

Functional morphemes in Japanese mothers' speech input to their infants

Division of Educational Psychology Etsuko HARYU
Tamagawa University Sachiyo KAJIKAWA

Etsuko HARYU, Sachiyo KAJIKAWA

In Japanese, *joshi* or functional morphemes attached to nouns are often omitted in colloquial Japanese speech, while verbs never occur without functional morphemes, i.e., verb suffixes. The present study explored how frequently Japanese-learning infants hear *joshi* (or noun particles) and verb suffixes in the speech input. Twenty mothers' speech to their infants aged 10 and 15 months was analyzed. The results showed that more than half of the noun tokens in mothers' utterances were followed by no particle, and that the most frequent particle *ga* followed 12% of the noun tokens. Verbs were followed by a limited number of verb suffixes at an extremely high frequency, and the most frequent verb suffix *-te-iru* occurred more frequently than the most frequent noun particle *ga*. The occurrence of the remaining verb suffixes was quite infrequent. Influences of such distributions of functional morphemes on children's learning of syntactic categories of nouns and verbs were discussed.

1. Introduction

It is well known that young children omit functional morphemes in their early speech (Brown, 1973). These morphemes, or functors, include verb inflections, articles, and noun particles, known in Japanese as *joshi*. However, recent studies have shown that even young children who omit functional morphemes in their production do recognize these elements in speech. For instance, when English-speaking 2-year-olds imitate sentences containing either functors (e.g., "Pete pushes the dog") or non-functor syllables (e.g., "Pete pusho na dog"), they omit functors but not non-functor syllables (Gerken, Landau, & Remez, 1990). Furthermore, eighteen-month-olds find the referent of a familiar noun more successfully and rapidly when the noun is preceded by an appropriate article, as in the utterance "Find the bird", than when the article is omitted, as in "Find bird" (Fernald & Hurtado, 2006; Kedar, Casasola & Lust, 2006). Even preverbal infants use functional morphemes to segment and syntactically categorize adjacent words in the speech stream (Höhle & Weissenborn, 2000; Höhle, Weissenborn, Kiefer, Schluz, & Schmitz, 2004; Shi, Culter, Werker, & Cruickshank,

2006; Shi & Lepage, 2008; Shi & Melançon, 2010).

Functional morphemes are necessary to establish the structural skeleton of a sentence. They are organized into a small, closed class of word types, each member of which has a high token frequency compared with content words. As a result, a limited number of class members appear repeatedly in utterances. This frequent occurrence of functional morphemes contributes to children's early recognition of these elements. In fact, infants come to use frequent determiners as a cue to word segmentation earlier than infrequent ones (Shi, Cutler et al., 2006; Shi & Lepage 2008). Further, infants' use of functional morphemes reflects the occurrence of functors in the language they are learning. French-learning infants begin to show the ability to use determiners for word segmentation earlier (8 months; Shi & Lepage 2008) than do English-learning infants (11 months; Shi, Cutler et al., 2006). This is probably because determiners are more frequent in French than in English: Not only singular nouns but also plural and mass nouns must be preceded by some determiner in French, whereas plural and mass nouns need not be preceded by a determiner in English.

Furthermore, both German- and French-learning infants

at around the age of 15 months are able to categorize novel words in the speech stream into a noun class by attending to preceding determiners, while at the same time they fail to syntactically categorize verbs by using preceding pronouns as a cue (Höhle et al., 2004; Shi & Melançon, 2010). This may be attributed to the fact that, in both German and French, the co-occurrence relationship between content words and functional morphemes are stronger for nouns than for verbs. In fact, Höhle et al., (2004) reported that 71% of appearances of the indefinite article *ein* in a corpus of German child-directed speech were immediately followed by a noun, whereas only 31% of appearances of the pronoun *sie* were followed by a verb in the same corpus. Likewise, 71% of determiners in Canadian French-speaking mothers' utterances were followed by a noun, whereas subject pronouns were followed by a verb 59% of the time (Shi & Melançon, 2010).

Previous studies have thus suggested that the distribution of functional morphemes in the speech input plays an important role in infants' learning. From this point of view, the Japanese language has at least two interesting characteristics. First, the Japanese language has particles or *joshi* as functional morphemes attached to nouns, which are frequently omitted in casual speech. Cook (1985) analyzed a mother's speech to her 2-year-old child and reported that the subject-marking particle *ga* was dropped 16% of the time, while the object-marking particle *o* was omitted 82% of the time. Our previous study of 25 Japanese mothers' speech to their infants around the age of 12 months also demonstrated that the subject-marking particle *ga* was dropped an average of 35% of the time, and the object-marking particle *o* was omitted an average of 91% of the time (Kajikawa & Haryu,

2008). This pervasive omission of noun particles must lower the frequency with which Japanese-learning infants hear particles in the speech input, making it difficult for them to learn to use particles to segment and syntactically categorize adjacent words.

A second characteristic of Japanese is that verbs never occur without verb suffixes, in contrast to nouns for which particles are often omitted. According to Shibatani (1990), a verb stem consists of a root and its inflectional ending, which is followed by an auxiliary, a particle, or both, as shown in (1).

$$\begin{array}{c} \text{Root} + \text{Inflectional ending} (+ \text{Auxiliary and/or Particle}) \\ \hline \text{Verb stem} \end{array} \quad (1)$$

Auxiliaries and particles mark preceding verbs for tense, aspect, voice, negation, and so on. An inflectional ending is a conjunctive element that enables the verb stem to be attached to an auxiliary and/or a particle. A verb has several forms of inflectional ending, each of which corresponds to one of five inflectional categories. Table 1 shows inflected forms of two verbs, *hashi-ru* ('run') and *tabe-ru* ('eat'), which exemplify two major types of inflection in Japanese verbs. Although inflected forms of a verb do not correspond to inflectional categories in one-to-one fashion, the inflectional ending that is attached to a verb root is determined by the auxiliary/particle that follows the verb stem. That is, these inflectional categories are quasi-semantic and quasi-functional. Because a verb root is followed by a combination of an inflectional ending and an auxiliary/particle in utterances, we may also regard the whole sequence of an inflectional ending and a following

Table 1 Verb inflectional categories, inflected forms, and auxiliaries/particles that can follow each ending.

Inflectional categories	Inflected forms (root – inflectional ending)		Following auxiliaries and particles
	'run'	'eat'	
<i>Mizen</i> (Irrealis)	<i>hashi – ra</i> <i>hashi – ro</i>	<i>tabe –</i>	<i>nai</i> (Negative), <i>su, seru</i> (Causative) <i>u, yō</i> ('will')
<i>Ren'yō</i> (Adverbial)	<i>hashi – ri</i> <i>hashi – t</i>	<i>tabe –</i>	<i>masu</i> (Polite), <i>tai</i> (Desiderative) <i>ta</i> (Past), <i>te</i> (Conjunctive), -
<i>Shūshi</i> (Conclusive) ¹⁾	<i>hashi – ru</i>	<i>tabe – ru</i>	<i>rashii</i> ('hearsay'), <i>mitaida</i> ('appear')
<i>Katei</i> (Hypothetical)	<i>hashi – re</i>	<i>tabe – re</i>	<i>ba</i> ('if')
<i>Meirei</i> (Imperative) ²⁾	<i>hashi – re</i>	<i>tabe – ro</i>	

1) Conclusive forms can occur without any following auxiliary/particle.

2) Neither auxiliary nor particle follows an imperative verb form.

auxiliary/particle as a verb suffix, as shown in (2).

$$\begin{array}{c} \text{Root} + \text{Inflectional ending (+ Auxiliary and/or Particle)} \\ \hline \text{Verb suffix} \end{array} \quad (2)$$

In sum, noun particles are likely to be omitted in colloquial Japanese speech, but a verb is unlikely to occur without an inflectional ending or a verb suffix. These structural characteristics of Japanese lead us to expect that Japanese-learning infants may have more difficulty in learning to use particles for noun segmentation and categorization than in learning to use verb inflectional endings for verb segmentation and categorization. However, it is unknown how frequently Japanese-learning infants in fact hear each noun particle and verb inflectional ending in the speech input. Therefore, we investigated the frequency and distribution of functional morphemes such as noun particles and verb suffixes in speech directed at Japanese-learning infants. More specifically, we analyzed Japanese mothers' speech to their preverbal infants by counting and categorizing nouns (concrete nouns) and verbs according to the functional morphemes that follow them. In so doing, we asked the following specific questions:

1. How frequently do Japanese-learning infants hear each noun particle in the speech input?
2. How frequently does each verb suffix follow verb roots in the speech input to Japanese-learning infants?
3. In what way are these functional morphemes for nouns and verbs distributed in utterances addressed to Japanese-learning infants?

2. Method

Participants.

Ten mothers whose infants were 10 months old (mean age: 321 days, range: 306 – 355 days) and 10 mothers whose infants were 15 months old (mean age: 460 days, range: 436 – 481 days) participated. Mothers of 10- and 15-month-olds were focused on in the current study, because previous studies have reported that (a) infants learning German, English, and French begin to show the ability to use determiners for segmentation and syntactic categorization of adjacent words between the ages of 8 and 15 months (Höhle & Weissenborn, 2000; Höhle et al., 2004; Shi, Culter et al., 2006; Shi & Lepage, 2008; Shi

& Melançon, 2010); and (b) Japanese children begin to use *joshi* 'particles' in their speech production at 18 to 22 months of age (Okubo, 1967).

Speech sample collection procedure.

The mother-infant interactions were video-recorded in a university playroom. The mothers were asked to describe 10 pictures to their infants, each of which depicted an everyday activity such as a boy reading a picture book, a woman making tea for guests, and so on.

Transcription and analyses.

Mothers' utterances across all 10 pictures were transcribed and the number of utterances was counted. Utterances comprising singing, counting, interjections and sound effects alone were excluded. Across all utterances, types (the number of different words used) and tokens (the number of times those words appeared) were counted for concrete nouns and verbs. As concrete nouns, we counted common nouns, either in conventional form (e.g., *neko* for 'cat') or in baby form (e.g., *nyanko* for 'kitty'), and proper nouns including the infants' own names. Locative nouns (e.g., *tonari*, 'next door'), formal nouns (e.g., *koto*, 'that'), pronouns, and number words were excluded. Verbs that were counted referred to activity (e.g., *nomu*, 'drink'), change of location (e.g., *iku*, 'go') and existence (e.g., *aru*, 'be'). For each noun extracted, we coded whether or not it was followed by a particle, and which particle, if any, followed it. Functional morphemes for verbs were analyzed by counting the frequency of each type of inflectional ending and verb suffix.

To determine the variability of the occurrence of functional morphemes with nouns and verbs in the speech input, scores of morphological variability were calculated for each word type (nouns, verbs) overall within each mother. That is, we first counted the numbers of different morphemes that occurred with individual nouns and verbs within each mother, and the mean number was calculated for each word type (nouns, verbs) overall within each mother. For nouns, we counted the number of different particles that occurred with each noun in mothers' speech across all 10 pictures. For instance, when a mother used the noun *nyanko* in uttering such phrases as *nyanko-ga* (*nyanko*, 'cat' in baby form, followed by the subject-marking particle *ga*), *nyanko-da* (sentence-final particle), and *nyanko* (no particle, zero form), the noun was regarded as having been used in three different morphological contexts. Therefore, the morphological variability score for the word *nyanko* was three. In addition,

if the noun *kaba* ('hippopotamus') was used only in zero form, that is, without being followed by any particle, its morphological variability score was one. If the two words *nyanko* and *kaba* were the only nouns the mother used while describing all 10 pictures, her mean morphological variability score for noun was two. Likewise, for verb morphological variability, the number of different verb suffixes that followed each verb root was counted. For instance, if a mother used two verb forms *tsuku-ra-nai* and *tsuku-t-te* for one verb root *tsuku* 'make' during the description of all 10 pictures, the morphological variability score for the verb was two. In addition, if the same mother used only one form *tabe-te* (verb root *tabe* 'eat' followed by verb suffix *te*) for another verb, the score for the verb was one. If the two were the only verbs the mother used during the interaction with her infant, her mean morphological variability score for verb was 1.5.

3. Results

Preliminary analyses showed no significant difference between mothers of 10- and 15-month-olds. Therefore, we report the results obtained by combining the two data sets. The mean number of utterances mothers produced in describing all the ten pictures was 45.8 (SD 21.9). As shown in Table 2, each mother produced an average (SD) of 23.1 (7.3) noun types and 47.4 (20.0) noun tokens. Of these, 26 noun tokens (54.9%) occurred in zero form. That is, they were uttered in isolation or followed directly by another content word without any particle inserted. Each of the remaining 21.4 noun tokens (45.1%) was produced with one of 20 different noun particles. The 8 most frequently occurring particles, each of which followed more than 2 percent of noun tokens in mothers' utterances, are shown in Table 2. The subject-marking particle *ga* was the most frequent. It occurred after 12.0% of concrete noun tokens, although this particle is often omitted in colloquial Japanese speech. The next four most frequent particles appeared 4.2 to 6.5% of the time. Thus, although infants heard the subject-marking particle *ga* more frequently than any other particle, the distribution of noun particles was relatively uniform.

As for verbs, mothers produced an average (SD) of 16.3 (6.1) verb types and 27.9 (13.1) tokens. The verb tokens were categorized according to inflectional ending or category. Table 2 shows the distribution. The inflectional ending *-t* which is categorized as *ren'yō* (adverbial) occurred with

extremely high frequency, with 79.6% of verb tokens. The second most frequent was the inflectional ending *-ru* of the *shūshi* (conclusive) category, which occurred with an average of 18.0% of verb tokens and was as frequent as the most frequent noun particle *ga*, $t(19) = 0.64$, $p = .531$, $d = 0.14$. The occurrence of other inflectional endings was rare. Interestingly, the most frequent *ren'yō* inflectional ending *-t* (22.2 verb tokens) was followed by the complex suffix *te-iru* consisting of the conjunctive particle *te* and the auxiliary verb *iru* 66.3% of the time (14.7 tokens). Thus the distribution of verb inflectional endings was biased.

The type-token ratio, which is calculated by dividing the number of tokens by the number of types, was 2.03 for nouns and 1.69 for verbs. The difference was statistically significant, $t(19) = 3.73$, $p < .001$, $d = 0.84$. Thus, a noun was more likely to be repeated than a verb in mother's speech. Scores of morphological variability were calculated for nouns and verbs per mother. The mean (SD) score of morphological variability was 1.37 (0.14) for nouns and 1.08 (0.08) for verbs. A paired t test revealed that nouns were morphologically more variable than verbs in mothers' utterances, $t(19) = 7.48$, $p < .001$, $d = 1.67$. In fact, mothers produced an average of 26.2% of noun types with different particles in multiple contexts, whereas only 7.8% of verb types were used with multiple suffixes.

4. Discussion

In Japanese, noun particles are known to be often omitted in casual speech, whereas a verb never occurs without a verb suffix. In the present study, Japanese mothers' speech to their infants was analyzed by focusing on the frequencies and distribution of noun particles and verb suffixes. The analyses demonstrated that more than half the concrete nouns in mothers' utterance were not accompanied by a particle. The most frequent noun particle was the subject marker *ga*, which followed 12.0% of concrete nouns. The frequency with which the particle appeared in utterances was significantly lower than that of the most frequent verb suffix, *-te-iru*, $t(19) = 4.45$, $p < .001$, $d = 1.00$. The results indicate that Japanese-learning infants hear one verb suffix more frequently than any noun particle.

However, to be able to use functional morphemes as cues to word segmentation, children need not only to hear those elements frequently in speech, but also to notice that

Table 2 Frequency and percentage of occurrence of each functional morpheme per mother

		Frequency	Percentage
Concrete nouns			
Type		23.1	
Token		47.4	100.0
Noun particles/auxiliaries			
<i>ga</i>	subject marker	5.7	12.0
<i>da</i>	auxiliary	3.1	6.5
<i>kana</i>	sentence-final particle, 'uncertain'	2.3	4.9
<i>o</i>	object marker	2.2	4.6
<i>mo</i>	topic/subject marker, 'also'	2.0	4.2
<i>no</i>	possessive marker	1.5	3.2
<i>ni</i>	indirect object / locative marker	1.3	2.7
<i>to</i>	conjunctive particle, 'and'	1.1	2.3
12 other types		2.2	4.6
No particle		26.0	54.9
Verbs			
Type		16.3	
Token		27.9	100.0
Inflectional categories			
<i>Mizen</i>	Irrealis	0.5	1.8
<i>Ren'yō</i>	Adverbial	22.2	79.6
	further followed by		
	<i>te</i> (conjunctive)	3.3	
	<i>te-iru</i> (present progressive)	14.7	
	<i>te-masu</i> (present progressive, polite)	1.7	
	<i>ta</i> (past tense)	1.1	
<i>Shūshi</i>	Conclusive	5.0	18.0
<i>Meirei</i>	Imperative	0.2	0.7

those elements are separable from the adjacent content word. Furthermore, to use functional morphemes for syntactic categorization of adjacent words, children need to understand that the same content word may occur with a different functional morpheme of the same grammatical category in a different context. For instance, a word which is followed by the subject-marking particle *ga* in one context may be followed in another context by the object marker *o*. Thus, in order for children to learn to use functional morphemes in word learning, it is not enough for them to observe a particular functional morpheme very frequently. Children also need to observe that the same content word can occur with multiple members of the same grammatical category of noun particles or verb suffixes. In this respect, not only frequency but also variability of functional morphemes that occur with a noun or a verb may affect how easily children learn functional morphemes during language development.

From this point of view, the distribution of noun particles

was relatively uniform. Besides the most frequent particle *ga*, there were four frequent particles (*da*, *kana*, *o*, and *mo*), each of which occurred with 4 to 7% of noun tokens. Such a distribution may help children learn the syntactic category of noun particles. In contrast, an extremely high proportion of verb tokens were followed by the complex verb suffix *-te-iru*, which marks the preceding verb for present progressive aspect. In addition, the average score of morphological variability for verbs was around one. That is, a particular verb root tended to always occur with a particular verb suffix in the speech input. This distribution of verb suffixes may make it difficult for children to learn that verb suffixes are separable from the adjacent verb root and to use them to syntactically categorize an adjacent word into a verb class. Future research should explore the influence of such distributions by comparing the ages at which infants begin to use functional morphemes to segment and syntactically categorize nouns and verbs.

Acknowledgements

The authors are grateful to the mothers and their infants who participated in the study. This research was supported by JSPS KAKENHI (Grant Number 25285183).

References

- Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.
- Cook, H. M. (1985). Frequency of nominal markers in the speech of a Japanese child and his caretakers: A case study. *Descriptive and Applied Linguistics*, 18, 13-24.
- Fernald, A. & Hurtado, N. (2006). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, 9, F33-F40.
- Gerken, L. A., Landau, B., & Remez, R. E. (1990). Function morphemes in young children's speech perception and production. *Developmental Psychology*, 26, 204-216.
- Höhle, B., & Weissenborn, J. (2000). The origins of syntactic knowledge: Recognition of determiners in one year old German children. In C. Howell, S. A. Fish, & T. Keith-Lucas (Eds.), *BUCLD 24 Proceedings* (pp.418-429). Somerville, MA: Cascadia.
- Höhle, B., Weissenborn, J., Kiefer, D., Schluz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, 5, 341-353.
- Kajikawa, S., & Haryu, E. (2008). Nyūji ni okeru joshi 'ga' no ninshiki. [Recognition of a function morpheme, ga, in infants]. *Tamagawadaigaku Nōkagakukenyūjokiyō*, 2, 13-21.
- Kedor, Y., Casasola, M., & Lust, B. (2006). Getting there faster: 18- and 24-month-old infants' use of function words to determine reference. *Child Development*, 77, 325-338.
- Okubo, A. 1967 *Yooji gengo no hattatsu* [Language development during young childhood]. Tokyo-do Shuppan.
- Shi, R., Culter, A., Werker, J. F., & Cruickshank, M. (2006). Frequency and form as determinants of functor sensitivity in English-acquiring infants. *Journal of Acoustical Society of America*, 119, EL61-67.
- Shi, R., & Lepage, M. (2008). The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science*, 11, 407-413.
- Shi, R. & Melançon, A. (2010). Syntactic categorization in French-learning infants. *Infancy*, 15, 517-533.
- Shibatani, M. (1990). *The language of Japan*. Cambridge: Cambridge University Press.