Pragmatic Factors Influencing Word Choice: An Examination Based on Corpus and the IR Method

*Yu WANG¹, Wenzhen SHI² and Tetsuhisa ODA²
1)Nagoya Univ. 2)Aichi Institute of Technology
Address: 〒464-0036 Senri Heights 201, Motoyama Chou, Chikusa Ku, Aichi Prefecture
Email: youwan@hotmail.com

Abstract—This study argues that the pragmatic factors which motivate the choice of the four reflexive pronouns jibun, jisin, mizukara, and jiko. Since the differences of these words are subtle and the speakers’ attitudes are complex, we introduce Improved Rating (IR) method which is developed from Oda’s Fuzzy-set Concurrent Rating (FCR) method. The analysis consists of two parts. First, a corpus-based investigation suggests that the choice of these words may be motivated by a range of factors including attitudes, conversation style, and contents, which can all be related to the “psychological distance”. Second, in order to reconfirm the psychological distance hypothesis, we executed a questionnaire using the IR method. As a result, the “psychological distance” turns out to be an effective factor influencing the speakers’ attitudes and thus word choice.

I. VAGUENESS

Vagueness seems to be a fundamental and pervasive feature of natural language. Although the prototypical vague predicates are often regarded to be gradable adjectives such as old or tall, other lexical categories are equally susceptible to vagueness. In Japanese even reflexive pronouns can be vague. Reflexive pronoun is a pronoun with a reflexive relationship with its self-identical antecedent. Unlike English, there are several words which can be used as reflexive pronouns in Japanese such as jibun, jisin, mizukara, etc. If contexts allow, they can substitute each other without changing the reference, i.e. truth value.

(1) 〔自分／自ら／自身／自己〕を見つけた。
‘I found myself.’

What matters for word choice is which word is used more appropriately. Channell [1] argues that “vagueness in language is neither all ‘bad’ nor all ‘good’. What matters is that vague language is used appropriately”. It goes without saying that to be “appropriate” concerns interactive relation between a word and its context.

II. PSYCHOLOGICAL DISTANCE

We suppose that the differences between the four reflexive pronouns can be accounted for by the psychological distance from the speaker to the referent. We argue that jibun is the closest one, while jibun, jisin, mizukara are comparatively far. It is quiet difficult to discriminate the implied distance between the speaker and referent. First, reflexive pronouns do not have a fixed representation. Their reference totally depends on the contexts. The meaning of reflexive pronouns is so vague that we might construe that it is the contexts rather than the pronouns that bring out the implication of psychological distance. Considering this fact, we think it is reasonable and feasible to describe the implication of reflexives by analyzing the contexts.

Second, psychological distance is correlated closely with sense-experience, and is necessarily vague. It is impossible to be measured out exactly. Still there are other problems about the notion of discriminability: the distance is discriminated by whom? How about the man who has never used jisin or jiko? Perhaps the question above is too accidental to mention, but it is essential for a general notion of psychological distance as it is ultimately a type of individual and subjective case.

Considering these problems, we will take two methods to prove the hypothesis of psychological distance. One is a
corpus-based method, and the other is the IR questionnaire method. The former provides us with a considerable amount of data, and enables us to describe the features of the co-occurring contexts objectively. The latter is to allow us to know the degree of confidence and the psychological state of the speakers. Because the result of the latter is pertinent to psychology and the investigation differs individually, we regard it as a subjective method.

III. A CORPUS-BASED INVESTIGATION

Four kinds of genres are selected in order to find the features of different contexts. They are:

- News from Asahi newspaper
- Novels from Aozora Bunko
- Formal conversations from records of Diet
- Informal conversations from TV series

Each genre contains around 20,000 words.

Table 1. Frequency of the reflexive pronouns

<table>
<thead>
<tr>
<th>Pronoun</th>
<th>News</th>
<th>Novels</th>
<th>Formal conversations</th>
<th>Informal conversations</th>
</tr>
</thead>
<tbody>
<tr>
<td>jibun</td>
<td>57</td>
<td>71</td>
<td>71</td>
<td>66</td>
</tr>
<tr>
<td>jisin</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>mizukara</td>
<td>21</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>jiko</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>0</td>
</tr>
</tbody>
</table>

Among the four reflexive pronouns, jibun is most often used through all the genres, whereas the frequencies of jisin, mizukara and jiko are very low. They are used only in certain occasions. The frequencies of jisin, mizukara and jiko are decreasing along the following cline.

News 〉 Novels, Formal conversations 〉 Informal conversations

Table 2. Pragmatic Factors and Psychological Distance

<table>
<thead>
<tr>
<th>Anticipants</th>
<th>Style</th>
<th>Contents</th>
<th>PD</th>
</tr>
</thead>
<tbody>
<tr>
<td>one to one, acquaintance</td>
<td>Inter–active</td>
<td>many oral expressions</td>
<td>Near</td>
</tr>
<tr>
<td>one to multiple, not necessarily knowing each other</td>
<td>Inter–active</td>
<td>less oral expressions</td>
<td></td>
</tr>
<tr>
<td>one to multiple, not knowing each other</td>
<td>one–way</td>
<td>oral expressions possible</td>
<td></td>
</tr>
<tr>
<td>one to multiple, not knowing each other</td>
<td>one–way</td>
<td>no oral expressions</td>
<td>Far</td>
</tr>
</tbody>
</table>

These three factors can be related to our hypothesis. The one-to-one interactive communication in which oral expressions are often used shows the most ideal context-model for speakers to be relaxed and not need to worry about the distance he/she should take. On the other hand, in the one-to-multiple one-way communication, in which oral expressions are seldom used, speaker should be careful about his/her expressions, and pay attention to take the appropriate psychological distance to the referent.

To sum up, through a corpus-based investigation, jibun is found to be the most widely and the most often used one, while the other three are of low frequency and appear mainly in formal contexts. What varies here is the mental representation evoked by the speaker, i.e. the psychological distance taken by him/her according to the context.

IV. THE IR METHOD

A. General introduction of the IR method

The Improved Rating method was developed from Oda’s Fuzzy-set Concurrent Rating (FCR) method [2] [3]. It incorporates a list of four reasons that have to be selected by the subject when the degree of confidence chosen is not 100%. We apply it to linguistic analysis to investigate the psychological state of subject. Here is an example.

(4) Do you think the reflexive pronoun is used appropriately?

[Context] When someone said to himself:

「自ら が はずがしい。’

‘I am ashamed of myself.’

- Degree of Suitability 0——100%
- Degree of Confidence 0——100%
- Reason:
  1. I agree and disagree at the same time
  2. My answer would be a range of the scale, not only a point.
  3. The information given on the question is not enough for me.
  4. Others.
The four reasons are to provide the details to know why this person is not sure about his/her response, i.e. whether he/she is in a state of contradiction (in case reason 1 is chosen) or of irrelevance (in case reason 2, 3 are chosen). Here the state of irrelevance can also be regarded as unfamiliarity with the expressions in question.

We take degree of suitability as X, degree of confidence as Y, and reason as Z. With these values we can present the result in a two-dimension model which is named Hyper Logic Space (HLS) [4]. The algorithm of transition can be referred to [5], [6].

Here is the definition of the characteristic points, lines and areas in Figure 1.

- 0: “Empty point”, perfectly irrelevant. Its coordinates are (0,0).
- 1: “Contradicting point”, perfectly contradictory. Its coordinates are (1, 1).
- Y=0: Y is the value of confidence. It shows subject is perfectly unconfident. It is a point and it has two positions (0,0) and (1,1).
- V & Y=1: “Numeral truth value space”, (“truth” in classical logic). This line shows the subject is perfectly confident.
- T: “True point”, (“truth” in classical logic). Its coordinates are (1,0)
- F: “False point”, (“false” in classical logic). Its coordinates are (0,1)
- Contradiction Area: Where the observed point is put in case reason 1 is chosen.
- Irrelevance Area: Where the observed point is put in case reason 2 and 3 are chosen.

B. Merits of the IR method

The traditional evaluating method of sentences in linguistics is to use the marks such as “?”, “??”, “?””. Although it is a convenient way, it is too rough to tell the subtle distinction among synonyms, and it fails to provide any information of the psychological state of the speaker. Because the differences among the four reflexive pronouns are so subtle, and the subject’s attitude is to be investigated, we apply the IR method. Our purpose is to describe the speaker’s psychological state, that is, whether he/she is confident with the answer, and whether he/she thinks his/her answer is contradictory or irrelevant (unfamiliar). By doing that, further evidence is expected to be found to understand the difference of the four words.

C. The IR questionnaire

The IR questionnaire is designed to represent different genres. It arranges from informal communication such as daily conversations to rather formal ones as news reports and questions in court. We distributed the questionnaires in April, 2010. The people who took the questionnaire are sixty Japanese native speakers of 20~21 years old. Removing the questionnaires with some unfinished answers, we collected 45 valid questionnaires.

D. Analysis

We set eight contexts. Here we focus on three of them. Context1 is a daily conversation, context 2 news report, while context3 a question in court. First, we present values of suitability of the four reflexive pronouns (see Table 3).

It is interesting to note again that the subjects treat informal communications differently from formal ones. They prefer to jibun in daily conversations, and tend to switch their expressions to jisin, mizukara, jiko in more formal contexts. We can tell that they often want to convey their propositional attitude along with content so that when contexts change, they have to re-select the expressions and choose the most appropriate ones for the contexts.

Second, the degree of confidence and the reasons may be different for each reflexive pronoun even though their values of suitability are similar. Generally speaking, when the degree of confidence is not 100%, the reason for choosing jibun tends to be the first reason listed in (4), i.e. the one of contradiction, while the reason of jisin, mizukara, and jiko tends to be the second and third reasons, i.e. the ones of irrelevance.

For example, values of suitability of jibun in context1 and jiko in context 3 are similar. However, if we look into their values of confidence and the reasons, it’s easy to find that the two situations are actually quite different from each other.
For *jibun* in context 1, many of the observed points are distributed in contradiction area, while for *jiko* in context 3, the observed points are apparently concentrated in irrelevance area. We suppose that the concept of contradiction arises from high frequency and high degree of familiarity of *jibun*, while the concept of irrelevance may be caused by the lack of knowledge.

We argue that the more familiar the expressions are, the more individuation and focus are needed, and in addition such expressions evoke many similar contexts and yield a wide range of contextual and analytical implications that may cause misunderstanding. On the other hand, less familiar expressions imply that a referent can remain in the background and thus the contextual implications are not easy to be evoked. The idea that implication may rely on resemblance relations is indebted to Sperber and Wilson [7]. Sperber and Wilson argue:

> [g]enerally speaking, an utterance can be used to represent any representation which it resembles in content, whether a public representation such as another utterance, or a mental representation such as a thought.

In other words, an utterance can be used to represent any phenomenon which it resembles in some respects. In order to avoid excessive implications and thus the risk of being misunderstood in virtue of resemblance relations, the less familiar reflexive pronouns are chosen. Whether to use more familiar expressions or not is an interactional strategy.

Also, we have to mention that the argument of familiarity matches our hypothesis of psychological distance. Since *jibun* is the most familiar expression, it must be the closest one to the referent. While the degree of familiarity of *jisin, mizukara*, and *jiko* are low, so they must be farther to the referent.

In conclusion, the IR questionnaires reconfirms the point that speakers tend to choose different reflexive pronouns according to different contexts in order to convey their attitude of referent and contexts appropriately. The notion of familiarity presents another aspect of psychological distance and thus provides further evidence for the hypothesis.

### V. CONCLUSION

The differences of four reflexive pronouns we focused on lie in psychological distance. This argument has been proved in two ways. First, a corpus-based investigation shows different distribution of the four pronouns. Through a further comparison of anticipant, style, and content, the three crucial factors which can be the characteristics of contexts, we find these factors can be related to the psychological distance. This relation provides a new proof for the hypothesis of psychological distance.

Through our investigation, we also find that word choice may thus be seen as managing conversational implication. That is to say, word choice may convey the speaker’s attitude such as detachment from or closeness to the referent and it may be understood as conversational strategies to serve social functions such as engendering camaraderie and avoiding excessive implication.

### REFERENCES