

Construction of Japanese Sign Language Dialogue Corpus: KOSIGN

Yuji Nagashima, Mina Terauchi, Kaoru Nakazono

Kogakuin University, Polytechnic University, NTT Network Innovation Laboratories,

2665-1 Nakano, Hachioji, Tokyo 192-0015, Japan

Hashimoto-dai 4-1-1, Sagami-hara, Kanagawa, 229-1196, Japan

3-9-11 Midoricho, Musashino-shi, Tokyo 180-0012, Japan

E-mail: nagasima@cc.kogakuin.ac.jp, terauchi@uitech.ac.jp, nkzn@mac.com

Abstract This report presents a method of building corpuses of dialogue in Japanese Sign Language (JSL) and the results of the analysis in co- occurrences of manual and non-manual signals using the corpus. We have built the sign dialogue corpus by video recording the dialogues between native JSL signers. The purpose of building corpus is deriving electronic dictionaries such as morphological dictionary, different meaning word dictionary, allomorph dictionary and example dictionary. Example sentences are recorded for every word (key sign) those were recorded in the sign language word data base KOSIGN Ver.2. We were able to confirm a correlation of manual and non-manual signals or a characteristic appearance of sign language dialogue. As a result of the analysis, the pointing occurred to the end of sentence at high frequency. It suggested that pointing be one of the ends of sentence, and clarified the role as the conjunctive pronoun. The co-occurrence relation between the manual and non-manual signals acquired confirmed an important role to make the meaning of the expression sign language limited was achieved. Moreover, "Roll shift" and "Sandwich construction" that was the linguistic feature of sign language were confirmed, too.

1. Introduction

The engineering field has been applied to sign language by means of the development of a translation system between Japanese and sign language, construction of an electronic dictionary system, and study of a potential sign language recognition system. Moreover, linguistic analysis has been performed based on phonology, morphology, syntax, and structural analysis. In comparison with American Sign Language (ASL), Japanese Sign Language (JSL) has been studied much less and there is little data on the structural analysis of conversation and communication; further study is required.

Unlike spoken languages, which are expressed in one-dimensional fashion, sign languages are visual, expressed in a three-dimensional space. Sign language consists of manual signals and non-manual signals. Manual signals consist of hand shapes, palm directions, exhibited positions, and overall movements; sign language is mainly constructed by showing these motions simultaneously or continuously. Non-manual signals consist of facial expressions, mouth formations, nods, and line-of-sight motions, all of which are considered to work syntactically and semantically. Thus, various body motions are compounded to express information.

It is uncertain that which information is meaningful to which vocabulary word, and which information is a grammatical marker; in other words, the linguistic structure of sign language is unknown. Since a sign language notation system has not yet been systematically constructed, linguistic studies have made little progress. However, the construction of a sign language notation system has been tried by the SIGINDEX Workgroup. In order to explicate the linguistic structure, semantic structure, and cognitive mechanism of sign language

using linguistic analysis, accumulation of linguistic data called the corpus is essential. By the end of last year, the authors had photographed JSL dialogue samples in the participation of native signers and children of deaf adults (CODA) based on the sign language word database (KOSIGN Ver. 2), which contains 1,096 signs. One of our purposes was to digitalize dictionaries, such as the morpheme dictionary, homograph dictionary, synonym dictionary, and citation dictionary. Moreover, the authors have been analyzing inflectional endings and assimilation of words, grammatical functions of non-manual motions, and cognitive mechanism in a dialogue. Collocation in dialogue corpora was analyzed, and its characteristics were reported.

2. Collection method for dialogue corpora

Here, the outline of a dialogue corpus collection method is described. The dialogists participated in making dialogue corpora were the following two females: (1) Signer A, a Child of Deaf Adult (CODA) whose first language was sign language, and (2) Signer B, a congenitally hearing-impaired child whose parents were normal listeners. A dialogue sample was created against each vocabulary word (hereafter referred to as a key vocabulary word) in KOSIGN Ver. 2, according to the following procedure:

- (1) Before recording, signers A and B discussed the dialogue's scenario while looking at an image of the key vocabulary word.
- (2) The scenario was created, in which signer A asked a question (hereafter referred to as C1) and signer B gave an answer including the key vocabulary word (hereafter referred to as C2).
- (3) Videotaping was initiated.
- (4) Signer A asked C1 and signer B answered C2.

(5) In order to create a natural dialogue, signer A immediately gave an answer (hereafter referred to as C3) to C2, without any scenarios after the dialogue between C1 and C2.

(6) Signer B gave an answer (hereafter referred to as C4) to C3 without any scenarios.

Thus, a dialogue sample, with two correspondences regarding a key sign, was created. While the conversation was active, the dialogue continued. With this method, a natural dialogue could be created in the second correspondence, that is C3 and C4. When the sign language system adopted in KOSIGN Ver. 2 key differed from that used daily by signer A or B, that of signer A or B was preferentially adopted. Vocabulary words with low frequency of use in KOSIGN Ver.2, such as proper nouns, were deleted as candidates for key vocabulary words.

In the recording signers A and B did not directly dialogue. Two cameras recorded two signers, respectively, and each signer dialogued while looking at the projected front image of the other signer. Actually, signer A looked at signer B's life-size image, which was projected by a projector, and signer B looked at signer A's image, which was projected on a prompter. The size and position of each projected image were adjusted so that the signer's line-of-sight position was almost the same as that of a normal dialogue.

A camera (BVP-550, SONY Corporation, Japan) was used for recording at a shutter speed of 1/125 s. The front image of each signer was synchronously recorded using a recorder (DVW-500, SONY Corporation, Japan). During recording, in order to reduce the Japanese influence as much as possible, the prepared sample manuscript was not exhibited. The dialogue was initiated with expression of a KOSIGN sign, which would become a key sign, in sign language by one signer. Approximately 3,800 samples have so far been recorded, which include 944 KOSIGN signs.

3. Analysis method for dialogues

With reference to the dialogues created by the participants, and roughly extracted sign information, manual signals and non-manual signals were temporally extracted, together with the two signers. This extraction was performed using a dialogue analysis support system called Movie Analysis Tool (MAT). The characteristics of MAT are that multiple animated images can be controlled, each image can be tagged, and statistical information for each tag can be obtained in real time. Since tags of multiple images can be displayed on the same window, the relationship between the manual or non-manual signals of both signers can be grasped.

As part of the analysis, the meaning (reason) of a manual or non-manual signal during a dialogue was checked and noted as subsidiary information. The signal, expressed by a signer against a matter, could become to be important in interpretation of sign language expression. So far dialogue samples for 125 key signs have been

completely analyzed. Figure 1 shows an analysis image of MAT.

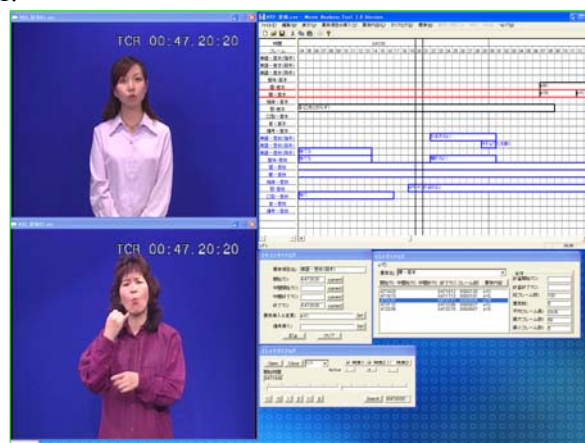


Figure 1: Analysis image of MAT

4. Results

Here, the characteristics are described of dialogue samples which have so far been completely analyzed are described. The classification and the numbers of words and samples, which have been recorded, are shown in Table 1.

Number of key signs	Number of samples	Number of signs	Number of sign types
125	563	4260	964

Table 1: Classification and numbers of signs and examples

The average number of vocabulary words used in one sample by each signer is shown in Table 2.

	Signer A	Signer B
Average number of signs used in one sample	14.5	19.5

Table 2: Each signer's average number of vocabulary words used in one sample

As shown in Table 2, the average number of vocabulary words used for one sample of signer B was larger than that for signer A. This was because the both sentences of signer A (C1 and C3) were questions, and a question usually consists of a short sentence. In the dialogue, signer B gave an answer (C2) against a question of signer A (C1), and signer A gave a question (C3) against C2. The frequency of hand usage (dominant hand, non-dominant hand, and both hands) used in one sample by each signer was investigated. The result is shown in Table 3.

	Dominant hand	Non-dominant hand	Both hands
Signer A	8.8	0.7	5
Signer B	10.4	0.9	8.3

Table3: Each signer's frequency of hands

When the dominant hand pointed at “that man,” third person finger-pointing (pT3), while the non-dominant hand simultaneously pointed at the man (oTOKO), this case was not classified into a group of both hands, and a vocabulary word was assigned to each hand.

sIGNDEX: pT3[oTOKO]

Most of the non-dominant hands were remainders of previous sign language. Non-dominant hand expressed alone was finger-pointing (pT) such as I, first person finger-pointing (pT1), or you, second person finger-pointing (pT2). Regarding this result, the signers were interviewed. They answered that they did not intentionally use their non-dominant hands. They unconsciously used their non-dominant hands under the influence of the final hand and finger positions of the previous sign language.

5. Collocation in sign language

Collocation is generated temporally or simultaneously in the same sentence. In linguistics, temporal collocation is called a syntagmatic relation and simultaneous collocation is called a paradigmatic relation. Here, the former is called temporal collocation and the latter is called simultaneous collocation following the literature. In this study, several collocations were observed.

5.1 Temporal collocation in pT

In JSL, pT is often observed. The frequency of pT in dialogue samples analyzed in this study is shown in table 4.

Sign	Frequency	Ratio (%)
pT2 (you)	465	12.51
pT1 (I)	345	9.28
pT3	201	5.41

Table 4: Frequency of pT

In JSL, it is said that a signer often finger-points to him/herself or a third person at the beginning of a sentence, and finger-points again at the end of a sentence. It is uncertain whether the second pT is always performed at the end of a sentence. However, since pT is often performed together with non-manual signals, such as nodding and winking, pT is considered to be a marker at the end of a sentence. Then, we noticed a repeated expression of pT. Among 563 samples against 125 key vocabulary words, pT was repeated in 129 samples (approximately 23%). The distribution of repeated pT is shown in Table 5.

	pT1	pT2	pT3
Signer A	17	50	15
Signer B	30	6	11
Total	47	56	25
Distribution	36.4%	43.4%	20.2%

Table 5: Distribution of repeated pT

In 96 samples of JSL as described in sIGNDEX V. 2, pT

was repeated in 20 samples (approximately 20.8%)[7]. Among the 20 samples, pT1 was observed in 17 samples, and pT2, pT3, or specific third person finger-pointing (pTR) were observed in 1 sample each. Since this analysis was not performed for dialogue samples, signers mainly finger-pointed themselves (pT1).

On the other hand, since dialogue samples were analyzed in the current study, the distribution of pT2 was the highest. However, the distribution significantly differed according to the signer; the distribution of pT2 was highest in signer A but the distribution of pT1 was the highest in signer B. The reason was considered to be that signer A finger-pointed signer B, who was a second person, in order to give a question (C1) and to express the completion of C1 to signer B. On the other hand, since signer B gave an answer (C2) regarding herself, pT1 was frequently performed. Further study is required to investigate the function of the first pT in temporal collocation.

5.2 Simultaneous collocation in pT

In spoken language, the meaning of a sentence can be changed by adding prosodic factors to the sentence. When intonation at the end of a sentence is changed, the grammatical meaning of the sentence may change. However, intonational changes cannot be expressed by written language.

In sign language, multiple words can be simultaneously expressed using both hands. Moreover, various meanings can be also expressed by the directions of body, face, and line-of-sight. Therefore, complicated content can be simultaneously expressed by adding non-manual signals to manual signals.

When both signals are simultaneously expressed, functions similar to prosodic factors in spoken language become effective. Moreover, by changes in speed, magnitude, appearance, iteration, and shape, many adverbial functions become effective. Thus, collocation in sign language is more complicated than that in spoken language.

We noticed the end of a sentence, and tried to determine the factors shown at the end of a sentence. In 563 samples analyzed in this study, pT was confirmed to exist at the end of a sentence in 197 samples (35%). The details are shown in Table 6.

	Question sentence	Answer sentence	Total
Signer A	82	28	110
Signer B	4	83	83
Total	86	111	193

Table 6: Frequency of pT at the end of a sentence

Signer A's pT were mostly observed in question sentences to signer B, and pT3 and pTR was occasionally performed. On the other hand, signer B's pT were mostly observed in answer sentences, in which the signer defined herself (pT1), and pT3 or pTR was performed only in one sample. Since the frequency of pT shown at the end of a sentence

was high, pT was suggested to be one of sentence end factors. Moreover, since dialogue data were used in this study, it was clarified that signers often used pT to confirm each other.

6. Synchronization phenomenon and collocation between signers

In order to synchronize the same manual signals, signers often confirm the contents of the conversation with each other. In sign language, they often exchange an agreement response {onaji (oNAJI)}. When confirming a size, signers often express the size mutually and simultaneously. For example, signer A confirms the size by a manual signal, which is shown by signer B, who also sometimes corrects it.

Many phenomena were observed in which signers' non-manual signals were synchronized. Among the non-manual signals, hDN was the most frequently observed signal.

Many phenomena were observed in which signers' non-manual signals were synchronized. Among the non-manual signals, hDN was the most frequently observed signal.

In order to express {katadukeru (kATADUKERU)} in sign language, hDN is indispensable. Speaker A understood speaker B's intention from the flow of dialogue and predicted that {katadukeru (kATADUKERU)} would come next. Therefore, their hDN rhythms were perfectly synchronized. In the dialogue corpora analyzed in this study, synchronization of non-manual signals, such as eyebrow motions, was also often observed in addition to hDN. Synchronized non-manual signals include emotional displays resulting from the other speaker's intention. However, collocation between manual and non-manual signals, which is indispensable in sign language expression, is considered to be strongly related to the synchronization phenomenon. Whether this synchronization phenomenon due to collocation between manual and non-manual signals occurs only in a dialogue between native signers is uncertain. Therefore, further investigation is required.

7. Conclusion

This study analyzed dialogue images of the sample corpora. Consequently, collocation and synchronization phenomenon were observed in the dialogues. By using a dialogue analysis support tool, correlation between manual and non-manual signals, and the relationship between two signers, could be analyzed, and characteristic phenomena in sign language could be investigated.

Collocation between manual and non-manual signals, which is obtained by native signers, was confirmed to play an important role in defining the meaning of sign language expression. Such information is indispensable in order for normal listeners to learn sign language. Therefore, further hearing investigation is

required. Moreover, collocation in sign language was suggested to influence the synchronization phenomenon between signers.

In the future, the grammatical functions and cognitive mechanism will be analyzed in addition to the further analyses of sample corpora.

Kazuyuki KANDA, et al. (2001). Notation System and Statistical Analysis of NMS in JSL, *Proceedings of HCI*, Vol.1, pp.25--27.

D. Mutoh and Y. Nagashima, (2001). Conversation analysis support system for sign video images, *Transactions of Human Interface Society*, Vol.3, No.3, pp. 153-162(in Japanese).

Prillwitz, S., Leven, R., Zienert, H., Hamke, T., Henning, J.(1989). HamNoSys Version 2.0;Hamburg Notation System for Sign Languages; *An Introductory Guide, volume 5 of International Studies on Sign Language and Communication of the Deaf*, Signum Press.

Kanda, K., Ichikawa, A., Nagashima, Y., Kato, Y., Terauchi, M.(2000). Signdex V.1 & V.2 The Computerized Dictionary of Japanese Sign Language; *Proceedings of TISLR(7th International Conference on Theoretical Issues in Sign Language Research)*, p.195.