

ERROR ANALYSIS OF JAPANESE ADJECTIVAL COLLOCATIONS FOR AN ERROR DATABASE

Hongquan Cao, Kikuko Nishina, Tokyo Institute of Technology

Abstract: The aim of this study is to construct an annotated error database based on survey data. Focusing on adjective collocations, a collocation production survey consisting of 70 basic Japanese nouns was administered to 190 learners of Japanese. Error analysis identified 1,449 errors (tokens) in total among the 23,525 collocations (tokens) produced. As a part of the database design, the collocation errors are classified according to eight error categories based on various production strategies. The error classification provides useful insights into the characteristics of learner expressions.

Keyword: error analysis, adjective, collocation, error category

1. INTRODUCTION

Error analysis can be helpful for structuring syllabus design and teaching techniques in ways that minimize the occurrence of errors. Similarly, the incorporation of an error database within a composition support system would be particularly useful for minimizing errors and for error checking. Focusing on adjective collocations, we have previously analyzed collocation errors in a bilingual composition corpus of Japanese and other languages (Cao & Nishina, 2006a). However, because the incidence of collocations in the corpus is limited, it has been necessary to gather more extensive collocation data in order to improve error analysis reliability. This study reports on the error analyses conducted in constructing a database of errors. The data consists of adjectival collocations produced by Japanese learners in a collocation production survey.

2. COLLOCATION PRODUCTION SURVEY

A survey questionnaire was developed consisting of a 70-item syntagmatic association task. In compiling the list of basic stimulus words, selection was based on whether the target experimental population could be expected to know the words. After extracting all the nouns from the Level 4 (lowest) vocabulary list for the Japanese Language Proficiency Test (2004), a list of common items was compiled by referring to fundamental vocabulary for Japanese language teaching (1984). The final list of 70 target nouns for the questionnaire was selected to balance for semantic categories. The instructions asked the participants to look at each printed item and to write down in a blank space a Japanese word (adjective) that would form a collocation with the printed word (noun), and multiple responses were encouraged. The survey was designed to collect large-scale data for modification phrases (adjective + noun) and predicate phrases (noun + adjective). While the full list of 70 items was administered to 73 learners, due to time constraints shorter questionnaires of 35 items were administered to 117 participants. In both cases the presentation order was randomized (20 different orders for the full list and 12 for the half lists). The participants included 165 Chinese and 25 non-Chinese. The general level of the participants' Japanese proficiency was level 3 or higher on the Japanese Language Proficiency Test. Based on responses to the questionnaire items, the participants were grouped under three levels: 52 upper, 78 middle, and 60 lower level.

3. ERROR CLASSIFICATION

From the collocation production survey, 23,525 collocations responses were collected. These were examined with the assistance of three native Japanese speakers in judging the naturalness of a collocation expression, with 1,449 unnatural collocation expressions being identified. The natural expressions are hereafter referred to as ‘correct usages’, while the unnatural expressions are referred to as ‘error usages’. Drawing on James’ (1998) discussion of learner errors, we have previously developed an error classification scheme (Cao & Nishina, 2006b) in terms of what various kinds of errors can reveal about the L2 knowledge of the learner. The present error usages were classified according to this classification scheme and by what the errors can tell us about learner production strategies. Accordingly, the learner-produced errors were classified into three groups: A semantic group (S) [802 (55.3%)], a formal group (F) [595 (41.1%)], and a small indeterminable group (I) [52 (3.6%)]. The semantic and formal errors were further subdivided. These groups are described below, with the numbers and ratios for each sub-group.

A. Semantic group (S)

Errors under the semantic group relate to learners using existing Japanese forms, but in ways which do not represent their intended meanings. There are four types of semantic errors.

i. S1 Collocational errors [496 (34.2%)]

The term collocations refer to the set of words that frequently appear in the company of a given word. In the case of collocation errors, while the noun and adjective constituents of the combination are used in appropriate forms, the combination itself is unnatural. For example, *厚い顔 /atsui kao/ (thick face), *病気がふかい /byoki ga fukai/ (illness is deep). Some of these errors would be natural if other words are also added, such as *はげしい川 /hageshii kawa/ (violent river), which is natural as part of the longer phrase 流れのはげしい川/nagare no hageshii kawa/ (a river which flows violently), or *頭がはやい /atama ga hayai/ (head is quick), which would be appropriate as part of a longer phrase such as 頭の回転が速い /atama no kaiten ga hayai/ (quick-witted). Some of these errors are due to misreading or misinterpreting the target noun, such as *体が高い /karada ga takai/ (body is tall).

ii. S2 Confusion of sense relations [166 (11.5%)]

It is generally assumed that learners store words in their mental lexicons in terms of sense-relations. Although this suggests that it would be useful for learners to learn words based on sense relations, they can often be a source of confusion. Errors of this kind are the result of misunderstandings about a word's semantic scope, such as using a more general term where a more specific one is needed, or conversely, using a term that is too specific, as well as inappropriate selections of a word from a set of near-synonyms. Examples include *若い子供 /wakai kodomo/ (young child), *秋が浅い /aki ga asai/ (autumn is shallow), and *意味が明るい /imi ga akarui/ (meaning is bright).

iii. S3 Inappropriate kanji [68 (4.7%)]

Errors within this category probably reflect the tendencies of some learners to

prioritize visual information in their language study. Learners use kanji forms that exist in Japanese but these forms do not represent the meaning they wish to express. Examples here would include *細かい木 /komakai ki/ (細かい /komakai/ 'fine, delicate' is used when 細い/hosoi/ would be appropriate for 'slender tree'), *粗い木/arai ki/ (Chinese learners using 粗 which means 'thick' in Chinese, but having the meaning of 'rough' in Japanese), and *下流な絵 /karyuu na e/ (Chinese learners using 下流 which means 'vulgar' in Chinese to describe a picture, but having the meaning of 'lower stream' in Japanese).

iv. S4 Literary or cultural usages [72 (5.0%)]

These are special expressions that are only found in literary works such as *黄色い秋 /kiroi aki/ (yellow autumn, an image from ginkgo or maple trees). While these usages are not completely wrong, as a meaning can be understood, they are rarely ever used. Metaphorical expressions would also be included within this group such as *時間が正直だ /jikan ga shojikida/ (time is honest). Some expressions have special cultural significance such as *おいしい犬 /oishii inu/ (delicious dog).

B. Formal group (F)

The formal group encompasses errors of learners using forms that simply do not exist in Japanese. There are also four types within the formal group.

i. F1 Non-word errors [119 (8.2%)]

These are errors where the 'written' response does not exist in Japanese. They can originate either in the native language or be created by the learner from target language resources (Japanese). Interlingual non-word errors refer to items created based on Chinese resources. One type of interlingual non-word error is borrowing, where the native language word is used in Japanese as it is, such as *新建な建物 (new-built building, where 新建 is a Chinese word but does not exist in Japanese). Another type is creations, where a new combination of elements is formed by inappropriate analogy to existing Japanese words, such as *水がにごい (water is muddy, where the learner attempts to create an adjectival form of にごい /nigoi/ from the existing verb にごる /nigoru/ which means 'get muddy'). The last type is where the new combination is created according to literal translation of an L1 phrase, such as *時間が急い (in a hurry, being a literal translation from the Chinese 時間急 'in a hurry').

ii. F2 Word class errors [219 (15.1%)]

This group refers to parts of speech conversions. For example, when a noun is substituted for an adjective or a verb is substituted for an adjective resulting in a syntactic error. Errors of this kind are particularly prevalent for na-adjectives; one reason is that na-adjectives are very similar to nouns and some na-adjectives can be used as real nouns. Another reason is there are verbs which have adjectival meanings in Japanese. Accordingly, na-adjectives can be quite confusing for Japanese learners. Examples include *雑種な犬 (mongrel dog, → 雑種の), *充実な授業 (full lesson → 充実した), *国が繁栄だ (country is prospering → 繁栄している).

iii. F3 Restrictive use [200 (13.8%)]

While adjectives are used as predicates and noun modifiers, there are, however, differences between predication use and modification use. It is not simply possible to make a modifying pattern from a predicate pattern by changing the order of the elements, because the two usages are not mutually interchangeable. Some adjectives can be used as predicates but not as noun modifiers; while, conversely, some adjectives can be used as noun modifiers but as not predicates. As the Japanese word 多い (many) can not be used as a noun modifier, there are restrictive use errors involving 多い, such as *多い雨 (many rain), *多いお金 (many money) and *多い家庭 (many family). However, it should be noted that there are also some marginal expressions, which could be regarded as natural under some conditions, like *少ない問題 (few problems) and *もったいないお金 (wasteful money).

iv. F4 TEKI[的] suffixation errors [57 (3.9%)]

TEKI is a suffix attached to nouns to form new na-adjectives. It means something like 'assume the characteristics of' or 'the state of', and is similar to the English '-ic, -ive, -al'. Although Japanese TEKI came from Chinese, it has changed in meaning in Japanese. As there are still many TEKI words in Chinese, it is especially confusing for Chinese learners. Accordingly, TEKI suffixation errors are treated as a separate error category, although they could equally be interpreted as non-word errors and word class errors. TEKI errors can be further sub-divided into 1) omissions (*抽象な絵 → 抽象的), 2) inclusions (*友達が善良のだ → 善良), and 3) redundancy (*希望的な春, *体格的な体).

4. CONCLUSIONS

This paper has presented the results of a detailed analysis of adjectival collocation errors produced by learners of Japanese as a foreign language. These results have been used to construct an error database which is a part of a composition support system. As an important aspect of the database, the collocation errors are classified according to an error classification scheme which includes eight categories. The eight categories are organized according to two groups based on their semantic and formal features. The developed error classification, which drew on qualitative and quantitative analyses of learner errors, provides useful insights into both the influence of L1 transfer and aspects of the Japanese language that are particularly problematic for learners. Based on these findings, the results have been used to construct an error database that allows teachers and learners to search for error examples. It could also be used for an error checker or a composition correction system.

5. REFERENCES

- Carl James (1998). Errors in language learning and use -exploring error analysis, Longman.
- Hongquan Cao, and Kikuko Nishina. (2006a). Chinese learners' acquisition of Japanese adjectival collocations: A composition error analysis and pedagogical implications (in Japanese). Journal of Japanese Language Teaching, Vol. 130, 50-59.
- Hongquan Cao, and Kikuko Nishina. (2006b). The development of Japanese adjectival collocations within Chinese learners: Results from a collocation production survey (in Japanese). Japanese Language Education Methods, Vol.13 No.2, 46-47.