An Association Between Phonetic Complexity of Infant Vocalizations and Parent Vowel Hyperarticulation

Ellen Marklund[1]*, Ulrika Marklund[2,3,4] and Lisa Gustavsson[1,4]

[1] Phonetics Laboratory, Stockholm Babylab, Department of Linguistics, Stockholm University, Stockholm, Sweden, [2] Division of Sensory Organs and Communication, Department of Biomedical and Clinical Sciences, Linköping University, Linköping, Sweden, [3] Speech and Language Clinic, Department of Neurology, Danderyd Hospital, Stockholm, Sweden, [4] Division of Speech and Language Pathology, Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden

Extreme or exaggerated articulation of vowels, or vowel hyperarticulation, is a characteristic commonly found in infant-directed speech (IDS). High degrees of vowel hyperarticulation in parent IDS has been tied to better speech sound category development and bigger vocabulary size in infants. In the present study, the relationship between vowel hyperarticulation in Swedish IDS to 12-month-old and phonetic complexity of infant vocalizations is investigated. Articulatory adaptation toward hyperarticulation is quantified as difference in vowel space area between IDS and adult-directed speech (ADS). Phonetic complexity is estimated using the Word Complexity Measure for Swedish (WCM-SE). The results show that vowels in IDS was more hyperarticulated than vowels in ADS, and that parents' articulatory adaptation in terms of hyperarticulation correlates with phonetic complexity of infant vocalizations. This can be explained either by the parents' articulatory behavior impacting the infants' vocalization behavior, the infants' social and communicative cues eliciting hyperarticulation in the parents' speech, or the two variables being impacted by a third, underlying variable such as parents' general communicative adaptiveness.

Keywords: vowel hyperarticulation, vowel space area, infant-directed speech, word-complexity measure for Swedish, WCM-SE, phonetic complexity, communicative adaptiveness, parent-infant interaction

INTRODUCTION

Infant-directed speech (IDS), the speech style commonly used when speaking to infants and small children (e.g., Soderstrom, 2007; Golinkoff et al., 2015), has been reported to facilitate or promote various aspects of early language development (e.g., Trainor and Desjardins, 2002; Singh et al., 2009; Ma et al., 2011; Bosseler et al., 2016; Foursha-Stevenson et al., 2017). One specific characteristic commonly found in IDS, clear or exaggerated vowel articulation (for a review, see Marklund and Gustavsson, 2020) compared to adult-directed speech (ADS), has been linked to better speech sound discrimination (Liu et al., 2003; García-Sierra et al., 2021), better word-recognition (Song et al., 2010), as well as larger receptive and productive vocabulary sizes (Hartman et al., 2017; Kalashnikova and Burnham, 2018). The present study focuses on the relationship between parents' articulation and children's own productions. Since phonetic detail of parent vocalizations can directly impact the phonetic form of child vocalizations (Goldstein and Schwade, 2008), it is hypothesized that a relationship will be found between degree of parents' vowel hyperarticulation and phonetic complexity of infant vocalizations.

Vowel Hyperarticulation and Language Development

The way in which speech is realized is highly variable. Speakers adapt their articulation to the perceived demands of the listener (Lindblom, 1990), for example speaking more clearly in noisy environments (Šimko et al., 2016) or using more reductions when uttering semantically predictable words than when uttering unpredictable words (Clopper and Pierrehumbert, 2008). These adaptations mean that individual speech sounds are realized on a continuum that ranges from exaggerated articulation (i.e., hyperarticulation) to relaxed articulation (i.e., hypoarticulation). IDS has been reported to be more hyperarticulated overall than ADS in some cases (vowels: Kuhl et al., 1997; see Marklund and Gustavsson, 2020 for a review; consonants: Cristià, 2010; McMurray et al., 2013) and more hypoarticulated than ADS overall in other cases (vowels: Englund and Behne, 2005, 2006; Benders, 2013; Englund, 2018; consonants: Sundberg and Lacerda, 1999; Benders et al., 2019). In this study, the focus is articulatory adaptation of vowels, and it will be referred to as vowel hyperarticulation (VH), since vowels in IDS to Swedish 12-month-old infants has previously been shown to be more hyperarticulated overall than vowels in Swedish ADS (Marklund and Gustavsson, 2020).

Vowel hyperarticulation has been suggested to promote infant learning of speech sound categories by exaggerating the distance between different vowel categories (e.g., Bernstein Ratner, 1984; Kuhl et al., 1997). However, detailed investigation of variability within and between different vowel categories in IDS contra in ADS suggests instead that IDS provides higher within-category variability than ADS (Cristia and Seidl, 2014), which, based on findings from second language learning, might contribute to more robust speech sound categories (Wong, 2014). Testing the relationship between VH in parent IDS and infant speech sound perception, a positive correlational relationship has been reported. Specifically, more hyperarticulation in mothers' IDS correlates with better discrimination of a native fricative contrast in Mandarin-learning 6–8-month-old and 10–12-month-old infants (Liu et al., 2003). Similar findings have been reported for English-Spanish bilingual infants at 11 and 14 months, but only for a native Spanish contrast, not a non-native Mandarin contrast (García-Sierra et al., 2021).

Vowel hyperarticulation in mothers' IDS also predicts infant vocabulary. In a longitudinal study, Australian English IDS to infants from 9 months and up predicted parent-reported expressive vocabulary size at ages 15 and 19 months (Kalashnikova and Burnham, 2018). Similarly, American English-learning children, whose mothers speak to them with high degrees of VH at 18 months, score higher in standardized vocabulary tests at 24 months than do children of mothers who use less hyperarticulation (Hartman et al., 2017). One potential explanation for an impact of VH on vocabulary size is that it helps infants more readily identify the words they hear. In favor of that notion, 19-month-old infants recognize known words faster when they are spoken with hyperarticulated vowels than when they are not (Song et al., 2010).

To summarize, previous research has reported a positive relationship between VH and other language skills/outcomes, specifically speech sound category perception (Liu et al., 2003; García-Sierra et al., 2021) and vocabulary development (Hartman et al., 2017; Kalashnikova and Burnham, 2018). This study expands the scope of language skills investigated in relation to VH to that of infants' own productions.

Parent Feedback and Infant Vocalizations

Infant vocalizations are influenced by parent social and vocal behavior. For example, removing social feedback from adult-infant interaction leads to fewer vocalizations from 5-month-old infants (Goldstein et al., 2009), and providing social feedback in response to 8-month-old's vocalizations results both in a higher number of vocalizations and more mature vocalizations (syllabic rather than vocalic, more canonical syllables, fully voiced) compared to social feedback non-contingent to infant productions (Goldstein et al., 2003). When it comes to vocal behavior, amount of parent IDS in parent-infant interactions correlates with amount of infant speech output (Dunst et al., 2012; Ramírez-Esparza et al., 2014; Spinelli et al., 2017). Vocal feedback from adults also influences the phonetic form of infant vocalizations. Just the presence of contingent vocal feedback leads to a higher ratio of mature vocalizations (syllabic rather than vocalic) in 3-month-olds (Bloom et al., 1987). In interactions between mothers and their 9.5-month-olds, the specific type of vocal feedback infants received impacted their vocal production. Infants whose mothers responded to vocalizations with a long vowel sound produced more fully voiced vocalizations, while infants whose mothers responded with words produced more consonant-vowel syllables (Goldstein and Schwade, 2008).

These findings demonstrate that the phonetic content of parent utterances can have an impact on the phonetic realization of infant vocalizations. This study therefore investigates whether degree of parent VH in IDS correlates with phonetic complexity of infant vocalizations.

This Study

In the present study, parents' articulatory adaptations in terms of VH is quantified using the difference in of vowel space area (VSA), that is, the area between the point vowel formant means in F_1-F_2 space, between IDS and ADS (Kuhl et al., 1997; Marklund and Gustavsson, 2020). Phonetic complexity of children's productions is measured using the Swedish adaptation of the Word-Complexity Measure (WCM-SE). This measure assigns a complexity score to each vocalization based on number of syllables, stress position, as well as number, position and combinations of articulatory complex speech sounds (Marklund et al., 2018). To determine whether VH could be established in the present IDS sample, the VSA was compared between IDS and ADS. To investigate whether there is a link between parents' articulatory adaptations and phonetic complexity of infants' vocalization, the relationship between VSA difference

between IDS and ADS, and infants' mean WCM-SE score, was tested for correlation.

MATERIALS AND METHODS

Participants

Nineteen infants (mean age = 12 months, range = 11.5–12.3, SD = 0.2) and their parents participated in the study (9 girls, 10 boys; 12 mothers, 7 fathers). All infants were born full-term (within 3 weeks of due date) and monolingual (defined as both parents speaking only Swedish with the infant). The majority of the parents ($n = 15$) had university education, and all had completed high school (which entails three non-obligatory years of education after the mandatory 9–10 years of basic education in Sweden). The participants constitute a subset of a larger group of subjects ($n = 72$), taking part in a longitudinal study in which parent-child interactions were recorded every 3–6 months, when the child was between 3 months and 3 years[1]. Free-play interaction sessions were recorded at each lab-visit, but additional tests and tasks varied between visits. An ADS sample was recorded at the 27-month-old visit. Inclusion criteria for participants in the present study were (a) there was a recording from the 12-month visit, (b) the infant was monolingual, and (c) there was sufficient ADS material from the same parent as in the 12-month visit.

The study was approved by the Regional Ethics Review Board (2015/63-31). Recruitment of infants in the appropriate age and their parents living in the greater Stockholm area was conducted via mail with invitation to participate in the study. Addresses were obtained via the Swedish Tax Agency. Parents received memory-sticks with all their audio and video recordings as thanks for their participation in the longitudinal study.

Recordings

Audio and video recordings of parent-infant interaction were made at Stockholm Babylab, the Phonetics Laboratory, Stockholm University in a comfortable carpeted studio equipped with age-appropriate furniture and toys. The recordings were made with three wall mounted cameras (Canon XA10) to capture all angles of the parent interacting with the infant. A fourth camera (GoPro Hero3), attached to the parent's chest, enabled close up video uptake of the infant from the parent perspective. To capture high-quality audio, omni-directional wireless lavalier microphones (Sennheiser EW 100 G2) were mounted on parent and infant, and one room microphone (AKG SE 300 B) was mounted on a high shelf. In the present study, audio from the lavalier microphones was used, since this setup enabled high-quality close-up recordings of the adult's speech and the infant's vocalizations with minimal interference from the other speaker.

The recording sessions lasted for approximately 10 min. The experimenter instructed parents to interact, play and talk with their infant as they typically would at home. After instructions and equipment arrangements, the researchers left the studio and

monitored the recording from the adjacent control room. For the ADS material, a conversation between the parent and the experimenter about the infant and their participation in the study was recorded. This recording took place at the beginning of the lab visit when the infant was around 27 months old. Parents were encouraged to speak as much and as freely as possible, and the conversation typically lasted around 2–5 min.

Vowel Hyperarticulation Estimations in Parent Speech

Estimation of VH in parent's IDS and ADS was performed as a part of a previous study, and detailed information about the procedure can be found there (Marklund and Gustavsson, 2020). In brief, parent speech was quasi-orthographically transcribed using ELAN 4.6.2-5.3 (Sloetjes and Wittenburg, 2008). The transcriptions (for annotation protocol see Gerholm, 2018) were automatically segmented, converted to IPA and aligned with their audio files using the web service WebMAUS General 5.33 of the Bavarian Archive for Speech Signals at the University of Munich (Schiel, 1999; Kisler et al., 2017).

Formants were estimated for the audio recordings using Praat 6.0.37 and 6.0.40 (Boersma and Weenink, 2018). Default settings were applied, except for formant ceiling and maximum number of expected formants (Escudero et al., 2009). Since reliability of formant estimations decreases with higher fundamental frequency (f_o), vowels with a median f_o exceeding 350 Hz were excluded (Lindblom, 1962; Monsen and Engebretson, 1983). To reduce the impact of coarticulation, the mid 40% of the vowel was used in the analysis (Marklund and Gustavsson, 2020). Each parents' average F_1 and F_2 for the point vowels /i/, /ɑ/, and /u/ in IDS and ADS were extracted and used for calculations of VSA. VSA was calculated in R 3.5.0 (R Core Team, 2018), separately for IDS (VSA$_{IDS}$) and ADS (VSA$_{ADS}$), using the following formula (Liu et al., 2003):

$$ABS((iF_1 \times (aF_2 - uF_2) + aF_1 \times (uF_2 - iF_2) +$$
$$uF_1 \times (iF_2 - aF_2)) / 2) \qquad (1)$$

This study uses the difference in VSA between IDS and ADS as the measure of parents' tendency to adapt their articulation when speaking to infants.

Phonetic Complexity Estimations of Infant Vocalizations

Infant vocalizations were phonetically transcribed in ELAN 5.8–5.9 (Sloetjes and Wittenburg, 2008) by two experienced phoneticians (authors UM and LG) according to an annotation protocol developed for compatibility with WCM-SE (Marklund et al., 2018). The protocol entailed transcribing all sounds present in the Swedish phoneme inventory as described in Engstrand (1999), with the addition of a number of other common allophones (**Table 1** and **Figure 1**). Segments not recognizable as any of those phonemes were marked as "C" (if consonant-like) or "V" (if vowel-like) in the transcription. If not possible to determine whether the sound was a consonant or a vowel,

TABLE 1 | The Swedish consonants used in the transcription of infant vocalizations.

	Bilabial	Labiodental	Dental	Retroflex	Alveolar	Palatal[†]	Velar	Uvular	Glottal
Plosive	p b		t d	ʈ ɖ			k g		
Nasal		m	n	ɳ			ŋ		
Trill			r					ʀ	
Tap/flap			ɾ						
Fricative		f v	s	ʂ ʐ		ʝ		ʁ	h
Approximant					ɹ				
Lat. Approximant			l	ɭ					

[†] *Also /ɧ/ (voiceless dorso-palatal/velar fricative) and /ɕ/ (voiceless alveolo-palatal fricative). Table adapted from IPA Chart from International Phonetic Association to include phonemes specified in* Engstrand (1999).
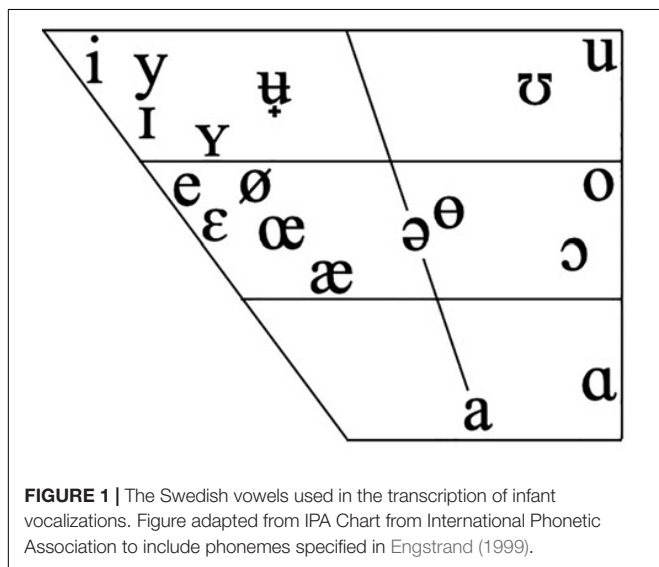

FIGURE 1 | The Swedish vowels used in the transcription of infant vocalizations. Figure adapted from IPA Chart from International Phonetic Association to include phonemes specified in Engstrand (1999).

TABLE 2 | The WCM-SE measure as implemented in the present study, based on Marklund et al. (2018).

Domains	Complexity parameter	N points
Word patterns	>2 syllables	1 per vocalization
	Non-initial stress	1 per vocalization
Syllable structures	Word-final consonant	1 per vocalization
	Consonant cluster[†]	1 per occurrence
Sound classes	Velar consonant [k], [g], [ŋ], [ɧ]	1 per occurrence
	Liquid [l], [ɭ], [ɹ]	1 per occurrence
	Fricative[‡] [f], [v], [s], [ʐ], [ʁ], [ʂ], [j], [h], [ɧ], [ɕ]	1 per occurrence
	Voiced fricative [v], [ʐ], [ʁ], [j]	1 per occurrence
	Trill [r], [ʀ]	3 per occurrence
	Long, front, rounded vowel [y], [ø], [ʉ]	1 per occurrence

Each transcribed vocalization was given a WCM-SE score, after which a subject mean score was calculated.
[†] *Only consonant clusters within syllables were counted.*
[‡] *Vocalization-final [h] excluded, since that is likely to be release of breath.*

it was denoted by a square. Syllable boundaries and primary stress were also marked up in each vocalization. Boundaries between vocalizations were based on silence (pause or breath), and thus not dependent on semantic interpretation. All infant vocalizations consisting of words, syllables, babbling or isolated speech sounds were transcribed. Overlapping or distorted speech, laughter, crying, fuzzing, coughing, effort sounds and vegetative sounds such as breathing, sneezes and hiccups were excluded.

Two recordings were transcribed by both annotators independently to check inter-transcriber agreement. Percentage of matching characters for each transcribed vocalization was compared. Inter-transcriber agreement of which vocalizations were transcribed was 70% and out of those the average transcription inter-transcriber agreement was 78%.

To estimate complexity in infant vocalizations, the WCM-SE was used (Marklund et al., 2018). Based on a number of phonological/phonetic complexity parameters in three domains, a complexity score is calculated for each vocalization based on how many of the parameters are present in it (**Table 2**). For example, the Swedish word *elefant* ("elephant") produced as /eleˈfant/ results in a WCM-SE score of 6, the word *sko* ("shoe") produced as /skuː/ results in a score of 3, and the word *mamma* ("mommy") produced as /màma/

results in a score of 0. WCM-SE points were calculated for each vocalization based on the transcriptions using a script written in R 3.5.0–4.0.2 (R Core Team, 2018). Although the WCM-SE is originally a phonological measure, the principles parameters are based on phonetic complexity (Marklund et al., 2018), so it can also be used for estimation of phonetic complexity.
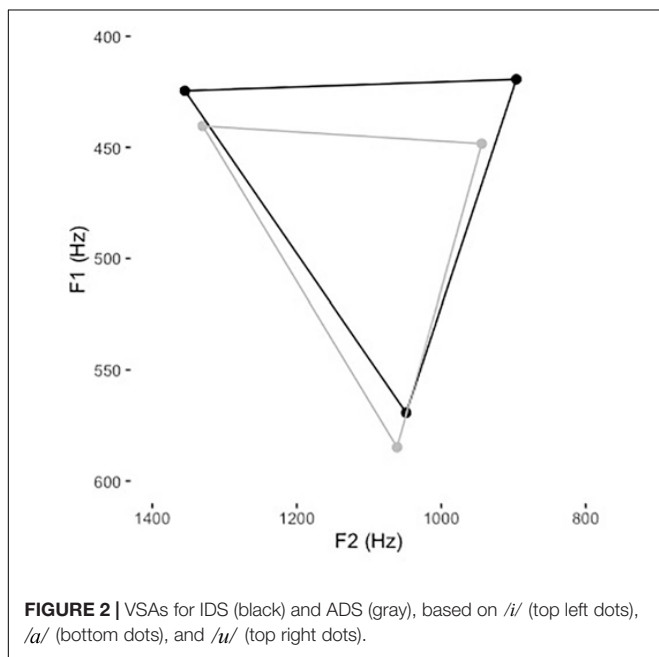
Analyses

In order to establish that IDS vowels were hyperarticulated relative to ADS vowels, parents' VSA_{IDS} and VSA_{ADS} were compared. Second, it was tested whether parents' adaptation in terms of hyperarticulation, described as VSA difference ($VSA_{IDS}-VSA_{ADS}$), predicted infants' average WCM-SE scores. The analysis was performed in R 3.5.0 (R Core Team, 2018).

RESULTS

Formant estimations for point vowels in IDS and ADS can be found in **Table 3** and **Figure 2**.

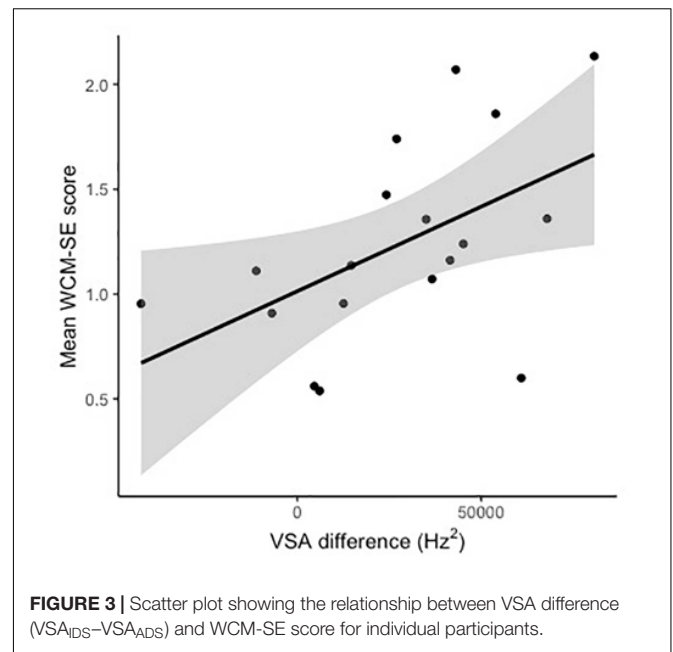**TABLE 3 |** Descriptive statistics of formant estimations of point vowels in IDS and ADS.

Speech sample	Vowel type	Mean F_1 (SD) in Hz	Mean F_2 (SD) in Hz	N included tokens
IDS	i	453 (88)	1,972 (305)	507
	u	430 (90)	934 (195)	223
	ɑ	656 (117)	1,383 (263)	936
ADS	i	472 (81)	1,904 (272)	282
	u	473 (75)	1,032 (229)	119
	ɑ	661 (96)	1,307 (234)	667

FIGURE 2 | VSAs for IDS (black) and ADS (gray), based on /i/ (top left dots), /ɑ/ (bottom dots), and /u/ (top right dots).

A two-tailed paired-samples t-test revealed a significantly larger VSA in IDS than in ADS [$t(18) = 3.87$, $p = 0.001$], with a mean of 101,021 Hz^2 in IDS (SD = 35,831) and 74,304 Hz^2 in ADS (SD = 35,639).

Subjects' mean WCM-SE scores ranged from 0.54 to 2.13 (mean = 1.23, SD = 0.48, after excluding one outlier with a score of 4.69). The mean VSA difference of the included participants was 27,405 Hz^2 (SD = 30,810). A linear regression showed that VSA difference significantly predicted WCM-SE score [$F(1,16) = 5.89$, $\beta = 8.07e\text{-}6$, $R^2 = 0.223$, $p = 0.027$], see **Figure 3**. For a 10000 Hz^2 increase in VSA difference, the model predicts an 0.08 increase in WCM-SE score.

Frequencies are reported in Hz for the results to be comparable with previous similar studies (Liu et al., 2003; García-Sierra et al., 2021), however, the same pattern of results was found when converting the Hz values to Bark (Winkelmann et al., 2021) for each vowel token prior to calculating the VSA [t-test: $t(18) = 3.75$, $p = 0.001$; linear regression: $F(1,16) = 5.89$, $\beta = 0.19$, $R^2 = 0.228$, $p = 0.045$].

FIGURE 3 | Scatter plot showing the relationship between VSA difference (VSA$_{IDS}$–VSA$_{ADS}$) and WCM-SE score for individual participants.

DISCUSSION

Swedish IDS to 12-month-old infants show more VH than ADS, in terms of VSA. Difference in hyperarticulation between IDS and ADS, that is, parents' degree of articulatory adaptation in terms of hyperarticulation, correlates positively with phonetic complexity of infant vocalizations.

Previous studies have reported VH in IDS to infants around 1 year of age, both in Swedish (Marklund and Gustavsson, 2020)[2] and in IDS of several other languages (e.g., Liu et al., 2003; Kondaurova et al., 2012; Kalashnikova et al., 2017, 2018; Tang et al., 2017; Kalashnikova and Burnham, 2018). However, IDS to Dutch-learning 11-month-old instead report hypoarticulation in IDS compared to in ADS (Benders, 2013), and in some cases, no difference was found between IDS and ADS (Dodane and Al-Tamimi, 2007; Kondaurova et al., 2012; Xu et al., 2013; Burnham et al., 2015). These disparate findings suggest that the occurrence of VH in IDS might be dependent on context more fine-grained than language and infant age. For example, in conversations between adults, more hyperarticulation is found in rare words than in more common words (Munson and Solomon, 2004). This means that parents potentially adapt their articulation to different degrees for different words, based on expectations of whether or not their child understands them.

Positive relationships between VH and different language skills and outcomes have previously been reported, specifically speech sound category development (Liu et al., 2003; García-Sierra et al., 2021) and vocabulary (Hartman et al., 2017; Kalashnikova and Burnham, 2018). The present findings add phonetic complexity of vocalizations to the list. Just like the previous studies, the present study does not provide information

---

[2]Note, however, that this is the same group of parent-infant dyads as in the present study.

about causality. Although one study has shown that it is possible for parent articulatory behavior to impact infant vocalizations (Goldstein and Schwade, 2008), it is equally possible for parent VH to be elicited by interactional cues from the infant. When mothers interact with their infant without the infant being able to hear them, they do not hyperarticulate their vowels, indicating that their articulatory adaptation is a response to in-the-moment cues from the infant (Lam and Kitamura, 2012). It might also be the case that there is no causal relationship between parent VH and the various language skills and outcomes of the child studied so far, but that they are instead impacted by other, underlying, variables. One possible such underlying variable is general communicative adaptiveness of the parent. Articulatory adaptation can be considered a specific type of realization of general communicative adaptiveness. Other types of communicative adaptive behaviors include temporal and conceptual contingency as well as prosodic entrainment. One broad concept that captures communicative adaptation, parent responsiveness (either in terms of a general sensitivity to the infant's communicative needs or in terms of contingent responses), has previously been linked to speech sound category development (Elsabbagh et al., 2013), vocabulary development (Tamis-LeMonda et al., 2014; Marklund et al., 2015), and maturity of infant vocalizations (Bloom et al., 1987; Goldstein et al., 2003). All of these behaviors, VH, parent responsiveness, a propensity to provide temporally and conceptually contingent responses, as well as prosodic entrainment, may well be different realizations of the same underlying communicative adaptiveness, which could be beneficial to infant language development and/or an inherited trait.

Limitations in this study include a relatively small sample size, although in line with previous similar studies (Liu et al., 2003; Hartman et al., 2017; Kalashnikova and Burnham, 2018; García-Sierra et al., 2021). Further, the WCM-SE measure is originally intended for assessment of phonological rather than phonetic complexity, and the infants in this study are too young to be expected to have a phonological system fully in place. However, the WCM-SE is based on phonetic principles, and can therefore be used independently of phonological development. The rationale for using WCM-SE specifically is that previous findings have shown that the phonetic content of parents' utterances impacts the phonetic content of infant vocalizations (Goldstein and Schwade, 2008), and WCM-SE is more informative about the infants' articulation than other, less detailed estimates of infant production (number of CV syllables, etc.).

The present study contributes to the field of language development by discussing VH in IDS in the context of broader phonetic and speech communication principles. Rather than viewing VH as a characteristic specific to IDS, it is here treated as part of speech communication in general, and this perspective is applied to the discussion on how it relates to early language development. From this perspective, future studies on VH in IDS should focus on variations in the more fine-grained linguistic context, in order to provide clarity on why it is realized the way it is in IDS compared to in ADS. For example, it is reasonable to assume that more novel words are introduced in interaction with an infant than in conversation with another adult, and so a

higher degree of VH in IDS than in ADS may be a side effect of this contextual difference rather than vowels overall being more hyperarticulated in IDS. As for the potential impact of VH on different aspects of language development, experimental studies in which it can be separated from other adaptive behaviors should be conducted. This study is followed up by investigating the relationship between infant vocalization complexity and parent VH on the level of conversational turns, rather than on a subject level, in order to be able to provide some information about potential causality (Marklund et al., accepted).

In conclusion, the phonetic complexity of 12-month-old's vocalizations is correlated with the degree of articulatory adaptation toward hyperarticulation in their parents' IDS. The findings are in line with previous research in terms of a positive relationship between VH in IDS and infant language skills, as well as in that they do not provide information about causality of the relationship.

DATA AVAILABILITY STATEMENT

Tabular data generated for this study are available at the Open Science Framework (https://osf.io/rc8v4/).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Regional Ethics Review Board in Stockholm (2015/63-31). Written informed consent to participate in this study was provided by the adult participants and the infant participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Benders, T. (2013). Mommy is only happy! Dutch mothers' realisation of speech sounds in infant-directed speech expresses emotion, not didactic intent. *Infant Behav. Dev.* 36, 847–862. doi: 10.1016/j.infbeh.2013.09.001

Benders, T., Pokharel, S., and Demuth, K. (2019). Hypo-articulation of the four-way voicing contrast in Nepali infant-directed speech. *Lang. Learn. Dev.* 15, 232–254. doi: 10.1080/15475441.2019.1577139

Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *J. Child Lang.* 11, 557–578. doi: 10.1017/s030500090000595x

Bloom, K., Russell, A., and Wassenberg, K. (1987). Turn taking affects the quality of infant vocalizations. *J. Child Lang.* 14, 211–227. doi: 10.1017/s0305000900012897

Boersma, P., and Weenink, D. (2018). *Praat: Doing phonetics by computer [Computer program]. Version 6.0.37 (February 3, 2018) to 6.0.40 (May 11, 2018).* Available online at: http://www.praat.org/ (accessed May 11, 2018).

Bosseler, A. N., Teinonen, T., Tervaniemi, M., and Huotilainen, M. (2016). Infant directed speech enhances statistical learning in newborn infants: An ERP study. *PLoS One* 11:e0162177. doi: 10.1371/journal.pone.0162177

Burnham, E. B., Wieland, E. A., Kondaurova, M. V., McAuley, J. D., Bergeson, T. R., and Dilley, L. C. (2015). Phonetic modification of vowel space in storybook speech to infants up to 2 years of age. *J. Speech Lang. Hear. Res.* 58, 241–253. doi: 10.1044/2015_jslhr-s-13-0205

Clopper, C. G., and Pierrehumbert, J. B. (2008). Effects of semantic predictability and regional dialect on vowel space reduction. *J. Acoust. Soc. Am.* 124, 1682–1688. doi: 10.1121/1.2953322

Cristià, A. (2010). Phonetic enhancement of sibilants in infant-directed speech. *J. Acoust. Soc. Am.* 128, 424–434. doi: 10.1121/1.3436529

Cristia, A., and Seidl, A. (2014). The hyperarticulation hypothesis of infant-directed speech. *J. Child Lang.* 41, 913–934. doi: 10.1017/s0305000912000669

Dodane, C., and Al-Tamimi, J. (2007). "An acoustic comparison of vowel systems in adult-directed speech and child-directed speech: evidence from French, English and Japanese," in *Proceedings of the 16th International Congress of Phonetics Sciences*, (Saarbrücken), 6–10.

Dunst, C., Gorman, E., and Hamby, D. (2012). Preference for infant-directed speech in preverbal young children. *Early Lit. Learn.* 5, 1–13.

Elsabbagh, M., Hohenberger, A., Campos, R., Van Herwegen, J., Serres, J., de Schonen, S., et al. (2013). Narrowing perceptual sensitivity to the native language in infancy: Exogenous influences on developmental timing. *Behav. Sci.* 3, 120–132. doi: 10.3390/bs3010120

Englund, K. T. (2018). Hypoarticulation in infant-directed speech. *Appl. Psycholing.* 39, 67–87. doi: 10.1017/s0142716417000480

Englund, K., and Behne, D. (2006). Changes in infant directed speech in the first six months. *Infant Child Dev.* 15, 139–160. doi: 10.1002/icd.445

Englund, K. T., and Behne, D. M. (2005). Infant directed speech in natural interaction Norwegian vowel quantity and quality. *J. Psycholing. Res.* 34, 259–280. doi: 10.1007/s10936-005-3640-7

Engstrand, O. (1999). *Swedish, in Handbook of the International Phonetic Association: A Guide to the usage of the International Phonetic Alphabet.* Cambridge, MA: Cambridge University Press, 140–142.

Escudero, P., Boersma, P., Rauber, A. S., and Bion, R. A. (2009). A cross-dialect acoustic description of vowels: Brazilian and European Portuguese. *J. Acoust. Soc. Am.* 126, 1379–1393. doi: 10.1121/1.3180321

Foursha-Stevenson, C., Schembri, T., Nicoladis, E., and Eriksen, C. (2017). The influence of child-directed speech on word learning and comprehension. *J. Psycholing. Res.* 46, 329–343. doi: 10.1007/s10936-016-9441-3

García-Sierra, A., Ramírez-Esparza, N., Wig, N., and Robertson, D. (2021). Language learning as a function of infant directed speech (IDS) in Spanish: Testing neural commitment using the positive-MMR. *Brain Lang.* 212:104890. doi: 10.1016/j.bandl.2020.104890

Gerholm, T. (2018). Conventions for annotation and transcription of the MINT-project: Modulating child language acquisition through parent-child interaction, MAW:2011.0070. Available online at: http://su.diva-portal.org/smash/get/diva2:1204492/FULLTEXT04.pdf (accessed March 3, 2021).

Goldstein, M. H., King, A. P., and West, M. J. (2003). Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc. Natl. Acad. Sci.* 100, 8030–8035. doi: 10.1073/pnas.1332441100

Goldstein, M. H., and Schwade, J. A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.* 19, 515–523. doi: 10.1111/j.1467-9280.2008.02117.x

Goldstein, M. H., Schwade, J. A., and Bornstein, M. H. (2009). The value of vocalizing: Five−month−old infants associate their own noncry vocalizations with responses from caregivers. *Child Dev.* 80, 636–644. doi: 10.1111/j.1467-8624.2009.01287.x

Golinkoff, R. M., Can, D. D., Soderstrom, M., and Hirsh-Pasek, K. (2015). (Baby) talk to me: the social context of infant-directed speech and its effects on early language acquisition. *Curr. Dir. Psychol. Sci.* 24, 339–344. doi: 10.1177/0963721415595345

Hartman, K. M., Ratner, N. B., and Newman, R. S. (2017). Infant- directed speech (IDS) vowel clarity and child language outcomes. *J. Child Lang.* 44, 1140–1162. doi: 10.1017/s0305000916000520

Kalashnikova, M., and Burnham, D. (2018). Infant-directed speech from seven to nineteen months has similar acoustic properties but different functions. *J. Child Lang.* 45, 1035–1053. doi: 10.1017/s0305000917000629

Kalashnikova, M., Carignan, C., and Burnham, D. (2017). The origins of babytalk: Smiling, teaching or social convergence? *R. Soc. Open Sci.* 4:170306. doi: 10.1098/rsos.170306

Kalashnikova, M., Goswami, U., and Burnham, D. (2018). Mothers speak differently to infants at-risk for dyslexia. *Dev. Sci.* 21:e12487.

Kisler, T., Reichel, U., and Schiel, F. (2017). Multilingual processing of speech via web services. *Comput. Speech Lang.* 45, 326–347. doi: 10.1016/j.csl.2017.01.005

Kondaurova, M. V., Bergeson, T. R., and Dilley, L. C. (2012). Effects of deafness on acoustic characteristics of American English tense/lax vowels in maternal speech to infants. *J. Acoust. Soc. Am.* 132, 1039–1049. doi: 10.1121/1.4728169

Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., et al. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* 277, 684–686. doi: 10.1126/science.277.5326.684

Lam, C., and Kitamura, C. (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Dev. Sci.* 15, 212–221. doi: 10.1111/j.1467-7687.2011.01118.x

Lindblom, B. (1962). "Accuracy and limitations of sona-graph measurements," in *Proceedings of the 4th International Congress of Phonetic Sciences*, (The Hague: Mouton), 188–202.

Lindblom, B. (1990). "Explaining phonetic variation: A sketch of the H&H theory," in *Speech production and speech modelling*, eds W. J, Hardcastle, A, Marchal (Dordrecht: Springer), 403–439. doi: 10.1007/978-94-009-2037-8_16

Liu, H. M., Kuhl, P. K., and Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Dev. Sci.* 6, F1–F10.

Ma, W., Golinkoff, R. M., Houston, D. M., and Hirsh-Pasek, K. (2011). Word learning in infant-and adult-directed speech. *Lang. Lear. Dev.* 7, 185–201. doi: 10.1080/15475441.2011.579839

Marklund, E., and Gustavsson, L. (2020). The dynamics of vowel hypo-and hyperarticulation in Swedish infant-directed speech to 12-month-olds. *Front. Commun.* 5:79. doi: 10.3389/fcomm.2020.523768

Marklund, U., Marklund, E., and Gustavsson, L. (accepted). Relationship between parents' vowel hyperarticulation in infant-directed speech and infants' phonetic complexity on the level of conversational turns. *Front. Psychol.* 688242. doi: 10.3389/fpsyg.2021.688242

Marklund, U., Marklund, E., Lacerda, F., and Schwarz, I.-C. (2015). Pause and utterance duration in child-directed speech in relation to child vocabulary size. *J. Child Lang.* 42, 1158–1171. doi: 10.1017/s0305000914000609

Marklund, U., Marklund, E., Schwarz, I.-C., and Lacerda, F. (2018). Introducing WCM-SE: The word complexity measure phonetically justified and adapted to Swedish. *Clin. Linguist. Phon.* 32, 1042–1053. doi: 10.1080/02699206.2018.1493620

McMurray, B., Kovack-Lesh, K. A., Goodwin, D., and McEchron, W. (2013). Infant directed speech and the development of speech perception: Enhancing development or an unintended consequence? *Cognition* 129, 362–378. doi: 10.1016/j.cognition.2013.07.015

Monsen, R. B., and Engebretson, A. M. (1983). The accuracy of formant frequency measurements: A comparison of spectrographic analysis and linear prediction. *J. Speech Lang. Hearing Res.* 26, 89–97. doi: 10.1044/jshr.2601.89

Munson, B., and Solomon, N. P. (2004). The effect of phonological neighborhood density on vowel articulation. *J. Speech Lang. Hear. Res.* 47, 1048–1058. doi: 10.1044/1092-4388(2004/078)

Core Team, R. (2018). *R: A language and environment for statistical computing. R Foundation for Statistical Computing.* Vienna: R Core Team.

Ramírez-Esparza, N., García-Sierra, A., and Kuhl, P. K. (2014). Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Dev. Sci.* 17, 880–891. doi: 10.1111/desc.12172

Schiel, F. (1999). "Automatic phonetic transcription of non-prompted speech," in *Proceedings of the 14th International Congress on Phonetic Sciences*, (Berkley,CA: University of California), 106–110.

Šimko, J., Beòuš, Š, and Vainio, M. (2016). Hyperarticulation in Lombard speech: Global coordination of the jaw, lips and the tongue. *J. Acoust. Soc. Am.* 139, 151–162. doi: 10.1121/1.4939495

Singh, L., Nestor, S., Parikh, C., and Yull, A. (2009). Influences of infant-directed speech on early word recognition. *Infancy* 14, 654–666. doi: 10.1080/15250000903263973

Sloetjes, H., and Wittenburg, P. (2008). "Annotation by category – ELAN and ISO DCR," in *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, (Marrakech: European Language Resources Association (ELRA)).

Soderstrom, M. (2007). Beyond babytalk: Re-evaluating the nature and content of speech input to preverbal infants. *Dev. Rev.* 27, 501–532. doi: 10.1016/j.dr.2007.06.002

Song, J. Y., Demuth, K., and Morgan, J. (2010). Effects of the acoustic properties of infant-directed speech on infant word recognition. *J. Acoust. Soc. Am.* 128, 389–400. doi: 10.1121/1.3419786

Spinelli, M., Fasolo, M., and Mesman, J. (2017). Does prosody make the difference?. *Dev. Rev.* 44, 1–18.

Sundberg, U., and Lacerda, F. (1999). Voice onset time in speech to infants and adults. *Phonetica* 56, 186–199. doi: 10.1159/000028450

Tamis-LeMonda, C. S., Kuchirko, Y., and Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Curr. Dir. Psych. Sci.* 23, 121–126. doi: 10.1177/0963721414522813

Tang, P., Xu Rattanasone, N., Yuen, I., and Demuth, K. (2017). Phonetic enhancement of Mandarin vowels and tones: Infant-directed speech and Lombard speech. *J. Acoust. Soc. Am.* 142, 493–503. doi: 10.1121/1.4995998

Trainor, L. J., and Desjardins, R. N. (2002). Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. *Psychon. Bull. Rev.* 9, 335–340. doi: 10.3758/bf03196290

Winkelmann, R., Jaensch, K., Cassidy, S., and Harrington, J. (2021). *emuR: Main Package of the EMU Speech Database Management System. R package version 2.2.0 (March 30, 2021).* Available online at: https://cran.r-project.org/web/packages/emuR/index.html (accessed June 9, 2021).

Wong, J. (2014). "The Effects of High and Low Variability Phonetic Training on the Perception and Production of English Vowels /e/-/æ/ by Cantonese ESL Learners with High and Low L2 Proficiency Levels," in *Proceedings of the 15 th Annual Conference of the International Speech Communication Association*, (Singapore: International Speech Communication Association), 524–528.

Xu, N., Burnham, D., Kitamura, C., and Vollmer-Conna, U. (2013). Vowel hyperarticulation in parrot-, dog- and infant-directed speech. *Anthrozoös* 26, 373–380. doi: 10.2752/175303713x13697429463592