Journal of Autism and Developmental Disorders
https://doi.org/10.1007/s10803-021-05087-5

BRIEF REPORT

Brief Report: Autistic Traits Predict Spectral Correlates of Vowel Intelligibility for Female Speakers

Jason Bishop[1,2,3] · Chen Zhou[2] · Katarina Antolovic[1] · Lauren Grebe[1] · Kyung Hae Hwang[4] · Gerald Imaezue[1] · Ekaterina Kistanova[2] · Kyung Eun Lee[1] · Katherine Paulino[1] · Sichen Zhang[4]

Accepted: 14 May 2021

Abstract

A growing body of research finds that neurotypical autistic traits are predictive of speech perception and language comprehension patterns, but considerably less is known about the influence of these traits on speech production. In this brief report, we present an analysis of vowel productions from 74 American English speakers who participated in a communicative speaking task. Results show higher autistic trait load to be broadly and inversely related to spectral correlates of vowel intelligibility. However, the statistical significance of this relationship is specific to autistic traits along the pragmatic communication dimension, and limited to female speakers.

Keywords Speech production · Autism-spectrum quotient · Broad autism phenotype · Pragmatic communication · Sex differences · Intelligibility

Introduction

A growing body of research suggests that some aspects of neurotypical language processing may be modulated by subclinical autistic traits, i.e. characteristics associated with the broad autism phenotype. For example, recent findings link higher levels of autistic traits—at least traits related to pragmatic communication skills—to decreased lateralization of language processing (Jouravlev et al., 2020) and to reduced sensitivity to both prosodic cues (Bishop, 2017; Bishop, Kuo & Kim 2020) and pragmatic context (Nieuwland et al., 2010; Xiang et al., 2013) during speech perception and language comprehension tasks. In this brief report, we investigate the role of the broad autism phenotype in speech production, which has thus far been the subject of less study than speech perception and comprehension. The question we asked was whether measures of autistic traits are systematically related to acoustic correlates of vowel intelligibility. Given the relation between autistic traits and pragmatic communication skills, our basic prediction was that individuals with higher levels of autistic traits should tend to produce less acoustically distinct vowel categories during communicative speaking tasks.

We present this study in the context of another recent investigation along similar lines. In a set of carefully designed experiments with American English speakers, Turnbull (2019) found that scores on the Autism Spectrum Quotient (AQ; Baron-Cohen et al., 2001) were only modestly predictive of the enhancements speakers made when instructed to read scripted materials aloud for the benefit of an imaginary hearing-impaired listener (a common method for eliciting clear speech in the laboratory). In particular, speakers with lower AQ scores (indicating lower autistic trait load) made larger adjustments to word durations than speakers with higher AQ scores. However, as just mentioned, the effect was modest in size, and was only marginally significant when vowel duration rather than word duration was the acoustic measure. Moreover, and of primary interest to the present report, AQ was found to have no effect on the extent of speakers' vowel formant dispersion, a well-established correlate of vowel intelligibility reflecting spectral rather than temporal enhancement (e.g. Carl et al., 2020).

✉ Jason Bishop
jbishop@gc.cuny.edu

[1] Program in Speech-Language-Hearing Sciences, City University of New York-Graduate, Center, 365 5th Ave., New York, NY 10016, USA

[2] Program in Linguistics, City University of New York-Graduate Center, New York, NY, USA

[3] Linguistics Program, City University of New York-College of Staten Island, New York, NY, USA

[4] Teachers College, Columbia University, New York, NY, USA

Published online: 26 May 2021

Table 1 Carrier sentences used to elicit targets in the four conditions based on prosodic position and ordering relative to filler items

Prosodic position/order	Carrier sentence
Phrase initial/target-filler:	___**[Target]**___ is the first word, and __[Filler]____ is the second word
Phrase final/target-filler:	The first word is ___**[Target]**___, and the second word is __[Filler]___
Phrase-initial/filler-target:	___[Filler]___ is the first word, and __**[Target]**____ is the second word
Phrase final/filler-target:	The first word is ___[Filler]___, and the second word is __**[Target]**___

There are a few reasons to suspect that Turnbull's findings may have underestimated the relationship between autistic trait load and acoustic correlates of intelligibility in vowel production, and we sought to address these in the experiment below. First, if autistic traits influence acoustic correlates of vowel intelligibility, we assumed this would be most apparent for speech produced in communicative, socially-interactive situations. Therefore, although the speech we sampled was produced in a laboratory setting using read materials, it was elicited from speakers in a communicative task involving an actual listener. Second, we assumed that autistic traits related specifically to pragmatic communication skills would be the ones most relevant to the speech phenomenon in question. We therefore predicted that measures intended to estimate these traits specifically—such as the Communication subscale of the AQ and the Pragmatic Language subscale of the Broad Autism Phenotype Questionnaire (BAPQ; Hurley et al., 2007)—would have a stronger relationship with vowel production than their respective composite measures. Third, some previous work has found speaker sex to modulate the effects of autistic trait load on speech production (e.g. Yu, 2016), although the underlying mechanisms are not yet understood and the number of studies remains quite small. We therefore collected speech samples from a group of native English speakers well-balanced for sex, and from an overall larger number of speakers than tested in Turnbull's (2019) investigation. Finally, in assessing the acoustic properties of speakers' vowel productions, we also controlled for aspects of prosodic context, such as accentuation and prosodic phrase position, both of which influence the articulation of speech segments in relevant ways (Keating, 2006). Thus, while our study is not, strictly speaking, a follow-up to Turnbull's (for example, we do not carry out a direct comparison between communicative and non-communicative speech styles, nor test lexical variables that Turnbull tested), we took a second look at the role that neurotypical autistic traits might play in the acoustic realization of vowel categories.

Methods

Stimuli

Materials were chosen for a speech production experiment intended to explore the distinctiveness of speakers' vowel categories in a two-dimensional acoustic space (based on the frequencies of the first and second formants). To this end, two sets of English CVC wordforms (differing in terms of surrounding consonantal context) were chosen to elicit the peripheral vowels /ɑ, ɛ, i, u/: *hod*, *head*, *heed*, *who'd* and *pot*, *bet*, *beat*, *boot*. The vowel /ɛ/ was chosen to represent the lower left corner of the vowel space in order to avoid confounding dialect-related variation involving words with /æ/ (Becker & Wong, 2010). In addition to these 8 CVC target words, the materials also included (a) 8 CVC wordforms intended to probe the acoustic properties of /s/ and (b) 28 unrelated filler CVC wordforms, neither of which are discussed further. All CVC wordforms were to be produced by participants in carrier sentences that were designed to control for certain prosodic properties as follows. First, to control for prosodic prominence (i.e. sentence stress), and more specifically to ensure that all CVC words would be emphatically stressed, target words were read in carrier sentences that contrasted the target CVC word with a filler CVC word. Second, the utterance position and the relative ordering of the target and filler word in each utterance were varied to control for prosodic effects related to utterance/phrase position, and were made to occur either early or late in a prosodic phrase. The carrier sentences and the four resulting prosodic conditions are shown in Table 1. Two repetitions of a pseudo-randomized list based on these four conditions resulted in 64 vowel samples per speaker (4 vowels $\times$ 2 consonantal contexts $\times$ 2 phrase positions $\times$ 2 target-filler orderings $\times$ 2 repetitions). These 64 vowel samples would be used to derive 16 vowel spaces (two for each prosodic condition), described further below.

Participants

Participants were 74 (36 male/38 female) monolingual speakers of American English from metropolitan New York City, drawn from a university population (mean age = 21.8 years; s.d. = 3.7; range = 18–38). All speakers confirmed that they had no known history of a speech, hearing or communication disorder and all were nominally compensated for their time.

Procedure

Production Task

Participants served as speakers in a production task that involved a communicative elicitation of the experimental

Table 2 Descriptive statistics related to autistic trait profiles for the participants in terms of composite scores (AQ and BAPQ) and their respective pragmatic/communication subscale scores (AQ-Comm and BAPQ-PragLang)

	AQ		AQ-Comm		BAPQ		BAPQ-PragLang	
	Males	Females	Males	Females	Males	Females	Males	Females
Mean	111.8	108.3	19.5	18.5	100.3	101.9	32.1	32
SD	9.8	15.2	3.8	5	19.2	22	8.1	7.6
Min	93	78	12	10	40	64	12	19
Max	129	146	27	32	133	155	47	49

materials described above. Recordings took place in a sound-attenuating audiological booth, with participants situated in front of a computer screen that presented the target and filler words. The experiment was blocked by carrier sentence, which (being short and simple) was memorized by participants and did not appear on the screen. Rather than simply producing these utterances aloud, the participants were asked to produce them for the benefit a listener (a research assistant in the study) who, sitting behind a curtain in the sound booth, would write down the target and filler CVCs in each sentence produced. After each sentence, the research assistant would then confirm comprehension to the speaker by replying with "got it." Participants' productions were recorded using a head-mounted Shure SM-10A microphone at a 41.1 kHz sampling rate and stored as *wav* files for later acoustic analysis, described below. Following the speech production task, participants also completed two questionnaires designed to estimate autistic trait load in neurotypical adults: the AQ (Baron-Cohen et al., 2001; Likert scoring of items was used) and the Broad Autism Phenotype Questionnaire (BAPQ; Hurley et al., 2007).

Acoustic Analysis

Formant measurements were taken at the midpoint of the vowel in target CVCs automatically by a script in Praat (Boersma & Weenink, 2020). Inspection of all formant values outside of two standard deviations of vowel-specific means (for each sex) was then carried out, with manual corrections made where clear formant tracking errors had occurred. From these productions, 16 quadrilateral vowel spaces were then derived based on the (bark-transformed) F1 and F2 values of /ɑ, ɛ, i, u/. These vowel spaces—not the individual vowels—were the primary unit of analysis, and were the basis for calculating two common acoustic correlates of vowel intelligibility: *vowel space area* (the geometric area of the vowel quadrilateral in an F1 × F2 space) and *vowel dispersion* (the mean Euclidean distance of each of the four peripheral vowels from the center of the vowel space, defined by the mean F1 and F2 of all four vowels). However, because these two measures were highly correlated, and especially because models of them were equivalent in terms of predictors, direction of effects and significance levels, we limit our discussion to vowel space area.

Results

Descriptive statistics related to the AQ, the BAPQ, and their respective pragmatic/communication-related subscales are shown in Table 2, separately for male and female speakers. The relationships these measures have with each other and with vowel space area are shown in the correlation matrix in Table 3, also separately for each sex. As can be seen in the correlation matrix, and as predicted, autistic traits were broadly and inversely related to vowel space area, indicating the tendency for speakers with higher autistic trait load to produce more compact spaces (and thus less acoustically distinct vowel categories). However, the relationship was stronger for female speakers and strongest when the measure of autistic traits was a communication-related subscale, in this case BAPQ-PragLang (illustrated in Fig. 1), to which the significant effects were limited.

Given that BAPQ-PragLang was the only measure significantly correlated with vowel space area, this was the relationship we tested further in a more rigorous statistical model. In particular, mixed-effects linear regression was used to predict vowel space area based on the following fixed-effect factors (with contrast coding for binary variables and continuous variables centered on their means): CVC context (h_d wordforms or bilabial_alveolar wordforms); prosodic position (phrase-initial or phrase-final); self-declared speaker sex (male or female); and BAPQ-PragLang scores. Based on the above patterns, a crucial fixed-effects term was also included for the interaction between sex and BAPQ-PragLang. Finally, the model included random intercepts and by-participant slopes for each fixed-effects factor. The results of the model are shown in Table 4; the significant interaction between sex and BAPQ-PragLang scores (and the non-significant simple effect for BAPQ-PragLang) confirm that higher BAPQ-PragLang scores were in fact associated with smaller vowel spaces for female speakers only. Larger vowel space areas were also associated with the h_d consonant context,

Table 3 Pearson correlation matrix for measures of autistic traits and vowel space area, with coefficients for males on the left and those for females on the right of each cell

	AQ-Comm		BAPQ		BAPQ-PragLang		V-space area	
AQ	**0.681** ***	**0.826** ***	**0.428** **	**0.808** ***	**0.435** **	**0.673** ***	− 0.131 ns	− 0.284 ·
AQ-Comm			**0.364** *	**0.735** ***	**0.515** **	**0.732** ***	− 0.152 ns	− 0.295 ·
BAPQ					**0.811** ***	**0.780** ***	− 0.116 ns	− 0.310 ·
BAPQ-PragLang							− 0.077 ns	**– 0.545** ***

***p < 0.001

**p < 0.01

*p < 0.05

· p < 0.1

ns p > 0.1

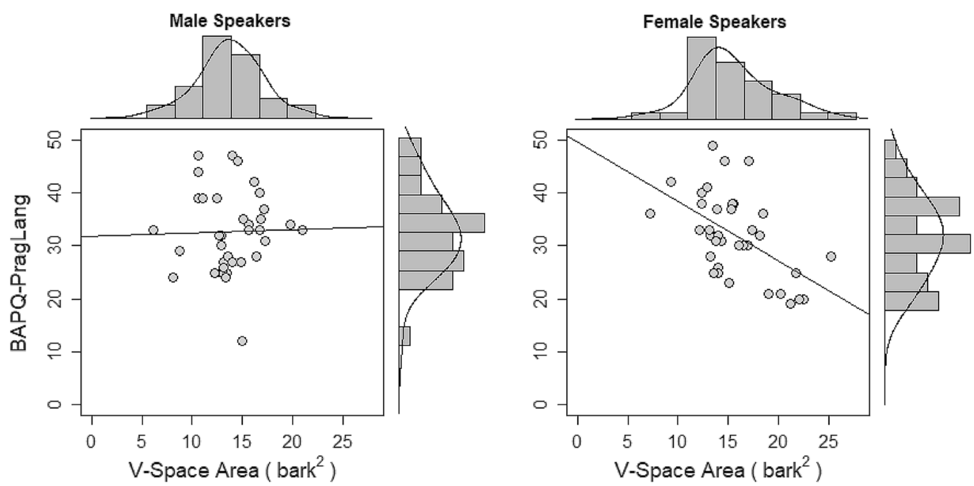

Fig. 1 Scatterplot-histograms showing the relationship between vowel space area and scores on the Pragmatic Language subscale of the Broad Autism Phenotype Questionnaire (BAPQ-PragLang) for male and female speakers

the phrase-initial prosodic position, and female sex—all expected patterns.[1]

While our primary interest was in spectral correlates of vowel intelligibility, it is possible that some temporal correlates may also be associated with BAPQ-PragLang, as pointed out to us by an anonymous reviewer. Important to remember is that longer acoustic duration can result from a number of different articulatory maneuvers, not all of which involve reaching more extreme articulatory targets (e.g. Edwards et al., 1992). Recall also that Turnbull (2019) found higher autistic trait load (as measured by composite AQ scores) to be significantly associated with shorter word durations in careful/communicative speech, but not shorter or less dispersed vowels (see also discussion of duration and speech rate in Yu, 2016). Nonetheless, Patel et al. (2020) recently found global speech rate to distinguish neurotypical speakers from those with clinically-diagnosed autism, the latter producing slower speech. Indeed, they also report tentative evidence that unaffected parents of individuals with autism spectrum disorder—especially mothers with features of the broad autism phenotype—produced slower speech than controls. Thus, slower speech may characterize the broader autism phenotype, but in sex-specific ways. We therefore explored whether vowel durations had significant predictive power in the model of vowel space area presented above. Durations of vowels in all test words were extracted and mean

[1] However, studies have often found utterance/phrase-initial position to have unreliable effects on articulatory targets for vowels (e.g. Cho & Keating, 2009). One possibility is that the effect for phrase position here reflects our speakers' interpreting utterance-initial target words as the focused element in a focus-topic construction; such contrasts in information structure have been shown to influence articulatory dynamics independent of prosodic structure (Mücke & Grice, 2014). An interesting question that we leave for future research is whether autistic trait load might predict the robustness with which neurotypical speakers encode information structure in fine-grained articulatory or acoustic measures. (See Krüger et al., 2018 for relevant findings from speakers drawn from the clinical population).

Table 4 Results for fixed-effects factors in the regression model used to predict vowel space area

	Estimate	Std. error	df	t-value	p-value
(Intercept)	– 1.0230	0.533	35.77	– 1.92	< 0.1
Consonantal context (h_d)	0.6709	0.160	73.05	4.19	**< 0.001**
Phrase position (final)	– 0.5287	0.210	73.27	– 2.52	**< 0.05**
Sex (female)	1.8191	0.740	69.76	2.46	**< 0.05**
BAPQ-PragLang	– 0.0313	0.065	31.73	– 0.48	> 0.1
Sex (female) × BAPQ-PragLang	– 0.2255	0.095	64.71	– 2.37	**< 0.05**

Parentheses show the non-default contrast for each factor

vowel duration was calculated for each speaker. In fact, we found these durations to have somewhat different relationships with BAPQ-PragLang depending on the speaker's sex; vowel duration and BAPQ-PragLang were positively and significantly correlated for male speakers ($r = 0.36$, $p < 0.05$) but negatively and only marginally for female speakers ($r = -0.30$, $p < 0.1$). However, these temporal differences were not significantly predictive of vowel space area size; adding a simple effect for vowel duration in the model presented above did not improve its fit to the data as determined by a log-likelihood ratio test ($\chi = 3.527$, $p > 0.1$). Moreover, nested models that tested a two-way interaction between vowel duration and BAPQ-PragLang ($\chi = 0.391$, $p > 0.1$) and a three-way interaction between vowel duration, BAPQ-PragLang and sex ($\chi = 0.5203$, $p > 0.1$) showed similar non-improvements. Thus, there appear to be some possibly interesting sex-specific temporal differences associated with autistic traits along the pragmatic communication dimension, but these temporal differences are not closely linked to the spectral ones that are of primary interest here.

Discussion

As described in the introduction, a number of studies have linked autistic trait load to language processing in neurotypical populations, but relatively little is known about how the broad autism phenotype appears in speech production patterns. Here we tested the simple hypothesis that individuals with a higher autistic trait load would produce less intelligible speech, with speech intelligibility being operationalized as well-established spectral measures of vowel distinctiveness. We noted that Turnbull (2019) found higher composite scores on the AQ to be associated with weaker enhancements to careful speech in the temporal domain (word durations, though not vowel durations) but not the spectral domain (vowel dispersion). The present findings suggest that reliable spectral effects can be detected, but are likely dependent on at least two crucial factors: the presence of communicative goals (a property of the speaking context) and autistic trait load along the pragmatic communication dimension (a property of the individual speaker). Moreover, the importance of these factors appears to be limited to female speakers, a pattern that has been reported in other recent studies of speech production (Yu, 2016). Given the large number of differences across a small number of relevant studies, we do not think it is currently possible to posit a mechanism for this apparent sex difference, although there are some avenues that should be considered in future research. For example, our findings may support the notion of a female-specific phenotype for autism that is characterized, in part, by stronger deficits in pragmatic/communicative language use (e.g. Frazier et al., 2014). Another possibility is more artifactual (and not mutually exclusive): due to gender-related differences in "camouflaging" (e.g. Lai et al., 2015), scores on self-report measures like the BAPQ and the AQ may be artificially low for some female participants, such that mid-to-high scores on such measures actually indicate high-to-very high "true" autistic trait load. What does seem clear, however, is that there is much work to be done if we are to understand the implications of the broad autism phenotype for adult neurotypical speech patterns, and this work will likely need to take possible sex (and/or gender) differences into account.

Acknowledgments The authors are grateful to Samantha Randazzo, Boram Kim, Alexandra Dia, Krissy Dellecave, and Maha Elcharfa for technical assistance with the study.

Author Contributions JB carried out collection of speech samples and JB and CZ carried out the main statistical analyses. All authors contributed to the study conception, design, the acoustic analysis, manuscript preparation and approval of final version.

Funding This study was supported in part by funding from Traditional Grant #63281-00-51 from the Professional Staff Congress of the City University of New York to the first author.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest to disclose.

Ethical Approval This research study was approved by the institutional IRB at the City University of New York and informed consent was obtained for all participants.

References

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17. https://doi.org/10.1023/a:1005653411471

Becker, K., Wong, A. (2010). The short—a system of New York City english: An update. In: Kyle Gorman & Laurel MacKenzie (Eds.), *University of Pennsylvania working papers in linguistics*, 15(2), Article 3. https://repository.upenn.edu/pwpl/vol15/iss2/1. Accessed 21 May 2021

Bishop, J. (2017). Focus projection and prenuclear accents: Evidence from lexical processing. *Language, Cognition and Neuroscience, 32*(2), 236–253. https://doi.org/10.1080/23273798.2016.1246745

Bishop, J., Kuo, G., & Kim, B. (2020). Phonology, phonetics, and signal-extrinsic factors in the perception of prosodic prominence: Evidence from rapid prosody transcription. *Journal of Phonetics, 82*, 100977. https://doi.org/10.1016/j.wocn.2020.100977

Boersma, P., Weenink, D. (2020). Praat: doing phonetics by computer [Computer program]. Version 6.1.16. Available at http://www.praat.org. Accessed 20 June 2020

Carl, M., Kent, R., Levy, E., & Whalen, D. (2020). Vowel acoustics and speech intelligibility in young adults with Down Syndrome. *Journal of Speech, Language, and Hearing Research, 63*(3), 674–687. https://doi.org/10.1044/2019_JSLHR-19-00204

Cho, T., & Keating, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics, 37*(4), 466–485. https://doi.org/10.1016/j.wocn.2009.08.001

Edwards, J., Beckman, M., & Fletcher, J. (1992). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America, 89*(1), 369–382. https://doi.org/10.1121/1.400674

Frazier, T., Georgiades, S., Bishop, S., & Hardan, A. (2014). Behavioral and cognitive characteristics of females and males with autism in the Simons simplex collection. *Journal of the American Academy of Child & Adolescent Psychiatry, 53*(3), 329-340.e3. https://doi.org/10.1016/j.jaac.2013.12.004

Hurley, R., Losh, M., Parlier, M., Reznick, J., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders, 37*(9), 1679–1690. https://doi.org/10.1007/s10803-006-0299-3

Jouravlev, O., Kell, A., Mineroff, Z., Haskins, A., Ayyash, D., Kanwisher, N., & Fedorenko, E. (2020). Reduced language lateralization in autism and the broader autism phenotype as assessed with robust individual-subjects analyses. *Autism Research*. https://doi.org/10.1002/aur.2393 Published online 15 September, 2020.

Keating, P. (2006). Phonetic encoding of prosodic structure. In J. Harrington & M. Tabain (Eds.), *Speech production: Models, phonetic processes, and techniques* (pp. 167–186). Psychology Press.

Krüger, M., Cangemi, F., Vogeley, K., & Grice, M. (2018). Prosodic marking of information status in adults with autism spectrum disorders. In: K. Klessa, J. Bachan, A. Wagner, M. Karpiński & D. Śledziński (Eds.), *Proceedings of the 9th International Conference on Speech Prosody*. (pp. 182–186). https://doi.org/10.21437/SpeechProsody.2018-37. Accessed 21 May 2021

Lai, M.-C., Lombardo, M., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/gender differences and autism: Setting the scene for future research. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*(1), 11–24. https://doi.org/10.1016/j.jaac.2014.10.003

Mücke, D., & Grice, M. (2014). The effect of focus marking on supralaryngeal articulation—is it mediated by accentuation? *Journal of Phonetics, 44*, 47–61. https://doi.org/10.1016/j.wocn.2014.02.003

Nieuwland, M., Ditman, T., & Kuperberg, G. (2010). On the incrementality of pragmatic processing: An ERP investigation of informativeness and pragmatic abilities. *Journal of Memory and Language, 63*(3), 324–346. https://doi.org/10.1016/j.jml.2010.06.005

Patel, S., Nayar, K., Martin, G., Franich, K., Crawford, S., Diehl, J., & Losh, M. (2020). An acoustic characterization of prosodic differences in autism spectrum disorder and first-degree relatives. *Journal of Autism and Developmental Disorders*, *50*, 3032–3045. https://doi.org/10.1007/s10803-020-04392-9

Turnbull, R. (2019). Listener-oriented phonetic reduction and theory of mind. *Language, Cognition and Neuroscience, 34*(6), 747–768. https://doi.org/10.1080/23273798.2019.1579349

Xiang, M., Grove, J., & Giannakidou, A. (2013). Dependency dependent interference: NPI interference, agreement attraction, and global pragmatic inferences. *Frontiers in Psychology, 4*, 708. https://doi.org/10.3389/fpsyg.2013.00708

Yu, A. C. L. (2016). Vowel-dependent variation in Cantonese /s/ from an individual-difference perspective. *Journal of the Acoustical Society of America, 139*(4), 1672–1690. https://doi.org/10.1121/1.4944992