Investigating some Chinese dialects using Functional Data Analysis

John Aston

Department of Statistics

University of Warwick

17 February 2012

Acknowledgements

Joint work with

- Jeng-Min Chiou
  Institute of Statistical Science, Academia Sinica
- Jonathan Evans
  Institute of Linguistics, Academia Sinica
- Pantelis Hadjipantelis
  Complexity Science and Dept of Statistics, University of Warwick

Many thanks to SAMSI and EPSRC for support for the work

Outline

Phonetic Analysis of Pitch

Mixed Models
- Example: Median Analysis of Qiang

Functional Principal Component Analysis

Analysis of Qiang

Analysis of Mandarin

Outline

Phonetic Analysis of Pitch

Mixed Models
  Example: Median Analysis of Qiang

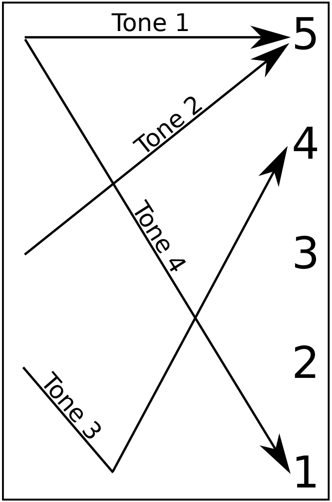
Functional Principal Component Analysis

Analysis of Qiang

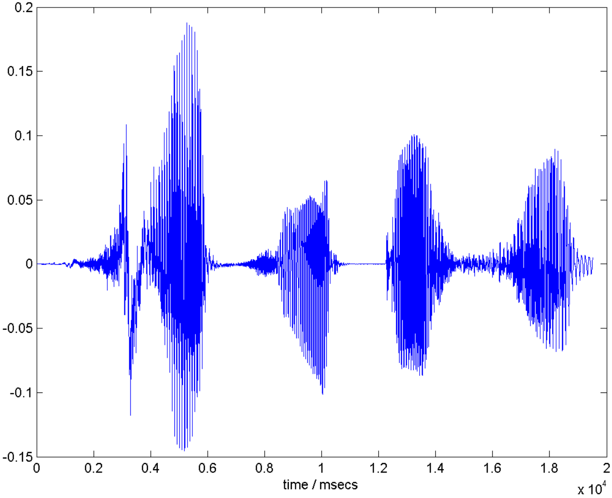
Analysis of Mandarin

Tones in Mandarin

Tone1 媽 "mother"
Tone2 痲 "hemp"
Tone3 馬 "horse"
Tone4 罵 "scold"
Tone5 嗎 question particle

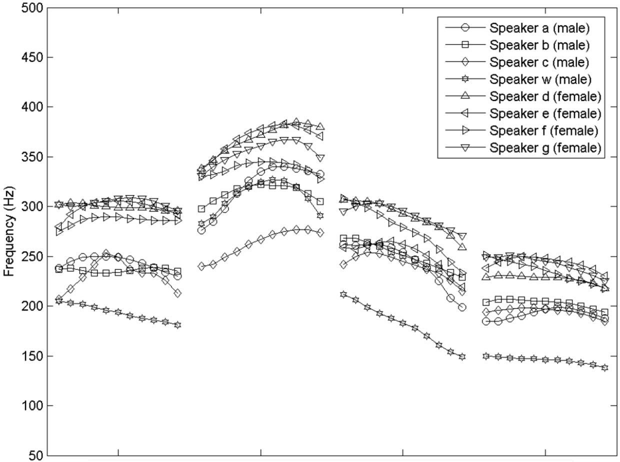


Qiang Speech Signal



Storage Room Door

Example of F0 data



Vowel F0 plot for /tshà tʂú qò qò/, 'storage room door'.

a b c w d e f g

Outline

Phonetic Analysis of Pitch

Mixed Models

Example: Median Analysis of Qiang

Functional Principal Component Analysis

Analysis of Qiang

Analysis of Mandarin

Linear Models

For a

- set of observations y,
- a set of covariates X,

The data is modelled as

$$y = X\beta + \epsilon$$

where

- β is the effect of the covariates,
- ϵ is the error assumed to be distributed $N(0, \sigma^2)$.

Linear Mixed Models

For a

- set of observations y,
- a set of fixed covariates X,
- a set of (mean-zero Gaussian) random covariates Z,

The data is modelled as

$$y = X\beta + Z\gamma + \epsilon$$

where

- β is the effect of the covariates,
- γ is assumed to be distributed $N(0, \sigma_\gamma^2)$
- ϵ is the error assumed to be distributed $N(0, \sigma^2)$.

Parameter Estimation and Inference

- Maximum Likelihood (ML)
  - Estimates of Fixed Effects and Random Effects
  - Biased Estimates of the Random Effect Variances
- Restricted Maximum Likelihood (ML)
  - Estimates of Random Effects
  - Unbiased Estimates of the Random Effect Variances

Procedure

1. ML and Parametric Bootstrap to choose variables in model
2. ReML to find variance estimates in chosen model

Qiang Language and People

Qiang Language and People

Qiang Language and People

Qiang Language and People

Qiang Language

- Tibeto-Burman Language
- Spoken by about 100,000 people
- Two distinct varieties (Northern and Southern)
- Many dialects - the one under investigation is Luobuzhai Qiang (a southern variety)
- Tonal language with 2 known tones, but very little studied
- No written form

Words in Study

No	Form	tones	Glossary
1	/dʐù ˈbè/	LL	star
2	/dʐə́ ˈɕí/	HH	day before yesterday
3	/ˈlí phò gè/	HLL	trumpet
4	/tɕè ˈpjá ʁè gè/	LHLL	corn cake
5	/pù qhà pà (gé)/	LLLH	large intestine
6	/dʐò dʐò gé/	LLH	ruler
7	/pú ʂú ʂtʂà (gè)/	HHLL	youth (n.)
8	/mù tɕhàn ˈthá mí/	LLHH	robber
9	/ˈtshà tʂú qò qò/	LHLL	storage room door
10	/ɕì ˈphú grə̀/	LHL	root fibers
11	/ɕí tʂú ˈpiàn tsə̀/	HHLL	river bank
12	/biá ȵú pì ˈqhuá/	HHLH	female panda
13	/pù qhà ˈpà/	LLL	large intestine
14	/ptú/	H	flail
15	/tɕè ˈpjá/	LH	corn cake
16	/tʂhə̀ (ʂ)tə́ ˈquá/	LHH	stomach
17	/biá ȵú ʐdó/	HHH	male panda
18	/pˁɹ/	L	snow
19	/lì ˈχà sò gé/	LLLH	tenderness

Forms are given in International Phonetic Alphabet. No local writing system is available for Luobuzhai Qiang.

Linguistic Effects

Factors	Expected effects
Tone	F0 of tone category.
Lexical stress	Broader F0 range in a tone language.
Vowel height	Higher vowels have higher F0.
Consonant voicing	Voiceless consonants raise F0 on neighbouring vowel.
Previous/following tone	Assimilatory effects.
Location of syllable within word	Downtrend is expected.
Sex of speaker	Men typically have lower F0.
Age of speaker	Decrease in F0 with age.
Sentence condition	F0 may be shifted.
Word	Allows lexical F0 to vary.
Speaker idiosyncrasies	F0 affected by speaker characteristics

Median Model

Initial Model:

$F0_{\text{med}} \sim$ previous$*$tone$*$following + gender$*$tone$*$condition + previous$*$voice + tone$*$intensity$*$condition + tone$*$syllable + vowel + age$*$gender + *word* + *speaker$*$gender*

Final Model:

$F0_{\text{med}} \sim$ previous$*$tone$*$following + gender$*$condition + tone$*$condition + tone$*$syllable + vowel + *word* + *speaker*

Median Model Results

	Est (Hz)	Std (Hz)	t value	HPD l(Hz)	HPD u(Hz)
(Intercept)	377.02	9.31	40.5	355.16	398.45
previousH	6.64	4.55	1.5	-2.40	15.49
previousL	9.84	3.88	2.5	2.12	17.44
toneL	-69.60	4.07	-17.1	-77.73	-61.77
followingL	30.98	4.84	6.4	21.31	40.89
male	-78.34	11.65	-6.7	-106.71	-49.91
constrast	-25.11	2.06	-12.2	-29.27	-21.19
question	-45.00	2.05	-21.9	-48.96	-40.93
syllable	-25.52	1.58	-16.2	-28.67	-22.41
/a/	4.61	2.62	1.8	-0.59	9.80
/e/	11.93	3.01	4.0	5.93	17.83
/i/	23.26	2.99	7.8	17.41	29.13
/u/	20.28	3.05	6.7	14.14	26.42

Random Effects

Groups	Name	Std.Dev.	HPD l 95%	HPD u 95%
word	(Intercept)	9.3 Hz	6.9 Hz	14.9 Hz
speaker	(Intercept)	16.9 Hz	9.8 Hz	33.7 Hz
Residual		17.2	16.6 Hz	17.9 Hz

Linguistic Effects

Factors	Expected effects
Tone	F0 of tone category.
Lexical stress	Broader F0 range in a tone language.
Vowel height	Higher vowels have higher F0.
Consonant voicing	Voiceless consonants raise F0 on neighbouring vowel.
Previous/following tone	Assimilatory effects.
Location of syllable within word	Downtrend is expected.
Sex of speaker	Men typically have lower F0.
Age of speaker	Decrease in F0 with age.
Sentence condition	F0 may be shifted.
Word	Allows lexical F0 to vary.
Speaker idiosyncrasies	F0 affected by speaker characteristics

Outline

Phonetic Analysis of Pitch

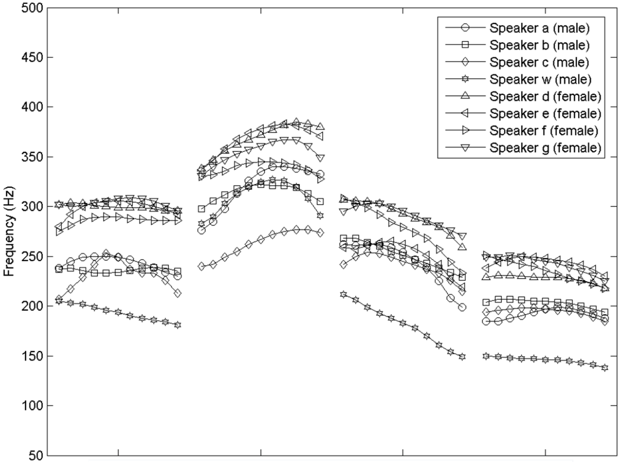
Mixed Models
Example: Median Analysis of Qiang

Functional Principal Component Analysis

Analysis of Qiang

Analysis of Mandarin

Curves as Data



Vowel F0 plot for /tshà tʂú qò qò/, 'storage room door'.

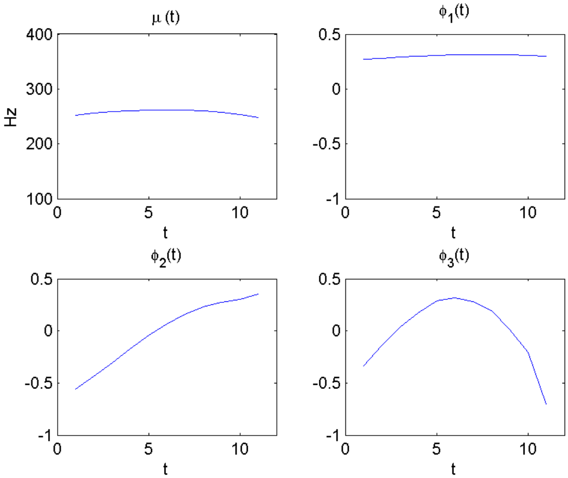
Functional Principal Component Analysis

Let $Y_i(t)$, $t \in T = [0, 1]$, $i = 1, \dots, n$ be data sampled from a Gaussian stochastic process on the domain T.

$$Y_i(t) = \mu(t) + \sum_{j=1}^{\infty} A_{i,j}\phi_j(t)$$

where $\phi_j(t)$ is the jth basis function and $A_{i,j}$ is the weight associated with the ith curve and the jth basis function. $\mu(t)$ is the overall mean of the sampled processes. Essentially the process is modelled as a mean function coupled with a stochastic basis expansion component.

Qiang Functional Principal Components



Estimated eigenfunctions for Qiang.

Mean and Covariance

An estimate of the mean function $\hat{\mu}(t_j)$ can be simply calculated from the mean of the data.
The eigenfunctions are then determined from a spectral analysis of the estimated covariance matrix

$$\hat{C}(t_k, t_l) = \frac{1}{n}\sum_{i=1}^{n}\left(Y_i(t_j) - \hat{\mu}(t_j)\right)\left(Y_i(t_l) - \hat{\mu}(t_l)\right), \quad k, l = 1, \ldots, m$$

Eigenvalues and Eigenfunctions

This yields the estimated eigenfunctions $\hat{\phi}_j(t)$ as

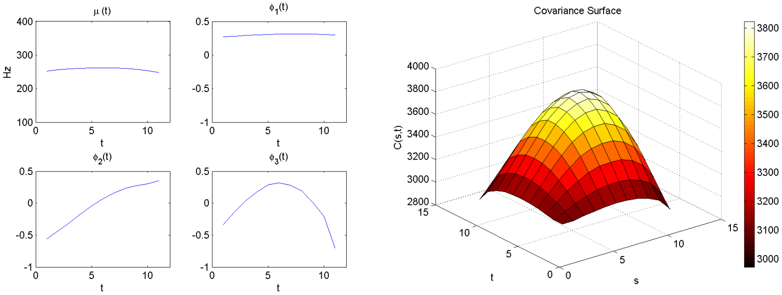
$$\hat{C}(t_k, t_l) = \sum_{j=1}^{m} \hat{\lambda}_j \hat{\phi}_j(t_k) \hat{\phi}_j(t_l)$$

with ordered eigenvalues $\hat{\lambda}_1, \hat{\lambda}_2, \ldots, \hat{\lambda}_m$. The FPC scores $A_{i,j}$ are then estimated by discrete approximation

$$\hat{A}_{i,j} = \sum_{k=1}^{m} \left(Y_i(t_k) - \hat{\mu}(t_k)\right) \hat{\phi}_j(t_k) \Delta_k$$

where $\Delta_k = t_k - t_{k-1}$.

Covariance and Components



Mean function and eigenfunctions for the first 3 components which account for 99.8% variance in the data, along with the estimated covariance function of the data.

FPCA and LME

For each sample process Y_i, two sets of scalar covariates X_i and Z_i are available. X_i are fixed effects, such as tone, while Z_i are zero-mean Gaussian random effects, such as speaker. The following model is proposed:

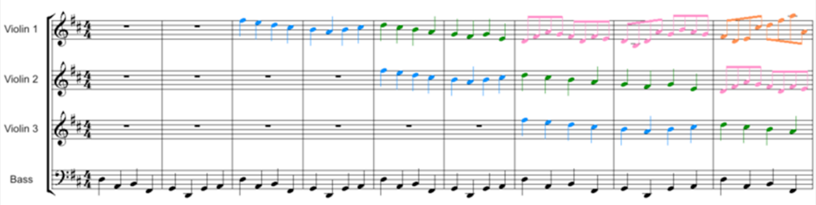
$$
\begin{aligned}
E(Y_i(t)|X_i, Z_i) &= \mu(t) + \sum_{j=1}^{\infty} E(A_{i,j}|X_i, Z_i)\phi_j(t) \\
E(A_{i,j}|X_i, Z_i) &= X_i\beta^{(j)} + Z_i\gamma^{(j)}, \quad \gamma^{(j)} \sim N(0, \Sigma_{\gamma^{(j)}})
\end{aligned}
$$

where $\phi_j(t)$ is the jth basis function and $A_{i,j}$ is the weight associated with the ith curve and the jth basis function. $\mu(t)$ is the overall mean of the sampled processes. In practice the summation is taken to K rather than ∞.

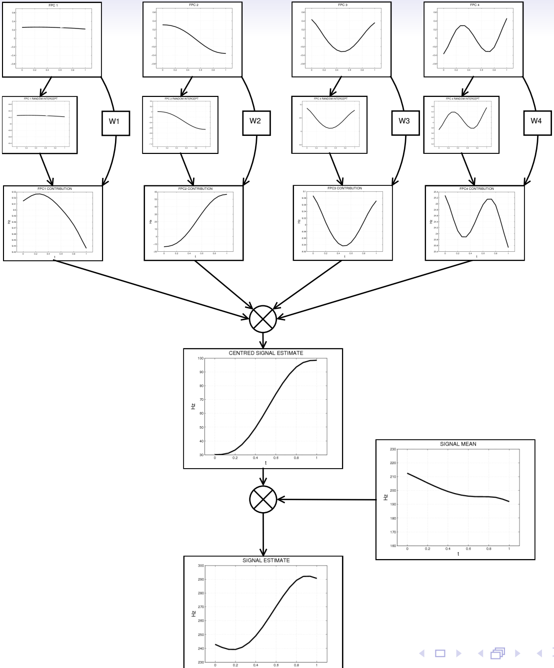
Intuitive idea behind their use

Pachelbel's Canon

From Wikipedia, the free encyclopedia



[Play]


FPC 1
FPC 2
FPC 3
FPC 4
W1
W2
W3
W4
FPC1 CONTRIBUTION
FPC2 CONTRIBUTION
FPC3 CONTRIBUTION
FPC4 CONTRIBUTION
CENTRED SIGNAL ESTIMATE
SIGNAL MEAN
SIGNAL ESTIMATE
Hz
t

Outline

Phonetic Analysis of Pitch

Mixed Models
Example: Median Analysis of Qiang

Functional Principal Component Analysis

Analysis of Qiang

Analysis of Mandarin

FPCA-LME Qiang Model

$$
\begin{aligned}
E(Y_i(t)|X_i, Z_i) &= \mu(t) + \sum_{j=1}^{K} E(A_{i,j}|X_i, Z_i)\phi_j(t) \\
E(A_{i,j}|X_i, Z_i) &= X_i\beta^{(j)} + Z_i\gamma^{(j)}, \quad \gamma^{(j)} \sim N(0, \Sigma_{\gamma^{(j)}})
\end{aligned}
$$

but how to choose K - the number of components.

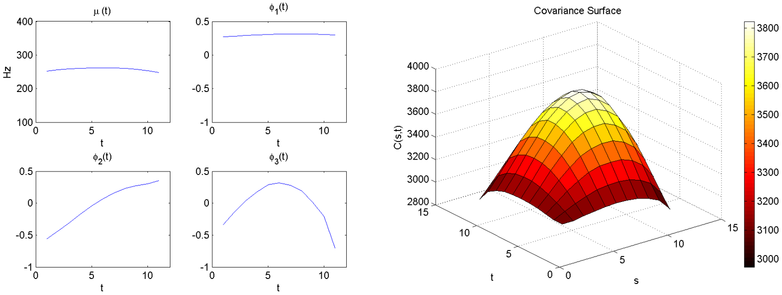
FPCA-LME Qiang Model

$$
\begin{aligned}
E(Y_i(t)|X_i, Z_i) &= \mu(t) + \sum_{j=1}^{K} E(A_{i,j}|X_i, Z_i)\phi_j(t) \\
E(A_{i,j}|X_i, Z_i) &= X_i\beta^{(j)} + Z_i\gamma^{(j)}, \quad \gamma^{(j)} \sim N(0, \Sigma_{\gamma^{(j)}})
\end{aligned}
$$

K chosen using these three criteria:

- Look at amount of variance explained by the component (given by eigenvalue)
- Look at the whether any regression components are related to the scores
- See whether component is below "just noticeable difference" threshold

Covariance and Components



Mean function and eigenfunctions for the first 3 components which account for 99.8% variance in the data, along with the estimated covariance function of the data.

FPC Models

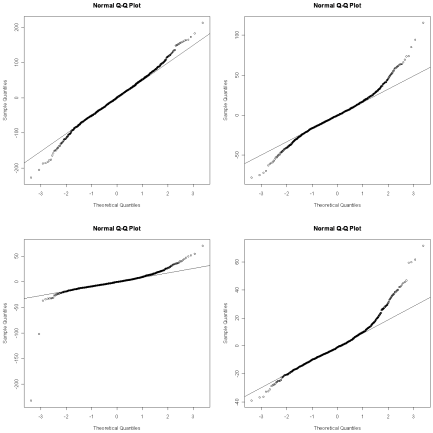
FPC1 $\sim$ previous$*$tone$*$following + gender$*$condition + tone$*$condition + tone$*$syllable + vowel + *word* + *speaker*

(Recall:
$F0_{\text{med}}$ $\sim$ previous$*$tone$*$following + gender$*$condition + tone$*$condition + tone$*$syllable + vowel + *word* + *speaker*)

FPC2 $\sim$ previous$*$tone$*$following + previous$*$voice + tone$*$intensity + vowel + *word* + *speaker*

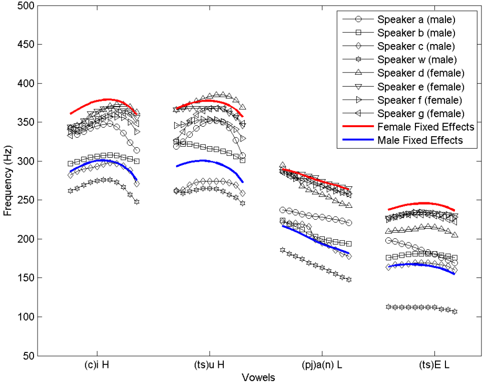
FPC3 $\sim$ previous$*$tone$*$following + gender$*$tone + previous$*$voice + vowel + tone$*$syllable + *word* + *speaker*

Diagnostics



Diagnostic QQ plots for the different functional principal component score linear mixed effect models.

Data and Fitted Values



F0 contour curves for the quadrisyllable of /ɕí tʂú 'piàn tsə̀/ ("riverbank"), with tonal pattern *HHLL* and as a declarative statement. The third syllable of the word is stressed. Also shown is the estimated functional response model curves for males and females.

Outline

Phonetic Analysis of Pitch

Mixed Models
  Example: Median Analysis of Qiang

Functional Principal Component Analysis

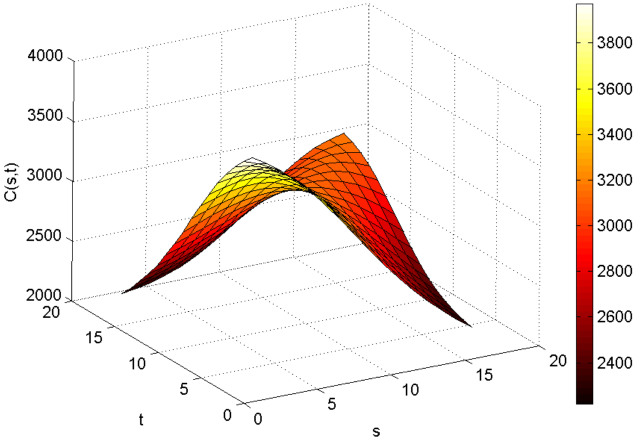
Analysis of Qiang

Analysis of Mandarin

Mandarin Data Characteristics

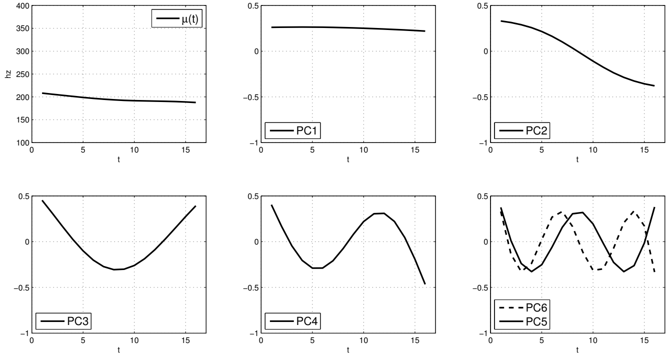
- COSPRO Data Set
- 5 Speakers
- 599 Paragraphs (mainly news items)
- Approx 70,000 individual syllables (curves)
- Designed to include all tonal combinations
- Complete recordings utilised not just frame words

Mandarin Covariance Function



Estimated covariance function for Mandarin.

Mandarin Functional Principal Components



Estimated eigenfunctions for Mandarin.

Percentage Variance Explained

FPC #	Individ. Variation	Cummul. Variation
FPC1	88.23	88.23
FPC2	9.78	98.01
FPC3	1.42	99.43
FPC4	0.32	99.75
FPC5	0.11	99.86
FPC6	0.05	99.91
FPC7	0.03	99.94
FPC8	0.02	99.96
FPC9	0.01	99.97
FPC10	0.01	99.98
FPC11	0.01	99.99
FPC12	0.01	99.99

FPC Hz range

FPC#	Hz (99%)	Hz (95%)	FPC#	Hz (99%)	Hz (95%)
FPC1	133.3	101.3	FPC7	3.6	1.7
FPC2	55.3	38.3	FPC8	2.9	1.2
FPC3	35.8	20.7	FPC9	2.4	1.1
FPC4	19.1	9.1	FPC10	1.8	.85
FPC5	8.9	4.2	FPC11	1.7	.68
FPC6	5.7	2.5	FPC12	1.3	.45

Effects	Values	Meaning
Fixed effects		
previous tone	0:5	previous tone, 0 - no previous tone
current tone	1:5	tone
following tone	0:5	following tone, 0 - no following tone
previous consonant	0:3	0-voiceless, 1-voiced
next consonant	0:3	2-no const, 3-sil/short pause
B2	linear	Position of the B2 break
B3	linear	Position of the B3 break
B4	linear	Position of the B4 break
B5	linear	Position of the B5 break
Sex	0:1	1 for male, 0 for female
rhyme type	1:37	Rhyme of syllable
Random Effects		
Speaker	$N(0, \sigma^2_{speaker})$	Speaker Effect
Sentence	$N(0,\sigma^2_{sentence})$	Sentence Effect

Break Type	Meaning
Break 1	Normal syllable boundary. In written Chinese, this corresponds to one character. (As this is our experimental data unit, B1 is equivalent to the mean value in the regressions and thus not included separately).
Break 2	Prosodic word boundary. Syllables group together into a word, which may or may not correspond to a lexical word.
Break 3	Prosodic phrase boundary. This break is marked by an audible pause.
Break 4	Breath group boundary. The speaker inhales.
Break 5	Prosodic group boundary. A complete speech paragraph.

Table: COSPRO Break Annotation

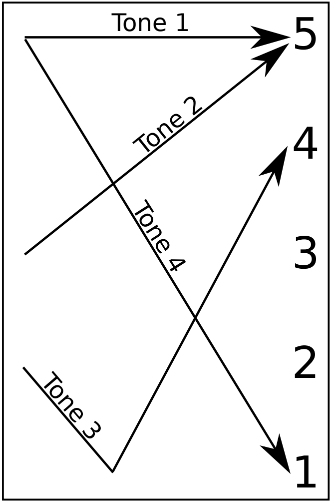
Initial and Final Models

$$
\begin{aligned}
FPC_X = \{[tn_{previous} * tn_{current} * tn_{next}] + [cn_{previous} * tn_{current} * cn_{next}] + \\
[(B2) + (B2)^2 + (B2)^3 + (B3) + (B3)^2 + (B3)^3 + \\
(B4) + (B4)^2 + (B4)^3 + \\
(B5) + (B5)^2 + (B5)^3] * Sex + [rhyme_t]\}\beta \\
+ \{[Sentence] + [SpkrID]\}\gamma + \epsilon
\end{aligned}
$$

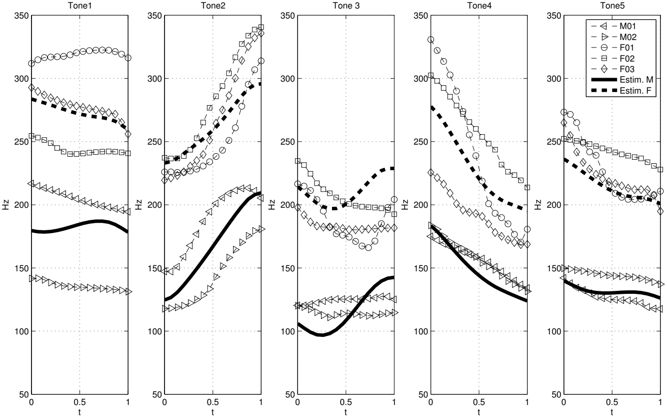
- FPC1 - 2nd order tone and consonant interactions
- FPC2 - 3rd order interactions
- FPC3 - 2nd order tone, 3rd order consonant, no longer phrasal dependence
- FPC4 - 2nd order tone and consonant interactions, no complete phrasal dependence

Tones in Mandarin

Tone1 媽 "mother"
Tone2 痲 "hemp"
Tone3 馬 "horse"
Tone4 罵 "scold"
Tone5 嗎 question particle

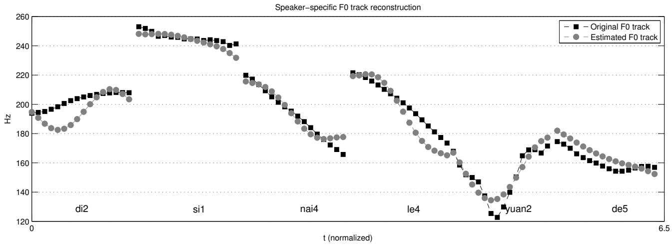


Mandarin Estimates



Estimates of specific utterances in Mandarin.

Mandarin Estimates as Time Series



Time Tracked Estimates for Mandarin.

Summary

- Presented a functional response model with can account for mixed effects in Phonetics
- Incorporates LME through the FPC scores
- Can be used to analyse pitch curve data
- Showed its application to Qiang and Mandarin