

Rikke L. Bundgaard-Nielsen*, Carmel O'Shannessy, Yizhou Wang,
Alice Nelson, Jessie Bartlett and Vanessa Davis

Two-part vowel modifications in Child Directed Speech in Warlpiri may enhance child attention to speech and scaffold noun acquisition

<https://doi.org/10.1515/phon-2022-0039>

Received October 28, 2022; accepted May 8, 2023; published online June 14, 2023

Abstract: Study 1 compared vowels in Child Directed Speech (CDS; child ages 25–46 months) to vowels in Adult Directed Speech (ADS) in natural conversation in the Australian Indigenous language Warlpiri, which has three vowels (/i/, /a/, /u/). Study 2 compared the vowels of the child interlocutors from Study 1 to caregiver ADS and CDS. Study 1 indicates that Warlpiri CDS vowels are characterised by fronting, /a/-lowering, f_0 -raising, and increased duration, but not vowel space expansion. Vowels in CDS nouns, however, show increased between-contrast differentiation and reduced within-contrast variation, similar to what has been reported for other languages. We argue that this two-part CDS modification process serves a dual purpose: Vowel space shifting induces IDS/CDS that sounds more child-like, which may enhance child attention to speech, while increased between-contrast differentiation and reduced within-contrast variation in nouns may serve didactic purposes by providing high-quality information about lexical specifications. Study 2 indicates that Warlpiri CDS vowels are more like child vowels, providing indirect evidence that aspects of CDS may serve non-linguistic purposes simultaneously with other aspects serving linguistic-didactic purposes. The studies have novel implications for the way

***Corresponding author: Rikke L. Bundgaard-Nielsen**, MARCS Institute of Brain, Behaviour and Development, Western Sydney University, Westmead, NSW, Australia; and School of Literature, Languages and Linguistics, Australian National University, Canberra, ACT, Australia,
E-mail: rikkellou@gmail.com

Carmel O'Shannessy and Yizhou Wang, School of Literature, Languages and Linguistics, Australian National University, Canberra, ACT, Australia, E-mail: carmel.o'shannessy@anu.edu.au (C. O'Shannessy), yizhou.wang@anu.edu.au (Y. Wang)

Alice Nelson and Jessie Bartlett, Red Dust Role Models, Alice Springs, NT, Australia,
E-mail: alice@reddust.org.au (A. Nelson), jessie@reddust.org.au (J. Bartlett)

Vanessa Davis, School of Literature, Languages and Linguistics, Australian National University, Canberra, ACT, Australia; and Tangentyere Council Research Hub, Alice Springs, NT, Australia,
E-mail: vanessa.davis@tangentyere.org.au

CDS vowel modifications are considered and highlight the necessity of naturalistic data collection, novel analyses, and typological diversity.

Keywords: ADS; CDS; nouns; vowels; Warlpiri

1 Introduction

The use and characteristics of special Infant and Child Directed Speech registers (IDS; CDS; or in the older literature Baby Talk; Motherese) has been argued to be a supportive strategy adopted by caregivers (mothers, fathers, other caregivers, and older children) to scaffold language acquisition (Fernald and Mazzie 1991; Gleitman et al. 1988; Hirsh-Pasek et al. 1987; Nelson et al. 1989), assisting emotional regulation and socialisation (Fernald 1989, 1992; Werker and McLeod 1989), and regulating infant attention (Fernald and Simon 1984; Papoušek et al. 1990; Stern et al. 1982). Questions about the universality of IDS/CDS have received significant attention in the literature, and cross-linguistic research has demonstrated differences across languages/cultures, and changes in the characteristics of IDS/CDS across development (e.g., Buchan and Jones 2014; Kitamura and Burnham 2003; Liu et al. 2009; Ratner 1984). The present studies add to this literature in two ways. First, we compare the spectral and temporal characteristics of CDS and ADS vowels in the Indigenous Australian language Warlpiri, and secondly, we examine the acoustic characteristics of the vowels produced by the children who participated in the research alongside their caregivers. The results show that Warlpiri-speaking caregivers modify their vowels in two ways: firstly, by systematically shifting the vowel space in the F1 and to some extent F2 dimension, potentially to sound more child-like, and secondly, to ensure that vowels in nouns are as clear as possible, potentially to scaffold and support lexical acquisition. In the literature review that follows, we refer to IDS and CDS relatively interchangeably, reflecting the label of choice in the research that we review. We assume IDS/CDS is a continuous phenomenon of speech style modification to children, changing dynamically in response to child development (with child age often used as a proxy).

1.1 Characteristics of IDS/CDS

Speech to infants and young children (IDS/CDS) is characterised by a slower speech rate, higher fundamental frequency, greater pitch variations (Fernald et al. 1989), longer pauses, often repetitive intonational structures (Katz et al. 1996), and shorter sentences with a more limited lexicon than Adult Directed Speech (ADS) (Cameron-

Faulkner et al. 2003). IDS/CDS is also often characterised by hyperarticulation of speech segments: vowels (and consonants) are often given articulatorily/acoustically extreme realisations, resulting in an expanded articulatory/acoustic vowel space, and increased acoustic/articulatory differentiation (see e.g., Burnham et al. 2002; Kalashnikova et al. 2017; Kitamura and Burnham 2003; Kitamura et al. 2001; Kuhl et al. 1997; Liu et al. 2003, 2009; Werker et al. 2007).

Young infants prefer to listen to IDS over ADS from soon after birth, a phenomenon which has often been argued to be largely carried by differences in pitch height and range between the registers, and slower speaking rate (Cooper and Aslin 1990; Fernald 1985; Fernald and Kuhl 1987; Pegg et al. 1992; Werker and McLeod 1989). Whether this is also a preference that reflects attention to certain linguistic choices is an open question, and the strong preference for IDS pitch has been argued to reflect not so much a preference for the pitch patterns/range in itself, but for the affective salience of IDS (Kitamura and Burnham 1998; Singh et al. 2002). It is possible, though, that the developmental stage of a child plays a role in determining the reasons for this preference. For instance, it is possible that infants' awareness of being a likely recipient of IDS utterances shapes this preference as research shows that by 12 months of age, infants have learned that IDS is a register used for infants, while ADS is used for adults (Soley and Sebastian-Galles 2020).

Substantial other research has, however, demonstrated that, in addition to being a feature of IDS, vowel hyperarticulation is also used in clear speech (Lindblom 1990) and in Foreigner Directed Speech (FDS: Uther et al. 2007), but not Pet Directed Speech (PDS: Burnham et al. 2002), unless the pet is a parrot, an animal with well known ability to copy and produce human speech (Xu et al. 2015). Hyperarticulation is also reduced or absent in speech to infants with hearing or cognitive impairment (Kalashnikova et al. 2016; Lam and Kitamura 2010, 2012). The broader applications of vowel hyperarticulation may suggest that the phenomenon is used as a didactic strategy in communication with individuals or entities who are perceived to be in the process of acquiring a language (or at least capable of some language learning).

Other research has suggested that vowel modifications in IDS may be a product of laryngeal raising in studies where there was no evidence for vowel hyperarticulation but of higher formant values (see e.g., Kalashnikova et al. 2017). This is argued to be the effect of a deliberate shortening of the vocal tract, motivated by caregiver desire to appear less threatening to the infants as higher formant frequencies are associated with smaller (and therefore less threatening) bodies, and low formant frequencies with larger (and potentially more threatening) bodies. Importantly (and as the data reported here shows), it is entirely possible for a parent to raise all their formants to sound smaller (and more child-like) and also hyperarticulate vowels. These are not mutually exclusive or competing processes.

IDS vowel modifications, and in particular vowel hyperarticulation, however, are not uniformly observed in language or for all vowels within a language. Englund and Behne (2005), for instance, found that Norwegian mothers hypoarticulate back vowels in IDS, that is, produce them with a higher F2 than for ADS. This observed hypoarticulation was interpreted by the researchers to make the vowel articulation more visibly accessible to infants than they are in ADS, and rather than representing a goal of target undershoot, it thus indicated enhancement in the visual domain (as opposed to the acoustic), also bringing this finding into line with the interpretation that IDS is didactic.

IDS/CDS vowel hyperarticulation has been especially highlighted as facilitating segmental acquisition (the learning of vowels and consonants; see Liu et al. 2003; García-Sierra et al. 2021 for perception and Marklund et al. 2021a, 2021b for production), and in word-learning (Hartman et al. 2017; Ma et al. 2011; Song et al. 2010). In the first year of life, vowel hyperarticulation has been argued to enhance segmental learning by providing infants with high-quality, articulatorily extreme, and maximally differentiated vowel tokens that infants might attend to preferentially, perhaps due to its prosodic characteristics (Adriaans and Swingley 2017). The effect of IDS in the first year of life also extends into very early word-learning: IDS has been found to increase neural activity in 6- and 13-month-old infants compared to ADS, and this is argued to assist very young word-learners with ‘word spotting’ (Zangl and Mills 2007) by making individual words ‘stand out’ from the speech stream, a function of IDS/CSD that may only increase in the second year of life (see below).

In the second year of life, vowel hyperarticulation is argued to be particularly helpful in terms of word-learning with research showing that the degree of vowel hyperarticulation in the maternal input at 18 months of age is positively correlated with the size of the receptive and productive language at 24 months of age (Hartman et al. 2017). Other research has shown that vowel hyperarticulation helps 21-month-olds learn new words faster unless they already have large vocabularies, while 27-month-old toddlers learn new words in both ADS and IDS/CDS (Ma et al. 2011), and vowel hyperarticulation (but not an enhanced pitch range) plays an important role in word recognition for 19-months-old toddlers (Song et al. 2010). Recent work has also demonstrated that caregivers systematically fine-tune their input to children (also known as the fine-tuning hypothesis; Snow and Ferguson 1977), both in terms of the lexicon (Odijk and Gillis 2021), and in terms of vowel hyperarticulation, with one cleverly designed study showing that caregivers hyperarticulate *exactly those words* that the children were beginning to produce (and not vowels in those same words three months prior or three months after ‘word birth’) (Odijk and Gillis 2022). This latter observation is particularly relevant to the studies presented here and suggests that extra clarity in the phonemic specifications of words in CDS supports the

acquisition of new vocabulary items, as well as the recognition of already known words.

While one recent meta study found that vowel hyperarticulation did not change with age of the infant/child (Cox et al. 2023), other studies have shown that the acoustic characteristics of IDS, including the degree of vowel hyperarticulation, does change across early development (Buchan and Jones 2014; Kitamura and Burnham 2003; Liu et al. 2009). This is generally taken as evidence of IDS fine-tuning to the developmental stage of the child, including its linguistic development, as has also been argued for other linguistic domains (e.g., Huttenlocher et al. 2007). In a cross-sectional design, Ratner (1984) showed that the use of vowel space expansion and vowel target precision in IDS changes from infancy into toddlerhood, such that vowels are more hyperarticulated to early word-learners than prelingual infants. Vowels are also more hyperarticulated to young children who are combining words into short phrases and sentences—in both content words and function words, again suggesting that vowel hyperarticulation past the first year of life is didactic and tied to word-teaching or learning.

Longitudinal studies of IDS/CDS have further demonstrated that segmental modifications, including hyperarticulation, and the avoidance of segmental reduction patterns found in adult speech, change across early development. In one study, mothers were found to hyperarticulate vowels to pre-linguistic children (<12 months of age) much more than to 5-year-old children and to adults (Liu et al. 2009), and another study shows that mothers' use of segmental deletion increases between 1;6 and 2;0 years of age only to decrease again between 2;0 and 2;6 years of age (Buchan and Jones 2014), suggesting deliberate avoidance of deletion in speech to children during segmental acquisition and again in the early multiword stage of language' acquisition. It is plausible that these changes in input characteristics reflect a 'first wave' of hyperarticulation to support vowel learning, and a 'second wave' to support rapid word-learning after the second half of the second year of life.

The research findings reported above, however, are predominantly based on studies undertaken in just a few of the world's languages (Kidd and Garcia 2022), leaving large gaps in our knowledge of cross-linguistic IDS/CDS patterns. And while most studies indicate that caregivers do use special registers with children, and a recently published mega-study of more than 50,000 individuals from 187 countries reliably distinguishes IDS/CDS from ADS across a sample from 21 languages/cultures (none of them Australian) (Hilton et al. 2022), careful study of IDS/CDS in a more diverse set of cultures and languages is certainly called for.

Indeed, our current understanding reflects insights primarily from a group of languages that are, in fact, *atypical*. A survey of the vowel inventories of the 451 languages in the UPSID database (<https://phoible.org/contributors/UPSID#inventories>) reveals, for instance, that the vowel profile of languages like English, French and German

are the odd ones out: According to the UPSID database, the *Mean* number of vowels in the 451 languages included is 8.5 (a number skewed by typologically very unusual languages such as !Xu with 46 vowels), a *Median* number of vowels of seven, and a *Mode* of just five vowels. More than 50 % of the languages in the database have seven or fewer vowels, and 23 of the languages in the database, like Warlpiri, have only three vowels. Languages of 11 or more vowels fall in the fourth quartile.

We need to study typologically diverse languages to fully understand the role, characteristics and prevalence of vowel hyperarticulation. This is particularly so given that some studies have found that caregivers do not hyperarticulate vowels at all to 18-month-old Danish children (Bohn 2013) or to two Nungon children aged 2;2 and 3;2 respectively, in a preliminary study (Sarvasy et al. 2019), though the CDS and ADS compared in the latter study came from different speakers. Likewise, cross-linguistic comparisons have demonstrated differences in the implementation or degree of hyperarticulation and pitch raising across different languages, when hyperarticulation is present (e.g., Kitamura et al. 2001). The prosodic system of the target adult language has also been found to influence the prosodic modifications made in IDS (Igarashi et al. 2013), as do sociolinguistic considerations (Bernstein Ratner and Pye 1984; Pye 1986). This suggests that the ‘shape’ of IDS/CDS in each language is subject to significant variability and reflects linguistic and sociolinguistic demands outside of the caregiver-child interaction and the developmental characteristics of the child.

1.2 IDS/CDS in Warlpiri

Warlpiri is a Pama-Nyungan language spoken by approximately 2,624 people (ABS 2021 Census), mostly in the remote communities of Yuendumu, Lajamanu, Nyirrpi, and Willowra, and in regional towns and cities, including Alice Springs and Darwin, in the Northern Territory of Australia (see Figure 1). Although endangered, Warlpiri continues to be learned by children as their primary language in Yuendumu, Nyirrpi, and Willowra, and as one of their first languages in Lajamanu. The families in this study were recorded while in Alice Springs and Yuendumu.

Phonologically, Warlpiri is characterized by a single series of stops /p t t̪ c k/ with five main places of articulation. This is repeated in the nasal series /m n ŋ j n̪/, and Warlpiri further has three laterals /l ɭ ɭ̪/, two approximants /w j/, and two or three rhotic phonemes; trill /r/ approximant /ɻ/, and a retroflex flap /ɽ/, though the phoneme status of the latter has been recently questioned (Bundgaard-Nielsen and O'Shannessy 2021). In terms of vowels, Warlpiri has only three: /i/, /a/, and /u/, with a phonemic length contrast (a very limited set of words are reported to have long vowels but no long vowels were included in the study here). The Warlpiri formant

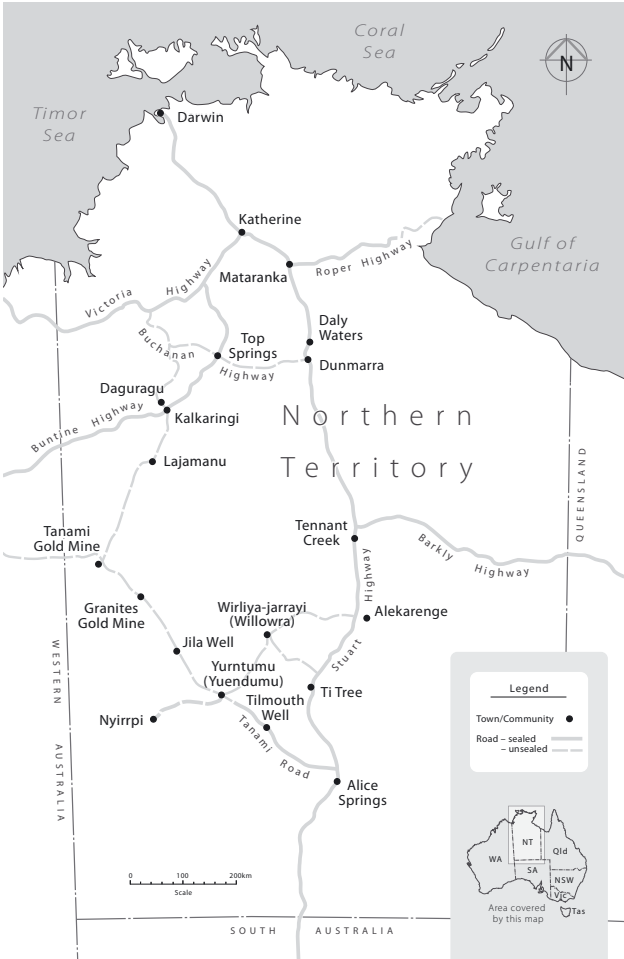


Figure 1: Map of Warlpiri communities, Northern Territory, Australia. © Brenda Thornley 2019.

space is compact, with a relatively compressed F1 range: Butcher (1994) reports the F1/F2 values of a single female Warlpiri speaker, indicating an F1 range of appx. 470–600 Hz (/i/ and /u/ vs /a/), and an F2 range of 1,200–2,400 Hz (/u/ vs /i/). Such a compressed vowel system is typical of Australian Indigenous languages, as they are often characterised by ‘corner vowels’ that are less peripheral than in, for instance, English (see Fletcher and Butcher 2014 for discussion). The three Warlpiri vowels are distributed unevenly in the lexicon according to a vowel count of all entries in PanLex online dictionary (<https://vocab.panlex.org>), with /a/ contributing 45 % of

vowels, /i/ vowels contributing 33 %, and /u/ only 22 %, making /u/ only half as frequent as /a/. This is also typical for many Australian Indigenous languages (Busby 1979; p. 79).

The existing literature on the characteristics of ‘Baby Talk’ in Warlpiri (and the limited literature on IDS/CDS in other Australian languages; see Davidson et al. 2023) has focused on consonants and changes to adult wordforms: Laughren (1984) identifies patterns of substitution of coronal consonants (stops /t t/, nasals /n ŋ/, and laminals /l l/) with the corresponding lamino-palatal consonants (/c ɲ Δ), and the use of special IDS/CDS wordforms, for instance, ‘apa’ for ‘ngapa’ (water). Such segmental modifications are very common cross-linguistically in IDS and in children’s speech development.

1.3 The present studies

In the following, we report first on a study of the acoustic characteristics of vowels in Warlpiri CDS and, secondly on a study (Study 2) of vowels produced by the young Warlpiri-acquiring children whose caregivers participated in the CDS study. The CDS study (Study 1) examines CDS to young Warlpiri-acquiring children by comparing ADS and CDS vowel quality, quantity, and pitch within a group of four adult speakers. The study provides a first comparison of vowel space size and f_0 in CDS versus ADS vowels in a three-vowel system, where pressures for contrast enhancement might differ from those in languages with more crowded vowel inventories, such as English, and vowel systems (such as Japanese) where vowel duration is phonemic (vowel length is contrastive in only a very restricted set of words in Warlpiri: Butcher and Anderson 2008). We further compare the degree of acoustic overlap between vowels in CDS and ADS, addressing the question of whether vowel space expansion is (necessarily) associated with improved acoustic differentiation (less category overlap), as is sometimes assumed to be the case (Werker et al. 2007). This aspect of our study has typological ramifications, given the fact that we examine the acoustic characteristics in a three-vowel system, which is subject to different constraints in terms of contrast maintenance than the larger vowel systems found in most of the world’s other languages, and indeed, in the predominantly European languages that have been the focus of IDS/CDS research to date. Finally, we examine the characteristics of vowels in CDS nouns relative to vowels from other parts of speech in CDS, testing the hypothesis that CDS is (at least in part) a didactic strategy for scaffolding vocabulary development in young word-learners. This is consistent with the observations of Odijk and Gillis (2021, 2022) that adult noun usage peaks at the time of ‘word birth’ (i.e., productive acquisition). This aspect of our study is also of methodological import given that lab-based IDS/CDS vowel studies tend to make use of

more restricted materials (typically nouns labelling toys provided in the laboratory setting) than the general CDS material examined here. By taking this methodological approach, our study allows for a better understanding of whether the typically observed vowel hyperarticulation is particular to the materials included (nouns).

This comparison provides an opportunity to test the idea that IDS involves phonetic convergence, a hypothesis raised by Polka and Ruan (2021), and Kalashnikova et al. (2017) by comparing the acoustic characteristics of ADS and CDS vowel production to the vowel productions of the children who are the addressees of the CDS. This comparison tests the hypothesis that CDS modifications exploit child listening preferences to voices with similar F1 and F2 acoustic properties as their own (Polka et al. 2022), in addition to reflecting a preference for higher f_0 (pitch), a modification that has been argued to serve the purpose of emotional and social regulation, rather than linguistic scaffolding.

2 Study 1: Warlpiri CDS

Study 1 examines the acoustic properties of vowels in Warlpiri CDS and ADS and, to our knowledge, provides the first systematic examination of the characteristics of CDS vowels in an Indigenous Australian language, as well as in any language with a three-vowel system.

Our study also uses the CDS data to investigate the hypothesis that CDS is a dynamic didactic modification, suited to the developmental needs of an infant/child, and we consequently separated the CDS data into two word classes: CDS Nouns (all concrete), and CDS Non-Nouns, which consisted of a large proportion of verbs (often imperatives), adjectives (often colours, sizes; these words are syntactically nominals in Warlpiri, with the semantics of attributes), pronouns and demonstratives ('that one', 'this one'). In the present study, three of the children were in the early multi-word stage of language acquisition, and one child a little older, and we hypothesise that caregiver CDS modifications may (among other things) be particularly targeted to scaffold word learning, and perhaps in particular to the teaching of 'names for things', in line with the persuasive results from Odijk and Gillis (2021, 2022). We also base this speculation on the supportive effect that CDS has been demonstrated to have on toddlers' word learning and word recognition (Hartman et al. 2017; Ma et al. 2011; Song et al. 2010; Zangl and Mills 2007), and the typical developmental stage of children around the age of 24 months, who are becoming expert word-learners. If that is the case, we would expect vowel hyperarticulation and temporal expansion to be at least partially driven by the hyperarticulated characteristics of nouns rather than reflecting hyperarticulation in CDS in general.

2.1 Methods

2.1.1 Participants

We report on data from three Warlpiri-speaking women (RA03, RA06, RA07) and one Warlpiri-speaking man (RA05). At the time of the recordings, RA03 was a grandmother in her 50s, and RA06 was a mother of appx. 30 years of age, RA07 was a 60+ grandmother, and RA05 (the male participant) was a 50+ years old grandfather.¹ The adult Warlpiri participants were recorded interacting with four Warlpiri-acquiring children: RA03, RA06 and RA07 interacted with various constellations of *RC01 (boy; 28 months)*, *RC11 (boy; 30 months)*, and *RC15 (boy; 25 months)*, while RA05 primarily interacted with *RC10 (girl; 46 months)*. These children are included in the second study (Section 3).

All four participants were L1 speakers of Warlpiri, and additionally spoke one or more varieties of English. One caregiver also spoke Yankunytjatjara and Western Arrarnta, and one understood Western Arrarnta. The participants (and the children in Study 2) were recruited by two of the authors through personal relationships, as part of the Little Kids Learning Languages project.² Adult caregivers completed an informed consent process and were compensated with a \$AUD50 supermarket voucher after each recording session. All participants were recorded (video, audio) at their homes in either Alice Springs or Yuendumu, in the Northern Territory (NT) of Australia.

2.1.2 Materials

The three female participants were recorded interacting with each other and other female community members, and three of the young children (RC01, RC11, and RC15) in various social constellations in three different recording sessions. During the sessions, the women and children took part in play, story-telling (from picture book stimuli; O'Shannessy 2004) and discussion activities, centred on day-to-day child-rearing activities and interactions with the children, who ranged in age from 25 to 30 months. The male participant (RA05) was likewise recorded while taking part in a story-telling activity with the fourth young child (RC10; 46 months) in the company of

¹ Child rearing in Indigenous communities in Australia is often communal, with grandmothers and aunties (biological or classificatory) in particular, playing significant roles, and sometimes taking the role of main caregiver for extended periods.

² ARC Future Fellowship project #FT190100243, awarded to Dr. Carmel O'Shannessy: <https://little-kids-learning-languages.net/>. In this project the authors are Aboriginal and non-Aboriginal researchers. Collaboration of this kind is an important aspect of working with languages spoken in Aboriginal families in Australia.

the child's mother and other relatives. The recording sessions differ from the typical one-on-one setup in IDS research, as multiple adults were present and the children were relatively free to come and go as they pleased, and other children likewise free to join the play, in accordance with the cultural practices of the community. Non-participating adults also occasionally joined the groups, but interactions with 'visiting' children and adults were not included in the study.

Except for RA05, a Sennheiser EW112 PG4-GB Portable Wireless Lapel Microphone System and Zoom Q8 camera with a microphone extension were used to record. For RA05, only a Sony 4K FDR-AX33 handycam camera was used. Participants sat on a blanket or an outdoor bench in their yard. During each recording session, the camera was placed near the participants so that all of them were visible, the lapel microphone was positioned on the collar of an adult's clothing, and the Zoom Q8 extended microphone was placed near the participants. In addition, RA06 wore a Rode lavalier microphone connected to a H2 zoom recorder.

The video/audio recordings of the four adult participants (RA03, RA05, RA06, RA07) were transcribed, glossed, and translated into English, and coded for the intended addressee (CDS or ADS on the basis of close viewing of the video recordings) in ELAN 6.3. We further coded all CDS vowels from concrete nouns in the CDS of the four speakers, creating two subsets of CDS vowel data: Vowels from concrete nouns (CDS Nouns) and vowels from everywhere else (CDS Non-Nouns). No abstract nouns occurred in the dataset.

Target vowels were then hand-segmented and labelled in *Praat* 6.2.12. Vowels degraded by environmental noise, overlapping talkers, etc., were excluded from the dataset, and vowel duration, f_o , F1, and F2 extracted using an automatic script (Kroos et al. 2010). Unusual and missing values were checked by hand. In all individual datasets, /a/ was much more frequent than /i/ and /u/, consistent with the general distribution in Warlpiri. We extracted target measurements from as many CDS vowels as possible from each of the female participants, as well as a roughly matching number of ADS vowels. The male participant provided 247 CDS vowels in his recording but had only one brief ADS interaction during the recording session, providing just nine /i/, /a/, and /u/ tokens. In total, we extracted 1,599 vowel tokens from the four speakers. The male voice was not treated differently to the female voices in our analyses as the male speaker's f_o , F1 and F2 values were intermediate to the women's ranges. We acknowledge that this is not standard practice (separating male and female speakers is typically done), but we do not find it necessary for the present data set. We did not normalise the data. The number of vowels from each participant is presented in Appendix 1.

2.2 Warlpiri CDS results

2.2.1 Warlpiri vowels in ADS and CDS

The combined data points from each of the four speakers (RA03, RA05, RA06, RA07) by speech style (ADS vs CDS) are presented in Figure 2 as scatter plots and estimated (/i/, /a/, /u/) vowel ellipses, the latter with superimposed vowel triangles defined by the /i/, /a/, and /u/ vowel centroids. Individual descriptive statistics are found in Appendix 2. Mean values are presented in Table 3: these values are relatively consistent with the values reported by Butcher for a single speaker of Warlpiri (Butcher 1994).

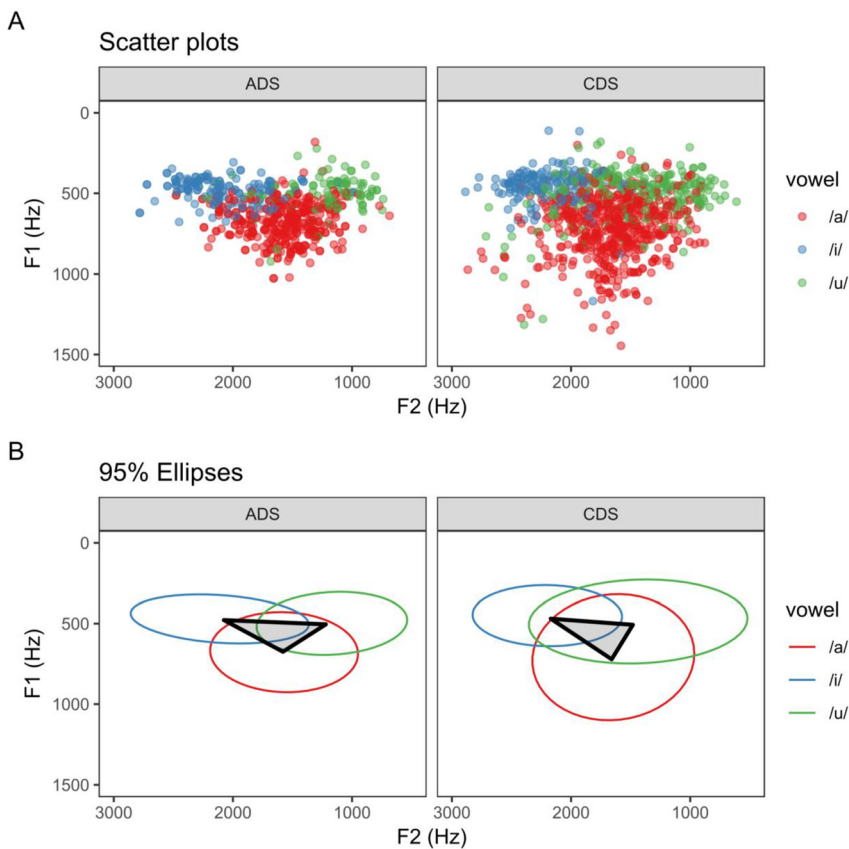


Figure 2: Spectral characteristics of Warlpiri vowels in adult-directed speech (ADS) and child-directed speech (CDS). (A) Scatter plots of individual tokens. (B) Estimated vowel ellipses from the same dataset. Black triangles represent the centroids of the ellipses, i.e., mean values of F1/F2.

In order to test whether the four Warlpiri speakers produce CDS vowels differing from ADS in the spectral quality (F1 and F2), and f_o (fundamental frequency), we first built a linear mixed-effects model (LMM, fixed effects: Vowel [i/, /a/, /u/] and Speech Style [ADS vs CDS]; random effect: Speaker), see Table 1. We checked the model using a Wald Chi-squared test, which (unsurprisingly) revealed a significant main effect of Vowel in terms of F1, F2, and in Duration. We also observed a significant main effect of Speech Style in terms of F1, F2, Duration and f_o , and a significant Vowel $\times$ Speech Style interaction in terms of F1, F2 and f_o . We further analysed the effect of Vowel and Speech Style by carrying out a series of *post-hoc* tests (see Table 2). The summarised results in Table 3 indicate clearly that Warlpiri CDS is characterised by higher F2 values (vowel fronting) for /i/, /a/ and /u/, while CDS /a/ also has a higher F1 (indicating vowel lowering). The results also indicate that CDS is characterised by longer vowel durations (12 ms on average) and higher f_o (21 Hz on average) than ADS.

Given consistent reports of vowel space expansion in CDS (e.g., Burnham et al. 2002; Kalashnikova et al. 2017; Kitamura and Burnham 2003; Kitamura et al. 2001; Werker et al. 2007), we also compared the vowel spaces denoted by /i/, /a/, and /u/ in Warlpiri ADS and CDS (see Table 4). As is clear from Table 4, and as is visually appreciable from Figure 2 (and for the speakers individually in Appendix 2), the space denoted by the vowel centroids in /i/, /a/, and /u/ in Warlpiri ADS and CDS are of almost identical size (ADS = 77,783 vs CDS = 77,105 Hz squared; see Table 4). While this result differs from previous work identifying more extreme realisations of (corner) vowel centroids as a feature of CDS, it is also clear from the results presented in Tables 3 and 4, along with visual inspection of Figure 2 and Appendix 1, that similar vowel spaces in ADS and CDS does not necessarily suggest that there is no modification of the *shape* of the CDS vowel space in Warlpiri, even though the size of the vowel triangles is similar. In fact, the results here indicate that CDS modifications in Warlpiri are best described as system shifting through fronting and lowering, resulting in a system with more dispersed vowel centroids and a larger vowel space; these changes are indicative of a smaller vocal tract instead of hyperarticulation.

Table 1: Effect of vowel, style, and their interaction on acoustic measures in Warlpiri, based on linear mixed-effects modelling (LMM).

Measure	Vowel		Style		Vowel $\times$ style	
	Wald χ^2	<i>p</i> Value	Wald χ^2	<i>p</i> Value	Wald χ^2	<i>p</i> Value
F1	776.7 ^a	<0.0001	19.4 ^a	<0.0001	11.5 ^b	0.0032
F2	889.6 ^a	<0.0001	54.2 ^a	<0.0001	16.1 ^a	0.0003
Duration	11.0 ^b	0.0040	11.3 ^a	0.0008	4.4	0.1134
f_o	0.2	0.8956	46.1 ^a	<0.0001	6.2 ^c	0.0452

^a $p < 0.001$, ^b $p < 0.01$, ^c $p < 0.05$.

Table 2: Acoustic characteristics of Warlpiri vowels in adult- and child-directed speech.

Warlpiri vowel	ADS			CDS			Diff.	t Value	p Value
	Mean	SD	N	Mean	SD	N			
F1 (Hz)									
/a/	675	116	378	721	195	537	46	5.536 ^a	<0.0001
/i/	478	72	155	469	142	164	−9	0.069	0.9450
/u/	504	103	109	507	145	256	3	0.531	0.5957
F2 (Hz)									
/a/	1,580	290	378	1,659	326	537	79	4.102 ^a	<0.0001
/i/	2,075	355	155	2,170	303	164	95	2.823 ^b	0.0048
/u/	1,220	311	109	1,481	432	256	261	6.977 ^a	<0.0001
Dur. (ms)									
/a/	83	48	378	89	59	537	6	1.196	0.2321
/i/	89	60	155	112	168	164	23	2.679 ^b	0.0075
/u/	70	44	109	93	77	256	23	2.597 ^b	0.0095
Combined	82	51	642	94	92	957	12	3.833 ^a	0.0001
f₀ (Hz)									
/a/	198	38	150	213	53	149	15	3.058 ^b	0.0023
/i/	203	45	90	209	48	229	6	1.051	0.2933
/u/	198	40	587	214	47	859	16	5.057 ^a	<0.0001
Combined	196	39	347	217	44	481	21	6.597 ^a	<0.0001

^a $p < 0.001$, ^b $p < 0.05$.

Table 3: Summary of CDS features as compared to ADS in Warlpiri.

Warlpiri vowels	Lowering	Fronting	Lengthening	Higher f_0
/a/	✓	✓		✓
/i/		✓	✓	
/u/		✓	✓	✓

Table 4: Acoustic area (in Hz-squared) of Warlpiri vowels in adult- and child-directed speech.

Vowel pair	Area: ADS	Area: CDS	Diff.	CDS/ADS
/a/ ellipsis	485,347	845,927	360,580	174 %
/i/ ellipsis	347,466	384,450	36,984	111 %
/u/ ellipsis	386,034	748,686	362,652	194 %
Vowel triangle	77,783	77,105	−678	99 %

An exclusive focus on the Euclidian space (vowel triangles) denoted by vowel centroids in CDS versus ADS may also result in a failure to capture CDS modifications pertaining to differences in spread and overlap between vowel categories. This may be particularly salient where CDS modification is in the form of (unidirectional) system shifting rather than (multidirectional) hyperarticulation resulting in more peripheral vowel realisations, at least for the corner vowels /i/, /a/, and /u/ typically examined in studies of hyperarticulation in CDS. To further address the question of vowel differentiation, but also of enhanced exploitation of more extreme/peripheral realisations, we also calculated the difference between the size of the vowel ellipses of each vowel in ADS and CDS. As is clear from Table 4, Warlpiri CDS vowel ellipses are much larger than ADS vowel ellipses, indicating both greater within-category variation and a much larger vowel space for each vowel.

Finally, in order to examine the question of whether CDS vowel modifications are motivated by a desire to decrease vowel category overlaps (and presumably confusion about the vowel targets in question), and particularly the question of whether the expansion of the individual CDS vowel spaces in Warlpiri results in increased acoustic overlap, we calculated and compared Pillai scores for each vowel pair in ADS and CDS. A Pillai score, or Pillai–Bartlett trace, is a statistic computed in MANOVA and MANCOVA tests. Pillai scores range from 0 to 1, such that a score of 1.0 indicates complete separation of datasets while 0.0 indicates complete overlap. This statistic thus allows us to examine whether observed values predict category identity without phonological labels: Overlapping values will make it harder to predict the category identity, leading to a lower Pillai score, while categories can be predicted easily with non-overlapping values, thus leading to a high Pillai score. The Pillai score has been used in phonetics studies for measuring vowel mergers (e.g., Hay et al. 2006). For a methodological review and computational derivations, see Hall-Lew (2010) and Nycz and Hall-Lew (2013). In the case of our study, the Pillai scores were calculated. on F1/F2 measurements at the group level, similar to the calculation of overlap areas and proportions, because we only have a small dataset and the token numbers are not balanced for each speaker. If CDS vowel modifications increase acoustic differentiation/clarity by decreasing category overlap, we would expect Pillai scores to be higher in CDS than in ADS. As is clear from Table 5 below however,

Table 5: Overlap between vowels in adult Warlpiri by speech style. Pillai at 1.0 indicates complete separation while Pillai at 0.0 indicates overlap. Overlap areas are shown in Hz-squared, and percentage overlap score (%) is defined as the ratio between the conjunction over the disjunction of two ellipsis areas.

Vowel pair	Pillai: ADS	Pillai: CDS	Overlap: ADS	Overlap: CDS	% Overlap: ADS	% Overlap: CDS
/a/-/i/	0.54	0.47	114,805	151,562	16.0 %	14.0 %
/a/-/u/	0.40	0.25	165,440	430,957	23.4 %	37.0 %
/i/-/u/	0.61	0.46	71,705	209,409	10.8 %	22.7 %

the Pillai scores are higher for ADS than for CDS for each vowel comparison (/a/-/i/, /a/-/u/, /i/-/u/), indicating that rather than decreasing overlap between the three vowels, the observed CDS modification increases the overlap between categories. This is particularly the case for comparisons involving /u/, which encroaches on the vowel spaces of /i/ and /a/ as it is fronted: /a/-/u/ increases from 23.4 to 37 % overlap, while /i/-/u/ increases from 10.8 to 22.7 % overlap. This latter observation, we suggest, reflects the fact that it is possible to front /u/ vowels to a much larger extent than it is to front /i/ in a system shift.

2.2.2 Warlpiri vowels in CDS Nouns and CDS Non-Nouns

The combined data points from each of the four speakers (RA03, RA05, RA06, RA07) by Word Class (CDS Nouns vs CDS Non-Nouns) are presented in Figure 3 as scatter plots and estimated /i/, /a/, /u/ vowel ellipses, the latter with superimposed vowel triangles defined by the /i/, /a/, and /u/ vowel centroids. Individual descriptive statistics are found in Appendix 2. Mean values are presented in Table 6.

To test whether the four Warlpiri speakers produce CDS Nouns and CDS Non-Nouns which differ in spectral quality (F1 and F2), f_0 (fundamental frequency), we again built a linear mixed-effects model (LMM, fixed effects: Vowel [/i/, /a/, /u/] and Word Class [CDS Nouns vs CDS Non-Nouns]; random effect: Speaker), see Table 7. We checked the model using a Wald Chi-squared test, which (again unsurprisingly) revealed a significant main effect of Vowel for F1 and F2, as well as marginally for Duration. We also observed a significant main effect of Word Class in terms of vowel Duration, with vowels in CDS Nouns having longer durations than vowels in CDS Non-Nouns, as well as a Vowel $\times$ Word Class interaction in terms of F1. We further analysed the effect of Vowel and Speech Style by carrying out a series of *post hoc* tests (see Table 6), which indicate that vowels in CDS Nouns are longer than vowels in CDS Non-Nouns in general, and in vowels /a/ and /u/ in particular. We also observe a near-significant difference in terms of F1 (CDS Nouns may have slightly higher F1 values than CDS Non-Nouns).

We also compared the triangular vowel spaces denoted by /i/, /a/, and /u/ in Warlpiri CDS Nouns and CDS Non-Nouns (see Table 8 below) to determine whether CDS Nouns and CDS Non-Nouns differ in terms of vowel space expansion or hyper-articulation. As is clear from Table 8, and as is visually appreciable from Figure 3 (and for the speakers individually in Appendix 2), the space denoted by the vowel centroids in /i/, /a/, and /u/ in Warlpiri CDS Nouns is 54 % larger than the triangle denoted by the vowels in Warlpiri CDS Non-Nouns, a pattern consistent with observations from other studies where only vowel tokens from concrete (lab-provided) toys eliciting corner vowels /i/, /a/, and /u/, were compared across CDS and ADS.

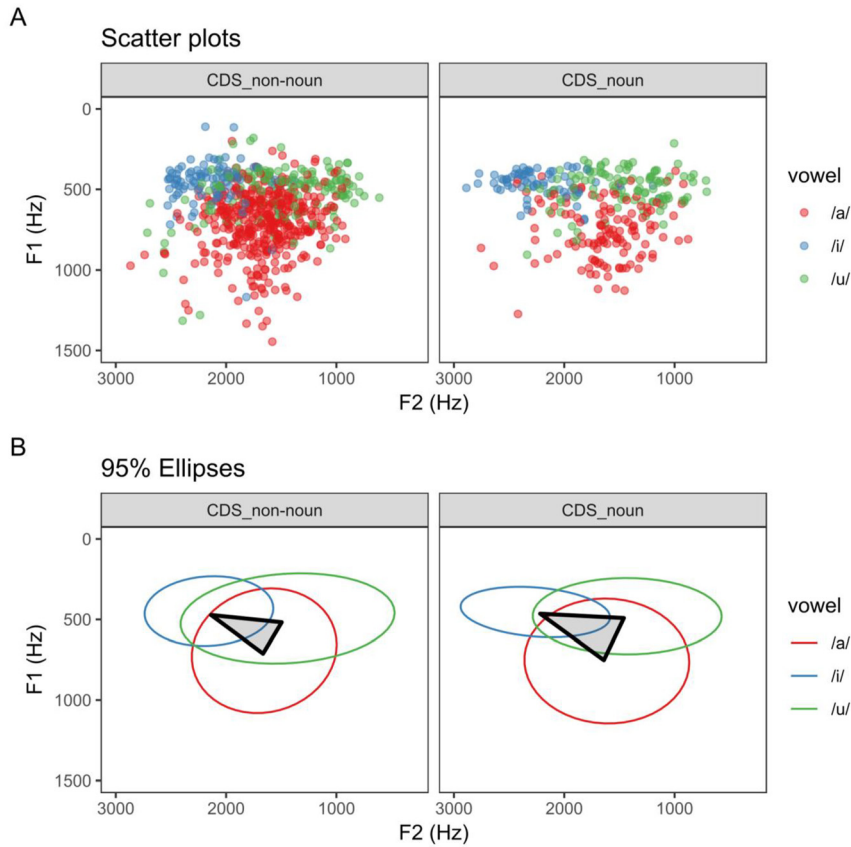


Figure 3: Spectral characteristics of adult Warlpiri vowels by Word Class, i.e., CDS Non-Nouns versus CDS Nouns. (A) Scatter plots of individual tokens. (B) Estimated vowel ellipses from the same dataset. Black triangles represent the centroids of the ellipses, i.e., mean values of F1/F2.

To further address the question of potential differences in vowel differentiation between Word Classes (CDS Nouns vs CDS Non-Nouns), but also of enhanced exploitation of more extreme/peripheral realisations of the three Warlpiri vowels, we again calculated the difference between the size of the vowel ellipses of each vowel in CDS Nouns and CDS Non-Nouns. Here, the results again differ from those in the ADS versus CDS comparison above: While the ellipsis for /a/ vowels in CDS Nouns expands, as observed also in the comparison between ADS and CDS, both /i/ and /u/ vowels show reduced within-category variation in CDS Nouns, as indicated by smaller CDS Noun vowel ellipses (a reduction in both cases of around 25 %).

Table 6: Acoustic characteristics of Warlpiri vowels in CDS Non-Nouns (CDS-NN) and CDS Nouns (CDS-N) in child-directed speech.

Warlpiri vowel	CDS-NN			CDS-N			Diff.	t Value	p Value
	Mean	SD	N	Mean	SD	N			
F1									
/a/	712	199	408	752	178	129	40	1.796	0.0728
/i/	472	174	96	464	77	68	−8	−0.780	0.4354
/u/	518	161	154	491	116	102	−27	−1.509	0.1316
F2									
/a/	1,665	312	408	1,641	365	129	−24	−0.399	0.6902
/i/	2,136	275	96	2,218	334	68	82	1.286	0.1987
/u/	1,496	459	154	1,460	389	102	−36	−0.872	0.3836
Duration									
/a/	80	42	408	117	89	129	37	4.069 ^a	0.0001
/i/	109	213	96	117	67	68	8	0.445	0.6564
/u/	76	50	154	118	101	102	42	3.587 ^a	0.0004
Combined	84	91	658	117	89	299	33	4.128 ^a	<0.0001
f_o									
/a/	216	45	359	219	42	122	3	0.257	0.7973
/i/	216	47	82	209	58	67	−7	−0.952	0.3414
/u/	206	47	138	213	48	91	7	0.910	0.3632
Combined	214	46	579	215	48	280	1	0.029	0.9770

^a*p* < 0.001.

Table 7: Effect of vowel, word class, and their interaction on acoustic measures in Warlpiri child-directed speech, based on linear mixed-effects modelling (LMM).

Measure	Vowel		Word class		Vowel × word class	
	Wald χ^2	<i>p</i> Value	Wald χ^2	<i>p</i> Value	Wald χ^2	<i>p</i> Value
F1	397.8 ^a	<0.0001	0.005	0.9461	6.1 ^b	0.0466
F2	423.2 ^a	<0.0001	0.04	0.8491	2.5	0.2802
Duration	5.6	0.0603	25.5 ^a	<0.0001	4.2	0.1219
f _o	0.06	0.8012	2.4	0.3076	1.7	0.4194

^a*p* < 0.001, ^b*p* < 0.05.

Table 8: Ellipsis area (in Hz-squared) of Warlpiri vowels in non-nouns and nouns in child-directed speech.

Vowel pair	Area: CDS-NN	Area: CDS-N	Diff.	N/NN
/a/ ellipsis	809,959	910,120	100,161	112 %
/i/ ellipsis	415,807	325,613	−90,194	78 %
/u/ ellipsis	847,743	637,512	−210,231	75 %
Triangle	65,967	101,363	35,396	154 %

Table 9: Overlap between vowels in Warlpiri child-directed speech by word class. Pillai at 1.0 indicates complete separation while Pillai at 0.0 indicate complete overlap. Overlap areas are shown in Hz-squared, and percentage overlap score (%) is defined as the ratio between the conjunction over the disjunction of two ellipsis areas.

Vowel pair	Pillai: CDS-NN	Pillai: CDS-N	Overlap: CDS-NN	Overlap: CDS-N	% Overlap: CDS-NN	% Overlap: CDS-N
/a/-/i/	0.416	0.582	168,469	108,877	16 %	10 %
/a/-/u/	0.184	0.427	472,664	334,888	40 %	28 %
/i/-/u/	0.416	0.515	254,376	150,596	25 %	19 %

Finally, to examine whether vowels in CDS Nouns differ from vowels in CDS Non-Nouns in terms of category overlaps, we again calculated Pillai scores for each pair of vowels by Word Class. As indicated in Table 9, the overlapping areas for vowels in CDS Nouns are smaller than for vowels in CDS Non-Nouns, indicating better acoustic differentiation and resulting in a lower % overlap score. This is consistent with the vowels being realized as more acoustically distinct units, as indicated by the results in Table 9 above, and together the results suggest that CDS Nouns represent clearer speech than CDS Non-Nouns (and ADS /a/-/i/ and /a/-/u/).

2.3 Study 1: Discussion

In Study 1, we presented first a comparison of the spectral and durational characteristics of Warlpiri vowels /a/ /i/, and /u/ in ADS and CDS produced by adults interacting with children in the same recording session. Secondly, we presented a comparison of the spectral and durational qualities of vowels in CDS Nouns and vowels in CDS Non-Nouns. The results show that Warlpiri caregivers modify their speech to young children by producing vowels that are characterised by higher F1 and F2 values, higher pitch/fundamental frequency (f_0), and longer durations (see summary in Table 3), and that vowels further, as a secondary process, undergo some degree of hyperarticulation and spectral focussing when they are produced as part of CDS Nouns. In what follows, we argue that the modifications to Warlpiri CDS vowels in general and the modifications specific to CDS Nouns constitute two different processes, with two different motivations. We address the first process—vowel system fronting and pitch raising in CDS in general—first, and then turn to the second process—vowel space expansion and vowel category differentiation.

The systemic fronting and lowering in Warlpiri CDS in general constitutes a novel finding and deserves explanation. While we cannot of course exclude that

these general CDS modifications in Warlpiri serve a linguistic purpose, unlike CDS modifications specifically targeting vowels in CDS Nouns (discussed below), we tentatively propose that the modifications to general CDS are not primarily linguistic. We favour this interpretation of the data for several reasons. Firstly, general CDS vowel modifications fail to result in improved segmental ‘clarity’, as demonstrated by larger (less focused) vowel ellipses (see Table 4), increased category overlaps (see Table 5), and an absence of increased vowel space (also Table 4). Negative evidence is, of course, at best indicative, but we draw also on evidence that prosodic modifications (higher pitch, and increased vowel durations included) of CDS often serve non-linguistic purposes such as infant emotional and attentional regulation. It is conceivable that CDS pitch modulations are used to convey affect and other non-linguistic information to young children, including what we might call ‘caregiver stances’ providing indications of which behaviours are desirable, and which are not, rather than assisting language learning directly.

Two recent studies are relevant to this latter point. The first, Kalashnikova et al. (2016), suggests that CDS modifications may reflect caregiver motivations to appear smaller and less threatening to children, and the second, Polka et al. (2022), demonstrates that infants prefer to listen to vowels with spectral characteristics (not just f_o) similar to their own. Interpreting the Warlpiri CDS vowel data in the light of these positions would suggest that it is plausible that Warlpiri caregivers

1. use prosodic modifications in CDS in general to appear less intimidating to children (rather than as a linguistic teaching tool), and;
2. systematically shift their vowel space up in Hz in the F2 dimension, as well as for /a/, up in Hz in the F1 dimension, producing vowels that are more consistent with the vowels produced by young children.

We examine this hypothesis in Study 2 below, along with the typological characteristics of Warlpiri that may allow caregivers to pursue such a strategy.

The second set of analyses reported above (Section 2.2.2) focus on the question of whether CDS modifications to young word-learners serve the didactic purpose of teaching ‘names for things’ as previous research (Odijk and Gillis 2021, 2022) has demonstrated caregiver sensitivity to word acquisition, and modification of vowel quality specifically in response to word acquisition (‘word birth’). This suggests that caregivers actively exploit this potential of CDS to assist young children to learn new words. The results from this analysis—a comparison of the spectral and durational properties of CDS nouns (‘words for things’) and CDS Non-Nouns (‘the rest’)—indicate that vowels in CDS nouns are longer in duration than vowels in other word classes in CDS. This finding is reminiscent of some of the results from a typologically diverse cross-linguistic study of noun and verb duration in Baure, Bora, Chintang, Dutch, English, Even, Hoocak Nlɨŋg, Sakha, and Texistepec (Strunk et al. 2020). This study

indicated that, when effects of word length and of morphological complexity are accounted for, nouns have longer durations than verbs in most languages. In a related study of pausing before nouns and verbs in Baure, Bora, Chintang, Dutch, English, Even, Hoocak Nlɨŋ, and Texistepec, six languages showed greater probability of pausing prior to a noun than prior to a verb (Seifart et al. 2018). The authors ascribe this finding to an inherently high processing demand for nouns, and argue that nouns typically introduce new information, or are replaced by pronouns or omitted.

The CDS data analysed here, however, does not follow that pattern: the Warlpiri CDS was reasonably rich with concrete nouns, while the ADS had a much lower proportion of nouns.³ We tentatively take this to suggest that the use of concrete nouns for established referents may be a choice made by caregivers in Warlpiri CDS, and that nouns are unlikely to be introducing new information at every mention. It is also plausible that the observed repetition is part of teaching adult Warlpiri discourse style, as proposed by Bavin (2000, p. 574):

A feature of Warlpiri adult discourse is repetition, which is reported to be a common feature of narratives in oral tradition (Brewer 1985). [...] A “build-up” style is often used: information from one sentence is partly repeated in the next with something new added. [...] For example, in telling the frog story a speaker might give the information that someone fell, then someone fell to the water, then someone fell down to the water, and then specify that it was the child and dog who fell. Not all is revealed at once. This could be a way of holding the attention of the listeners, although this is speculative. Repetition is noted even in the talk from three-year-old Warlpiri children.

Importantly, the results reported here go beyond durational enhancements such as those reported by Strunk et al. (2020) and show that vowels in CDS nouns are characterised by substantial vowel space expansion (see Table 8), increasing their acoustic differentiation. Vowels in CDS Nouns are also showing reduced within-category variance: In other words, while CDS nouns, like CDS in general, show evidence of systemic fronting relative to ADS, vowels from CDS nouns *also* demonstrate some of the more frequently reported CDS characteristics (hyperarticulation, enhanced acoustic separation, and reduced within-category variance) relative to other CDS vowels.

We argue that the results from the ADS versus CDS and the CDS Nouns versus CDS Non-Noun-analyses reported here are consistent with the interpretation that

³ We were unable to perform an ADS Noun versus ADS Non-Noun analysis due to the limited size of the ADS Noun dataset; there were fewer than five tokens and these included loan-words such as English ‘plane’ and ‘kimby’, (English and/or Kriol for a baby’s nappy/diaper, from a brand name). The elision of nouns in adult-to-adult speech when the noun referent is recoverable from the discourse is a well-documented aspect of Warlpiri information structure (Hale 1992; Hale et al. 1995; Swartz 1991).

Warlpiri CDS vowel modifications have two separate components: The first component applies to CDS in general and is characterised by vowel system shifting (fronting), pitch raising, and longer vowel durations. CDS Nouns appear to undergo a second process of enhancement in addition to the fronting/lowering of vowels in CDS in general. This type of modification may have separate motivations from the initial fronting/lowering, and we argue that it constitutes a didactic component tailor-made to young word-learners, whose caregivers may (consciously or not) realise that the child productions are noun-biased.

We base our argument that CDS Noun modifications in Warlpiri are didactically motivated on the following reasoning: Firstly, we note that the style of modifications (vowel space expansion, reduced within-category variance) observed specifically in Noun CDS is similar to what has been reported for vowels in IDS in English and other languages in input to infants under 12 months (e.g., Werker et al. 2007). One possible explanation for this style of modification is that IDS modifications serve the purpose of facilitating phonological acquisition in the first year of life, by providing high quality input. We propose here, by analogy, that the modifications observed in Warlpiri are similarly didactically motivated though the focus is not phoneme acquisition, but word learning, through the provision of clear and unambiguously produced target words, that provide ‘best possible’ information about the lexical specifications of the words (nouns) in question, much like what we might think is the objective of the behaviour of the Dutch caregivers described by (Odiijk and Gillis 2022). This is also consistent with patterns of avoidance of lenition processes in speech to children of a similar age to those participating in the present studies (Buchan and Jones 2014): reducing speech clarity is not helpful at an age where clarity of speech may assist a child to learn and recognise new words.

Secondly, and in lieu of formal assessment of the vocabularies and language use of the participating children,⁴ we base the interpretation of the Noun-specific modifications as a didactic strategy on a rich literature on noun acquisition. This literature has demonstrated that nouns tend to be acquired before verbs by young children (e.g., Clark 1993; Gentner 1978, 1982), and that young children use more nouns than verbs in all play activities (Gelman and Tardiff 1998), while older children (around 36 months of age), still use more nouns in book-play/reading, than in playing with mechanical toys (Gelman and Tardiff 1998; Tardiff et al. 1999). These are activities like those in which the participating children were involved during the data

⁴ An authorized Macarthur-Bates Communicative Development Inventory for Warlpiri, Eastern & Central Arrernte, Western Arrarnta and English has been developed from the recordings analyzed here and other recordings in the Little Kids Learning Languages project. Prior to this there was no assessment instrument available for these languages.

recording. Evidence also suggests that verbs remain more difficult to retrieve than nouns, even for 5-year-old children (D'Amico et al. 2002).

In making this argument, we acknowledge the possibility that the differences between CDS Nouns and CDS Non-Nouns reflect properties of the two defined Word Classes rather than an CDS-related phenomenon as such. We cannot rule this out with the current dataset, but we would like to make several observations on the data set and the existing literature on acoustic differences between vowels in nouns and other word classes. Firstly, we highlight that the CDS Non-Noun word class contains a range of grammatical classes (including a large proportion of open class words); particularly (imperative) verbs, pronouns and demonstratives, as well adjectives. The CDS Non-Noun class is thus not a class consisting of only grammatical elements or morphemes; it has a substantial proportion of lexical words. Warlpiri is an agglutinative language and nouns often host case-markers and other bound morphemes. Nouns hosting bound morphemes are included in the nouns measured, but non-nouns were only included when they were not part of a word with a noun stem. We also highlight that Warlpiri is not characterised by vowel reduction in unstressed syllables, a phenomenon that contributes significantly to differences between, in particular, closed and open word-classes in English (Monaghan et al. 2005, 2007; Morgan et al. 1996; Shi 1995), with closed word classes having shorter and more centralised vowels than open word classes, including nouns and verbs. This finding has, to some extent, also been demonstrated for other languages (Turkish and Mandarin: Shi et al. 1998), as well as Dutch, Japanese and French (Monaghan et al. 2007), but equally importantly for this study, Monaghan et al. (2007) have also demonstrated that Nouns have been found to contain a higher ratio of reduced vowels than verbs in child-directed speech in English, but not in Dutch. This latter combination of findings is important because it indicates that we cannot assume that nouns in Warlpiri (in CDS or in general) are necessarily produced with enhanced (or unreduced) vowel clarity relative to other word classes, including those encompassed by the CDS Non-Noun class in the present study. Finally, as we discuss below, we observe no effect in terms of f_0 between the two Word Classes (CDS nouns vs CDS Non-Nouns), and we direct the reader's attention to research suggesting that prosodic cues do not distinguish noun phrases from verb phrases (Cooper and Paccia-Cooper 1980).

As indicated above, the results reported on here suggest that vowels from CDS nouns are not characterised by a higher f_0 than Non-Noun CDS vowels, suggesting that f_0 is not used for the same didactic purposes as vowel hyperarticulation and reduced within-category variance, and vowel space expansion. This appears consistent with our analysis of the modifications in CDS in general, including the argument that prosodic modifications (vowel duration, pitch) may reflect caregiver attitudes and serve non-linguistic communicative purposes, like maintaining child

attention, and appearing ‘friendly’ or unthreatening (Kalashnikova et al. 2016), as pitch is not used to ‘highlight’ nouns in CDS.

In Study 2, we return to the question of what might motivate CDS modifications in Warlpiri, beyond the didactic intentions demonstrated for CDS Nouns. We do so with a focus on the possibility that caregivers exploit child preferences for speech with vowels with similar spectral characteristics to child vowels (Polka et al. 2022), through a study of the relative spectral similarity of Warlpiri ADS and CDS vowels to the Warlpiri vowel productions of the children participating in Study 1.

3 Study 2: Warlpiri child vowels

The results from Study 1 indicate that Warlpiri caregivers modify the quality and duration of vowels in CDS, as well as their f_0 . They do so in a manner that suggests

1. that the CDS vowel space shifts upwards in the F2 dimension relative to the ADS vowel space, that CDS is characterised by increased F1/F2 vowel dispersion, increased vowel overlap, increased vowel duration, and a higher f_0 in general. This implies that fronting of the vowels in CDS emulate a smaller vocal tract, while the f_0 raising may imply larynx raising, i.e., shorter tract tube length, and;
2. CDS Nouns exhibit enhanced vowel differentiation (indicated by a larger vowel space), and reduced within-category variance/contrast overlap relative to other parts of CDS speech, but CDS Nouns do not differ from CDS Non-Nouns in terms of their F1/F2 values, indicating that all CDS Nouns and CDS Non-Nouns shift in a similar manner relative to ADS (primarily in the F2 dimension).

As discussed in Section 2.3, we interpret this pattern to indicate a two-part process of CDS modification consisting of general vowel shifting (upwards in Hz in the F2 dimension), as well as duration and f_0 increases between ADS and CDS, and an additional process of enhanced vowel differentiation and reduced within-category variance associated exclusively with CDS Nouns. We argue this secondary stage to be evidence of a didactic caregiver intention that reflects the developmental stage of the children in the study.

Study 2 addresses the motivation underpinning the first of the two processes—the general shifting of the vowel space in conjunction with duration and pitch modification. It tests the hypothesis that the ‘general’ Warlpiri CDS vowel modifications, excluding those that pertain only to nouns, serve the purpose of enhancing child attention by harnessing well-documented child preferences for IDS, and particularly their preferences for speech with higher f_0 (Fernald and Kuhl 1987) and speech with F1/F2 modifications (Polka et al. 2022) that bring adult speech production closer to child speech production: this is an innovation, rarely if ever done in

previous IDS/CDS studies which have focused solely on the adult IDS versus ADS speech and not the child's vowels. The study does so by first describing the acoustic characteristics of vowels produced by the four Warlpiri-acquiring children that participated in Study 1, and secondly, by comparing the acoustic characteristics of the child vowels to the adult CDS and ADS vowels from Study 1.

3.1 Methods

3.1.1 Participants

We recorded the four children who participated with their caregivers in Study 1; this included three boys; *RC01* (boy; 28 months), *RC10* (girl; 46 months), *RC11* (boy; 30 months), and *RC15* (boy; 25 months). No atypical development or hearing loss was reported. All four children are acquiring Warlpiri as (one of) their L1(s), and all grow up in multilingual households and have exposure to varieties of English and English-lexified contact varieties, and some are exposed to other Indigenous Australian languages (e.g., Arrernte/Arrarnta). The children were recorded at the same time as their caregivers (Study 1) in interaction with close family members, during play and story-telling activities in their homes in either Alice Springs or Yuendumu in the Northern Territory of Australia.

3.1.2 Materials

The recording equipment used to collect the child data is described in Study 1. The video/audio recordings of the four child participants (*RC01* [28 months], *RC10* [46 months], *RC11* [30 months], and *RC15* [25 months]) were transcribed, glossed, and translated into English, in ELAN 6.3, in the same manner as the adult data in Study 1. The children's productions were generally short, and contained few verbs, but many nouns such as *jarntu* (dog), *mamu* (monster), and *ngipi/ipi/nipi* (various realisations of egg), typically in isolation. The child utterances were remarkably similar to what has been reported elsewhere for Warlpiri-acquiring children at the same age: see Bavin (2000) for glossed transcriptions of utterances from a child aged 24 months and a child aged 30 months. We extracted 436 vowel tokens from the four children. Appendix 3 shows participant contributions.

Target vowels were hand-segmented and labelled in *Praat* 6.2.12., and vowel duration, F1, and F2 extracted using the same automatic script as in Study 1. We manually estimated f_o in *praat*, as the automatic script (Kroos et al. 2010) was unable to estimate f_o values for 25 % of the child vowels; manual estimation of f_o allowed us to include 96 % of the vowels in our study. We estimated f_o on the basis of the first

formant peak, or the highest point on the peak where two points were identifiable, in the spectral slice, using the *view spectral slice* function in *praat*, with pitch range set to 75–500 Hz. Vowels degraded by environmental noise, overlapping talkers, etc., were excluded. As in the adult data, /a/ was more frequent in the data than /i/ and /u/.

3.2 Results

The Warlpiri child vowel data from the four child participants (*RC01*, *RC10*, *RC11*, and *RC15*) are presented in Figure 4 as a scatter plot and as estimated /i/, /a/, /u/ vowel ellipses, the latter with a superimposed vowel triangle defined by the /i/, /a/, and /u/ vowel centroids. Mean values in Hz are presented in Table 10, and individual descriptive statistics are presented in figures in Appendix 4. We do not include f_o in the statistical analyses, or in formal comparisons with adult ADS or CDS.

Although the /i/, /a/, and /u/ vowels produced by Warlpiri-learning children in the study were highly overlapping, the ellipses still indicate three distinctive vowel categories, /i/, /a/, and /u/ (see Figure 3B). To confirm the contrastiveness in the vowel

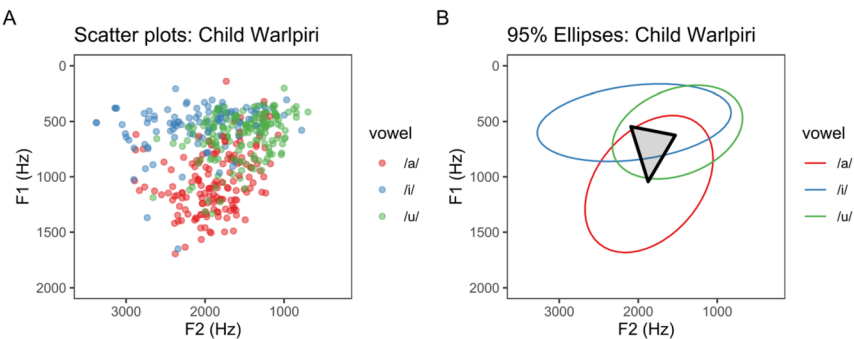


Figure 4: Vowels produced by child Warlpiri speakers.

Table 10: Acoustic characteristics of vowels in child Warlpiri.

Warlpiri vowel	F1 (Hz)		F2 (Hz)		Duration (ms)		f_o (Hz)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
/a/	1,044	282	1,875	384	132	74	359	72
/i/	547	210	2,093	560	160	102	356	90
/u/	625	209	1,531	373	171	86	363	86
Average					155	88	360	82

Table 11: Acoustic area (in Hz-squared) of vowels in child Warlpiri. We include the values reported for ADS, CDS, CDS Nouns, and CDS Non-Nouns presented above for easy comparison.

Data	/a/	/i/	/u/	Triangle
Child	1,464,255	1,300,625	1,042,150	131,155
ADS	485,347	347,466	386,034	77,783
CDS	845,927	384,450	748,686	77,105
CDS NN	809,959	415,807	847,743	65,967
CDS N	910,120	325,613	637,512	101,363

Table 12: Child vowel Pillai score, Overlap size (in Hz-squared) and Overlap % of vowel pairs in child Warlpiri, with adult ADS and CDS values included for direct comparison. See Table 5 for overlap between vowels in adult Warlpiri by speech style (ADS vs CDS), and Table 9 for overlap between vowels in Warlpiri CDS by word class (Nouns vs Non-Nouns).

	/a/-/i/	/a/-/u/	/i/-/u/	Average
Pillai score (child)	0.541	0.432	0.334	0.436
Overlap (child)	341,424	552,412	689,569	527,569
% Overlap (child)	14.1 %	28.3 %	41.7 %	28.0 %
Pillai score (ADS)	0.541	0.396	0.614	0.517
Pillai score (CDS)	0.470	0.248	0.456	0.391

distributions, we built another two linear mixed-effects models (LMMs) to compare the mean formant values. When checked by a Wald Chi-squared test, we found a significant effect of vowel for both F1 ($\chi^2 = 370.7$, $df = 2$, $p < 0.0001$), and F2 ($\chi^2 = 132.88$, $df = 2$, $p < 0.0001$). A series of Bonferroni adjusted *post hoc* comparisons further revealed that for F1, /a/ was significantly higher than /u/ ($p < 0.0001$), and the F1 for vowel /i/ was also significantly higher than for vowel /u/ ($p = 0.0160$), i.e., /a/ > /u/ > /i/. For F2, the post hoc comparisons revealed that /u/ is lower than /a/ ($p < 0.0001$), and /i/ is higher than /a/ ($p = 0.0003$), i.e., /i/ > /a/ > /u/. Taken together, these results confirm that all three vowels are contrastive acoustically in Child Warlpiri, despite the high level of overall spectral overlap. As we did with the adult Warlpiri data in Study 1, we also calculated both the vowel triangle denoted by the /i/, /a/, and /u/ centroids and the size of each individual vowel ellipse (see Table 11), and the overlap between each pair of vowel categories (see Table 12).

In order to test the hypothesis that Warlpiri CDS is characterized by modifications that serve the dual purpose of maintaining child interest/attention, in addition to scaffolding linguistic development, we subsequently compared the child Warlpiri vowel data (vowel triangle size, and vowel category overlap in terms of Pillai scores) with the adult ADS and CDS data to determine whether CDS modifications in Warlpiri

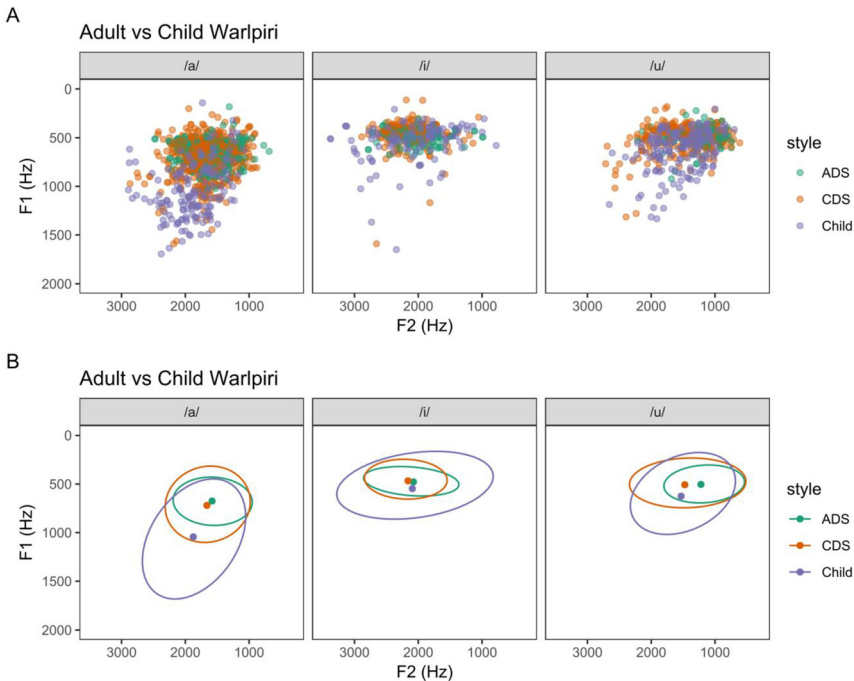


Figure 5: Comparison of adult and child Warlpiri vowels. For adult speakers, vowels are further distinguished by speech style, i.e., adult- and child-directed speech. Dots in ellipses indicate centroids.

result in vowel qualities that reflect the vowel qualities of child productions. Figure 5 presents visual comparisons of the child data and the adult CDS and ADS data, respectively, first in scatter plots, and then in terms of individual vowel ellipses.

A numerical comparison of the size of CDS, ADS and child vowel triangles and vowel ellipses (Table 12) suggest that Child productions are more like CDS than ADS, and in particular like CDS Nouns. Lower Pillai overlap-scores for /a/ and /u/ for CDS

Table 13: Overlap between Warlpiri vowels between adult and child speakers. Pillai at 1.0 indicates complete separation while 0.0 indicates complete overlap. Acoustic overlap areas are shown in Hz-squared. Percentage (%) overlap score is defined by the ratio between the conjunction and the disjunction of the two ellipses.

Warlpiri vowel	Pillai: ADS-child	Pillai: CDS-child	Overlap: ADS-child	Overlap: CDS-child	% Overlap: ADS-child	% Overlap: CDS-child
/a/	0.484	0.286	417,315	629,396	27.2 %	38.0 %
/i/	0.052	0.060	347,469	418,079	26.7 %	32.1 %
/u/	0.189	0.101	358,414	652,691	33.5 %	58.0 %

and Child vowels than ADS and Child vowels (see Table 13) similarly indicate that the general CDS vowel modification in Warlpiri results in much greater F1/F2 overlap between the adult and child vowels. This greater similarity between CDS and child vowels is further illustrated by the % overlap calculations, also in Table 13.

3.3 Study 2: Discussion

The results from Study 2 provide a first acoustic study of the vowels produced by children acquiring Warlpiri. The results show, unsurprisingly, that Warlpiri-acquiring children in their third and fourth year of life produce the Warlpiri /i/, /a/, and /u/ vowels with relatively wide distributions (vowel ellipses; Table 11) with some category overlap (Table 12), and with relatively small vowel triangles denoted by the /i/, /a/, and /u/ centroids (also Table 11). This result is consistent with reports indicating that children acquiring a range of languages achieve a high degree of vowel accuracy by the age of 36 months (for an excellent review, see Donegan 2002).

Figure 4 also indicates that child Warlpiri /i/, /a/, and /u/ vowel *centroids* are more central (lower F2 for /i/, slightly higher F2 for /u/) than those reported for children speaking French (Ménard et al. 2007) and English (Vorperian and Kent 2007), likely reflecting the narrower F1 and F2 ranges observed in adult vowel productions in Australian Indigenous languages with smaller vowel inventories (Butcher 1994). This aspect may reflect differences in vowel implementation borne out of the characteristics of the vowel typologies of the languages in question: the successful implementation of large vowel inventories may demand smaller ellipses for each vowel, and perhaps also more systematic exploitation of all parts of the articulatory space, in service of contrast maintenance, a pressure less likely to influence the realisation of vowels in three-vowel inventories.

Comparisons of the overlap between the child productions of /i/, /a/, and /u/, and those of adult ADS and CDS in terms of vowel triangle size (the area denoted by the centroid of each vowel category), and the vowel distribution (i.e., the vowel ellipses for each vowel) indicate that CDS modifications, beyond f_o , result in adult productions increasing in spectral similarity to child productions. This may be particularly the case for vowels in CDS nouns.

Finally, we acknowledge that we do not have direct evidence that the observed CDS modifications induce greater child attention to their caregivers' speech, nor explicit evidence that the modifications are intended to make the caregivers sound more child-like. The results are however consistent with the notion that spectral modifications in CDS (F1–F2 modifications) reflect multiple concurrent caregiver goals, including linguistic-didactic, bonding and attention regulation goals, as also suggested by Polka and Ruan (2021).

4 General discussion

The present paper reports on two closely linked studies of CDS and child vowel production in the Australian Indigenous language Warlpiri. Study 1 represents the first examination of the acoustic comparison of CDS vowels in an Indigenous Australian language (Warlpiri), and, to our knowledge, the first study of the characteristics of CDS in a three-vowel system. The second study examines the acoustic characteristics of child vowel production in Warlpiri and compares this data with the adult data from Study 1, and to our knowledge, is also the first study to conduct such a comparison.

The results from Study 1 show that adult Warlpiri caregivers modify their speech to children by shifting their vowel space to higher values in Hz in the F2 dimension (and also higher values in Hz in the F1 dimension), by extending vowel duration and by increasing their f_0 , possibly to enhance child attention to, or preference for, the speech input. Indeed, the results presented here contribute to growing evidence, much needed from lesser- and under-studied, typologically diverse languages from culturally diverse (non-WEIRD; Henrich et al. 2010) and geographically distributed populations (Kidd and Garcia 2022), that some form of IDS/CDS modification is implemented in many parts of the world (and recent research indicates it is also recognised across cultures: Hilton et al. 2022).

The results however also add to a growing literature that demonstrates that IDS/CDS modifications may not be identical across different languages and language typologies (Bernstein Ratner and Pye 1984; Fernald et al. 1989; Igarashi et al. 2013; Kitamura et al. 2001), and raise questions about the underlying learning mechanisms. Indeed, while the durational and f_0 characteristics of Warlpiri CDS may be consistent with reports from many languages (but not all: e.g., Bernstein Ratner and Pye 1984), the spectral characteristics of Warlpiri CDS in general (we return to the modifications specifically for vowels in nouns in CDS) *differ* from what has been described for many other languages (and sometimes suggested as near-universal), namely that IDS/CDS is characterised by hyperarticulation of vowels. Instead, our results suggest that Warlpiri CDS modifications take the shape of a *systemic fronting and lowering* of the vowel system—importantly, F2 was increased for ALL vowels, including /u/ which would have lowered F2 values (i.e. more back) if produced with hyperarticulation. We argue that this CDS adaptation reflects an effort on behalf of the caregivers to bring their production (in terms of F1 and F2) closer to the productions of the children addressed, potentially to enhance child attention to the speech input (Polka et al. 2022). One of the Warlpiri co-authors commented that when speaking to very young children Warlpiri adults aim to sound like the children (Nelson, p.c. 2022), in keeping with the idea that Child Directed Speech in Warlpiri is

an adult construct that imitates aspects of child production (Laughren 1984), however we would think that the nuances of vowel realization would be below the level of conscious awareness. The adaptation is in addition to an increase in f_o in CDS relative to ADS, a modification that has elsewhere been argued to serve the purpose of emotional regulation (as opposed to linguistic purposes) by enhancing the affective salience of the utterances (Kitamura and Burnham 1998), or indicating positive affect (Singh et al. 2002).

The study also indicates that the didactic role of CDS is tailored to the developmental stage of the child addressees. In the present study, the children were all at an age where vocabulary growth is typically exponential, and the results reported here suggest that vowels in CDS Nouns are special in that they undergo a potential additional process of hyperarticulation in Warlpiri: While CDS vowels *shift* in the manner described above, vowels in CDS Nouns also show reduced acoustic overlaps (greater contrast separation) and reduced within-category variance (smaller vowel ellipses), as well as an increased duration, compared to vowels in other words (including verbs and other open classes). We argue that this may be a didactic strategy adopted by caregivers to ensure high-quality lexical information (segmental specification) to young word-learners. This is not a finding that has been reported elsewhere (most CDS studies use only vowels from nouns), though IDS/CDS has been shown to assist word recognition and word learning in the laboratory studies (Ma et al. 2011; Song et al. 2010), and as indirectly supported by observations that the degree of maternal IDS is positively correlated with toddler productive vocabulary (Hartman et al. 2017). It is also a finding that is consistent with evidence of caregiver vowel modifications to young word-learners centred specifically around emerging vocabulary items (Odijk and Gillis 2022).

Additionally, the studies reported here present preliminary evidence that it is possible that some IDS/CDS modifications, even spectral modifications, serve two different purposes concurrently: linguistic-didactic purposes, and attention-regulating purposes by recruiting child preferences for child-like productions. This is a novel finding, as is our tentative conclusion that the CDS modifications observed in Study 1 reflect not a single process but rather two CDS processes:

1. A shifting and fronting of the vowel space in terms of F2, a lowering of /a/ (increase in F1), and an associated increase in f_o , and
2. A process of vowel space expansion (by 54 %) and reduction of within-category variance associated specifically with vowels in CDS Nouns, and which does not involve further f_o raising, but which does involve increased vowel durations.

We have argued here that the first process applies to CDS in general, and taking these results in conjunction with those from Study 2, we suggest that this initial modification may be viewed as phonetic convergence which may confer multiple benefits

(increasing attention, positive affect, social bonding, building familiarity with their own vocal patterns) (Polka et al. 2022), and potentially also, in particular with respect to the observed CDS f_0 raising, with a desire to sound small and less threatening (Kalashnikova et al. 2016). We have also argued here that the second process (reduced within-category variance and vowel space expansion for CDS Nouns) reflects the developmental stage of the child addressees in the studies, as well as the special status of nouns in early language acquisition (Clark 1993; D'Amico et al. 2002; Gelman and Tardiff 1998; Gentner 1978, 1982; Tardiff et al. 1999).

As pointed out in the discussion sections above, the observed vowel system fronting and additional vowel space expansion and reduced within-category variance of vowels in nouns in Warlpiri CDS differs from many other reports of CDS hyperarticulation and vowel space expansion as important characteristics of CDS vowels. We are not aware of research on the acoustics of IDS/CDS in languages with three-vowel systems other than Warlpiri with which to compare our results, but we offer two typologically-based reflections that may be relevant to understanding the phenomenon of CDS fronting (and lowering) in Warlpiri CDS, and our proposal that we are observing a two-part process of modification (one intended to enhance communicative attention, and one intended to assist word acquisition). It is possible here to link back to the findings of Seifart et al. (2018) and the claim that nouns pose processing challenges and increased processing load because they introduce new material (familiar objects are less likely to be named, compared to novel objects, increasing the cognitive load as comprehension first requires an association to be made). While we argue that the CDS noun use here is different in that it does not necessarily introduce new information but rather instances of rephrasing and repetition, e.g., *'Look at the egg. Here is the egg. Can you see the egg?'*, it is plausible that this additional cognitive load of nouns (as per Seifart et al. 2018) is also relevant to CDS. It does however not account for the observed differences in CDS and ADS, nor for the additional reduced within-category variance and vowel triangle expansion observed between CDS nouns and CDS Non-nouns in the present study; the claims by Seifart et al. (2018), Strunk et al. (2020), and Lester et al. (2019) all pertain to duration, not the spectral characteristics of vowels in nouns.

The Warlpiri three-vowel system is typologically very different from that of English and other European languages with large vowel inventories, the languages on which many, if not most, IDS/CDS studies have been based. Relative to English, the three-vowel system of Warlpiri is typologically very *different*—but this difference should not lead to the assumption that languages like English with large vowel inventories are in fact the norm, and that languages like Warlpiri are *unusual*. Rather, as discussed in the Introduction, it is the other way around: typologically, large vowel inventories like that of English are unusual. It is possible that Warlpiri speakers, and speakers of many other languages in the world with simple 3–5 vowel systems have a great deal more freedom to exploit in their IDS/CDS than speakers of

languages like English. We see support for this assumption in the fact that Warlpiri CDS is characterised by big vowel ellipses and small vowel triangles—the opposite of what we see in languages such as English, which tend to be characterised by big vowel triangles and small vowel ellipses. These typological differences are easy to comprehend intuitively, as are perhaps the resulting system-specific constraints on the implementation of IDS/CDS modifications, and we present a schematic illustration of the effect of Warlpiri-like CDS modifications on an imaginary three-vowel and an imaginary eight-vowel language in Figure 6. Indeed, it may not seem feasible to modify a large vowel inventory through systematic fronting and/or lowering (whether to sound more child-like or not), as well as expand the vowel ellipses for CDS in general (and reduce within-category variance just for CDS Nouns). In fact, we hypothesise that such an approach would result in increased overcrowding in the front half of the vowel space and induce serious problems with category overlaps, even when the system is characterised by (unusually) equidistant vowels and fewer vowels than, for instance, English (12 monophthongs in Australian English) and Norwegian (20+ monophthongs). It may also be a point to consider in relation to studies (particularly of languages with rather large vowel inventories) where no evidence for vowel hyperarticulation has been found (e.g., Danish: Bohn 2013). Conversely, it may be unnecessary to implement multidirectional vowel hyperarticulation in general in a three-vowel system, and plausible that vowel hyperarticulation can be harnessed in a more refined manner to accommodate the developmental stage of a child interlocutor; here, we argue, to assist young word-learners with vocabulary development.

Focusing on research from lesser-studied, under-represented languages, as well as typologically diverse languages in child language research (Kidd and Garcia 2022)

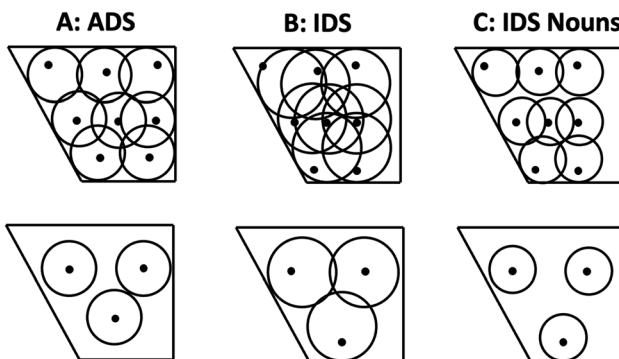


Figure 6: Schematic illustration of Warlpiri-style CDS modifications applied to an imaginary three- and an imaginary eight-vowel system, with relatively equidistant vowel centroids. Panel A shows the unmodified ADS, while Panel B shows the system fronted with expanded vowel ellipses. Panel C shows the fronted system with reduced within-category variance.

is thus necessary. We argue this not only for the sake of diversity or representation alone, but because diverse languages may hold the key to different questions in language acquisition research: The conclusions from the studies presented here would have been very unlikely to have arisen from a similar study of English, French or Norwegian, for obvious typological reasons. The results highlight that not all research questions can be answered through the investigation of a small set of languages, and further, in this case, that processes of IDS/CDS may be much more diverse and complex, and even nimble, than volumes of studies of languages with large vowel inventories may have led many to assume.

Finally, we wish to offer a comment on the number of speakers and the question of reliability in the studies presented here. While we fully acknowledge that four adult and four child participants is a small number of individuals (within 25–50 % of the number of participants of many studies with speakers of the main European languages like English), it is important to recognise that this group of eight represents a much larger proportion of the Warlpiri speaking population, as well as a severely understudied non-WEIRD population, than what is typically sampled. With these observations in mind, we hope that the findings and methodological innovations trialled here will inspire similar and further examinations of IDS/CDS in other languages with small vowel inventories.

Acknowledgements: We thank the families who participated in the studies, as well as Brett Baker and Barbara Kelly[†] for careful comments on the manuscript.

Statement of ethics: The participants of the study presented provided written informed consent. The research presented was approved by the Central Australian Human Research Ethics Committee #CA-20-3633 and the Australian National University Human Research Ethics Committee #2019-183.

Conflict of interest statement: The authors have no conflicts of interest to declare.

Funding sources: The research was funded by the Australian Research Council Future Fellowship project #FT190100243, and Australian National University Futures Scheme award, both awarded to Carmel O'Shannessy.

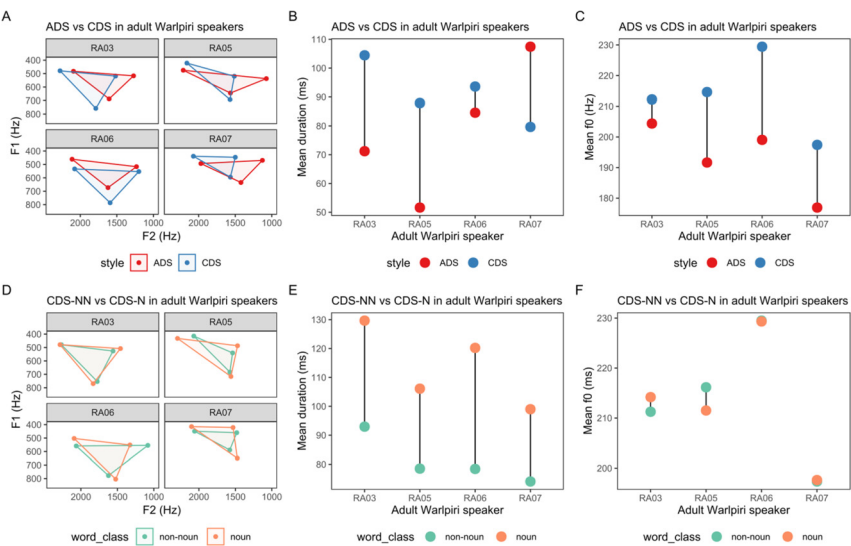
Author contributions: Rikke Bundgaard-Nielsen (RBN) conceptualised the studies. Carmel O'Shannessy (CO'S), Alice Nelson (AN), Jessie Bartlett (JB), and Vanessa Davis (VD) designed the data collection protocol. AN, JB and CO'S recruited and interacted with the participating families, recorded the data, transcribed, translated, and coded the recorded data. RBN undertook the acoustic segmentation, labelling, and acoustic measurements. RBN, Yizhou Wang (YW), and CO'S jointly decided on the statistical analyses, and YW conducted the analyses. Interpretation of the statistical results was undertaken by RBN and YW. RBN drafted the manuscript, and CO'S and YW provided comments and suggestions.

Appendix 1: Number of tokens from each adult Warlpiri speaker in the dataset (N = 1599).

ADS = Adult Directed Speech; CDS = Child Directed Speech; CDS NNs = Child Directed Speech Non-Nouns; CDS Ns = Child Directed Speech Nouns.

Warlpiri speaker	Vowel			Total
	/a/	/i/	/u/	
ADS				
RA03	196	71	53	320
RA05	5	2	2	9
RA06	115	50	24	189
RA07	62	32	30	124
Total	378	155	109	642
CDS				
RA03	210	64	109	383
RA05	147	32	68	247
RA06	101	31	25	157
RA07	79	37	54	170
Total	537	164	256	957
CDS-NN				
RA03	165	34	64	263
RA05	102	19	42	163
RA06	70	17	13	100
RA07	71	26	35	132
Total	408	96	154	658
CDS-N				
RA03	45	30	45	120
RA05	45	13	26	84
RA06	31	14	12	57
RA07	8	11	19	38
Total	129	68	102	299

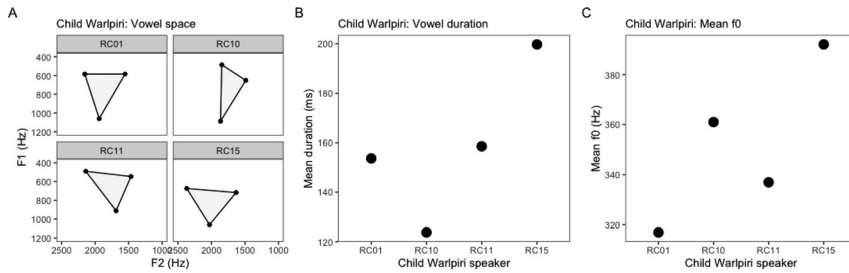
Appendix 2: Individual results for adult Warlpiri speakers. Values are averaged across all valid observations.



Appendix 3: Number of tokens from each child Warlpiri speaker in the dataset ($N = 436$).

Child speaker	Vowel			Total
	/a/	/i/	/u/	
RC01	30	16	36	82
RC10	67	37	47	151
RC11	30	27	49	106
RC15	28	25	44	97

Appendix 4: I Individual results for child Warlpiri speakers. Values are averaged across all valid observations.



References

- Adriaans, Frans & Daniel Swingley. 2017. Prosodic exaggeration within infant-directed speech: Consequences for vowel learnability. *Journal of the Acoustical Society of America* 141(5). 3070–3078. Australian Bureau of Statistics. <https://www.abs.gov.au/census>.
- Bavin, Edith. 2000. Ellipsis in Warlpiri children's narratives: An analysis of frog stories. *Linguistics* 38(3). 569–589.
- Bernstein Ratner, Nan & Clifton Pye. 1984. Higher pitch in BT is not universal: Acoustic evidence from Quiche Mayan. *Journal of Child Language* 11(3). 515–522.
- Bohn, Ocke-Schwen. 2013. Acoustic characteristics of Danish infant directed speech. *Proceedings of Meetings on Acoustics (POMA)* 19. 060055.
- Brewer, William F. 1985. The story schema: Universal and culture-specific properties. In D.R. Olson, N. Torrance, A. Hildyard & R. Scollon (eds.), *Literacy, language and learning: The nature and consequences of reading and writing*, 167–194. Cambridge: Cambridge University Press.
- Buchan, Heather & Caroline Jones. 2014. Phonological reduction in maternal speech in northern Australian English: Change over time. *Journal of Child Language* 41(4). 725–755.
- Bundgaard-Nielsen, Rikke L. & Carmel O'Shannessy. 2021. Voice onset time and constriction duration in Warlpiri stops (Australia). *Phonetica* 78. 113–140.
- Burnham, Denis, Christine Kitamura & Ute Vollmer-Conna. 2002. What's new, pussycat? On talking to babies and animals. *Science* 296(5572). 1435.
- Busby, Peter A. 1979. *A classificatory study of phonemic systems in Australian aboriginal languages*. Canberra: Australian National University Unpublished Master's thesis.
- Butcher, Andrew R. 1994. On the phonetics of small vowel systems: Evidence from Australian languages. In *Proceedings of the 5th SST*, vol. 1, 28–33. Canberra: ASSTA.

- Butcher, Andrew & Victoria Anderson. 2008. The vowels of Australian Aboriginal English. *Proceedings of InterSpeech, the 9th Annual Conference of the International Speech Communication Association*, 347–350. Australasian Speech Science and Technology Association (ASSTA).
- Cameron-Faulkner, Thea, Elena Lieven & Michael Tomasello. 2003. A construction-based analysis of child directed speech. *Cognitive Science* 27(6). 843–873.
- Clark, Eve V. 1993. *The Lexicon in acquisition*. Cambridge: Cambridge University Press.
- Cooper, Robin P. & Richard N. Aslin. 1990. Preference for infant-directed speech in the first month after birth. *Child Development* 61(5). 1584–1595.
- Cooper, William E. & Jeanne Paccia-Cooper. 1980. *Syntax and speech*. Cambridge, MA: Harvard University Press.
- Cox, Christopher, Christina Bergmann, Emma Fowler, Tamar Keren-Portnoy, Andreas Roepstorff, Greg Bryant & Riccardo Fusaroli. 2023. A systematic review and Bayesian meta-analysis of the acoustic features of infant-directed speech. *Nature Human Behaviour* 7(1). 114–133.
- D'Amico, Simonetta, Simone Bentrovato, Manual Gasparini, Daniela Costabile & Elizabeth Bates. 2002. Timed picture naming in Italian-speaking children and adults: Differences between nouns and verbs. Talk presented at the *International Congress for the Study of Child Language and the Symposium on Research in Child Language Disorders* (IASCL/SRCLD) 2002.
- Davidson, Lucinda, Barbara Kelly, Gillian Wigglesworth & Rachel Nordlinger. 2023. Chapter 60: Language input and child directed-speech. In Claire Bown (ed.), *The Oxford guide to Australian languages*, 697–703. Oxford: Oxford University Press.
- Donegan, Patricia. 2002. Normal vowel development. In Martin J. Ball & Fiona E. Gibbon (eds.), *Vowel disorders*, 1–35. Boston: Butterworth/Heinemann.
- Englund, Kjellrun T. & Dawn M. Behne. 2005. Infant directed speech in natural interaction—Norwegian vowel quantity and quality. *Journal of Psycholinguistic Research* 34(3). 259–280.
- Fernald, Anne. 1985. Four-month-old infants prefer to listen to motherese. *Infant Behavior & Development* 8(2). 181–195.
- Fernald, Anne. 1989. Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development* 60. 1497–1510.
- Fernald, Anne. 1992. Meaningful melodies in mothers' speech to infants. In Hanu Papoušek, Uwe Jürgens & Mechthild Papoušek (eds.), *Nonverbal vocal communication: Comparative and developmental approaches*, 262–282. Paris: Cambridge University Press; Editions de la Maison des Sciences de l'Homme.
- Fernald, Anne & Claudia Mazzie. 1991. Prosody and focus in speech to infants and adults. *Developmental Psychology* 27(2). 209.
- Fernald, Anne & Patricia K. Kuhl. 1987. Acoustic determinants of infant preference for motherese speech. *Infant Behavior & Development* 10(3). 279–293.
- Fernald, Anne & Thomas Simon. 1984. Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* 20(1). 104.
- Fernald, Anne, Traute Taeschner, Judy Dunn, Mechthild Papousek, Bénédicte de Boysson-Bardies & Ikuko Fukui. 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language* 16(3). 477–501.
- Fletcher, Janet & Andrew Butcher. 2014. Sound patterns of Australian languages. In Harold Koch & Rachel Nordlinger (eds.), *The languages and linguistics of Australia: A comprehensive guide*, 91–138. Berlin: De Gruyter Mouton.
- García-Sierra, Adrián, Nairán Ramírez-Esparza, Noelle Wig & Dylan Robertson. 2021. Language learning as a function of infant directed speech (IDS) in Spanish: Testing neural commitment using the positive-MMR. *Brain & Language*. <https://doi.org/10.1016/j.bandl.2020.104890>.

- Gelman, Susan A. & Twila Tardiff. 1998. Acquisition of nouns and verbs in Mandarin and English. In Eve V. Clark (ed.), *The proceedings of the twenty-ninth annual child language research forum*, 27–36. Stanford: Center for the Study of Language and Information.
- Gentner, Dedre. 1978. On relational meaning: The acquisition of verb meaning. *Child Development* 49(4). 988–998.
- Gentner, Dedre. 1982. Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Language* 2. 301–334.
- Gleitman, Lila R., Henry Gleitman, Barbara Landau & Eric Wanner. (1988). Where learning begins: Initial representations for language learning. In F. J. Newmeyer (ed.), *Language: Psychological and biological aspects*. 150–193. New York: Cambridge University Press.
- Hale, Kenneth. 1992. Basic word order in two ‘free word order’ languages. In Doris L. Payne (ed.), *Pragmatics of word order flexibility*, 63–82. Amsterdam: John Benjamins.
- Hale, Kenneth, Mary Laughren & Jane Simpson. 1995. Warlpiri. In Jacobs Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *An international handbook of contemporary research*, 2, 1430–1449. Berlin: Walter de Gruyter.
- Hall-Lew, Lauren. 2010. Improved representation of variance in measures of vowel merger. *Proceedings of Meetings on Acoustics* 9. 060002.
- Hartman, Kelly M., Nan B. Ratner & Rochelle S. Newman. 2017. Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language* 44(5). 1140–1162.
- Hay, Jennifer, Paul Warren & Katie Drager. 2006. Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics* 34(4). 458–484.
- Henrich, Joseph, Steven J. Heine & Ara Norenzayan. 2010. Most people are not WEIRD. *Nature* 466(7302). 29.
- Hilton, Courtney B., Cody J. Moser, Mila Bertolo, Harry Lee-Rubin, Dorsa Amir, Constance M. Bainbridge, Jan Simson, Dean Knox, Luke Glowacki, Elias Alemu, Andrzej Galbarczyk, Grazyna Jasienska, Mary Beth Neff, Alia Martin, Laura K. Cirelli, Sandra E. Trehub, Jinqi Song, Minju Kim, Adena Schachner, Tom A. Vardy, Quentin D. Atkinson, Amanda Salenius, Jannik Andelin, Jan Antfolk, Purnima Madhivanan, Anand Siddaiah, Caitlyn D. Placek, Gul Deniz Salali, Sarai Keestra, Manvir Singh, Scott A. Collins, John Q. Patton, Camila Scaff, Jonathan Stieglitz, Silvia Ccari Cutipa, Cristina Moya, Rohan R. Sagar, Mariamu Anyawire, Audax Mabulla, Brian M. Wood, Max M. Krasnow & Samuel A. Mehr. 2022. Acoustic regularities in infant-directed speech and song across cultures. *Nature Human Behaviour* 6. 1545–1556.
- Hirsh-Pasek, Kathy, Deborah G. K. Nelson, Peter W. Jusczyk, Kimberly W. Cassidy, Benjamin Druss & Lori Kennedy. 1987. Clauses are perceptual units for young infants. *Cognition* 26(3). 269–286.
- Huttenlocher, Janellen, Marina Vasilyeva, Heidi R. Waterfall, Jack L. Vevea & Larry V. Hedges. 2007. The varieties of speech to young children. *Developmental Psychology* 43(5). 1062–1083.
- Igarashi, Yosuke, Ken'ya Y. Nishikawa, Kuniyoshi Tanaka & Reiko Mazuka. 2013. Phonological theory informs the analysis of intonational exaggeration in Japanese infant-directed speech. *Journal of the Acoustical Society of America* 134(2). 1283–1294.
- Kalashnikova, Marina, Christopher Carignan & Denis Burnham. 2017. The origins of babytalk: Smiling, teaching or social convergence? *Royal Society Open Science* 4(8). 170306.
- Kalashnikova, Marina, Usha Goswami & Denis Burnham. 2016. Mothers speak differently to infants at-risk for dyslexia. *Developmental Science* 21. e12487.
- Katz, Gary S., Jeffrey F. Cohn & Christopher A. Moore. 1996. A combination of vocal f_0 dynamic and summary features discriminates between three pragmatic categories of infant-directed speech. *Child Development* 67(1). 205–217.

- Kidd, Evan & Rowena Garcia. 2022. How diverse is child language acquisition research? *First Language* 42(6). 703–735.
- Kitamura, Christine, Chayada Thanavishuth, Denis Burnham & Sudaporn Luksaneeyanawin. 2001. Universality and specificity in infant-directed speech: Pitch modifications as a function of infant age and sex in a tonal and non-tonal language. *Infant Behavior and Development* 24. 372–392.
- Kitamura, Christine & Denis Burnham. 1998. The infant's response to maternal vocal affect. In Carolyn Rovee-Collier, Lewis Paeff Lipsitt & Harlene Hayne (eds.), *Advances in infancy research*, vol. 12, 221–236. Stamford, CT: Ablex.
- Kitamura, Christine & Denis Burnham. 2003. Pitch and communicative intent in mother's speech: Adjustments for age and sex in the first year. *Infancy* 4(1). 85–110.
- Kroos, Christian, Rikke Bundgaard-Nielsen & Michael Tyler. Modified by Mark Antoniou. 2010. *Get measurements (PRAAT script)*. Available at: <http://markantoniou.blogspot.ca/2011/09/do-you-have-praatscript-that-can.html>.
- Kuhl, Patricia K., Jean E. Andruski, Inna A. Chistovich, Ludmilla A. Chistovich, Elena V. Kozhevnikova, Viktoria L. Ryskina, Elvira I. Stolyarova, Ulla Sundberg & Francisco Lacerda. 1997. Cross-language analysis of phonetic units in language addressed to infants. *Science* 277(5326). 684–686.
- Lam, Christa & Christine Kitamura. 2010. Maternal interactions with a hearing and hearing impaired twin: Similarities and differences in speech input, interaction, and word production. *Journal of Speech, Language, and Hearing Research* 53. 543–555.
- Lam, Christa & Christine Kitamura. 2012. Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Developmental Science* 15. 212–221.
- Laughren, Mary. 1984. Warlpiri baby talk. *Australian Journal of Linguistics* 4(1). 73–88.
- Lester, Nicholas, Balthasar Bickel, Steven Moran & Sabine Stoll. 2019. Speech rates differentiate nouns and verbs in child-surrounding and child-produced speech: Evidence from Chintang. *Proceedings of the 44th annual Boston University conference on language development, Boston, 7 November 2019 – 10 November 2019*, 280–293. Somerville, MA: Cascadilla Press.
- Lindblom, Björn. 1990. Explaining phonetic variation: A sketch of the H&H theory. In William Hardcastle, J. & Alain Marchal (eds.), *Speech production and speech modelling. NATO ASI series*, 55. Dordrecht: Springer.
- Liu, Huei-Mei, Feng-Ming Tsao & Patricia K. Kuhl. 2009. Age-related changes in acoustic modifications of Mandarin maternal speech to preverbal infants and five-year-old children: A longitudinal study. *Journal of Child Language* 36(4). 909–922.
- Liu, Huei-Mei, Patricia K. Kuhl & Feng-Ming Tsao. 2003. An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* 6. F1–F10.
- Ma, Weiyi, Roberta Michnick Golinkoff, Derek M. Houston & Kathy Hirsh-Pasek. 2011. Word learning in infant- and adult-directed speech. *Language Learning & Development* 7(3). 185–201.
- Marklund, Ellen, Ulrika Marklund & Lisa Gustavsson. 2021a. An association between phonetic complexity of infant vocalizations and parent vowel hyperarticulation. *Frontiers in Psychology* 12. 693866.
- Marklund, Ulrika, Ellen Marklund & Lisa Gustavsson. 2021b. Relationship between parents' vowel hyperarticulation in infant-directed speech and infants' phonetic complexity on the level of conversational turns. *Frontiers in Psychology* 12. 688242.
- Ménard, Lucie, Jean-Luc Schwartz, Louis-Jean Boë & Jérôme Aubin. 2007. Articulatory–acoustic relationships during vocal tract growth for French vowels: Analysis of real data and simulations with an articulatory model. *Journal of Phonetics* 35(1). 1–19.
- Monaghan, Padraic, Morten H. Christiansen & Nick Chater. 2007. The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology* 55(4). 259–305.

- Monaghan, Padraic, Nick Chater & Morten H. Christiansen. 2005. The differential role of phonological cues in grammatical categorisation. *Cognition* 96. 143–182.
- Morgan, James L., Rushen Shi & Paul Allopenna. 1996. Perceptual bases of grammatical categories. In James L. Morgan & Katherine Demuth (eds.), *From signal to syntax*, 263–283. Mahwah, NJ: Lawrence Erlbaum Associates.
- Nelson, Deborah G. K., Kathy Hirsh-Pasek, Peter W. Jusczyk & Kimberly W. Cassidy. 1989. How the prosodic cues in motherese might assist language learning. *Journal of Child Language* 16(1). 55–68.
- Nycz, Jennifer & Lauren Hall-Lew. 2013. Best practices in measuring vowel merger. *Proceedings of Meetings on Acoustics* 20. 060008.
- Odijk, Lotte & Steven Gillis. 2021. Fine lexical tuning in infant directed speech to typically developing children. *Journal of Child Language* 48(3). 591–604.
- Odijk, Lotte & Steven Gillis. 2022. Parents tune their vowels to the emergence of children's words. *Journal of Child Language* 1–20. <https://doi.org/10.1017/S0305000922000289>. <https://www.cambridge.org/core/journals/journal-of-child-language/article/parents-tune-their-vowels-to-the-emergence-of-childrens-words/811BA277305EFE294D79DFCBDA6DF0BD#> (Accessed 25th of May 2023).
- O'Shannessy, Carmel. 2004. *The monster stories: Picture stimuli for elicited production*. (Unpublished series). Nijmegen: Max Planck Institute for Psycholinguistics.
- Papoušek, Mechthild, Marc H. Bornstein, Chiara Nuzzo, Hanuš Papoušek & David Symmes. 1990. Infant responses to prototypical melodic contours in parental speech. *Infant Behavior & Development* 13(4). 539–545.
- Pegg, Judith E., Janet F. Werker & Peter J. McLeod. 1992. Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior & Development* 15(3). 325–345.
- Polka, Linda, Matthew Masapollo & Lucie Ménard. 2022. Setting the stage for speech production: Infants prefer listening to speech sounds with infant vocal resonances. *Journal of Speech, Language, and Hearing Research* 65(1). 109–120.
- Polka, Linda & Yufang Ruan. 2021. The ins and outs of baby talk. *Acoustics Today* 17(1). 26–34.
- Pye, Clifton. 1986. Quiché Mayan speech to children. *Journal of Child Language* 13(1). 85–100.
- Ratner, Nan. 1984. Patterns of vowel modification in mother–child speech. *Journal of Child Language* 11(3). 557–578.
- Sarvasy, Hannah, Jaydene Elvin, Weicong Li & Paola Escudero. 2019. An acoustic analysis of Nungon vowels in child-versus adult-directed speech. In Sasha Calhoun, Paola Escudero, Marija Tabain & Paul Warren (eds.), *Proceedings of the 19th international congress of phonetic sciences, Melbourne, Australia 2019*, 3155–3159. Canberra, Australia: Australasian Speech Science and Technology Association Inc.
- Seifart, Frank, Jan Strunk, Swintha Danielsen, Iren Hartmann, Birgitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich, Nivja H. de Jong & Balthasar Bickel. 2018. Nouns slow down speech across structurally and culturally diverse languages. *Proceedings of the National Academy of Sciences* 115(22). 5720–5725.
- Shi, Rushen. 1995. *Perceptual correlates of content words and function words in early language input*. Providence, RI: Brown University PhD Dissertation.
- Shi, Rushen, James Morgan & Paul Allopenna. 1998. Phonological and acoustic bases for earliest grammatical category assignment: A cross-linguistic perspective. *Journal of Child Language* 25. 169–201.
- Singh, Leher, James L. Morgan & Catherine T. Best. 2002. Infants' listening preferences: Baby talk or happy talk? *Infancy* 3(3). 365–394.
- Snow, Catherine E. & Charles A. Ferguson (eds.). 1977. *Talking to children: Language input and acquisition*. Cambridge: Cambridge University Press.

- Soley, Gaye & Nuria Sebastian-Galles. 2020. Infants' expectations about the recipients of infant-directed and adult-directed speech. *Cognition* 198. 104214.
- Song, Jae Yung, Katherine Demuth & James Morgan. 2010. Effects of the acoustic properties of infant-directed speech on infant word recognition. *Journal of the Acoustical Society of America* 128(1). 389–400.
- Stern, Daniel N., Susan Spieker & Kristine MacKain. 1982. Intonation contours as signals in maternal speech to prelinguistic infants. *Developmental Psychology* 18(5). 727.
- Strunk, Jan, Frank Seifart, Swintha Danielsen, Iren Hartmann, Birgitte Pakendorf, Søren Wichmann, Alena Witzlack-Makarevich & Balthasar Bickel. 2020. Determinants of phonetic word duration in ten language documentation corpora: Word frequency, complexity, position, and part of speech. *Language Documentation & Conservation* 14. 423–461.
- Swartz, Stephen. 1991. *Constraints on zero anaphora and word order in Warlpiri narrative text*. Darwin: Summer Institute of Linguistics.
- Tardiff, Twila, Susan A. Gelman & Fan Xu. 1999. Putting the “noun bias” in context: A comparison of English and Mandarin. *Child Development* 70. 620–635.
- UPSID. <https://phoible.org/contributors/UPSID#tinventories>.
- Uther, Maria, Monja A. Knoll & Denis Burnham. 2007. Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech. *Speech Communication* 49(1). 2–7.
- Vorperian, Hourii K. & Ray D. Kent. 2007. Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research* 50(6). 1510–1545.
- Werker, Janet F., Ferran Pons, Christiane Dietrich, Sachiyo Kajikawa, Laurel Fais & Shigeaki Amano. 2007. Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* 103(1). 147–162.
- Werker, Janet F. & Peter J. McLeod. 1989. Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology/Revue canadienne de psychologie* 43(2). 230–246.
- Xu, Nan, Denis Burnham, Christine Kitamura & Ute Vollmer-Conna. 2015. Vowel hyperarticulation in parrot-dog- and infant-directed speech. *Anthrozoos* 26(3). 373–380.
- Zangl, Renate & Debra L. Mills. 2007. Increased brain activity to infant-directed speech in 6- and 13-month-old infants. *Infancy* 11(1). 31–62.