Production and perception of listener-oriented clear speech in child language

KRISTEN SYRETT and SHIGETO KAWAHARA

Journal of Child Language / *FirstView* Article / November 2013, pp 1 - 17
DOI: 10.1017/S0305000913000482, Published online: 14 November 2013

J. Child Lang., Page 1 of 17.
doi:10.1017/S0305000913000482

BRIEF RESEARCH REPORT

Production and perception of listener-oriented clear speech in child language*

KRISTEN SYRETT

Rutgers University – New Brunswick

AND

SHIGETO KAWAHARA

Keio University, Japan

(*Received* 16 *July* 2012 – *Revised* 27 *January* 2013 – *Accepted* 8 *October* 2013)

ABSTRACT

In this paper, we ask whether children are sensitive to the needs of their interlocutor, and, if so, whether they – like adults – modify acoustic characteristics of their speech as part of a communicative goal. In a production task, preschoolers participated in a word learning task that favored the use of clear speech. Children produced vowels that were longer, more intense, more dispersed in the vowel space, and had a more expanded F0 range than normal speech. Two perception studies with adults showed that these acoustic differences were perceptible and were used to distinguish normal and clear speech styles. We conclude that preschoolers are sensitive to aspects of the speaker–hearer relationship calling upon them to modify their speech in ways that benefit their listener.

[*] This work was made possible from funding provided by NIH grant HD-057699 to K.S., a postdoctoral fellowship from the Rutgers Center for Cognitive Science to K.S., and Rutgers startup grants to K.S. and S. K. We received invaluable research assistance from Lara Greenberg, Melanie Kelliher, Sarah Korostoff, Michelle Marron, and Georgia Simon, and are grateful for participation from the children and staff at Pennington Presbyterian Nursery School, Pine Grove Cooperative Nursery School, and the Douglass Psychology Child Study Center. Finally, this work benefited from comments from two anonymous reviewers, discussions with the audiences at RUMMIT 2010, LSA 2011, the Rutgers Psycholinguistics and Phonetics lab groups, and suggestions on data analysis from Mafuyu Kitahara, Mary Beckman, Britta Lintfert, and Jalal-eddin Al-Tamimi. Address for correspondence: Kristen Syrett, Department of Linguistics and Center for Cognitive Science, Rutgers, The State University of New Jersey – New Brunswick, 18 Seminary Place, New Brunswick, NJ 08901. e-mail: k-syrett@ruccs.rutgers.edu

INTRODUCTION

There is by now substantial empirical evidence demonstrating that adults adjust both segmental and suprasegmental properties of their speech in response to aspects of the speaker–hearer relationship. For example, relative to adult-directed speech, infant-directed speech displays fewer words per utterance, more repetitions, longer pauses, higher pitch (F0), exaggerated prosody in utterance- or phrase-final position, possibly delayed pitch accents, and a slower rate of speech (Cruttenden, 1994; Davis & Lindblom, 2001; Ferguson, 1977; Fernald, 1992, 2000; Fisher & Tokura, 1995, 1996; Lee & Davis, 2009; Lee, Davis & MacNeilage 2008; Swanson & Leonard, 1994). In addition, vowels in infant- and child-directed speech display an expanded F0 range (Fernald *et al.*, 1989; Fernald & Mazzie, 1991; Fernald & Simon, 1984; Kitahara, Nishikawa, Igarashi, Shinya & Mazuka, 2009), and an expanded vowel space (as evidenced in the point vowels: /i/, /u/, and /ɑ/) (Kuhl *et al.*, 1997; Liu, Kuhl & Tsao, 2003). While many of these characteristics appear to hold universally, there may be cross-linguistic variability (Igarashi & Mazuka, 2008; Wassink, Wright & Franklin, 2007). Child-directed speech shares many of these same qualities, again in contrast to adult-directed speech, along with the possibility of less complex syntactic structures (Foulkes, Docherty & Watt, 2005; Newport, Gleitman & Gleitman, 1977; Shockey & Bond, 1980; Snow, 1972, 1995).

Even when speaking to other adults, adults modify their speech to compensate for factors such as hearing impairment, background noise, or native speaker status. This style of speech is known as 'clear' speech, because it is the result of speakers attempting to provide their listener with more salient acoustic cues in the speech signal with the communicative goal of becoming more intelligible (Lindblom, 1990; Smiljanić & Bradlow, 2009). As with child-directed speech, 'clear' speech has been shown to differ from 'normal' speech by a number of features. For example, it may be spoken at a slower tempo, or with an increased intensity. Vowels produced in a clear speech style exhibit increased F0 maximum, F0 average, F0 range, and the vowel space is more expanded (Krause & Braida, 2004). As with infant- and child-directed speech, the addressee may benefit from such modifications (Bradlow, Torretta & Pisoni, 1996). See Smiljanić and Bradlow (2009) and Uchanski (2005) for comprehensive overviews of clear speech research.

While it is well established that adults who wish to be better understood modify the acoustic properties of their speech to take their listener into consideration, it is currently an open question how soon or in what manner the ability to produce listener-oriented clear speech emerges. There are reasons to think that this ability to produce listener-oriented clear speech might be delayed. First, the ability to modify speech style in response to considerations of the listener relies upon an ability to shift perspectives and take into account the needs of another individual. Since the work of Piaget (1951), a number of researchers have proposed that children in preschool may be subject to a level of egocentrism preventing them from successfully adopting another's perspective. Certainly, children notoriously fail at so-called 'Sally-Anne' false-belief tasks (cf. Wimmer & Perner, 1983) and 'false contents' or 'appearance-reality' tasks (cf. Gopnik & Astington, 1988; Perner Leekam & Wimmer, 1987) until at least four years of age, and often older. Second, experimental research by Redford and Gildersleeve-Neumann (2009) suggests that the ability to modulate acoustic characteristics of speech style in response to changes in discourse context may develop in a protracted and piecemeal fashion, and not fully evidence itself until after five to six years of age.

However, there are reasons to suspect that this ability may develop early. First, while children do consistently demonstrate non-adult-like performance on false-belief tasks, their performance can be improved given certain manipulations (Rubio-Fernandez & Geurts, 2013; Surian & Leslie, 1999). Moreover, poor performance on such tasks does not necessarily entail an impoverished theory of mind (Leslie, 1987; Onishi & Baillargeon, 2005).

Second, three-year-olds are able to switch perspectives in order to judge the relative size of an object and are not locked into assessing whether or not something counts as 'big' or 'tall' based on their own size. For example, they are able to judge whether an article of clothing is big for a doll, or how it compares to others in its comparison class (Ebeling & Gelman, 1994; Gelman & Ebeling, 1989). They can also shift judgments of size given contextual shift, recalibrating the standard of comparison accordingly. Moreover, when their interlocutor uses an adjective such as *full* to refer to a container filled to some degree with some substance that children would otherwise not label as *full* (because it does not meet the maximal standard of fullness), they appear to loosen the standard and allow the label to hold so that the speaker's statement is felicitous in that context (Syrett, Kennedy & Lidz, 2010).

Third, children have been shown to be sensitive to certain characteristics of their listener, such as age or social status, and alter their speech as a result. For example, preschoolers have been shown to exhibit a longer mean length of utterance (MLU) and increase their use of clausal complements when engaging with adults than when they engage with two-year-olds (Shatz & Gelman, 1973, 1977). Moreover, by four, children not only adjust their speaking style in response to the status of their interlocutor (i.e., a puppet representing different professional roles such as a doctor or nurse; Andersen, 1992, 1996; Andersen, Brizuela, DePuy & Gonnerman, 1999), but can also identify the addressee based on the register that was used by the speaker (Wagner, Greene-Havas & Gillespie, 2010). Four-year-olds also distinguish new versus given information and mark contrastive focus prosodically with greater pitch and intensity on specific constituents (Wonnacott & Watson, 2008).

Finally, there is reason to reevaluate possible evidence to the contrary coming from Redford and Gildersleeve-Neumann (2009), whose studies suggested that children do not produce acoustically speech styles that covary with the discourse context. Their 'casual' (normal) speech task elicited spontaneous speech in a free-play task in which the linguistic environment of the target word was not controlled for, and the number and distribution of tokens across the vowel space varied significantly. These tokens were also compared to words produced in a task in which children were told to use 'big girl/boy' voice – instructions that did not necessarily signal a listener-oriented purpose for the speech style. In fact, four- and five-year-olds in their 'clear speech' task produced vowels that were shorter in duration and had a lower F0 in comparison with the vowels of words produced in their 'casual speech' task. These differences may demonstrate that they did change their speech style, speaking more like 'big kids' would (for example, faster and with a lower voice), but did not necessarily do so to speak more clearly. Not surprisingly, the adult listeners in their perception task were unable to distinguish the two styles.

It thus remains an open question whether preschoolers can adjust acoustic characteristics of their speech style in order to be more intelligible for the benefit of their interlocutor. If the answer to this question is affirmative, a series of questions then arise. Are there differences between 'normal' and 'clear' speech by preschoolers? If so, along what parameters do they differ, and are these differences perceptible to adults? In this paper, we provide affirmative evidence demonstrating that this ability begins to be evidenced in preschool, with many of the same acoustic characteristics observed in adults. Moreover, differences between clear and normal speech styles in preschoolers are perceptible to adults. Our results therefore demonstrate that preschoolers are not only aware of key aspects of the speaker–hearer relationship, but can also take advantage of acoustic manipulations of their speech to accomplish their communicative goal.

EXPERIMENT 1

The aim of this study was to determine whether preschoolers distinguish between normal and clear speech in their production.

METHOD

Participants

Thirteen children (8 boys, 5 girls; range 3;8–5;6; mean 4;8) participated. Data from two children (boys) were excluded due to native language (1) and inattention during the task (1).

TABLE 1. *Stimuli used in Experiment 1 (production study)*

Vowels	Words			
front	word 1	word 2	word 3	word 4
/i/	peas	Cheese	sheep	
/ɪ/	pig	Fish	bib	
/ɛ/	bell	Bread	pen	bed
/æ/	cat	Hat	bat	
back	word 1	word 2	word 3	
/u/	moon	Spoon	boot	
/ʊ/	book	Cook	foot	
/o/	boat	Nose	phone	
/ʌ/	duck	Sun	cup	
/ɑ/	dog	Frog	(n/a)	

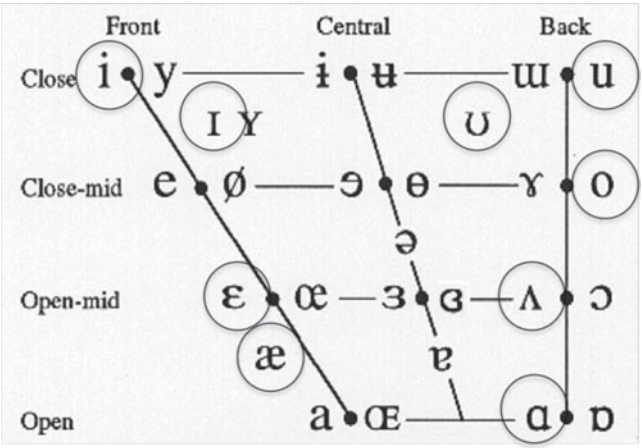

Fig. 1. Vowels represented in the stimuli for Experiment 1. The nine target vowels, situated in the vowel space, are circled. Two to four words corresponded to each vowel. Chart adapted from the IPA vowel chart (http://www.langsci.ucl.ac.uk/ipa/vowels.html).

Procedure

Stimuli consisted of twenty-seven monosyllabic words (24 CVC, 3 CCVC). The vowels represented nine vowels distributed throughout the vowel space. In most cases, there were three words per target (with the exception of /ɛ/, which had four words, and /ɑ/, which had two words; see Table 1 and Figure 1). The words chosen as stimuli are highly familiar and imageable (Wilson, 1988), and/or produced by a majority of children at age 2;6 (Dale & Fenson, 1996; see 'Appendix').

Children were recorded in two preschools. In both locations, they were tested in a quiet room on their preschool premises separate from their classroom and any activity. The recording set-up was highly similar in both locations. Two experimenters and a child were seated at a child-size table. One experimenter was in charge of directing the experimental session with the child and manipulating the technology. The second experimenter played the role of a puppet.

The experimental session took place in two parts. During the first part of the experimental session, the first experimenter explained to the child that they would view some images on a computer (a Macbook Pro laptop with a 17″ display using Powerpoint to present images on individual slides), and the child should name the picture s/he saw. This experimenter then showed the child each image sequentially and asked the child to name the object s/he saw. Whenever the child supplied a name that was not the target (e.g., *shoe* for *boot*) or did not appear to know the word, the experimenter supplied the word, and asked the child to say it (but not necessarily repeat it as the experimenter said it). This only happened on a rare occasion both across children and across images. In this way, we were sure that the child was aware of the intended label for the image and provided a baseline production for the word in the first task. If there was unexpected background noise interrupting the first production, the experimenter asked the child to say the label again. Words were presented in pseudo-randomized order, controlling for vowel features and semantic association. Sound was recorded using a portable Edirol mp3 audio recorder with a 44·1 sampling rate, placed on the table directly in front of the child.

The experimenter then explained that since the child was so good at knowing the labels for the objects, s/he was going to help the experimenter's friend Mr Rabbit learn those same words. This thus began the second part of the experimental session. The experimenter then introduced the puppet (played by the second experimenter), explaining that the child was going to see the exact same images as before, and would say the exact same words, only this time, s/he would be helping Mr Rabbit learn these words. Thus, the premise of the second session was that the child was helping the puppet learn new words. The experimenter demonstrated briefly that the puppet would benefit from the child's assistance by asking the puppet to name some article of clothing or color of clothing the child was wearing; the puppet would subsequently supply the wrong label (e.g., *shirt* for *pants, red* for *yellow*), and the child was asked to supply the right label.

The second part proceeded in a manner similar to that of the first, with the child naming the objects on the screen – this time with the puppet asking about the objects (e.g., 'What is that?', 'What do you do with that?') or making comments about the objects to elicit responses from the child (e.g., 'That looks familiar, but I don't think I know what it is'). The puppet also occasionally attempted to say the word on his own after the child supplied the label, but mispronounced it. These interactions helped to reinforce the premise of the puppet learning words, and the child's role in helping him learn.

During the task, the puppet and first experimenter periodically gave the child encouragement in order to help them remain engaged, maintain a suitable speech volume throughout the task, and keep pace. Children received comments such as 'You're doing such a good job!', 'Remember, the puppet needs your help to learn these words!', or 'Let's keep going!' The entire experimental session lasted approximately 10 minutes.

Under the direction of a professional phonetician (the second author), trained research assistants using Praat (Boersma & Weenink, 2007) excised individual words from the longer files of the individual sessions, automatically generated text grids for each sound file using a script, and annotated the segments by hand. Two additional RAs trained in acoustic analysis and the second author independently checked the annotations to ensure the boundary locations and labeling. Scripts were written and run to extract a number of key values for the vowels: average intensity, duration, F0 maximum and minimum (used to calculate F0 range), and average F0. (We exclude analysis of consonants here, for reasons of space, and because the vowels hosted the key differences between the two conditions.) We calculated the difference between the clear and normal speech styles for the relevant parameter for each minimal pair, then averaged this difference across all items. In addition, we extracted five formant values within 7,000 Hz, in a 10 ms window centered at the midpoints of the vowels to obtain F1 and F2 of the three point vowels (/i/, /u/, /ɑ/,). We then plotted these values in the vowel space, and performed a comparison of F1 and F2 between the two speech styles.

RESULTS

Our acoustic analyses (using two-tailed *t*-tests) revealed highly significant differences along all vowel parameters, holding for most or all of the child participants in each case. Average differences between clear and normal speech styles for Experiment 1 are presented in Table 2. Positive differences in the second vowel indicate that the values for the clear speech style were greater. The number of children exhibiting a positive difference is also indicated.

Vowels in the 'clear speech' style were more intense and longer, and had a higher F0, a wider F0 range, and a greater average F0 (average intensity: $t(10) = 3{\cdot}67$, $p < {\cdot}005$; duration: $t(10) = 4{\cdot}59$, $p < {\cdot}001$; maximum F0: $t(10) = 8{\cdot}99$, $p < {\cdot}001$; F0 range: $t(10) = 5{\cdot}16$, $p = {\cdot}001$; F0 average: $t(10) = 3{\cdot}75$, $p < {\cdot}005$).These differences held for most or all children. There were no correlations with either age or gender. An example of the contrast in the overall vowel quality between the 'normal' and 'clear' conditions for one child's production of the word *boot* ([but]) is presented in Figure 2.

TABLE 2. *Averages for each speech style, average difference between clear and normal minimal pairs for each word ('clear'–'normal'), and number of children displaying a positive difference for key vowel parameters in Experiment 1 (production study). Positive value in the second row indicates greater value for clear speech style*

avg. intensity (dB)	duration (ms)	max F0 (Hz)	F0 range (Hz)	F0 avg. (Hz)
66·8 v. 60·7	309·3 v. 237·7	359·8 v. 315·4	102·9 v. 67·8	303·8 v. 278·1
6·2	68·8	43·3	37·0	23·0
10/11	11/11	11/11	11/11	10/11

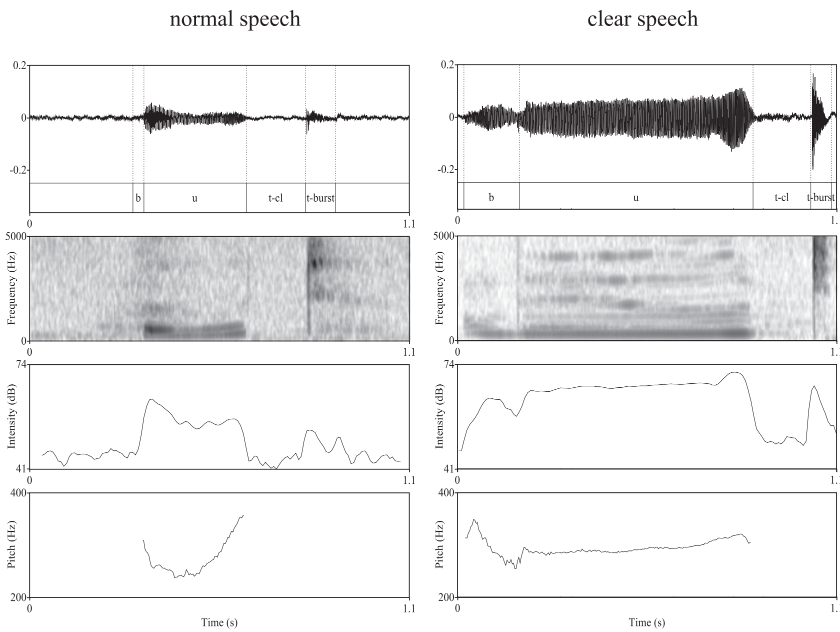

Fig. 2. Difference in overall vowel quality between two conditions of Experiment 1 for the word *boot* [but], produced by a child whose files were used in Experiments 2 and 3. The time scale is 1·1 seconds (1100 ms).

In addition, calculations of formant values indicated that the vowel space was stretched. In particular, the F1 value for /ɑ/ was significantly higher and the F2 lower, possibly indicating that children opened their mouths wider and moved their tongue further back when pronouncing these vowels. Clear and normal speech styles for three point vowels were as follows, respectively: /ɑ/(F1(1022·95 v. 884·12): $t(21)=5{\cdot}64$, $p<{\cdot}001$; F2(1497·50 v. 1593·53): $t(21)=-1{\cdot}96$, $p={\cdot}06$); /i/ (F1(506·40 v. 520·84): $t(32)=-0{\cdot}6$,

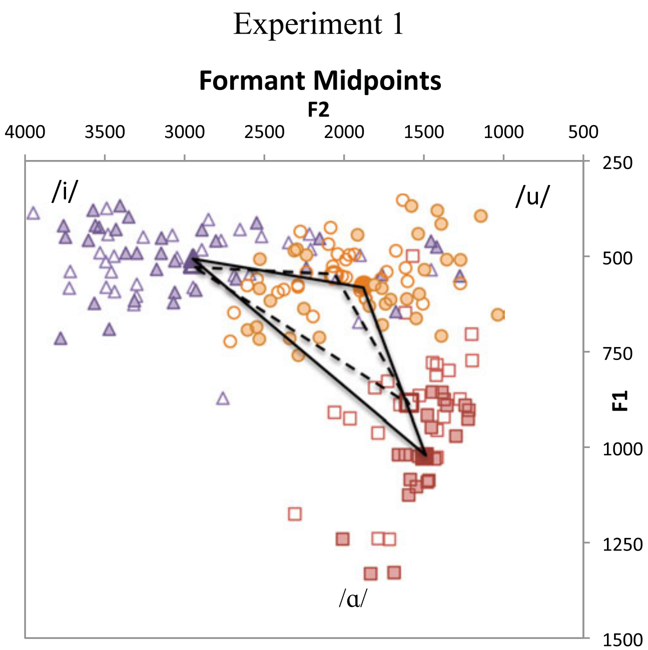

Fig. 3. Averages of formant midpoints for three point vowels (normal speech: no fill; clear speech; fill) produced by child participants in Experiment 1. The dark line represents clear speech; the dotted line represents normal speech.

$p = \cdot 55$; F2(2946·67 v. 2947·46): $t(32) = -0\cdot 01$, $p = \cdot 99$); /u/ (F1(577·18 v. 550·06): $t(32) = 1\cdot 42$, $p = \cdot 17$; F2(1877·30 v. 2064·48): $t(32) = -1\cdot 9$, $p = \cdot 07$; see Figure 3).

DISCUSSION

The results of Experiment 1 demonstrate that children produce distinct speech styles, depending on the task in which they are engaged, and the communicative goals of that task. When they are part of a word learning task that calls upon them to be more intelligible for the benefit of the listener, they produce vowels that are significantly different from those produced when they were asked to simply list the words. Much like infant- and child-directed speech and clear speech produced by adults, the vowels in the preschoolers' 'clear speech' were longer, more intense, had a higher F0 and wider F0 range, and a greater average F0. Given these positive findings demonstrating children's ability to produce clear speech resembling that of adults for the benefit of their interlocutor, we now ask whether the difference between the 'normal' and 'clear' styles is perceptible to adults.

EXPERIMENT 2

The aim of Experiments 2 and 3 was to determine whether the differences between preschoolers' clear and normal speech are perceptible to adults. Experiment 2 was an identification/categorization task using individual words. Experiment 3 was a two-alternative forced-choice task, in which participants identified the order of pairs of 'normal' and 'clear' versions of each word.

METHOD

Participants

Thirty-five adults (all native speakers of English) participated in return for extra credit in an undergraduate Linguistics course. Data from two participants were excluded, because of non-native English status (1) or a bias to respond 'clear' to all tokens (1).

Procedure

All twenty-seven sound files from six children (2 boys, 4 girls; range: 3;8–5;6; M: 4;10) in Experiment 1 were selected for presentation to adults in two perception experiments: Experiments 2 and 3. These children were chosen somewhat at random, taking into consideration age, gender, and the clarity of the production of their tokens.

Participants were tested in a quiet room, seated individually at a testing station with an iMac and high-quality headphones. Stimuli were presented using SuperLab stimulus presentation software (v. 4.2) and a response pad. In both perception experiments, sound files were blocked by children, with a short break between each block. Files were randomized by SuperLab within each block. Adults were randomly assigned to one of two presentation orders. There was no effect of order. Taking into account a small number of sound files that could not be used because of background noise in one or the other minimal pair member ($n = 12$), the entire session for each perception experiment involved 312 sound files. The experiment lasted less than 30 minutes.

Participants were told that they would hear words spoken by preschoolers, and that some would be spoken in a normal speech style, and others spoken in a clear speech style. They were instructed to identify the speech style of the individual words they heard, given the choices 'clear' and 'normal'. A short practice session preceded the test session, during which participants heard a series of examples of both speech styles spoken by children who were not in the test session.

RESULTS

Using signal detection theory (Macmillan & Creelman, 2005), which is a measure of sensitivity, hits and false alarm rates were used to calculate a d-prime score for each participant (d′ = (z(Hits)–z(False Alarms); scores can range from 0 (no detectability) to roughly 4.) The average d′ scores for all blocks ranged from 0 to 2·03 (median: 1·09, mean: 1·04). The mean was significantly above 0 ($t(32) = 11{\cdot}97$, $p < {\cdot}001$), indicating that the two speech styles were perceptually distinguishable (above chance level). Average d′ scores for the individual children ranged from 0·84 to 1·66. There was no correlation between child age and d′ score ($r^2 = {\cdot}009$).

DISCUSSION

The results of this experiment demonstrate that the acoustic differences between preschoolers' normal and clear speech were perceptible to adult listeners, as adults used these differences to distinguish between the two speech styles. In Experiment 3, we sought to determine whether the juxtaposition of the two speech styles for each word facilitated discrimination.

EXPERIMENT 3

METHOD

Participants

Twenty-nine adults (all native speakers of English) participated in return for extra credit. The data from one participant were excluded, because of non-native English status (1).

Procedure

As before, sound files were blocked by child, and sound files randomly presented within each block. This time, participants all saw the same block order. Because of a technical error with the final block of stimuli, only the first five of the six children were included in the analysis.

In this experiment, participants were presented with both 'clear' and 'normal' members of a minimal pair for each word. A Praat script was written to automatically concatenate the two tokens for each word spoken by each child in two orders ('normal > clear' and 'clear > normal') with 250 ms of silence in between tokens, thereby doubling the number of sound files from all six children. Participants were asked to choose the order of the two speech styles. The two choices were presented on the computer screen, always on the same sides: 'normal > clear' and 'clear > normal'. Within each block, there were therefore two versions of the concatenated pair (a clear–normal and a normal–clear version) for each target word.

Following the experimental session, participants were asked to report in a paper-based survey the factors they found most helpful in discriminating

TABLE 3. *Ratio and percentage of participants in Experiment 3 providing responses concerning clear/normal speech style differences and strategy for decision-making that fell into at least one of four categories*

Question	duration	stress	intensity	intonation
description of difference	19/27 70·4%	14/27 51·9%	15/27 55·6%	7/27 25·9%
self-reported strategy	16/27 59·3%	10/27 37·0%	14/27 51·9%	5/27 18·5%

speech styles. Twenty-seven of the twenty-eight participants provided responses. The motivation for this part of the task was to gain insight on the level of adults' metalinguistic awareness of the features of children's speech that contributed to 'normal' and 'clear' speech styles. Participants were asked to describe the difference between the speech styles in their own words and describe their strategy for listening to the sound files in this experiment. The experiment took approximately 30 minutes.

RESULTS

To determine whether the results of Experiment 3 were above chance (0), we divided the d′ scores by $\sqrt{2}$ to take into account the advantage of a 2AFC task ($d' = (z(H)-z(FA))/\sqrt{2}$) (Macmillan & Creelman, 2005). These values ranged from 0·33 to 1·82 (median: 1·35, mean: 1·22). The mean was significantly above 0 ($t(27) = 16{\cdot}45$, $p < {\cdot}001$), but was not significantly different that of Experiment 2 ($t(59) = 1{\cdot}55$, $p = {\cdot}13$, two-tailed). Average d′ scores divided by $\sqrt{2}$ for the individual children ranged from 0·96 to 1·77. As before, there was no correlation between child age and d′ score ($r^2 = {\cdot}048$).Thus, in both perception experiments, adults were able to successfully distinguish the speech styles.

Participants' comments on the factors influencing their decisions were transcribed by three Research Assistants who were not involved in the experiment design and analysis, and classified according to the four categories in Table 3. Apart from responses such as 'fast/slow', 'longer', 'volume', 'louder', or 'pitch', which were easily coded into a category, other responses were assigned based on relevance. For example, expressions such as 'more drawn out', and 'dragged out' were classified as 'duration'; 'quieter' or 'timid' as 'intensity'; 'enunciation', 'stressed syllables', or 'more pronounced' as 'stress'; and 'monotone' or 'went up at the end' as 'intonation'. Only two responses could not be assigned to a category, as these basically indicated that the participants listened to the words to make a decision.

DISCUSSION

The results of Experiments 2 and 3 combine to demonstrate that the acoustic differences between preschoolers' normal and clear speech were perceptible and discriminable to adult listeners. Furthermore, adults' post-experiment survey responses in Experiment 3 indicated that they were to a certain degree aware on a metalinguistic level of some of the acoustic characteristics between normal and clear speech in preschoolers. Participants indicated that they listened for factors such as vowel duration, consonant burst and aspiration, intensity, and pitch when discriminating between the two speech styles.

GENERAL DISCUSSION

We began the 'Introduction' with the question of whether preschoolers, like adults, could adjust their speech style accordingly based on the dynamics of the speaker–hearer relationship and if so, what characteristics their version of 'clear speech' would exhibit. Specifically, we asked whether they could produce clear speech, given a goal of being more intelligible to their listener. We conclude by answering these questions affirmatively. Children engaged in a word learning scenario produce words using a 'clear speech' style that is statistically distinct from their 'normal speech' style. Vowels were more intense, longer, had an expanded F0 range, and were more dispersed in the vowel space. Such differences demonstrate that children are aware at some level of the pragmatics of the context and modulate their speech accordingly. Moreover, these acoustic differences are perceptible to adults, and can be recruited to distinguish between the two speech styles. We thus conclude that even by preschool age, children are able to adjust the acoustic characteristics of their speech style for a listener-oriented communicative goal. Thus, young children are able to take into account the needs of their conversational partner and modify their speech style to accommodate these constraints accordingly.

Now, while we have shown that children DO modify their speech style in this experimental task, we cannot at this point say definitively WHY they did so. Relating our findings to the adult literature on shifting of speech styles (such as Bell, 1984. or Pickering & Garrod, 2004, 2006), we note that it is possible that children modified their style for a variety of reasons connected to their interlocutor. (See Pardo, 2012, for a review of research in this area, and McFarland, Jurafsky, & Rawlings, 2013, for related work.) Although we offer one interpretation of the findings related to their sensitivity of the listener status, the current experimental design does not allow us to pin down precisely what it was about the discourse context or the relationship between the interlocutors that led children to modify their speech style.

We think this is an exciting area for continued research, because of the connections between synchrony in dialogue, theory of mind, pragmatically informed speaker–hearer interactions, and sociophonetics. Future researchers may choose to extend investigations of this phenomenon beyond the production of single words, taking into account a broader range of listener-oriented constraints, such as speaker–hearer common ground, native language status, and language impairments and disorders. In addition, children should be invited to participate in more interactive exchanges with their conversational partners that serve to highlight the motivations for and the range of linguistic aspects targeted in speech accommodation.

REFERENCES

Andersen, E. S. (1992). *Speaking with style*. New York: Routledge.

Andersen, E. S. (1996). A cross-cultural study of children's register knowledge. In D. I. Slobin, J. Gerhardt, A. Kyratzis & J. Guo (eds.), *Social interaction, social context, and language: essays in honor of Susan Ervin-Tripp* (pp. 125–42). Mahwah, NJ: Lawrence Erlbaum.

Andersen, E. S., Brizuela, M., DePuy, B. & Gonnerman, L. (1999). Cross-linguistic evidence of the early acquisition of discourse markers as register variables. *Journal of Pragmatics* **31**, 1339–51.

Bell, A. (1984). Language style as audience design. *Language in Society* **13**, 145–204.

Boersma, P. & Weenink, D. (2007). Praat: doing phonetics by computer [computer program]. Retrieved from <http://www.praat.org/>.

Bradlow, A., Torretta, G. & Pisoni, D. (1996). Intelligibility of normal speech I: global and fine-grained acoustic-phonetic talker characteristics. *Speech Communication* **20**, 255–72.

Cruttenden, A. (1994). Phonetic and prosodic aspects of baby talk. In C. Gallaway & B. J. Richards (eds.), *Input and interaction in language acquisition* (pp. 135–52). Cambridge: Cambridge University Press.

Dale, P. S. & Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, & Computers* **28**, 125–27. Searchable database: <http://www.sci.sdsu.edu/lexical/>.

Davis, B. & Lindblom, B. (2001). Phonetic variability in baby talk and development of vowel categories. In F. Lacerda, C. von Hofsten & M. Heimann (eds.), *Emerging cognitive abilities in early infancy* (pp. 135–71). Mahwah, NJ: Lawrence Erlbaum.

Ebeling, K. S. & Gelman, S. A. (1994). Children's use of context in interpreting big and little. *Child Development* **65**, 1178–92.

Ferguson, C. A. (1977). Baby talk as a simplified register. In K. Snow & C. A. Ferguson (eds.), *Talking to children: language input and acquisition* (pp. 219–36). Cambridge: Cambridge University Press.

Fernald, A. (1992). Meaningful melodies in mothers' speech to infants. In H. Papoušek, U. Jürgens & M. Papoušek (eds.), *Nonverbal vocal communication: comparative and developmental approaches* (pp. 262–82). Cambridge: Cambridge University Press.

Fernald, A. (2000). Speech to infants as hyperspeech: knowledge-driven processes in early word recognition. *Phonetica* **57**, 242–54.

Fernald, A. & Mazzie, C. (1991). Prosody and focus in speech to infants and adults. *Developmental Psychology* **27**, 209–21.

Fernald, A. & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* **20**, 104–13.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., Boysson-Bardies, D. D. & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language* **16**, 477–501.

Fisher, C. & Tokura, H. (1995). The given–new contract in speech to infants. *Journal of Memory and Language* **34**, 287–310.
Fisher, C. & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: cross-linguistic evidence. *Child Development* **67**, 3192–218.
Foulkes, P., Docherty, G. & Watt, D. (2005). Phonological variation in child-directed speech. *Language* **81**, 177–206.
Gelman, S. A. & Ebeling, K. S. (1989). Children's use of nonegocentric standards in judgments of functional size. *Child Development* **60**, 920–32.
Gopnik, A. & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance–reality distinction. *Child Development* **59**, 26–37.
Igarashi, Y. & Mazuka, R. (2008). Exaggerated prosody in infant-directed speech? Intonational phonological analysis of Japanese infant-directed speech. In H. Chan, H. Jacob, and E. Kapia (eds.), *Proceedings of the 32nd Annual Boston University Conference on Language Development* (pp. 177–88). Somerville, MA: Cascadilla Press.
Kitahara, M., Nishikawa, K., Igarashi, Y., Shinya, T. & Mazuka, R. (2009). Characteristics of pitch accents in infant-directed speech. Presentation at the *Workshop on Prosody and Information Structure*, Shiga, Japan.
Krause, J. & Braida, L. (2004). Acoustic properties of naturally produced clear-speech normal speaking rates. *Journal of the Acoustical Society of America* **115**, 362–78.
Kuhl, P., Andruski, J., Chistovich, I., Chistovich, L., Kozhevnikova, E., Ryskina, V., Stolyarova, E., Sundberg, U. & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* **277**, 684–86.
Lee, S. & Davis, B. L. (2009). Segmental distribution patterns of English infant- and adult-directed speech. *Journal of Child Language* **37**, 767–91.
Lee, S., Davis, B. L. & MacNeilage, P. F. (2008). Segmental properties of input to infants: a study of Korean. *Journal of Child Language* **35**, 591–617.
Leslie, A. (1987). Pretense and representation: the origins of 'theory of mind'. *Psychological Review* **94**, 412–26.
Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (eds.), *Speech production and speech modelling* (pp. 403–39). Dordrecht: Kluwer Academic.
Liu, H. M., Kuhl, P. & Tsao, F. M. (2003). An association between mothers' speech clarity and infants' speech discrimination skills. *Developmental Science* **6**, F1–F10.
Macmillan, N. A. & Creelman, C. D. (2005). *Detection theory: a user's guide*. Mahwah, NJ: Lawrence Erlbaum.
McFarland, D., Jurafsky, D. & Rawlings, C. M. (2013). Making the connection: social bonding in courtship situations. *American Journal of Sociology* **118**, 1596–649.
Newport, E., Gleitman, H. & Gleitman, L. R. (1977). Mother, I'd rather do it myself: some effects and non-effects of maternal speech style. In C. Snow & C. A. Ferguson (eds.), *Talking to children: language input and acquisition* (pp. 109–149). Cambridge: Cambridge University Press.
Onishi, K. H. & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* **308**, 255–58.
Pardo, J. (2012). Reflections on phonetic convergence: speech perception does not mirror speech production. *Language and Linguistics Compass* **6**, 743–67.
Perner, J., Leekam, S. R. & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: the case for a conceptual deficit. *British Journal of Developmental Psychology* **5**, 125–37.
Piaget, J. (1951). *The child's conception of the world*. London: Routledge & Kegan Paul Ltd.
Pickering, M. J. & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* **27**, 169–225.
Pickering, M. J. & Garrod, S. (2006). Alignment as the basis for successful communication. *Research on Language and Computation* **4**, 203–28.

Redford, M. & Gildersleeve-Neumann, C. (2009). The development of distinct speaking styles in preschool children. *Journal of Speech, Language, and Hearing Research* **52**, 1434–48.
Rubio-Fernandez, P. & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science* **24**, 27–33.
Shatz, M. & Gelman, R. (1973). The development of communication skills: modifications in the speech of young children as a function of listener. *Monographs of the Society for Research in Child Development* **38** (5, Serial No. 152).
Shatz, M. & Gelman, R. (1977). Beyond syntax: the influence of conversational constraints on speech modification. In C. Ferguson & C. Snow (eds.), *Talking to children: language input and acquisition* (pp. 189–98). Cambridge: Cambridge University Press.
Shockey, L. & Bond, Z. S. (1980). Phonological processes in speech addressed to children. *Phonetica* **37**, 267–74.
Smiljanić, R. & Bradlow, A. (2009). Speaking and hearing clearly: talker and listener factors in speaking style changes. *Language and Linguistics Compass* **3**, 236–64.
Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development* **43**, 549–65.
Snow, C. E. (1995). Issues in the study of input: fine-tuning, universality, individual and developmental differences, and necessary causes. In P. Fletcher & B. MacWhinney (eds.), *The handbook of child language* (pp. 180–93). Oxford: Blackwell.
Surian, L. & Leslie, A. (1999). Competence and performance in false belief understanding: a comparison of autistic and normal 3-year-old children. *British Journal of Developmental Psychology* **17**, 141–55.
Swanson, L. & Leonard, L. (1994). Duration of function-word vowels in mothers' speech to young children. *Journal of Speech and Hearing Research* **37**, 1394–405.
Syrett, K., Kennedy, C. & Lidz, J. (2010). Meaning and context in children's understanding of gradable adjectives. *Journal of Semantics* **27**, 1–35.
Uchanski, R. M. (2005). Clear speech. In D. B. Pisoni and R. Remez (eds.), *The handbook of speech perception* (pp. 207–35). Malden, MA/Oxford: Blackwell.
Wagner, L., Greene-Havas, M. & Gillespie, R. (2010). Development in children's comprehension of linguistic register. *Child Development* **81**, 1678–86.
Wassink, A. B., Wright, R. & Franklin, A. (2007). Intraspeaker variability in vowel production: an investigation of motherese, hyperspeech, and Lombard speech in Jamaican speakers. *Journal of Phonetics* **35**, 363–79.
Wilson, M. D. (1988). The MRC psycholinguistic database: machine readable dictionary, Version 2. *Behavioural Research Methods, Instruments, & Computers* **20**, 6–11. Searchable database: <http://www.psy.uwa.edu.au/MRCDataBase/uwa_mrc.htm>.
Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* **13**, 103–28.
Wonnacott, E. & Watson, D. (2008). Acoustic emphasis in four year olds. *Cognition* **107**, 1092–101.

APPENDIX A

Words used in the experiment. Values for familiarity and imageability are from Wilson (1988). Proportion of children understanding or producing each word at age 2;6 comes from Dale and Fenson (1996); 'n/a' indicates that the data are not available.

Word	Vowel	Familiarity	Imageability	Proportion of children
cheese	/i/	588	592	100
pea(s)	/i/	524	568	78·6
sheep	/i/	507	596	82·9
bib	/ɪ/	380	488	88·6
fish	/ɪ/	548	615	95·7
pig	/ɪ/	509	635	92·9
bell	/ɛ/	543	610	n/a
bread	/ɛ/	611	619	92·9
pen	/ɛ/	554	576	n/a
bed	/ɛ/	636	635	80
bat	/æ/	514	586	74·3
cat	/æ/	582	617	92·9
hat	/æ/	580	562	94·3
boot	/u/	566	604	81·4
moon	/u/	585	585	87·1
spoon	/u/	612	584	97·1
book	/ʊ/	643	591	98·6
cook	/ʊ/	568	504	84·3
foot	/ʊ/	583	597	100
boat	/o/	584	631	95·7
nose	/o/	584	605	100
phone	/o/	550	587	n/a
cup	/ʌ/	595	558	98·6
duck	/ʌ/	529	632	95·7
sun	/ʌ/	635	639	92·9
dog	/ɑ/	598	636	97·1
frog	/ɑ/	507	617	90