Vowel Acoustics in Dysarthria: Speech Disorder Diagnosis and Classification

Kaitlin L Lansford & Julie M Liss

Journal of Speech, Language and Hearing Research (2014) 57: 57-67

PRESENTED BY GILLIAN DE BOER MA MSC
SPEECH-LANGUAGE PATHOLOGY

Study Questions

- Can vowel acoustics metrics distinguish dysarthric speech from healthy speech?

- Can they distinguish between the types of dysarthric speech?
  - Ataxic (Cerebellar ataxia)
  - Hypokinetic (Parkinson's disease)
  - Hyperkinetic (Huntington's disease)
  - Mixed flaccid-spastic (amyotrophic lateral sclerosis)

Dysarthric speech

- Distorted vowels
  - Centralised vowel frequencies
  - Reduced vowel space
  - Abnormal formant frequencies (Kent et al. 1999)
  - Reduced F2 slope (Kent et al., 1989)

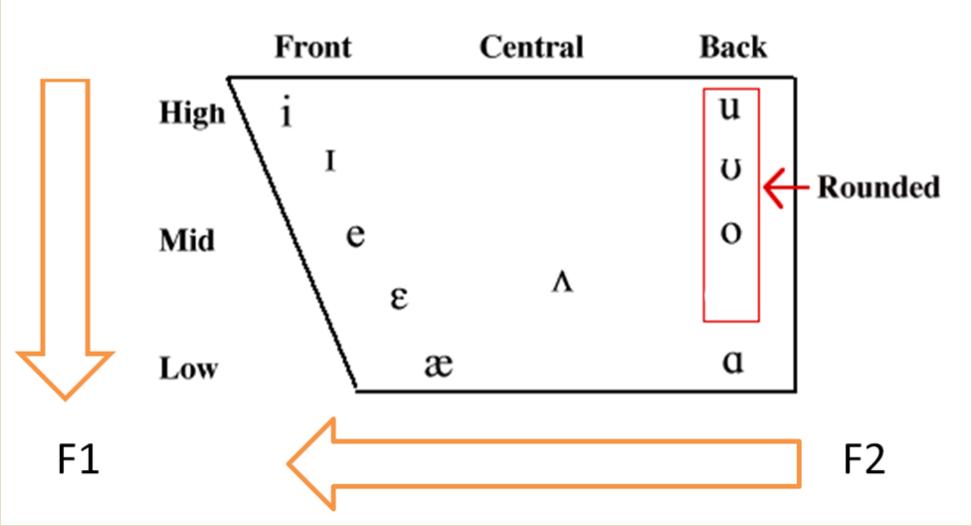
- Mechanism
  - Limited and slower lip, tongue & jaw movements
  - Aberrant timing (Yunusova et al. 2008)

Previous studies (1)

- Reduced F2 transitions for males with ALS (Weismer et al. 2001, Weismer & Martin, 1992)
- Reduced F2 transitions in Parkinson's disease, stroke (Kim et al, 2009) and multiple sclerosis (Rosen et al, 2008)

Previous studies (2)

- Quadrilateral vowel space area (VSA)
  - Mixed results (Weismer et al., 2001; Tjaden & Wilding, 2004)
- Triangular VSA
  - Sapir et al. (2007)
- Lax VSA
  - Tjaden et al. (2005)



Previous studies (3)

- Dispersion
  - Intelligibility in cerebral palsy (Kim et al. 2011)
- Formant centralisation ratio (FCR)
  - Vowel centralisation : FCR > 1
  - Hypokinetic (parkinson's) > healthy controls (Sapir et al., 2010)

Study aim

- Evaluate multiple vowel space metrics to distinguish
  - Normal vs dysarthric speech
  - Dysarthria subtypes

T-tests, Anovas and discriminant function analysis

Participants

- 57 speakers (29 M, 28 F)
  - 12 healthy controls
  - 12 ataxic dysarthria
  - 12 hypokinetic - Parkinson's disease (PD)
  - 10 hyperkinetic - Huntington's disease (HD)
  - 11 mixed flaccid-spastic (ALS)
- American English
- 2 SLPs rated speech mild, moderate, severe

Data collection

- One session
- Head mounted microphone, sound booth, stimuli read from computer screen in “normal conversational voice”
- Recording: TF32 script (Milenkovic 2004, 16-bit, 44 kHz) and saved to disk
- Editing: SoundForge (Sony Corp, Palo Alto CA)

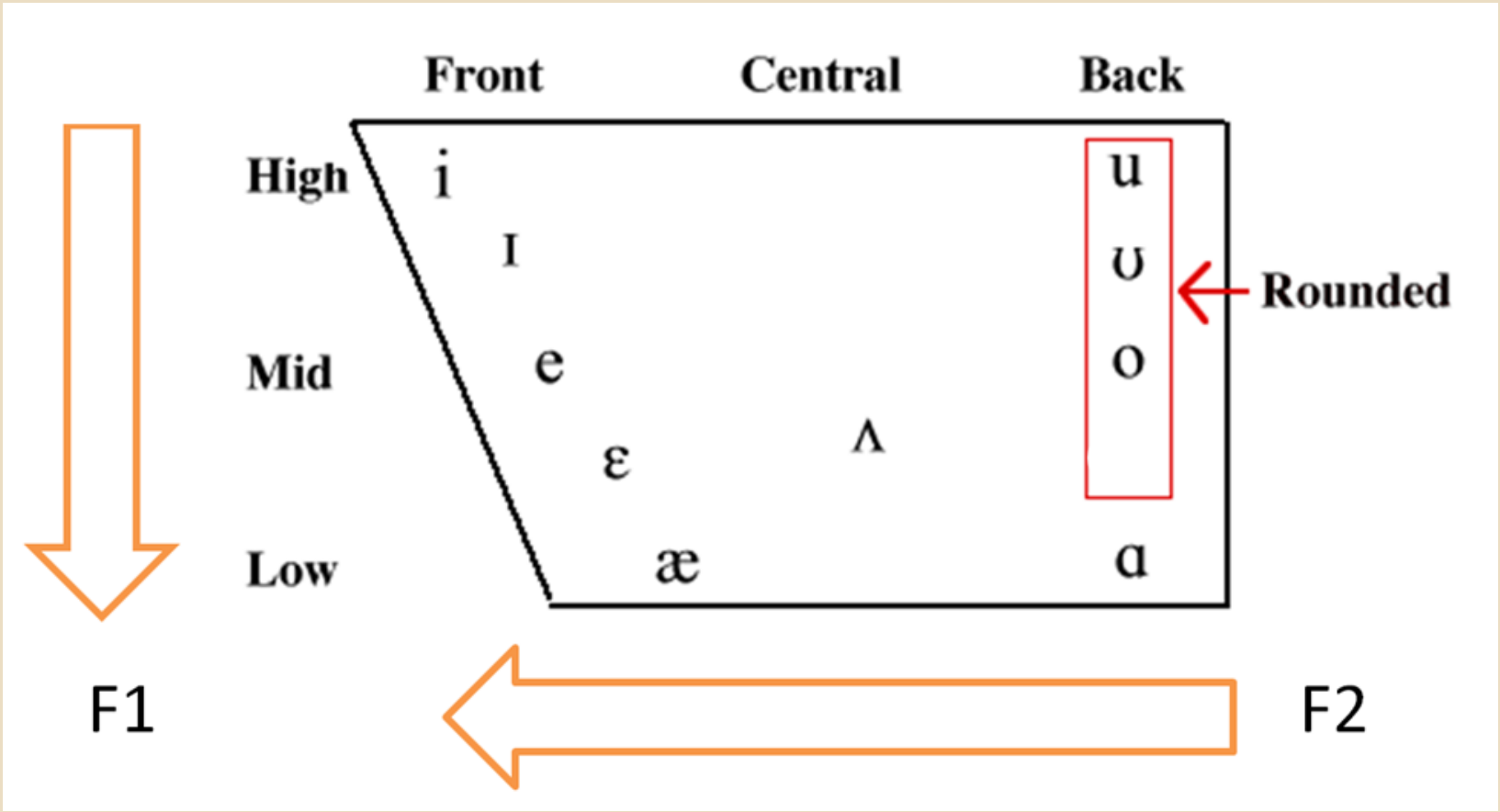
Stimuli

- 80 short phrases (6 syllables each)
  - “push her equal culture”
- Mix of strong and weak syllables
- Analysis : 36 phrases & strong syllables
- Four instances of 9 vowels (/i/, /I/, /e/, /ɛ/, /æ/, /u/, /o/, [/ɑ/], /^/)
- 3 instances of /ʊ/

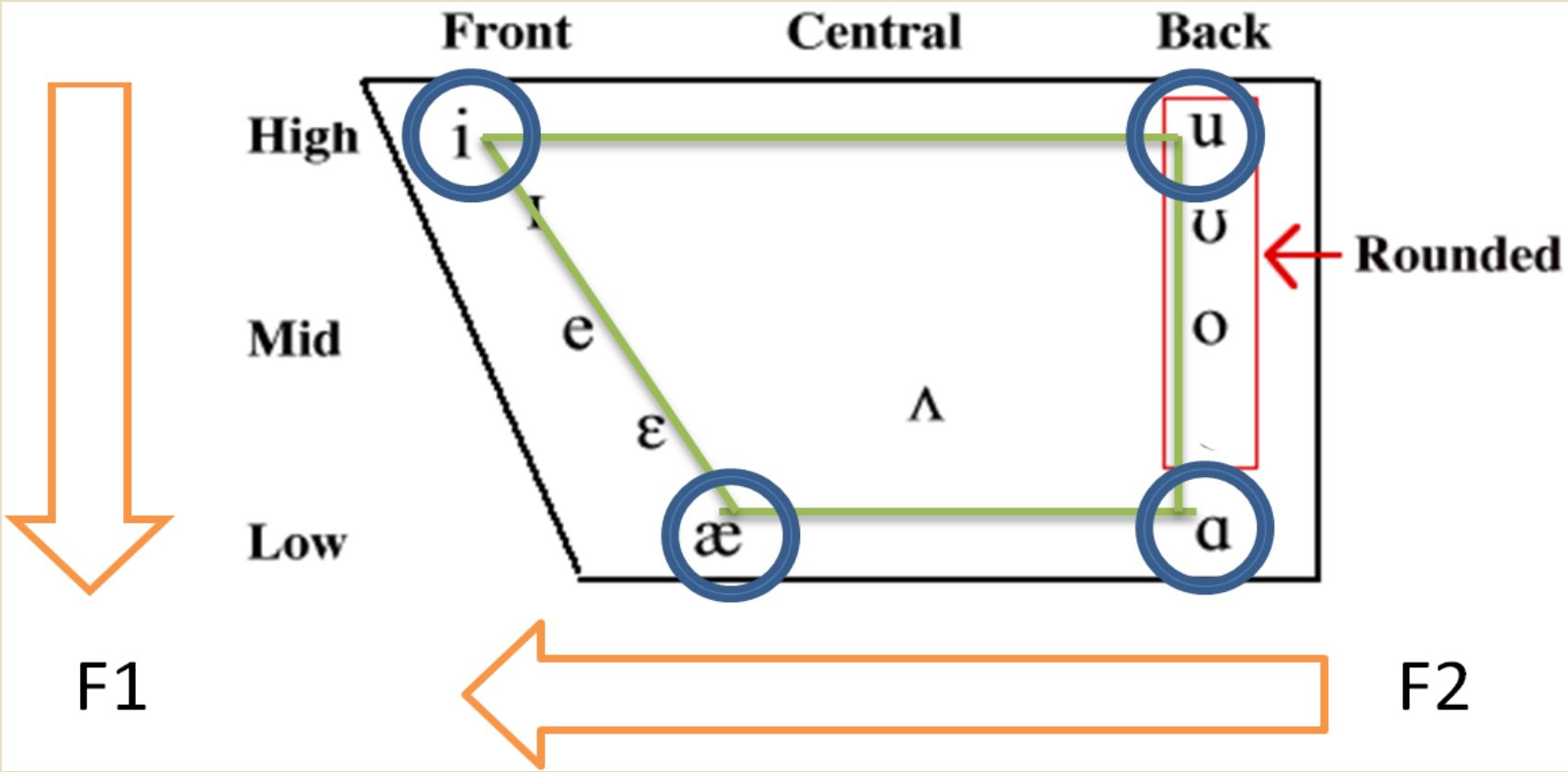
Spectral and Temporal Measurements

- Praat (Boersma & Weenik, 2006)
- F1 & F2 (Hz) : onset (20% duration), midpoint (50% duration), offset (80% duration)
  - Midpoint = steady state
- Two individuals made the measurements
- Intra-rater reliability = .886 (F1), .819 (F2)
- Inter-rater reliability = .889 (F2), .884 (F2)

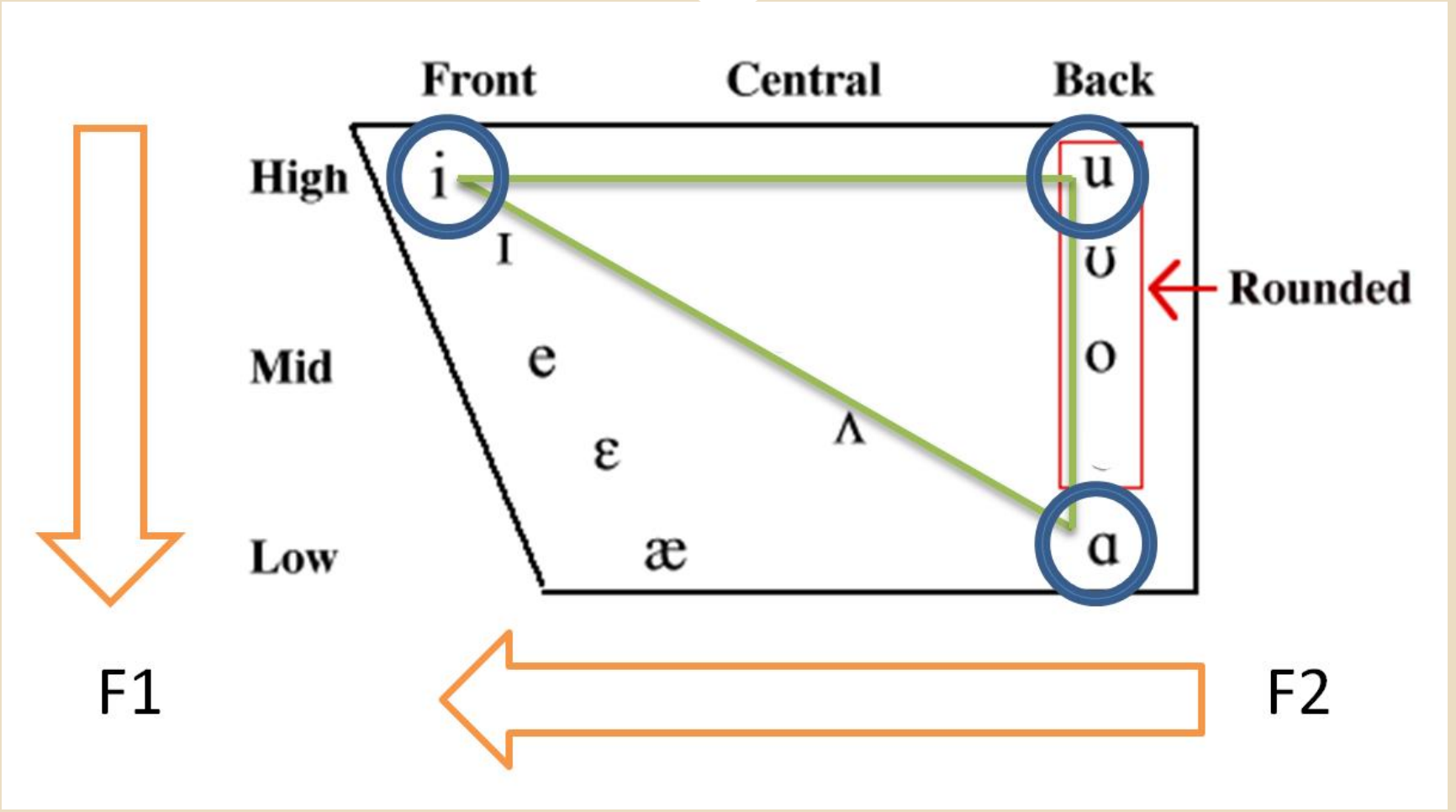
Derived vowel metrics



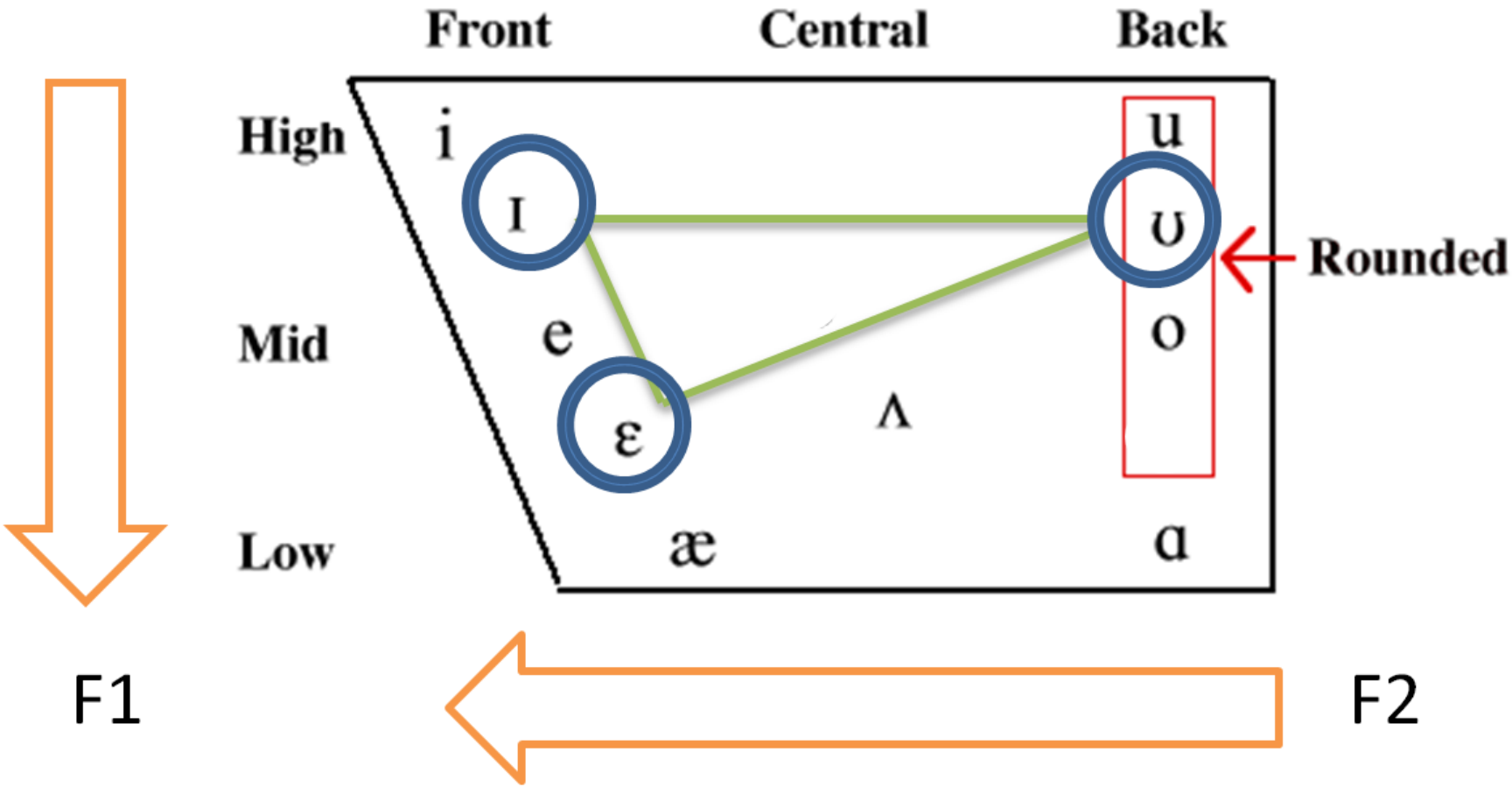
Quadrilateral Vowel Space Area



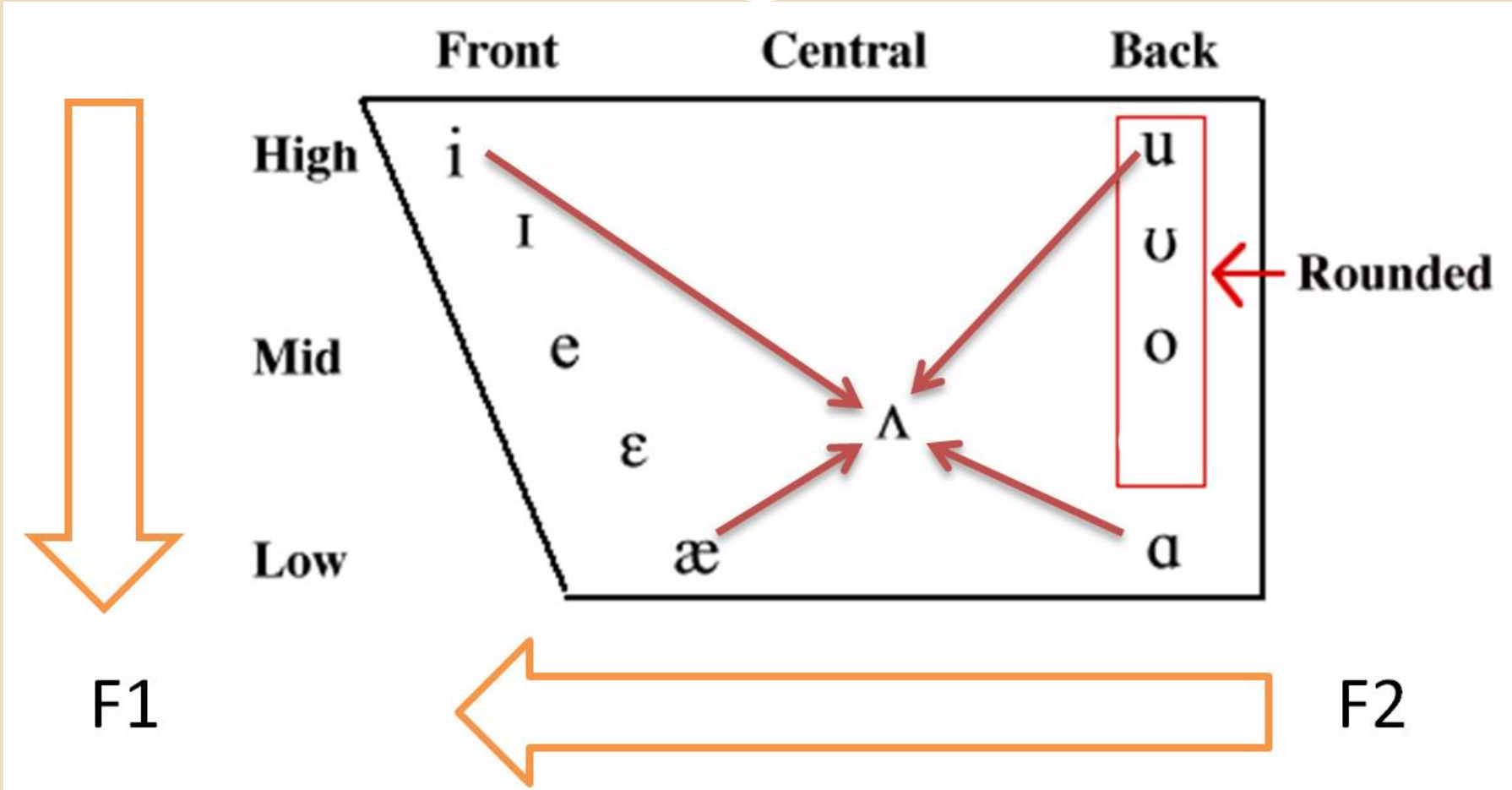
Triangular VSA



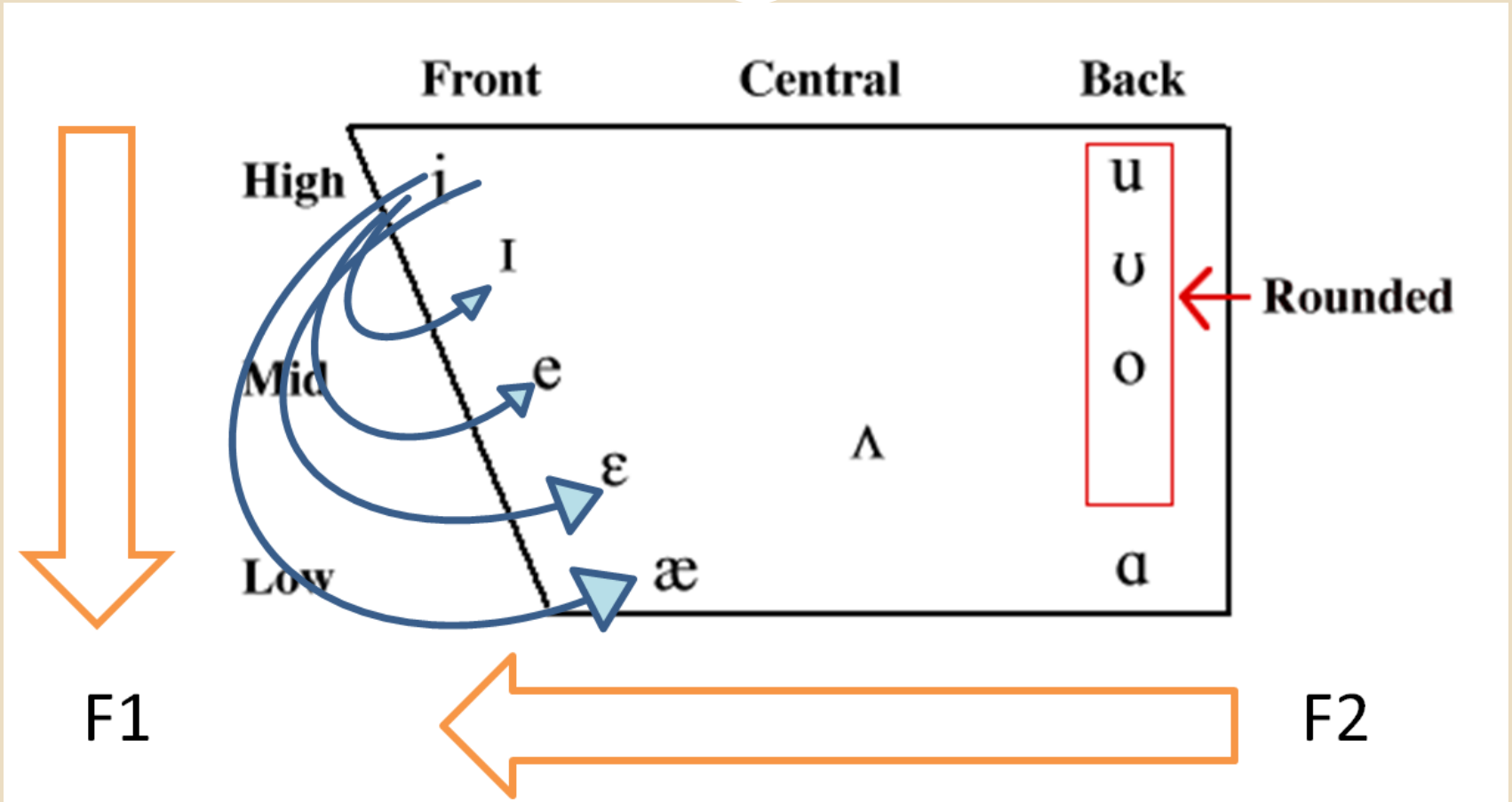
Lax VSA



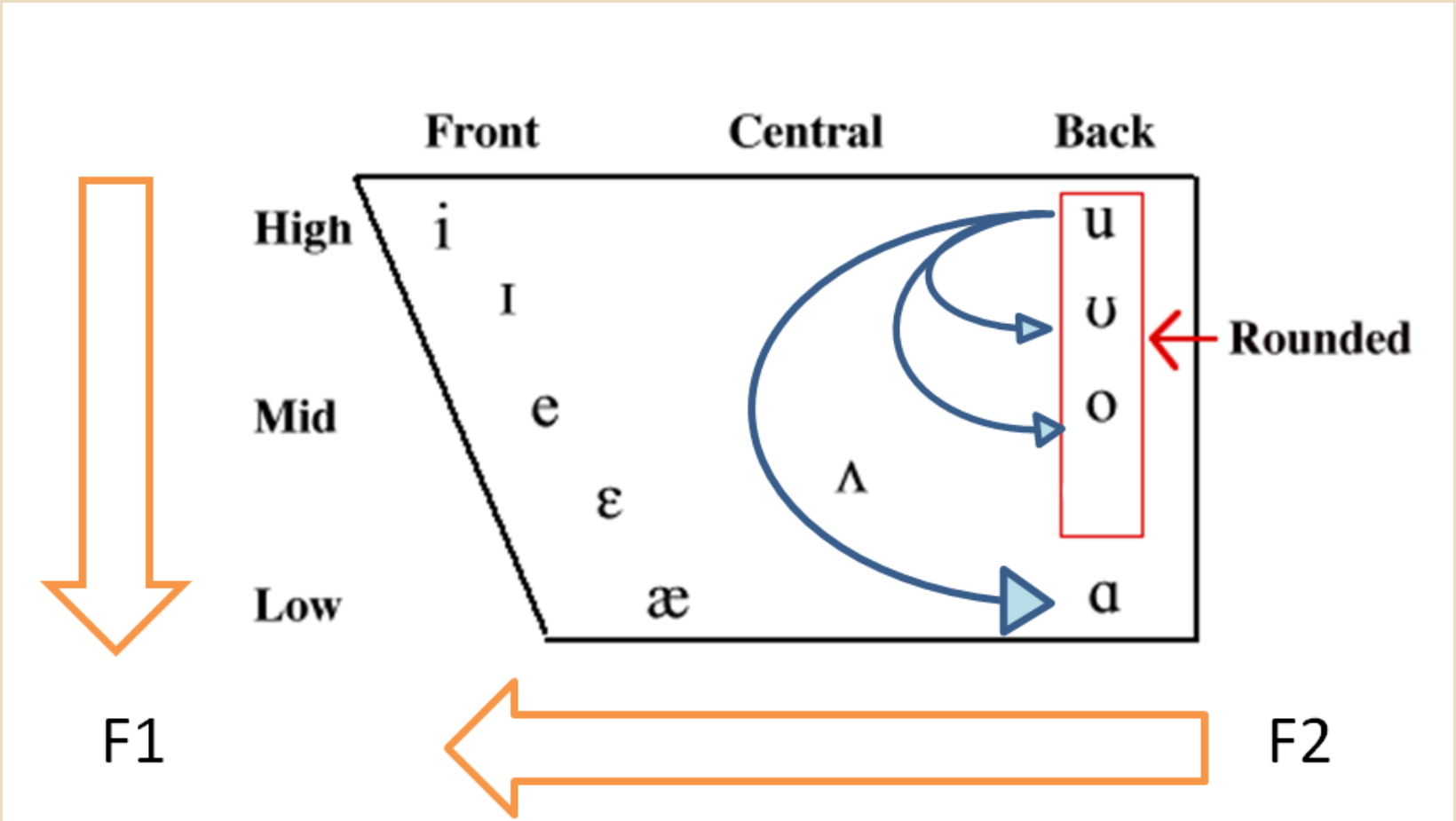
Corner dispersion



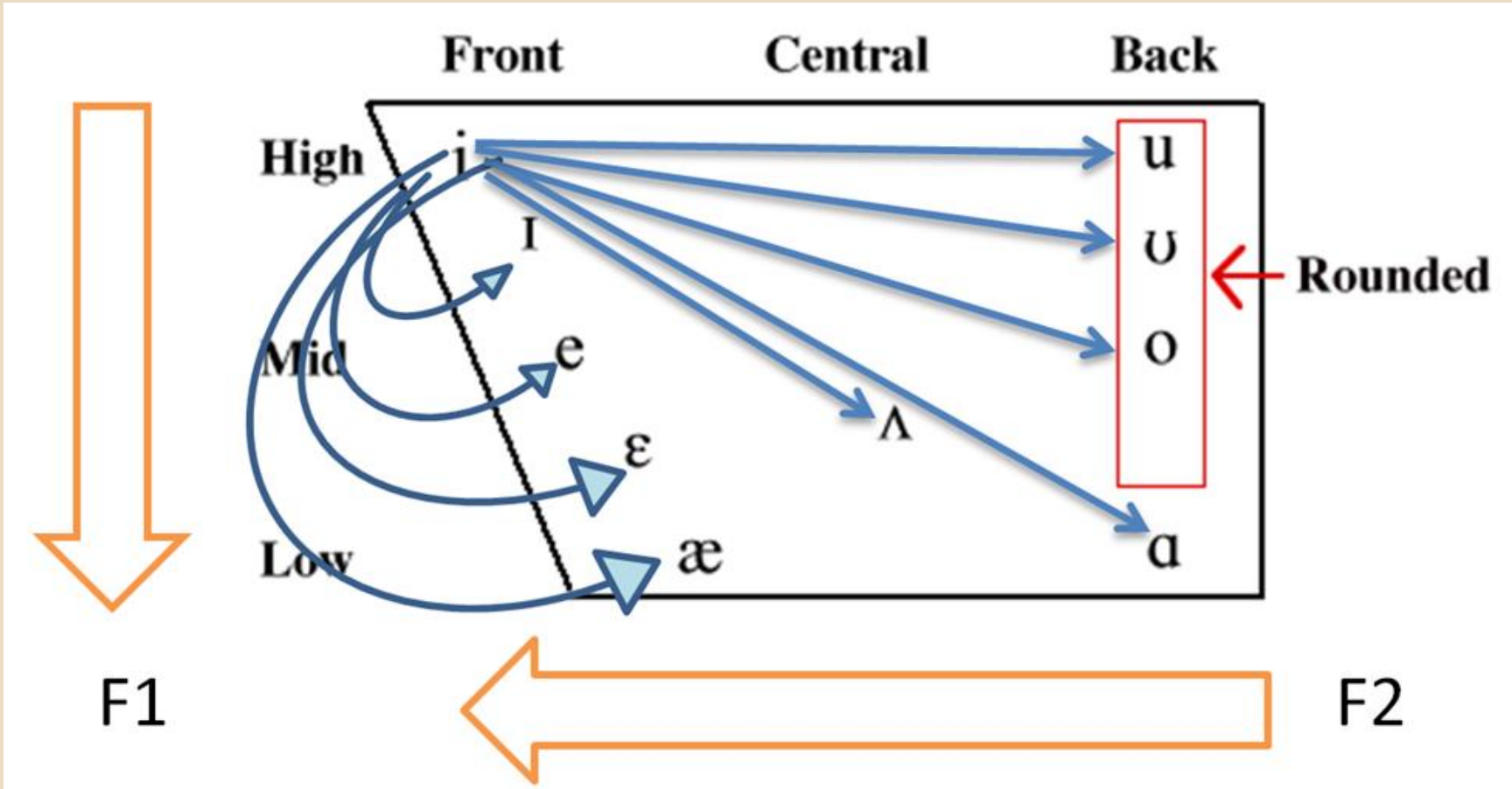
Front dispersion



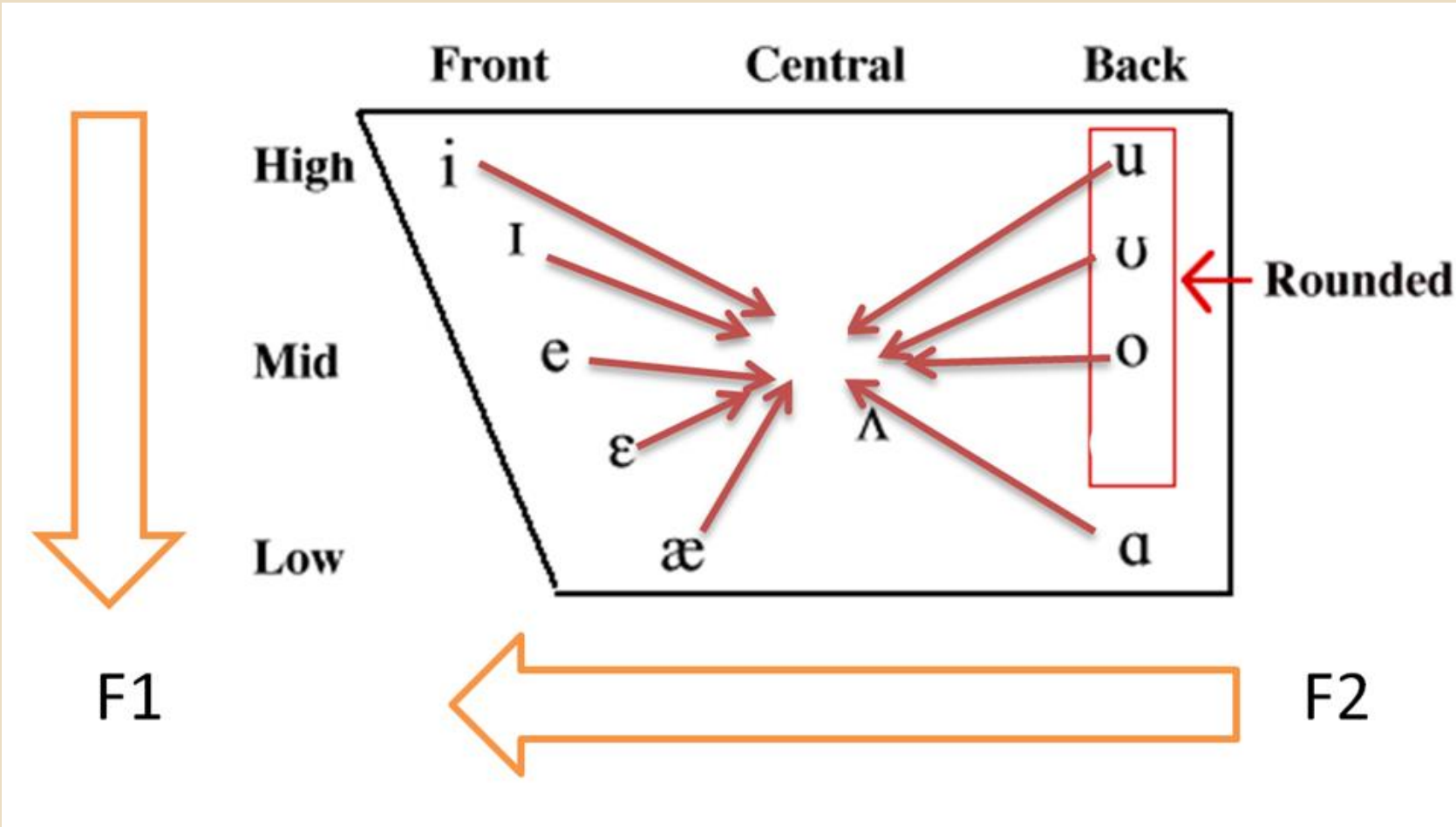
Back dispersion



Mean dispersion



Global dispersion



Formant centralisation ratio (FCR)

$$FCR = \frac{F1i + F1a + F2u + F2a}{F1a + F2i}$$

Vowel centralisation = FCR > 1

F2 Slope

$$\frac{|\text{F2onset} - \text{F2offset}|\ (\text{Hz})}{\text{duration (ms)}}$$

- Average F2 slope (Hz/ms) of all vowels
- Dynamic F2 slope (Hz/ms) of /æ/, /^/ & /ʊ/

Data Analysis

- T-tests
  - p value = .0045 (Bonferroni .05/11)
- Discriminant Function Analysis (DFA)
  - Predicts a categorical dependent variable by one or more continuous or binary independent variables
- 11 one-way Anovas (subtype)
- Subtype DFA

Results(1a)

Table 3. Healthy control and dysarthric group means and results of independent samples *t* tests.

Vowel metric and group	*n*	*M*	*SD*	*t*(55)	*p*	Cohen's *d*
Quadrilateral VSA				5.056	.000*	1.62
HC	12	286,213.07	71,217.41			
D	45	174,822.17	66,928.04			
Triangular VSA				2.745	.008	0.96
HC	12	175,285.55	49,012.16			
D	45	120,378.89	64,311.64			
Lax VSA				2.202	.032	0.33
HC	12	312,88.86	19,208.13			
D	45	18,659.61	17,240.48			
FCR				–5.098	.000*	1.31
HC	12	1.07	0.05			
D	45	1.19	0.12			
Mean dispersion				3.283	.002*	1.04
HC	12	400.54	69.31			
D	45	330.46	64.76			
Front dispersion				5.503	.000*	1.82
HC	12	503.32	83.38			
D	45	345.65	89.34			
Back dispersion				3.916	.000*	1.25
HC	12	368.45	75.32			
D	45	276.13	71.86			
Corner dispersion				4.051	.000*	1.22
HC	12	563.45	120.48			
D	45	432.14	93.89			
Global dispersion				3.756	.000*	1.18
HC	12	597.56	101.37			
D	45	484.11	90.76			
Average F2 slope (Hz/ms)				4.271	.000*	1.11
HC	12	2.08	0.29			
D	45	1.55	0.61			
Dynamic F2 slope (Hz/ms)				2.927	.005	1.04
HC	12	3.21	0.70			
D	45	2.32	0.99			

Note. HC = healthy control; D = dysarthric.

*$p < .0045$.

Results (1b)

Vowel metric and group	*n*	*M*	*SD*	*t*(55)	*p*
Quadrilateral VSA				5.056	.000*
HC	12	286,213.07	71,217.41		
D	45	174,822.17	66,928.04		
Triangular VSA				2.745	.008
HC	12	175,285.55	49,012.16		
D	45	120,378.89	64,311.64		
Lax VSA				2.202	.032
HC	12	312,88.86	19,208.13		
D	45	18,659.61	17,240.48		
FCR				–5.098	.000*
HC	12	1.07	0.05		
D	45	1.19	0.12		
Average F2 slope (Hz/ms)				4.271	.000*
HC	12	2.08	0.29		
D	45	1.55	0.61		
Dynamic F2 slope (Hz/ms)				2.927	.005
HC	12	3.21	0.70		
D	45	2.32	0.99		

Note. HC = healthy control; D = dysarthric.

*$p < .0045$.

Results (1c)

Vowel metric and group	*n*	*M*	*SD*	*t* (55)	*p*
Mean dispersion				3.283	.002*
HC	12	400.54	69.31		
D	45	330.46	64.76		
Front dispersion				5.503	.000*
HC	12	503.32	83.38		
D	45	345.65	89.34		
Back dispersion				3.916	.000*
HC	12	368.45	75.32		
D	45	276.13	71.86		
Corner dispersion				4.051	.000*
HC	12	563.45	120.48		
D	45	432.14	93.89		
Global dispersion				3.756	.000*
HC	12	597.56	101.37		
D	45	484.11	90.76		

Results (2a)

Table 4. Healthy control and dysarthric speaker classification accuracy by vowel metric.

Vowel metric and group	Predicted group HC	Predicted group D	Overall accuracy (%)	Mild accuracy (%)	Moderate accuracy (%)	Severe accuracy (%)
Front dispersion			84.2	100.0	73.1	92.3
HC	11	1				
D	8	37				
Quadrilateral VSA			80.7	83.3	80.8	92.3
HC	8	4				
D	7	38				
Back dispersion			73.7	66.7	73.1	92.3
HC	7	5				
D	10	35				
Corner dispersion			73.7	33.3	73.1	92.3
HC	9	3				
D	12	33				
Global dispersion			71.9	66.7	69.2	84.6
HC	8	4				
D	12	33				
Average F2 slope			71.9	33.3	73.1	100.0
HC	9	3				
D	13	32				
FCR			70.2	50.0	65.4	76.9
HC	11	1				
D	15	30				
Mean dispersion			70.2	50.0	69.2	84.6
HC	8	4				
D	13	32				
Triangular VSA			66.7	50.0	61.5	84.6
HC	8	4				
D	15	30				

Results (2b)

Vowel metric and group	Predicted group HC	Predicted group D	Overall accuracy (%)
Front dispersion			84.2
HC	11	1	
D	8	37	
Quadrilateral VSA			80.7
HC	8	4	
D	7	38	
Back dispersion			73.7
HC	7	5	
D	10	35	
Corner dispersion			73.7
HC	9	3	
D	12	33	
Global dispersion			71.9
HC	8	4	
D	12	33	
Average F2 slope			71.9
HC	9	3	
D	13	32	

Results (2c)

Vowel metric and group	Overall accuracy (%)	Mild accuracy (%)	Moderate accuracy (%)	Severe accuracy (%)
Front dispersion HC D	84.2	100.0	73.1	92.3
Quadrilateral VSA HC D	80.7	83.3	80.8	92.3
Back dispersion HC D	73.7	66.7	73.1	92.3
Corner dispersion HC D	73.7	33.3	73.1	92.3
Global dispersion HC D	71.9	66.7	69.2	84.6
Average F2 slope HC D	71.9	33.3	73.1	100.0

Results (3)

Table 5. Analysis of variance testing equality of means of vowel metrics for the dysarthria subtypes.

Vowel metric	$F(3, 41)$	p	η^2
Quadrilateral VSA	0.358	.783	
Triangular VSA	1.403	.256	
Lax VSA	0.208	.890	
FCR	0.672	.574	
Mean dispersion	0.436	.728	
Front dispersion	1.634	.196	
Back dispersion	0.614	.610	
Corner dispersion	0.974	.414	
Global dispersion	0.669	.576	
Average F2 slope	14.327	.000	.512
Dynamic F2 slope	12.270	.000	.473

Results (4)

Table 6. Group means of significant variables revealed by analysis of variance.

Vowel metric	*n*	*M*	*SD*	95% CI LL	95% CI UL
		Average F2 slope (Hz/ms)			
Ataxic	12	1.32	0.34	1.10	1.54
ALS	11	1.01	0.45	0.71	1.31
HD	10	1.70	0.32	1.47	1.93
PD	12	2.16	0.59	1.78	2.54
		Dynamic F2 slope (Hz/ms)			
Ataxic	12	1.90	0.80	1.40	2.41
ALS	11	1.51	0.81	0.97	2.05
HD	10	2.59	0.32	2.37	2.82
PD	12	3.25	0.87	2.69	3.81

Note. CI = confidence interval; LL = lower limit; UL = upper limit; HD = Huntington's disease; PD = Parkinson's disease.

Results (5)

Table 7. Classification summary by dysarthria subtype.

Group	Predicted group: Ataxic	ALS	HD	PD
		Average F2 slope[a]		
Ataxic	4	5	3	
ALS	3	6	2	
HD	3		3	4
PD	2		3	7
		Dynamic F2 slope[b]		
Ataxic	4	5	2	1
ALS	1	7	3	
HD	1		7	2
PD			6	6

[a]Overall accuracy = 44.4%. [b]Overall accuracy = 53.3%.

Discussion (1)

- 8 of 11 vowel metrics differed significantly between HC and dysarthria group.
- Front dispersion and quadrilateral VSA best distinguish dysarthria vs control (84%, 80%)

Discussion (2)

- ANOVA within dysarthric groups
  - F2 slope (avg & dynamic)
- DFA of F2 slope
  - F2 slope avg 44.4%, F2 slope dynamic (53.3%)
- F2 slope unit = Hz/ms
- F2 monophthongs vs dipthongs

Discussion (3)

- Vowel space compression
- Perceptual similarity across subtypes

- Future line of investigation

Conclusion

- Acoustic metrics can aid in the diagnosis of dysarthria
- How the acoustics map to perception is the topic of a companion piece (Landsford & Liss, 2014)

Quadrilateral VSA Formula

"Heron's formula was used to calculate the area [...] formed by the corner vowels (i, æ, a, u) in F1 × F2 space. Toward this end, the area (as calculated by Heron's formula) of the two triangles formed by the sets of vowels /i/, /æ/, /u/ and /u/, /æ/, /a/ are summed. Heron's formula is as follows:

$$\sqrt{s(s-a)(s-b)(s-c)}$$

where s is the semiperimeter of each triangle, expressed as s=½(a+b+c), and a, b, and c each represent the Euclidean distance in F1 × F2 space between each vowel pair (e.g., /i/ to /æ/)" (Lansford & Liss, 2014)

Triangular VSA formula

"Triangular vowel space area was constructed with the corner vowels (i, a, u). It was derived using the equation outlined by Sapir and colleagues (2010) and is expressed as

ABS{[F1i× (F2a–F2u)+F1a×(F2u–F2i)+F1u×(F2i–F2a)]/2}.

ABS in this equation refers to absolute value."