From Text to Sign Language: Exploiting the Spatial and Motioning Dimension

Ji-Won Choi
Computer Science Division,
KAIST
373-1 Guseong-dong, Yuseong-gu,
Daejeon 305-701, South Korea
identity@nlp.kaist.ac.kr

Hee-Jin Lee
Computer Science Division,
KAIST
373-1 Guseong-dong, Yuseong-gu,
Daejeon 305-701, South Korea
heejin@nlp.kaist.ac.kr

Jong C. Park
Computer Science Division,
KAIST
373-1 Guseong-dong, Yuseong-gu,
Daejeon 305-701, South Korea
park@nlp.kaist.ac.kr

Abstract

In this paper, we address the problem of automatically converting information in the Korean language to one in a sign language as used in Korea. First, we discuss the differences between sign language and natural language, and in particular between the sign language in Korea and the Korean language. Then, we focus on issues that are relevant to the process of converting expressions in Korean into their counterparts in the sign language, including: 1) making explicit elided subjects of expressions in Korean, 2) omitting some expressions in Korean, and 3) reordering some expressions. We argue that it is important to utilize the spatial and motioning dimensionality of a sign language in order to minimize information loss and distortion. We also argue that the right decision to omit, or to merge some expressions in Korean plays a key role in exploiting this dimensionality. Finally, we present a system that converts sentences in Korean into corresponding animations in the sign language as proof of evidence for our claim.

1. Introduction

The abundance of information in the present day, mostly in the form of written and spoken languages, has inevitably deepened the information divide among the people with different backgrounds. In this paper, we address the problems of automatically converting information in natural language to one in sign language, so as to lessen the information divide that people with different levels of hearing difficulty experience. For concreteness, we study the differences between the Korean language and the sign language as used in Korea.

Earlier work in the field has mostly defined one-to-one relations between unit expressions in spoken and sign languages, along with simple translation rules. However, they suffer from the problem of not really reflecting the differences in the information space that the two types of language have. It is thus not surprising that an expression in spoken language becomes either unnecessarily verbose or severely restricted in its counterpart in sign language. In this paper, we argue that it is important to utilize the spatial and motioning dimensionality of a sign language to minimize information loss and distortion. For this reason, we also present a system that converts sentences in Korean into corresponding animations in the sign language, as proof of evidence for our claim.

The rest of this paper is organized as follows. Section 2 describes other approaches to converting a natural language sentence into the corresponding sign language expression. Section 3 discusses the

differences between sign language and natural language, Section 4 explores issues that arise when natural language sentences are converted into sign language expressions, using characteristics of the sign language as identified in Section 3. Then, Section 5 presents an automatic translation system.

2. Related Work

In this section, we examine three other approaches to automatic translation from natural language to sign language.

(Greive-Smith, 1999) takes weather forecasting information as a natural language sentence, and tags each word according to its meaning, and analyzes the sentence with a recursive-descent parser customized to the domain, to fill out a preformatted table. The system then applies each word to translation rules, and rephrases the sentence. However, this approach works well only for specific domains where the number of used words is limited and the structure of sentences is standardized.

(Kwon, Woo, and Min, 2000) analyze an input sentence with a morpheme-level analyzer, and applies translation rules to rephrase the original sentence, producing a sign language expression as a result. The system then searches for a database for a GIF movie clip that corresponds to each sign language word. However, this work can not deal with directional or morphological variations in a sign language, as it just searches for a database for each word in the sentence.

(Lee, Choi, 2002) propose a method that converts sign language sentences that conform to the Korean natural language grammar to those that conform to the grammar that is applicable solely to the sign language. However, this work makes minimal changes in the Korean natural language grammar only to the extent that eliminates postpositions, and the result is often too much bound by the source sentence.

To overcome the limitation as observed from the earlier work, we perform the data analysis not on the word level but on the sentence level, and apply rules that exploit the differences between the Korean language and sign language.

3. Sign Language vs. Natural Language

Sign language is not just the medium for substituting other natural languages, but is also uniquely specialized for the representation of thoughts. In this section, we first discuss the differences between sign language and natural language, and then those between the sign language in Korea and the Korean language.

3.1. Sign languages as distinct from other natural languages

Sign languages employ human hands and fingers for various movements, and use the human body and its surrounding space as an active environment, forming a vision-movement system (Davis and Silverman, 1978), which can be contrasted from the audio-vocal system of spoken languages. The basic units of information for sign language are hand-shape, hand-position, hand-motion and hand-direction (Seok, 1989), as opposed to phonemes in spoken languages. While phonemes are arranged serially to form a sentence or a word, these basic units may be employed simultaneously to form an expression compound. The fact that an expression compound is constructed in this manner in sign language has the effect of reducing the dimensionality of interactions among the basic units, thereby restricting the number of available words, compared to the nearly infinite possibility of the serial combinations of phonemes. It also gives sign language the expressive power to convey several pieces of information at the same time. For example, it is possible for sign language users to refer to the subject, the verb, and the object, thus a full usually sentence in other languages, with a single hand movement, as noted also by (Peng, 1975).

3.2. Sign languages in Korea as distinct from the Korean language

Korean is an agglutinative language, whose words and grammatical relations are formed and represented by joining morphemes together. Sign languages in Korea can be classified as isolating, along with Chinese and Vietnamese, in which 'morphemes' are considered syntactically full-fledged expressions and their variations seldom occur (Eom, 1996). Grammatical relations can be expressed by changing the order of morphemes or words. While the Chinese language is the representative case of an isolating language, sign languages show a freer degree of expressions than Chinese, since they utilize syntactic markers simultaneously as well as serially.

4. Converting Expressions

We have examined in the previous section the characteristics of a sign language as a distinctive and independent language apart from natural language. The unique characteristics of a sign language suggest that a translation process beyond a simple word-to-word match is required. In this section, we examine some issues that we need to address in the process of converting expressions in Korean into their counterparts in the sign language.

4.1. Elided expressions to explicit expressions

Elided information in a Korean sentence may have to be explicitly realized in a sign language sentence. Relevant cases include realizing functional phonemes, restoring elided subjects and objects, and rephrasing unknown words.

   Distinct endings of a word or postpositions often indicate syntactic information in the Korean language. However, expressions in sign language represent their syntactic information by varying word order rather than by varying words themselves, which is the very reason why they are classified as an isolating language. Therefore, it is important to find and realize implicit syntactic information in Korean to transfer the information to sign language more effectively.

   Tense information in Korean is usually represented by distinct endings of a verb. Therefore, sign language must utilize additional units of information to represent tense. For example, for a past-tense sentence as shown in Example 1 (a), its corresponding sign language expression is to explicitly represent a hand-motion meaning 'end' after the hand-motion meaning 'to come', as shown in Example 1 (b).

   Example 1 (a) 영희가 왔다. (Yenghuy came)
           (b) [영희 (Yenghuy) + 오다 (to come) + 끝 (end)]

   Certain postpositions in Korean also need to find their explicit counterparts in the sign language. The possessive marker '~의' and the coordination marker '~와' are converted directly to their corresponding hand-motions, as sign language has explicit words that mark possessiveness and coordination. Realizing the adverbial marker in the sign language is more complicated, however, as there are various kinds of adverbial marker in Korean. For example, '~로서' is a postposition that functions as a syntactic marker representing the position, status or qualification. Sign language can explicitly use its word that means 'position' to convey the same meaning. Likewise, '~로써' can be represented with a sign language word that means 'to use', as it represents the syntactic sense of means or tool, and '~라고' with a word that means 'quotation mark', as it represents citation. There are other adverbial markers that can be represented in sign language by a combination of sign language words. They are explained further in Section 4.2.

   Both Korean and sign language in Korea often elide subjects or objects in their sentences. However, to make sign language expressions clearer, it is sometimes necessary to restore these elided components of Korean sentences in corresponding components in sign language expressions. For example, if we do not express the sense of 'I' of the sign language expression explicitly in

Example 2 (b), the sign language word 'child' may incorrectly mean 'young', as the hand-motion for the two meanings are the same, inducing a different interpretation from the source Korean sentence in Example 2 (a).

Example 2 (a) 어렸을 때 어른들께서 장래 희망이 무엇이냐고 물었다.
(When I was a child, elders asked about my hope for the future)
(b) [나 (I) + 아이 (child) + 때 (time) + 어른들 (elders) + 묻다 (ask about) + 앞으로 (future) + 희망 (hope) + 무엇 (what)]

The sign language vocabulary is much smaller than that of the Korean language. Therefore, it is desirable to rephrase unknown words in an explanatory manner or to express them with the use of fingers for exact spelling. In this paper, we choose to use the finger expression for exact spelling, as finding an analogical expression for unknown words is beyond the scope of the present work.

4.2. Explicit expressions to implicit expressions

There are cases where some explicit words in Korean may have to be converted implicitly, either as a mere composition of units of sign language words or simply omitted, and can still have their meanings conveyed. This is due to the simultaneity of syntactic rules in the sign language, which are considered to deliver meanings of various verb endings, auxiliary verbs and transmutations of word functions (Hwang, 1996). Such rules are realized either by altering spatial properties such as the direction of movements or arrangements, or by changing hand motions such as the way of ending the motion, or the speed of the motion. We examine these cases in further detail.

Formal morphemes, or morphemes that deliver syntactic information, can be omitted entirely or integrated into other components. The Korean language indicates case information by adding an appropriate postposition to the relevant word. '~이' and '~가' are the representative postpositions for nominative case, and '~을' and '~를' for objective case. Adverbial case information can also be represented with postpositions, such as '~에' or '~로'. These postpositions are usually omitted in sign language translations, because the sign language represents syntactic relations by the fixed order of words. Example 3 (a) shows the source Korean sentence, and 3 (b) its possible sign language translation with omitted postpositions.

Example 3. (a) 나는 학교에 간다 (I go to school)
(b) [나 (I) + 학교 (School) + 간다 (to go)]

Postpositions that transform noun, pronoun and numeral to either an epithet or adverb may add meanings other than syntactic information to words, but they can also be omitted entirely if they are included visually through other words. Example 4 (a) shows the adverbial postposition '~로부터', which means 'from', together with the verb '오다', which means 'to come'. As the sign language expression for '오다' is to use the hand motion to signify the directionality with the original and ending positions, if we start the hand motion for '오다' from the position where we represented '어머니', which means 'mother', the additional expression for '~로부터' becomes unnecessary. The resulting sign language translation is represented in Example 4 (b).

Example 4. (a) 어머니로부터 온 편지 (A letter from mother)
(b) [어머니 (mother) + 오다 (to come) + 편지 (letter)]

Nouns in the sign language often convey verbal semantics together with their nominal semantics. Nouns with their typical verbs show this property whose verbs can simply be omitted. The sign language expression for 'meal' implies the sense of 'to eat', therefore the appropriate translation in

Example 5 (b) for Example 5 (a) does not require the additional explicit expression for 'to eat'. Similarly, the sign language expression for 'piano' implies the sense of 'to play', and the expression for 'angry' that of 'to get'.

Example 5 (a) 밥을 먹었다. ((I) ate meal)
       (b) [밥 (meal) + 끝 (end)]

Furthermore, by changing the way of expressing nouns, we can also signify their verbal, adverbial, or adjectival characteristics. For instance, using a hand motion for the bus with a backward movement is enough to express the idea that '(the) bus moves backward' in Example 6, and undoing the link made with the thumbs and the index fingers of both hands is enough to express the idea that 'there is no relation' in Example 7, as the link itself represents 'relation'.

Example 6 (a) 버스가 뒤로 가다. ((the) bus moves backward)
   (b) [버스 (bus, with backward movement)]

Example 7 (a) 관계 없다. ((there is) no relation)
       (b) [관계 (relation, undoing the link)]

Verbs or adjectives have a similar degree of freedom for their adverbial characteristics. The function of an adverb is to modify the expressions of verbs and adjectives. For example 'walk fast' can be expressed by moving hands for 'walk' faster than that for 'walk' alone, and 'walk slowly' can be expressed by moving hands more slowly than that for 'walk' alone. More complex sentences as in Example 8 (a) can be expressed by moving hands slowly for the first '걷다', and moving hands more slowly for the second '걷다', where '걷다' means 'to walk'.

Example 8 (a) 철수는 느리게 걷고, 순이는 더 느리게 걷는다.
       (Chelswu walks slowly, and Swuni walks more slowly)

The directions of a hand or the hand movement that represents verbs can be adjusted to convey meanings of a personal pronoun or passiveness and causativeness.

Sign language as a visual language has its own scheme to perform similarly to the demonstrative pronoun schemes of natural languages, that is, to take note of the position where a certain noun is expressed and to point at that position to refer to the noun. The scheme for a personal pronouns is somewhat different, as there are reserved positions in space to signify specific personal pronouns. Reserved positions and corresponding personal pronouns are shown in Figure 1.
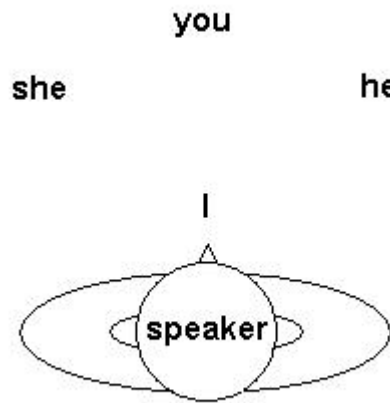
Figure 1: Reserved positions for specific personal pronouns

Thus, pointing at these positions while expressing verbs can effectively signal the respective personal pronouns. For instance, performing a hand action that means 'to see' from the right to the left represents the sentence 'He sees her', as the right-hand side represents a male third person and the left-hand side a female third person. The more complex case in Example 9 (a) can be represented in sign language by moving the right hand shaped for 'to see' to the right-hand side of the speaker, and the left hand also shaped for 'to see' to the left-hand side of the speaker.

Example 9 (a) 나는 그와 그녀를 보았다. (I saw him and her)

Passiveness and causativeness can be expressed similarly. The Korean language uses verbs such as '시키다' or '보내다' to represent causativeness and '되다', '받다' or '당하다' to represent passiveness. However, in sign language, information of this kind can be conveyed by performing a verb-representing action from the position that represents the causative subject to the position for the passive subject. For example, the sentence in Example 10 (a) can be converted into sign language by first representing 'Minswu' and '(the) teacher', and then performing 'to hate' from the position '(the) teacher' was represented to the position 'Minswu' was represented.

Example 10 (a) 민수는 선생님으로부터 미움을 받았어요. (Minswu was hated by (the) teacher)

Similarly, the sentence in Example 11 (a) can be converted by performing '~를 하라고 하셨다', which means 'to tell to do', toward the speaker, whereas the conventional direction for the hand motion is to the opposite of the speaker.

Example 11 (a) 어머니께서 나에게 빨래를 하라고 하셨다. (Mother told me to do laundry)

This use of directionality for representing causativeness and passiveness may also resolve possible confusion caused by subject-object reversal. Examples 12 (a) and (b) both have the preceding word '아버지', meaning 'father', and the following word '아들', meaning 'son'. However, the meanings they convey are different, as '아버지' in (b) is a fronted object whereas that of (a) is the subject. The directionality of sign language can be utilized to discern the difference sense of this kind. That is, we can first represent '아버지' and '아들', and then perform '사랑하다', meaning 'to

love', from the position of '아버지' to the position of '아들' to represent the meaning of (a), and use the opposite directionality to represent the meaning of (b).

Example 12 (a) 아버지가 아들을 사랑한다. (Father loves son)
(b) 아들이 아버지를 사랑한다. (Father, son loves)

4.3. Reordering expressions

We may need to reorder words in the expression in Korean to that in the sign language. If the subject has too long an epithet, it is likely that the listener forgets information about the subject during the long delivery of the epithet. In Example 13 (a), the appropriate sign language translation for avoiding confusion can be the representation for '분주하다', meaning 'busy', right after the subject '딸', which means 'daughter'. Example 13 (b) represents the sign language translation.

Example 13 (a) 딸은 결혼식이 다음주로 다가오자 분주해졌습니다.
(Daughter as the wedding came to next week got busy)
(b) [딸 (daughter) + 분주하다 (busy) + 결혼식 (wedding) + 다음주 (next week) + 다가오다 (come to)]

Reordering may also be necessary if points of time in a sentence are mixed. In Example 14, if we convert the source Korean sentence in (a) into one as in (b), a misleading interpretation such as 'where will I be if I transfer from the subway line number three?' may result in. To avoid such confusion, we may exchange the order of '바꾸다', meaning 'transfer', and '3호', meaning 'line number three', according to the time line of the involved events. Example 13 (c) represents the intended representation.

Example 14 (a) 3호선으로 바꿔 타는 곳은 어디입니까?
(Where can I transfer to the line number three?)
(b) [지하철 (subway) + 3 (three) + 호 (line number) + 바꾸다 (to transfer) + 어디 (where) + 입니까? (question)]
(c) [지하철 (subway) + 바꾸다 (to transfer) + 3 (three) + 호 (line number) + 어디 (where) + 입니까? (question)]
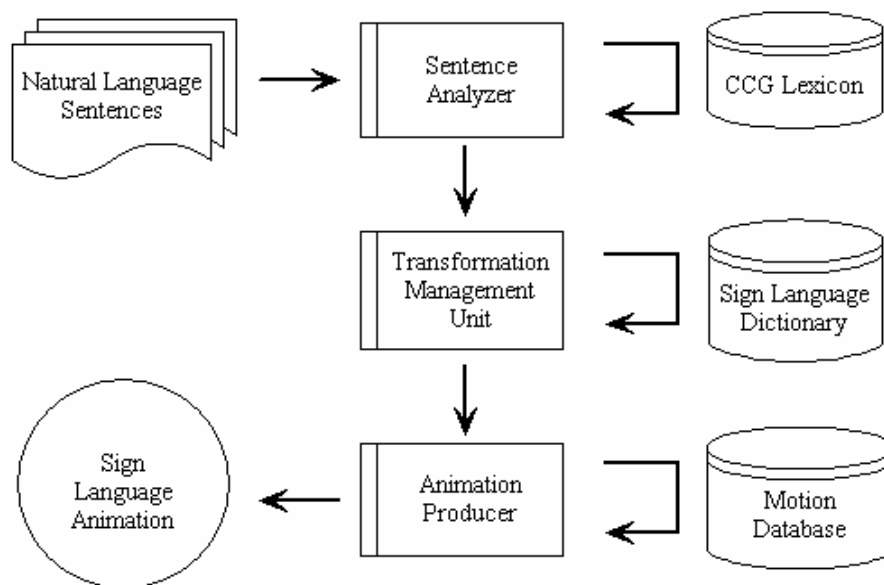
We may also use reordering to translate the sense of emphasis in the source sentence. Example 15 shows an example where two sentences put emphasis on the reason for the fighting.

Example 15 (a) 왜 학생들이 싸움을 합니까? (Why are (the) students fighting?)
(b) [학생 (student) + 싸움 (to fight) + 왜? (why?)]

However, it is difficult to identify the point of emphasis only from the text information without intonation or gesture. We leave this as future work.

5. The Translation System

We have implemented a system that converts sentences in Korean into sign language animations. We restrict the scope of Korean sentences to be practical and concrete, avoiding the use of words beyond everyday usage, and the use of sentences with highly complex structures. The translation system constructed with this restriction is enough to aid people with hearing difficulties in their everyday lives. The system is composed of three parts: the sentence analyzer, the transformation management unit, and the animation producer, as shown in Figure 2.

Natural Language Sentences → Sentence Analyzer ⇄ CCG Lexicon

Transformation Management Unit ⇄ Sign Language Dictionary

Sign Language Animation ← Animation Producer ⇄ Motion Database

Figure 2: The system structure

The sentence analyzer uses a combinatory categorial grammar (CCG) to obtain syntactic and semantic information from the input sentence in text format. CCGs are an extended version of categorial grammars (CGs) with a limited set of combinators, such as function composition, function substitution, and type raising (Steedman, 2000), where CGs assign categories to each lexical item, for instance each word in a natural language such as Korean or English. CCGs can model difficult linguistic phenomena, such as coordination, long distance extraction and topicalization, without further stipulation. CCGs can also analyze semantic or discourse information by adding information of this kind to their lexicons. We use a CKY parser in the Prolog programming language for the CCG framework.

Information obtained by the sentence analyzer is then passed to the transformation management unit, which applies rules we have discussed in Section 4, consulting the sign language dictionary. Rules include positioning the objects in the space, determining directionalities of words, reordering words, dealing with phrases composed of Chinese characters, omitting verbs that come with typical noun, omitting verbs and adverbs with a movement property, and dealing with adverbs representing speed or power. Phrases composed of Chinese characters are considered as unknown words. Thus we first search the dictionary for the substitution into corresponding explanatory phrases. If there is no appropriate result, we use fingers for exact spelling. The transformation management unit produces as its result an 'animation script', and hands it over to the animation producer.

The animation producer takes the animation script to give rise to sign language animation. It searches for an appropriate sign language animation clip for each sign language unit expression from a motion database, and then plays the clips sequentially. Animation producer is in the Visual Basic programming language.

The system can produce an appropriate animation for each of the source sentences with subtle variations, such as reordering of words, or differentiation of similar verbs and adjectives. For example, the produced animaition for Example 16 (a) is: representing 'teacher' in the right-hand lower side of the speaker, 'I' in front of the speaker, and then 'cleaning', followed by 'to command' with directionality from the right lower side of the speaker to the front of the speaker, and finally 'end' to signal the past tense. Example 16 (b) shows a new sentence produced by reordering words of the sentence in (a). The animation produced by the system for this sentence is similar to that of the sentence in (a), but with a reversed order of the representation of 'I' and that of 'cleaning', thereby conveying the sense of the reordered sequence of wordss. Example 16 (c) shows another

variation of the source sentence in (a), where the predicate has changed from '시키셨다', meaning 'to command' with the past tense, to '하라고 하셨다', meaning 'to tell to do' with the past tense. The produced animation is the same as that of the sentence in (a), except that it has the additional representation for '하다' meaning 'to do' at the end of the sign language expression. This additional representation further clarifies the sense as shown in (c).

Example 16 (a) 선생님께서 나에게 청소를 시키셨다. (Teacher commanded cleaning to me)
(b) 선생님께서 청소를 나에게 시키셨다. (Teacher commanded me cleaning)
(c) 선생님께서 나에게 청소를 하라고 하셨다. (Teacher told me to do cleaning)

The system obtains these differentiations by working on the sentence level, rather than on the word level as in most of the previous studies.

6. Conclusion

We have analyzed properties of a sign language as an independent language different from other spoken languages. As an isolating language, sign language employs hand-shape, hand-position, hand-motion and hand-direction simultaneously to form an expression compound. The simultaneity utilizes the spatial and motioning dimension, inducing effective conveyance of information.

We proposed translation rules that signify the characteristics of sign language with the effect of minimizing information loss, mainly by focusing on the dimensionalities. The rules include making explicit expressions implicit, implicit expressions explicit, and reordering certain expressions.

We also described an automatic translation system that uses space and motioning as the key dimensionality. The system is capable of detecting subtle differences between sentences with similar semantics.

However, as the input sentences of the system are in text form, it is difficult to find where they are stressed without additional forms of information such as annotation or gesture. And the system represents unknown words with exact spelling using fingers, rather than with explanatory phrases. Enhancing the system with speech recognition to process non-textual information and implementing an inference system for unknown words are among the future work to do.

7. Acknowledgements

8. References

Eom, M.-S. 1996. *Syntactic analysis of Korean sign languages*. Master's thesis, Daegu University, Korea.

Davis, H. Silverman, R. 1978. *Hearing and Deafness*. N.Y. Holt Rinehart and Winston.

Grieve-Smith, A. B. 1999. English to American sign language machine translation of weather reports. *High Desert Linguistics Conference*. pp.23-30.

Hwang, D.-S. 1996. The length development of the sign language sentences by Korean hearing-impaired students, *Research on Special Education*, the 3rd series.

Kwon, K.-H. Woo, Y.-S. Min, H.-K. 2000. Design and implementation of a Korea text to sign language translation system. *The Transactions of the Korea Information Processing Society*. pp.756-765.

Lee, Y.-W. Choi, C.-S. 2002. Methodology for translating Korean sentence to sign language sentence for auto-produced sign language animation. *HCI conference*.

Peng, F.C. 1975. Sign language and language acquisition in man map. *AAAs selected symposium*.

Seok, D.-I. 1989. *Linguistic analysis of Korean sign languages*. Ph.D. thesis, Daegu University, Korea.

Steedman, M. 2000. *The Syntactic Process*, The MIT Press.