The Formantive Years: Vowel Change in a Longitudinal Study of LDS Talks

O'Reilly Miani
ormiani95@gmail.com

Colin Wilson
colin@cogsci.jhu.edu

Johns Hopkins University

NWAV46 November 4, 2017 UW-Madison

How do vowels change during adulthood?

Physiology of laryngeal and supralaryngeal aging

Ossification/calcification of cartilages, vocal fold shortening, change in epithelial thickness

Craniofacial growth, muscle atrophy, degeneration of the temporomandibular joint

---

- How are vowel acoustics affected by physiol changes?
- Are the effects similar across (~matched) speakers?
- Do the effects degrade vowel contrast/dispersion?

How do vowels change during adulthood?

Global slowing of speech rate

(Mysack & Hanley 1958; Ryan & Burk 1974; Smith et al. 1987; Westrop 2000)

Lowering of f0 (esp. women), many other voice source changes (e.g., increased jitter, shimmer)

(Ptacek et al. 1966; Benjamin 1981; Mueller et al. 1984; Russell & Penny 1995, Ramig et al., 2001; Xue & Deliyski 2001; Gorham-Rowan & Laures-Gore 2006)

Formant lowering and vowel centralization

(Endres et al. 1971; Linville 1987; Rastatter & Jacques 1990; Linville & Rens 2001; Xue & Hao 2003; Reubold et al. 2010; Romensko et al. 2016; Vorperian et al. 2016)

(General reviews of age-related phonetic changes: Schötz 2007, Torre & Barlow 2009; Hazan 2017)

Case study

White educated/professional males in Utah

(phonetics and phonology of Utah dialect: Helqvist 1970; Di Paolo & Faber 1990; Di Paolo 1992; Lillie 1998; Bowie 2003, 2008; Reeves 2009)

Longitudinal data covering 30+ yrs per speaker

(real-time phonetics: Harrington et al. 2000; Boberg 2004; Sankoff et al. 2006; Pope et al., 2007; Reubold et al. 2010, 2015; Gahl et al. 2012; Hunter et al. 2012; José & Stuart-Smith 2014; Bowie 2015; Harrington & Rebuild 2015; Stanley 2015; Yu et al. 2015; Stanley & Renwick 2016; Fouquet et al. 2016; Rathcke et al. 2017)

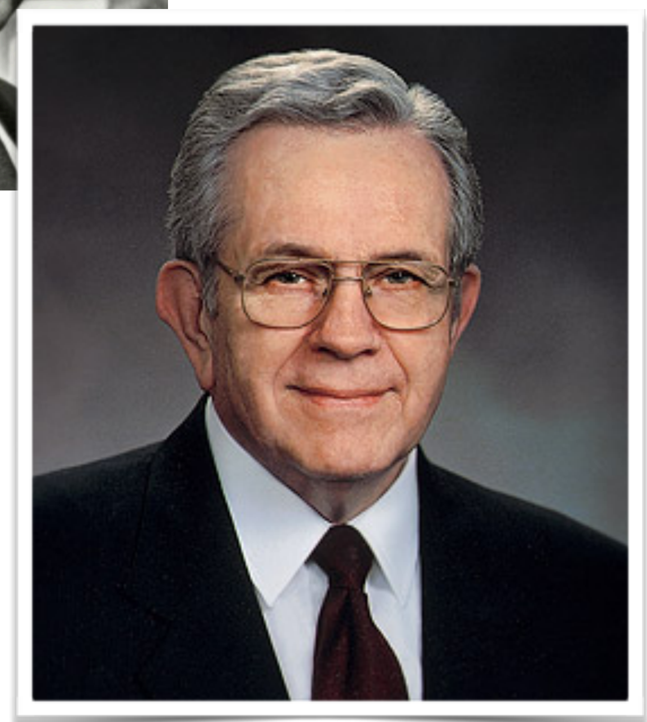
Prepared religious talks delivered at biannual General Conference of the LDS (Mormon) Church

(religious speech: Crystal 1976; Baker & Bowie 2010; Castro et al. 2010; Bedford 2013 — other studies of this speech style?)

Thomas S. Monson

- August 1927 - present
- Born Salt Lake City, UT
- Business background
- LDS general authority since 1963
- Current President of the Church

Boyd K. Packer



- September 1924 - July 2015
- Born Brigham City, UT
- Educator by profession
- LDS general authority since 1970
- Former President of the Quorum of the Twelve Apostles

Additional speakers

Gordon B. **Hinckley** (Hunter et al. 2012, Ferguson et al. 2013)
1910 - 2008, UT, journalism, 37 yrs of General Conference talks

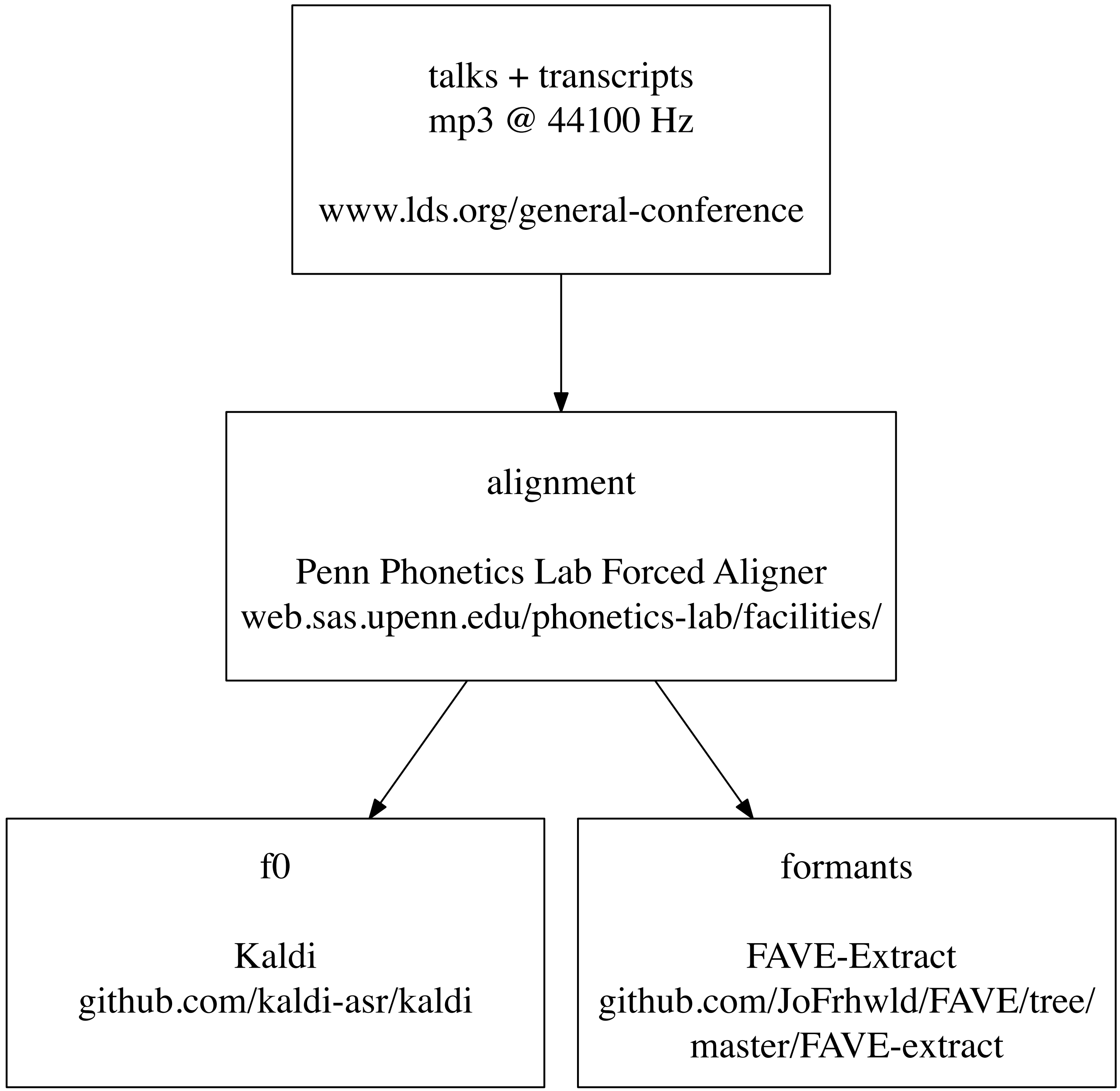
L. Tom **Perry** (Stanley 2015, Stanley & Renwick 2016)
1922 - 2015, UT, businessman, 43 yrs of talks

Russel M. **Nelson**
1924 - present, UT, cardio-thoracic surgeon, 33 yrs of talks

M. Russell **Ballard**
1928 - present, UT, real estate and investment, 32 years of talks

Robert D. **Hales**
1932 - 2017, NY, business executive, 32 yrs of talks

Dallin H. **Oaks**
1932 - present, UT, lawyer and judge, 33 yrs of talks


talks + transcripts
mp3 @ 44100 Hz
www.lds.org/general-conference
alignment
Penn Phonetics Lab Forced Aligner
web.sas.upenn.edu/phonetics-lab/facilities/
f0
Kaldi
github.com/kaldi-asr/kaldi
formants
FAVE-Extract
github.com/JoFrhwld/FAVE/tree/
master/FAVE-extract

Vowel measures

Fundamental frequency (f0)
Median of central three f0 values as measured by compute-kaldi-pitch-feats

Vowel formants (F1 and F2)
Measured by FAVE-Extract (near vowel midpoint)

- Only stressed vowels included in analysis
- Formant outliers determined from ellipse fit to each vowel for each speaker (collapsed across ages)
- Duration outliers ≥ 750 ms
- Data aggregated into half-decade bins for plots

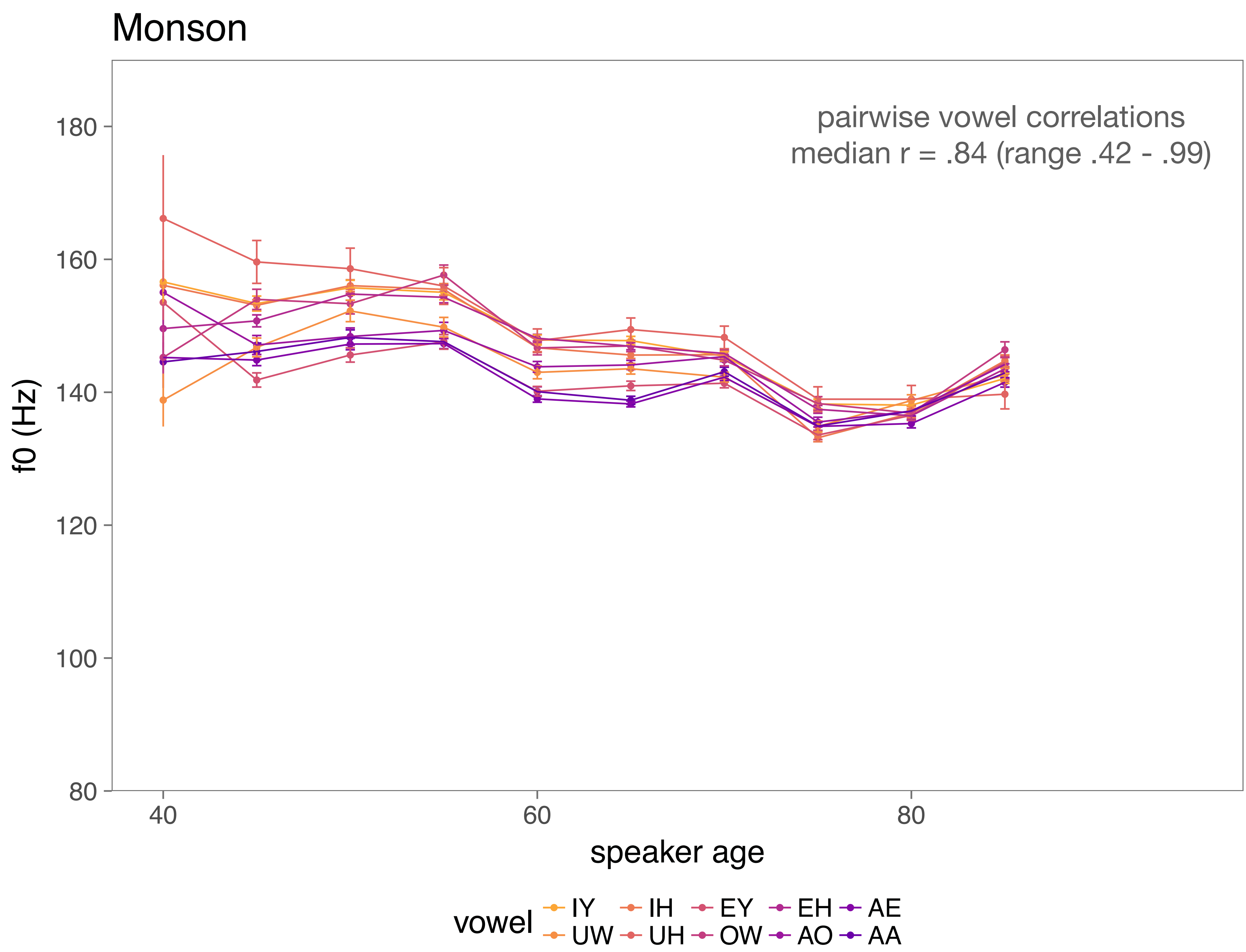
Vowel corpora

Monson

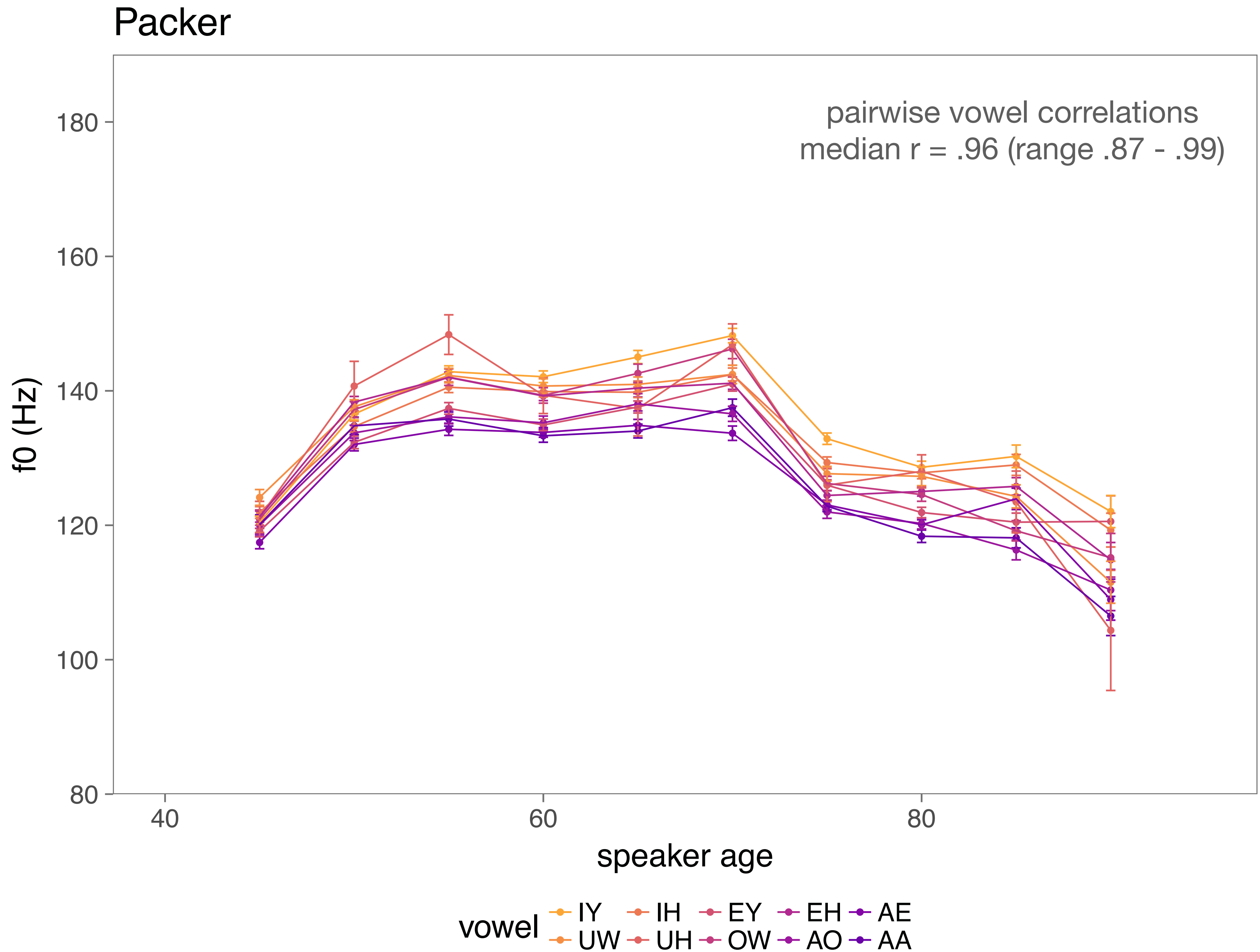
- 179 talks over 46 years, 54+ hours of speech
- Age 44-89 (1971-2016)
- Talk durations: 3-27 min (median = 19 min)
- Stressed monophthongs: 169,749

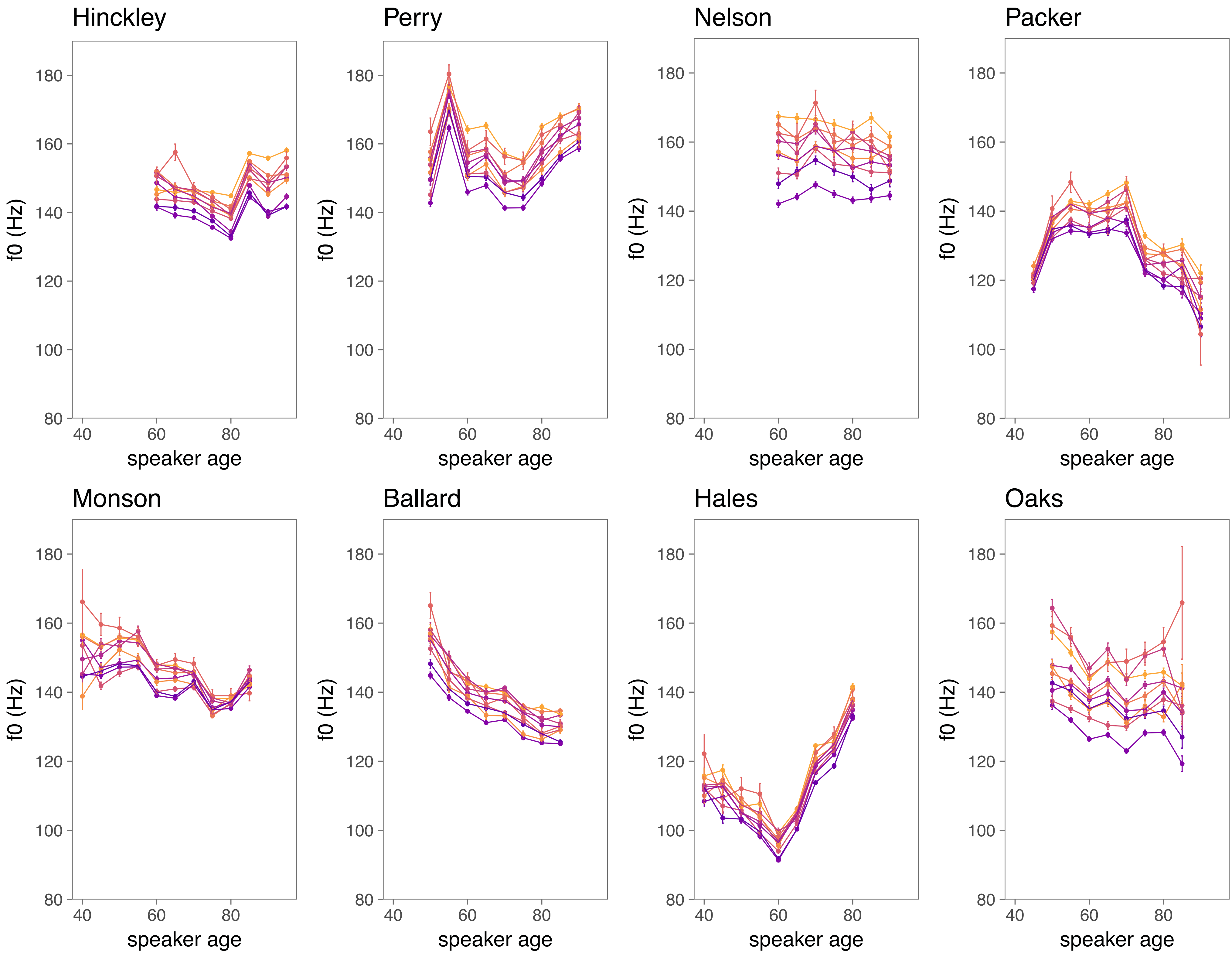
Packer

- 90 talks over 44 years, 25+ hours of speech
- Age 47-90 (1971-2014)
- Talk duration: 13-20 min (median = 16 min)
- Stressed monophthongs: 93,774



Monson
pairwise vowel correlations
median r = .84 (range .42 - .99)
f0 (Hz)
180
160
140
120
100
80
40
60
80
speaker age
vowel
IY
UW
IH
UH
EY
OW
EH
AO
AE
AA


Packer
pairwise vowel correlations
median r = .96 (range .87 - .99)
f0 (Hz)
180
160
140
120
100
80
40
60
80
speaker age
vowel
IY
UW
IH
UH
EY
OW
EH
AO
AE
AA


Hinckley
Perry
Nelson
Packer
Monson
Ballard
Hales
Oaks
f0 (Hz)
speaker age
80
100
120
140
160
180
40
60
80

Statistical analysis of f0

Fixed	age_bin * vowel_height (high, mid, low)
Random	(1 | talk_id) + (1 | word)

Monson

age_bin	-0.38 (*t* = -5.98)		
high	1.67 (*t* = 6.60)	age_bin:high	-0.10 (*t* = -8.29)
mid	-0.38 (*t* = -1.85)	age_bin:mid	0.02 (*t* = 1.88)

Packer

age_bin	-0.27 (*t* = -3.40)		
high	3.94 (*t* = 13.57)	age_bin:high	0.08 (*t* = 6.56)
mid	-1.18 (*t* = -5.14)	age_bin:mid	-0.05 (*t* = -4.91)

Dependent variable and numerical age predictor mean-centered
Weighted effect coding used for all categorical predictor(s)

Summary of f0 changes

Both speakers show a **negative linear effect of age** on vowel f0, approximately constant across vowel categories

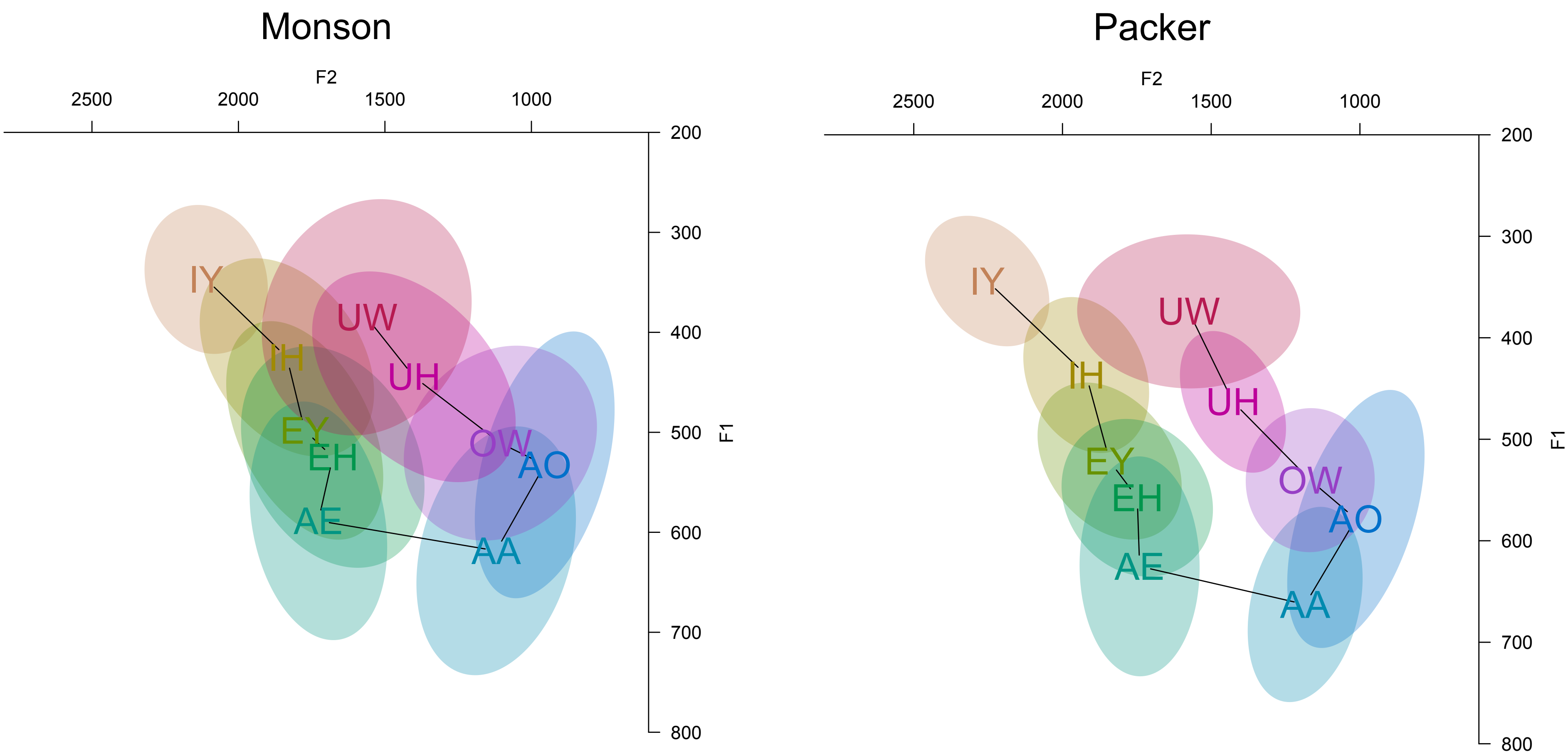
- But change with age is clearly more monotonic for some speakers (Monson, Ballard) than for others (Packer, Hales)

- Majority of speakers significantly better fit by models incorporating higher-order age terms (at least cubic)

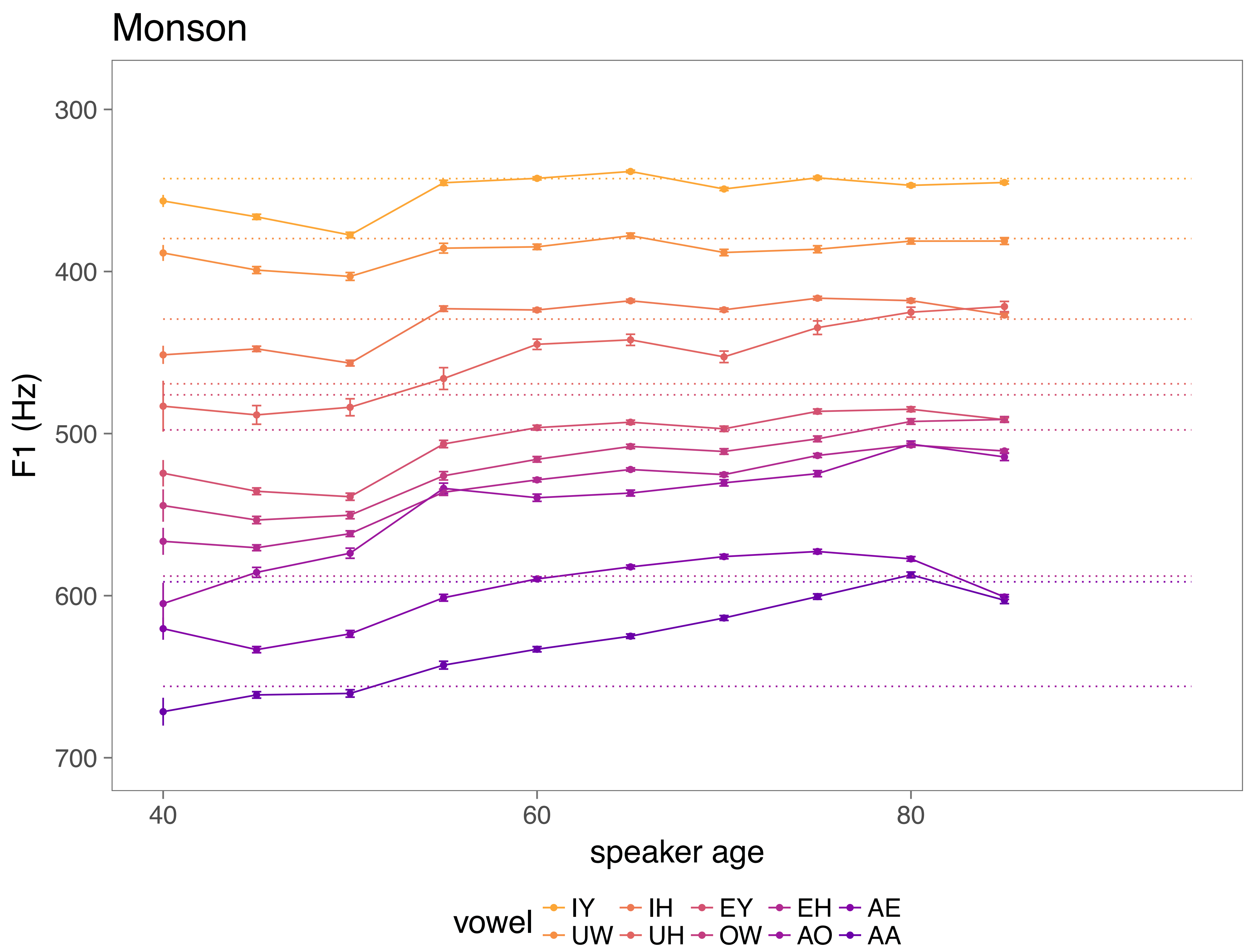
Both speakers also show the expected **intrinsic f0** pattern, high vowels having higher f0 than low vowels (e.g., Whalen & Levitt 1995)

- Preservation of intrinsic f0 indicates relatively constant phonetic targets -and- ‘automatic’ effects across ages

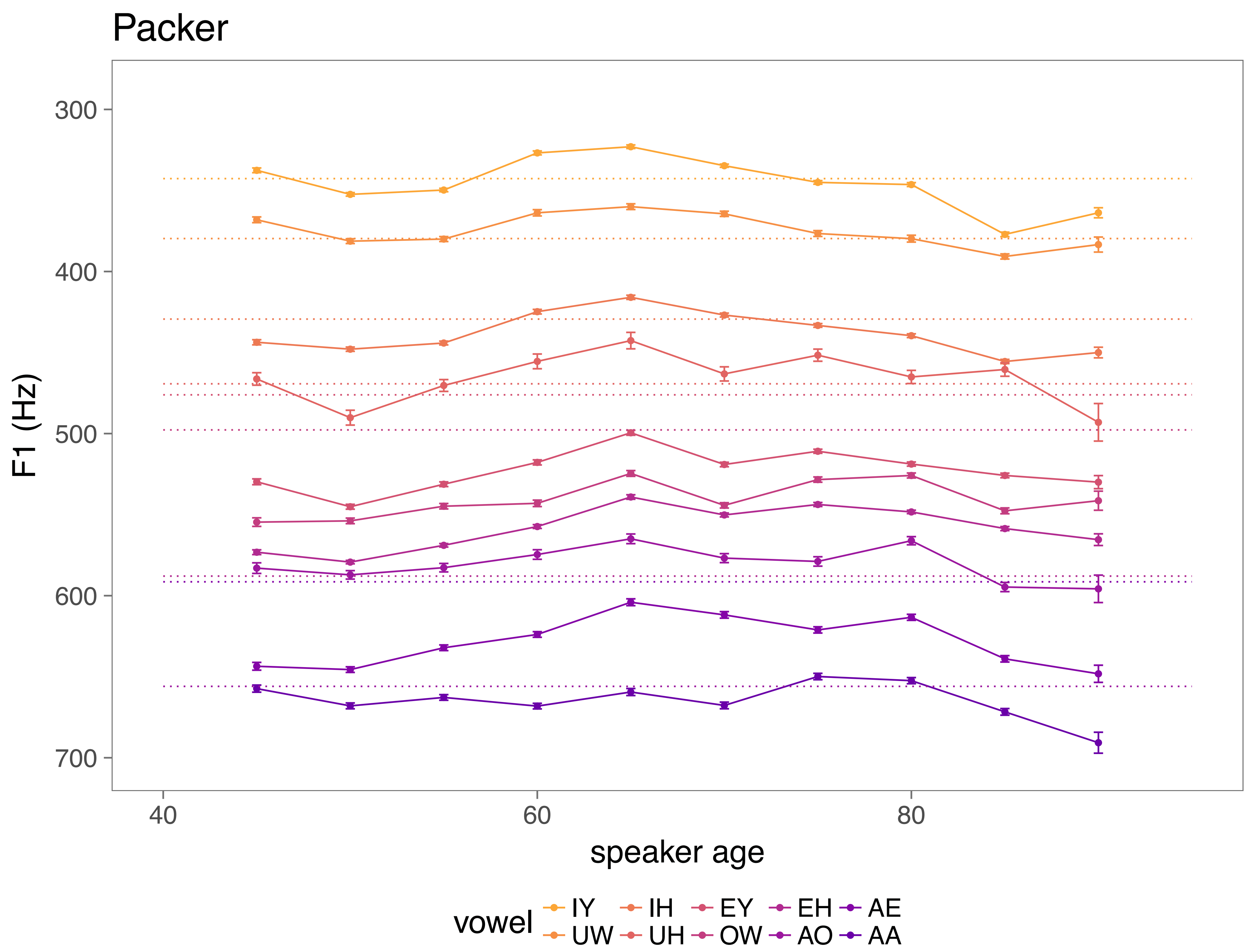
F1 × F2 vowel spaces



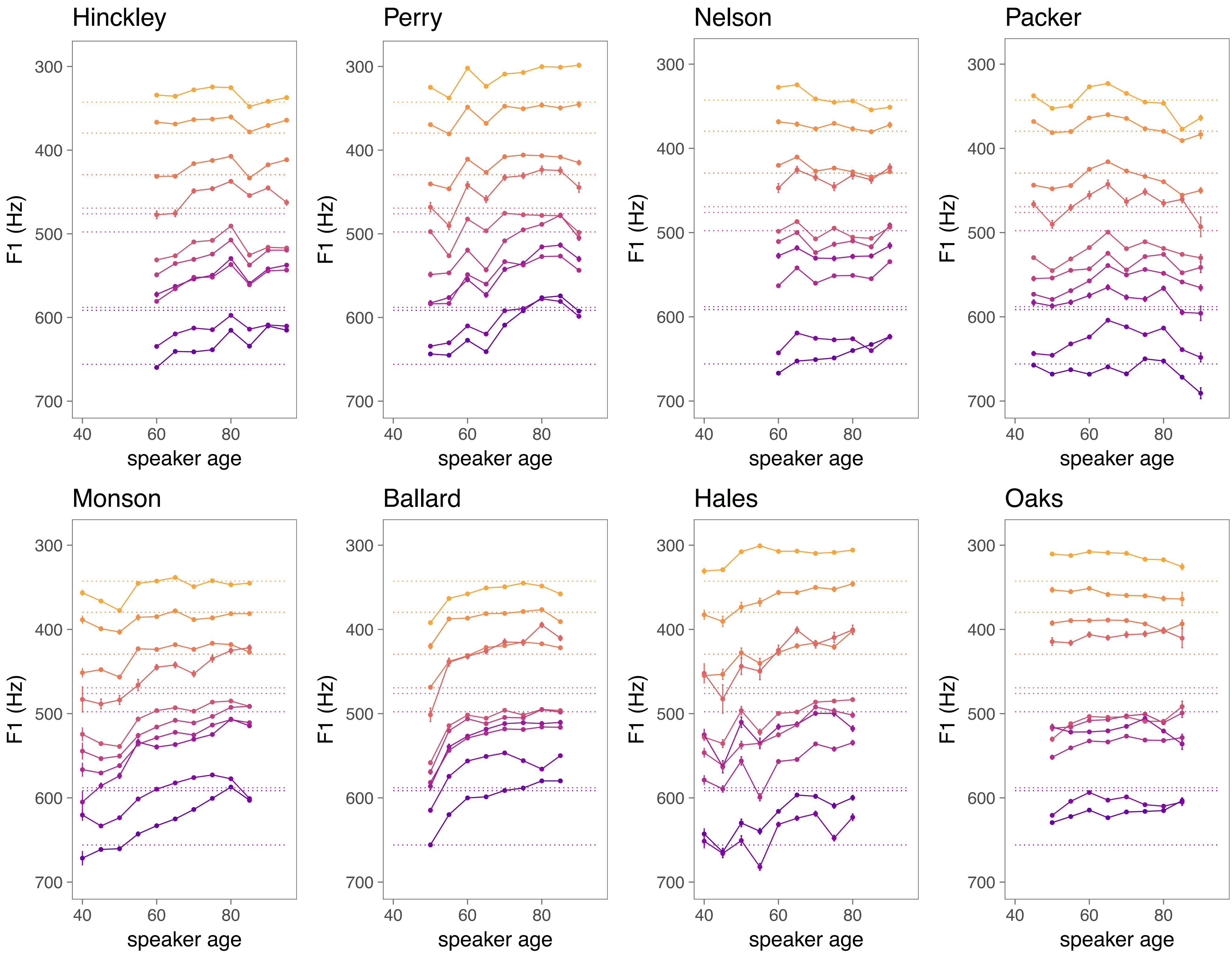
Data collapsed over all ages for each speaker



Monson
F1 (Hz)
300
400
500
600
700
40
60
80
speaker age
vowel
IY
UW
IH
UH
EY
OW
EH
AO
AE
AA



Packer
300
400
500
600
700
F1 (Hz)
40
60
80
speaker age
vowel
IY
UW
IH
UH
EY
OW
EH
AO
AE
AA


Hinckley
Perry
Nelson
Packer
Monson
Ballard
Hales
Oaks
F1 (Hz)
speaker age
300
400
500
600
700
40
60
80

Statistical analysis of F1

Fixed   age_bin * vowel_height (low, mid, high)
Random (1 | talk_id) + (1 | word)

Monson

age_bin	-0.01 (*t* = -13.74)		
low	0.86 (*t* = 123.99)	age_bin:low	-0.001 (*t* = -8.07)
mid	0.23 (*t* = 52.07)	age_bin:mid	-0.002 (*t* = -13.85)

Packer

age_bin	n.s.		
low	1.01 (*t* = 115.94)	age_bin:low	-0.001 (*t* = -3.82)
mid	0.31 (*t* = 57.27)	age_bin:mid	-0.002 (*t* = -17.73)

F1 in Bark. Full analysis also included fixed effects of vowel backness and tense vs. lax

Summary of F1 changes

Monson shows a **negative linear effect of age** on vowel F1 (i.e., vowel raising), greater effect on non-high vowels

Packer displays a more complicated **vowel raising then lowering** pattern over age, with non-high vowels raising overall

- Some of the additional speakers have across the board vowel raising (Ballard), others have raising-falling (Perry)

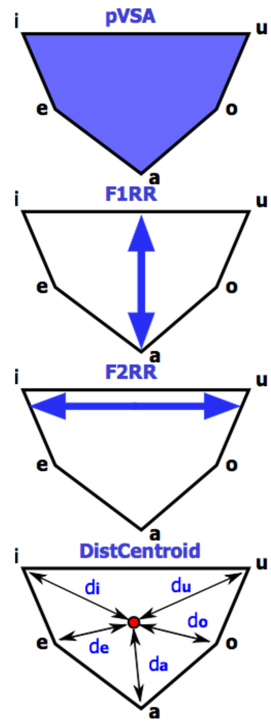
Both speakers also show expected effects of vowel height, backness (F1 lower for front), and tenseness (F1 lower for tense)

- Are degrees of vowel contrast/dispersion preserved?

Acoustic measures of vowel contrast/dispersion

Holistic measures

(figure from Audibert et al. 2015)



pVSA: *Vocalic Space Area (pentagon).* Surface of vocalic shape defined by the /i,e,a,o,u/ pentagon in F1/F2 plane. Vowel positions are computed as weighted centroids of measured tokens.

F1RR: *First formant range ratio.* Designed to measure the ambitus in the F1 dimension (jaw/tongue height), see [22].

F2RR: *Second formant range ratio.* Designed to measure the ambitus in the F2 dimension (front-back tongue dimension and/or lip rounding), see [22].

DistCentroid: Vowel Distance to Centroid A common metric computing the average distance of the 5 vowels' centroids to the overall centroid of the speaker's vowel space.

Vowel-pair measure

Bhattacharyya distance between $\mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$, $\mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$

$$D_B = \frac{1}{8}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) + \frac{1}{2}\log\left(\frac{|\boldsymbol{\Sigma}|}{\sqrt{|\boldsymbol{\Sigma_1}||\boldsymbol{\Sigma_2}|}}\right)$$

Acoustic measures of vowel contrast/dispersion

Holistic measures are highly correlated and generally indicate **reduced contrast** over age

- Vowel dispersion is not perfectly retained under physiol changes to the vocal tract in adulthood

Bhattacharyya distance decreases (i.e., overlap increases) over half decades for certain vowel pairs

Monson			/u/-/ʊ/	/u/-/o/	/o/-/ɔ/,/a/
Packer	/i/-/ɪ/	/i/-/e/		/u/-/o/	

Summary

Vowel f0 and formants (F1, F2) generally **lower** over the course of middle/late adulthood

- Differences in degree of change and monotonicity across matched male speech samples
- Approximately parallel effects on all vowels within a speaker, perhaps greater F1 raising for non-high

Perfectly parallel shifts would preserve degree of vowel dispersion despite age-related physiol changes

- Observe general trend of slowly shrinking vowel space
- Distance between certain vowel pairs decreases

Applications

Middle/late adulthood baseline for acquired speech disorders

(Liu et al. 2005; Tjaden & Rivera 2005; Sapir et al. 2010; Fougeron & Audibert 2011; Kim et al. 2011; Skodda et al. 2010; Soares 2011; Audibert & Fougeron 2012; Lansford & Liss 2014ab; Luan et al. 2014; Whitfield & Goberman 2014; Connaghan & Patel 2016; den Ouden et al. 2017; Fletcher et al. 2014, 2017; Lee et al. 2017)

Vowel overlap/dispersion measures and vowel intelligibility

(Wassink 2006; Ferguson & Kewley-Port 2007; Neel et al. 2008; Karlsson & Doorn 2012; McCloy et al. 2012; Ferguson & Quené 2014; Haynes 2014; Nycz & Hall-Leuw 2015; Schulz et al. 2016; Kelley et al. 2017; Story & Button 2017)

Effect of aging on ASR and speaker identification/verification

(Wolters et al. 2009; Vipperla 2010; Kelly 2012; Pellegrini et al. 2013; Kelly et al. 2014; Kelly & Harte 2015)

Acoustic predictors of perceived vocal age

(Schötze 2006; Harnsberger et al., 2008, 2010; Hillenbrand & Clark 2009; Ferguson et al. 2013)

Future directions

Further data cleaning and model cross-validation

Additional speakers and similar talks by women!

Direct tests of age effects on vowel intelligibility for human (and automatic) listeners

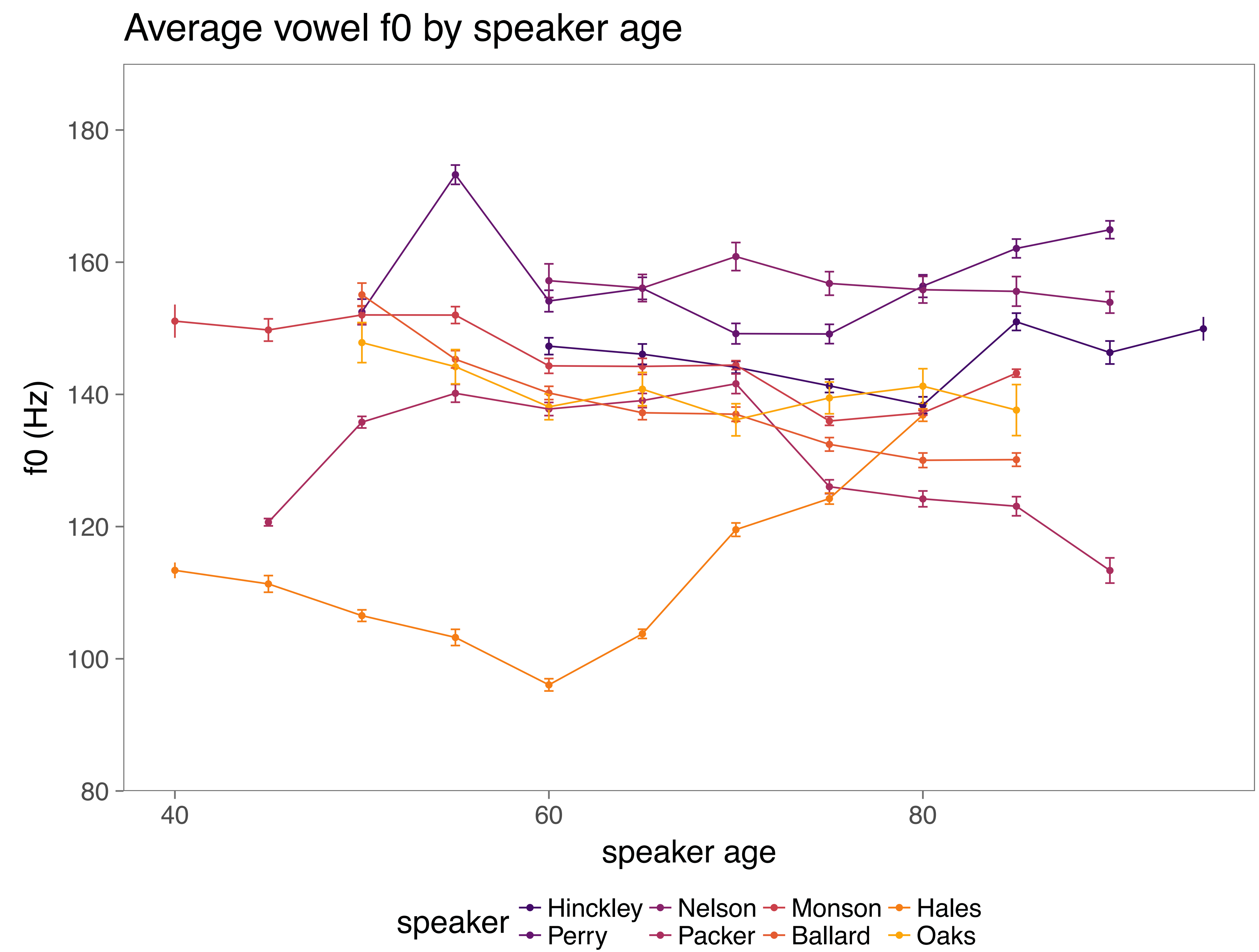
Study change of other phonetic variables, such as stop Voice Onset Time, in middle/late adulthood

(Benjamin 1982; Sweeting & Baken 1982; Thomas 1985; Liss et al. 1990; Brazeal 1992; Morris & Brown 1994; Singh et al. 2016)

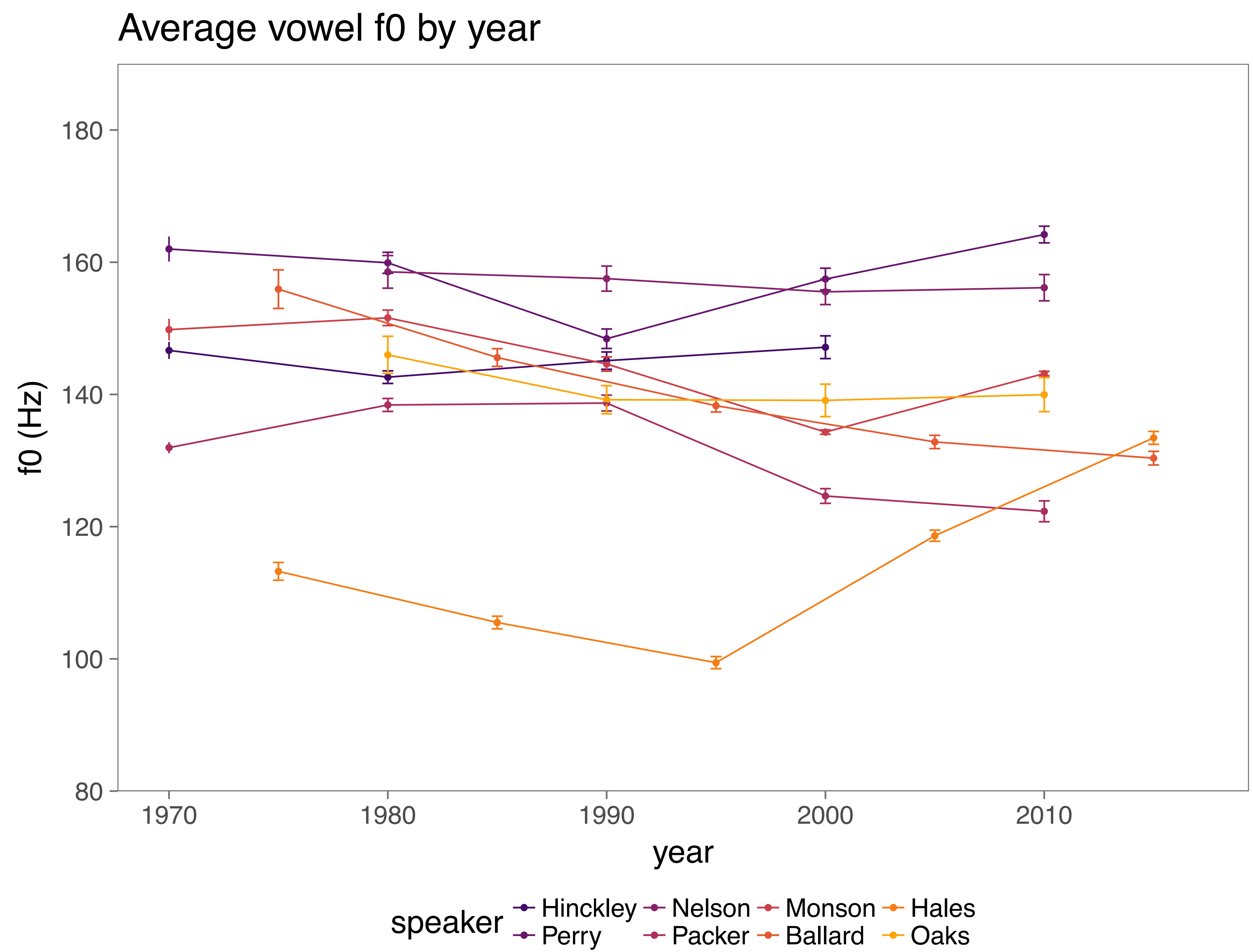
Thank you!

Thanks to Eleanor Chodroff, Coleman Haley, Emma Maxwell, Joseph Stanley, Zhenglong Zhou for helpful comments and questions

Thanks to the LDS Church for making the talks available, and to the developers of all of the open-source tools used in our analysis



Average vowel f0 by speaker age
f0 (Hz)
180
160
140
120
100
80
40
60
80
speaker age
speaker
Hinckley
Perry
Nelson
Packer
Monson
Ballard
Hales
Oaks



Average vowel f0 by year
f0 (Hz)
180
160
140
120
100
80
1970
1980
1990
2000
2010
year
speaker
Hinckley
Perry
Nelson
Packer
Monson
Ballard
Hales
Oaks

References xxx preliminary

Bowie, D. 2003. Early development of the card-cord merger in Utah. American Speech 78(1). 31–51.

Bowie, D. 2008. Acoustic characteristics of Utah's CARD-CORD merger. American Speech 83(1). 35–61.

Bowie, D. 2010. The Ageing Voice: Changing Identity Over Time. Language and Identities, 55–65. Edinburgh: Edinburgh University Press.

Endres, W., Bambach, W., & Flösser, G. (1971). Voice spectrograms as a function of age, voice disguise, and voice imitation. *JASA* 49 (4) , 1842-1848.

Harrington, J., S. Palethorpe & C. Watson. 2000. Monophthongal vowel changes in Received Pronunciation: An acoustic analysis of the Queen's Christmas broadcasts. *JIPA* 30(1-2). 63–78. (20 April, 2015).

Jurik, A. (1984). Ossification and calcification of the laryngeal skeleton," Acta Radiol Diagn., vol. 25, pp. 17–22.

Linville, S. E., & Rens, J. (2001). Vocal tract resonance analysis of aging voice using long-term average spectra. Journal of Voice 15 (3) , 323-330.

Reubold, U., Harrington, J., & Kleber, F. (2010). Vocal aging effects on F0 and the first formant: a longitudinal analysis in adult speakers. Speech Communication 52 (7-8) , 638-651.

Rosenfelder, Ingrid; Fruehwald, Joe; Evanini, Keelan and Jiahong Yuan. (2011). FAVE (Forced Alignment and Vowel Extraction) Program Suite. http://fave.ling.upenn.edu.

Xue, Steve An & Grace Jianping Hao. 2003. Changes in the human vocal tract due to aging and the acoustic correlates of speech production: a pilot study. Journal of Speech, Language & Hearing Research 46(3). 689–701.

How do vowels change during adulthood?

- Age-grading studies
- Longitudinal studies
- Effects of aging on intelligibility
- Baseline for acquired speech disorders