Common Themes and Cultural Variations in Japanese and American Mothers' Speech to Infants

Anne Fernald and Hiromi Morikawa

Stanford University

FERNALD, ANNE, and MORIKAWA, HIROMI. Common Themes and Cultural Variations in Japanese and American Mothers' Speech to Infants. CHILD DEVELOPMENT, 1993, 64, 637-656. This study explored both universal features and cultural variation in maternal speech. Japanese and American mothers' speech to infants at 6, 12, and 19 months was compared in a cross-sectional study of 60 dyads observed playing with toys at home. Mothers' speech in both cultures shared common characteristics, such as linguistic simplification and frequent repetition, and mothers made similar adjustments in their speech to infants of different ages. American mothers labeled objects more frequently and consistently than did Japanese mothers, while Japanese mothers used objects to engage infants in social routines more often than did American mothers. American infants had larger noun vocabularies than did Japanese infants, according to maternal report. The greater emphasis on object nouns in American mothers' speech is only partially attributable to structural differences between Japanese and English. Cultural differences in interactional style and beliefs about child rearing strongly influence the structure and content of speech to infants.

Adults in diverse cultures use a special speech style when interacting with infants. Certain characteristics of infant-directed speech appear to be universal, or at least very widespread across languages. In a review of anthropological evidence, Ferguson (1978) reported that phonological and syntactic modifications, a simplified lexicon, and exaggerated prosody were common features of speech to children in the 27 languages for which data were then available. Other studies have reported cross-linguistic differences as well as similarities in child-directed speech (e.g., Nwokah & Fogel, 1990). For example, although the exaggeration of intonation in infant-directed as compared to adult-directed speech may be universal, the extent of prosodic exaggeration differs to some degree among cultures (Fernald et al., 1989). Different sorts of explanations have been offered to account for the apparently universal features of infant-directed speech, on the one hand, and significant cultural variations in maternal speech style on the other. Mothers in different cultures may modify their speech in similar ways in order to accommodate the perceptual and cognitive immaturity of the preverbal infant (e.g., Fernald, 1984, 1992; Newport, Gleitman, & Gleitman, 1977). While the universal attributes of mothers' speech could thus be shaped primarily in response to universal characteristics of the human infant, variations in maternal speech style are more likely to be linguistic or cultural in origin. Structural differences among languages can constrain the form of early language input (e.g., Gopnik & Choi, 1990), while culture-specific beliefs and practices related to children and caretaking can also have a powerful influence on the nature of speech addressed to infants in different cultures (Shatz, 1991).

The goal of the research reported here was to investigate both common strategies and cultural variations in maternal speech style. This research was supported by grants from the National Institutes of Health (HD 24349 and MH 41511) to Anne Fernald. We are extremely grateful to Carla Herrera, Taeko Koshimizu, Gerald McRoberts, Deanne Perez, and Catherine Sanderson for assistance with data collection and analysis, and to Ellen Markman, Marilyn Shatz, and Dan Slobin for their helpful comments on an earlier version of the manuscript. Correspondence should be sent to Anne Fernald, Department of Psychology, Stanford University, Stanford, CA 94305.

This research was supported by grants from the National Institutes of Health (HD 24349 and MH 41511) to Anne Fernald. We are extremely grateful to Carla Herrera, Taeko Koshimizu, Gerald McRoberts, Deanne Perez, and Catherine Sanderson for assistance with data collection and analysis, and to Ellen Markman, Marilyn Shatz, and Dan Slobin for their helpful comments on an earlier version of the manuscript. Correspondence should be sent to Anne Fernald, Department of Psychology, Stanford University, Stanford, CA 94305.

[Child Development, 1993, 64, 637-656. © 1993 by the Society for Research in Child Development, Inc. All rights reserved. 0009-3920(93)6403-0001801.00]
and cultural variation in Japanese and American mothers' speech to infants at different levels of development. In a cross-sectional study with infants at 6, 12, and 19 months of age, we focused on Japanese and American mothers' use of language as they engaged their infants in play with objects. Four questions were of central interest: First, which features of maternal speech appear to be constrained by the developmental level of the addressee, varying with infant age irrespective of culture? Second, which features of maternal speech are constant across infant age but vary consistently between cultures? Third, are the differences in form and content between Japanese and American mothers' speech determined more by linguistic constraints or by cultural differences in communicative style? And, finally, in what ways might cultural variations in maternal speech have consequences for the course of language development in children learning Japanese and English?

Rationale for Focusing on Mothers' Talk about Objects

Our focus on talk about objects was motivated by several considerations. First, object play between mother and infant is assumed to be of special relevance to language learning (e.g., Adamson, Bakeman, & Smith, 1990). A large proportion of infants' early vocabulary consists of object names (e.g., Nelson, 1973), which typically account for more than 60% of the first 50 words acquired by children learning English (see Bridges, 1986). Research on lexical acquisition in Japanese has been less systematic, but also reflects a preponderance of nouns in children's early vocabularies (Iwabuchi & Muraishi, 1976; Murata, 1984).

Why is early lexical acquisition in both English and Japanese apparently biased toward nouns? Gentner (1982) proposes that object words are conceptually less complex than action words and thus easier to encode. Another possible explanation is that infants are biased to attend differentially to objects and to assume that new words are names for objects (see Markman, 1992). A third explanation, advocated by Bridges (1986), is that parents focus more often on objects than on actions, and that object labels are given special prominence in early language input. Nelson (1973), for example, found that more than half of American mothers' utterances to 13-month-old infants in a play situation were related to the toys and objects at hand, with a high frequency of object labeling. Cross-cultural data are obviously relevant in evaluating the contribution of linguistic input to children's early bias toward learning nouns.

One question to be addressed in this study is whether Japanese and American mothers are similar in their tendency to provide noun labels for objects when interacting with infants at different levels of language development.

Another reason for focusing on object play is that mothers' talk about objects affords several theoretically interesting measures that are comparable in Japanese and English. Much research on early language input has concentrated primarily on syntactic features of mothers' speech (e.g., Hoff-Ginsberg, 1985; Snow, 1977), structural characteristics that are often difficult to compare across unrelated languages. In this study, we focused instead on several measures related to object labeling, including the frequency of noun references to objects, the variety of different labels used to refer to a given object, and the consistency with which the same noun label was used in successive references to a given object. Also of interest was the extent to which mothers used the adult form of the object label, as opposed to an onomatopoeic word or other "babtalk" form. To explore ways in which mothers might make object labels especially salient in their speech to infants, we examined how often noun labels were spoken in isolation, how often a noun label was the only content word in a multiword utterance, and how often noun labels were presented in question utterances. These measures could all be clearly operationalized and meaningfully compared in Japanese and English. Moreover, the object play situation could be standardized and was appropriate for infants at all three ages in both Japanese and American cultures.

Features of Maternal Speech Influenced by the Developmental Level of the Child

Early mother-infant interactions are inherently asymmetrical, in that one partner is socially sophisticated and linguistically adept, while the other has only a rudimentary repertoire of social behaviors and limited cognitive and linguistic abilities. Many observers of mother-infant social interaction have called attention to the intuitive skill with which mothers compensate for this asymmetry, adjusting their behaviors to complement both the limitations and the developing capabilities of the infant (e.g., Stern, 1977). For example, American mothers use shorter utterances and more exaggerated pitch modulation with 4-month-old in-
Femald and Morikawa 639

fants than with infants beginning to speak (Stern, Spieker, Barnett, & MacKain, 1983), changes in maternal vocal behavior that are appropriate to the developmental level of the infant. If the mother's goal is to engage the 4-month-old infant in face-to-face play, then short, highly modulated vocalizations are especially effective. In the second year, however, mothers modify their speech in ways related less to eliciting attention and more to facilitating language comprehension and production (Adamson & Bakeman, 1984).

The claim that modifications in maternal speech serve different functions at different points in infant development is supported by experimental research. With young infants, prosodically exaggerated speech recruits greater attention (e.g., Fernald, 1985) and elicits more positive affect (Fernald, 1993; Werker & McLeod, 1989) than adult-directed speech. In the second year, when infants are beginning to learn language, exaggerated intonation facilitates word recognition (Fernald & McRoberts, 1991). In the present study, we predicted that both Japanese and American mothers would use more attention-directing words and nonsense sounds with younger infants than with older infants and would increase their use of noun labels with older infants. Such findings would be consistent with the idea that certain common features of maternal speech across cultures are driven by universal response tendencies and capabilities of the developing infant.

Features of Maternal Speech Influenced by the Linguistic Structure of the Input Language

Another important influence on the nature of maternal speech is the formal linguistic structure of the particular language spoken to the child. In five Korean infants observed by Gopnik and Choi (1990), the vocabulary spurt emerged several months later than in American infants. Gopnik and Choi suggest that structural differences in Korean and English could partially account for these developmental differences. For example, nouns can be deleted in Korean, but not in English; moreover, nouns typically occur at the ends of sentences in English, while verbs occur in sentence-final position in Korean. Since linguistic units in final position are potentially more perceptually prominent than those embedded within the sentence (Slobin, 1973), Gopnik and Choi argue that nouns may be more salient in English than in Korean, and that English-learning children may attend relatively more to nouns than do children learning Korean. This differential emphasis on nouns, enhanced both by frequent noun deletion in Korean and by the canonical subject-verb-object word order of English, could influence the onset of the vocabulary spurt.

Like Korean, Japanese is a verb-final language in which nominal ellipsis is permitted. One question of interest in the present study was motivated by Gopnik and Choi's (1990) hypothesis that nouns in maternal speech should receive less emphasis in languages like Korean and Japanese than in English, with relatively greater emphasis on verbs in the Asian languages. This question was addressed by comparing the overall frequency of nouns and verbs related to target objects in Japanese and American mothers' speech, as well as the extent to which mothers in these cultures emphasized nouns and verbs by presenting them in isolation. Another question of interest was whether any differences found in Japanese and American mothers' use of object words in speech to infants would be associated with differences in the speed of noun acquisition by infants learning Japanese and English.

Features of Maternal Speech Influenced by Cultural Values

The form and content of speech to children are also shaped by the values of the particular culture in which the child is being socialized. Schieffelin and Ochs (1986) have emphasized that early language input is a powerful medium of cultural transmission. For example, Shatz, Grimm, Wilcox, and Niemeier-Wind (1989) found that the use of modals such as must and can in German and American mothers' speech reflected different cultural values. German mothers talked more about necessity to their infants, consistent with the emphasis on obligation prevalent in German culture. American mothers focused more on agreement, intention, and possibility, consistent with the more egalitarian style of child rearing popular in American culture. Shatz et al. also found that the acquisition of modals by German and American children reflected these culturally appropriate biases in the input.

Cultural biases in maternal speech are also evident in the pragmatic contexts of language use. Clancy's (1986) research on the acquisition of communicative style illustrates how Japanese mothers emphasize the infant's role as a member of the social group. Japanese mothers frequently rehearse social
routines with their infants, teaching the child to respond politely and to consider the feelings of others. Clancy suggests that by continually articulating the thoughts and feelings of other people to the infant, and by encouraging empathy, Japanese mothers instill social values which are at the heart of Japanese culture.

In this study, cultural variations in infant-directed speech were investigated in several ways. First, we compared specific aspects of Japanese and American maternal speech in which other researchers have reported cultural differences, such as the use of nonsense sounds and the frequency with which mothers engage their infants in social routines (e.g., Clancy, 1986; Fischer, 1970). Second, we explored Fischer's (1970) claim that Japanese mothers continue to use “babytalk” speech modifications for a longer period than American mothers. Finally, following the observation session, we interviewed mothers informally about whether they thought a special speech style was appropriate when interacting with infants, and what functions might be served by this infant-directed speech style.

Comparative Research on Japanese and American Mother-Infant Interaction

Although a few studies in the Japanese literature (e.g., Chew, 1969; Ogino, 1981) report modifications in infant-directed speech similar to those in other languages, there have been no detailed studies of age-related changes. The nine previous studies systematically comparing Japanese and American mother-infant interaction (Bornstein, Azuma, Tamis-LeMonda, & Ogino, 1990; Bornstein, Toda, Azuma, Tamis-LeMonda, & Ogino, 1990; Caudill & Weinstein, 1970; Fogel, Toda, & Kawai, 1988; Otaki, Durrett, Richards, Nyquist, & Pennebaker, 1986; Sengoku, Davitz, & Davitz, 1982; Shand & Kosawa, 1985) and infant-directed speech (Morikawa, Shand, & Kosawa, 1988; Toda, Fogel, & Kawai, 1990) have all focused on 3–5-month-old infants.

Caudill and Weinstein's (1970) pioneering study of Japanese and American mother-infant interaction revealed both similarities and differences in caretaking styles between the two cultures. They found that American mothers talked more to the child and spent more time actively encouraging the infant to respond than did Japanese mothers, while Japanese mothers were more involved in rocking, carrying, and “lulling” the infant. More recent detailed analyses of Japanese and American mother-infant interaction have not found overall differences in maternal expressiveness, touching, or vocalizing (Fogel et al., 1988), or in the proportion of time mothers engaged their infants in social and didactic activities (Bornstein, Azuma, Tamis-LeMonda, & Ogino, 1990). However, Fogel et al. found cultural differences in the patterning of maternal behavior: American mothers responded to infant behaviors primarily with facial and vocal displays, while Japanese mothers were more likely to respond with touch and other nonverbal behaviors. Given the greater emphasis on tacit understanding and nonverbal communication in Japanese culture, the responses of Japanese mothers to their preverbal infants may reflect these cultural differences in communicative style (Clancy, 1986).

Although these comparative studies have all included maternal vocalization as a dependent measure, most were not concerned with the linguistic content of maternal speech. Two previous studies have provided detailed comparisons of the structure and content of Japanese and American mothers’ speech to 3-month-old infants. Morikawa et al. (1988) found that Japanese mothers used indirect speech styles more often than American mothers. Toda et al. (1990) found American mothers to be more information-oriented than Japanese mothers. Japanese mothers more often used nonsense sounds and onomatopoeic words to engage the infant’s attention, a finding consistent with other reports of the frequent use of “babytalk” sounds in Japanese infant-directed speech (e.g., Chew, 1969; Fischer, 1970; Ogino, 1981). These cross-linguistic studies indicate that Japanese and American maternal speech to preverbal infants shares common features such as syntactic simplicity and frequent repetition, but also that mothers’ speech in these two cultures reflects differences in communicative style. The present research extends these findings by providing data on age-related changes in Japanese and American mothers’ speech to infants beyond the preverbal period.

Method

Subjects

The subjects in this study were 30 American and 30 Japanese mother-infant dyads. In each language group, there were 10 infants at each of three ages: 6, 12, and 19 months. In the 6-month-old group, the mean age of Japanese infants was 5;23 (range =
The mean age of the 12-month-old group, the mean age of Japanese infants was 12.17 (range = 11.01 to 14.04), and the mean age of American infants was 12.23 (11.02 to 14.22). In the 19-month-old group, the mean age of Japanese infants was 19.06 (range = 18.03 to 20.28), and the mean age of American infants was 19.19 (range = 18.05 to 22.24). All infants were reported by their mothers to be healthy and developing normally. In each age group, half the infants were female.

The mothers in the American sample were white, middle-class residents of an affluent suburban community, recruited through a university hospital. All were native speakers of American English and all had 2–4 years of college education. The mothers in the Japanese sample were native speakers of Japanese, all Japanese citizens who were visiting the San Francisco area for a short time as wives of affiliates of Japanese companies. Although the period of their residence in the United States varied from 2 months to 4 years (M = 20 months), the Japanese mothers all reported that they socialized primarily or exclusively with other Japanese families, and most had only limited competence in English. These mothers were recruited through a Japanese pediatrician serving the community of visiting families.

Observation Procedures
Subjects were observed during visits to the family home. The same pair of female observers conducted all home visits, one a native speaker of Japanese, the other a native speaker of American English. During visits to Japanese families, the Japanese observer interacted with the mother and supervised the session, while the American observer operated the recording equipment; during visits to American families, the roles of the two observers were reversed. The session began with a 15–30 min introductory period during which the observers got acquainted with mother and infant and explained the observation procedure. Subjects were told that the study concerned infant play and mother-infant interaction, but not that the focus was on maternal speech. Mother and infant were then audio- and video-recorded for about 10 min as they played together with the infant’s own toys, to acclimate them to the presence of recording equipment. Audio-recordings were made on a Sony TC-D5M professional quality cassette recorder; video-recordings were made using a Panasonic WV-3250 camera and a Panasonic PV-9000 portable video recorder.

When the mother and infant appeared to be playing comfortably together, two sets of standardized toys were introduced, including a pair of stuffed animals (dog and pig), and a pair of wooden vehicles (car and truck). These particular toys were selected after pilot testing because they were judged by Japanese and American mothers to be appropriate for infants in both cultures and suitable for all three age groups. After the infant’s own toys had been removed from the play area, the mother was given one of the two pairs of standard toys and asked to play with the infant as she normally would. Mother and infant were observed playing with the first pair of toys for 3–5 min, depending on the infant’s interest. These toys were then removed, and the second pair was presented for another 3–5 min period. Order of presentation of the two pairs of toys was counterbalanced across subjects.

Data Reduction and Analysis
Data analysis focused on the first 2.5 min following the introduction of each object pair, for a total of 5 min for each subject. Japanese and English speech samples were segmented into utterances by native speakers of each language, on the basis of pauses and syntactic and prosodic information. For each subject in each language, all utterances were initially classified as either related or unrelated to one of the four toys, or target objects. Utterances related to target objects included those containing references to the toys or their parts, actions by or upon the toys, and characteristics, states, and sounds attributed to the toys. In cases of uncertainty, the video records were consulted and contextual information about the mother’s focus of visual attention was used to determine whether a particular utterance was related to one of the target objects.

Measures Related to Target Objects
Noun label types.—Within the category of utterances related to target objects, those utterances containing a noun label for the toy were of special interest. Utterances with target labels were classified into three categories according to the form of the noun label used to refer to the toy: (1) adult form of the target noun, for example, dog in English, and inu in Japanese; (2) adult-form-plus-affix, for example, doggie in English, and koinu-chan in Japanese; (3) onomatopoeic name, for example, woof-woof in English, and wanwan in Japanese. The adult-form-
Japanese and English could be marked as mother said “woof woof woof” in English, attributed to the object. For example, if the were not used as noun labels, but that were ances were coded for nonsense sounds that of each language. ically, or both, as judged by native speakers in Japanese. Yes/no questions in vroom?) Interrogatives either syntactically or prosod- Eubuu suki? (Is this a vroom- English, or Is that a car? No questions, for example. In Japanese, or as yes/ Kore in English, or What’s this? Questions pertaining to the identity, location, actions, and characteristics of target objects or their parts, as well as to the child’s attitude toward the target object, were classified either as Wh-questions, for example, What’s this? in English, or Kore nani? (What’s this?) in Japanese, or as yes/no questions, for example, Is that a car? in English, or Buubuu suki? (Is this a vroom-vroom?) in Japanese. Yes/no questions in Japanese and English could be marked as interrogatives either syntactically or prosodically, or both, as judged by native speakers of each language.

Object sounds.—Target-related utterances were coded for nonsense sounds that were not used as noun labels, but that were attributed to the object. For example, if the mother said “woof woof woof” in English, or “wanwanwan” in Japanese, while playing with the dog, these were classified as object sounds, whereas if she said “This is a woof woof” in English, or “Kore wa wanwan” in Japanese, these were classified as onomatopoeic labels.

Consistency of lexical usage in object labeling.—Given that mothers could use a variety of noun labels to name each of the four target objects, we were interested in mothers’ consistency in lexical choice when repeatedly labeling a particular toy. Labeling consistency was assessed by computing the percentage of labeling instances in which a given target object was labeled using the same lexical item as had been used in the immediately previous instance of labeling that object. For each mother, an average lexical consistency score was computed across the four target objects.

Linguistic complexity.—All utterances containing target-object labels were classified as either simple or complex utterances. In simple utterances, the target-object label was the only content word, while complex utterances had one or more content words in addition to the target-object label. Content words included all nouns, verbs, adjectives, and adverbs. This measure was used to compare relative linguistic complexity because it could be clearly operationalized in both Japanese and English, unlike other measures of complexity such as MLU, which are difficult to equate across languages.

Questions related to target objects.—Questions pertaining to the identity, location, actions, and characteristics of target objects or their parts, as well as to the child’s attitude toward the target object, were classified either as Wh-questions, for example, What’s this? in English, or Kore nani? (What’s this?) in Japanese, or as yes/no questions, for example, Is that a car? in English, or Buubuu suki? (Is this a vroom-vroom?) in Japanese. Yes/no questions in Japanese and English could be marked as interrogatives either syntactically or prosodically, or both, as judged by native speakers of each language.

Actions by or upon target objects.—Target-related utterances containing action verbs were coded according to whether the verb referred to an action performed by the target object, for example, “The doggy runs,” or to an action performed by the mother or the child upon the target object, for example, “Kiss the doggy.”

Other Measures

Content words in isolation.—Content words occurring in isolation were classified in three categories: (1) nouns; (2) action verbs, including Japanese “babytalk” noun/verb compounds such as dakko suru (literally hug do, glossed as do a hug in English); (3) adjectives and adverbs. Other single-word utterances were not included in this analysis. Japanese utterances consisting of a single noun plus a final particle were classified as isolated object words. Japanese utterances consisting of a single word that is typically part of a noun/verb compound, such as dakko without suru, were classified as isolated action verbs, even though dakko functions grammatically as a noun in this context.

Repetitions.—The entire corpus was coded for exact and partial repetitions of maternal utterances. Exact repetitions occurred when an entire utterance was repeated verbatim within three utterances of the original, excluding repetitions of the infant’s name, exclamations and interjections, and short responses such as yeah and no. Partial repetitions were those involving removal, substitution, or addition of a single content word in an original utterance containing at least three content words. (For this analysis, pronouns and interrogatives were counted as content words, as well as nouns, verbs, adjectives, and adverbs.) However, substitutions resulting in a substantial change in meaning (e.g., “Do you want to ride the car?” following “Do you want to push the car?”) were not counted as partial repetitions.

Social routines.—The entire corpus was coded for verbal formulas typically used in three kinds of social routine: (1) greeting routines, such as Hello and Bye bye in English, and Konnichiwa (Hello) and O-genki desu ka (Are you well?) in Japanese; (2) exchange routines, used in offering, request-
ing, and accepting things politely, such as Here you are, and Thank you in English, and Doozo (Please help yourself), Choodai (Give me), and Arigato (Thank you) in Japanese; (3) empathy routines, used to encourage the child to feel or behave positively toward a person or object, such as Give the doggy a love in English, and Kawaii kawaii shi-te age-te (literally: Do cute cute, glossed as Give it a love) in Japanese.

Content of utterances unrelated to target objects.—Utterances unrelated to a target object were classified in five content categories: (1) references to people, events, and objects other than target objects; (2) attentionals and vocatives, such as Hey! in English, and Hora (Hey) in Japanese; (3) interjections and backchannel vocalizations, including Hmm, Aah, etc.; (4) nonsense sounds used playfully but not attributed to a target object, such as Skootch! in English; (5) other infrequently occurring utterances unrelated to target objects, such as references to the infant's state, praise vocalizations, and game routines.

Infants' productive vocabulary.—Mothers of 12- and 19-month-old infants were asked to complete the productive vocabulary checklist developed by Bates and her colleagues as part of the Communicative Development Inventory (e.g., Snyder, Bates, & Bretherton, 1981). A Japanese version of the vocabulary checklist was prepared by translating the original questionnaire into Japanese with appropriate minor modifications. For example, the English words feet and leg were both represented by ashi in Japanese, while the English first-person pronoun was translated both as watashi (the generic first-person pronoun) and as boku (the first-person pronoun used by young males) in Japanese.

Mothers' attitudes toward infant-directed speech.—At the end of the observation session, each mother was interviewed informally about her beliefs and opinions about "useing a special speech style" with her infant. Mothers were asked to describe how and why their speech to the infant was different from their speech to adults, and whether they thought it was appropriate or inappropriate to speak in different ways to infants and adults.

Observer Agreement
A Japanese-English bilingual (H.M.) coded both the Japanese and American transcripts. Two Japanese dyads and two American dyads in each of the three age groups were also coded independently by monolingual speakers of Japanese and English, respectively. For all measures related to target object labels, linguistic complexity, questions, repetition, isolated content words, action verbs, backchannels, and attentionals, agreement was high. The mean kappa coefficients across these measures ranged from .87 to .97 in English (M = .92) and from .83 to .94 in Japanese (M = .87). A separate reliability analysis was run on the three potentially confusable categories of nonsense sounds: onomatopoeic labels, sounds attributed to target objects, and other nonsense sounds used playfully. Kappa coefficients for nonsense sounds ranged from .77 to 1.0 in English (M = .92), and from .92 to .95 in Japanese (M = .93). Social routines were also analysed separately, with kappa coefficients ranging from .87 to 1.0 in English (M = .95), and from .89 to 1.0 in Japanese (M = .98).

Results
The data analyses were motivated by two quite different goals. One goal was primarily descriptive: to provide detailed information about the content of mothers' speech to infants at different ages in Japanese and English. A second goal was comparative and inferential: to investigate the effects of both culture and infant age on the nature of maternal speech. With these two ends in mind, both descriptive and inferential statistics are used in presenting the results of this study. Moreover, in order to provide linguistically sensitive descriptions of mothers' speech to their infants across languages, the data on some measures are broken down in finer detail than would be necessary if our intention was only to document those differences that emerge as statistically significant.

Unless otherwise indicated, the data were analyzed using a 2 (language) × 3 (infant age) factorial design. Sex of the infant was not included as a factor because preliminary analyses revealed no significant effects of infant sex. To establish whether the 5-min samples of mother-infant interaction chosen for analysis were similar in total amount of speech, the first step was to compare the mean numbers of utterances in Japanese (M = 124.4, SD = 24.9) and English (M = 98.5, SD = 26.6). Japanese mothers spoke significantly more to their infants than did American mothers, F(1, 54) = 14.63, p < .001, contrary to the findings of Caudill and Weinstein (1970) with 3-month-old infants.
Neither the effect of age nor the age x language interaction was significant.

In Japanese mothers' speech, 60% of utterances were related in some way to the target toys, while 67% of American mothers' utterances were target-related. The numbers of utterances related to target objects were compared in a $2 \times 3$ ANCOVA, in which the total number of utterances for each subject was entered as the covariate. An ANCOVA was used in order to control for the differences found between language groups in the total amount of speech. This analysis revealed that the mean number of target-related utterances was significantly higher in American mothers' speech ($M = 73.9, SD = 19.7$) than in Japanese mothers' speech ($M = 66.6, SD = 17.8$), $F(1, 53) = 4.67, p < .05$. The effect of age was also significant, $F(2, 53) = 7.06, p < .01$, with no significant interaction. Follow-up tests showed that mothers used significantly fewer target-related utterances in speech to 6-month-old infants ($M = 62.4$) than to 12-month-old infants ($M = 75.7$), $F(1, 53) = 13.04, p < .001$, or to 19-month-old infants ($M = 72.5$), $F(1, 53) = 7.46, p < .01$.

Object Labels in Mothers' Speech

Given that the target toys were a central topic of conversation for both Japanese and American mothers, how often did mothers in these two cultures refer to the toys using noun labels when talking to infants of different ages? The numbers of noun labels referring either to whole target objects or to parts of target objects were compared in a $2 \times 3$ ANCOVA, with the total number of utterances as the covariate. The mean numbers of noun labels in Japanese and American mothers' speech, adjusted for the covariate, are shown in Figure 1. American mothers used noun labels for target objects in speech to infants significantly more often ($M = 38.2, SD = 16.6$) than did Japanese mothers ($M = 20.9, SD = 13.8$), $F(1, 53) = 21.07, p < .001$. The effect of infant age was also reliable, $F(2, 53) = 5.94, p < .01$. Follow-up tests on the age effect showed significantly fewer noun labels in mothers' speech to 6-month-old infants ($M = 22.4, SD = 18.1$) as compared to 19-month-old infants ($M = 36.6, SD = 13.8$), $F(1, 53) = 11.88, p < .001$. The number of noun labels in mothers' speech to 12-month-old infants ($M = 29.6, SD = 12.4$) was not significantly different from speech to 6- or 19-month-olds ($p < .10$ in both comparisons). The age x language interaction was not significant.

The previous analyses revealed that Japanese mothers produced more utterances overall in speech to their infants, and that American mothers used more noun labels for target objects than did Japanese mothers. For this reason, ANCOVAs were used in all subsequent analyses that might be influenced by these baseline differences between language groups. In analyses that included the entire corpus of speech, the total number of utterances was entered as the covariate. However, several analyses focused

![American Mothers vs Japanese Mothers](image)

**FIG. 1.**—Adjusted mean numbers of noun labels referring to whole target objects and parts of target objects in Japanese and American mothers' speech to infants. (Total number of utterances was the covariate used to adjust the means.)
on how mothers treated object labels in Japanese and English. Since these analyses were limited to utterances containing target noun labels, the total number of target noun labels was used as the covariate rather than the total number of utterances. In all analyses involving ANCOVAs, the means reported have been adjusted for the covariate.

**Noun Label Types**

As shown in Tables 1 and 2, both American and Japanese mothers used a variety of noun label types when referring to whole target objects, ranging from the adult form of the word to special diminutive forms and onomatopoeic words. The frequencies of adult-form noun labels for the whole target object in Japanese and American mothers’ speech were compared using a 2 x 3 ANCOVA with total number of noun labels as the covariate. American mothers used adult-form labels significantly more often ($M = 14.2, SD = 8.5$) than Japanese mothers ($M = 5.9, SD = 4.6$), $F(1, 53) = 38.11, p < .001$, with no significant effect of age or interaction. American mothers used the adult form of the noun 60% of the time when labeling target objects, while Japanese mothers used the adult form for only 24% of their target noun labels. Adult-word-plus-affix forms such as diminutives were also used significantly more often by American mothers ($M = 9.6, SD = 9.3$) than by Japanese mothers ($M = 5.7, SD = 4.3$), $F(1, 53) = 7.94, p < .01$, with no significant effect of age or interaction.

The most striking cultural difference was that Japanese mothers used onomatopoeic words more frequently than any other type of noun label type to infants at all three ages, while the American mothers in this sample almost never used this form to refer to objects. On average, 52% of Japanese mothers’ labels for target objects consisted of onomatopoeic words such as *buubuu*

<table>
<thead>
<tr>
<th>Target Object</th>
<th>American</th>
<th>Japanese</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dog</td>
<td>dog</td>
<td><em>inu</em> (dog)</td>
</tr>
<tr>
<td></td>
<td>doggy</td>
<td><em>wanwan</em> (woof-woof)</td>
</tr>
<tr>
<td></td>
<td>puppy dog</td>
<td><em>wan-chan</em> (woof-DIM)</td>
</tr>
<tr>
<td></td>
<td>Mr. Doggy</td>
<td><em>wanwan-san</em> (woof-woof-HON)</td>
</tr>
<tr>
<td></td>
<td><em>woof-woof</em></td>
<td><em>wanwan-chan</em> (woof-woof-DIM)</td>
</tr>
<tr>
<td></td>
<td>puppy</td>
<td><em>koinu-chan</em> (puppy-DIM)</td>
</tr>
<tr>
<td></td>
<td>toy</td>
<td></td>
</tr>
<tr>
<td>Pig</td>
<td>pig</td>
<td><em>buta</em> (pig)</td>
</tr>
<tr>
<td></td>
<td>piggy</td>
<td><em>butsan</em> (pig-HON)</td>
</tr>
<tr>
<td></td>
<td>puppet pig</td>
<td><em>buta-chan</em> (pig-DIM)</td>
</tr>
<tr>
<td></td>
<td>puppet</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>buubuu/buubuubuu</em> (oink-oink)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mr. Pig</td>
<td><em>buubuu buta-san</em> (pig-HON oink-oink)</td>
</tr>
<tr>
<td></td>
<td>oink-oink</td>
<td><em>buubuu buta-san</em> (oink-oink pig-HON)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>buu-san</em> (oink-HON)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>buu-chan</em> (oink-DIM)</td>
</tr>
<tr>
<td>Car</td>
<td>car</td>
<td><em>jidoosha</em> (automobile)</td>
</tr>
<tr>
<td></td>
<td>toy</td>
<td><em>kuruma</em> (car-wheel)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>buubuu</em> (vroom-vroom)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>buu</em> (vroom)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>buubuu-tan</em> (vroom-vroom-DIM)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>kaa/ka</em> (car)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>supootsu kaa</em> (sports car)</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>omocha</em> (toy)</td>
</tr>
<tr>
<td>Truck</td>
<td>truck</td>
<td><em>torakku</em> (truck)</td>
</tr>
<tr>
<td></td>
<td>truck-car</td>
<td><em>buubuu</em> (vroom-vroom)</td>
</tr>
<tr>
<td></td>
<td>car</td>
<td><em>kaa</em> (car)</td>
</tr>
<tr>
<td></td>
<td>circus truck</td>
<td></td>
</tr>
<tr>
<td></td>
<td>cookie truck</td>
<td></td>
</tr>
<tr>
<td></td>
<td>toy</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE 1**

NOUN LABELS USED BY AMERICAN AND JAPANESE MOTHERS TO REFER TO TARGET OBJECTS

*NOTE.*—Words in parentheses are English glosses for Japanese labels. DIM = diminutive suffix; HON = honorific suffix.
TABLE 2
DISTRIBUTION OF NOUN LABEL TYPES USED TO REFER TO WHOLE TARGET OBJECTS IN JAPANESE AND AMERICAN MOTHERS’ SPEECH TO INFANTS AT 6, 12, AND 19 MONTHS OF AGE

<table>
<thead>
<tr>
<th>LANGUAGE GROUP</th>
<th>NOUN LABEL TYPE</th>
<th>6 Months</th>
<th>12 Months</th>
<th>19 Months</th>
<th>6 Months</th>
<th>12 Months</th>
<th>19 Months</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Adult form</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.4</td>
<td>13.3</td>
<td>15.0</td>
<td>6.5</td>
<td>6.0</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(10.2)</td>
<td>(8.8)</td>
<td>(7.0)</td>
<td>(3.4)</td>
<td>(3.7)</td>
<td>(6.2)</td>
</tr>
<tr>
<td></td>
<td>Adult form plus affix</td>
<td>9.5</td>
<td>10.8</td>
<td>8.5</td>
<td>7.9</td>
<td>5.3</td>
<td>4.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(13.3)</td>
<td>(7.5)</td>
<td>(6.7)</td>
<td>(4.5)</td>
<td>(3.3)</td>
<td>(5.0)</td>
</tr>
<tr>
<td></td>
<td>Onomatopoeia</td>
<td>- .5 a</td>
<td>- .7 a</td>
<td>- .1 a</td>
<td>8.9</td>
<td>12.1</td>
<td>14.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.0)</td>
<td>(.0)</td>
<td>(1.9)</td>
<td>(7.2)</td>
<td>(7.9)</td>
<td>(8.0)</td>
</tr>
</tbody>
</table>

NOTE.—Figures represent adjusted mean numbers of each noun label type, with total number of noun labels used as the covariate to calculate the adjusted means. Standard deviations are in parentheses.

* The unadjusted means for American mothers’ onomatopoeic noun labels were .0 in speech to 6- and 12-month-old infants and .6 in speech to 18-month-old infants.

Consistency of Lexical Usage in Object Labeling
American mothers were more likely than Japanese mothers to be consistent in using the same lexical item for a particular target object when labeling a toy repeatedly. That is, rather than switching from dog to puppy to doggy when labeling the dog three times in succession, American mothers tended to repeat one label. The average lexical consistency score for American mothers, indicating the probability that they would use the same lexical item on successive references to a target object, was .87 (SD = .15). Japanese mothers alternated more frequently among different lexical options, with an average lexical consistency score of .66 (SD = .25). A 2 × 3 ANOVA on the consistency scores confirmed that this difference between American and Japanese mothers was significant, F(1, 54) = 15.50, p < .001. Neither the effect of infant age nor the language × age interaction was significant.

Linguistic Complexity of Utterances Containing Target Object Labels
Linguistically “simple” utterances were operationalized as those in which the object label was the sole content word in the utterance. The use of simple utterances in Japanese and American mothers’ speech was compared in a 2 × 3 ANCOVA, with the total number of noun labels as the covariate. Simple utterances were used significantly more often by Japanese mothers (M = 11.0, SD = 7.1) than by American mothers (M = 5.8, SD = 5.7), F(1, 53) = 13.59, p < .001, with no significant effect of age or interaction. On average, 52% of Japanese mothers’ target noun labels occurred in linguistically simple utterances, while only 27% of American mothers’ noun labels occurred in such linguistically simple utterances.

Questions Related to Target Objects
The total numbers of questions in Japanese and American mothers’ speech to infants were analyzed in a 2 × 3 ANCOVA, with total number of utterances as the covariate. Questions were used significantly more often by American mothers (M = 33.8, SD = 13.1) than by Japanese mothers (M = 20.4, SD = 9.3), F(1, 53) = 27.76, p < .001. The effect of infant age was also significant, F(2, 53) = 6.12, p < .01, with no significant interaction. Follow-up tests on the main effect of age revealed that mothers used significantly more questions in speech to 19-month-old infants (M = 32.3) than to 12-month-old infants (M = 26.2), F(1, 53) = 4.98, p < .05, or 6-month-old infants (M = 22.8), F(1, 53) = 11.87, p < .001. In both Japanese and English, mothers used more yes/no questions than wh-questions. On average, yes/no questions comprised 62.5% of the interrogatives in American mothers’ speech, and 57% in Japanese mothers’ speech.

The use of questions in relation to noun labels was of particular interest in this study. On average, American mothers presented 41% (SD = 20.3) of their labels for target objects in question form, while Japanese
mothers presented only 18% (SD = 12.1) of their target labels in question form. Does this difference merely reflect the higher percentage of questions in American mothers’ speech overall, or do American mothers use the question form preferentially as a means of presenting labels? To address this question, we analyzed the percentage of noun labels presented in question form in a 2 × 3 ANCOVA, using the mean percentage of questions overall as the covariate to control for the greater number of questions in American mothers’ speech. A main effect of language, $F(1, 53) = 5.88, p < .05$, with no significant effect of age or interaction, indicated that American mothers had a significantly greater tendency to label objects using questions than did Japanese mothers.

**Nonsense Sounds Attributed to Objects**

Japanese mothers made extensive use of nonsense sounds not only by labeling objects with onomatopoeic words, but also by attributing sounds to objects. The numbers of nonsense sounds attributed to objects in Japanese and American mothers’ speech were analyzed in a 2 × 3 ANCOVA, with total number of utterances as the covariate. Nonsense sounds were used in this way significantly more often in Japanese mothers’ speech ($M = 18.7, SD = 12.4$) than in American mothers’ speech ($M = 9.9, SD = 6.0$), $F(1, 53) = 10.65, p < .01$. The effect of infant age was also significant, $F(2, 53) = 4.33, p < .05$. Follow-up tests revealed that mothers attributed nonsense sounds to target objects significantly more often when interacting with 12-month-old infants ($M = 18.2, SD = 11.4$) than with 19-month-old infants ($M = 9.7, SD = 4.4$), $F(1, 53) = 8.47, p < .01$. Mothers also frequently attributed nonsense sounds to objects in speech to 6-month-old infants ($M = 13.1, SD = 8.8$), although the mean for this age group did not differ significantly from the mean for the oldest group ($p = .07$). The language × age interaction was not significant.

**Repetitions**

The numbers of exact and partial repetitions for each subject were entered into a language × infant age × repetition type (exact/partial) ANCOVA with repeated measures on the last factor. The total number of utterances was used as the covariate in this analysis. This analysis yielded significant main effects of language, $F(1, 107) = 10.10, p < .01$, and repetition type, $F(1, 107) = 17.21, p < .001$, with a significant age × repetition type interaction, $F(2, 107) = 12.05, p < .001$. These findings indicate that American mothers used repetitions significantly more often overall ($M = 17.8, SD = 10.6$) than did Japanese mothers ($M = 12.0, SD = 7.6$), and that mothers used significantly more exact repetitions ($M = 9.1, SD = 6.5$) than partial repetitions ($M = 5.8, SD = 3.7$).

To explore the age × repetition type interaction, we compared exact repetitions as a percentage of overall repetitions in speech to infants at different ages. In speech to 6-month-old infants, 79% of mothers’ repetitions were exact repetitions, while in speech to 12- and 19-month-old infants, only 51%–53% of mothers’ repetitions were exact. A 2 × 3 ANOVA on the percentages of exact repetitions yielded a significant effect of infant age, $F(2, 54) = 13.05, p < .001$, with no significant effect of language or interaction. Follow-up tests revealed significantly higher percentages of exact repetitions in mothers’ speech to 6-month-old infants, as compared to both 12-month-old and 19-month-old infants, $p < .001$.

**Actions By and Upon Target Objects**

Words referring to actions upon the object exceeded words referring to actions by the object in both Japanese and English. The mean percentages of action verbs describing actions upon the object ranged from 60% to 68% in American mothers’ speech, and from 69% to 73% in Japanese mothers’ speech. The mean numbers of utterances containing words referring to actions by or upon a target object were analyzed in a 2 × 3 ANCOVA, with total number of utterances as the covariate. In speech to infants, action verbs were used significantly more often by American mothers ($M = 31.1, SD = 13.4$) than by Japanese mothers ($M = 18.4, SD = 12.7$), $F(1, 53) = 18.44, p < .001$, with no significant effect of age. Expressed as a percentage of speech related to the target objects, 27.4% of American mothers’ target-related utterances contained action verbs, while 18% of Japanese mothers’ target-related utterances contained action verbs.

Japanese but not American mothers tended to increase their use of verbs with older infants, although the age × language interaction in this analysis was not significant, $F(2, 53) = 2.80, p < .07$. The mean number of action verbs in English was 31.8 in speech to 6-month-olds, 32.6 in speech to 12-month-olds, and 28.5 in speech to 19-month-olds. The mean number of action verbs in Japanese was 11.4 in speech to 6-month-olds, 20.8 in speech to 12-month-olds, and 23.1 in speech to 19-month-olds.
Post hoc tests were used to explore these differences in action verb use in Japanese and American mothers’ speech to infants of different ages. Japanese mothers used significantly more target-related action verbs with 19-month-old infants than with 6-month-old infants, $F(1, 26) = 5.57, p < .05$. No other age differences were significant.

**Content Words in Isolation**

The distribution of content words occurring in isolation is summarized in Figure 2. A $2 \times 3$ ANCOVA on the total numbers of isolated content words, with total number of utterances as the covariate, revealed main effects of both language group, $F(1, 53) = 6.69, p < .05$, and infant age, $F(1, 53) = 3.32, p < .05$, with no significant interaction. Japanese mothers used significantly more single-word utterances consisting of isolated content words ($M = 13.0, SD = 9$) than did American mothers ($M = 8.2, SD = 5.4$). The mean number of isolated content words was 8.0 in speech to 6-month-olds, 10.5 in speech to 12-month-olds, and 13.2 in speech to 19-month-olds. Follow-up tests on the age effect revealed that mothers used significantly more isolated content words in speech to 19-month-old infants ($M = 13.2$) than to 6-month-old infants ($M = 8.0$), $F(1, 53) = 6.63, p < .05$. The 12-month-old age group did not differ significantly from either of the other two.

**Social Routines**

Figure 3 shows the distribution of social expressions used in exchange, empathy, and greeting routines in Japanese and American mothers’ speech to infants at different ages. The total numbers of social routines were analyzed in a $2 \times 3$ ANCOVA, with total number of utterances as the covariate. Japanese mothers engaged their infants in social routines twice as often ($M = 10.7, SD = 8.2$) as did American mothers ($M = 4.7, SD = 4.6$), $F(1, 53) = 10.42, p < .01$. The main effect of infant age was also significant, $F(2, 53) = 5.53, p < .01$, with no significant interaction. Follow-up tests revealed that mothers used social routines significantly more often when interacting with 12-month-old infants ($M = 11.3$) than when interacting with either 6-month-olds ($M = 4.5$) or 19-month-olds ($M = 7.3$). Pairwise comparisons among the three age groups were all significant at the .05 level.

**Content of Utterances Unrelated to Target Objects**

ANCOVAs with total number of utterances as the covariate were performed on the numbers of utterances in four content categories unrelated to target objects: attentionals, backchannel/interjections, nonsense sounds, and references to other people and events. The only content category in which there was a main effect of language group was backchannel/interjections, which were used significantly more often by Japanese mothers ($M = 16.2, SD = 7.6$) than by American mothers ($M = 9.1, SD = 6.1$), $F(1, 53) = 11.27, p < .01$.

Effects of infant age were found in two content categories: attentionals, $F(2, 53) =$
American Mothers

Japanese Mothers

FIG. 3.—Social expressions in Japanese and American mothers' speech to infants: mean numbers of greeting, exchange, and empathy routines.

11.26, *p* < .001, and nonsense sounds, *F*(2, 53) = 38.63, *p* < .001. Follow-up tests revealed that both attentionals and nonsense sounds were used significantly more frequently with 6-month-old infants than with either group of older infants, *p* < .01. The mean numbers of nonsense sounds were 8.1 in speech to 6-month-old infants, 0.2 in speech to 12-month-old infants, and 0.05 in speech to 19-month-old infants. In the case of attentionals, a significant language group × infant age interaction, *F*(2, 53) = 8.16, *p* < .001, reflected the finding that the age-related changes were much more dramatic in the Japanese sample. In Japanese mothers’ speech, the mean numbers of attentionals were 18.7 at 6 months of age, 6.1 at 12 months, and 1.7 at 19 months. In American mothers’ speech, the means for attentionals were 8.1 at 6 months, 6.5 at 12 months, and 5.3 at 19 months. Post hoc tests revealed that the age effect was significant for Japanese mothers, *F*(2, 26) = 13.57, *p* < .001, but not for American mothers. Japanese mothers used attentionals significantly more often when interacting with 6-month-old infants than with 12- or 19-month-old infants, *p* < .001. The numbers of attentionals in Japanese mothers’ speech to the two older groups of infants did not differ significantly from each other.

**Productive Vocabulary in Japanese and American Infants**

Figure 4 summarizes the data on infants’ productive vocabularies at 12 and 19 months, as reported by mothers using Japanese and English versions of the Bates Vocabulary Inventory. The total numbers of object words reported in infants’ vocabularies were analyzed in a 2 × 3 ANOVA. American infants had significantly larger noun vocabularies (M = 49.8, SD = 60) than did Japanese infants (M = 25.6, SD = 28.3), *F*(1, 36) = 4.29, *p* < .05. Not surprisingly, the effect of age on object word vocabulary size was also significant, *F*(1, 36) = 24.02, *p* < .001. A similar analysis on the non-object words in infants’ vocabularies yielded a significant main effect of age, *F*(1, 36) = 18.08, *p* < .001, but no significant effect of language. The language × age interaction was not significant in either analysis.

**Discussion**

When observed playing with their infants at home, Japanese and American mothers accommodated their speech to the special needs of the infant in fundamentally similar ways, by simplifying their speech, repeating themselves frequently, and using interesting sounds to engage the infant’s attention. At the same time, Japanese and American mothers interacted with their infants in ways shaped by beliefs and practices specific to their cultures. American mothers focused more on the target objects overall and emphasized the names of the objects, providing labels frequently and consistently as they played with their infants. Japanese mothers labeled the toys less often than did American mothers, but used them more often in rituals of social exchange, emphasizing...
verbal politeness routines. These comparative data can help in identifying which features of infant-directed speech appear to be universal responses to the immaturity of the human infant, and which features reflect more directly the constraints of a particular language and the worldview of a particular culture.

**Common Strategies in Japanese and American Mothers' Speech to Infants**

Both Japanese and American mothers interacted with their infants using speech that was simplified in several respects. Infant-directed utterances in both language groups were linguistically simple, often containing only one content word. Moreover, 8%-10% of mothers' utterances consisted solely of an isolated content word, a form that rarely occurs in adult-directed conversation in either Japanese or English. The frequent use of nonsense words and verbatim repetition, other infant-directed speech forms rarely used in adult discourse, was also typical of Japanese and American mothers' speech to infants in this study. Linguistic simplification, the use of perceptually salient nonsense sounds, and frequent repetition are all highly appropriate communicative strategies, given the infant's limited attention and memory span and immature conceptual capacities (Newport et al., 1977). These are robust characteristics of speech to children that have been found in numerous other studies of maternal speech in English and other languages (e.g., Ferguson, 1978; Snow, 1977).

Japanese and American mothers also adjusted their speech in similar ways to infants of different ages. Mothers across language groups focused less on objects overall in speech to 6-month-old infants. They also used more attention-getting nonsense sounds in speech to 6-month-old infants, unskilled in deploying attention selectively, than in speech to older infants. Mothers across cultures also used exact repetitions more frequently with younger infants, whose memory is less well developed. With 19-month-old infants engaged in learning language, both Japanese and American mothers labeled objects and asked questions more frequently than in speech to younger infants, a shift in emphasis appropriate to the developing linguistic capabilities of the older infant.

Similar patterns of appropriate age-related adjustments were also evident in the varied uses of nonsense sounds by Japanese and American mothers. Nonsense sounds in mothers' speech were analyzed in three different contexts in this study: as onomatopoeic labels for target objects, as sounds attributed to target objects, and as playful, attention-getting sounds unrelated to target objects. Age-related changes in each of these contexts indicated that the functions of nonsense sounds varied with the developmental level of the infant. Nonsense sounds used as attentionals occurred most frequently with 6-month-olds across languages and very rarely with older infants. Sounds attributed playfully to target objects occurred more of-
ten with 12-month-old infants than with 6- or 19-month-old infants across languages. Although limited to Japanese mothers’ speech, the use of nonsense sounds as onomatopoeic labels for target objects was most common in speech to older infants. Thus different functions peaked at different ages, with the attentional and affective uses of nonsense sounds more prevalent in speech to younger than older infants across languages, and the use of nonsense sounds to label objects in Japanese increasing in speech to older infants beginning to learn language. The common characteristics observed in Japanese and American mothers’ speech, along with similar patterns of age-related adjustments in speech to infants at 6, 12, and 19 months, suggest that certain features of infant-directed speech are motivated by an intuitive sensitivity to the developmental level of the infant shared by mothers across cultures (Fernald, 1992).

**Differences between Japanese and American Mothers’ Talk about Objects**

This study also revealed striking differences in communicative style as Japanese and American mothers talked to their infants about objects. Japanese mothers were less likely to talk about the objects overall, and less likely to provide a noun label when referring to a target object. While more than half of all utterances related to the target objects contained noun labels in American mothers’ speech, only a third of Japanese mothers’ target-related utterances contained noun labels. Japanese and American mothers also differed in the consistency with which they labeled target objects. When they did provide a noun label for a toy, Japanese mothers were much less likely than American mothers to use the adult form of the word for the object. Because Japanese mothers typically alternated among several “babytalk” variants available as labels for each target object, there was a much higher probability that they would switch labels when referring to the same object in successive utterances than in American mothers’ speech. Moreover, Japanese mothers often used the onomatopoeic word *buubuu* to name the pig, the car, and the truck. Since *buubuu* in Japanese represents both the sound made by a pig (*oink oink* in English) and the sound made by vehicles (*vroom vroom* in English), the same word served as a label for three of the four target toys. Thus Japanese mothers labeled objects less frequently overall, less often in the correct adult form, and with greater lexical variability and potential ambiguity than did American mothers. Such differences in the availability, form, and consistency of noun labels in early linguistic input are factors that could serve to make object words relatively less salient to Japanese than to American infants.

Another difference in Japanese and American mothers’ treatment of noun labels emerged from the analysis of questions. Mothers in both language groups used questions frequently; on average, 20% of all utterances consisted of questions in Japanese, and 30% in English. However, American mothers used questions much more often in the context of labeling target objects than did Japanese mothers. Fernald and Mazzie (1991) have suggested that presenting an object label in a question format may have perceptual benefits. When a question is used to introduce a label to an infant in English, the noun is most likely to appear at the end of the utterance on an exaggerated pitch peak, a perceptually prominent position in the speech stream. Thus, the differential use of questions to label objects in American mothers’ speech to infants is another factor that could make nouns relatively more salient to English-learning than to Japanese-learning infants. However, it should also be noted that nouns occurred in isolation more often in Japanese than in American mothers’ speech, a labeling format that could also facilitate the infant’s perception of object words.

**Are Nouns and Verbs Differentially Emphasized in English and Japanese?**

Gopnik and Choi (1990) have argued that English is a “noun-dominant” language, relative to more “verb-dominant” languages like Korean and Japanese, and that such structural differences could influence the rate of acquisition in these languages. In particular, infants learning languages in which nouns are emphasized may acquire nouns more rapidly than infants learning languages in which nouns are less salient. Consistent with this prediction, we found that American infants had larger noun vocabularies than did Japanese infants at 19 months, according to maternal report. Cross-language production data based on maternal report must be regarded with considerable caution, however, for two reasons: First, the Bates Vocabulary Inventory has been standardized for English-learning infants, and may be less appropriate for infants learning Japanese. And, second, given cultural proscriptions against praising one’s own children in public, Japanese mothers may have tended to
underestimate their infants' verbal ability. However, the finding that American infants were reported to be advanced in productive noun vocabulary relative to Japanese infants is consistent with Gopnik and Choi's more reliable observational findings on advanced lexical development in American infants relative to Korean infants.

To what extent could structural features of English and Japanese account for the observed differences in the tendency of American and Japanese mothers to emphasize object words in speech to their infants? Gopnik and Choi (1990) suggest that the prevalence of noun deletion, or nominal ellipsis, in Korean and Japanese results in relatively less emphasis on nouns in speech to infants in these languages than in English. The implicit assumption in this argument is that typological differences between languages are primarily responsible for cross-linguistic differences in the nature of early language input. However, our results suggest that the greater emphasis on nouns in American mothers' speech was not simply a consequence of structural differences between English and Japanese, but reflected referential choices made by American and Japanese mothers. Although nominal ellipsis is not a grammatical option in English, nouns can be omitted by replacing them with pronouns after a noun has been established as "old" information. In fact, Clancy's (1980) comparison of the frequency of noun phrases, pronouns, and ellipsis in adult narratives in English and Japanese indicates that noun phrases are omitted even more frequently in English than in Japanese. If American mothers speaking to infants had replaced nouns with pronouns in repeated references to a target object, as is done typically in adult-directed speech, the frequency of labeling would have been more similar in English and Japanese. Instead, American mothers chose to name the target object repeatedly, while Japanese mothers were more likely to omit the noun label in repeated references to a target object.

Gopnik and Choi (1990) also hypothesize that a relatively greater emphasis on verbs in Japanese and Korean than in English could contribute to cross-language differences in the rate of noun acquisition. However, the mean numbers of action verbs in utterances related to target objects was actually significantly greater in American mothers' speech than in Japanese mothers' speech. Moreover, within the Japanese sample there was no evidence for a bias toward verbs, since Japanese mothers labeled objects (M = 24.5) as often as they labeled actions by or upon these objects (M = 18.4). Could the option of nominal ellipsis, which allows verbs to occur in isolation in Japanese, make verbs relatively more salient in Japanese mothers' speech? The analysis of content words occurring in isolation showed no significant differences in the frequency with which Japanese and American mothers isolated verbs (see Fig. 2). More importantly, both Japanese and American mothers used isolated nouns as often, on average, as they used isolated verbs, although verb ellipsis is not a grammatical option in either Japanese or English. One potentially important difference between Japanese and American mothers' use of isolated verbs in this study was that American mothers used only attentional verbs such as look and see in isolation, while Japanese mothers used a wider variety of verbs as single-word utterances, including asubo (play), suru (do), ai-ta (opened), it-chat-ta (gone), okkot-chat-ta (fell down), and nainai (put away), among others.

To summarize, the Japanese infants in this study heard object labels much less frequently and less consistently than did the American infants, but this difference in the nature of early language input was not dictated by the grammatical structure of Japanese and English. Nouns can be optionally omitted in both languages, either by substitution of a pronoun in English, or by nominal ellipsis in Japanese. American mothers exercised this option less often, providing a noun label for the target object in more than half of the utterances related to the toys. Moreover, American mothers preferred to use the adult word for the target object and switched labels for the same referent less often than did Japanese mothers. Such differences in the frequency, form, and consistency of object words in American and Japanese mothers' speech to infants could have contributed to the reported differences in rate of noun vocabulary growth in infants learning English and Japanese.

This is not to say that structural differences between Japanese and English have no influence on the relative salience of object words in early language input. Nouns in American mothers' speech may indeed be perceptually enhanced by virtue of the SVO word order of English and the use of exaggerated pitch (Fernald & Mazzie, 1991), al-
though Japanese mothers can achieve comparable effects by presenting nouns more frequently in isolation. Moreover, verbs may be relatively more accessible in Japanese because they typically occur in final position. The important point to be made here is that speakers have stylistic options within the grammar of their language. American and Japanese mothers chose to talk about objects in different ways, and these referential choices contributed to the differential emphasis on object words observed in English and Japanese language input. In this sense, it is the infant-directed speech register in American middle-class culture, rather than the English language per se, which emphasizes nouns, and which more plausibly accounts for any differences in rate of noun acquisition between American and Japanese infants.

Cultural Influences on Mothers' Speech to Infants

Cross-linguistic differences in how a mother speaks to her infant are influenced by many factors other than the grammar of the language she is speaking. Some features distinguishing Japanese and American mothers' speech appear to be directly related to cultural conventions of language use. For example, the more frequent use of backchannel vocalizations by Japanese mothers is not surprising, given that backchannels occur three times more often in conversations among Japanese adults than among Americans (Maynard, 1986). By signaling the listener’s ongoing attention to the speaker, backchannels are thought to be important in maintaining omoiyari, a Japanese concept referring to harmony in interactions (White 1989).

The influence of cultural values was also evident in the tendency of American and Japanese mothers to focus on different aspects of the social situation. When playing with toys with their infants, American mothers called attention to object names, while Japanese mothers used the toys more often to engage their infants in social routines. For example, a typical sequence of utterances related to a target object in American mothers’ speech was: That's a car. See the car? You like it? It's got nice wheels. In Japanese mothers’ speech, in contrast, polite verbal routines accompanying the exchange of the object were likely to be emphasized, often omitting the name of the object after the first mention: Hai buubuu. (Here! It’s a vroom vroom.) Hai doozo. (I give it to you.) Hai kore chooodai. (Now give this to me.) Choo-dai. (Give me.) Hai arigatoo (Yes! Thank you.) For Japanese mothers, the name of the object often appeared to be irrelevant. What was important was the polite exchange of the object in the ritual game of give and take with the infant.

Japanese mothers were also more likely than American mothers to engage their infants in empathy routines, encouraging the infant to show positive feelings toward a toy, as in the following example: Hai wan-chan (Here! It's a doggy.) Kawaii kawaii shi-te age-te. (Give it a love.) Kawaii kawaii kawaii. (Love love love.) As she said Kawaii kawaii, the Japanese mother encouraged the infant to pet the dog gently. These findings are consistent with Clancy’s (1986) observation that Japanese mothers are concerned with actively and explicitly teaching their infants cultural norms for polite speech. Hess, Kashiwagi, Azuma, Price, and Dickson (1980) also report that Japanese mothers expect children to master such social expressions at an earlier age than do American mothers.

Maternal speech is also inevitably influenced by culturally transmitted views about the nature of the mother-infant bond. In their relationship with the infant, Japanese mothers encourage amae, roughly translated as mutual dependence (Doi, 1973), while Americans place more value on fostering independence in their children. According to Fischer (1970), this cultural difference in child-rearing philosophies is reflected in the tendency of Japanese mothers to use “babytalk” words more extensively and for a longer period than do American mothers. Indeed, we found that even in speech to 19-month-old infants, 21% of Japanese mothers’ utterances contained nonsense or onomatopoeic words, compared to 7.5% in American mothers speech to older infants. Japanese mothers also used linguistically simplified speech with their infants to a greater extent overall than did American mothers. When interviewed informally about why they modified their speech to the infant, Japanese mothers typically explained that their goals were to talk gently and to use sounds the infant could imitate easily. American mothers were more likely to report that their goals were to attract the infant’s attention and to teach the infant words. These observations suggest that the Japanese mothers were more interested in directly fostering linguistic competence, and more
interested in establishing affective communication and accommodating their own speech to the immature productions of the infant. Such accommodation may also reflect the ideal of *omoiyari*, the maintenance of harmonious interactions that is valued so highly in Japanese culture (see Lebra, 1976).

The form and content of mothers' speech to infants are determined by a complex interaction of biological, linguistic, and cultural factors. Universal characteristics and response tendencies of the human infant motivate culturally widespread features of infant-directed speech, such as linguistic simplification and the use of exaggerated prosody and frequent repetition (Ferguson, 1978; Fernald, 1992). Other features of infant-directed speech, such as how frequently, how consistently, and in what format objects are named for the infant, are more variable across cultures. Whether or not the variability in how Japanese and American mothers talk about objects has consequences for infants' rate of language learning is a question that needs to be addressed by future research using larger samples and more reliable measures of language acquisition. A question of equal importance is how cross-cultural differences in maternal speech influence learning in other domains such as social cognition. From the mother's speech, the infant begins to acquire not only the rules of a language, but also the norms of a culture.

References


This document is a scanned copy of a printed document. No warranty is given about the accuracy of the copy. Users should refer to the original published version of the material.