Song and speech prosody influences VOT in stuttering and non-stuttering adolescents

Simone Falk[1, 2], *Elena Maslow*[3]

[1] Laboratoire Parole et Langage, CNRS UMR 7309, Aix-en-Provence, France
[2] Institut für dt. Philologie, LMU München, Germany
[3] Institut für Phonetik und Sprachverarbeitung, LMU München, Germany

simone.falk@lpl-aix.fr, elena.maslow@campus.lmu.de

Abstract

Since a long time, it is known that singing helps persons who stutter to produce their utterances more fluently. The prosodic characteristics of spoken and sung utterances differ considerably in their rhythmic and tonal structure. Therefore, it has been proposed that song prosody helps stutterers to improve their rhythmic planning of verbal material [1]. In order to investigate this idea, we examined temporal aspects, namely Voice Onset Time (henceforth, VOT) of voiceless plosives, in sung and spoken utterances of young German stutterers and non-stuttering controls. VOT tends to be reduced in song compared to speech. We expected a more important reduction in the stuttering group as voice onset timing should be facilitated in song compared to speech. Eight stuttering adolescents and eight normal fluent peers read and sang an altered version of “Happy Birthday” with test words containing the three voiceless stops /p/, /t/, /k/. Results showed that stuttering as well as non-stuttering adolescents reduced VOT during singing compared to speech. In contrast, only adolescents who stutter were less variable in their VOT production in song compared to speech. Additional analyses indicated further group differences in vowel duration following the stop consonant. These findings suggest that young stutterers benefit from sung prosody in their timing abilities.

Index Terms: Stuttering, Song and Speech Prosody, Voice Onset Time

1. Introduction

When stutterers sing, their disfluencies can reduce in a substantial way [2, 3, 4]. The reasons for this intriguing phenomenon are still unclear. In the 70ies, Wingate (1969) proposed that stuttering occurs in part because of a rhythmic deficit that can be attenuated by altering prosodic variables of the vocalization [1]. Both parts of this idea, the rhythmic deficit hypothesis [5] and the altered vocalization hypothesis [6] have been pursued in subsequent research. With respect to altered vocalization, it has been found that prolonged phases of phonation are good predictors of fluency in stutterers [7, 8]. In singing, phonation is fostered by prolongation of vocalic and sonorant portions of the speech signal [9]. However, recent research suggests that increased phonation time is not sufficient to explain all fluency-evoking conditions (e.g., paced speech [10]). This raises the question if rhythmic processes could play a more important role. In song, temporal intervals between syllables become more regular and hence, more predictable due to musical beat structure [11]. This could help stutterers by timing segmental material at the right moment during their production [12] and thereby enhance their fluency [1].

In this study, we aim to investigate temporal processes during singing in stuttering children and adolescents. Voice Onset Time is an indicator of fine-grained motor and timing control of laryngeal processes [13]. Stuttering (young) adults tend to show longer and more variable VOTs than non-stuttering adults, even in fluent speech [14, 15]. For children, the situation is less clear. In children of 4 and 9 years on average, no significant differences in VOT were found compared to age-matched peers, but their productions were overall more variable [16, 17].

In this paper, we examine how VOT of voiceless plosives changes in stuttering adolescents during singing compared to speaking, a question that was not addressed so far. An age-group was chosen that was intermediate between the groups tested in previous research (young stuttering children, adults). In general, VOTs are compressed in singing compared to spoken speech [18]. If singing enhances temporal planning and control, stuttering adolescents should benefit from this in their VOT productions. VOTs in spoken speech should be longer and more variable than in age-matched controls. On the other hand, VOTs in singing shouldn't differ from controls as far as timing processes are facilitated by song. Therefore, we hypothesize that VOTs will reduce more in the stuttering group than in controls when comparing spoken vs. sung productions. In addition, several other acoustic parameters such as stop closure, vowel duration and pitch were examined in order to better understand the influence of prosodic differences between sung and spoken utterances in both groups.

2. Method

2.1. Participants

Eight stuttering children aged from 11 to 15 years (*M* = 12.4, *SD* = 1.9, 6 males) participated in the study. Their stuttering symptoms ranged from mild to severe on the SSI-3 scale [19]. The participants were tested while attending the therapy course “SAS - Stärker als Stottern” in Starnberg near Munich. The control group consisted of 8 age-matched children (*M* = 13.0, *SD* = 2.3, 6 males). The control group was recorded at home or at the Institute of Phonetics and Speech processing, LMU Munich. All participants were native speakers of German and were untrained singers (no choral activity, no singing lessons).

2.2. Procedure

The experiment had two parts. First, participants were asked to read repeatedly a simple text consisting of three sentences which was based on the lyrics of the song "Happy Birthday". Participants were instructed to read the sentences at a

10.21437/SpeechProsody.2014-156

comfortable moderate reading pace. At this time, the participants did not know that the lyrics were derived from the “Happy Birthday” song. In the second part of the experiment, participants sang the text to the melody of "Happy Birthday". Again, they were asked to sing at a moderate tempo that was comfortable to them. The material was presented on handouts and participants had the opportunity to practice their reading and singing with one repetition or stanza before recording. All participants declared to be familiar with the song.

Each participant was recorded separately in a quiet room with the experimenter present who controlled the recording. The participant was seated at a table and had the visual handouts before him / her on the desk. The sung and spoken texts were repeated nine times with different test words per repetition (see Material). Data collection was done with a Beyerdynamic headset (TG H54c) and a H-4N Zoom recorder. The data was recorded at 44100Hz/24bit.

2.3. Material

The three German voiceless stops /p, t, k/ were inserted in bisyllabic nonsense words. Each test word had a trochaic strong-weak accent pattern. The critical stop was in the onset of the first, stressed syllable followed by a long vowel. The test words were /'pi:ta/, /'pa:ta/, /'pu:ta/, /'ti:ta/, /'ta:ta/, /'tu:ta/, /'ki:ta/, /'ka:ta/, /'ku:ta/. These words were inserted in the German version of "Happy Birthday". In the German version, the name of the birthday child is used in every sentence ("Dear X, all the best to you"). The test words were inserted five times in each repetition at the place where the name occurs as demonstrated for /'ki:ta/ in (1).

(1) *Liebe* /'ki:ta/, *viel Glück, liebe* /'ki:ta/, *viel Glück, liebe* /'ki:ta/, *liebe* /'ki:ta/, *liebe* /'ki:ta/, *viel Glück*.

"Dear /'ki:ta/, all the best to you, etc..."

Overall, 15 read / spoken samples were collected for each stop per participant. This resulted in overall 45 × 2 items per participant for analysis.

2.4. Analyses

VOT was measured manually for each plosive by inspecting the oscillogram and spectrogram of the audio signal in Praat [20]. VOT was annotated for each plosive [13]. The time interval was marked between the onset of the stop release and the beginning of vocal fold vibration of the following vowel onset. The second zero crossing of the glottal pulse was consistently used to mark the beginning of voicing. Furthermore, standard deviations of VOT were calculated for each participant in order to assess variability in sung and spoken VOT productions. Some additional measures were taken. First, closure duration was measured from the point where vocal fold vibration of the preceding Schwa-vowel stopped until the onset of the stop release. Second, vowel duration following the stops was extracted from the recordings. The left vowel boundary was determined by marking the beginning of periodic glottal pulses in the oscillogram (second zero crossing). The right boundary was assessed by inspecting the formant patterns and the subsequent spectrum of the consonantal context. Pitch targets were used for the evaluation of vocalic pitch. Maximum pitch per vowel was analyzed using the pitch analysis algorithm in Praat (range: 75-600 Hz) [20]. Furthermore, overall pitch characteristics (median pitch and range) as well as the total duration of each utterance (i.e., one stanza or repetition of the text / song) were assessed. Pauses (e.g., respirations, hesitations, stuttering) were subtracted from the total utterance duration.

3. Results

Test words containing disfluencies or errors of any kind (such as misreading, stuttering) were discarded from analysis. Overall, stuttering symptoms were not very frequent in the test group, and occurred equally often in speech and song (i.e., in 1.6 % of all words in the material). Mean duration and variability for VOT values per stop consonant as well as closure and vowel duration (all vowels confounded) were calculated. Before averaging, outliers were excluded (20 values of closure duration in both groups). These values as well as overall duration and pitch of the utterance and vocalic pitch are displayed for for both groups in Table 1 and Fig. 1.

	Adolescents who stutter		Control group	
Measures	Reading	Singing	Reading	Singing
Utterance				
Mean overall duration (s)	8.25 (1.56)	8.88 (1.63)	7.05 (0.88)	8.94 (0.88)
Median pitch (Hz)	198 (51)	213 (48)	200 (49)	232 (61)
Mean pitch range (st)	11.2 (2.8)	12.2 (3.4)	12.6 (5.9)	13.0 (2.9)
Stop and subsequent vowel				
Mean closure duration (s)	0.093 (0.016)	0.076 (0.011)	0.080 (0.013)	0.067 (0.013)
Mean vowel duration (s)	0.099 (0.033)	0.174 (0.042)	0.115 (0.022)	0.224 (0.039)
Maximum vocalic pitch (Hz)	222 (55)	241 (61)	216 (52)	255 (85)

Table 1. *Overall mean values of prosodic utterance, stop and vowel measures (in seconds (s), Hertz (Hz) and semitones (st)) of the sung and read performance in both groups of participants. Standard deviations are displayed in brackets.*

As can be seen in Fig. 1, sung stops displayed smaller VOT values than read stops in both groups. The VOT data were entered in a three-way mixed-design Analysis of Variance (ANOVA) with the within-subject factors Stop (/p/, /t/, /k/) and Vocalization (reading vs. singing) and the between-subject factor Group (test vs. control). Results showed main effects for Vocalization (F (1, 14) = 62.13, $p < .001$) and Stop (F (2, 28) = 34.55, $p < .001$). No interactions or group differences were found. Additional pairwise comparisons confirmed that VOT differed significantly in all three stop classes (with /k/ > /t/ > /p/). This is in line with previous research reporting differences between stop categories [13].

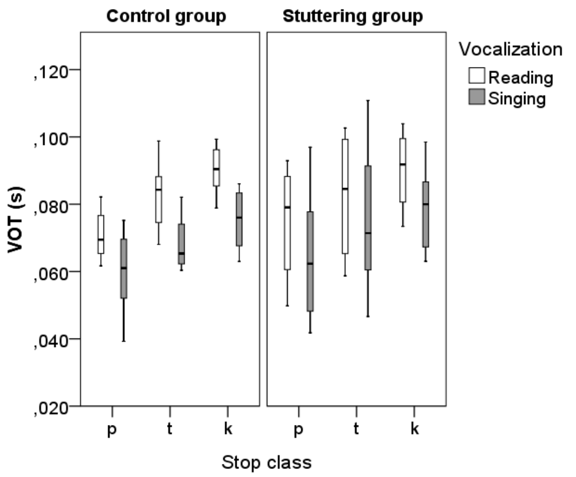

Figure 1. *Boxplots of VOT values of three German voiceless stops in sung and read speech displayed for adolescents who stutter and for the control group.*

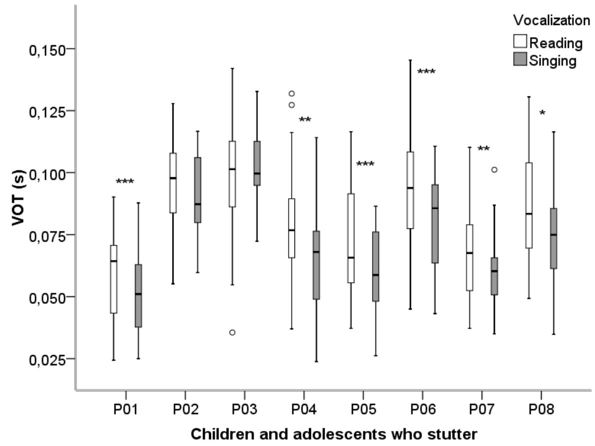

Figure 3. *Individual VOT reduction patterns in read and sung speech in participants who stutter. Stars indicate significance levels (*** $p < .001$, ** $p < .01$, * $p < .05$).*

Another ANOVA with the same factors reported above was performed for VOT variability (the standard deviation of the mean per participant, see 2.4). Main effects were found for Vocalization ($F\,(1,\,14) = 19.46$, $p < .005$) and Group ($F\,(1,\,14) = 6.68$, $p < .05$) as well as a significant interaction between both factors ($F\,(1,\,14) = 10.77$, $p < .01$). The interaction is displayed in Fig. 2. Two-sided t-tests (Bonferroni-corrected) confirmed that adolescents who stutter significantly reduced VOT variability from speech to song ($t\,(7) = 6.12$, $p < .001$). This was not the case for the control group.

In order to examine individual VOT patterns, especially in the test group (see Fig. 3), statistical analyses were performed on VOTs for each participant separately (two-sided paired-samples t-tests on VOTs in sung vs. read speech averaged over all three stop classes).

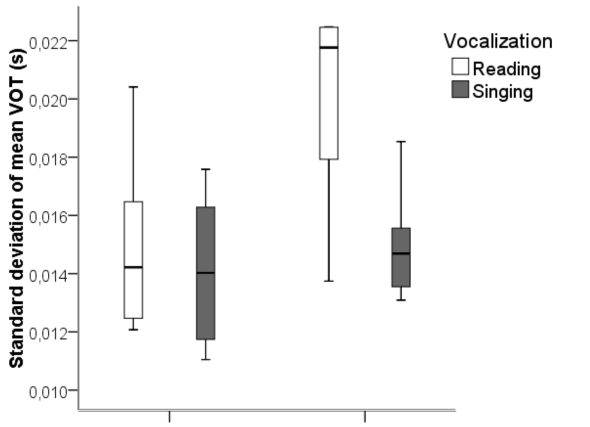

Figure 2. *Variability of VOT in sung and read speech displayed for adolescents who stutter and for the control group.*

Results revealed that VOT reduction was significant only in six out of eight stuttering children. In the control group, all children showed a significant reduction of VOT from speech to song (with at least $p < .01$). In order to examine further temporal differences between speech and song and their realizations by the test and control group, additional statistical analyses were performed on stop closure, vowel and utterance duration and their variability (see Table 1). The same ANOVA design as for the VOT analyses was used except that for vowel and utterance duration, no third factor (Stop) was needed.

For closure duration, a similar reduction pattern from speech to song was found as for VOT ($F\,(1,\,14) = 32.58$, $p < .001$). A main effect was also found for Stop class ($F\,(2,\,28) = 47.72$, $p < .001$). Groups did not differ, there was only a marginal trend for the test group showing longer closure than the control group ($p = .098$). Closure variability was also reduced from speech to song ($F\,(1,\,14) = 7.63$, $p < .05$) with again a marginal trend for the test group to be more variable than controls ($p = .088$). In contrast to VOT and closure duration, vowels were lengthened in song vs. speech ($F\,(1,\,14) = 87.15$, $p < .001$). Adolescents who stutter displayed overall shorter vowel duration than the control group ($F\,(1,\,14) = 4.89$, $p < .05$). Moreover, vowel variability was greater in song than in speech in both groups ($F\,(1,\,14) = 6.31$, $p < .05$). Finally, total utterance duration was longer in song than in speech ($F\,(1,\,14) = 14.86$, $p < .005$) and less variable in song than in speech ($F\,(1,\,14) = 7.66$, $p < .05$). Adolescents who stutter merely showed a trend to have longer ($p = .075$) and more variable utterance duration ($p = .063$).

As we did not control for tempo during the recordings between speech and song, we evaluated to what extent VOT or closure duration and reduction and vowel duration and lengthening was related to utterance duration differences. First, we correlated utterance duration with mean VOT and closure duration as well as mean vowel duration per utterance for sung and spoken speech separately. VOT as well as vowel duration, but not closure duration, were positively correlated with utterance duration in both speech (VOT: $r\,(142) = 0.23$, $p < .01$; vowel: $r\,(142) = 0.29$, $p < .001$) and song (VOT: $r\,(142) = 0.25$, $p < .01$; vowel: $r\,(142) = 0.38$, $p <$

.001). In other words, longer utterance duration was related to longer VOT and longer vowel duration in both reading and singing.

Second, we assessed if differences in utterance duration were likely to influence VOT reduction. As children and adolescents who stutter tended to show longer utterance duration, but still similar VOT values as the control group, it might be the case that we underestimated group differences without a temporal normalization. In order to answer this question, we calculated difference scores between speech and song by subtracting utterance duration in speech from utterance duration in song for each trial. The same calculation was performed for VOT reduction which was estimated by subtracting sung mean VOTs per trial from read VOTs. A linear regression with the predictor variable utterance difference and the dependent variable VOT reduction was performed. No significant relation was found between utterance duration differences and VOT reduction. Hence, we conclude that the results from the VOT reduction data are independent of utterance duration differences.

4. Discussion

In the present study, we investigated the impact of sung prosody on temporal aspects of young stutterers' verbal productions. In a control group, the VOT of voiceless German plosives was found to reduce significantly in sung speech compared to read speech. This result extends prior research as we show the same result for children and adolescents as done for adult populations [18]. VOT reduction was also found in a group of eight stuttering children and adolescents. Importantly, VOT variability (defined as the standard deviation of the mean) was a good indicator of group differences. VOT variability was significantly reduced during singing compared to reading in adolescents who stutter whereas the control group did not show differences in both tasks.

These findings suggest that singing influenced timing aspects in the present group of young stutterers. First, we demonstrated that VOT reduction occurs as well in stuttering adolescents as in age-matched peers, but against our initial hypothesis, this reduction was not more important in the test group. VOT values were comparable to age-matched peers in sung and, in particular, in spoken productions. This is in line with previous research on VOT in the speech of stuttering children (mean age < 10) [16, 17], but differs from results on adults that showed longer VOT in spoken speech [14, 15]. It is likely that VOT differences only arise consistently at an adult age. Children and adolescents follow individual patterns of development and are in general more variable in their timing capacities [21]. For instance, individual differences were found in our group of adolescents who stutter. Two participants did not show a VOT reduction. Future studies should investigate if the degree of VOT reduction is a predictor of other variables as for example stuttering severity or the efficiency of altered vocalization in reducing disfluencies.

Although VOT duration per se did not allow to distinguish between test group and controls, VOT variability did. Adolescents who stutter were significantly less variable in their VOTs during singing than during reading. It has been shown in previous studies [16], that persons who stutter have more variable VOT in speech production compared to controls. Similar results have been found for perception as well [22]. Our finding shows for the first time that singing has the potential to reduce VOT variability in the productions of adolescents who stutter. This lends support to the hypothesis that singing helps stutterers to improve their timing abilities. Consequently, as suggested by Wingate [1], disfluencies would reduce because rhythmical planning is facilitated. In our sample, disfluencies were rarely occurring in either spoken or sung speech, probably due to the simple and repetitive nature of the German text of "Happy Birthday". A more difficult text/song could be used in future studies in order to observe fluency enhancement during singing in relation to VOT variability reduction.

Further variables were examined in order to better understand the temporal differences between speech and song in our participants. It is well-known that vowel lengthening is very prominent in singing [9] and this tendency was also confirmed by our data. In contrast, previous research on VOT [18] suggests that consonantal parts of the segmental stream undergo a durational compression in song unless consonantal lengthening is part of the phonological system of a language [23]. Our data confirmed compression for plosives by showing the same temporal reduction of closure duration in song vs. speech as found in VOT. While no group differences appeared for closure duration, vowel duration was shorter in participants who stutter. This was unexpected as stutterers tended to show longer utterances which should have led to longer vowel duration. Previous studies have found shorter vowel duration in stutterers' speech [24]. The shorter vowels in our spoken and sung material are probably reflections of this segmental characteristic of stuttering.

In sum, our findings show that singing positively impacts on timing abilities in adolescents who stutter. Future research should address the question to what extent rhythmic characteristics interact with tonal aspects in singing to reduce disfluencies. In a previous study, Glover et al. (1996) discussed the possibility that fluency-enhancement in singing could also arise due to a better representation of global melodic structures, normally not present in speech [6]. For instance, in the domain of neurological rehabilitation, Melodic Intonation Therapy [25] combines both melodic and rhythmic alterations of verbal productions in order to enhance fluency in aphasic patients. In fact, both, rhythm and melody have been shown to impact on production in this task [26, 27]. Future research using neuro-imaging and electrophysiological methods could further unravel the role of melody and rhythm for fluent sung and spoken productions in persons who stutter [e.g., 10, 28].

5. Acknowledgements

The research leading to these results has received funding from a SSHRC-MCRI AIRS (www.airsplace.ca) research grant as well as from the European Union Seventh Framework Program [FP7/2007-2013; FP7-PEOPLE-2012-IEF] under grant agreement n° 327586 to Simone Falk. We thank Thilo Müller and the team of the SAS course for their help with recording the stuttering participants as well as Jonathan Harrington, Phil Hoole, Ulrich Reubold and Florian Schiel of the IPS Munich, Simone Dalla Bella, Euromov, Montpellier, and two anonymous reviewers for their advice and helpful comments.

6. References

[1] Wingate, M.E. (1969). Sound and pattern in “artificial” fluency. *Journal of Speech and Hearing Research*, *12*, 677-686.

[2] Andrews, G., Howie, P.M., Dozsa, M., & Guitar, B.E. (1982). Stuttering: speech pattern characteristics under fluency-inducing conditions. *Journal of Speech, Language, and Hearing Research, 25*, 208-216.

[3] Healey, E.C., Mallard, A.R., & Adams, M.R. (1976). Factors contributing to the reduction of stuttering during singing. *Journal of Speech and Hearing Research, 19*(3), 475-480.

[4] Johnson, W., & Rosen, L. (1937). Studies in the psychology of stuttering: VII. Effect of certain changes in speech pattern upon frequency of stuttering. *Journal of Speech disorders*, *2*, 105-109.

[5] Olander, L., Smith, N., & Zelaznik, H.N. (2010). Evidence that a motor timing deficit is a factor in the development of stuttering. *Journal of Speech, Language, and Hearing Research*, *53*, 867-886.

[6] Glover, H., Kalinowski, J., Rastatter, M., & Stuart, A. (1996). Effect of instruction to sing on stuttering frequency at normal and fast rates. *Perception and Motor Skills*, *83*(2), 511-522.

[7] Colcord R.D., & Adams, M.R. (1979). Voicing duration and vocal SPL changes associated with stuttering reduction during singing. *Journal of Speech and Hearing Research, 22*(3), 468-479.

[8] Davidow, J.H., Bothe, A.K., Andreatta, R.D., & Ye, J. (2009). Measurement of phonated intervals during four fluency-inducing conditions. *Journal of Speech and Hearing Research, 52*(1), 188-205.

[9] Eckardt, F. (1999). *Sprechen und Singen im Vergleich artikulatorischer Bewegungen* [Comparing the articulatory movements in speaking and singing]. Darmstadt: Thiasos Musikverlag.

[10] Stager, S.V., Jeffries, K.J., & Braun, A. R. (2003). Common features of fluency-evoking conditions studied in stuttering subjects and controls: an H215O PET study. *Journal of Fluency Disorders, 28*, 319-336.

[11] London, J. (2004). *Hearing in time*. New York, Oxford: Oxford University Press.

[12] Harrington, J. (1988). Stuttering, delayed auditory feedback, and linguistic rhythm. *Journal of Speech and Hearing Research, 31,* 36-47.

[13] Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical Measurements. *Word, 20*(3), 384-422.

[14] Hillman, R. E., & Gilbert, H. R. (1977). Voice onset time for voiceless stop consonants in the fluent reading of stutterers and nonstutterers. *Journal of the Acoustical Society of America, 61*(2), 610-611.

[15] Metz, D. E., Conture, E. G., & Caruso, A. (1979). Voice onset time, frication, and aspiration during stutterers’ fluent speech. *Journal of Speech and Hearing Research, 22*(3), 649-656.

[16] De Nil, L. F., & Brutten, G. J. (1991). Voice onset times of stuttering and nonstuttering children: The influence of externally and linguistically imposed time pressure. *Journal of Fluency Disorders, 16,* 143-158.

[17] Zebrowski, P. M., Conture, E. G., & Cudahy, E. A. (1985). Acoustic analysis of young stutterers' fluency: Preliminary observations. *Journal of Fluency Disorders, 10*, 173-192.

[18] McCrea, C. R. & Morris, R. J. (2007). Effects of vocal training and phonatory task on voice onset time. *Journal of Voice, 21*(1), 54-63.

[19] Riley, G. D. (1994). A stuttering severity instrument for children and adults (SSI-3). Austin: Pro Ed.

[20] Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5*(9/10), 341-345.

[21] Falk, S., Müller, T., & Dalla Bella, S. (in preparation). Poor sensorimotor timing in children and adolescents who stutter.

[22] Neef, N.E., Sommer, M., Neef, A., Paulus, W., von Gudenberg, A.W., Jung, K., & Wüstenberg, T. (2012). Reduced speech perceptual acuity for stop consonants in individuals who stutter. *Journal of Speech, Language and Hearing Research, 55*(1), 276-289.

[23] Falk, S. (2011). Temporal variability and stability in infant-directed sung speech: evidence for language-specific patterns. *Language and Speech, 54*(2), 167-180.

[24] Howell, P., & Vause, L. (1986). Acoustic analysis and perception of vowels in stuttered speech. *Journal of the Acoustical Society of America, 79*(5), 1571-1579.

[25] Albert, M.L., Sparks, R.W., & Helm, N.A. (1973). Melodic intonation therapy for aphasia. *Archives in Neurology*, *29*, 130-131.

[26] Schlaug, G., Norton, A., Marchina, S., Zipse, L., & Wan C.Y. (2010). From singing to speaking: facilitating recovery from nonfluent aphasia. *Future Neurology, 5*(5), 657-665.

[27] Stahl, B., Kotz, S. A., Henseler, I., Turner, R., & Geyer, S. (2011). Rhythm in disguise: Why singing may not hold the key to recovery from aphasia. *Brain*, *134*(10), 3083-3093.

[28] Toyomura, A., Fujii, T., & Kuriki, S. (2011). Effect of external auditory pacing on the neural activity of stuttering speakers. *Neuroimage*, *57*, 1507-1516.