

Now you hear it, now you don't: Vowel devoicing in Japanese infant-directed speech*

LAUREL FAIS

*Department of Psychology, University of British Columbia,
Vancouver, Canada*

SACHIYO KAJIKAWA

*Tamagawa University Research Institute, Tamagawa University,
Tokyo, Japan*

SHIGEAKI AMANO

NTT Communication Science Laboratories, NTT Corporation, Kyoto, Japan

AND

JANET F. WERKER

*Department of Psychology, University of British Columbia,
Vancouver, Canada*

*(Received 22 August 2008 – Revised 18 December 2008 – Accepted 18 February 2009 –
First published online 3 June 2009)*

ABSTRACT

In this work, we examine a context in which a conflict arises between two roles that infant-directed speech (IDS) plays: making language structure salient and modeling the adult form of a language. Vowel devoicing in fluent adult Japanese creates violations of the canonical Japanese consonant–vowel word structure pattern by systematically devoicing particular vowels, yielding surface consonant clusters. We measured vowel devoicing rates in a corpus of infant- and adult-directed Japanese speech, for both read and spontaneous speech, and found that the mothers in our study preserve the fluent adult form of the language and mask underlying phonological structure by devoicing vowels in infant-directed speech at virtually the same rates as those for adult-directed speech. The results highlight the complex

[*] The authors would like to thank Tomoko Kawaguchi for transcribing the data, Lorella Chan, Nozomi Kido, Amber Lo and Nicole Stringer for their hard work in coding the data, and Tracy Cassells for assistance with statistical analysis. Address for correspondence: Laurel Fais, 2136 West Mall, Vancouver, BC, V6T 1Z4, Canada. e-mail: jwlab@psych.ubc.ca.

interrelationships among the modifications to adult speech that comprise infant-directed speech, and that form the input from which infants begin to build the eventual mature form of their native language.

INTRODUCTION

Infants master their native language by virtue of exposure to language input from speakers in their linguistic communities. Whether that input provides triggers to set parameters in Universal Grammar or comprises statistically tractable distributions of language elements, a language context is essential to language learning. Some research has shown that infants do learn from overheard, adult-directed speech (ADS; Au, Knightly, Jun & Oh, 2002), but by far the most attention has been focused on the nature of speech specifically tailored for and addressed to infants: how it differs from that of speech addressed to adults, and how it facilitates language acquisition. The work reported here examines one particular aspect of infant-directed speech in Japanese, namely vowel devoicing, a process that actually masks the standard form of Japanese words. This process provides a test case for the relative importance of two basic functions of infant-directed speech: the presentation of canonical language forms and the modeling of the adult form of the language to be learned.

Background: infant-directed speech

When compared to adult-directed speech, infant-directed speech (IDS) contains exaggerated pitch contours, expanded pitch range, enlarged vowel space, longer pauses, greater amounts of repetition, shorter utterances and simplified syntax (for example, Andruski, Kuhl & Hayashi, 1999; Fernald & Simon, 1984; Fernald, Taeschner, Dunn, Papoušek, de Boysson-Bardies & Fukui, 1989; Kitamura, Thanavishuth, Burnham & Luksaneeyanawin, 2002; Kuhl *et al.*, 1997; Phillips, 1973), and has been shown to attract and hold infants' attention to speech stimuli (Fernald & Simon, 1984), regulate affective behavior (Fernald & Simon, 1984; Werker & McLeod, 1989) and facilitate language acquisition (Fernald & Simon, 1984; Thiessen, Hill & Saffran, 2005).

The intuition behind the role of IDS in making language acquisition more accessible to infants is that features of IDS, for example an expanded vowel space, make particular aspects of the speech signal to be learned more salient, in this case the point vowels in a language (Andruski *et al.*, 1999; Bernstein Ratner, 1984; Uther, Knoll & Burnham, 2007). Other aspects of IDS function similarly: simplified syntactic phrasal structure is thought to

reduce the challenge of acquiring grammatical structures (Phillips, 1973; Snow, 1972); exaggerated durational and pitch changes could make phrase boundaries salient (Fisher & Tokura, 1996); and increased pitch and amplitude might make word-medial phonetic distinctions more salient (Karzon, 1985). Note that these modifications to the standard parameters of adult-directed speech make the language material directed to and heard by infants different from that directed to adults. Further, the characteristics of IDS themselves may change over time, in step with the development of the infant, to more closely approximate adult-directed speech (Amano, Nakatani & Kondo, 2006; Bernstein Ratner, 1984; Kitamura *et al.*, 2002; Phillips, 1973). These accounts incorporate the observation that speech addressed to young infants does NOT present a transparent model of fluent, adult-directed speech, i.e. the speech form that infants must ultimately learn.

But there are contexts in which such modifications of the speech output could in theory conflict with the underlying structure of the language being learned by the infant. A case in point is the distinction between long and short vowels in Japanese. In Japanese, word meaning can be differentiated by the length of the vowel alone; e.g. /oba:san/ 'grandmother' and /obasan/ 'aunt' have two different meanings signaled only by the fact that the former has a long vowel /a:/ and the latter a short vowel /a/. If Japanese mothers lengthen the pronunciation of short vowels in IDS, perhaps making them more perceptually salient, might they also blur the distinction between the long and short versions of the vowels? In a study of productions of Japanese mothers to both their twelve-month-old infants and another adult, Werker, Pons, Dietrich, Kajikawa, Fais & Amano (2007) found that, in fact, the mothers' speech to their infants still contained two distinct, reliably identifiable categories for vowel length. Thus, this distinctive speech information is still readily available to infants in their language input, despite alterations in the speech signal typical of infant-directed speech. Though it may be the case that the average lengths of short and long vowels differ in Japanese IDS and ADS, their category status remains unchanged and salient (Werker *et al.*, 2007).

On the other hand, it is not always the case that canonical phonemic distinctions are preserved in IDS. Vowels are typically lengthened when they appear before voiced consonants in English; Bernstein Ratner and Luberoff suggested, in a study of the speech of nine mother-child pairs, that this difference in vowel length was so pronounced in the mothers' IDS, and the articulation of the following consonants so reduced, that vowel length could appear to be a more salient feature of the voiced/voiceless consonant alternation than the voicing of the consonants themselves. They proposed that this feature of IDS might be one reason why early child productions commonly consist of CV constructions, and why some children

produce vowel length alternations in place of standard consonant voicing specifications (Bernstein Ratner & Luberoﬀ, 1984). The same connection between distortions present in IDS and errors in children’s productions was drawn by Li & Thompson (1977) in their examination of the acquisition of lexical tone in Mandarin-learning children. They noted that the children’s mothers presented conflicting information regarding lexical tone, sometimes exaggerating tone for salience when teaching words for concrete objects, and otherwise obscuring lexical tone information in the articulation of IDS-typical intonation contours. The same disregard for canonical lexical tone information in the interests of maintaining the modulated, simplified F_0 contours of IDS has been demonstrated for Mandarin-speaking mothers who were role-playing talking to a young infant (Papoušek & Hwang, 1991).

The conflict between modifications made in IDS for salience and the production of well-formed language expressions is also seen in infant-directed American Sign Language (ASL) (Reilly & Bellugi, 1996). The grammatical form of *wh*-questions in ASL requires brow-furrowing in addition to a manual component; however, brow-furrowing is typical of facial expressions that convey puzzlement or anger. ASL mothers communicating with infants younger than 2;0 consistently produce ungrammatical *wh*-questions by refraining from wrinkling the eyebrow, maintaining a neutral or ‘mock surprised’ facial expression. They sacrifice the well-formedness of their linguistic output to the desire to avoid communicating anger or puzzlement to the infant.

These examples illustrate the kinds of tensions that can arise in balancing the positive functions of IDS – exaggeration of vowel length for salience, or maintaining positive affect, for example – with the communication of well-formed or canonical linguistic constructions. The work reported here examines a somewhat different, as yet unstudied, incompatibility, that is, between the presentation of a canonical form of linguistic expression and the operation of a regular phonological process within the language itself that results in apparent violations of that canonical form. This is the case of vowel devoicing in Japanese. While vowel devoicing is a systematic and tractable process in the adult language, as illustrated below, it results in acoustic output that obscures the underlying forms of the language. Examining whether Japanese IDS maintains canonical form at the expense of phonological process, or vice versa, further enhances our understanding of the subtle interplay of the various functions IDS serves. Specifically, in this case, we investigate whether Japanese IDS facilitates language learning through the use of canonical forms of expression, or represents to infants the fluent adult form of the language that they will one day be expected to master, at the expense of those canonical forms.

Background: vowel devoicing in Japanese

Japanese is a canonically consonant–vowel–consonant–vowel (CVCV) language: the underlying structure of Japanese words is primarily¹ (C)V~, in which each mora² consists of an optional initial consonant followed by a vowel (Shibatani, 1990). One consequence of this structure is that there are no initial or internal consonant clusters, and no final consonants at all in canonical Japanese forms. This constraint is so strong that Japanese adults report ‘hearing’ vowels between consonants in experimentally manipulated stimuli that actually contain no vocalic material at all (Dupoux, Kakehi, Hirose & Pallier, 1999). On the other hand, spoken Japanese also undergoes a process of vowel devoicing, in which the high vowels /i/ and /u/ occurring between voiceless consonants or before a pause may be devoiced, with frequencies ranging from about 17% to virtually 100% in fluent, casual speech, especially for particular morphemes and certain combinations of manner of articulation of the preceding and following consonants, among other factors (Maekawa & Kikuchi, 2005; Nagano-Madsen, 1994).³ The end result of this process, then, is the occurrence of acoustic consonant clusters and final consonants that obscure the canonical (C)V~ form of Japanese words in actual speech productions.

This phenomenon is quite well documented for adult speech. Effects of speaking rate, accent position, dialect, position within morphemes and the manner of the surrounding consonants have been investigated (Kondo, 2005; Maekawa & Kikuchi, 2005; Sugito, 2005; Tsuchida, 1997). However, very little is known about the rate of devoicing in speech to Japanese infants.

If we return to the intuition that one way IDS functions to facilitate language acquisition is by making particular features of language structure salient, we might hypothesize that mothers DECREASE the amount of vowel devoicing in their speech to infants, making the vowels more salient, and, in the process, highlighting the canonical phonotactics of Japanese words. In fact, this is exactly what has been found for the speech of teachers to hearing-impaired children (Imaizumi, Hayashi & Deguchi, 1998). Imaizumi and colleagues measured the rate of vowel devoicing in the speech of professional teachers when engaged in a picture search task with

[1] The canonical form is actually (C)V(N); that is, a vowel may be followed by a syllabic nasal. However, we will not be concerned with structures of this type, and omit their consideration to simplify the discussion.

[2] The ‘mora’ functions in Japanese in much the same way as the ‘syllable’ in English. However, there are some phenomena in Japanese – geminate consonants, long vowels and syllabic nasals – which necessitate making a distinction between syllable and mora. These phenomena will not affect the discussion in this paper, however.

[3] Vowels other than high vowels devoice as well, and vowels are also sometimes devoiced in contexts containing voiced consonants. However, these types of devoicing occur at a much lower rate (Maekawa & Kikuchi, 2005) and will not be considered here.

hearing-impaired or normal-hearing children. The teachers' devoicing rates were also measured when reading a list of sentences. The lowest rate of vowel devoicing (i.e. that condition in which the teachers most often VOICED high vowels that occurred in devoicing contexts) was found in the dialogues with hearing-impaired children, and the highest rate of devoicing was found in read speech. Perhaps mothers decrease their vowel devoicing in order to produce clearly articulated vowels for their infants much as these teachers did for their students.

On the other hand, we might hypothesize that in the case of vowel devoicing, caregivers speaking to infants MAINTAIN crucial adult-like productions and variability, faithfully representing to infants the mature form of the language. This possibility receives support from a study of utterance-final vowel devoicing in French speakers talking to native and non-native speakers of French (Smith, 2006). Though the intuition may be that native French speakers would accommodate a lower proficiency in the understanding of French by voicing final vowels when speaking to non-native listeners, in fact the speakers devoiced these vowels just as much to non-native as to native listeners, thus accurately representing the fluent form of the language to the non-native listeners. In a similar way, Thai mothers do not exaggerate pitch contours in speech to their infants as much as do mothers speaking Australian English; rather, they preserve Thai lexical tone information in IDS⁴ (Kitamura *et al.*, 2002).

We sought to determine whether Japanese mothers speaking to normally developing infants actually do devoice vowels less frequently to present words in their standard form, or if they instead maintain the integrity of the vowel devoicing process as it occurs in the target, adult, language. We analyzed the database of Japanese IDS and ADS speech collected for a separate study (Werker *et al.*, 2007): ten mothers were recorded reading sentences in a picture book and talking about the pictures spontaneously with their twelve-month-old infants. Those same mothers were also recorded reading a list of sentences to, and interacting spontaneously with, an adult. We identified each instance of possible devoicing morae, that is, each mora that consisted of a voiceless initial consonant followed by a high vowel /i/ or /u/, followed by either a voiceless consonant or a sentence boundary (i.e. silence), that was contained in a subset of the sentences generated.⁵ We examined the vowel in each such mora to determine whether it was voiced or voiceless and labeled it as such. In this way, we were able to

[4] Notice that, in this way, they differ from the Mandarin mothers discussed above (Li & Thompson, 1977; Papoušek & Hwang, 1991).

[5] The Werker *et al.* corpus was designed to elicit production of a set of sixteen nonsense words in order to examine the production of long and short vowels in Japanese IDS and ADS. The subset analyzed here consisted of all the sentences in which these nonsense words appeared.

calculate devoicing rates for high vowels in both read and spontaneous speech, addressed to both infants and adults.

METHODS

Participants

Participants were ten monolingual Japanese-speaking mothers (Mean age = 31 years, range = 25–37, $SD = 4;2$) accompanied by their twelve-month-old infants (all healthy, first-born infants; five males and five females, Mean age = 1;0:10, range = 0;11:20–1;0:27, $SD = 11.4$ days). The mothers had been brought up and were currently living in the eastern region of Japan, i.e. Tokyo, Kanagawa, Saitama and Chiba prefectures, where rates of vowel devoicing are said to be higher than in western dialects (Tsuchida, 1997).

Recording procedure

The mothers were told simply that we wanted to record their voices talking to different people. They were recorded in three different contexts: each mother first sat in a comfortable chair with her infant and read and talked about a constructed picture book while the experimenter monitored the recording process from another room. The picture book consisted of sixteen pairs of pictures: one pair for each of sixteen nonsense words. For each word, the mother read three sentences printed below a picture of a nonsense object. On the second page for each word, the nonsense object appeared in a real-life setting (in a garden, on a table, in a car, etc.), and the mother was instructed to talk about the object in the picture. Thus, in this setting, mothers produced both read and spontaneous speech, which was differentiated in subsequent coding. Cases in which there was any ambiguity as to whether the speech was being read or produced spontaneously were not included in the sample. In the second recording context, mothers read a similar set of sentences to the adult experimenter, except that these sentences contained more adult-oriented vocabulary in addition to the nonsense words. In the third context, mothers participated with the adult experimenter in two different types of board games that were designed to elicit spontaneous, adult-directed speech. Recordings were made on a DAT recorder at a sampling rate of 48 kHz and later resampled at a rate of 16 kHz.

General characteristics of the data

An analysis of the speech samples using Praat software (Boersma & Weenink, 2005), revealed differences typical of those found in the literature

TABLE 1. *F₀ comparisons for the IDS and ADS samples; measurements (in Hz) represent averages over all mothers*

		Maximum F ₀	Average F ₀	F ₀ variability	F ₀ range
IDS	Read	414	282	29	241
	Spontaneous	395	278	45	213
ADS	Read	359	240	49	213
	Spontaneous	376	218	20	261

for IDS and ADS speech. Japanese IDS showed a higher maximum F₀ and higher average F₀ for both read and spontaneous speech, and much greater F₀ variability in spontaneous speech, than Japanese ADS. Research comparing IDS and ADS is primarily done with samples of spontaneous speech; thus, there is no precedent for comparison to the higher variability in read ADS speech in our sample. The greater range in ADS spontaneous speech than in IDS spontaneous speech is consistent with work by Fernald and colleagues (Fernald *et al.*, 1989) who found that Japanese mothers did not expand their F₀ range when speaking spontaneously to their infants (Table 1).

Mora coding

The mothers' speech was transcribed by a native speaker of Japanese, and all morae that constituted devoicing contexts were identified, that is, all morae that consisted of a voiceless consonant followed by /i/ or /u/, which were followed by a voiceless consonant or a sentence-final boundary, and thus a period of silence. These morae were identified on the basis of the transcriptions, regardless of whether the vowels in those contexts were voiced or voiceless. A total of five coders examined all the morae identified in this way and labeled them for voicing; four of these coders were native speakers of languages other than Japanese (languages which do not permit vowel devoicing) and one was a native speaker of Japanese living in Canada and doing graduate work in English. Even though the criteria used for vowel devoicing (described below) were explicit and objective, it might still have been the case that this coder was biased toward labeling any vocalic material in these contexts as 'voiced', in the same way that the adults in studies by Dupoux and colleagues 'heard' devoiced vowels (Dupoux *et al.*, 1999). Therefore, the first 25% of her coding was checked by a trained phonetician; no evidence of interference from native speaker intuitions was found. Approximately 10% of all coding done by the remaining four coders was also checked for reliability by a second trained coder, and all

questionable cases encountered in the corpus were reviewed by at least two trained coders. The latter cases made up no more than 3% of the tokens.

The voicing of the vowel was determined by a combination of visual inspection of the waveform and spectrogram associated with the vowel, and audio check of the isolated vowel segment. For the most part, the voiced/voiceless determination was easy to make based on the presence or absence of F₀ and of the regular pulses in the waveform associated with voicing; however, a small number of tokens were labeled ‘short voicing’ (21 in IDS; 8 in ADS). These were cases in which audio checks seemed to indicate voicing, though visual inspection showed a lack of F₀, or in which only one or two voicing pulses were present in the vowel portion of the mora. These tokens were counted as ‘voiced’ in the final analysis in order to make the devoicing determination as conservative as possible. Vowels following palatalized consonants were not labeled due to difficulties in separating the voicing of the palatalization from the voicing of the vowel in the mora.

The identified morae were also labeled for position: word-initial, -medial or -final, or sentence-final. In the final analysis, the numbers of tokens occurring in some word positions were too small to allow for a reliable analysis of the effects of word position, and thus the data for all positions were collapsed.

Morpheme types

Recall that the corpus was originally collected in order to examine vowel length in Japanese IDS and ADS (Werker *et al.*, 2007) and thus contained, in addition to ‘real’ words, nonsense forms embedded in the picture book, in the adult-directed read sentences, and in the board games played by the mothers and the experimenter. The devoiceable morae in these nonsense forms were coded similarly to those in real words, but were labeled as NONSENSE and counted separately from the REAL morae. The presence of these morae, and their separation into a distinct category, allowed us to test whether adults devoice vowels similarly in familiar and unfamiliar words. In addition, some morae in particularly frequent morphemes in Japanese undergo such consistent devoicing that it has been proposed that devoiced vowels have become lexicalized in these morphemes (Maekawa & Kikuchi, 2005). The tokens of all such morphemes occurring in the corpus were also identified and the devoiceable morae they contained were labeled separately from those in both the REAL and the NONSENSE words. They consisted of the /su/ morae occurring in the polite copula *desu* ‘to be’, the polite non-past verb ending ~*masu*, and the word *suki* ‘like.’⁶ We labeled these morphemes

[6] An item-specific study of vowel devoicing in Japanese has not been done; it may be the case that other morphemes should be included in this list. However, these are the most

SPECIFIC in order to capture the intuition that, while the devoiced morae they contain undergo regular devoicing processes, they appear to do so in a significantly more consistent fashion than other morae. Another piece of evidence for the lexicalization of devoicing in these frequent morphemes comes from the examination of their voicing in non-devoicing contexts. In our corpus, even when *desu* and *~masu* are followed by the sentence-final particles *ne* ‘confirmation-seeking’ and *yo* ‘expressing assertion’ (both of which begin with voiced material, and thus place the preceding /su/ in a non-devoicing context), the /su/ morpheme is devoiced in 6 out of 7 instances for *desu* and in 24 out of 26 instances for *~masu*. Treating these morphemes as a separate group allows us both to empirically test the notion that the devoicing of these morphemes may be the result of a lexical property rather than a phonological process, and to avoid inflating devoicing rates for words in which devoicing does occur as a phonological process.

Once labeling was complete, results were summed across all mothers to yield proportions for voiced and voiceless morae occurring in devoicing contexts for four conditions: READ IDS, SPONTANEOUS IDS, READ ADS and SPONTANEOUS ADS, and three word types: REGULAR, SPECIFIC, and NONSENSE.

RESULTS

The raw data were dichotomous in nature (i.e. either the vowel was voiced or voiceless) and thus proportions were calculated for each condition. The numbers of tokens across mothers (for each condition) constituted the *N*s for the analyses and two-sample, independent proportions were conducted to determine differences by condition. We compared the amount of devoicing in each of the three categories: regular words, specific morphemes and nonsense words. *Z*-scores were computed in all three word categories for adult-directed speech and infant-directed speech in both read and spontaneous conditions. We also examined devoicing rates across the categories, comparing rates of devoicing for regular words and specific morphemes, and for regular words and nonsense words, in both IDS and ADS, and in read and spontaneous speech. Table 2 lists the numbers of morae examined in each condition.

Because some of the *N*s are large, and thus can give rise to significant, yet meaningless, results, only values that remain significant when computed assuming an *N* of 100 for each proportion are reported below as significant. An *N* of 100 (per group) was chosen as it is large enough to detect reliable

frequently suggested cases and thus the forms that we isolate here as containing ‘specific’ morae.

VOWEL DEVOICING IN JAPANESE IDS

TABLE 2. *Number of devoicing contexts, devoiced vowels and percentage of devoicing for three types of mora, in infant- and adult-directed, read and spontaneous speech*

	Devoicing contexts (#)	Devoiced vowels (#)	Devoicing rate (%)
<i>Regular words</i>			
IDS	941	788	84
ADS	847	716	85
IDS, read	455	406	89
IDS, spontaneous	486	382	79
ADS, read	268	244	91
ADS, spontaneous	579	472	82
<i>Specific morphemes</i>			
IDS	262	260	99
ADS	800	744	93
IDS, read	115	114	99
IDS, spontaneous	147	146	99
ADS, read	385	381	99
ADS, spontaneous	415	363	87
<i>Nonsense words</i>			
IDS	453	81	18
ADS	306	65	21
IDS, read	156	17	11
IDS, spontaneous	297	64	22
ADS, read	67	12	18
ADS, spontaneous	239	53	22

differences, yet not so large so as to yield meaningless significant results. *Z* values reported below are adjusted to an *N* of 100 where appropriate.

Regular words

There was no difference in the overall rates of vowel devoicing in regular words in IDS and ADS ($z=0.16$, $p=0.87$), nor were there differences between IDS and ADS read speech or between IDS and ADS spontaneous speech. However, rates of read and spontaneous speech differed significantly for both IDS and ADS. For regular words in both IDS and ADS, there was greater devoicing in read speech (ADS: $z=1.96$, $p=0.05$; IDS: $z=2.01$, $p=0.04$; Figure 1).

Specific morphemes

The appearance of specific morphemes in the corpus depended upon the mothers' use of particular lexical items and word forms; some of these

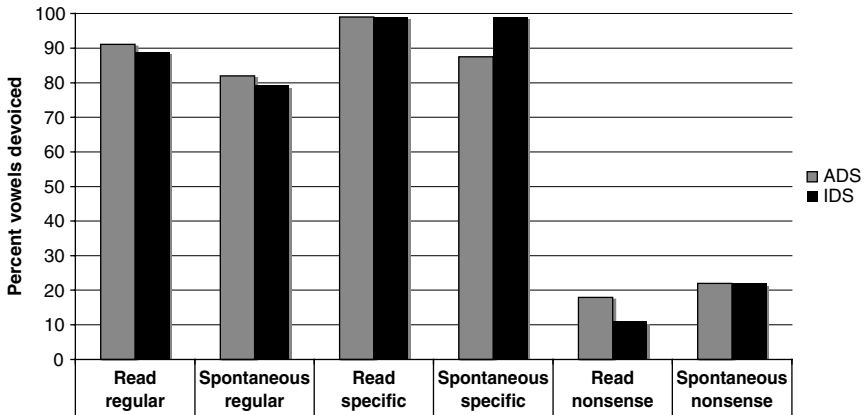


Fig. 1. Rates of vowel devoicing in regular, specific and nonsense words, in read and spontaneous speech, for ADS and IDS.

happened to have been included in the read sentences in ADS (though not in the IDS read sentences), but others occurred primarily as a result of each mother's particular style. The greater use of polite endings by the mothers to the experimenter than to their infants accounts for the greater number of *~masu* and *desu* forms in ADS than in IDS, and mothers seemed to be far more likely to talk about 'liking', *suki*, in the context of the nonsense figures (many of which were anthropomorphized shapes) in the IDS picture book task than in ADS puzzle tasks. Thus, the numbers of particular specific morphemes found in each condition are highly variable (Table 3). The overall rate for devoicing in specific morphemes was higher in IDS than in ADS ($z=2.48$, $p=0.01$), due to significantly higher devoicing rates in IDS spontaneous (but not read) speech than in ADS spontaneous (but not read) speech ($z=3.66$, $p<0.001$). In ADS, specific morphemes were devoiced at a higher rate in read speech than in spontaneous speech ($z=3.23$, $p=0.001$; Figure 1). Visual inspection of Table 3 suggests that, in general, the devoicing rate for these morphemes is virtually 100%, with two morae as exceptions to this generalization: /su/ of *~masu* and *desu* in spontaneous ADS.

Nonsense words

For nonsense words, there were no differences between IDS and ADS, or within ADS. However, devoicing rates were significantly greater in IDS spontaneous speech than read speech ($z=2.04$, $p=0.04$; Figure 1).

TABLE 3. Numbers of each mora in specific morphemes and vowel devoicing rates for all speech conditions

	Voiced (#)	Voiceless (#)	Devoicing rate (%)
IDS, read speech			
/su/ (~suki)	1	114	99
IDS, spontaneous speech			
/su/ (~masu)	1	92	99
/su/ (~suki)	0	43	100
/su/ (desu)	0	12	100
ADS, read speech			
/su/ (~masu)	3	336	99
/su/ (~suki)	1	23	96
/su/ (desu)	3	219	99
ADS, spontaneous speech			
/su/ (~masu)	52	349	87
/shi/ (~mashita)	0	14	100
/su/ (desu)	13	136	91

Regular words vs. specific morphemes

In IDS, all comparisons revealed significantly greater devoicing for specific morphemes than for regular words: overall ($z=4.07$, $p<0.001$); when read ($z=3.03$, $p=0.002$); and when spontaneous ($z=4.91$, $p<0.001$). In ADS, there was a trend toward greater devoicing overall for specific morphemes when N was lowered to 100 ($z=1.90$, $p=0.057$), and there was no difference in voicing rates for regular words and specific morphemes in spontaneous speech, though there was significantly greater devoicing of specific morphemes in read speech ($z=2.57$, $p=0.01$; Figure 2).

Regular words vs. nonsense words

All comparisons revealed significantly greater devoicing for regular words than for nonsense words: overall effects for IDS ($z=9.29$, $p<0.001$) and ADS ($z=9.01$, $p<0.001$); effects for IDS when read ($z=11.05$, $p<0.001$) and when spontaneous ($z=8.06$, $p<0.001$); and effects for ADS when read ($z=12.61$, $p<0.001$) and when spontaneous ($z=8.46$, $p<0.001$; Figure 3).

DISCUSSION

Regular words

Our results for regular words reveal that Japanese mothers do not, in fact, accommodate to infants aged 1;0 by changing the rate of vowel devoicing in their infant-directed speech. Instead, these mothers maintain a level of

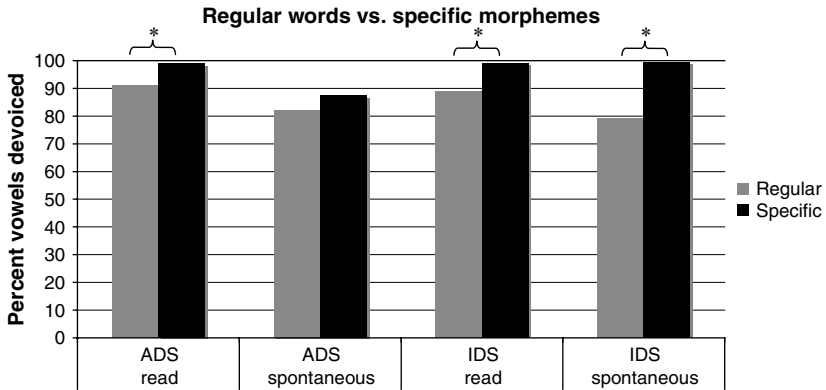


Fig. 2. Rates of vowel devoicing in regular words and specific morphemes for ADS and IDS; asterisks denote differences significant at $p < 0.02$.

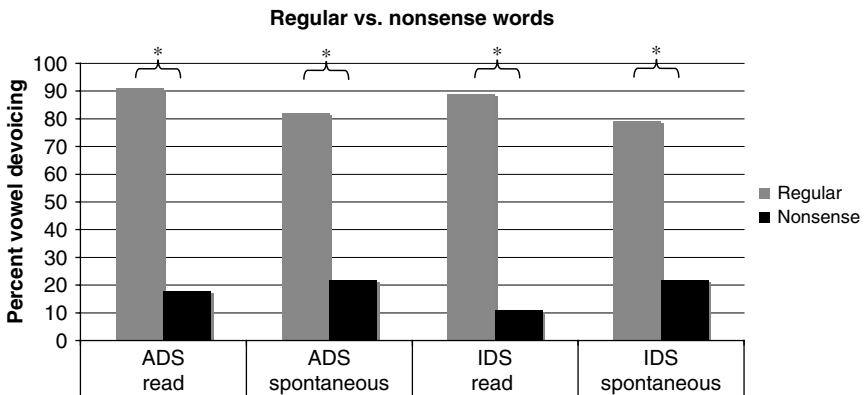


Fig. 3. Rates of vowel devoicing in regular words and nonsense words for ADS and IDS; asterisks denote differences significant at $p < 0.001$.

variability in vowel devoicing in speech to their infants similar to the level they use in speech to an adult. Thus, Japanese mothers do not attempt to manifest the acoustic properties of devoiceable vowels in the same way as the teachers of hearing-impaired children do (Imaizumi *et al.*, 1998). Of course, these situations are not strictly comparable; teachers are conscious of their students' difficulties and, because their function is specifically to help these students perceive language, they accommodate accordingly. Mothers of infants aged 1;0 in our study might have unconsciously incorporated such a 'clear speech' strategy in their speech to their infants, much

as mothers do by exaggerating intonation and vowel space, but in fact they did not. Instead, they accurately reflect fluent, adult-like devoicing variability in their speech to their infants.

One explanation for the lack of difference in devoicing rates in Japanese IDS and ADS might be that Japanese mothers, in general, do not produce as exaggerated a form of IDS as do North American mothers, and the lack of a difference in devoicing in IDS and ADS simply reflects this more thoroughgoing similarity in Japanese IDS and ADS. Some studies have shown that, unlike North American mothers, Japanese mothers do not significantly expand their pitch range in their initial vocalizations to infants, with respect to the size of the range they use to speak to adults (Fernald *et al.*, 1989; Masataka, 1992), though they do if initial utterances are unsuccessful in obtaining the infant's attention (Masataka, 1992). On the other hand, other studies HAVE shown a significant increase in both mothers' Fo and their Fo range (Amano *et al.*, 2006), especially under certain conditions: when the women were 'experienced with children' (Masataka, 2002) or had siblings themselves (Ikeda & Masataka, 1999). In addition, Japanese mothers' speech to infants does show greater Fo variation, shorter utterances and longer pauses than their speech to adults (Fernald *et al.*, 1989). Certainly the mothers in the present study 'were experienced with children', and clearly Japanese IDS, including the sample examined here, does differ from ADS along a number of prosodic and grammatical parameters, if not all. For these reasons, then, we suggest that the lack of difference in devoicing rate between Japanese IDS and ADS is not a result of a general lack of significant differences between Japanese IDS and ADS; instead we claim that it reflects the fact that Japanese mothers are producing speech to their infants that accurately incorporates the devoicing process in the adult language.

It should be no surprise that mothers speak in ways that precisely represent the target language, and yet this finding is at odds with the intuition that 'speakers adjust their productions to the needs of their listeners' (Imaizumi *et al.*, 1998: 776). Presumably, these needs would include learning the canonical (C)V~ form of Japanese morae, which is obscured by the vowel devoicing process. Yet we have seen that Japanese mothers do not adapt the devoicing process to allow the canonical form of Japanese morae to be more accessible. Studies with Japanese adults show that the high vowel /i/ in devoicing contexts in nonsense words is detected equally quickly and accurately whether the vowel is voiced or voiceless; indeed, real words produced with devoiced vowels in devoicing contexts are recognized more quickly and accurately than those words produced with voiced vowels in the same position (Ogasawara, 2005). These results are consistent with the finding that Japanese adults rate as 'better', nonsense forms containing voiceless vowels in devoicing contexts than those containing voiced vowels

in the same context (Fais, Kajikawa, Werker & Amano, 2005). By maintaining a rate of devoicing consistent with the adult rate, then, mothers are providing language input to their infants that matches what adult users of the language find more perceptually tractable.

But what do INFANTS find 'perceptually tractable?' This question has been investigated for Japanese infants aged 0;6, 1;0 and 1;6, who were tested on their ability to discriminate nonsense forms such as *keetsu* from a devoiced counterpart, *keets*. The study showed that Japanese infants aged 1;0 can, in fact, discriminate between nonsense words containing voiced and unvoiced vowels in devoicing contexts, raising the possibility that they may not recognize the forms as functionally interchangeable variants of one another, as Japanese adults do. However, by age 1;6, Japanese infants, though still discriminating the two forms, do so less robustly than do English infants for whom the two forms are not phonologically linked (Kajikawa, Fais, Mugitani, Werker & Amano, 2006; Mugitani, Fais, Kajikawa, Werker & Amano, 2007). What we do not know is whether infants are learning to prefer one variant to the other, as adults do, at this early an age. But these results do suggest that, at least by age 1;6, infants have begun to acquire the phonological process that eventually results in the perceptual advantage and preference for devoiced vowels in devoicing position seen in adults. In our present study, we have shown that infant-directed speech preserves the acoustic information that underlies this development.

Read vs. spontaneous speech

In both IDS and ADS, the mothers in this study devoiced vowels more in read speech than in spontaneous speech, a finding that is compatible with the work of Imaizumi and colleagues, in which read speech had a higher rate of devoicing than spontaneous dialogue (Imaizumi *et al.*, 1998). These results are consistent as well with more general findings concerning the speaking rates of read and spontaneous speech, and the effects of speaking rate on devoicing. Spontaneous speech is generally slower than read speech (Barik, 1977) and, more specifically, story telling is slower than story reading (Levin, Schaffer & Snow, 1982). Further, devoicing rates for slow speech tend to be lower than for fast speech (e.g. Sugito, 2005). Thus, the lower rate of devoicing for spontaneous speech may reflect the lower speed of production for spontaneous speech than for read speech. Such an explanation does not differentiate between ADS and IDS, and is supported by the fact that devoicing in ADS and IDS does in fact exhibit the same pattern of effects. Although Japanese vowel devoicing has been investigated in spontaneous speech (Maekawa & Kikuchi, 2005), in read speech (Nagano-Madsen, 1994) and under constrained experimental conditions

(Imaizumi *et al.*, 1998; Sugito, 2005), this study is the first to make a direct comparison of read and spontaneous speech by the same speakers to individuals under comparable situational constraints.

Specific morphemes

Devoicing rates for specific morphemes also support the notion that mothers are modeling adult-like devoicing processes in their speech to infants aged 1;0. In the two cases in which data exist for both IDS and ADS, the /su/ of *~masu* and of *~suki*, the devoicing rates are virtually identical, and furthermore, they are virtually 100%. The reliably greater devoicing in specific morphemes than in regular words reported here supports the suggestion that devoicing is lexicalized in specific, very frequent morphemes. Mothers present infants with evidence for the lexicalization of devoicing that occurs in the adult-directed versions of these particular morphemes by consistently devoicing them in IDS as well.

The apparent exception to this conclusion is the /su/ of *~masu* and *desu* that appears in spontaneous ADS. There is a stylistic variant of Japanese speech in which this final mora is explicitly voiced in non-past verb forms for politeness in formal situations; the fact that voiced tokens of this mora were found in our corpus only in spontaneous speech addressed to another adult in a face-to-face situation, and not in the read condition in which interactive communication is not taking place, suggests that these voiced tokens are examples of this stylistic variation (Inazuka & Inazuka, 2003). By definition, stylistic variation is susceptible to individual differences, and an examination of the use of voiced tokens by individual mothers reveals further evidence that stylistic factors are at work. There were 52 tokens of /su/ that were voiced in the polite non-past verb ending *~masu*; 23 of these tokens were produced by one mother, 16 others were produced by a second mother and the remaining 13 were distributed among five other mothers. Of the 13 voiced tokens of the /su/ in *desu*, 7 were produced by the same mother, who voiced 23 instances of the /su/ in *~masu*; another 2 were voiced by the mother who produced 16 instances of voiced /su/ in *~masu*. A third mother produced 3 tokens and one other mother the final token. If it is indeed the case, then, that voiced tokens represent individual stylistic variation on the part of at least these two mothers, and if we set aside these mothers' voiced tokens (39 tokens for *~masu* and nine tokens for *desu*), the devoicing rate in spontaneous ADS for the /su/ of *~masu* becomes 96%, and for *desu*, 97%, much closer to the near 100% rate for other specific morphemes. This suggests that, excluding some individuals' stylistic use of voiced final /su/ in polite non-past verb endings, rates for devoicing in ADS spontaneous speech are not, in fact, different from those in ADS read speech, and thus that devoicing rates of specific

morphemes in ADS and in IDS, like those of regular words, are not significantly different.

Nonsense words

Some of the bimoraic nonsense words included in this corpus for the analysis of vowel length in IDS (Werker *et al.*, 2007) contain /i/ in devoicing environments. This allowed us to examine how the process of vowel devoicing interacts with the production of unfamiliar, in this case nonsense, forms, whose perception relies on auditory information alone, unsupported by the kind of syntactic and semantic information available in the case of real, familiar words. Because the nonsense forms were not actual lexical items, the mothers producing them were likely aware that accurate perception of each nonsense word depended solely upon its acoustic properties. This awareness put the mothers in a position similar to that of the teachers of hearing-impaired children described by Imaizumi and colleagues (1998), or that of the Chinese mothers role-playing teaching foreigners described by Papoušek & Hwang (1991); that is, they might be expected to be somewhat more conscious of the effect of vowel devoicing on the perception of nonsense forms by their listeners. This consciousness is reflected in the vastly reduced level of devoicing seen for nonsense forms across all four conditions (Figure 1).

However, mothers did not reduce their devoicing rate as much in their spontaneous speech to their infants as they did in their read speech. It is not likely that this simply reflects a difference between read and spontaneous speech; as seen above, devoicing in read speech tends to be GREATER than that in spontaneous speech. Recall that in ADS, for both read and spontaneous nonsense words, the devoicing rate was low; we have suggested that this indicates that mothers were accommodating to the needs of their adult listeners to perceive the nonsense word. In the read condition of the IDS task, this accommodation was even more extreme. However, mothers significantly relaxed their attention to clarity in the spontaneous condition of the IDS task as compared to the read condition.

We suggest that the reason for this difference lies in the structure of the tasks employed in data collection. The ADS tasks were blocked by speech type: first, participants read a list of sentences, three for each nonsense word; then, following a short break, they engaged in cooperative 'games', which involved identifying to the experimenter puzzle pieces or positions that had been labeled with the nonsense words. In both reading and game-playing, we propose that the mothers were careful to make their utterances understood by reducing vowel devoicing that could obscure the perception of the nonsense form. The IDS tasks, on the other hand, were blocked by word. That is, as the mothers looked through a picture book with their

infant, they first read the three sentences that described the nonsense object bearing the nonsense label, and immediately afterwards talked in a spontaneous manner about the object as it was represented in a caption-less picture. We suggest that in the read portion of the task, mothers refrained from devoicing vowels in the nonsense words for the sake of representing the words to their infants in an intelligible fashion, just as they did for adults. This type of clarification of key lexical items has been observed in speech between adults as well as in caretakers' speech to infants (Papoušek & Hwang, 1991; Li & Thompson, 1977). However, once these forms had been accurately represented in the infant-directed read speech, mothers felt freer to incorporate them into the phonological processes that they normally apply to real word forms. A similar progression from the clear articulation of key words for concrete objects to intonational patterns that mask canonical lexical tone was observed for Mandarin-speaking caretakers of infants aged 1;6 to 3;0 (Li & Thompson, 1977). In the case of vowel devoicing, the incorporation of the nonsense forms into adult-like devoicing processes was not complete; the rate for devoicing in nonsense words in the spontaneous IDS speech is still far below that of spontaneous regular words in IDS (Figure 1). However, we suggest that the mothers' greater inclination to treat the newly 'learned' nonsense words as regular words in their subsequent use in spontaneous speech contributes to the higher rate of devoicing for IDS nonsense words in spontaneous speech than in read speech (Figure 1).

The clear difference between devoicing rates for regular words and nonsense words supports the argument that mothers in this study were representing to their infants the devoicing phenomena typical of adult language, rather than simply failing to exhibit patterns of IDS. If mothers were just not adapting devoicing rates for regular words to their infant perceivers, then it could be expected that, similarly, they might not adapt devoicing rates for nonsense words either. However, the mothers in this study significantly reduced the rate of vowel devoicing in those forms, and they did so with exquisite sensitivity to a progression from clearly articulating, i.e. voicing devoiceable vowels in newly introduced words, to producing more adult-like forms of newly 'learned' words. Thus, the lack of difference between Japanese ADS and IDS regular words does not reflect an inability to modify speech to accommodate perceivers; rather this lack results from a process of providing a model of fluent adult speech to infant perceivers.

CONCLUSION AND FUTURE DIRECTIONS

Although infant-directed speech serves the crucial function of making some aspects of the target language more salient and thus more accessible to

infant language learners, it must also model adult language to these learners. In the case of vowel devoicing, which obscures the canonical consonant–vowel form of the Japanese language, mothers nevertheless model the adult-like process of devoicing, including cases of near-uniform devoicing in specific morphemes, and adult-like distinctions between devoicing rates in read and spontaneous speech, in their speech to their infants aged 1;0. Mothers are clearly ABLE to reduce vowel devoicing when recognition of the word form depends crucially upon acoustic output, as in the case of nonsense words; however, they model the fluent adult speech process of devoicing for regular words and specific morphemes.

An immediate question presents itself: do mothers continue to represent adult-like devoicing as infants get older? It is well-known that characteristics of infant-directed speech change with the age of the infant being addressed (e.g. Amano *et al.*, 2006; Bernstein Ratner, 1984; Kitamura *et al.*, 2002; Papoušek & Hwang, 1991; Reilly & Bellugi, 1996); it might be the case that mothers decrease devoicing rates and thus make the canonical consonant–vowel form of the Japanese mora more salient to older infants as they themselves begin to produce word-like utterances. Bernstein Ratner (1984) demonstrated greater exaggeration of vowel exemplars, what she calls ‘more canonical phoneme articulation’ (p. 573), in English maternal speech to ‘advanced child listeners’, i.e. those producing utterances of a mean length of 2.5–4.0 words, than to holophrastic or preverbal infants. An investigation of devoicing rates in speech to infants at a more advanced stage of linguistic development than the infants aged 1;0 involved in this study might reveal age-related changes resulting in a similarly ‘more canonical’ pronunciation of devoiced contexts.

Another question of interest is whether and how the presence of vowel devoicing in IDS affects infants’ acquisition of the phonological process of vowel devoicing. We have noted above that infants can discriminate voiced and voiceless versions of morae in devoicing contexts (Kajikawa *et al.*, 2006; Mugitani *et al.*, 2007), and that adults prefer devoiced vowels in devoicing contexts (Fais *et al.*, 2005), but we do not know if infants show a similar preference. Especially since there is evidence that older infants (aged 1;6) are beginning to understand the phonological equivalence of devoiced and voiced vowels in devoicing contexts (Kajikawa *et al.*, 2006; Mugitani *et al.*, 2007), it would be of interest to examine the question of preference at both the age of the infants in this study and at age 1;6 to reveal possible developmental trends.

This study has provided evidence that, though they can reduce devoicing rates when necessary for the understanding of nonsense forms, mothers of Japanese infants aged 1;0 faithfully reproduce to their infants devoicing rates typical of adult speech for native Japanese word forms. This is true across both read and spontaneous speech, and despite the fact that

devoicing obscures the canonical form of Japanese morae. These results highlight the complexity of the interactions amongst apparently conflicting functions of infant-directed speech: IDS must at the same time make aspects of the adult language accessible to infant learners and accurately model the target language to be learned. While a great deal of research has been devoted to documenting the ways in which IDS serves the former purpose, the present research isolates one important way in which IDS does the latter, representing the adult processes of devoicing despite the differences between the resulting acoustic forms and the underlying phonological structure of the language.

REFERENCES

- Amano, S., Nakatani, T. & Kondo, T. (2006). Fundamental frequency of infants' and parents' utterances in longitudinal recordings. *Journal of the Acoustical Society of America* **119**(3), 1636–47.
- Andruski, J. E., Kuhl, P. K. & Hayashi, A. (1999). The acoustics of vowels in Japanese women's speech to infants and adults. *Proceedings of the 14th International Congress on Phonetic Sciences 3*, 2177–79. San Francisco, CA: Regents of the University of California.
- Au, T. K., Knightly, L. M., Jun, S. A. & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science* **13**(3), 238–43.
- Barik, H. C. (1977). Cross-linguistic study of temporal characteristics of different types of speech materials. *Language and Speech* **20**, 116–26.
- Bernstein Ratner, N. (1984). Patterns of vowel modification in mother–child speech. *Journal of Child Language* **11**, 557–78.
- Bernstein Ratner, N. & Luberoff, A. (1984). Cues to post-vocalic voicing in mother–child speech. *Journal of Phonetics* **12**, 285–89.
- Boersma, P. & Weenink, D. (2005). Praat: Doing phonetics by computer (Version 4.3.02) [Computer program; retrieved 27 February 2005 from <http://www.praat.org/>].
- Dupoux, E., Kakehi, K., Hirose, Y. & Pallier, C. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* **25**(6), 1568–78.
- Fais, L., Kajikawa, S., Werker, J. F. & Amano, S. (2005). Japanese listeners' perceptions of phonotactic violations. *Language and Speech* **48**(2), 185–202.
- Fernald, A. & Simon, T. (1984). Expanded intonation contours in mothers' speech to newborns. *Developmental Psychology* **20**(1), 104–113.
- Fernald, A., Taeschner, T., Dunn, J., Papoušek, M., de Boysson-Bardies, B. & 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. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development* **67**, 3192–218.
- Ikeda, Y. & Masataka, N. (1999). A variable that may affect individual differences in the child-directed speech of Japanese women. *Japanese Psychological Research* **41**(4), 203–208.
- Imaizumi, S., Hayashi, A. & Deguchi, T. (1998). Listener adaptive characteristics of vowel devoicing in Japanese dialogue. *Journal of the Acoustical Society of America* **98**(2), 768–78.
- Inazuka, E. & Inazuka, H. (2003). *An introduction to Japanese phonetics: Explanations and exercises [Nihongo no onsei nyumon: kaisetsu to enshu]*. Tokyo: Baberu.
- Kajikawa, S., Fais, L., Mugitani, R., Werker, J. F. & Amano, S. (2006). Cross-language sensitivity to phonotactic patterns in infants. *Journal of the Acoustical Society of America* **120**(4), 2278–84.
- Karzon, R. G. (1985). Discrimination of polysyllabic sequences by one- to four-month-old infants. *Journal of Experimental Child Psychology* **39**, 326–42.

- Kitamura, C., Thanavishuth, C., Burnham, D. & Luksaneeyanawin, S. (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**, 372–93.
- Kondo, M. (2005). Syllable structure and its acoustic effects on vowels in devoicing environments. In J. van de Weijer, K. Nanjo & T. Nishihara (eds), *Voicing in Japanese*, 229–45. Berlin: Mouton de Gruyter.
- Kuhl, P. K., Andruski, J. E., Chistovich, I. A., Chistovich, L. A., Kozhevnikova, E. V., Ryskina, V. L., Stolyarova, E. I., Sundberg, U. & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science* **277**, 684–86.
- Levin, H., Schaffer, C. A. & Snow, C. (1982). The prosodic and paralinguistic features of reading and telling stories. *Language and Speech* **25**, 43–54.
- Li, C. N. & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language* **4**(2), 185–99.
- Maekawa, K. & Kikuchi, H. (2005). Corpus-based analysis of vowel devoicing in spontaneous Japanese – An interim report. In J. van de Weijer, K. Nanjo & T. Nishihara (eds), *Voicing in Japanese*, 205–227. Berlin: Mouton de Gruyter.
- Masataka, N. (1992). Pitch characteristics of Japanese maternal speech to infants. *Journal of Child Language* **19**, 213–23.
- Masataka, N. (2002). Pitch modifications when interacting with elders: Japanese women with and without experience with infants. *Journal of Child Language* **29**, 939–51.
- Mugitani, R., Fais, L., Kajikawa, S., Werker, J. F. & Amano, S. (2007). Age-related changes in sensitivity to native phonotactics in Japanese infants. *Journal of the Acoustical Society of America* **122**, 1332–35.
- Nagano-Madsen, Y. (1994). Vowel devoicing in Japanese. *Lund University Department of Linguistics Working Papers* **42**, 119–29.
- Ogasawara, N. (2005). Processing voiceless vowels in Japanese: Effects of language-specific phonological knowledge. Poster presented at the 149th meeting of the Acoustical Society of America, Vancouver.
- Papoušek, M. & Hwang, S.-F. C. (1991). Tone and intonation in Mandarin babytalk to presyllabic infants: Comparison with registers of adult conversation and foreign language instruction. *Applied Psycholinguistics* **12**, 481–504.
- Phillips, J. R. (1973). Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development* **44**, 182–85.
- Reilly, J. & Bellugi, U. (1996). Competition on the face: Affect and language in ASL motherese. *Journal of Child Language* **23**, 219–39.
- Shibatani, M. (1990). *The languages of Japan*. New York: Cambridge University Press.
- Smith, C. (2006). French conversations do not show evidence of phonetic accommodation to a non-native interlocutor. *Journal of the Acoustical Society of America* **120**(5), 3170.
- Snow, C. (1972). Mothers' speech to children learning language. *Child Development* **43**, 549–65.
- Sugito, M. (2005). Devoiced accented vowels in Japanese. In J. van de Weijer, K. Nanjo & T. Nishihara (eds), *Voicing in Japanese*, 247–59. Berlin: Mouton de Gruyter.
- Thiessen, E. D., Hill, E. A. & Saffran, J. R. (2005). Infant-directed speech facilitates word segmentation. *Infancy* **7**(1), 53–71.
- Tsuchida, A. (1997). Phonetics and phonology of Japanese vowel devoicing. Unpublished doctoral dissertation, Cornell University, New York.
- Uther, M., Knoll, M. A. & Burnham, D. (2007). Do you speak E-N-G-L-I-S-H? A comparison of foreigner- and infant-directed speech. *Speech Communication* **49**, 2–7.
- Werker, J. & McLeod, P. J. (1989). Infant preference for both male and female infant-directed talk: A developmental study of attentional and affective responsiveness. *Canadian Journal of Psychology* **43**(2), 230–46.
- Werker, J., Pons, F., Dietrich, C., Kajikawa, S., Fais, L. & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition* **103**, 147–62.