

Perceived closeness and autistic traits modulate interpersonal vocal communication

Article

Published Version

Creative Commons: Attribution 3.0 (CC-BY)

Open Access

Sumathi, T. A., Spinola, O., Singh, N. C. and Chakrabarti, B.
ORCID: <https://orcid.org/0000-0002-6649-7895> (2020)
Perceived closeness and autistic traits modulate interpersonal
vocal communication. *Frontiers in Psychiatry*, 11. 50. ISSN
1664-0640 doi: 10.3389/fpsy.2020.00050 Available at
<https://centaur.reading.ac.uk/89315/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.3389/fpsy.2020.00050>

Publisher: Frontiers

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online



Perceived Closeness and Autistic Traits Modulate Interpersonal Vocal Communication

T. A. Sumathi^{1†}, Olivia Spinola^{2,3,4†}, Nandini Chatterjee Singh^{1*}
and Bhismadev Chakrabarti^{3,5,6*}

¹ National Brain Research Centre, Language, Literacy and Music Laboratory, Manesar, India, ² Department of Psychology, Università degli Studi di Milano Bicocca, Milan, Italy, ³ Centre for Autism, School of Psychology & Clinical Language Sciences, University of Reading, Reading, United Kingdom, ⁴ Department of Psychology, Sapienza University of Rome, Rome, Italy, ⁵ Inter University Centre for Biomedical Research, Mahatma Gandhi University, Kottayam, India, ⁶ India Autism Center, Kolkata, India

OPEN ACCESS

Edited by:

Frieder Michel Paulus,
Universität zu Lübeck, Germany

Reviewed by:

Carolyn McGettigan,
University of London, United Kingdom
Marian E. Williams,
University of Southern California,
United States

*Correspondence:

Nandini Chatterjee Singh
nandini@nbc.ac.in
Bhismadev Chakrabarti
b.chakrabarti@reading.ac.uk

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Social Cognition,
a section of the journal
Frontiers in Psychiatry

Received: 12 April 2019

Accepted: 21 January 2020

Published: 28 February 2020

Citation:

Sumathi TA, Spinola O, Singh NC and
Chakrabarti B (2020) Perceived
Closeness and Autistic Traits
Modulate Interpersonal
Vocal Communication.
Front. Psychiatry 11:50.
doi: 10.3389/fpsy.2020.00050

Vocal modulation is a critical component of interpersonal communication. It not only serves as a dynamic and flexible tool for self-expression and linguistic information but also plays a key role in social behavior. Variation in vocal modulation can be driven by individual traits of interlocutors as well as factors relating to the dyad, such as the perceived closeness between interlocutors. In this study we examine both of these sources of variation. At an individual level, we examine the impact of autistic traits, since lack of appropriate vocal modulation has often been associated with Autism Spectrum Disorders. At a dyadic level, we examine the role of perceived closeness between interlocutors on vocal modulation. The study was conducted in three separate samples from India, Italy, and the UK. Articulatory features were extracted from recorded conversations between a total of 85 same-sex pairs of participants, and the articulation space calculated. A larger articulation space corresponds to greater number of spectro-temporal modulations (articulatory variations) sampled by the speaker. Articulation space showed a positive association with interpersonal closeness and a weak negative association with autistic traits. This study thus provides novel insights into individual and dyadic variation that can influence interpersonal vocal communication.

Keywords: interpersonal closeness, dyad, vocal modulation, autism, social behavior

INTRODUCTION

The human voice is unique in its repertoire and functional utility. In addition to its role in expressing oneself through linguistic and non-linguistic routes, it acts as a crucial tool for social communication from early in development (1). Humans routinely and volitionally modulate their voice in social contexts like striking up a new friendship, arguing or making an emotionally charged

Abbreviations: CR, closeness rating; AQ, autism spectrum quotient; SMS, speech modulation spectrum; ASD, autism spectrum disorders; UK, United Kingdom; PCM, pulse-code modulation; NAQ, normalized amplitude quotient.

speech. A well-modulated voice carries considerable information about the message, speaker, language, and even on the emotions of the speaker (2–4). The dynamic nature of voice is vital for its social function, as vocal modulation can affect both the speaker and the addressee (5). Impairments in the recognition of socially communicative vocal modulation, such as emotional prosody, have been associated with impaired psychosocial functioning, as seen in autism or schizophrenia (6).

Several factors have been found to influence context-specific vocal modulation. In acoustic terms, voice modulation is defined as the manipulation of any non-verbal property of the voice including, but not limited to, pitch (F0) and formant frequencies. A recent account has distinguished two different types of vocal modulations. One of these types is involuntary, elicited automatically by environmental stimuli or different levels of endogenous arousal, and the other is a more controlled vocal modulation, which is goal-directed and less dependent on external stimuli, though not necessarily voluntary (1).

In its most basic form, the properties of the voice modulated speech signal are constrained by the physical and physiological properties of the speech production apparatus such as the thickness and characteristics of the vocal folds, variance in the shape and dimensions of a person's palate, and the dynamic use of the vocal tract (7). Thus, some acoustic variability can be attributed to the physical constraints and capabilities of the speaker, which vary with age, gender, and hormonal factors (8). Apart from this, speakers vary greatly in their articulatory habits and speaking style, both of which are functions of linguistic background, emotion-related state and trait measures of the speaker, acoustic environment, as well as the social context (7).

The social context-dependency of controlled vocal modulation remains relatively under researched. People communicate differently depending on their type of relationship, e.g., the way a person talks to friends is different from the way s/he talks to a stranger, or to a pet, or to a police officer (7). Consistent with this heuristic, Pisanski and others reported pitch modulation to be predictive of mate choice behavior (8). In another study by the same group, human males were shown to be able to volitionally exaggerate their body size by suitably modulating specific acoustic parameters of their voice (9). The role of social context and the relevance of relationship closeness in acoustic vocal modulation were studied by Katerenchuk and colleagues by extracting low-level acoustic features from a corpus of native English speech telephone conversations, which were then used to distinguish a conversation between friends or members of a family (10). Their results indicated that it is possible to distinguish friends from family on the basis of some low level lexical and acoustic signals. Similar results were reported from another study which examined the speech of a single Japanese speaker as she spoke to different conversational partners, finding significant differences in pitch (F0) and normalized amplitude quotient (NAQ) depending on the relationship between the speaker and her conversational partner (11). These preliminary studies provide important evidence on how voice modulation can be affected by context-specific factors such as dyadic relationship. However,

“friend pairs” or “family pairs” are somewhat arbitrary categories, within which the relationship between the interlocutors can vary greatly, depending on the individuals involved. To get around this variability, an alternative approach involves asking each member of a conversing pair to rate the perceived closeness toward the other member of the pair. This dimensional approach to evaluate the dyadic relationship is akin to the widely-used metric of “social distance” (12). Social distance in this sense, relates more to subjective closeness (perception of relationship quality) rather than to degree and variety of social interaction (13).

Modulation of vocal communication can also be influenced by individual-level factors, and not just on those specific to the dyad. Autism-related traits might constitute one such dimension of individual variability. Anecdotal reports have suggested an “autistic monotone” in describing an absence of context-appropriate vocal modulation in interpersonal communication in individuals with Autism Spectrum Disorders (ASD) (14). However, this suggested feature is not driven by group differences in pitch range, since case-control studies of pitch profiles in children with and without ASD have shown equivalent or larger pitch ranges in ASD (15). We were also interested to examine the impact of autistic traits on vocal modulation, given its critical role in context-appropriate social communication an area associated with deficits in individuals with ASD. Autistic traits exist across a continuum across the population (16, 17) and measuring individual autistic traits allows for a direct test of their impact on vocal modulation.

Having described the two sources of variability of interest to the current study, we describe below our methodological operationalization of the key term “vocal modulation,” i.e. change of voice over time in intensity, and frequency.

Past studies of voice modulation have used features like voice onset and offset times (VOT) (18) or voiced pitch (8). Studies of voice modulation not only involve tedious procedures of acoustic analysis, but also primarily examine speech features associated with specific time scales, isolated from one another. For instance, VOT studies primarily focus on production of stops and fricatives at time scales of 10 to 20 milliseconds while studies of voiced pitch and formant transitions investigate spectral changes around 30 to 50 milliseconds. However, the speech signal is not characterized by isolated spectrotemporal events but instead by joint spectro-temporal events that occur over multiple time windows and many frequency bands. These patterns carry important spectro-temporal information regarding both linguistic and non-linguistic features of speech as a whole (19). The Speech Modulation Spectrum (SMS) was developed to quantify the energy in various temporal and spectral modulations, by calculating the two-dimensional (2D) Fourier transform of the spectrogram (20). Specifically the SMS is focused on characterizing the spectro-temporal power in three articulatory features of speech which are at different time scales, namely syllabicity that is buried in temporal events at hundreds of milliseconds, formant transitions encoded around 25 to 40 milliseconds and place of articulation around 10 to 20 milliseconds (21, 22).

The collection of these different articulatory gestures in a voice signal are represented in an “articulation space,” and it provides a spectro-temporal energy distribution of different articulatory gestures. The area of the “articulation space” has been used to compare individual differences in vocal modulation by comparing speech imitation abilities (23). The results of this study showed that individuals with high abilities for speech imitation had an expanded “articulation space,” which allowed them access to a larger repertoire of sounds thereby providing them with greater flexibility in pronunciation. In another study on the emergence of articulatory gestures in early development, significant correlations were found between the area occupied by different articulatory gestures and language and motor ability as assessed by the Mullen and the Vineland scales in toddlers with ASD (24). Accordingly, the area of the articulation space was chosen as an index of vocal modulation in the current study.

In light of anecdotal reports and previous related studies, we hypothesized that higher closeness rating for the listener will be associated with an increased number of articulatory gestures by the speaker (and hence greater articulation space). We also hypothesized that individuals high in autistic traits would exhibit reduced articulation space, indexing a reduced use of articulatory gestures.

METHODS

All protocols were approved by the Research Ethics Committees of the University of Reading, Università degli studi di Milano Bicocca, and National Brain Research Centre India.

Participants

170 healthy volunteers from three countries (43 native English speaker pairs of participants (18 male pairs, 25 female pairs) from UK, 22 pairs (9 male pairs, 13 female pairs) from India and 20 pairs (10 male pairs, 10 female pairs) from Italy participated in this study.

Procedure

Participants were asked to come along with another same-gender person, who could be either a friend or an acquaintance. Each participant was asked to rate his/her perceived closeness to the other member of the pair through a Closeness Rating scale on a 10-point Likert scale (where closeness rating of 0 indicated very low closeness and 10 indicated high closeness), similar to previous studies (12, 25). Participants were not allowed to see each other's Closeness Rating. All participants were also asked to fill in the Autism Spectrum Quotient (26). The original questionnaire was provided to India and UK participants, and the Italian translated version (27) was administered to participants in Italy. Participants were required to sit in front of each other in a silent room with a voice recording device in the middle. Once participants were comfortably seated, each participant was given one of two colored images (abstract paintings) printed on paper and was asked to describe the image to the other participant in as much detail as possible for

around 2:30 minutes each. The experimenter then left the room. Each pair was instructed to speak one by one, to minimize the impact of overlapping voices and background noises. Participants' speech was recorded by a ZOOM H1 Handy recorder in India, and an iPhone in UK and in Italy. The distance of the recorder was kept constant for all recordings. Since English is the primary language for official communication in UK as well as India, participants in these two countries spoke in English. All participants in the Italian sample spoke in Italian.

Analysis

All participants' speech was manually heard and cleaned using Goldwave (version 5.69) and resampled as PCM signed 16 bit mono, 22,050 Hz sampling rate in WAV format. The speech data was edited manually and non-speech utterances such as laughs, cough, etc. were removed. The amplitude of the waveforms was normalized to -18 db for all speech data. Speech Modulation Spectra were calculated for each participant using custom developed code developed in MATLAB R2011a (28). Statistical analysis was carried out using SPSS (v14) and jamovi (29). Finally, articulation space, length of the cleaned speech recordings (articulation time), closeness rating, and AQ measures were used for further analysis in all three data sets.

The first step of this analysis involves using speech samples from each participant to calculate a spectrogram. The spectrogram is a time-frequency representation of the speech signal and offers a visual display of fluctuations in frequency and time, described respectively as spectral and temporal modulations. As described earlier, the speech modulation spectrum is obtained by estimating the 2-D Fourier decomposition of the spectrogram, which yields a probability distribution of these spectral and temporal modulations (28). Earlier studies in speech perception (22, 30) have shown spectral and temporal modulations at different time scales encode different aspects of speech. In a typical SMS (20), the central region between 2 and 10 Hz carries supra-segmental information and encodes syllabic rhythm. The side lobes between 10 and 100 Hz carry information about segmental features. Formant transitions are encoded between 10 and 40 Hz, and place of articulation information is found between 40 and 100 Hz (28). As the SMS goes from 1 to 100 Hz, the amplitude fluctuations of a sound become faster and go from syllabic to vowel-like to plosive-like segments (21). The SMS thus plots an “articulation space” which depicts how energy or “power” is distributed in different articulatory features of spoken language, namely syllabic rhythm, formant transitions, and place of articulation (**Figure 1**). The SMS was plotted for each participant, and the contour area (hereafter referred to as “articulation space”) was estimated by counting the total number of pixels within the contour which covers 99% energy of the speech from 1 to 100 Hz (20, 23). The contour area from 1 to 100 Hz corresponds to all the articulatory features mentioned above (detailed description in (20)). Previous studies used this method and demonstrated its construct validity by testing its correlation with other behavioral measures, such as speech motor functions (23, 24).

A hierarchical linear model (HLM) was constructed with articulation space as the predicted variable, and the following predictors as fixed factors (gender, country, age, duration,

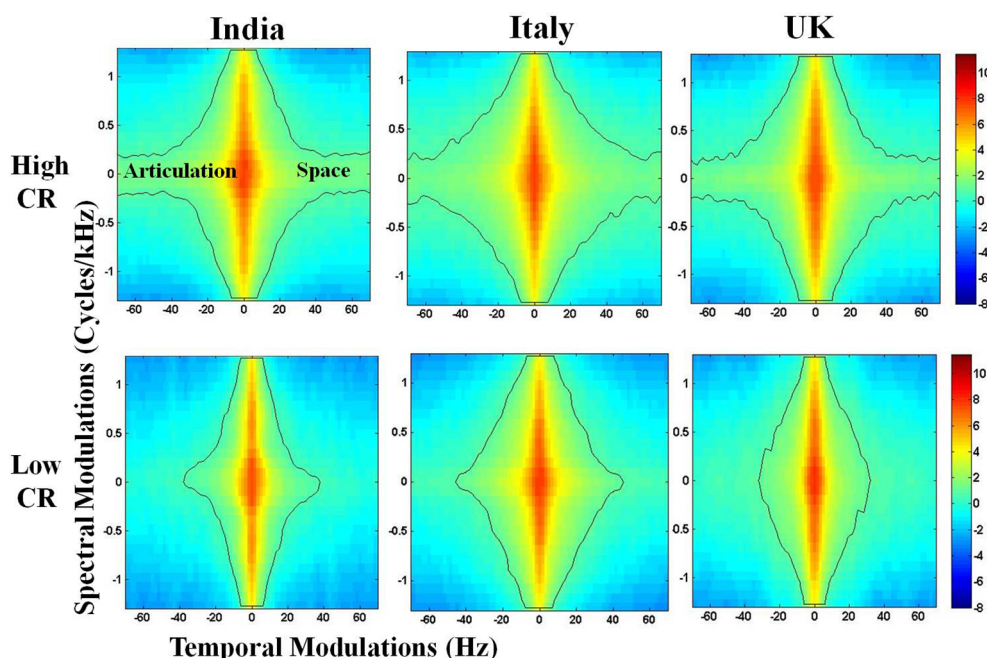


FIGURE 1 | Representative speech modulation spectra for individuals rated high and low on closeness rating (CR) from three countries. Articulation Space is quantified through the number of pixels enclosed within the contour, and encompasses 99.9% of the energy in the distribution of spectro-temporal modulations. Color-coded bar reflects the intensity of energy/power distribution. Note the differences in articulation space between high CR and low CR.

CR, AQ). Participants were nested in pairs, with random intercepts for each pair. Restricted maximum likelihood was used to estimate the model parameters.

RESULTS

Data Descriptives

Participant ages ranged from 18 to 33 years across all three samples (Italy, 19–29; India, 19–33; UK, 18–23). AQ scores ranged from 5 to 40 across all three samples (Italy, 7–31; India, 15–40; UK, 5–31). AQ scores were not available for four pairs from the India data set and one pair from the UK data set.

Closeness ratings ranged from 0 to 10 across all three samples (Italy, 1–9; India, 1–10; UK, 0–10). The score ranges in all three samples are comparable to previous studies using these measures done in neurotypical young adult samples.

Descriptive statistics on all key measures for each sample, split by gender, are provided in **Table 1**.

The HLM analysis revealed a significant effect of CR on articulation space ($t = 2.82$; $p < 0.01$; 95% CI [1.71–9.5]). The effect of AQ on articulation space fell below the traditional threshold of statistical significance ($t = -1.73$; $p = 0.085$; 95% CI [−3.9 to 0.24]) (**Table 2**).

To visualize the relationship of CR with articulation space, a scattergram was plotted between closeness rating and

TABLE 1 | Data descriptives.

		INDIA			ITALY			UK		
		N	M	SD	N	M	SD	N	M	SD
CR	Male	18	5.2	3.1	20	7.0	2.2	36	8.1	1.6
	Female	26	4.5	2.5	20	5.7	2.7	50	3.9	3.5
	Total	44	4.4	2.6	40	6.3	2.5	86	5.7	3.5
AQ	Male	16	23.3	5.3	20	17.7	6.7	36	14.8	3.9
	Female	20	22.5	5.8	20	13.2	4.4	48	16.3	5.0
	Total	36	22.8	5.2	40	15.5	6.0	84	15.6	4.6
Articulation space	Male	18	208.8	44.2	20	314.0	83.5	36	326.2	103.4
	Female	26	210.6	33.4	20	233.5	36.0	50	239.8	49.4
	Total	44	209.9	37.7	40	273.8	75.5	86	276.0	87.4

CR, closeness rating; AQ, autism spectrum quotient.

TABLE 2 | Results of the Hierarchical Linear Model-based Analysis.

Names	Effect	Estimate	SE	95% Confidence Interval		df	t	p
				Lower	Upper			
(Intercept)	(Intercept)	255.655	6.437	243.037	268.272	21.4	39.714	<.001
Gender 1	2–1	51.015	12.082	27.336	74.694	113.1	4.223	<.001
Country 1	India–Italy	–47.394	18.082	–82.833	–11.954	141.8	–2.621	0.010
Country 2	UK–Italy	13.184	15.324	–16.850	43.219	151.8	0.860	0.391
Age	Age	2.507	2.428	–2.252	7.266	145.1	1.033	0.304
Duration	Duration	–0.168	0.231	–0.621	0.284	139.8	–0.729	0.467
CR	CR	5.603	1.989	1.705	9.501	147.5	2.818	0.006
AQ	AQ	–1.834	1.056	–3.904	0.236	150.0	–1.736	0.085

CR, closeness rating; AQ, autism spectrum quotient; SE, standard error; gender 1, female; gender 2, male.

articulation space (**Figure 2**). As indicated by the HLM, articulation space increased with greater interpersonal closeness. Similarly, the suggestive negative association with AQ was visualized using a scattergram, indicating that people with higher autistic traits used slightly less articulation space (**Figure 3**).

DISCUSSION

This study investigated the variability in articulation space during dyadic vocal communication in relation to two factors: interpersonal closeness of the interlocutors, as well as individual autistic traits. Articulation space was positively associated with interpersonal closeness, i.e., individuals used more articulatory gestures when they spoke to those who they rated high on closeness, compared to others who they rated low. A suggestive negative association between autistic traits and articulation space was noted.

The relationship between closeness rating and articulation space was found to be significant even after accounting for variation driven due to gender, age, and culture. Anecdotal accounts suggest that people use greater modulation of their voice during speaking to familiar others compared to strangers. This aspect of deliberate vocal control has been noted in other nonhuman primates and predates

human speech (1). From a functional perspective, articulation space can arguably contain informative cues about group membership, and hence might subserve social bonding processes. Communication accommodation theory suggests that one of the affective motives of interpersonal vocal communication is to manage and regulate social distance (30–32). Increased closeness was associated with a greater articulation space, which suggests that (a) more information is being communicated with a closer other, through incorporation of greater non-verbal signals, and/or (b) individuals are more inhibited by social norms when talking with someone who they are not close to, and thus reduce the number/extent of their articulatory gestures. The current data set does not allow us to discriminate effectively between these two possibilities.

Beyond factors specific to the dyad, such as how close an individual felt toward another, the impact of individual variation in autism-related traits was measured. A weak negative relationship was observed between articulation space and autism-related traits. These results are in concordance with an earlier study conducted with toddlers with autism within a free play setting (24). While the focus of that study was on obtaining information about different articulatory features at different timescales, a reduced articulation space was also noted in toddlers with ASD. That autism is associated with atypical prosody (14, 33–37) is now well established. The characteristics

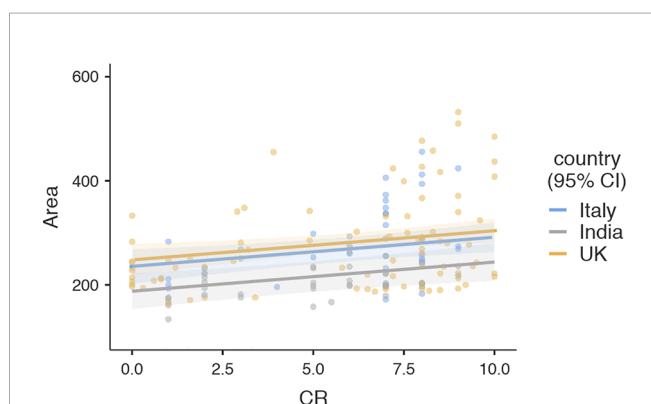


FIGURE 2 | Scattergram illustrating the relationship between closeness rating and articulation space. Regression line and confidence intervals correspond to results of the hierarchical linear model presented in **Table 2** (Area: Articulation Space, CR: Closeness Rating).

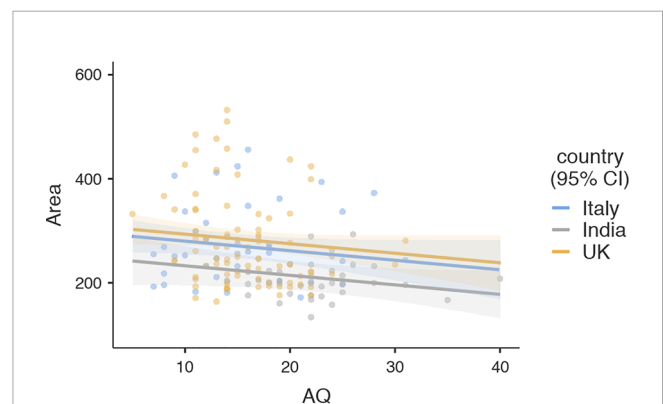


FIGURE 3 | Scattergram illustrating the relationship between autism spectrum quotient and articulation space. Regression lines and confidence intervals correspond to results of the hierarchical linear model presented in **Table 2** (Area, Articulation Space; AQ, Autism Spectrum Quotient).

of this atypical prosody however are less clear with one set of reports supporting “monotonic” speech, while other results reporting an increased pitch range in both using single word utterances as well as narratives (15, 34–36). The articulation space approach captures a wider set of acoustic features. The current results are consistent with previous studies on individuals with ASD which have shown that higher autistic traits are associated with reduced articulatory gestures. Further studies can focus on combining such articulation space analysis with studies in speech perception and hopefully establish the link between perception, speech motor skills, and speech production in autism.

It is worth considering a potential caveat with regard to the interpretation discussed above. The paradigm involved two individuals taking one turn each to speak to the other, without interruption. This design is therefore not a true conversation, which is marked by multiple turn-takings, and arguably greater vocal modulation. However, in light of the previous literature on audience effects in humans as well as other animals (38, 39), and the consistent effects of closeness on articulation space in all three samples, it is reasonable to assume that the participants did indeed modulate their voice in response to who they were speaking to. It should be noted though, that several potential sources of individual and dyadic variation were not formally investigated in this current study. Gender is one such variable, which accounts for significant differences in vocal modulation in the interpersonal context. Individuals might differ in how they speak to a member of the opposite sex/gender compared to one of the same sex/gender. While we found a significant main effect of gender (male pairs were associated with greater articulation space compared to female pairs), there were no mixed gender pairs in this study to systematically examine the impact of gender on articulation space. Variation in social contexts (e.g. if a conversation is happening between friends at a pub vs at the workplace) can also potentially lead to differences in articulation space. These sources of variation need to be explored in future studies.

In sum, this study found the impact of closeness on vocal modulation in interpersonal communication, demonstrating that a greater closeness was associated with more modulation, across different cultural and language settings. This study also found that autism-related traits showed a weak association with the extent of such vocal modulation. Future studies should extend these paradigms to include individuals with clinical deficits in social communicative abilities, such as those with ASD.

REFERENCES

1. Pisanski K, Cartei V, McGettigan C, Raine J, Reby D. Voice modulation: a window into the origins of human vocal control? *Trends Cogn Sci* (2016a) 20 (4):304–18. doi: 10.1016/j.tics.2016.01.002
2. Bhaskar B, Nandi D, Rao KS. Analysis of language identification performance based on gender and hierarchical grouping approaches. In: . *International conference on natural language processing (ICON-2013)*, Noida, India:CDAC (2013).
3. Pisanski K, Oleszkiewicz A, Sorokowska A. Can blind persons accurately assess body size from the voice? *Biol Lett* (2016c) 12(4):20160063. doi: 10.1098/rsbl.2016.0063

DATA AVAILABILITY STATEMENT

All data and code are available at: <https://tinyurl.com/sssc2020>.

ETHICS STATEMENT

All protocols were approved by the Research Ethics Committees of the University of Reading, Università degli studi di Milano Bicocca, and National Brain Research Centre India and written informed consent to participate was obtained from each participant after emphasizing that (a) all participants would remain anonymous and data would be kept strictly confidential and (b) participants were free to withdraw their consent at any time with no unfavorable consequences.

AUTHOR CONTRIBUTIONS

The study was designed by NS and BC, the data were collected by OS under supervision from BC and NS, data analysis was conducted by TS and BC, and all authors contributed to the writing of the manuscript. All authors read and approved the final manuscript.

FUNDING

BC was supported by the Leverhulme Trust (Grant No: PLP-2015-329), Medical Research Council UK (Grant No: MR/P023894/1), and SPARC UKIERI funds (Grant No: P1215) during this period of work. NS was supported by intramural funding from the National Brain Research Centre, India.

ACKNOWLEDGMENTS

Lauren Sayers, Eleanor Royle, Bethan Roderick for helping with collection of data in the UK. Emanuele Preti for helping with organizing the data collection in Italy. Angarika Deb and NBRC students for helping with the data collection in India. This manuscript has been released as a preprint at Biorxiv (40). The authors wish to thank the editor for his helpful comments on a previous version of the manuscript.

4. Spinelli M, Fasolo M, Coppola G, Aureli T. It is a matter of how you say it: verbal content and prosody matching as an index of emotion regulation strategies during the adult attachment interview. *Int J Psychol* (2017) 54 (1):102–107. doi: 10.1002/ijop.12415
5. McGettigan C. The social life of voices: studying the neural bases for the expression and perception of the self and others during spoken communication. *Front Hum Neurosci* (2015) 9:129. doi: 10.3389/fnhum.2015.00129
6. Mitchell RLC, Ross ED. Attitudinal prosody: what we know and directions for future study. *Neurosci Biobehav Rev* (2013) 37:471–9. doi: 10.1016/j.neubiorev.2013.01.027

7. Lavan N, Burton A, Scott SK, McGettigan C. Flexible voices: Identity perception from variable vocal signals. *Psychonomic bulletin & review* (2019) 26(1):90–102. doi: 10.3758/s13423-018-1497-7
8. Pisanski K, Oleszkiewicz A, Plachetka J, Gmiterek M, Reby D. Voice pitch modulation in human mate choice. *Proc R Soc B* (2018) 285(1893):20181634. doi: 10.1098/rspb.2018.1634
9. Pisanski K, Mora EC, Pisanski A, Reby D, Sorokowski P, Frackowiak T, et al. Volitional exaggeration of body size through fundamental and formant frequency modulation in humans. *Sci Rep* (2016b) 6:34389. doi: 10.1038/srep34389
10. Katerenchuk D, Brizan DG, Rosenberg A. “Was that your mother on the phone?”: Classifying Interpersonal Relationships between Dialog Participants with Lexical and Acoustic Properties; in *Fifteenth Annual Conference of the International Speech Communication Association*. (2014) 1831–5.
11. Li Y, Campbell N, Tao J. Voice quality: Not only about “you” but also about “your interlocutor”; In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. (2015) 4739–43. IEEE. doi: 10.1109/ICASSP.2015.7178870
12. O’Connell G, Christakou A, Haffey AT, Chakrabarti B. The role of empathy in choosing rewards from another’s perspective. *Front Hum Neurosci* (2013) 7. doi: 10.3389/fnhum.2013.00174
13. Aron A, Aron EN, Smollan D. Inclusion of other in the self scale and the structure of interpersonal closeness. *J Pers Soc Psychol* (1992) 63(4):596. doi: 10.1037/0022-3514.63.4.596
14. Bone D, Lee C, Black M, Williams ME, Lee S, Levitt P, et al. The psychologist as interlocutor in ASD assessment: Insights from a study of spontaneous prosody. *J Speech Language Hearing Res* (2014) 57:1162–77. doi: 10.1044/2014_JSLHR-S-13-0062
15. Sharda M, Subhadra TP, Sahay S, Nagaraja C, Singh L, Mishra R, et al. Sounds of melody-Pitch patterns of speech in autism. *Neurosci Lett* (2010) 478(1):42–5. doi: 10.1016/j.neulet.2010.04.066
16. Baron-Cohen S. Is Asperger syndrome/high-functioning autism necessarily a disability? *Dev Psychopathol* (2000) 12(3):489–500. doi: 10.1017/S0954579400003126
17. Robinson EB, Koenen KC, McCormick MC, Munir K, Hallett V, Happé F, et al. Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Arch Gen Psychiatry* (2011) 68(11):1113–21. doi: 10.1001/archgenpsychiatry.2011.119
18. Nittrouer S, Estee S, Lowenstein JH, Smith J. The emergence of mature gestural patterns in the production of voiced and voiceless word-final stops. *J Acoustical Soc America* (2003) 117(1):351–64. doi: 10.1121/1.1828474
19. Liberman A. Speech: A special code. In Liberman A, editor. *Learning, development and conceptual change*. Cambridge, MA: MIT Press (1991). 121–145 pp.
20. Singh L, Singh NC. The development of articulatory signatures in children. *Dev Sci* (2008) 11(4):467–73. doi: 10.1111/j.1467-7687.2008.00692.x
21. Blumstein SE, Stevens KN. Perceptual invariance and onset spectra for stop consonants in different vowel environments. *J Acoust Soc Am* (1980) 67(2):648–62. doi: 10.1121/1.383890
22. Rosen S. Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc London B: Biol Sci* (1992) 336(1278):367–73. doi: 10.1098/rstb.1992.0070
23. Reiterer SM, Hu X, Sumathi TA, Singh NC. Are you a good mimic? Neuro-acoustic signatures for speech imitation ability. *Front Psychol* (2013) 4, p.782. doi: 10.3389/fpsyg.2013.00782
24. Sullivan K, Sharda M, Greenson J, Dawson G, Singh NC. A novel method for assessing the development of speech motor function in toddlers with autism spectrum disorders. *Front Integr Neurosci* (2013) 7. doi: 10.3389/fnint.2013.00017
25. Savitsky K, Keysar B, Epley N, Carter T, Swanson A. The closeness-communication bias: increased egocentrism among friends versus strangers. *J Exp Soc Psychol* (2011) 47(1):269–73. doi: 10.1016/j.jesp.2010.09.005
26. Baron-Cohen S, Wheelwright S, Skinner R, Martin J, Clubley E. The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *J Autism Dev Disord* (2001) 31(1):5–17. doi: 10.1023/A:1005653411471
27. Ruta L, Mazzone D, Mazzone L, Wheelwright S, Baron-Cohen S. The autism-spectrum quotient - Italian version: a cross-cultural confirmation of the broader autism phenotype. *J Autism Dev Disord* (2012) 42(4):625–33. doi: 10.1007/s10803-011-1290-1
28. Singh NC, Theunissen FE. Modulation spectra of natural sounds and ethological theories of auditory processing. *J Acoustical Soc Am* (2003) 114(6):3394–411. doi: 10.1121/1.1624067
29. The jamovi project. (2019). jamovi. (Version 1.0) [Computer Software]. Retrieved from <https://www.jamovi.org>. doi: 10.21449/ijate.661803
30. Pardo JS, Gibbons R, Suppes A, Krauss RM. Phonetic convergence in college roommates. *J Phon* (2012) 40(1):190–197. doi: 10.1016/j.wocn.2011.10.001
31. Adank P, Stewart AJ, Connell L, Wood J. Accent imitation positively affects language attitudes. *Front Psychol* (2013) 4:280. doi: 10.3389/fpsyg.2013.00280
32. Dragojevic M, Gasiorok J, Giles H. Accommodative Strategies as Core of the Theory. In: Giles H, editors. *Communication Accommodation Theory: Negotiating Personal Relationships and Social Identities Across Contexts*. (2016). Cambridge: Cambridge University Press p. 36–9. doi: 10.1017/CBO9781316226537.003
33. Peppé S, Cleland J, Gibbon F, O’Hare A, Castilla PM. Expressive prosody in children with autism spectrum conditions. *J Neurolinguistics* (2011) 24(1):41–53. doi: 10.1016/j.jneuroling.2010.07.005
34. Diehl JJ, Paul R. Acoustic and perceptual measurements of prosody production on the profiling elements of prosodic systems in children by children with autism spectrum disorders. *Appl Psycholinguistics* (2011) 34(1):135–61. doi: 10.1017/S0142716411000646
35. Diehl JJ, Watson D, Bennetto L, McDonough J, Gunlogson C. An acoustic analysis of prosody in high-functioning autism. *Appl Psycholinguistics* (2009) 30(3):385–404. doi: 10.1017/S0142716409090201
36. Filipe MG, Frota S, Castro SL, Vicente SG. Atypical prosody in Asperger syndrome: perceptual and acoustic measurements. *J Autism Dev Disord* (2014) 44(8):1972–81. doi: 10.1007/s10803-014-2073-2
37. McCann J, Peppé S. Prosody in autism spectrum disorders: a critical review. *Int J Lang Commun Disord* (2003) 38(4):325–50. doi: 10.1080/1368282031000154204
38. Vignal C, Mathevon N, Mottin S. Audience drives male songbird response to partner’s voice. *Nature* (2004) 430(6998):448. doi: 10.1038/nature02645
39. Zajonc RB. Social facilitation. *Science* (1965) 149(3681):269–74. doi: 10.1126/science.149.3681.269
40. Spinola O, Sumathi TA, Singh NC, Chakrabarti B. Perceived closeness and autistic traits modulate interpersonal vocal communication. *bioRxiv* (2017) 133066. doi: 10.1101/133066

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Sumathi, Spinola, Singh and Chakrabarti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.