Input to Language: The Phonetics and Perception of Infant-Directed Speech

Alejandrina Cristia[1,2]*

[1]Neurobiology of Language, Max Planck Institute for Psycholinguistics and [2]Laboratoire de Sciences Cognitives et Psycholinguistique, CNRS, IEC-ENS, EHESS

Abstract

Over the first year of life, infant perception changes radically as the child learns the phonology of the ambient language from the speech she is exposed to. Since infant-directed speech attracts the child's attention more than other registers, it is necessary to describe that input in order to understand language development, and to address questions of learnability. In this review, evidence from corpora analyses, experimental studies, and observational paradigms is brought together to outline the first comprehensive empirical picture of infant-directed speech and its effects on language acquisition. The ensuing landscape suggests that infant-directed speech provides an emotionally and linguistically rich input to language acquisition.

1. Introduction

Language acquisition is of long-standing interest to linguists, psychologists, and almost anyone who has ever seen a child grow up to speak fluently in just a few years. The speed with which infants learn their ambient language is nothing short of miraculous: After a mere 6 months of life *ex utero*, infants have learned to discriminate between dialects of their ambient language (Nazzi et al. 2000) and know what large prosodic units (Seidl 2007), stress patterns (Friederici et al. 2007), and vowel categories (Kuhl et al. 1992) should sound like in their language. By their first birthday, they have learned some tonal categories (Mattock and Burnham 2006), some consonantal categories (Werker and Tees 1984), and phonotactics (Jusczyk et al. 1994), and they build a protolexicon with frequent wordforms (Hallé and de Boysson-Bardies 1996).

Therefore, the child lays the foundations of her ambient language's sound system, its *phonology*, even before she utters her first word. Since each and every one of these phonological pieces of knowledge is specific to the ambient language, then the input must play a crucial role (Fernald 2000). This input will be composed of overheard, adult-directed speech as well as speech directly addressed to the child. Speech directly addressed to infants, or infant-directed speech (IDS), may be different from adult-directed speech (ADS), since talkers typically accommodate their speech depending on their conversational partner. The existence of a special register, IDS, raises a host of additional questions, such as to what extent IDS and ADS are fundamentally different, and what specific role IDS may play in shaping language acquisition. The present review focuses on these questions.[1]

Specifically, this review seeks to provide a comprehensive panorama of the acoustic-perceptual characteristics of speech addressed to prelinguistic infants. In view of the extensiveness of this literature, I focus on the aspects of IDS that are relevant to phonological acquisition; Soderstrom (2007) provides a brief overview of classical research on

other linguistic levels in IDS. In this overview, I linger over stable patterns of results and rush through cases that cannot be thus summarized (due to scarce literature or inconsistent results). In each case, I provide one or a few representative references that illustrates the point made in the main text; the interested reader can also consult the topic in the categorized, annotated bibliography, which is available for download as Supplementary Material.

The review of extant literature suggests the following four conclusions, which are expanded in the sections that follow:

1. Section 2: IDS and ADS have somewhat different acoustic profiles. Acoustic instantiation at the prosodic level is clearly different between IDS and ADS (section 2.1). Evidence on segmental instantiation is less straightforward (section 2.2).
2. Section 3: Infants attend preferentially to IDS over ADS, and in most age groups this preference is driven by prosody.
3. Section 4: IDS and IDS-like acoustic properties boost infants' processing abilities in non-linguistic tasks (e.g., learning arbitrary associations), and in linguistic tasks with a high cognitive load (e.g., a difficult word segmentation task).
4. Section 5: At least some of the acoustic characteristics of IDS respond to an interplay between the infant and her caregiver.

2. Acoustic-phonetic Characteristics of IDS

As mentioned above, a fundamental question is to what extent IDS and ADS differ. Corpora studies comparing IDS and ADS are rather varied in how they proceed. Some record infants' primary caregivers at home, others do so in the lab. In the latter case, a variety of individuals, having more or less experience with infant caretaking, have been recorded, at times in interaction with a completely unfamiliar child. Speakers are sometimes provided with 'props' to elicit specific sounds of interest. To focus on studies more likely to reflect the infants' natural linguistic environment, I have excluded those where talkers were addressing a picture, or an imagined partner, and where speech was scripted.

Prosodic differences between the registers are discussed in section 2.1; work on the segmental level (vowels, tones, and consonants) is reported in section 2.2.[2] Some basic acoustic dimensions along which IDS and ADS have been compared become evident by comparing Figure 1 and Figure 2. The spectrogram on the top suggests that /u/ in 'shoe' and /æ/ in 'sandal' may not be hyperarticulated in this particular case, (/u/: F1 and F2 have both increased by 40-50 Hz, suggesting a more front and open position, contrary to hyperarticulation predictions –see Figure 3 for further details; /æ/: little change in F1, F2 is higher by 20 Hz, suggesting a more open position that is consistent with hyperarticulation for this vowel). The pitch track in the middle panel clearly shows higher overall pitch, larger pitch changes even within single words, and that contours may be simpler. As evident from the transcription of the sentence shown on the bottom of the figure, phrases are shorter and they may be spoken more slowly. Notice additionally that, while the duration of the word phrase-medially has not changed a great deal (if anything, it is shorter in IDS than ADS), the final word has been considerably lengthened, which must be attributed to exaggerated preboundary lengthening. Incidentally, two additional acoustic-phonetic aspects of IDS have been studied but will not be covered in the present review, namely speech rate (e.g., Fernald et al. 1989) and pronunciation variation (e.g., 'did you' can also be pronounced something like 'didja'; Fais et al. 2010).
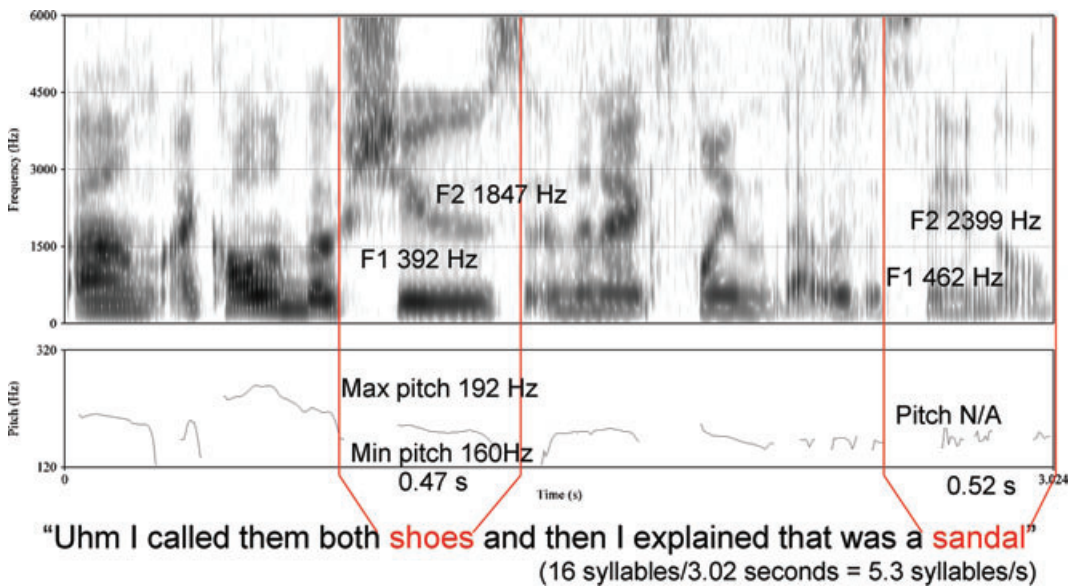
Fig 1. Acoustic analyses of a sample sentence spoken in adult-directed register. Top: A spectrogram (representation of intensity by frequency by time) of a sentence, used to measure acoustic implementation of vowels, such as their formants; for example, here F1 and F2 are noted for two words. Middle: A pitch track, which captures intonation as a function of time; here minimum and maximum pitch are noted for the same two words. Bottom: Transcription of the sentence, and calculation of speech rate.
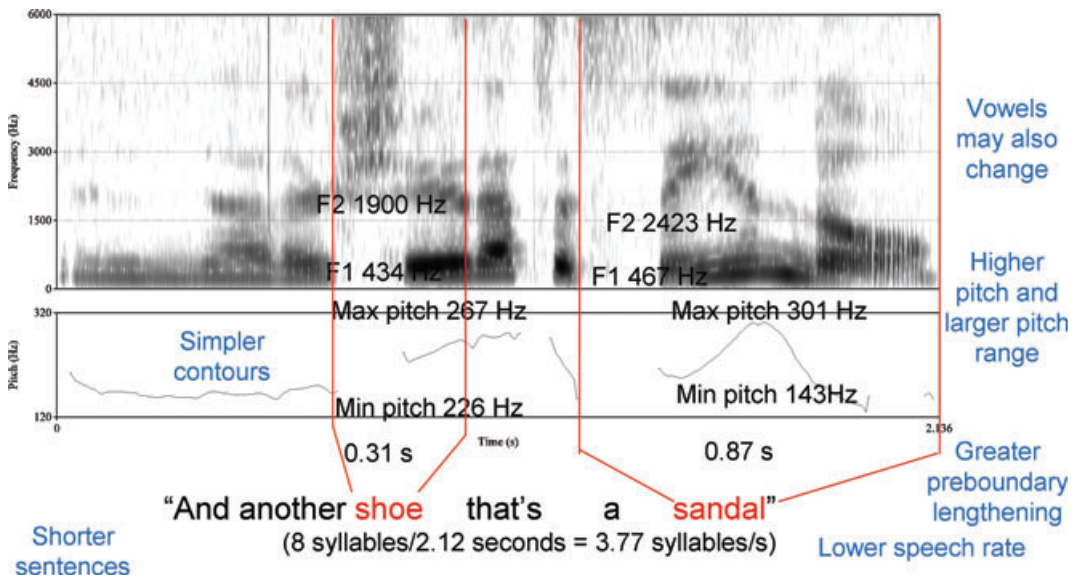
Fig 2. Acoustic analyses of a sample sentence spoken in infant-directed register uttered by the same caregiver in the same session, and comparison with Figure 1. See main text for interpretation.

2.1. SUPRASEGMENTALS IN IDS

2.1.1. General Prosodic Characteristics

The great majority of studies report that adults speak in a higher pitch in IDS across age groups, and even in tonal or pitch–accented languages (e.g., Fernald et al. 1989; Grieser
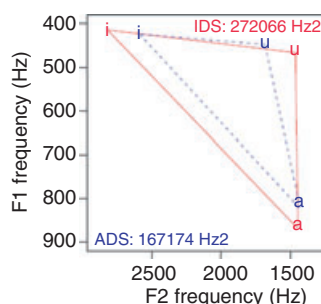
Fig 3. Vowel space size in IDS (red) and ADS (blue); the area of the triangles is given in the corners.

and Kuhl 1988). While American mothers tend to increase their pitch variability, this feature is somewhat less prevalent in other linguistic groups (e.g., Fernald et al. 1989). Several lines of work suggest that this general prosodic profile may be due to the heightened expression of *emotion*; for example, IDS and pet-directed speech display similar prosodic features in terms of pitch height and variability (Burnham et al. 2003), and little prosodic difference remains between IDS and ADS once the strength of emotion is controlled for (Trainor et al. 2000). It should be noted that this analysis focused on the strength of the emotions, and not in trying to determine whether IDS or specific IDS features were linked to specific emotions.

2.1.2. Communicative Intentions
The set of prosodic contours, or 'tunes', used in speech to infants is limited and less variable than the one found in ADS (Fernald and Simon 1984). There are also important changes in the frequency of occurrence of different contours across age groups, which is likely tied to variation in the occurrence of different communicative intents: mothers of younger infants tend to speak more comforting sentences, those of older infants utter more directives (Kitamura and Burnham 2003).

2.1.3. Prosodic Units
Acoustic cues to intonational phrase boundaries are systematically enhanced in speech by talkers of various languages (Fernald et al. 1989) addressed to infants at a variety of ages (Stern et al. 1983).

2.1.4. Informational Structure
IDS shows more consistent cues to focus, such that focused words are placed at the end of sentences and marked with exaggerated pitch peaks much more frequently than in ADS (Fernald and Mazzie 1991).

Thus, in comparison with ADS, IDS cues convey more clearly the talker's emotional state and communicative intent. In addition, large prosodic units and information structure are signaled more clearly in IDS.

2.2. SEGMENTS IN IDS

Since prosodic units are marked more carefully, may phonetic units be likewise hyperarticulated? One way of estimating hyperarticulation is by measuring vowel space size, the

acoustic area delimited by the triangle with vertices at the average location of the point vowels /i,a,u/ (e.g., Smiljianic and Bradlow 2005; note that there are other ways of measuring hyperarticulation). This is represented in Figure 3, where the vowel triangle in IDS is much larger than that in ADS, since /i/ had higher F2 (more front), /u/ had lower F2 (more back), and /a/ had higher F1 (more open). This resulted in a expansion by about 60% of the vowel space area in IDS compared to ADS (original data from Cristia 2010). Across several languages, talkers have been found to expand their vowel spaces in speech to infants (Kuhl et al. 1997). This is not a side effect of 'cuteness', since vowel triangles are not larger in the similarly cute-sounding speech to dogs and cats (Burnham et al. 2003). Since triangle expansion is also found in speech to second language learners (Uther et al. 2007), could talkers aim to facilitate language acquisition?

The answer to this question appears to be 'no.' To begin with, some papers report that the vowel space *is not* expanded in IDS compared to ADS (e.g., Englund and Behne 2005). Moreover, studies focusing on specific vocalic (e.g., /a:-a/) or consonantal contrasts (e.g., /t–d/) yields mixed results, with caregivers sometimes producing clearer contrasts between two sounds in IDS, and others in ADS (a recent discussion in Cristia 2010). To explain these null results, some researchers have proposed that caregivers only enhance phonemic contrasts at specific points in development (e.g., Bernstein Ratner 1984). I return to the question of the factors that may explain this variation below, in section 5.1.

3. Infants Prefer IDS Over ADS

In preference studies, the presentation of auditory stimuli is dependent on the behavior of the child; for example, the auditory sample will continue to be played from a speaker set to the right of the infant for as long as she orients towards it. Stimuli of two types (e.g., IDS, ADS) are then presented in alternating trials, and the experimenter measures how long the infant oriented during each of the two types. There is a sizable literature on this topic, typically revealing robust preferences for IDS over ADS (Werker and McLeod 1989). A recent meta-analysis on 34 studies carried out with infants under 9 months of age shows that this preference is greater for natural than simulated, synthesized, or filtered speech (Dunst et al. 2012). Interestingly, a preference is evinced even when when the samples are spoken in a language the child has never heard (Werker et al. 1994). Infants seem to form positive associations with individuals who address them in IDS, and negative ones with those that address them in ADS: 5-month-olds prefer to look at a silent picture of someone who has previously spoken IDS over a novel person, but they prefer a novel person over someone who addressed them in ADS (Schachner and Hannon 2011).

In addition, certain combinations of null and positive results reveal that the draw towards IDS is modulated by experience. First, newborns can actually prefer their mother's ADS over her IDS (Cooper et al. 1997; Hepper et al. 1993), which could be due to their prenatal exposure to maternal speech consisting mostly of ADS. Similarly, the preference for IDS is weaker for male voices (Werker and McLeod 1989), who tend to produce less prototypical IDS. Moreover, Dunst et al. (2012) find that effect sizes increase with age, in their meta-analysis of results on infants 0–9 months of age, and others propose the development of preference may be non-linear (Hayashi et al. 2001). Finally, some evidence suggests that infants between 6 and 13 months are listening for different things than younger or older children. For example, 6-month-olds' preference is affected to a greater extent by differences between IDS and ADS in terms of verbal

repetition, and less by differences in prosody, than other age groups (McRoberts et al. 2009). Additionally, it has been found that 7- to 9-month-olds' preference can be driven by acoustic-phonetic differences between IDS and ADS that go beyond prosody (Inoue et al. 2011).

A number of studies use highly simplified (single syllables) or synthetic stimuli to uncover which *particular* characteristics drive the preference. For example, four experiments compared listening preferences between pitch modulated (more typical of IDS) and flat pitch (more typical of ADS) that had been embedded in a synthetic voice or a frequency modulated tone. Although such designs allow a tight control on the information presented to the infant, they tend to yield variable results. Of those four studies, only one revealed a preference for the pitch modulation (Fernald and Kuhl 1987), one revealed the opposite preference (Sullivan and Horowitz 1983), and two no preference whatsoever (Colombo and Horowitz 1986). Simplifying the stimuli could alter the way in which infants process the stimuli, making it difficult to extrapolate such results to natural IDS. Nonetheless, it is clear that the preference is due to some acoustic property, not only because it can be obtained with audio–only, but also because in audiovideo conditions no preference is found for videos where the sound has been eliminated (Schachner and Hannon 2011) or replaced by music (Werker and McLeod 1989).

4. Effects of IDS on Performance

The evidence summarized in Section 3 revealed that infants attend preferentially to IDS, and in particular to IDS-like prosody (with the aforementioned caveats). It then comes as no surprise that results measuring infants' state of arousal, as well as performance in change detection or associative learning, are all boosted if they hear natural IDS segments (as in Section 3, exceptions typically involve less than ideal stimuli; see e.g., Kaplan et al. 1999). For example, in some associative learning studies, infants hear either IDS or ADS paired with the picture of a smiling face during conditioning; at test, they are presented with four trials in which the visuals are a simple checkerboard, and the auditory is the same speech segment for two trials and silence for the other two. Infants who successfully associated the smiling face to the auditory stimuli exhibit larger attention differences between the speech and silent test trials. The size of this difference, which indicates the strength of the associative learning, is much larger for IDS than ADS (Kaplan et al. 1995).

Additionally, corpora analyses reviewed in Section 2.1 show that prosody is consistently enhanced in IDS, be it to express emotion or communicative intent, or to cue large prosodic units. In line with those findings, both infants and adults are better at tracking emotion and communicative intent in IDS (e.g., Fernald 1990). Furthermore, 6-month-olds correctly parse speech into large prosodic units in IDS, but fail with ADS (Kemler Nelson et al. 1989).

Furthermore, it would be expected that IDS would boost performance in linguistic tasks with a heavy cognitive load, by virtue of enhancing the learners' attention. These predictions are largely met for demanding word learning tasks; for example, infants' long-term retention of wordforms (Singh et al. 2009), and encoding of novel words after just a few exposure trials (Ma et al. 2011) both improve with IDS.

In contrast, the corpora analyses summarized in section 2.2 revealed that sounds are not always produced as clearly in IDS as in ADS, which may render sound and word processing more difficult. Most evidence on sound processing aligns with this prediction, as adults tend to make more mistakes in sound classification when provided with IDS

than ADS samples (Kitamura et al. 2002) and infants find it harder to discriminate high-pitched vowels (Trainor and Desjardins 2002). Results from automatic speech recognizers or computer modeling of category learning show mixed results (e.g., an ADS advantage in Kirchhoff and Schimmel 2005, an IDS advantage in de Boer and Kuhl 2003). Similarly, wordform recognition studies, in which the maintenance of the superficial form of the word is crucial, yield somewhat mixed results (see a discussion in Kirchhoff and Schimmel 2005).

5. The Back-and-forth of IDS

A few pioneering studies explore the interplay between the child's language and behavior and caregivers' IDS. Much of this work is based on corpora, focusing on interactions between infant and adult to describe their turn-taking patterns or the emergence of child vocalizations. Other studies add a lab-based component, for example testing the child on her ability to discriminate between two sounds, in order to see if the child's performance relates to characteristics of her caregiver's speech.

5.1. CAREGIVERS AND INFANTS IN INTERACTION

Pitch height and contour are important in caregiver-child interaction from very early on. To begin with, these levels show convergence between the talkers, since measures in sequences of 'conversational' turns or vocalizations are correlated between caregivers and their 2- or 3-month-old infant (see, for example, Puyvelde et al. 2010). Further investigation suggests that it is usually the caregiver who does the accommodation, making their speech patterns more like the infant's immediately preceding vocalizations (McRoberts and Best 1997). Furthermore, specific prosodic patterns are used by the adult to attract and maintain the child's attention (e.g., Stern et al. 1982), although the patterns themselves, and their effectiveness, vary with the child's age (Niwano and Sugai 2002).

Indeed, as mentioned above, IDS appears to change as the infant ages. Some of the changes are easily traced back to extra-linguistic factors; this is the case of more directives being present in speech addressed to older than younger infants, who presumably do not have the motor skills necessary to execute them (Kitamura and Burnham 2003). In addition, age-related changes have been proposed to account for inconsistent reports in terms of phonological enhancement. For example, caregivers could hyperarticulate specific sounds when the child is attempting to produce them, in order to promote phonological acquisition; and/or they could speak more clearly when the child is learning words, to promote lexical acquisition. Alternatively, child age and phonological enhancement could be related through the mediation of a third factor. For example, if caregivers produce shorter sentences, more sounds are next to prosodic boundaries, and this prosodic strengthening and the generally lower speech rate could lead to vowels being pronounced in more peripheral positions (Cho 2005).

Understanding how the child may in turn shape caregiver's IDS is difficult through longitudinal studies. One recent line of work begins to shed light on how infant behavior can shape adult speech patterns using double video setups, where some natural conditions of the interaction can be experimentally manipulated. Braarud and Stormark (2008) documented that the quality (in terms of pitch height and variability) and quantity of IDS declines in desynchronization situations, when either of the interlocutors are shown an audiovideo from a different time in their conversation (where contingency cues are disrupted). A different manipulation was carried out in Smith and Trainor (2008); there,

unbeknownst to the mother, the child did not see her but saw an experimenter instead. The experimenter, responding to a cue delivered over headphones, got the baby to smile when the mother was using either a high pitch or a low pitch. Seeing these responses from her child, caregivers tended to raise their pitch significantly over the course of the study in the first manipulation, showing that the infant feedback can reinforce higher pitch. Contrastingly, no change occurred over the course of the study in the pitch-lowering condition. Thus, talkers seem to be biased against lowering their pitch when addressing their infant, a behavior that may relate to lowered pitch being perceived as threatening or dominant (see Ohala 1983 for a linguistically oriented summary of such effects).

Another line of study sheds light on how caregivers accommodate their IDS depending on their child's needs. If a child is profoundly deaf, the acoustic information of the spoken input is not available to her. Therefore, some researchers define the 'hearing age' as the amount of time that has elapsed since a child has received a cochlear implant, at which point the child will have *some* acoustic information (see Bergeson et al. 2006 for further details). This is a novel topic, on which only a few references are available, but at present it appears that caregivers tailor the prosodic profile of their IDS to the *hearing*, rather than the *chronological*, age of their child (Kondaurova and Bergeson 2011). Therefore, IDS may adapt more to linguistic development than general maturation. Additionally, some evidence suggests that vowel hyperarticulation may be less prevalent in speech to deaf children. In Lam and Kitamura (2010), a caregiver was recorded in interaction with her twins; one of them can hear and the other is hearing-impaired. She only showed an expanded vowel space for the hearing child.

This evidence could be read as an indication that caregivers are indeed producing hyperarticulated speech only if their child can profit from it, and thus specifically attempt to promote phonological development. However, I believe another interpretation should be favored. In the Lam and Kitamura (2010) study, the researchers also found increased dyadic availability between the hearing sibling and the mother, than between the hearing-impaired sibling and the mother.[3] The interpretation that it is this link with exaggerated emotional valence, rather than the caregiver's intention to promote language learning, which underlies vowel space expansion comes from a recent paper that replicates deaf/hearing differences through *online* manipulations of hearing status. Specifically, Lam and Kitamura (2012) asked caregivers to talk to their child through a closed television circuit. While the caregiver could see and hear the child perfectly well, the audibility of the caregiver's speech for the child was manipulated: Some children could not hear the caregiver at all or could barely hear her, whereas others could hear the caregiver normally. Caregivers in the latter group expanded their vowel space, whereas those in the barely audible or inaudible group did not expand their vowel space in IDS as compared to ADS. Clearly, the child's linguistic age cannot change across these conditions, casting a shadow of doubt over the phonological promotion interpretation offered above. Moreover, vowel space expansion was *not* modulated by caregivers' beliefs, because experimenters *saying* the child could or could not hear had no effect on hyperarticulation. This correlational evidence suggests that vowel space size is not purposefully controlled by caregivers, but arises as a function of the interaction between infant and caregiver. However, the physical or psychological mechanisms that sustain this correlation still need to be spelled out.

5.2. EFFECTS OF CAREGIVER SPEECH ON CHILD SPEECH, LANGUAGE, AND BEHAVIOR

Some recent work suggests that individual variation found in caregivers' speech affects the child. In a series of studies, Kaplan and colleagues have documented how experience

impacts the cognitively enhancing effects of IDS using the associative learning paradigm described in section 4. IDS spoken by very depressed mothers does not lead to improved cognitive performance, unlike IDS from non-depressed mothers (e.g., Kaplan et al. 2002). Moreover, infants who have one depressed caregiver are not able to benefit from IDS spoken by an unfamiliar talker of the depressed caregiver's gender, whereas their performance is boosted by IDS from an unfamiliar talker of the opposite gender (see, for example, Kaplan et al. 2010). Individual infants' performance correlated with measures of caregiver availability (Kaplan et al. 2009), but not with caregiver pitch variability (Kaplan et al. 2004, 2009). Unfortunately, no other prosodic or segmental parameters have been investigated in depressed IDS. Based on the work discussed in Section 2.2, one would predict that vowel space size should correlate with both infant performance and parental availability. Indeed, in support of the first half of this prediction, Liu et al. (2003) document a positive correlation between maternal vowel space size and infants' speed of habituation in a conditioned head-turn paradigm in 6-to-8- and 10-to-12-month-olds, a result that is expected if availability has a positive impact on infant performance and correlates with vowel space expansion in caregiver IDS.

In contrast to this considerable literature on cognitive effects of IDS, research on the direct effects of caregiver speech on child *language* is relatively sparse. It has been known for long that children whose parents talk to them less tend to have smaller vocabularies (Hart and Risley 1995). More recent work investigates *quality*, rather than quantity, of speech addressed to the child, reporting some positive correlations when the outcome variable is fairly proximal to the type of predictor studied in caregiver speech. For example, 5 and 13-month-olds whose caregivers produced clear, hyperarticulated /s/ were better able to discriminate /s/ from /ʃ/ than infants routinely exposed to less clear instances; performance was unrelated to vowel triangle size or prosodic characteristics of the caregivers' IDS (Cristia 2011).

These initial reports need to be augmented with additional evidence provided from independent labs, and stricter controls are needed to understand whether these correlations indicate a causal relationship, or rather are caused by a hidden variable. For example, some have interpreted Hart and Risley (1995) as suggesting a genetic advantage that is evident in both caregiver and child, and a similar argument may be made about the /s-ʃ/ results.[4] Additionally, most of this work focuses on relatively short-term, highly specific effects, seeking associations between verbal fluency in the caregiver and the child, or sibilants in the caregiver and the child. Longitudinal work that investigates cascades of effects on other linguistic levels, and long-term consequences on language acquisition, would be greatly welcome.

6. Future Directions and Conclusions

Studies mapping out the characteristics and effects of IDS are increasingly common, largely thanks to technological advances that allow cheaper and more accurate recordings and analyses. The most notable is likely the LENA device and AVA, its associated software (Paul and Paul 2011). LENA is a recording device that allows full-day recordings of children's input and output, and AVA classifies the recorded sounds in terms of their sources, quantifying amount and diversity of vocalizations uttered by the child and those around them (e.g., child of interest, child-directed speech by a female talker, overheard speech by a child talker, etc.) Thus, it becomes easier to map the interplay between caregiver and child longitudinally, using recordings of spontaneous interactions in the child's natural environment. Many of these new studies are confirming the results drawn from

small, lab–based samples that were reviewed above. This research is so new that they have not been yet published in journals, and only presentations are available to vouch for their likely results. For example, Fernald (2011) presents evidence that it is child-directed input, and not overall input, which affects language development. They measured child-directed and overheard speech through whole–day, home recordings in a large cohort of 18-month-olds. When the children were 24 months, vocabulary size and speed of word recognition were measured in the lab. Results suggest that only the child-directed speech was a significant predictor of the outcome measures, whereas the quantity of speech not addressed to the child was not related to outcomes at all.

Another exciting enterprise is the Human Speechome Project (Roy 2011), where a single child's language input and output was almost continuously monitored through videocameras and microphones between 9 and 24 months of age. By recording both auditory and visual input to language, this project can shed unprecedented light on the possible interactions between auditory and visual experience for word learning and language acquisition. As common in computational linguistics, this research has not been published in journals, but rather presented in highly competitive peer-reviewed conferences. The Speechome Project has uncovered a small, but significant association between suprasegmental characteristics of specific wordforms and age of first utterance; that is, words that were louder, longer, and marked with pitch excursions in the child's linguistic input tended to be produced by the child at an earlier age (Vosoughi et al. 2011). In addition to these prosodic factors, structural characteristics also play an important role (Vosoughi et al. 2010), such that the age of first utterance was inversely proportional to the length of the sentence the form was uttered in, the form's overall frequency, and the likelihood of being the topic (i.e., recurring within a 1-minute time window). It may be inappropriate to interpret these data as reflecting how language in general is learned, since they are based on a single child's life. However, as technology progresses, it may become feasible to apply these rich analyses to larger samples.

This review brought together evidence from corpora analyses, experimental paradigms, and longitudinal and correlational analyses on the value and characteristics of infant-directed speech. IDS is characterized by high pitch, slower speaking rate, and larger vowel space sizes, and although sound contrasts are not enunciated more clearly, they are not completely distorted either. This register has the power to command infant attention; it increases arousal, enhances cognitive performance, and facilitates long-term retention of words. Moreover, analysis of individual variation in caregivers and children suggests that both IDS quantity and quality affect infant speech perception. In the near future, we look forward to more research using richer and longer recordings of naturalistic interactions, which allow a new window on both input and output. This evidence begins to provide empirical answers to some of the questions that have long occupied students of learnability.

Short Biography

Alejandrina Cristia received her Profesorado en Letras (BA) from the Universidad Nacional de Rosario (Argentina), and her MA and PhD in Linguistics from Purdue University (USA), before doing a post-doctorate in neurolinguistics and cognitive science at the Laboratoire de Sciences Cognitives et Psycholinguistique (France). She served as a scientific staff member at the Max Planck Institute for Psycholinguistics (Netherlands), before returning to France as a researcher at the Centre National de la Recherche Scientifique. Her research combines insights and methods from linguistics, psychology, and cognitive science to shed light on the representations and processes involved in language processing

and acquisition. She specializes on young infants' speech perception abilities, and what aspects of the input shape normal language development.

Supporting Information

Additional supporting information may be found in the online version of this article:

Data S1. Annotated database of studies on the phonetics and perception of infant-directed speech, compared to adult-directed speech.

Notes

* Correspondence address: Alejandrina Cristia, 29, rue d'Ulm, 75005 Paris, France. E-mail: alecristia@gmail.com

[1] There have been claims in the literature of non-universality of IDS, although close inspection often reveals that such statements are based on scattered and contradictory evidence; see Fernald et al. 1989 for a discussion.

[2] Admittedly, prosodic characteristics, communicative intentions, etc. are not purely suprasegmental, since they affect segmental cues as well. For example, prosodic units are coded not only through pitch patterns but also in the length of syllabic nuclei next to a boundary (Cho 2005).

[3] Availability is a dyadic construct in psychology: coding focuses primarily on caregiver behavior, but also includes some items for child response (e.g., child involvement of parent in play; see Kaplan et al. 2009, pp. 148–149 for a short description of one test).

[4] Stromswold (2001) summarizes a range of research suggesting that genetic contributions to each and every domain of language competence and use (with the possible exception of phonology) are high. Some of the studies reviewed here and in section 6 cannot be accounted for by heritability; for example, only the quantity of IDS (not overall parental talkativeness, reflected in number of vocalizations) were predictive of vocabulary in Fernald (2011). In others, the relationship documented could be mediated by common genetic predispositions, such as /s-ʃ/ discrimination. Since no study on the effects of very early input considers heritability, and given that results for slightly older children are mixed (compare Alarcón et al. 1999; Snedeker et al. in press), any comments on my part would be premature. Nonetheless, I do hope future research will shed light on this important question.

Works Cited

Alarcón, Maricela, Robert Plomin, David W. Fulker, Robin Corley, and John C. DeFries. 1999. Molarity not modularity: multivariate genetic analysis of specific cognitive abilities in parents and their 16-year-old children in the Colorado Adoption Project. Cognitive Development 14(1). 175–93.

Bergeson, Tonya R., Rachel J. Miller, and Kasi McCune. 2006. Mothers' speech to hearing-impaired infants and children with cochlear implants. Infancy 10. 221–40.

Bernstein Ratner, N. 1984. Patterns of vowel modification in mother–child speech. Journal of Child Language 11. 557–78.

de Boer, Bart, and Patricia K. Kuhl. 2003. Investigating the role of infant-directed speech with a computer model. Acoustics Research Letters Online 4. 2238–46.

Braarud, Hanne Cecilie, and Kjell Morten Stormark. 2008. Prosodic modification and vocal adjustments in mothers' speech during face-to-face interaction with their two- to four-month-old infants: a double video study. Social Development 17. 1074–84.

Burnham, Denis, Christine Kitamura, and Uté Vollmer-Conna. 2003. What's new, pussycat? On talking to babies and animals. Science 296. 1435.

Cho, Taehong. 2005. Prosodic strengthening and featural enhancement: evidence from acoustic and articulatory realizations of /a,i/ in English. The Journal of The Acoustical Society of America 11. 3867–78.

Colombo, John, and Francis Degen Horowitz. 1986. Infants' attentional responses to frequency modulated sweeps. Child Development 57. 287–91.

Cooper, Robin Panneton, Jane Abraham, Sheryl Berman, and Margaret Staska. 1997. The development of infants' preference for Motherese. Infant Behavior and Development 20. 476–88.

Cristia, Alejandrina. 2010. Phonetic enhancement of sibilants in infant-directed speech. The Journal of The Acoustical Society of America 128. 424–34.

——. 2011. Fine-grained variation in caregivers' /s/ predicts their infants' /s/ category. The Journal of The Acoustical Society of America 129. 3271–80.

Dunst Carl J., Ellen Gorman, and Deborah W. Hamby. 2012. Preference for infant-directed speech in preverbal young children. Center for Early Literacy Learning reviews 5. 1–13.

Englund, Kjellrun, and Dawn Behne. 2005. Infant directed speech in natural interaction: norwegian vowel quantity and quality. Journal of Psycholinguistic Research 34. 259–80.

Fais, Laurel, Sachiyo Kajikawa, Shigeaki Amano, and Janet F. Werker. 2010. Now you hear it, now you don't: vowel devoicing in Japanese infant-directed speech. Journal of Child Language 37. 319–40.

Fernald, Anne. 1990. Intonation and communicative intent in mothers' speech to infants: is the melody the message? Child Development 60. 1497–510.

——. 2000. Speech to infants as hyperspeech: knowledge-driven processes in early word recognitions. Phonetica 57(2–4). 242–54.

——. 2011. How and why early experience is so crucial in learning language. [Online]. Retrieved on 30 July 2011 from: http://www.lenafoundation.org/pdf/LENA-Conf-2011/Presentations/LENA-Conference-2011-Anne-Fernald.pdf.

Fernald, Anne, and Patricia K. Kuhl. 1987. Acoustic determinants of infant preference for motherese speech. Infant Behavior and Development 10. 279–93.

——, and Claudia Mazzie. 1991. Prosody and focus in speech to infants and adults. Developmental Psychology 27. 209–21.

——, and Thomas Simon. 1984. Expanded intonation contours in mothers' speech to newborns. Developmental Psychology 20. 104–13.

——, Traute Taeschner, Judy Dunn, Mechthild Papousek, Bénédicte de Boysson-Bardies, and Ikuko Fukui. 1989. A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. Journal of Child Language 16. 477–501.

Friederici Angela, Manuela Friedrich, and Anne Christophe. 2007. Brain responses in 4-month-old infants are already language specific. Current Biology 17. 1–4.

Grieser, DiAnne L. and Patricia K. Kuhl. 1988. Maternal speech to infants in a tonal language: support for universal prosodic features of Motherese. Developmental Psychology 24. 14–20.

Hallé, Pierre, and Bénédicte de Boysson-Bardies. 1996. The format of representation of recognized words in infants' early receptive lexicon. Infant Behavior and Development 19. 463–81.

Hart, Betty, and Todd R. Risley. 1995. Meaningful differences in the everyday experience of young American children. Baltimore, MD: Brookes.

Hayashi, Akiko, Yuji Tamekawa, and Shigeru Kiritani. 2001. Developmental change in auditory preferences for speech stimuli in Japanese infants. Journal of Speech, Language, and Hearing Research 44. 1189–200.

Hepper, Peter G., D. Scott, and Sara Shahidullah. 1993. Newborn and fetal response to maternal voice. Journal of Reproductive and Infant Psychology 11. 147–53.

Inoue, Takao, Ryuta Nakagawa, Misa Kondou, Tadashi Koga, and Kazuyuki Shinohara. 2011. Discrimination between mothers' infant- and adult-directed speech using hidden Markov models. Neuroscience Research 70. 62–70.

Jusczyk, Peter W., Paul A. Luce, and Jan Charles-Luce. 1994. Infants' sensitivity to phonotactic patterns in the native language. Journal of Memory and Language 33. 630–45.

Kaplan, Peter S., Jo-Anne Bachorowski, Moria J. Smoski, and William J. Hudenko. 2002. Infants of depressed mothers, although competent learners, fail to learn in response to their own mothers' infant-directed speech. Psychological Science 13. 268–71.

——, ——, and Patricia Zarlengo-Strouse. 1999. Child directed speech produced by mothers with symptoms of depression fails to promote associative learning in four month old infants. Child Development 79. 560–70.

——, Aaron P. Burgess, Jessica K. Sliter, and Amanda J. Moreno. 2009. Maternal sensitivity and the learning-promoting effects of depressed and nondepressed mothers' infant-directed speech. Infancy 14. 143–61.

——, Christina M. Danko, and Andres Diaz. 2010. A privileged status for male infant-directed speech in infants of depressed mothers? Role of father involvement. Infancy 15. 151–75.

——, Jessica K. Dungan, and Michael C. Zinser. 2004. Infants of chronically depressed mothers learn in response to male, but not female, infant-directed speech. Developmental Psychology 40. 140–8.

——, Michael H. Goldstein, Elizabeth R. Huckeby, and Robin Panneton Cooper. 1995. Habituation, sensitization, and infants' responses to motherese speech. Developmental Psychobiology 28. 45–57.

Kemler Nelson, Debora, Kathy Hirsh-Pasek, Peter W. Jusczyk, and K. Wright Cassidy. 1989. How prosodic cues in Motherese might assist language learning. Journal of Child Language 16. 55–68.

Kirchhoff, Katrin, and Steven Schimmel. 2005. Statistical properties of infant-directed versus adult-directed speech: insights from speech recognition. The Journal of The Acoustical Society of America 117. 2238–46.

Kitamura, Christina, and Dennis Burnham. 2003. Pitch and communicative intent in mothers' speech: adjustments for age and sex in the first year. Infancy 4. 85–110.

——, C. Thanavishuth, Dennis Burnham, and S. Luksaneeyanawin. 2002. 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. 372–92.

Kondaurova Maria V., and Tonya R. Bergeson. 2011. The effects of age and infant hearing status on maternal use of prosodic cues for clause boundaries in speech. Journal of Speech, Language, and Hearing Research 54. 740–54.

Kuhl, Patricia K., Jean E. Andruski, Inna A. Chistovich, Ludmilla A. Chistovich, Elena V.Kozhevnikova, Viktoria L. Ryskina, Elvira I. Stolyarova, Ulla Sundberg, and Francisco Lacerda. 1997. Cross-language analysis of phonetic units in language addressed to infants. Science 277. 684–6.

——, Karen A. Williams, Francisco Lacerda, Kenneth N. Stevens, and Bjorn Lindblom. 1992. Linguistic experience alters phonetic perception in infants by 6 months of age. Science 255. 606–8.

Lam, Christa, and Christina Kitamura. 2010. Maternal interactions with a hearing and hearing-impaired twin: similarities and differences in speech input, interaction quality, and word production. Journal of Speech, Language, and Hearing Research 53. 543–55.

——, and ——. 2012. Mommy, speak clearly: induced hearing loss shapes vowel hyperarticulation. Developmental Science 15. 212–21.

Liu, Huei-Mei, Patricia K. Kuhl, and Feng-Ming Tsao. 2003. An association between mothers' speech clarity and infants' speech discrimination skills. Developmental Science 6. F1–10.

Ma, Weiyi, Roberta Michnick Golinkoff, Derek Houston, and Kathy Hirsh-Pasek. 2011. Word learning in infant- and adult-directed speech. Language Learning and Development 7. 209–25.

Mattock, Karen, and Denis Burnham. 2006. Chinese and English infants' tone perception: evidence for perceptual reorganization. Infancy 10. 241–65.

McRoberts, Gerald W., and Catherine T. Best. 1997. Accommodation in mean F0 during mother–infant and father–infant vocal interactions: a longitudinal case study. Journal of Child Language 24. 719–36.

——, Colleen McDonough, and Laura Lakusta. 2009. The role of verbal repetition in the development of infant speech preferences from 4 to 14 months of age. Infancy 14. 162–94.

Nazzi, Thierry, Peter W. Jusczyk, and Elizabeth K. Johnson. 2000. Language discrimination by English-learning 5-month-olds: effects of rhythm and familiarity. Journal of Memory and Language 43. 1–19.

Niwano, Katsuko, and Kuniaki Sugai. 2002. Intonation contour of Japanese maternal infant-directed speech and infant vocal response. Japanese Journal of Special Education 39. 59–68.

Ohala, John J. 1983. Cross-language use of pitch: an ethological view. Phonetica 40. 1–18.

Paul, Terrance, and Judith Paul. 2011. LENA research foundation. More information [Online]. Retrieved on 30 July 2011 from: http://www.lenafoundation.org.

Puyvelde, Martine Van, Pol Vanfleteren, Gerrit Loots, Sara Deschuyffeleer, Bart Vinck, and Werner Verhelst. 2010. Tonal synchrony in mother-infant interaction based on harmonic and pentatonic series. Infant Behavior and Development 33. 387–400.

Roy, Deb. 2011. The Human Speechome Project. More information [Online]. Retrieved on 30 July 2011 from: http://www.media.mit.edu/cogmac/projects/hsp.html.

Schachner, Adena, and Erin E. Hannon. 2011. Infant-directed speech drives social preferences in 5-month-old infants. Developmental Psychology 47. 19–25.

Seidl, A. 2007. Infants' use and weighting of prosodic cues in clause segmentation. Journal of Memory and Language 57. 24–48.

Singh, Leher, Sarah Nestor, Chandni Parikh, and Ashley Yull. 2009. Influences of infant-directed speech on early word recognition. Infancy 14. 654–66.

Smiljanic, Rajka, and Ann R. Bradlow. 2005. Production and perception of clear speech in Croatian and English. The Journal of The Acoustical Society of America 118. 1677–88.

Smith, Nicholas A. and Laurel J. Trainor. 2008. Infant-directed speech is modulated by infant feedback. Infancy 13. 410–20.

Snedeker, Jesse, Joy Celeste Geren, and Carissa L. Shafto. in press. Disentangling age and linguistic experience: a longitudinal study of the acquisition of english in internationally-adopted children. Cognitive Psychology 65. 39–76.

Soderstrom, Melanie. 2007. Beyond babytalk: re-evaluating the nature and content of speech input to preverbal infants. Developmental Review 27. 501–32.

Stern, Daniel N., S. Spieker, R. K. Barnett, and K. MacKain. 1983. The prosody of maternal speech: infant age and context related changes. Journal of Child Language 10. 1–15.

——, Susan Spieker, and Kristine MacKain. 1982. Intonation contours as signals in maternal speech to prelinguistic infants. Developmental Psychology 18. 727–35.

Stromswold, Karin. 2001. The heritability of language: a review and metaanalysis of twin, adoption, and linkage studies. Language 77(4). 647–723.

Sullivan, Joseph W., and Francis Degen Horowitz. 1983. The effects of intonation on infant attention: the role of the rising intonation contour. Journal of Child Language 10. 521–34.

Trainor Laurel J., Caren M. Austin, and Renée N. Desjardins. 2000. Is infant-directed speech prosody a result of the vocal expression of emotion. Psychological Science 11. 188–95.

——, and Renée N. Desjardins. 2002. Pitch characteristics of infant-directed speech affect infants' ability to discriminate vowels. Psychonomic Bulletin and Review 9. 335–40.

Uther, M., M. A. Knoll, and D. Burnham. 2007. Do you speak E-NG-L-I-SH? A comparison of foreigner- and infant-directed speech. Speech Communication 49. 2–7.

Vosoughi, Soroush, Brandon C. Roy, Michael C. Frank, and Deb Roy. 2010. Contributions of prosodic and distributional features of caregivers' speech in early word learning. [Online]. Retrieved on 25 August 2011 from: http://citeseerx.ist.psu.edu/viewdoc/download? doi=10.1.1.165.9795&rep=rep1&type=pdf.

——, ——, ——, and ——. 2011. Effects of caregiver prosody on child language acquisition. [Online]. Retrieved on 26 February 2011 from: http://speechprosody2010.illinois.edu/papers/100429.pdf.

Werker, Janet F. and 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 43. 230–46.

——, Richard Tees, and Peter J. McLeod. 1994. A cross-language lnvestigation of infant preference for infant-directed communication. Infant Behavior and Development 17. 323–33.

——, and ——. 1984. Cross-language speech perception: evidence for perceptual reorganization during the first year of life. Infant Behavior and Development 7. 49–63.