Predicting Speech Intelligibility With a Multiple Speech Subsystems Approach in Children With Cerebral Palsy

Jimin Lee, Katherine C. Hustad, & Gary Weismer

Journal of Speech, Language and Hearing Research (2014) 57:1666-78

PRESENTED BY GILLIAN DE BOER MA MSC SPEECH-LANGUAGE PATHOLOGY

Cerebral palsy

- Dysarthria in CP: 50%, 58%, 31%-88%
- Dysarthria: very mild → very severe
- Non-progressive, persistent
- Childhood dysarthria = adult dysarthria ?
- Language acquisition

Previous studies (1)

- Scarce and outdated (1950s-1980s)
- Fundamental frequency (F0) and “loudness” sound pressure level (SPL) → no relation to “speech proficiency” (Clarke & Hoops, 1980)
- Vowel space area, F2 slope contribute to adult speech intelligibility (review Weismer, 2008)
- VSA correlated with speech intelligibility in children (4 - 5.5 yrs) (Lee & Hustad, 2013)

Previous studies (2)

- Voice, nasality and formant movement potentially contribute to speech intelligibility (de Bodt, Huici & Van De Heyning, 2002)

- CP may affect multiple speech subsystems

Study question (1)

- “What are the segmental, voice, resonance and intelligibility characteristics of speech in children with speech motor impairment (SMI) secondary to CP and in children with CP and no diagnosed speech motor impairment (NSMI), and how do they compare to the same characteristics in [typically developing] children?”
  - Lee, Hustad & Weismer (2014)

Study question (2)

- “When including multiple acoustic variables reflecting different speech subsystems in a prediction model, which acoustic variables are the best predictors of intelligibility in children with CP? Also, what is the independent contribution of each acoustic variable to speech intelligibility among multiple acoustic variables reflecting different speech subsystem?” (Lee et al., 2014)

Participants (1)

- 22 children w/CP (11M, 11F), 67 months (SD 9.9)
- Clinical diagnosis of CP
  - 13 Speech Motor Impairment (SMI)
  - 9 No Speech Motor Impairment (NSMI)
- American English
- Normal hearing

Participants (2)

- 19 Typically developing (TD) children, 64 months (SD 10.4)
- No known disabilities
- American English
- Normal hearing
- Preschool Language Scale 4th ed. (Zimmerman et al., 2005)
- Arizona Articulation Proficiency Scale 3rd ed. (Fudala, 2000)
- Age & gender matched to CP group

Participants (3)

- 82 listeners (two per child)
- American English
- Hearing test (pure tone 25 dB HL)
- Age 18-40 years
- No language/learning/cognitive disabilities
- No experience w/ communication disorders

Stimuli

- 38 words from Test of Children's Speech (Hodge & Daniels, 2007)
  - 25 words one repetition
  - 13 words five repetitions
- 13 chosen for acoustic analysis
  - *sheet*, *seat*, *hoot*, *boot*, *top*, *hot*, *bad*, *hat*, *pipe*, *whip*, *toys* & *big*
- One nonsense word: /mIm/
  - Five repetitions

Data collection

- Delayed imitation
- Picture + recording
- Sound attenuating booth
- Professional quality recording equipment
- Sampling rate 44.1 Hz (16 bit)
- Condenser microphone (45 cm approx)

Intelligibility scores

- Listeners heard recordings at 75 dB SPL
- Played once (no playback)
- Transcribed words
- Presentation order randomized
- Each listener → one child
- Intelligibility = (# correct words/38) x 100%
- Score = (listener 1 + listener 2)/2
- Δ listeners > 10% → listener 3

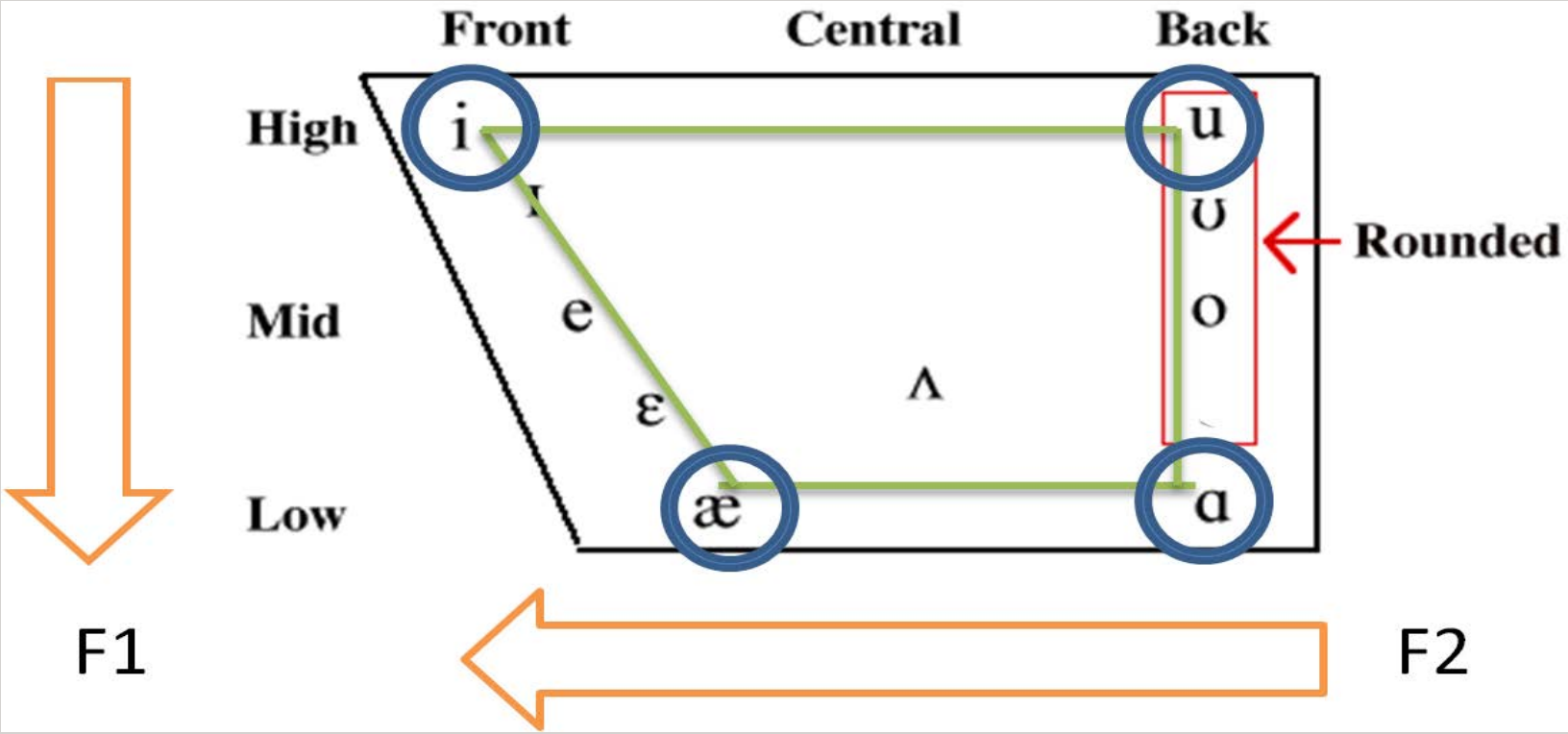
Acoustic Variables

- Articulatory, velopharyngeal and laryngeal subsystems
- No measure for respiratory subsystem

- Acoustic analysis in TF32 (Milenkovic, 2002)
  - Wideband spectograms
  - Fast Fourier transform
  - Linear predicative coding

Articulatory: vowel space

- 1422 tokens (8 words x 41 children x 4-5 reps)



Articulatory: vowel duration

- Duration of /i/, /u/, /æ/, & /ɑ/
- 1424 tokens

- Vowel duration increase → inteligibility decrease

Articulatory : F2 slope

- 515 tokens
- *Pipe, toys, whip*
- F2 slope = $\frac{|\text{F2onset} - \text{F2offset}|\ (\text{Hz})}{\text{duration (ms)}}$
- F2 transitional
  - duration → (ms)
  - extent = $|\text{F2onset} - \text{F2offset}|\ (\text{Hz})$

Velopharyngeal

- A1-P1 = measure of nasality, Chen (1995)
- /mIm/ → P1
- *big* (A1-P1) dB
- 315 tokens

- From Chen (1995)

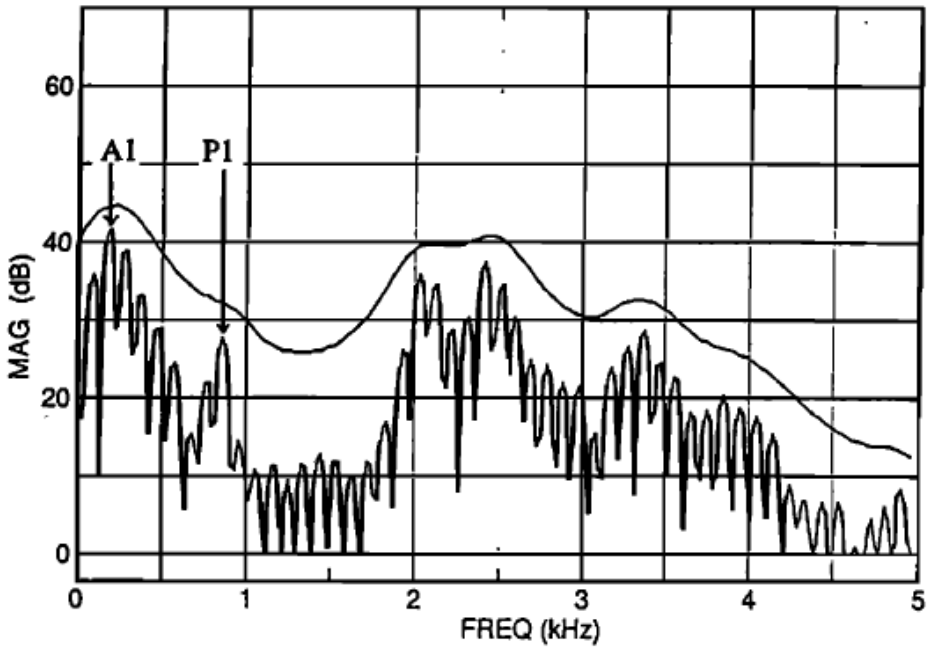


FIG. 2. The bandwidth of the first formant can be quantified by the amplitude of that formant, *A*1, and the prominence of the extra peak can be quantified by its amplitude, *P*1.

Laryngeal

- 186 tokens of /ɑ/ from *top*
- F0
- Signal to noise ratio (SNR)

Analysis

- Inter-judge reliability (10%) = .86 -.99

- Nine one-way ANOVAs across groups (TD, NSMI, SMI) →Fisher's LSD post hoc
- Alpha= .05 (exploratory)
- Simultaneous multiple linear regression
- Incremental R^2

Results: descriptive

Table 2. Group means and standard deviations (in parenthesis) of acoustic variables and word intelligibility.

Speech subsystem and variables	Acoustic variable	SMI	NSMI	TD
Articulatory				
Vowel space	Vowel space (Hz^2)	542,914.0 (307,412.0)	907,432.0 (193,302.0)	957,023.0 (159,125.0)
Duration	Vowel duration (ms)	250.6 (145.9)	173.7 (28.0)	176.8 (32.5)
F2 slope	*Pipe* (Hz/ms)	6.9 (3.6)	9.4 (0.9)	9.9 (1.8)
	Whip (Hz/ms)	7.1 (4.5)	12.7 (2.3)	11.4 (2.4)
	Toys (Hz/ms)	7.5 (3.5)	12.3 (5.2)	10.1 (1.8)
F2 transitional duration	*Pipe* (ms)	152.3 (197.5)	111.7 (25.4)	128.9 (26.2)
	Whip (ms)	143.5 (108.5)	105.2 (20.8)	107.7 (23.3)
	Toys (ms)	209.7 (63.2)	188.5 (49.2)	199.8 (19.3)
F2 transitional extent	*Pipe* (Hz)	771.0 (399.0)	1,068.0 (261.0)	1,237.0 (239.0)
	Whip (Hz)	884.0 (549.0)	1,335.0 (388.0)	1,197.0 (297.0)
	Toys (Hz)	1285.0 (553.0)	2,048.0 (263.0)	1,962.0 (289.0)
Velopharyngeal				
A1–P1	*Big* (dB)	19.0 (4.9)	22.2 (3.2)	19.7 (2.9)
Laryngeal				
F0	*Top* (Hz)	274.4 (53.9)	249.5 (47.5)	237.6 (24.8)
SNR	*Top* (dB)	14.4 (3.2)	13.0 (3.9)	14.2 (1.9)
Word intelligibility (%)		44.1 (27.4)	80.3 (8.7)	82.0 (9.6)

Note. SNR = signal-to-noise ratio.

Results: ANOVAs

Table 3. ANOVA results of group main effects for each variable, plus effect sizes (η^2).

Source (group difference) and acoustic variable	Sum of squares	*df*	*F*	*p*	η^2
Vowel space (Hz^2)	1,422,537,649,612	2	14.310	< .0001**	.43
Error	1,888,722,377,633	38			
Vowel duration (ms)	49,731	2	3.368	.0450*	.15
Error	280,553	38			
F2 slope					
Pipe (Hz/ms)	75	2	6.507	.0037**	.26
Error	218	38			
Whip (Hz/ms)	208	2	10.158	.0003**	.35
Error	388	38			
Toys (Hz/ms)	122	2	5.518	.0079**	.23
Error	421	38			
A1–P1 (dB)					
Big (dB)	59	2	2.137	.1320	.10
Error	525	38			
F0					
Top (Hz)	10,556	2	3.313	.0551	.14
Error	64,067	38			
SNR					
Top (dB)	11	2	0.710	.4978	.04
Error	305	38			
Word intelligibility (%)	12,383	2	20.860	< .0001**	.52
Error	11,279	38			

*$p < .05$. **$p < .01$.

Results: ANOVAs

Source (group difference) and acoustic variable	*df*	*F*	*p*
Vowel space (Hz^2)	2	14.310	< .0001**
Error	38		
Vowel duration (ms)	2	3.368	.0450*
Error	38		
F2 slope			
Pipe (Hz/ms)	2	6.507	.0037**
Error	38		
Whip (Hz/ms)	2	10.158	.0003**
Error	38		
Toys (Hz/ms)	2	5.518	.0079**
Error	38		
A1–P1 (dB)			
Big (dB)	2	2.137	.1320
Error	38		
F0			
Top (Hz)	2	3.313	.0551
Error	38		
SNR			
Top (dB)	2	0.710	.4978
Error	38		
Word intelligibility (%)	2	20.860	< .0001**
Error	38		

*$p < .05$. **$p < .01$.

Results: Fisher's LSD

Table 4. Summary of significant variables per each group contrast based on Fisher's least significant difference post hoc test results.

Group and statistically significant variables	Direction of effect	Mean difference	*SE*	*p*
NSMI vs. TD				
None				
SMI vs. TD				
Vowel space	SMI < TD	–372,739.0	83749.0	< .0001
Vowel duration	SMI > TD	67.5	30.6	.0337
F2 slope *pipe*	SMI < TD	–2.7	0.9	.0043
F2 slope *whip*	SMI < TD	–4.1	1.1	.0007
Word intelligibility	SMI < TD	–37.9	6.2	< .0001
SMI vs. NSMI				
Vowel space	SMI < NSMI	296,313.0	105382.0	.0078
Vowel duration	SMI > NSMI	–76.9	37.3	.0459
F2 slope *pipe*	SMI < NSMI	2.1	1.1	.0480
F2 slope *whip*	SMI < NSMI	5.9	1.4	.0001
F2 slope *toys*	SMI < NSMI	4.6	1.5	.0042
Word intelligibility	SMI < NSMI	36.2	7.5	< .0001

Note. The "Mean difference" column shows the actual magnitudes of the effects. The sign of the effect reflects the mean value of the second group in each pair subtracted from the mean value of the first group in the pair.

Results: Fisher's LSD

Group and statistically	Direction of effect	*p*
NSMI vs. TD		
None		
SMI vs. TD		
Vowel space	SMI < TD	< .0001
→ Vowel duration	SMI > TD	.0337
F2 slope *pipe*	SMI < TD	.0043
F2 slope *whip*	SMI < TD	.0007
Word intelligibility	SMI < TD	< .0001
SMI vs. NSMI		
Vowel space	SMI < NSMI	.0078
→ Vowel duration	SMI > NSMI	.0459
F2 slope *pipe*	SMI < NSMI	.0480
F2 slope *whip*	SMI < NSMI	.0001
F2 slope *toys*	SMI < NSMI	.0042
Word intelligibility	SMI < NSMI	< .0001

Criteria for multiple linear regression

- “Six predictor variables were chosen according to the following three criteria:
  - a) at least one measure to represent each of the three subsystems,
  - b) low correlations with other potential predictor variables, and
  - c) previous evidence in the literature of sensitivity of the variable to dysarthria.” (Lee et al. 2014)

Results: Correlations

Table 5. Correlation coefficient matrix of the six selected acoustic variables for children with CP ($n = 22$).

Variable	1	2	3	4	5	6
1. Vowel space	—	−.599**	.603**	.002	−.419	−.345
2. Vowel duration		—	−.728**	−.311	−.064	.528*
3. Average F2 slope			—	.427*	.002	−.427*
4. A1–P1				—	.093	−.264
5. F0					—	.325
6. SNR						—

*$p < .05$. **$p < .01$.

Results: Simultaneous MLR

Table 6. Beta coefficients of the simultaneous multiple linear regression model against word intelligibility for children with CP (n = 22).

Predictor variable	Unstandardized coefficients (*B*)	Standardized coefficients (β)	*t*	*p*	VIF
Vowel space	< 0.0001	.200	1.137	.273	3.268
Vowel duration	–0.074	–.313	–1.743	.102	3.410
A1–P1	0.657	.106	.901	.382	1.453
Average F2 slope	4.350	.509	3.035	.008	2.969
F0 of *top*	–0.225	–.414	–2.968	.010	2.058
SNR of *top*	2.240	.275	2.066	.057	1.873

Note. VIF = variance inflation factor.

Results: Simultaneous MLR

Predictor variable	*t*	*p*	VIF
Vowel space	1.137	.273	3.268
Vowel duration	−1.743	.102	3.410
A1–P1	.901	.382	1.453
Average F2 slope	3.035	.008	2.969
F0 of *top*	−2.968	.010	2.058
SNR of *top*	2.066	.057	1.873

Results: Subsystem MLR

Table 7. Incremental R^2 change results of the multiple linear regression model examining each speech subsystem's independent contribution to speech intelligibility in children with CP ($n = 22$).

Rank	Second block speech subsystem	R^2 changes with second block speech subsystem	R^2 with the remaining two speech subsystems in the first block
1	Articulatory	.579	.279
2	Laryngeal	.088	.770
3	Velopharyngeal	.008	.850

Note. Total $R^2 = .858$, adjusted $R^2 = .801$.

Results: Subsystem MLR

Rank	Second block speech subsystem	R^2 changes with second block speech subsystem
1	Articulatory	.579
2	Laryngeal	.088
3	Velopharyngeal	.008

Note. Total R^2 = .858, adjusted R^2 = .801.

Results: Subsystem MLR

Rank	Second block speech subsystem	R^2 with the remaining two speech subsystems in the first block
1	Articulatory	.279
2	Laryngeal	.770
3	Velopharyngeal	.850

Note. Total R^2 = .858, adjusted R^2 = .801.

Results: Incremental R^2 MLR

Table 8. Incremental R^2 change results of the multiple linear regression model examining each variable's independent contribution to speech intelligibility in children with CP ($n = 22$).

Rank	Sixth acoustic variable in the second block	R^2 change with the sixth variable in the second block	R^2 with remaining five variables in the first block
1	Average F2 slope	.087	.771
2	F0	.083	.775
3	SNR	.040	.818
4	Vowel duration	.029	.829
5	Vowel space	.012	.846
6	A1–P1	.008	.850

Note. Total $R^2 = .858$, Adjusted $R^2 = .801$.

Results: Incremental R^2 MLR

Rank	Sixth acoustic variable in the second block	R^2 change with the sixth variable in the second block
1	Average F2 slope	.087
2	F0	.083
3	SNR	.040
4	Vowel duration	.029
5	Vowel space	.012
6	A1–P1	.008

Note. Total R^2 = .858, Adjusted R^2 = .801.

Results: Incremental R^2 MLR

Rank	Sixth acoustic variable in the second block	R^2 with remaining five variables in the first block
1	Average F2 slope	.771
2	F0	.775
3	SNR	.818
4	Vowel duration	.829
5	Vowel space	.846
6	A1–P1	.850

Note. Total R^2 = .858, Adjusted R^2 = .801.

Discussion

- SMI ≠ NSMI = TD articulatory subsystem
- SMI = NSMI = TD velopharyngeal & laryngeal
- TD data correspond with literature

Discussion

- Childhood dysarthria = adult dysarthria
  - Smaller vowel space area
  - Longer vowel durations
  - (higher mean F0)
  - Lower speech intelligibility scores
  - Shallower F2 slope (duration & extent)
    - *But no statistical tests for duration & extent !*

Discussion

- Multiple linear regression
- Acoustic variables → 80% of variation in speech intelligibility scores (???)
- Articulatory subsystem (F2 slope, VSA, vowel duration) → 58% (???)

Discussion

- Limitations
- Small sample size
- “wide” age range
- Single words for intelligibility scores
- Heterogeneous CP population
- Laryngeal and velopharyngeal measures
- No measure of respiratory subsystem

Conclusion

- Articulatory subsystem involved in childhood dysarthria
- F0 might contribute to speech intelligibility
- Childhood dysarthria = adult dysarthria
  - No evidence to suggest that the speech motor disorder interacts with developing speech motor capabilities

Questions?