

Similarity of Prosody Between Speech and Singing: A Methodological Study

JESSICA STEWART

FACULTY ADVISORS: DR. RONALD SCHERER AND DR. GEOFFREY
STEPHENSON



Purpose of the Study

The purpose of this study was to identify how the spoken and musical prosody of simple songs compare.

Research Question:

- Comparing the acoustic measures of rhythm, intensity, and intonation: **how does the spoken prosody of simple songs compare to the sung prosody of the same songs?**
- **Which song was the most speech like?**
- **Which song was the least speech like?**

Background: Neuro Pathologies

Neurogenic communication disorders: effects regions that control the movements necessary to produce speech

Impairment of Speech Prosody

Common Causes: stroke (aphasia) and Parkinson's Disease

Characteristics:

- disturbances in muscular control
- muscle weakness
- incoordination of the mechanisms responsible for speech

Prosodic Deviancies: abnormal variations of the intonation, stress, and duration of speech

Singing as Therapy: Parkinsong

Target Population: people with Parkinson's Disease (PD)

Characteristics of PD: Rigidity, limited pitch range, limited intensity range

Benefits of Singing for PD:

- **Increased respiratory support**
- **Increased phonation**
- **Increased pitch variation**
- **Clearer articulation**

The outcomes of Parkinsong is comparable to the gold standard treatment for PD voice treatment: LSVT LOUD



<https://www.swtimes.com.au/news/south-western-times/parkinsong-a-joyous-form-of-music-therapy-ng-b881009723z>

Singing as Therapy: Melodic Intonation Therapy (MIT)

Target Population: people with Broca's Aphasia

- Ability to naturally converse is **impaired**, ability to sing the lyrics of a familiar song is **intact**
- MIT: a speech therapy which attempts redevelop language and speech prosody through singing and tapping
- **Early stages:** client's work on singing target phrases to tune of familiar songs
- **Late stages:** Transition to speech singing or Sprechgesang to further develop speech prosody



What's for din - ner?



What did you say?

The Present Study

There is still little understanding of how the prosody of singing could be best utilized in speech therapy

- Knowing the differences between sung and spoken prosody may help clinicians choose better targets for therapy.

Ex. Goal: Normal prosody

have the patient sing songs with prosody similar to speech.

Ex. Goal: Exaggerated prosody

have the patient sing songs with enhanced or broadened prosodic elements.



Methods Participant: Selection

- Participants were recruited by flyers in the Health and Human Services Building and College of Musical Arts
- Prior to selection prospective participants underwent a short telephone screening to verify eligibility.

SPEECH ELEGIBILITY CRITERIA:

- Be 18 years or older
- physically healthy the day of recording
- no history of voice and speech disorders
- native speaker of English
- speak with a Standard American English dialect.

SINGING ELIGIBILITY CRITERIA:

- Be a current voice student or musical theater major
- be 18 years of age or older at time of participation
- physically healthy the day of recording
- no history of voice and speech disorders
- native speaker of English
- speak with a Standard American English Dialect
- read and perform selections accurately with minimal practice.

Methods: Participants

At the time of participation all participants were current students at BGSU.

Singing Group:

- Two females (designated as F2 and F3)
- Had soprano voices
- Met all eligibility criteria

Speech Group:

- One male and one female (designated as M1 and F1)
- Met all eligibility criteria



Methods: Song Selection

In consultation with an expert (Dr. Geoffrey Stephenson), the songs chosen for this project were based on:

- **Simplicity of melodic structure**
- **Familiarity across a wide range of ages**
- **Prior usage in studies comparing speech and musical prosody**

The genre of **children's music** was determined to be best suited for this project.

The songs chosen were

- "Row, Row, Row Your Boat"
- "Happy Birthday To You"
- "Twinkle, Twinkle Little Star"



Methods: Materials

- A soundproof recording booth
- A head mounted microphone
- The music or lyric sheet
- An Electronic keyboard
- Praat recording software



Methods: Procedure

1. Participants were seated in the recording booth along with a researcher (JS)
2. Prior to recording, the participants were instructed to perform vocal exercises served as a way for the experimenter make sure the equipment was working correctly
3. Participants received instructions on how speak or sing
 - **Speech Group:** comfortable volume in a casual manner
 - **Singing Group:** comfortable volume with neutral intent



Methods: Procedure

4. Participants were asked to either sing or read aloud each selection at least three times

- After the third trial, the experimenter (JS or RS) would ask for additional trials if needed.

The recordings and analysis were performed using:

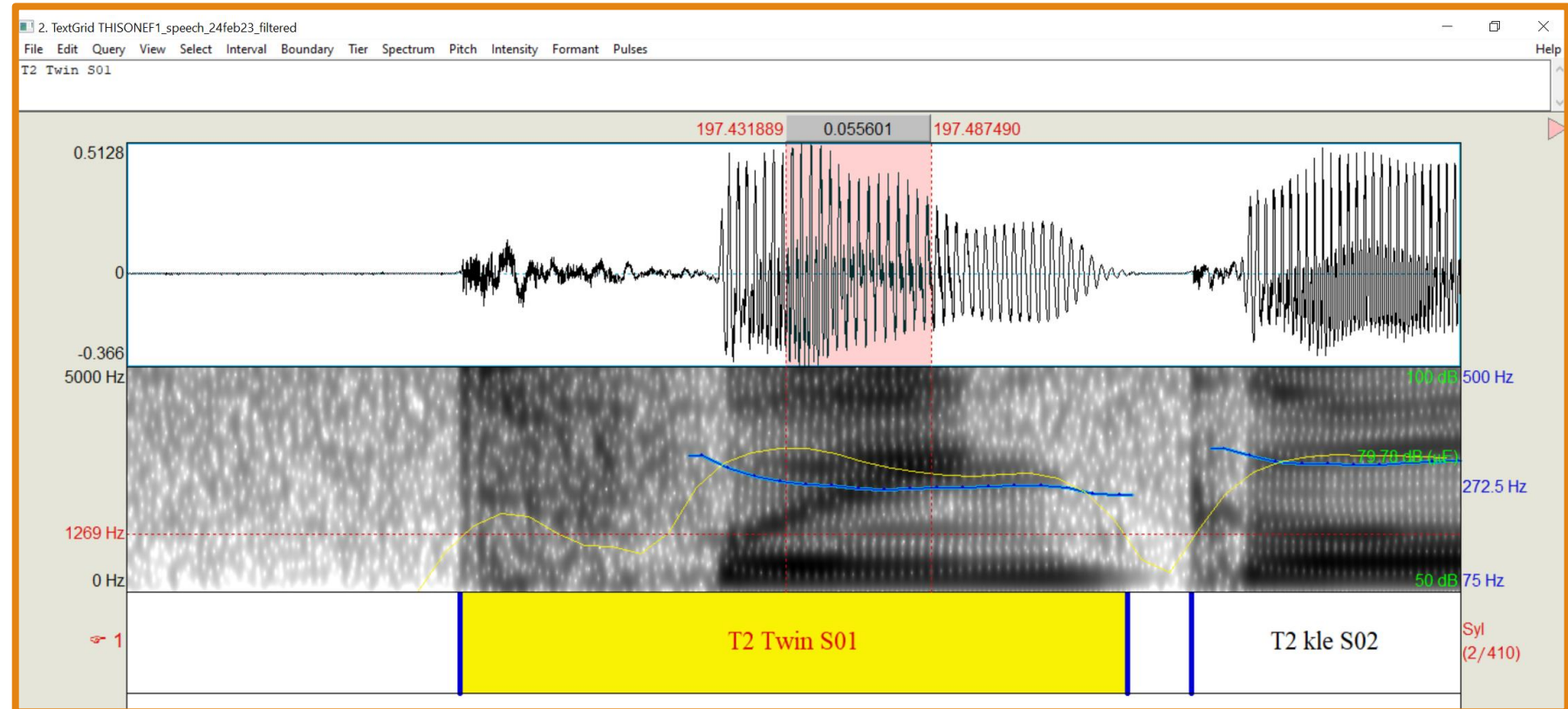
- electret microphone system
- Praat recording software

Analysis Stage 1: Preparing Data

Prosodic Measures:

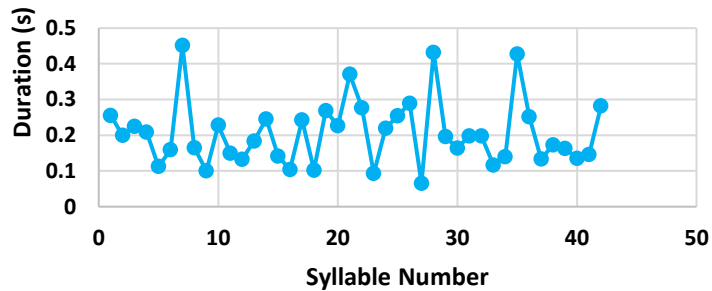
- syllabic duration
- fundamental frequency contour (blue line)
- dB contour (yellow line)

Ex. F1 Trial 2 of Twinkle

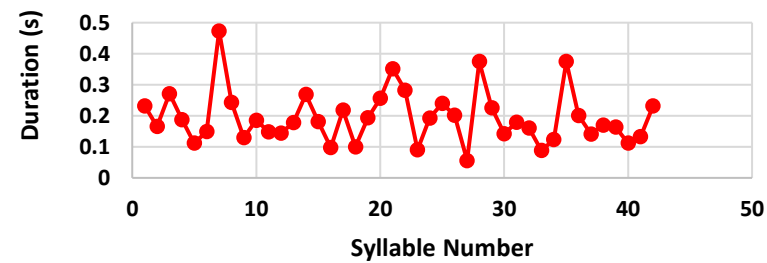


Analysis Stage 2: Averaging Trials

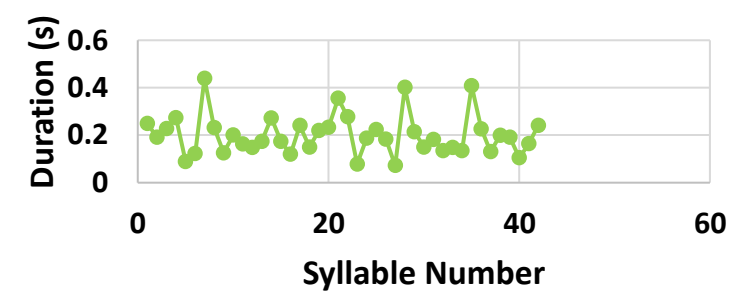
F1 Twinkle T2 Duration



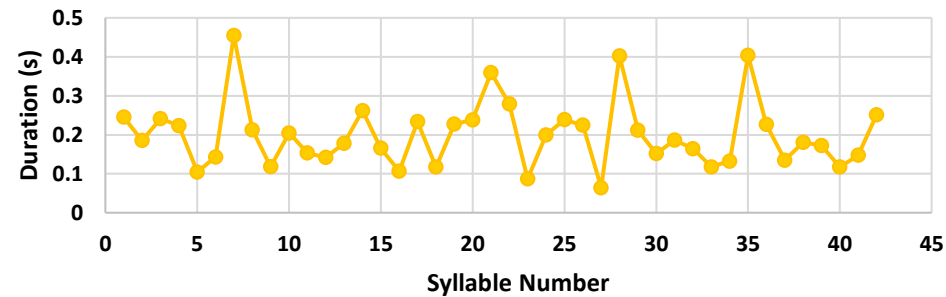
F1 Twinkle T4 Duration



F1 Twinkle T5 Duration

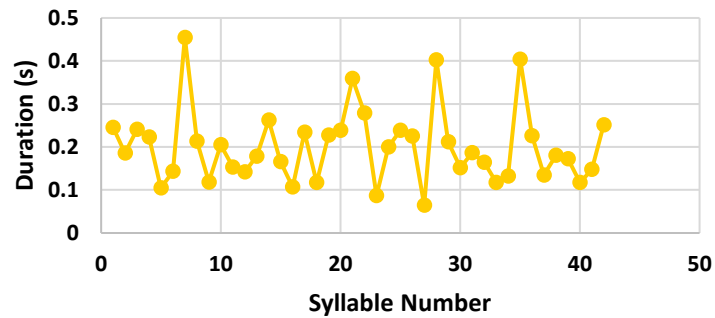


F1 Twinkle Average Duration

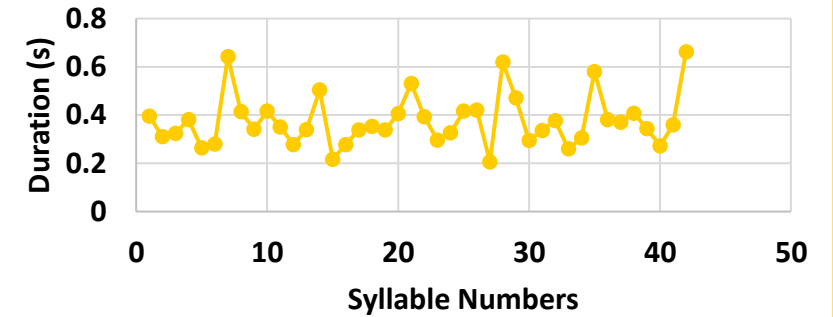


Analysis Stage 3: Averaging Participants

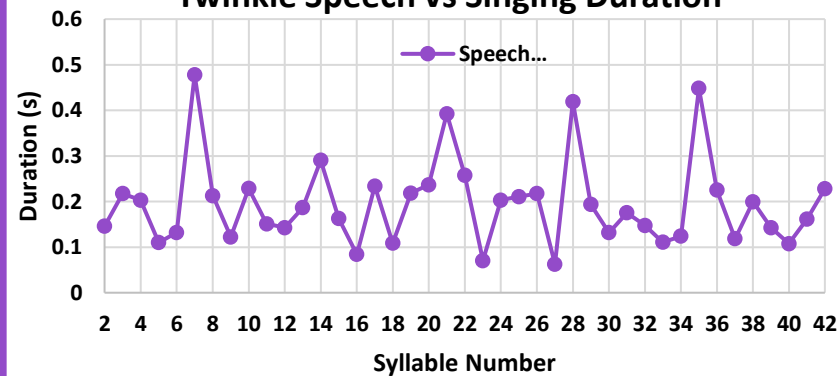
F1 Twinkle Average Duration



M1 Twinkle Average Duration



Twinkle Speech vs Singing Duration



Results: “Row, Row, Row Your Boat” Duration

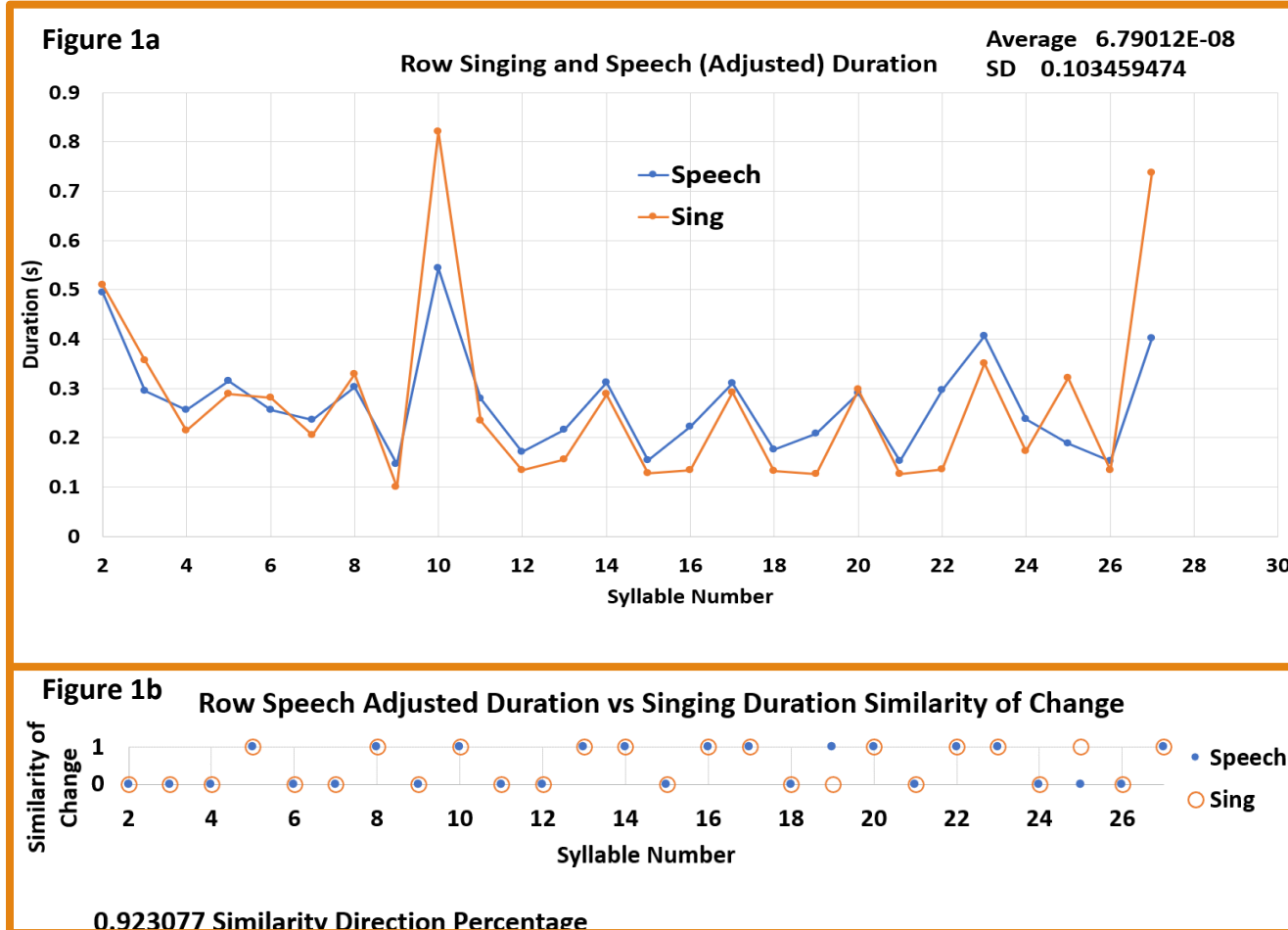


Figure 1a

- Change in Direction (greater duration or less duration) from one syllable to the next
- Singing and the speaking **overlap substantially**

Figure 1b

- Each of the duration changes and the percentage of changes that are in the same direction
- Similarity of Change Percentage = **92%**.

Results: “Row, Row, Row Your Boat” Duration

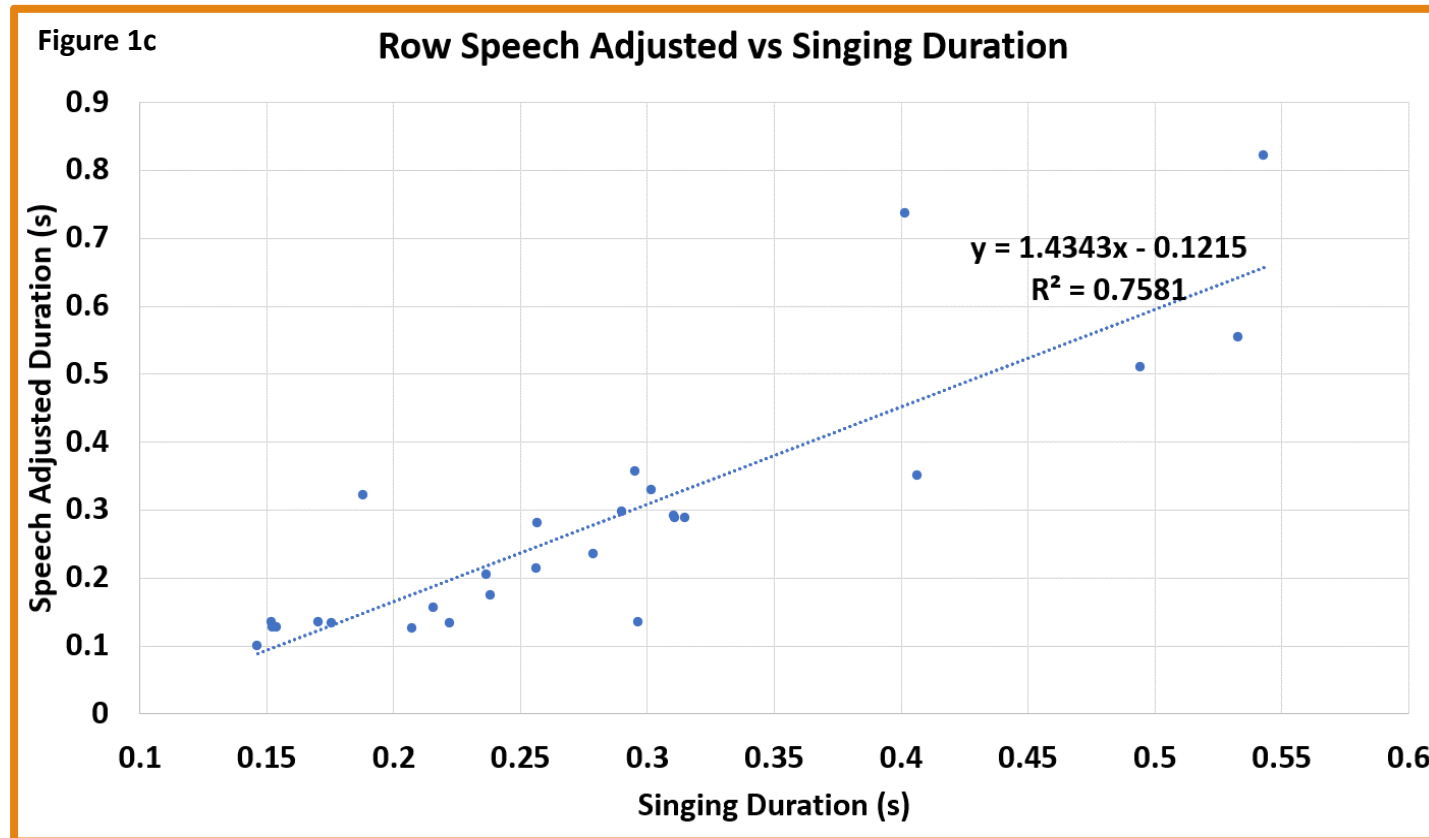


Figure 1c

- Relationship between the duration of the spoken syllables compared to sung syllables
- **High correlation** of $r = 0.87$
 - (the square root of the R^2 in the figure)

Results: “Row, Row, Row Your Boat” Intensity

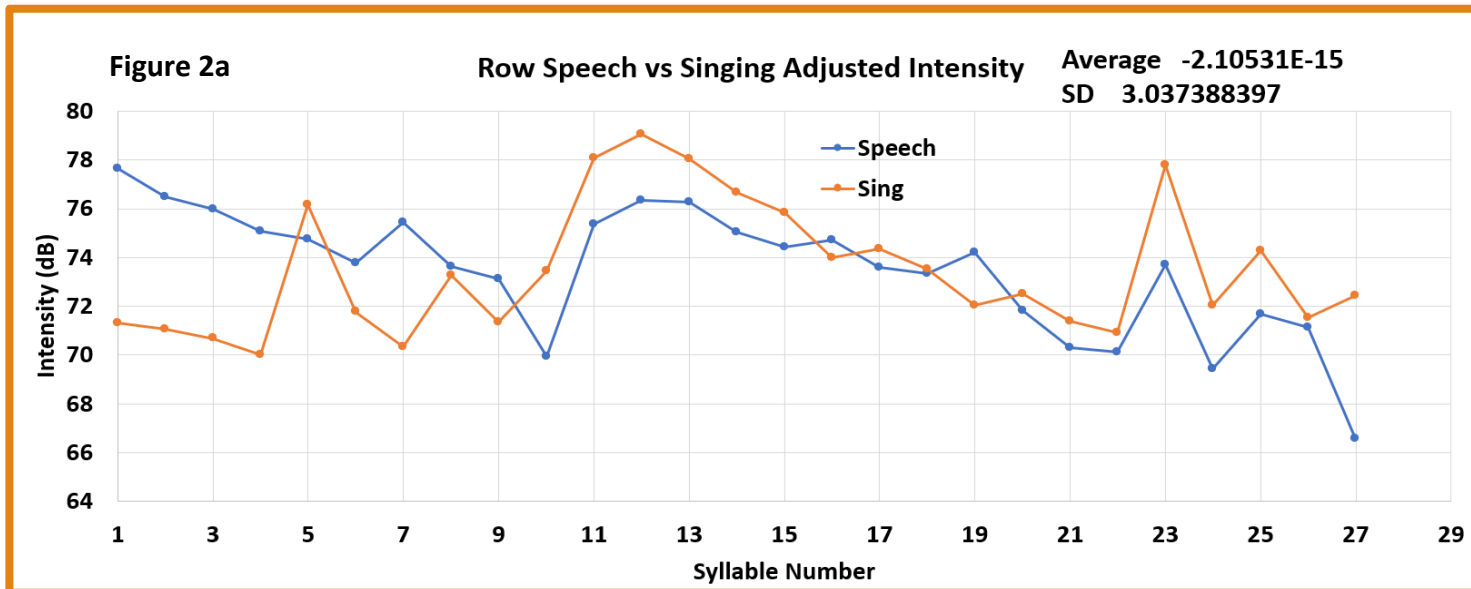


Figure 2a

- Singing and the speaking intensities **overlap somewhat**, with the change in direction

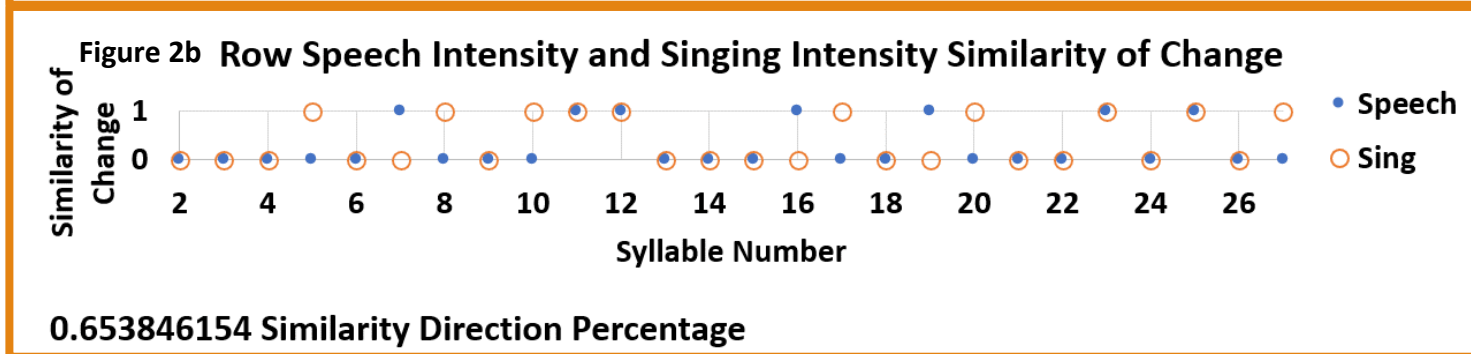


Figure 2b

- Similarity of Change Percentage = **65%**.
- Speaking and singing result in **somewhat similar** syllable intensities.

Results: “Row, Row, Row Your Boat” Intensity

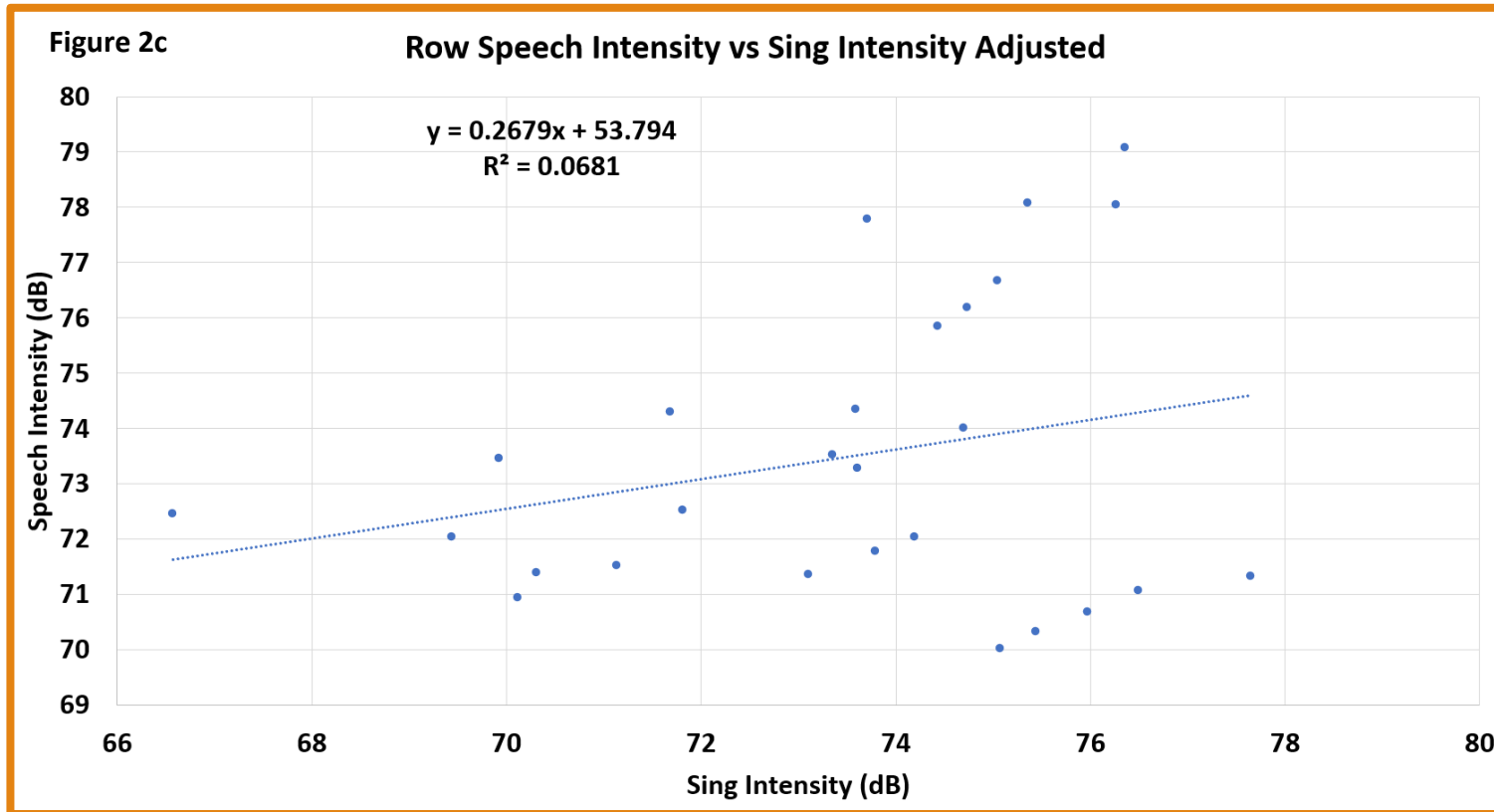


Figure 2c

- **Weak correlation** of $r = 0.26$ (the square root of the R^2 in the figure).
- The primary differences in intensity between spoken and sung occur at higher intensities (above 74 dB).

Results: “Row, Row, Row Your Boat” Semitones

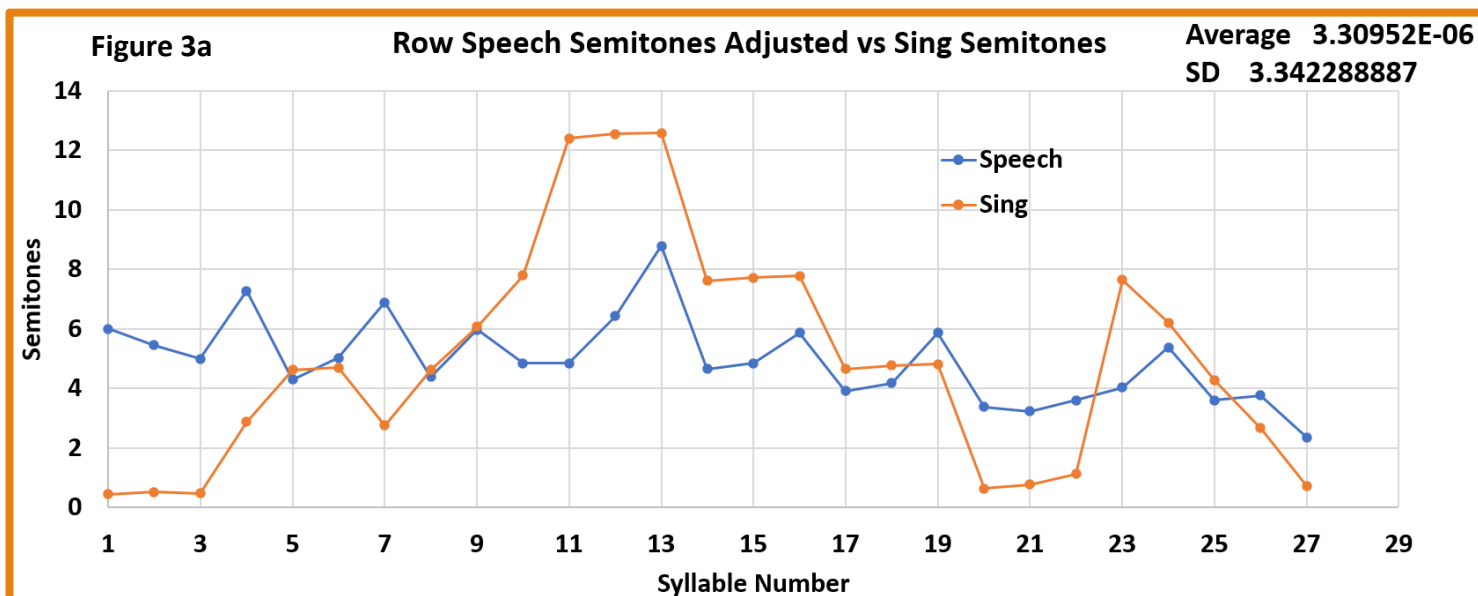


Figure 3a

- Singing and the speaking **overlap somewhat**
- Change in direction from one syllable to the next was **quite variable**.

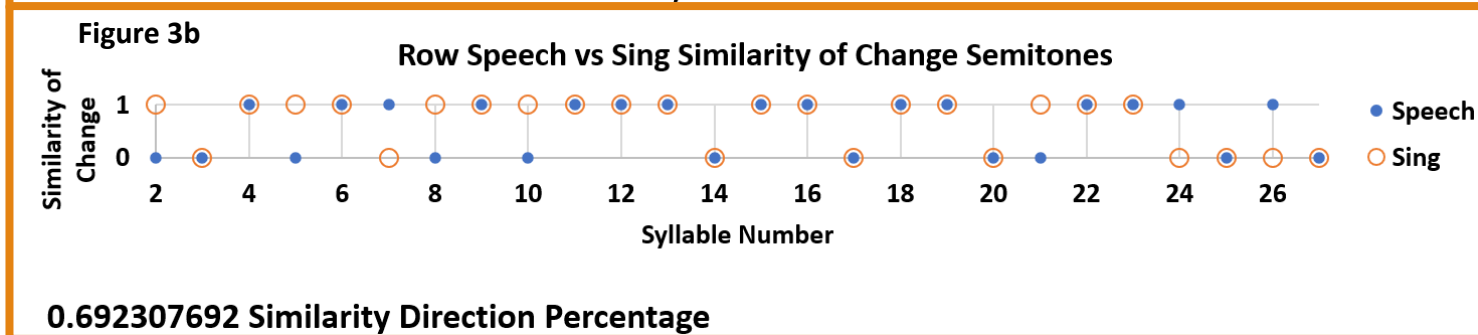


Figure 3b

- Similarity of Change Percentage = **69%**
- Speaking and singing result in **somewhat similar** syllable semitones

Results: “Row, Row, Row Your Boat” Semitones

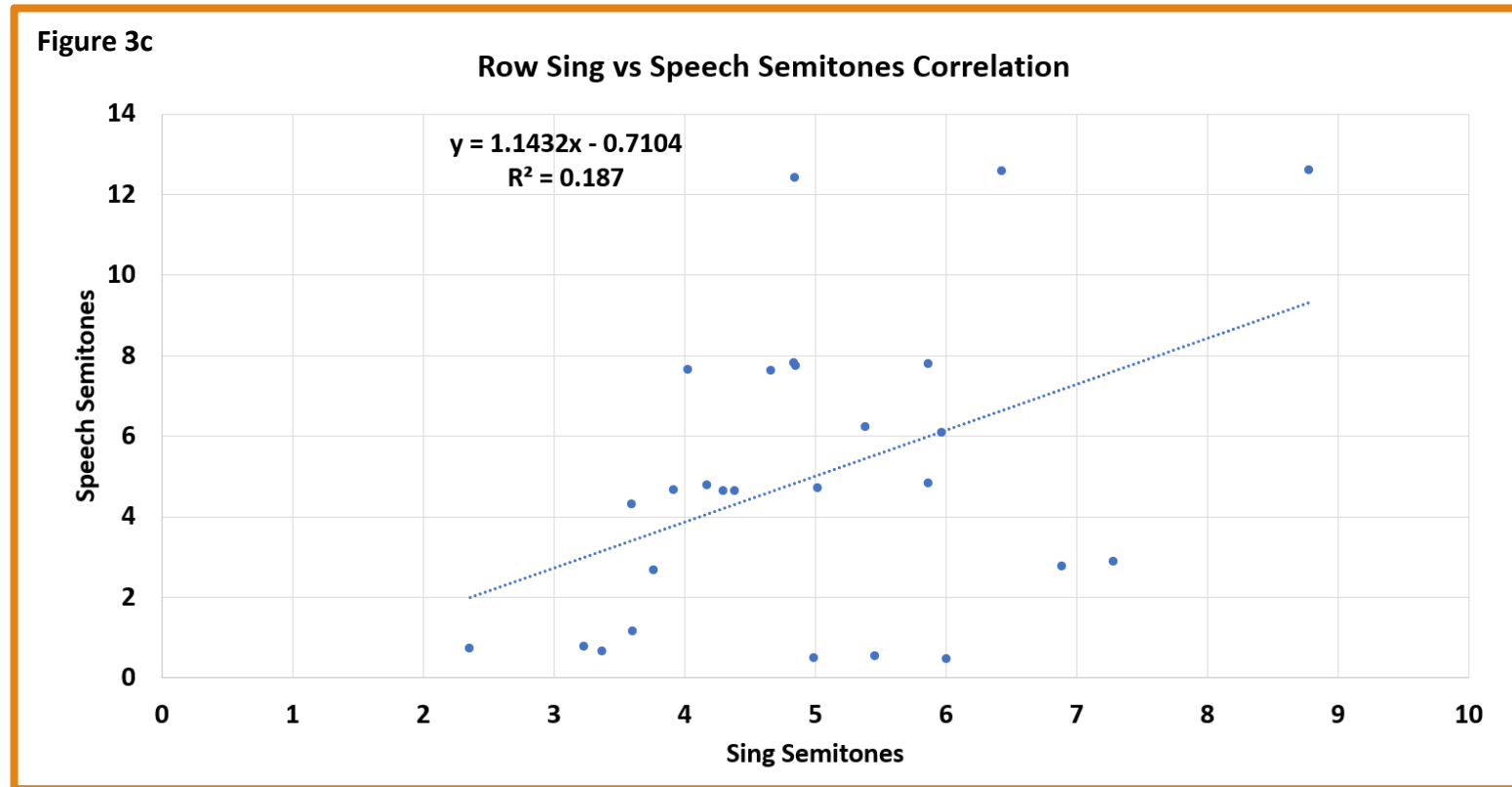


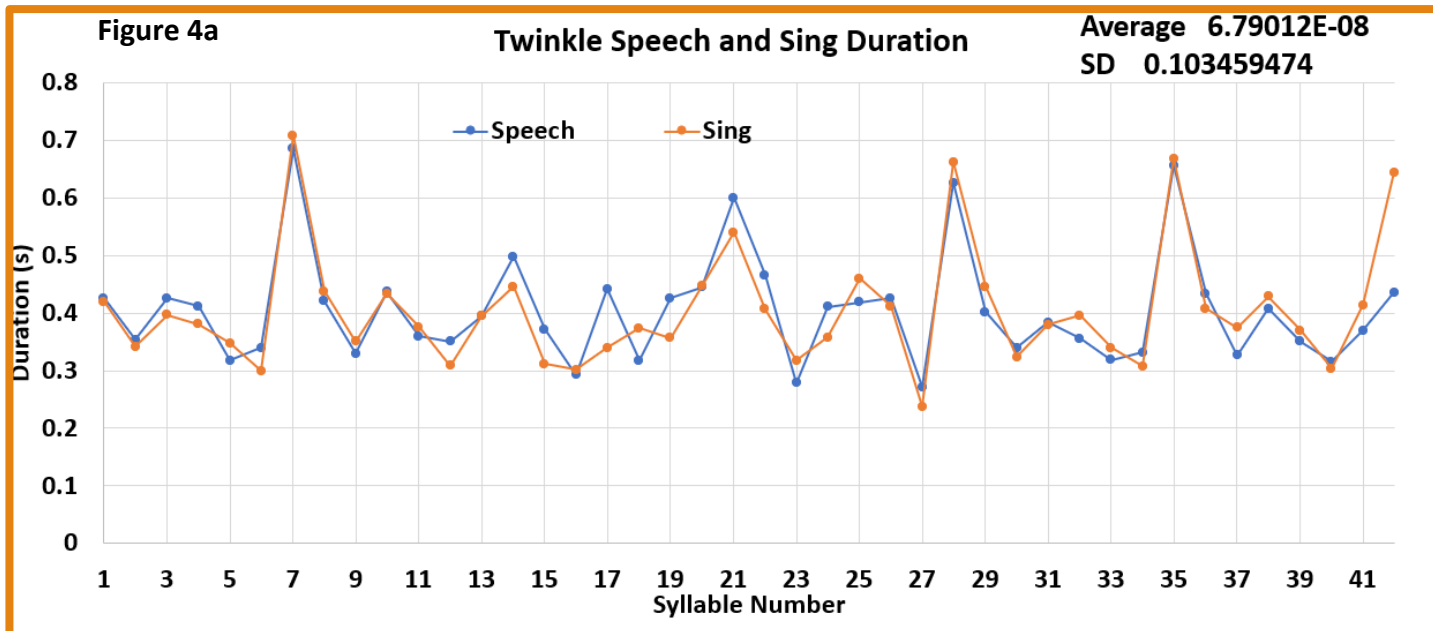
Figure 3c

Shows the relationship between the semitones of the spoken syllables compared to the semitones of the sung syllables,

- **Moderately low** correlation of $r = 0.43$ (the square root of the R^2 in the figure).

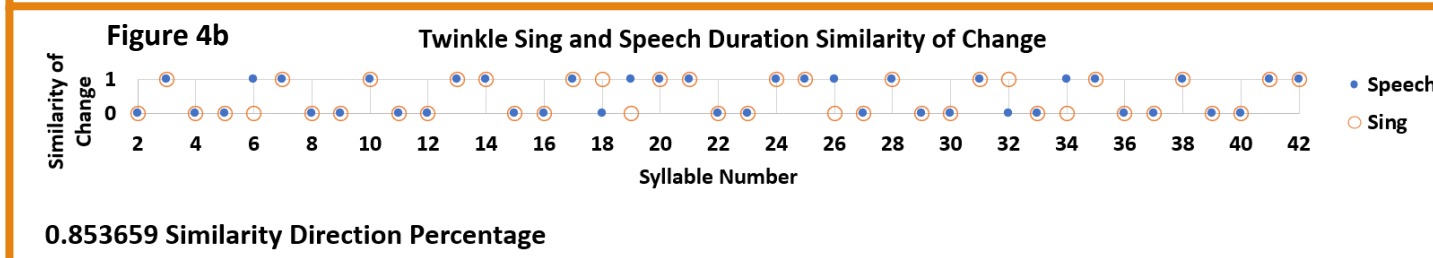
Results: “Twinkle, Twinkle Little Star” Duration

Figure 4a



- Singing and the speaking overlap **substantially**, with the change in direction

Figure 4b



- Similarity of Change Percentage = **85%**
- For duration and this song, speaking and singing result in **similar** syllable durations.

Results: “Twinkle, Twinkle Little Star” Duration

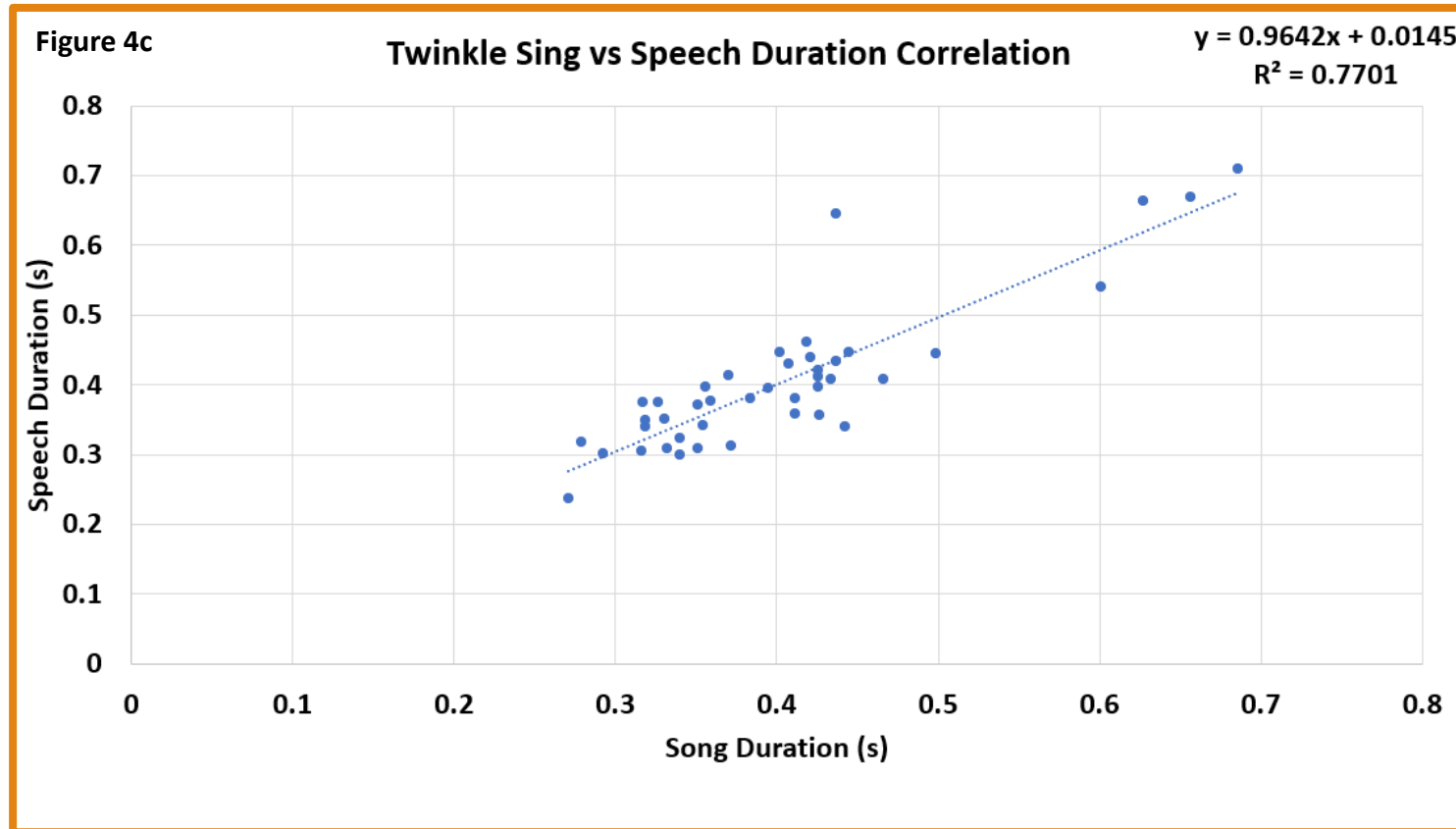


Figure 4c

Shows the relationship between the duration of the spoken vs sung syllables

- **High correlation** of $r = 0.88$ (the square root of the R^2 in the figure).

Results “Twinkle, Twinkle Little Star” Intensity

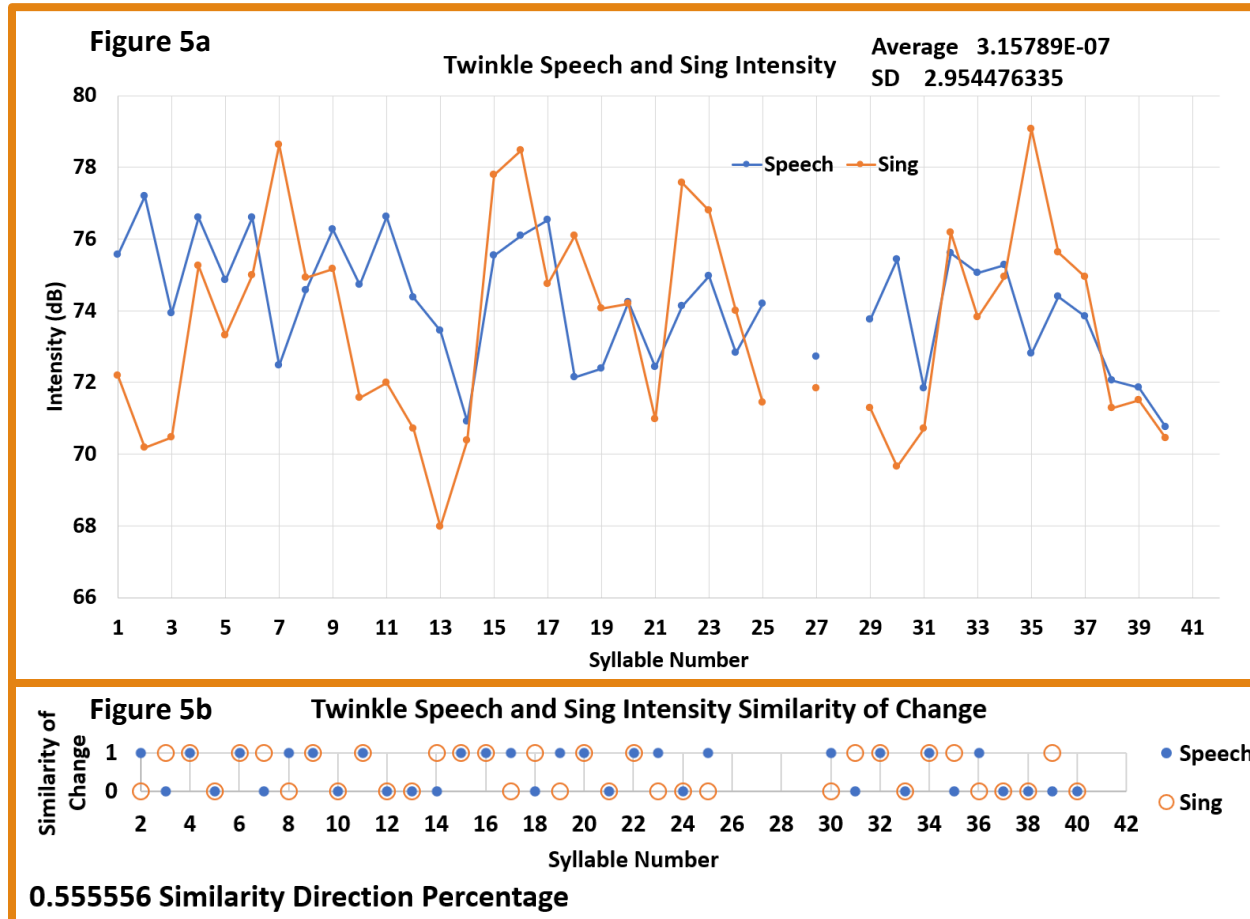


Figure 5a

- Singing and the speaking **overlap rarely**, with one syllable to the next being **quite variable**.

Figure 5b

- Similarity of Change Percentage = **56%**.
- Speaking and singing result in **somewhat dissimilar** syllable intensities.

Results: “Twinkle, Twinkle Little Star” Intensity

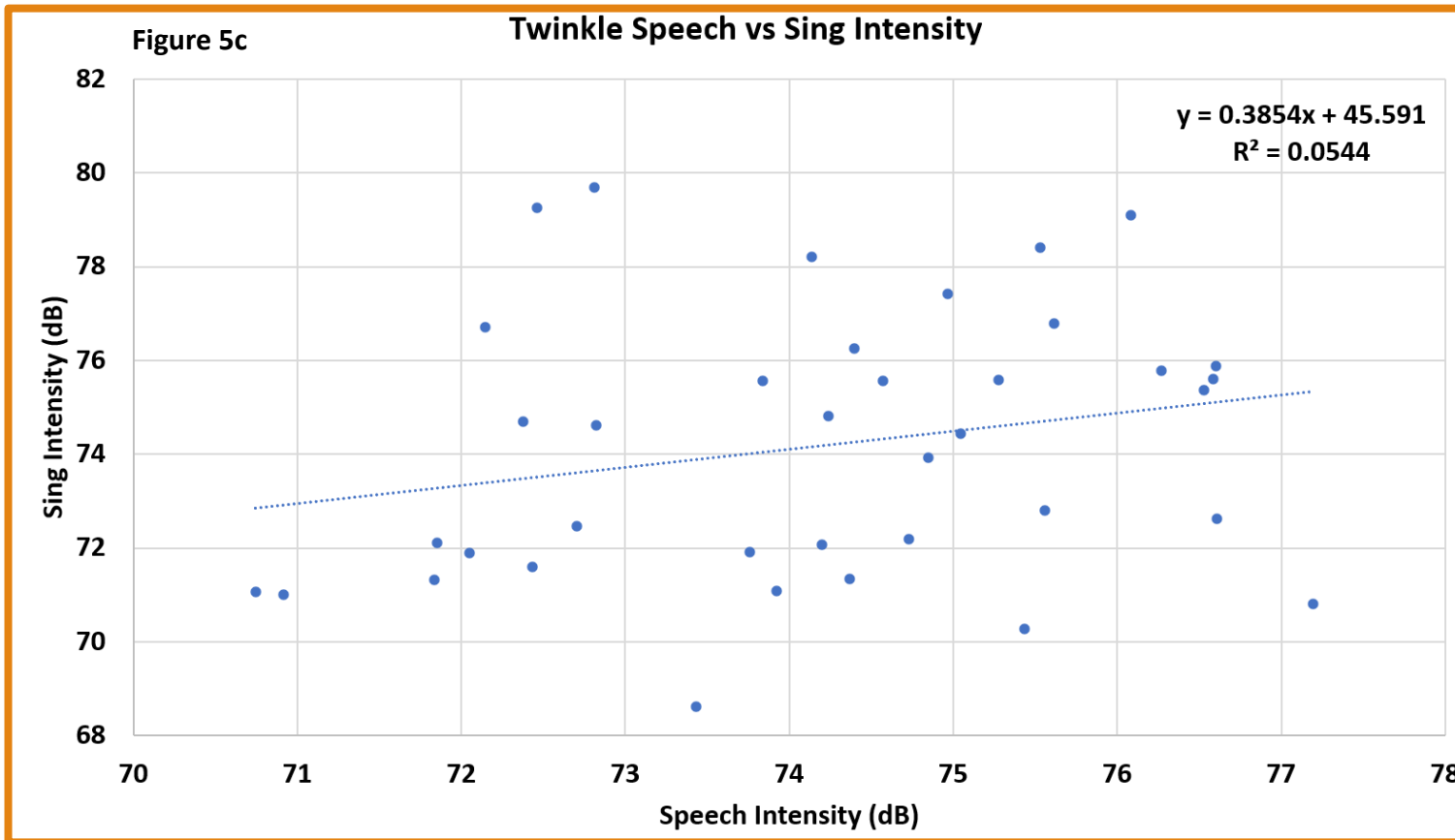


Figure 5c

Shows the relationship between the intensity of the spoken vs sung syllables

- **Weak correlation** of $r = 0.23$ (the square root of the R^2 in the figure).

Results: “Twinkle, Twinkle Little Star” Semitones

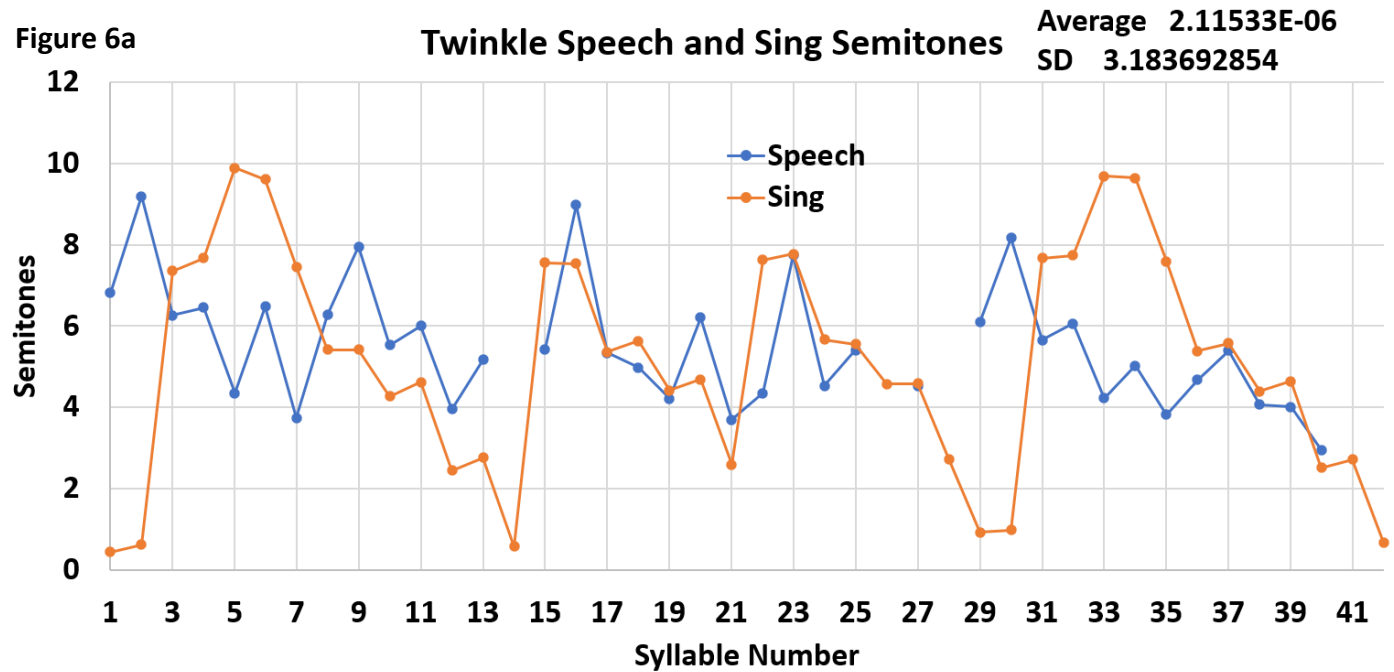


Figure 6a

- Indicates that the singing and the speaking **overlap rarely**
- **Quite variable**, although grossly in the same directions.

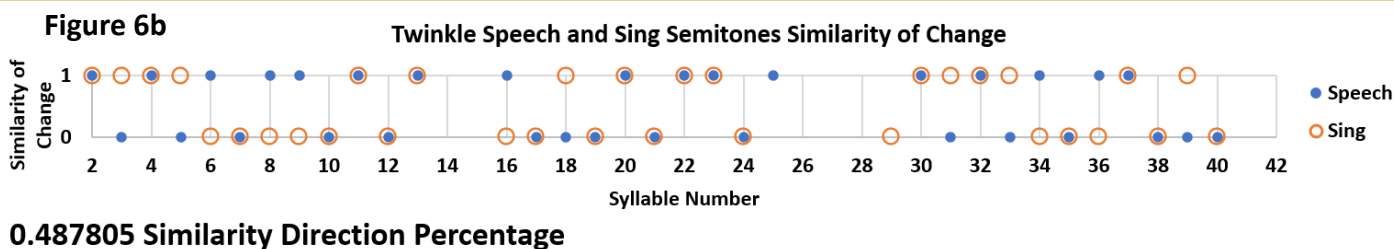


Figure 6b

- Similarity of Change Percentage = **49%**.
- Comparison between speaking and singing result in **dissimilar** pitch variation.

Results: “Twinkle, Twinkle Little Star” Semitones

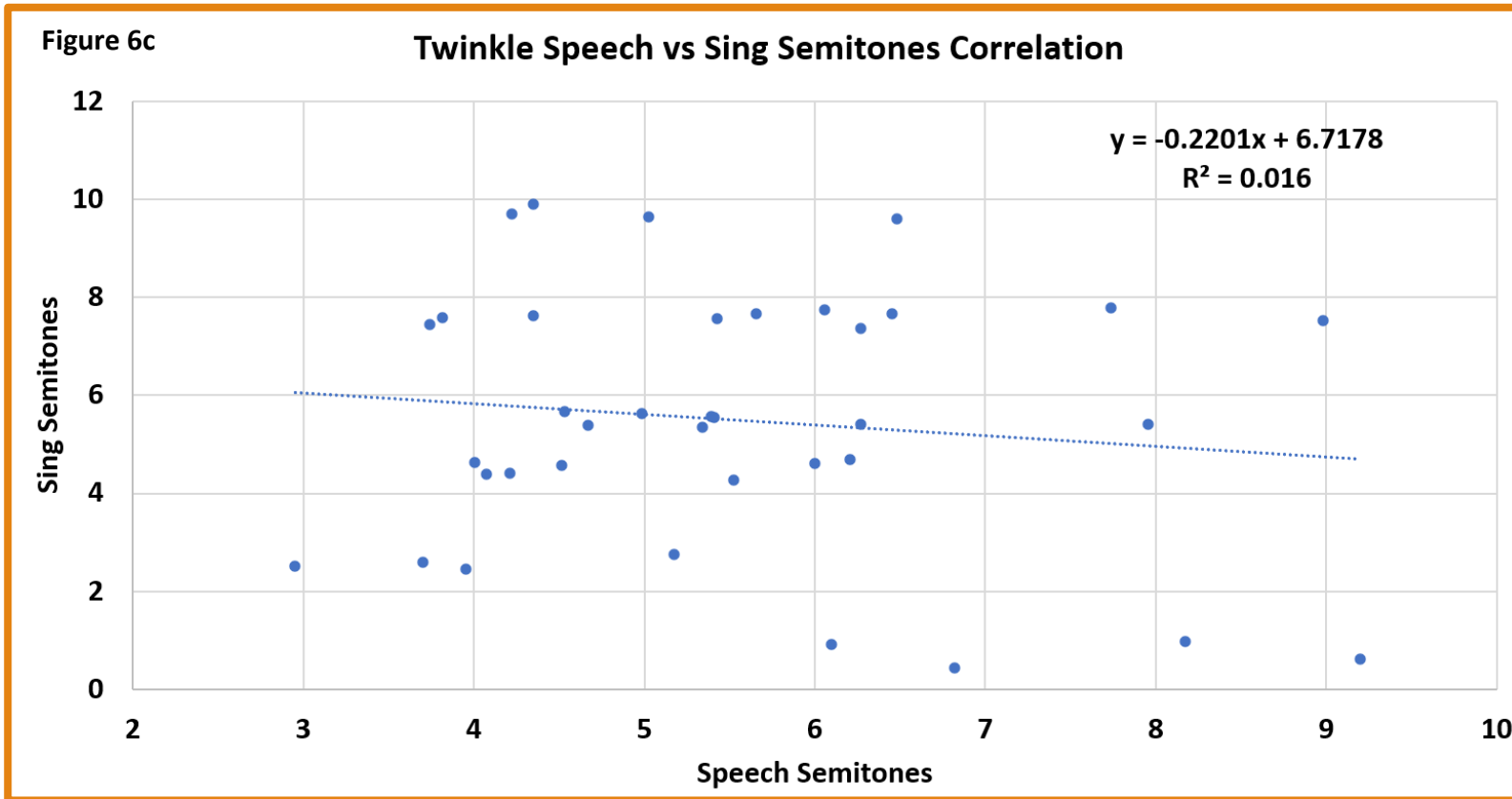


Figure 6c

- Relationship between the semitones of the spoken compared to the sung syllables.
- **Very weak** correlation of $r = 0.13$ (the square root of the R^2 in the figure),
- Essentially no relationship.

Results: “Happy Birthday To You” Duration

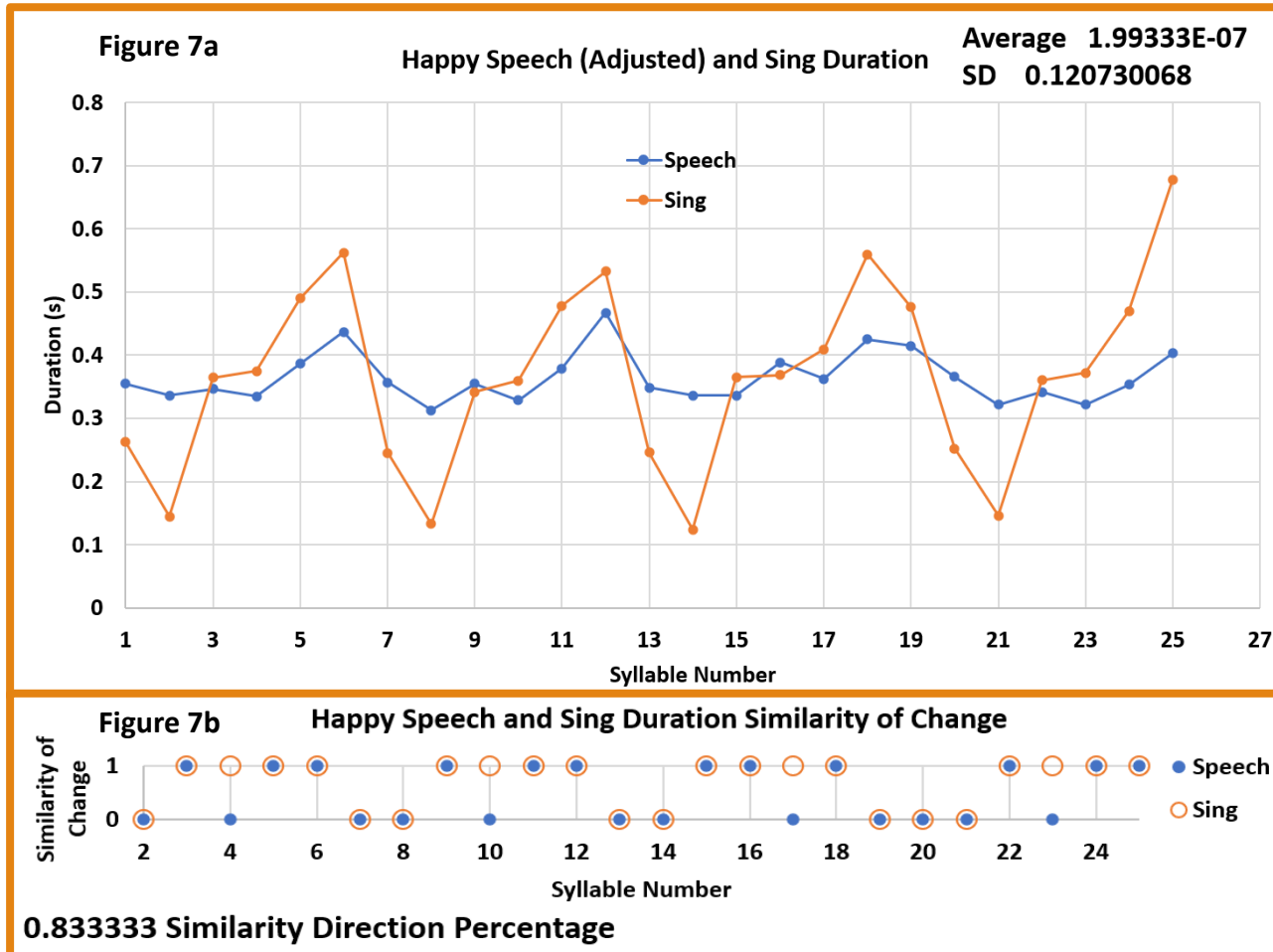


Figure 7a

- Singing and the speaking overlap **somewhat**
- Greater extension for the sung durations.

Figure 7b Similarity of Change
Percentage = **83%**

- Speaking and singing result in similar syllable duration directions.

Results: “Happy Birthday To You” Duration

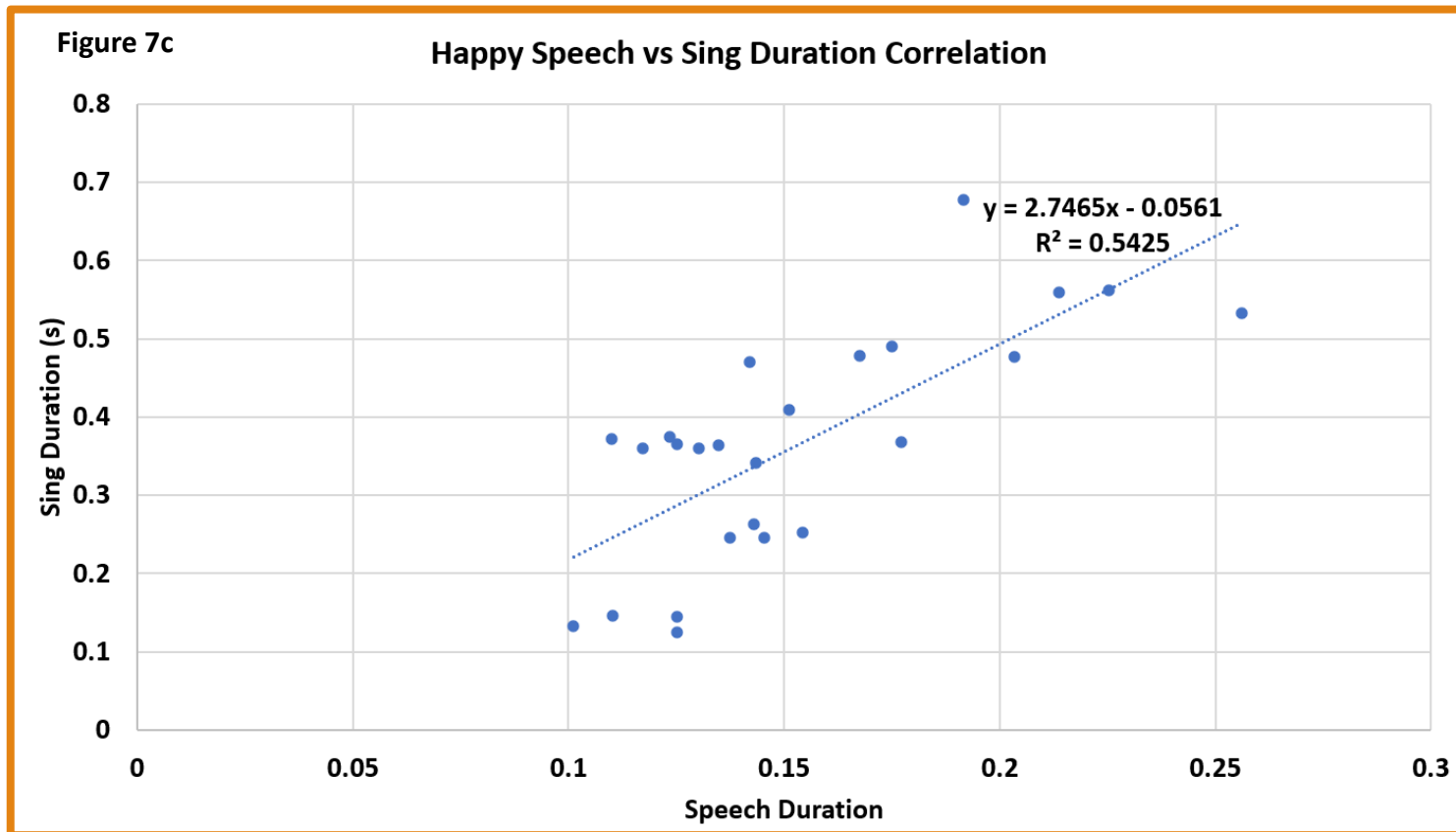


Figure 7c

Shows the relationship between the duration of the spoken vs sung syllables

- **High correlation** of $r = 0.73$ (the square root of the R^2 in the figure).

Results: “Happy Birthday To You” Intensity

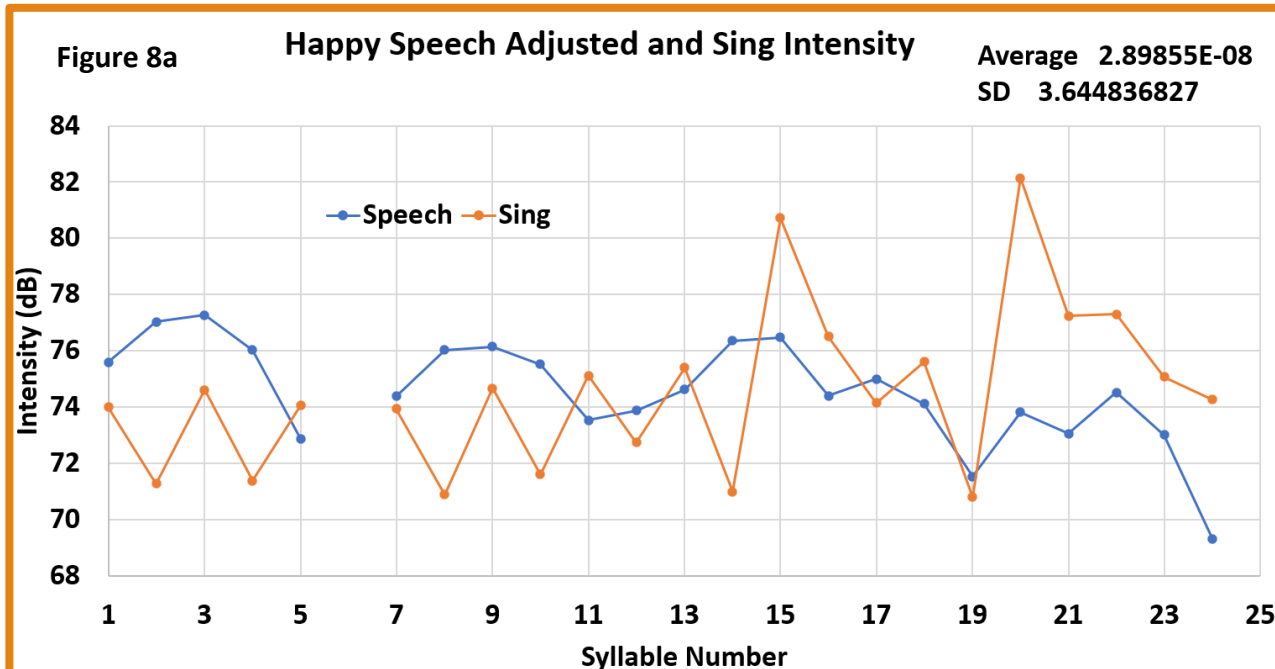


Figure 8a

- Singing and the speaking intensities overlap **rarely**, with one syllable to the next being **quite dissimilar**

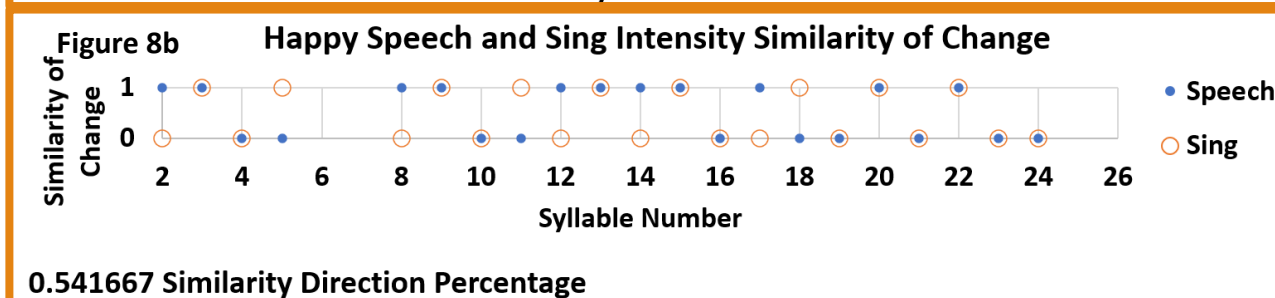


Figure 8b

Similarity of Change Percentage = **54%**

- Speaking and singing result in **dissimilar** syllable intensities

Results: “Happy Birthday To You” Intensity

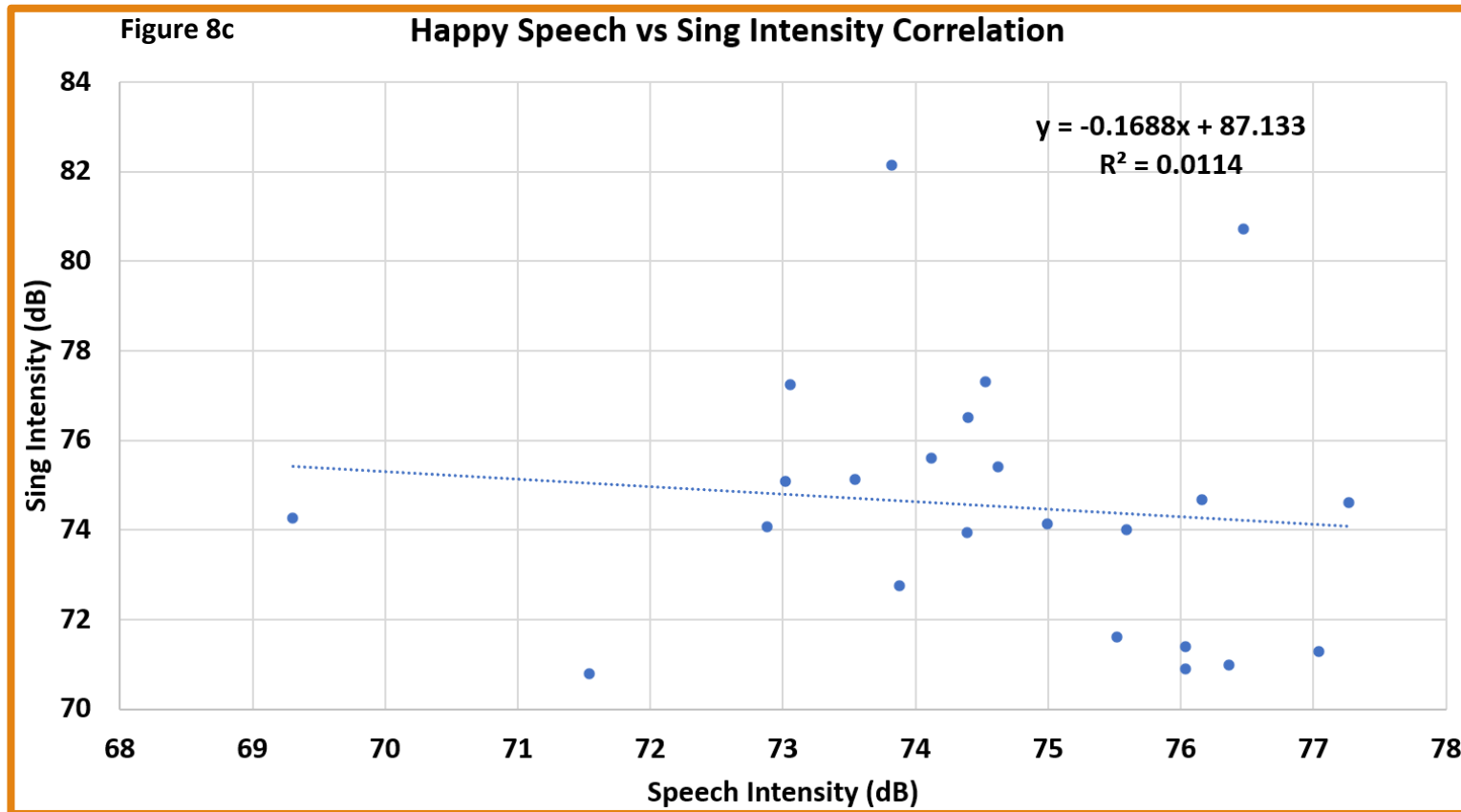


Figure 8c

shows the relationship between the intensity of the spoken vs sung syllables

- **Very weak** correlation of $r = 0.11$ (the square root of the R^2 in the figure), basically no relation between the two.

Results: “Happy Birthday To You” Semitones

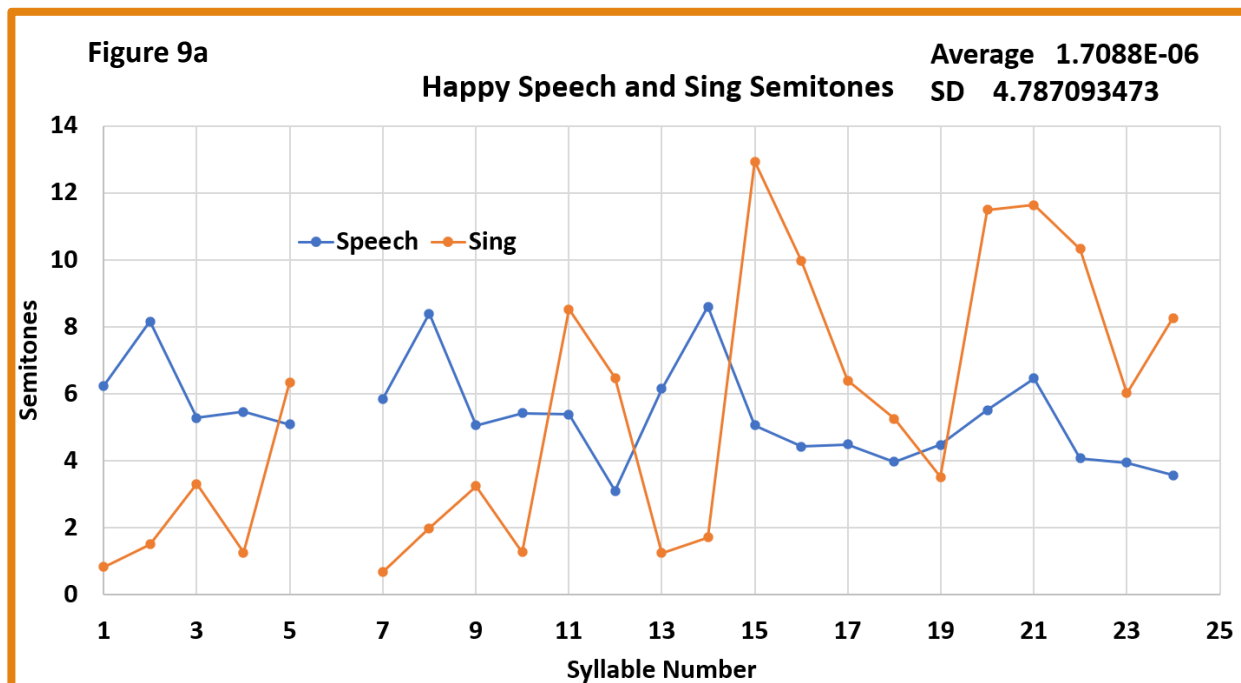


Figure 9a

- Singing and the speaking semitones overlap **rarely**, with the change in direction

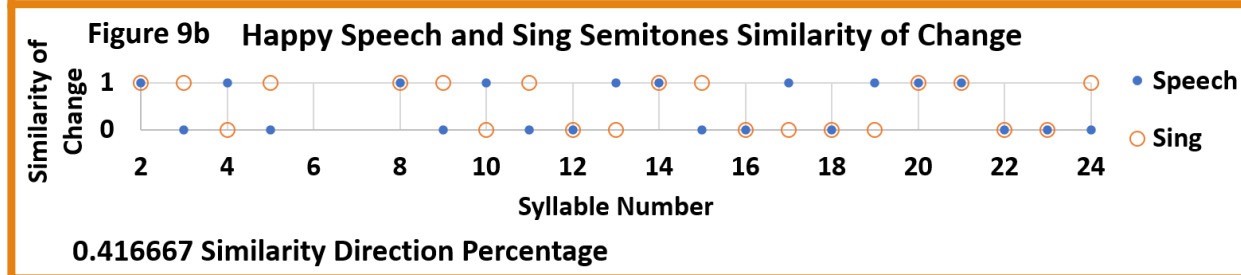


Figure 9b

- Similarity of Change Percentage = **41%**.
- Speaking and singing result in **dissimilar** syllable semitones.

Results: “Happy Birthday To You” Semitones

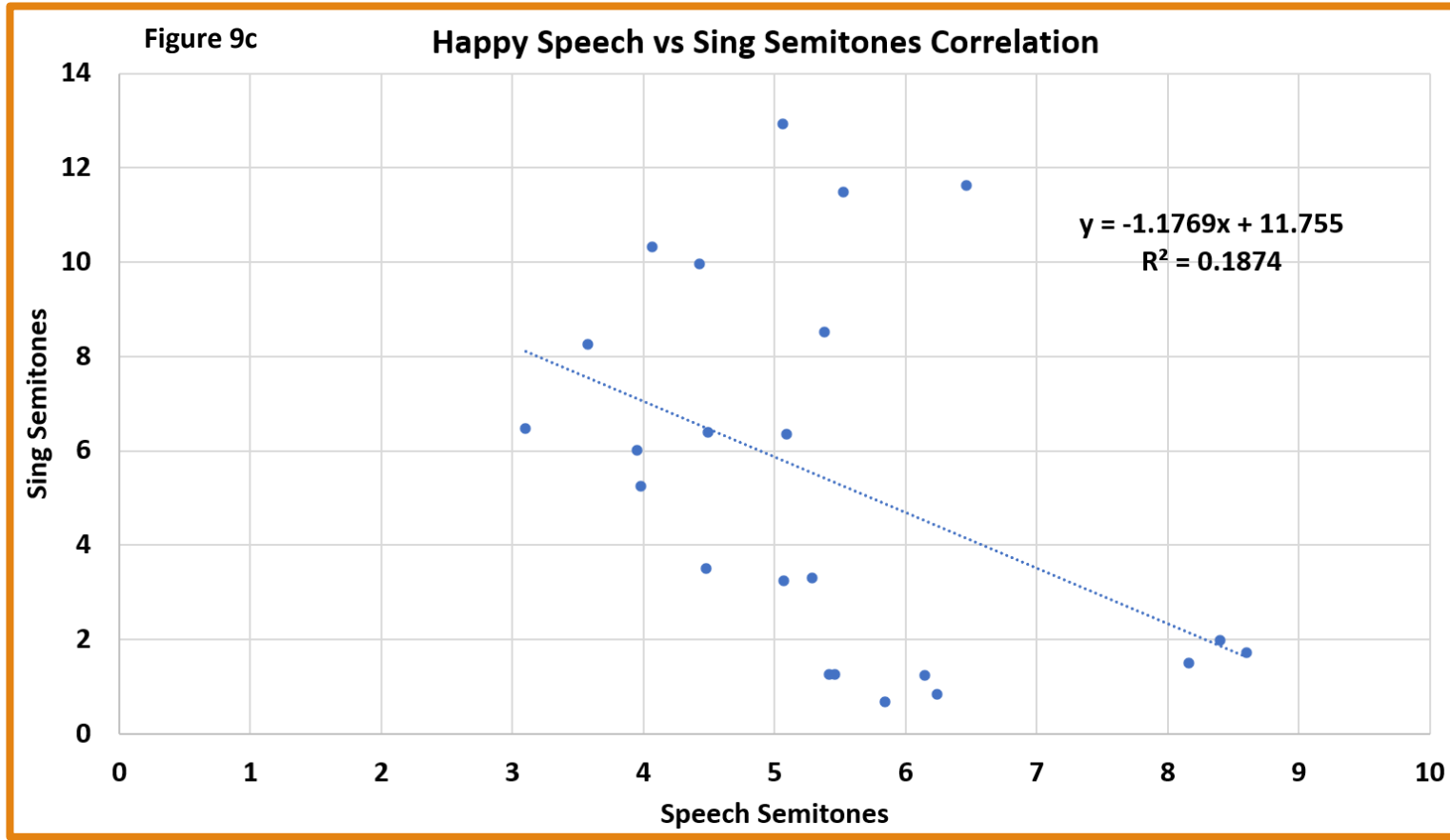


Figure 9c

- Shows the relationship between the semitones of the spoken vs sung syllables
- **Moderately low** correlation of $r = 0.43$ (the square root of the R^2 in the figure).

Discussion: Table

1a	Row, Row, Row Your Boat					
	SD	SCP	SPR	SPRD	SGR	SGRD
Duration (s)	0.103	92%	0.146 - 0.543	0.397	0.099 - 0.821	0.722
Intensity (dB)	3.04	65%	66.58 - 77.66	11.08	70.02 - 79.07	9.05
Semitones	3.34	69%	2.36 - 8.78	6.42	0.45 - 12.58	12.13

1b	Twinkle, Twinkle Little Star					
	SD	SCP	SPR	SPRD	SGR	SGRD
Duration (s)	0.103	85%	0.273 - 0.685	0.411	0.237 - 0.710	0.472
Intensity (dB)	2.95	56%	70.75 - 77.19	6.45	67.98 - 79.06	11.1
Semitones	3.18	49%	2.95 - 9.20	6.25	0.44 - 9.89	9.45

1c	Happy Birthday To You					
	SD	SCP	SPR	SPRD	SGR	SGRD
Duration (s)	0.12	83%	0.313 - 0.467	0.155	0.124 - 0.677	0.553
Intensity (dB)	3.64	54%	69.30 - 77.27	7.97	70.80 - 82.14	11.34
Semitones	4.79	42%	3.58 - 8.60	5.02	0.68 - 12.94	12.3

SD: Standard Deviation

SCP: Similarity of Change Percentage

SPR: Speech Range

SPRD: Speech Range Difference

SGR: Singing Range

SGRD: Singing Range Difference

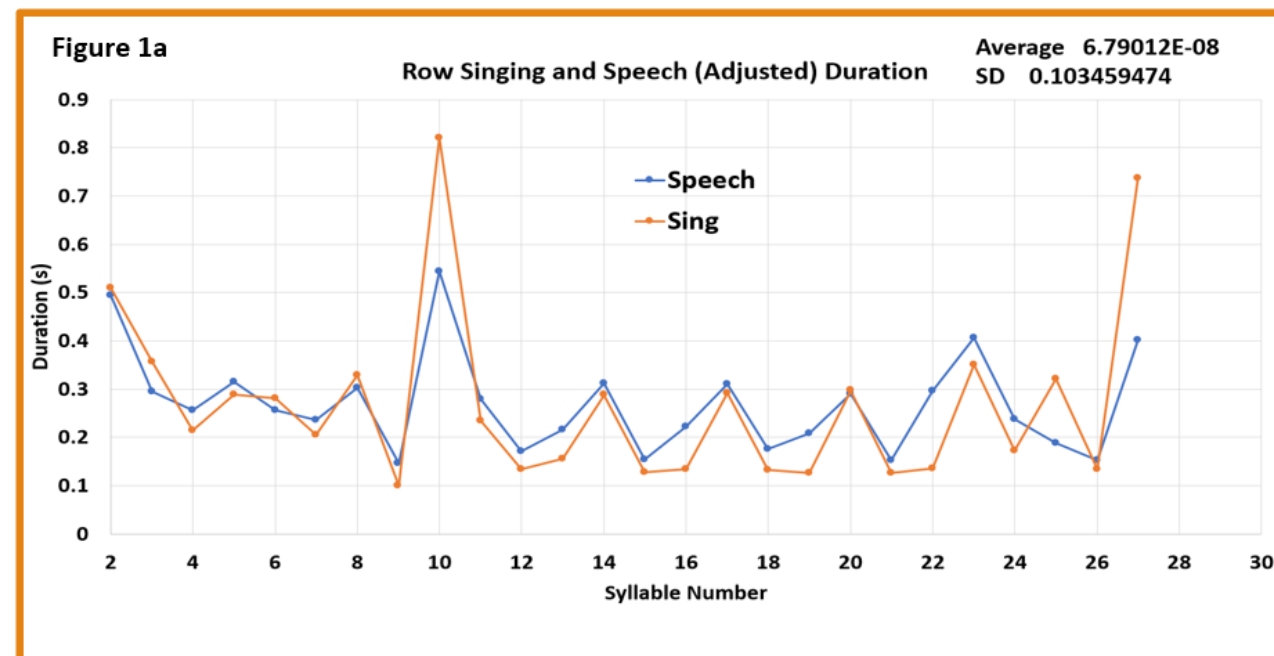
Discussion: “Row, Row, Row, Your Boat” Duration

Syllable duration was the closest value to speech in all three songs.

“Row, Row, Row Your Boat” had the highest similarity of change percentage score (SCP = 92%) for syllabic duration

- most variability occurs at syllables 10 “stream” and 27 “dream”
- “stream” and “dream” are notated with a half notes

Clinical Application: could be useful for exercises centered around elongating and exaggerating utterances



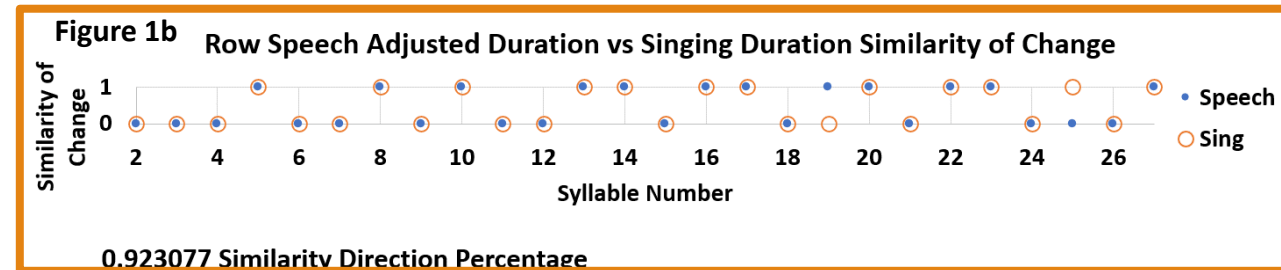
Discussion: “Twinkle, Twinkle Little Star” Duration

“Twinkle, Twinkle Little Star” (SCP = 85%) had second highest similarity of change percentage score.

- Smallest difference between the Speech and Sing range (SPRD = 0.411s vs SGRD= 0.472s).

- High durational similarity to speech

Clinical Application: May be useful for Clinicians working to normalize syllable duration.



1a	Row, Row, Row Your Boat					
	SD	SCP	SPR	SPRD	SGR	SGRD
Duration (s)	0.103	92%	0.146 - 0.543	0.397	0.099 - 0.821	0.722
Intensity (dB)	3.04	65%	66.58 - 77.66	11.08	70.02 - 79.07	9.05
Semitones	3.34	69%	2.36 - 8.78	6.42	0.45 - 12.58	12.13

Discussion: “Happy Birthday To You” Duration

“Happy Birthday To You” (SCP=83%) had the lowest change of percentage score.

higher variability than the other selections (SD= .120 vs SD= .103)

Durational range triple for sung vs spoken version (SPRD = 0.155 vs SGRD = 0.553)

Clinical Application: extend the duration of the production of syllables.

	Happy Birthday To You					
	SD	SCP	SPR	SPRD	SGR	SGRD
Duration (s)	0.12	83%	0.313 - 0.467	0.155	0.124 - 0.677	0.553
Intensity (dB)	3.64	54%	69.30 - 77.27	7.97	70.80 - 82.14	11.34
Semitones	4.79	42%	3.58 - 8.60	5.02	0.68 -12.94	12.3

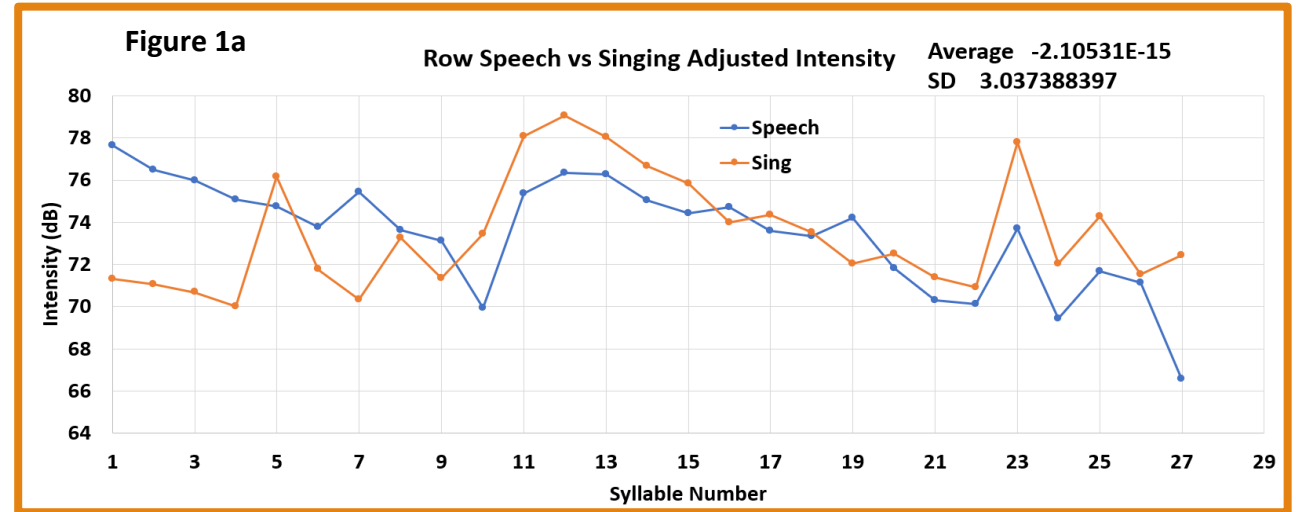
Discussion: “Row, Row, Row Your Boat” Intensity

Intensity had the second most similarity of change between speech and singing; however, the comparison is moderate

Out of the selections, “**Row, Row, Row Your Boat**” had the highest similarity percentage for intensity (SCP=65%)

Most overlap between the Singing and Speech groups occurred from syllables 11 to 22

- “merrily, merrily, merrily, merrily”
- Triplet section



Discussion: “Row, Row, Row Your Boat” Intensity

The Speech intensity range (SPRD = 11.08 dB) for “**Row, Row, Row Your Boat**”

within 2 dB of singing intensity range (SGRD = 9.1 dB)

- This may be due to the final syllable in Row, syllable 27
- “end of phrase prosody”

Contrary the finding for the other two songs

- 72% higher for “Twinkle, Twinkle Little Star” (6.45 dB vs 11.1 d)
- 42% higher for “Happy Birthday to You” (7.97 dB vs 11.34 dB)

	Row, Row, Row Your Boat					
	SD	SCP	SPR	SPRD	SGR	SGRD
Duration (s)	0.103	92%	0.146 - 0.543	0.397	0.099 - 0.821	0.722
Intensity (dB)	3.04	65%	66.58 - 77.66	11.08	70.02 - 79.07	9.05
Semitones	3.34	69%	2.36 - 8.78	6.42	0.45 - 12.58	12.13

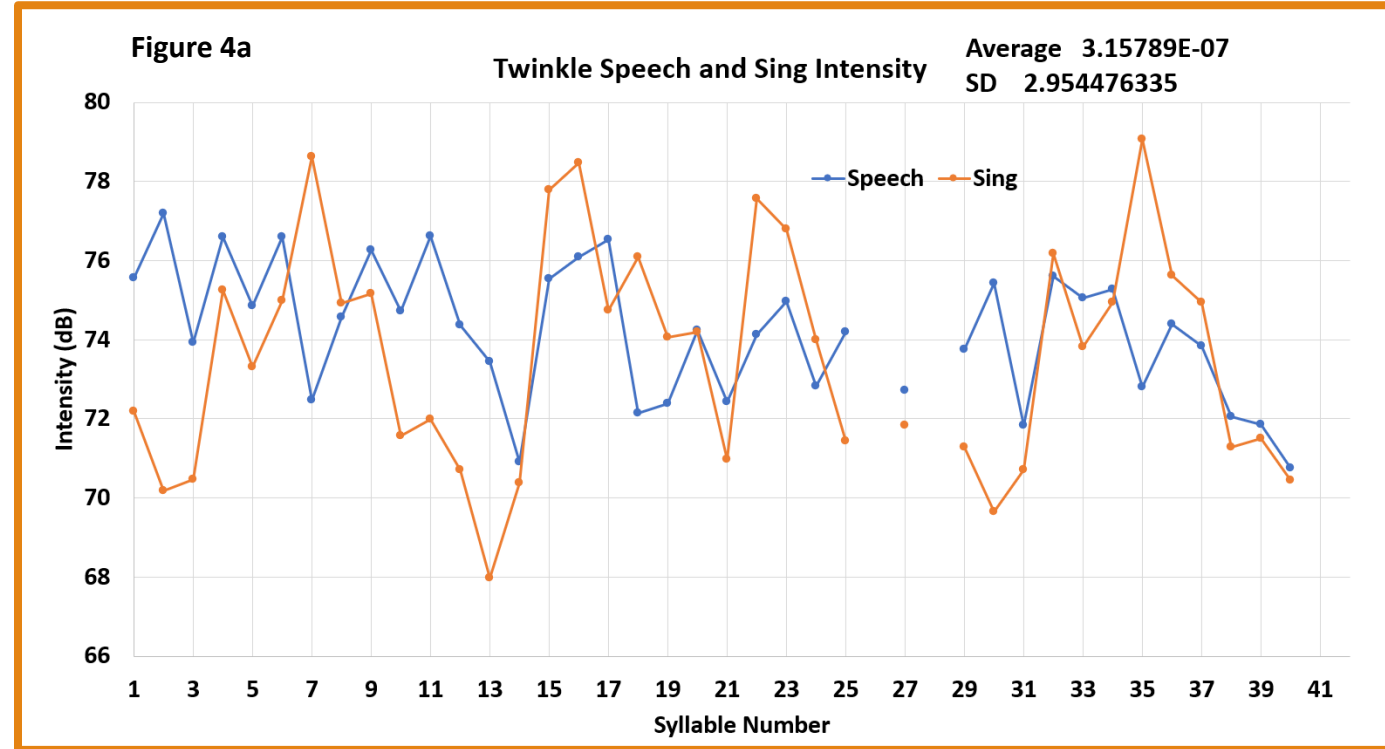
Clinical Application: may be a good for working on intensity in a more limited range

Discussion: “Twinkle, Twinkle Little Star” Intensity

“Twinkle, Twinkle Little Star” (SCP= 56%) had second highest similarity of change score.

Less overlap with intensity compared to “Row”

Clinical Application: could serve as a transitional step for clinicians wanting to widen range of intensity



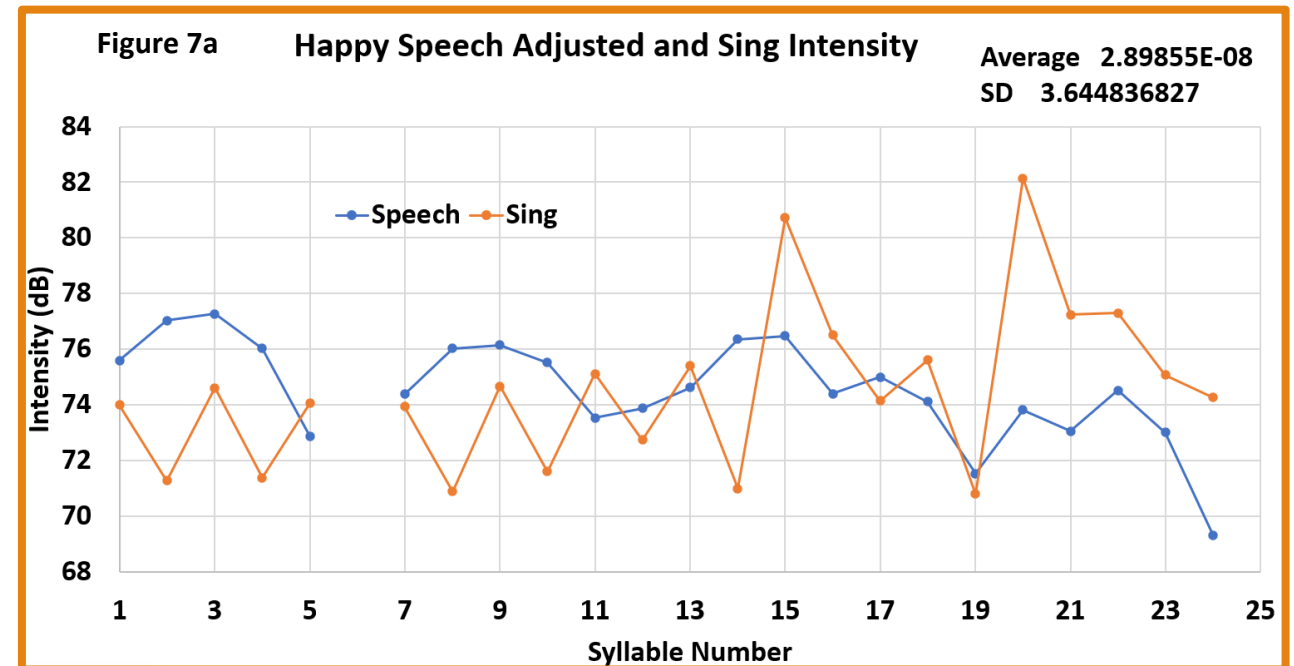
Discussion: “Happy Birthday To You” Intensity

“Happy Birthday To You” (SCP= 54%) had the lowest similarity of change percentage

- Selection starts with a limited dynamic range and then builds to much larger variations

“Happy Birthday To You” was the loudest song (SPR = 69.30 – 77.27; SGR = 70.80 – 82.14)

Clinical Application: could be helpful for clinicians trying to increase intensity at which clients speak.



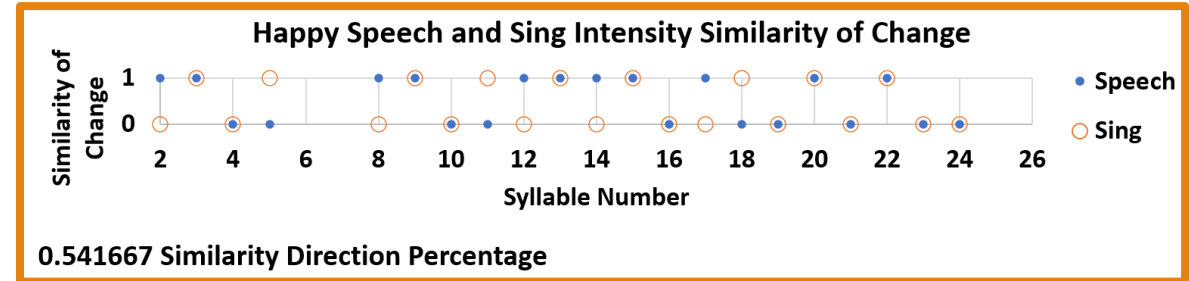
Discussion: “Happy Birthday To You” Intensity

Low SCP score (54%) may be attributed to the greater variability in beginning and middle and middle sections of the song

End section: change in syllable intensity was very similar

may indicate that the composition end of prosody phrasing

Clinical Application: good candidate if a clinician has a facilitatory goal for normalizing the prosody of end phrases.



Discussion: “Row, Row, Row Your Boat” Semitones

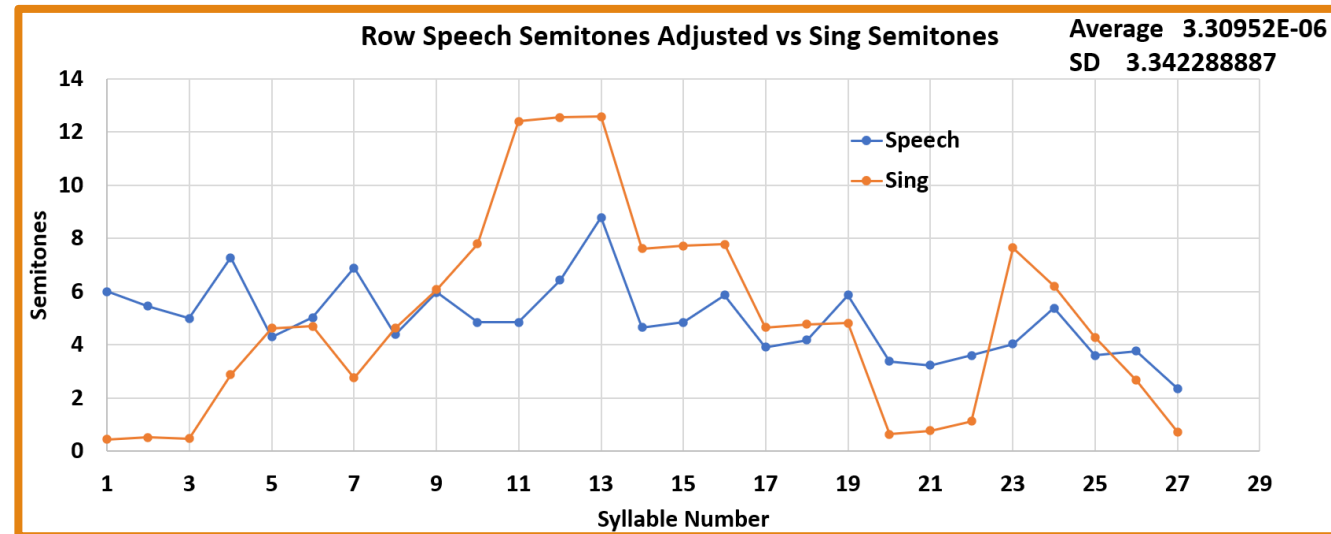
Semitones had the least similarity between speech and singing

“Row, Row, Row Your Boat” had the highest similarity percentage for semitones (SCP=69%)

Descent in middle and final section can be observed in both the Speech and Singing group

Semitones range for singing is nearly double that of speaking (SGRD = 6.42 vs SPRD =12.13)

Clinical Applications: Useful for increasing singing range beyond that of normal speech



Discussion: “Twinkle, Twinkle Little Star” Semitones

“**Twinkle, Twinkle Little Star**” (SCP:49%) has second highest similarity of percentage score

Out of the three selections its range is closest to speech (SPR = 2.95 – 9.20; SGR = 0.45 – 12.58)

	Twinkle, Twinkle Little Star					
	SD	SCP	SPR	SPRD	SGR	SGRD
Duration (s)	0.103	85%	0.273 - 0.685	0.411	0.237 - 0.710	0.472
Intensity (dB)	2.95	56%	70.75 - 77.19	6.45	67.98 - 79.06	11.1
Semitones	3.18	49%	2.95 - 9.20	6.25	0.44 - 9.89	9.45

Clinical Application: may be useful for working with clients who have a limited vocal range or difficulty with variable melodies.

Discussion: “Happy Birthday To You” Semitones

“Happy Birthday To You” (SCP:42%) had the lowest similarity percentage for semitones.

No apparent overlap between the singing and speech group’s trials

- was not created to emulate the intonational prosody of speech.

Greatest difference between SPRD and SGRD scores (7.28 semitones)

- twice as large as the difference between Twinkle, Twinkle Little Star (3.2 semitones)

	Happy Birthday To You					
	SD	SCP	SPR	SPRD	SGR	SGRD
Duration (s)	0.12	83%	0.313 - 0.467	0.155	0.124 - 0.677	0.553
Intensity (dB)	3.64	54%	69.30 - 77.27	7.97	70.80 - 82.14	11.34
Semitones	4.79	42%	3.58 - 8.60	5.02	0.68 - 12.94	12.3

Clinical Applications: may be best utilized in achieving treatment goals with the outcomes of increasing a client’s range of speech

Most and Least Speech Like

Most Speech Like: Out of the three selections, “Row, Row, Row Your Boat” was the most similar on all measures of duration (SCP = 92%), intensity (SCP = 65%), and intonation (SCP = 69%).

Least Speech Like: Out of the three selections, “Happy Birthday To You” was the least similar on all measures of duration (SCP = 83%), intensity, (SCP = 54%), and intonation (SCP = 42%).

Limitations

Prevalence of vocal fry in the speaking group

- (all fry productions were eliminated from measurements)

Small selection size of songs

Possible interference in the Speech group due to familiarity

- difficult for them to not sing the words while reading them aloud

Future Studies:

In the future research could expand this study by:

- Analyzing other songs than the ones studied
- Analyze more complex music
- Utilize songs that are unfamiliar to the subjects in the Speech Group
- Examine the prosodic difference between speech and singing of people with PD or aphasia

Conclusion:

This study helps to establish a basis for examining how simple songs can be utilized as a tool in speech therapy.

1. “Row, Row, Row Your Boat” most similar relative to change for all three measures
2. All three songs studied had a wider pitch range when sung than when spoken
3. The average percentage of “similarity of change” across the three measures was:
 - “Row, Row, Row Your Boat”, 75.3%,
 - “Twinkle, Twinkle Little Star”, 63.3%,
 - “Happy Birthday to You”, 59.7%.

Conclusion: Therapeutic Goals

4. Goal: greater syllable durations than the spoken lyrics

“Row, Row, Row Your Boat” and “Happy Birthday to You”

5. Goal: similar intensity productions between sung and spoken

“Row, Row, Row Your Boat”

6. Goal: quite different intensity productions between sung and spoken

“Twinkle, Twinkle Little Star” or “Happy Birthday to You”

Conclusion: Continued

7. The song “Row, Row, Row Your Boat” was the most speech like out of the three selections.
8. The song “Happy Birthday To You” was the least speech like out of the three selections.

References:

Kershenbaum, A., Nicholas, M. L., Hunsaker, E., & Zipse, L. (2017). Speak along without the song: What promotes fluency in people with aphasia? *Aphasiology*, 33(4), 405–428.

<https://doi.org/10.1080/02687038.2017.1413487>

Lim, K.-B., Kim, Y.-K., Lee, H.-J., Yoo, J., Hwang, J. Y., Kim, J.-A., & Kim, S.-K. (2013). The therapeutic effect of neurologic music therapy and speech language therapy in post-stroke aphasic patients. *Annals of Rehabilitation Medicine*, 37(4), 556–562. <https://doi-org.ezproxy.bgsu.edu/10.5535/arm.2013.37.4.556>

Merrett, D. L., Zumbansen, A., & Peretz, I. (2019). A theoretical and clinical account of music and aphasia. *Aphasiology*, 33(4), 379–381.

<https://doi.org/10.1080/02687038.2018.1546468>

Monroe, P., Halaki, M., Kumfor, F., & Ballard, K. J. (2020). The effects of choral singing on communication impairments in acquired Brain Injury: A systematic review. *International Journal of Language & Communication Disorders*, 55(3), 303–319. <https://doi.org/10.1111/1460-6984.12527>

References:

Norton, A., Zipse, L., Marchina, S., & Schlaug, G. (2009). Melodic intonation therapy: shared insights on how it is done and why it might help. *Annals of the New York Academy of Sciences*, 1169, 431–436.

<https://doi.org/10.1111/j.1749-6632.2009.04859.x>

Stegemöller, E. L., Hurt, T. R., O'Connor, M. C., Camp, R. D., Green, C. W., Pattee, J. C., & Williams, E. K. (2017). Experiences of persons with Parkinson's disease engaged in group therapeutic singing. *Journal of Music Therapy*, 54(4), 405–431. <https://doi.org/10.1093/jmt/thx012>

Tamplin, J., Morris, M. E., Marigliani, C., Baker, F. A., & Vogel, A. P. (2019). ParkinSong: A controlled trial of singing-based therapy for Parkinson's disease. *Neurorehabilitation and Neural Repair*, 33(6), 453–463. <https://doi.org/10.1177/1545968319847948>

Zumbansen, A., Peretz, I., & Hébert, S. (2014). Melodic intonation therapy: back to basics for future research. *Frontiers in neurology*, 5, 7. <https://doi.org/10.3389/fneur.2014.00007>

<https://www.swtimes.com.au/news/south-western-times/parkinson-a-joyous-form-of-music-therapy-ng-b881009723z>