index
- Average fundamental frequency

Indirectly correlated:

- frequency perturbation – fitness
- spectral noise – fitness
- speech tempo – cognitive performance

Relevance of vowel onset

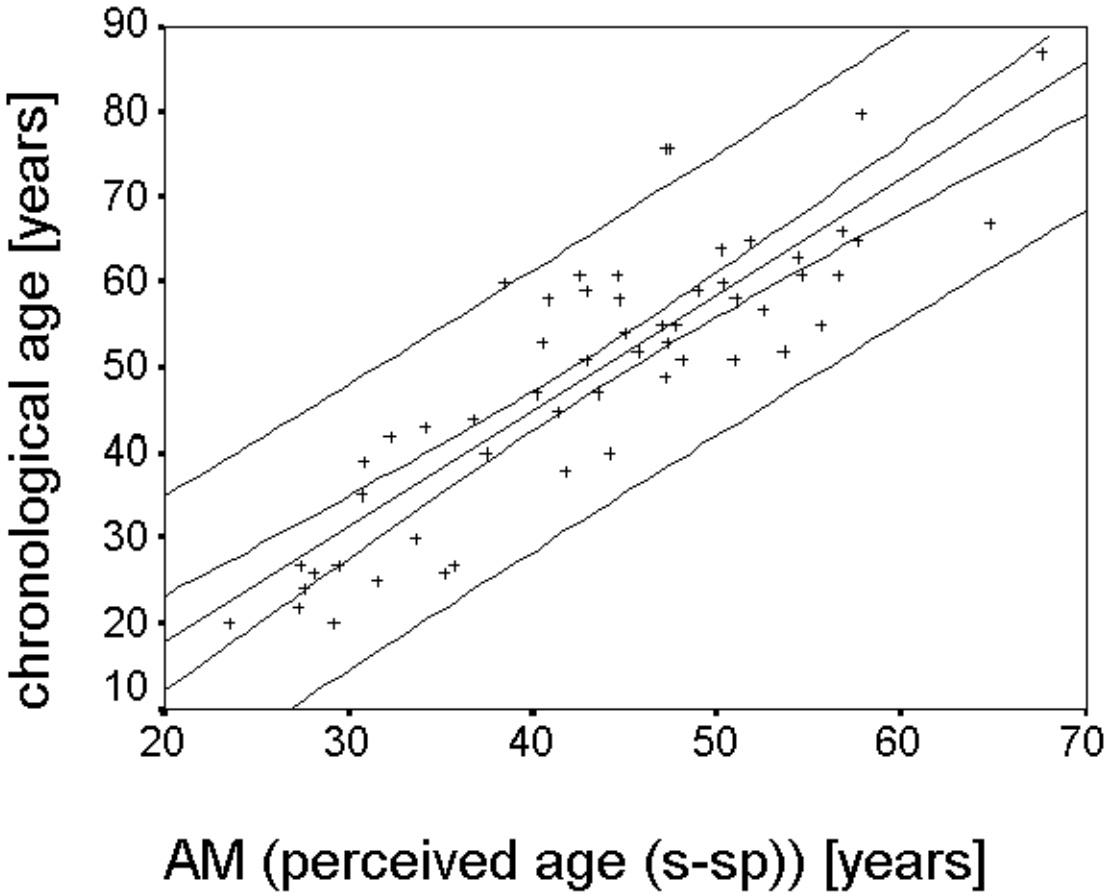
Questions

brueckl@kgw.tu-berlin.de
www.kgw.tu-berlin.de/KW/

Implications

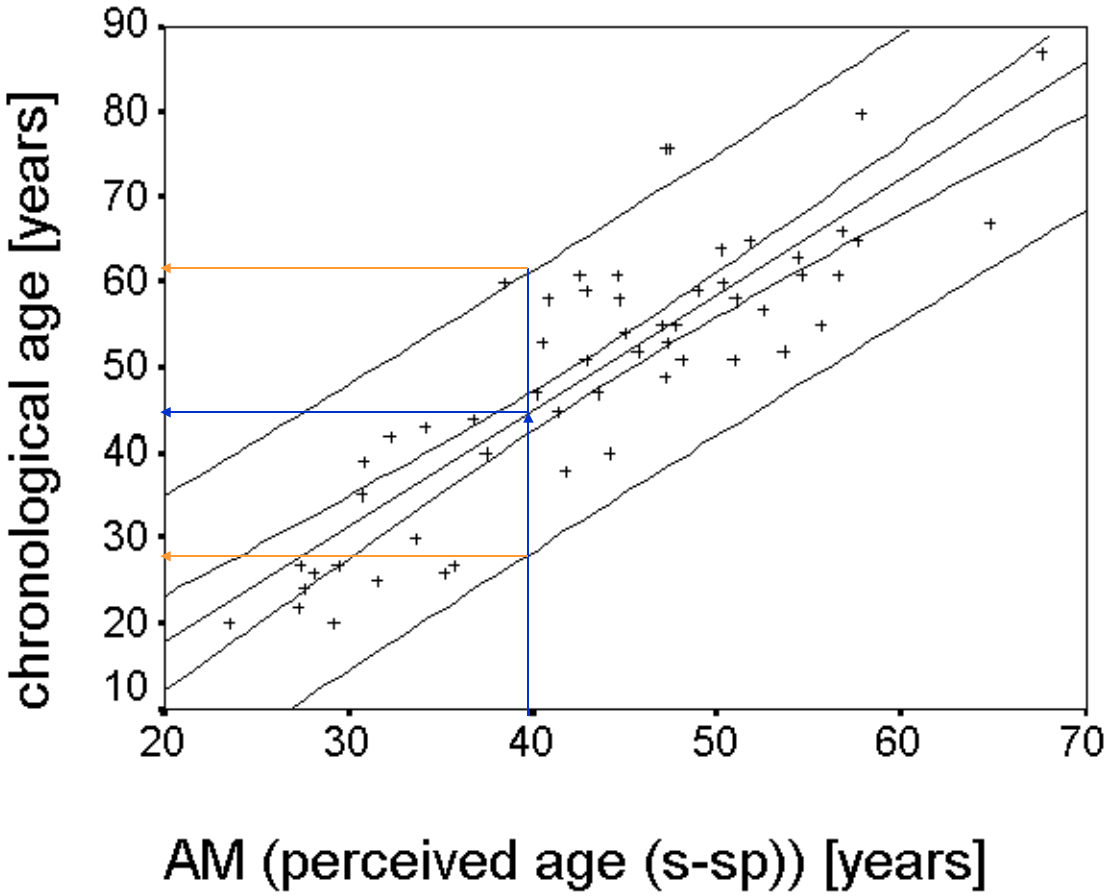
- Are the found acoustic correlates of age perceptively relevant? – synthesis
- Why do amplitude perturbation measurements of spontaneous speech correlate to age? – apply measure on segmented speech
- What measures the FTRI? Can it be improved? – reproduce and alter algorithm
- How is age information decoded in vowel onset? – analysis in different spectral bands

Accuracy of Perception



chron.age = 1.37(perc.age(s-sp)) – 9.63

Accuracy of Perception



chron.age = 1.37(perc.age(s-sp)) – 9.63

Acoustic Correlates

- The estimated age values are generally stronger correlated with the acoustic measures than the chronological age
- multiple regression explains up to
  - 47% of the variance of chronological age and
  - 40% (corrected? R^2) of the variance of perceived age (spontaneous speech)

Read Text

Ich bin zuerst einmal nur geradeaus gegangen. Und dann an der fünften Ampel rechts in die Grabenstraße rein. Die heißt übrigens nach einem halben Kilometer Steinmetzstraße. An der nächsten Ecke bin ich links in die Helenenstraße abgebogen und kurz danach gleich wieder links in die Schloßstraße – ach nein, falsch, da musste ich ja rechts in die Königsberger Straße. Dann lief ich am Schwimmbad vorbei bis zur Überführung – wie Du es mir gesagt hast.

(Example: )

Described Picture

W. E. Hill’s „My wife and my mother-in-law“, demonstrating perceptual ambiguity

(Example:)

Defining Vocal Age

Vocal Aging is the process of the long-term alteration of the biological subsystem speaking apparatus.

Vocal Age is the sum of information in the acoustical signal on a specific state during the process of vocal aging.