

CONCEPTUAL TAXONOMY OF JAPANESE ADJECTIVES FOR UNDERSTANDING
NATURAL LANGUAGE AND PICTURE PATTERNS

Naoyuki OKADA, Aiko MIURA

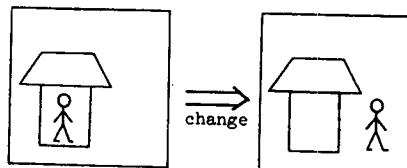
Information Science and Systems Engineering
Oita University, Oita 870-11, Japan

1. Introduction

This paper presents a conceptual taxonomy of Japanese adjectives, succeeding that on Japanese verbs¹. In this taxonomy, natural language is associated with real world things - matter, events, attributes - and mental activities - spiritual and sensual. Adjective concepts are divided into two large classes, simple and non-simple. Simple concepts cannot be reduced into further elementary adjective concepts, whereas non-simple ones can be. Roughly speaking, simple concepts are concrete and can be directly associated with physical and mental attributes, whereas non-simple ones are abstract and indirectly associated with them.

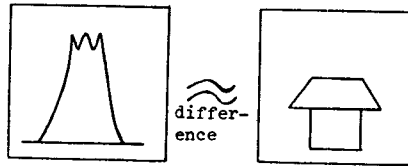
2. Simple Adjective Concepts

Verb concepts were well understood as "change" from state S_0 to state S_1 as shown in Fig. 1.¹ Adjective concepts are considered to be captured as the "difference" between objects O_0 and O_1 . Fig.2 shows how the difference in vertical length between O_0 and O_1 brings about the concept of "high". Notice that surface structures often lack the expression of O_0 like "yama-ga takai (the mountain is high)". Since the meaning of "high" cannot be expressed only by O_1 , deep structures need O_0 as an object for comparison.



State S₀ State S₁
otoko-ga ie-kara deru.
 The man goes out of the house.

Fig. 1 The structure of verb concept



Object O₁ Object O₀
yama-ga ie-yori takai.
 The mountain is higher than the house.

Fig. 2 The structure of adjective concept

2.1 Structural Patterns

Adjective concepts can be regarded as the concept of a relation among constituents which produce difference and can be represented as follows:

$$A_d(c, s, o, o_f, o_t, o_m, o_g, i, l, t, r, \dots) \quad (1)$$

where each symbol in parentheses stands for a constituent. For example, *c* and *s* stand for an object for comparison and a subject, respectively. Some constituents are obligatory in the sense that they are indispensable in the recognition of difference. Seven types of structural patterns are proposed, according to the combination of obligatory constituents.

2.2 Semantic Contents

The semantic contents of A_d in formula (1) are categorized into physical, mental and abstract attribute. We classify each of them in detail because we cannot detect any difference

without such information. In the case of verbs, twenty categories such as displacement, deformation, and spiritual activity were proposed to classify the semantic contents of "changes". We modify them slightly to apply to the classification of semantic contents of adjectives. For example, displacement and deformation are modified to location and shape, respectively.

3. Non-Simple Adjective Concepts

Non-simple adjective concepts are not directly associated with physical or mental attributes; therefore we emphasize an analysis of how they are composed of simple ones.

3.1 Complex Concepts-A

For example, there is a complex adjective "hoso-nagai (thin-long)" which is registered as one word in ordinary Japanese dictionaries. The connecting relation between two adjective concepts is a logical product in this case. The complex concept of A is defined as follows: if two adjective concepts are connected by a rule and the connected concept is represented by a complex word of those adjectives, it is called a "complex concept of A". Two types of connecting rules are proposed: logical product and syntactic relation.

3.2 Complex Concepts-B

This concept is usually represented by a simple word and is considered to consist of several elementary adjective concepts. No general rule can be found to connect elementary adjective concepts. Semantic contents of the concepts are classified into such categories as linguistic activity and character.

3.3 Derivative Concepts

Some concepts possess a function of deriving a new concept by operating others. They have two morphemic structures, prefix and suffix, which add certain information to operands. For example, suffix "-tarashi" adds the meaning "indeed" to its operand. Such derivative information is categorized into five classes.

4. Classification

There are approximately 530 adjectives considered sufficient for daily Japanese. We actually classified their concepts, numbering about 630, according to types and categories proposed in Sec.2 and 3. They were well classified into those types and categories. In the classification of simple concepts, "standard concepts" were chosen from among similar concepts, amounting to approximately 270. They might be considered "primitives" in the sense that every other adjective concept can be derived from them by some rule.

5. Extraction of Conceptual Features

We extract conceptual features from each A_d and its constituents, noticing the standard concepts with physical attributes obtained in Sec.4. The following analysis shows part of the feature extraction from takai (high)/ hikui (low).
takai(high) / hikui(low)

Polysemy (1) Structural pattern; $A_d(s,c)$, semantic contents; shape

s, c : thing

A_d : The vertical length of s is larger/smaller than that of c.

Example: Tōkyō tawā-ga Kyōto tawā-yori takai.

Tokyo Tower is higher than Kyoto Tower.

Polysemy (2) Structural pattern; $A_d(s,c)$, semantic contents; location

s, c : thing

A_d : The vertical distance of s from a standard plane is larger/smaller than that of c.

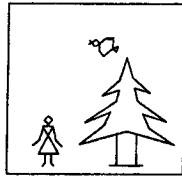
Example: Hikōki-ga Kumo-yori takai.

The airplane is higher than the cloud.

6. Sentence Production from Picture Patterns

System SUPP, which accepts moving picture patterns, interprets the meaning of changes in them, infers about similarity among changes, and describes the results in Japanese and English sentences, was constructed as an application of

the conceptual taxonomy of verbs. SUPP was improved to interpret physical attributes of picture patterns. Fourteen standard concepts were chosen for this improvement. An example of experimental results is shown in Fig.3.



(a) Input pattern

Note. In 1) through 3), input picture "onna" is compared to its model in machine.

- 1) *onna-ga nagai.*
The woman is long.
- 2) *onna-ga takai(1).*
The woman is "high[1]".
- 3) *onna-ga ōki.*
The woman is large.
- 4) *onna-ga ki-yori hikui(1).*
The woman is "lower[1]" than the tree.
- 5) *onna-ga ki-yori hikui(2).*
The woman is lower[2] than the tree.
- 6) *onna-ga ki-yori chīsai.*
The woman is smaller than the tree.
- 7) *onna-ga tori-yori takai(1).*
The woman is "higher[1]" than the bird.
- 8) *onna-ga tori-yori hikui(2).*
The woman is lower[2] than the bird.
- 9) *onna-ga tori-yori ōki.*
The woman is larger than the bird.

(b) Output sentences

Fig. 3 Sentence production from a picture pattern

7. Conclusions

This taxonomy has made clear an outline of the whole of Japanese adjective concepts. In Japanese, the number of adjectives is rather small compared to that of adjective verbs. This taxonomy is expected to give a basis to that of adjective verbs.

Reference

- (1) N. Okada: "Conceptual taxonomy of Japanese verbs for understanding natural language and picture patterns", Proc. COLING 80, Tokyo, 1980.