

## Use of English prepositions as Japanese predicates: A challenge to NLP\*

Masaharu Shimada (University of Tsukuba)  
shimada.masaharu.fu@u.tsukuba.ac.jp

Akiko Nagano (Tohoku University)  
nagano@ling.human.is.tohoku.ac.jp

### 1. Introduction

One of the differences between English and Japanese is that the former, but not the latter, has a category called a preposition, such as *in*, *on*, and so on. However, in contemporary Japanese, we often find them in brand names of living necessities and dishes. They seem to function as linkers, connecting two nouns in between. For example, (1) is the name of a kind of a liquid soap for washing hair. It contains the English preposition *in*, meaning ‘with.’

- (1) *rinse in shanpuu*  
rinse in shampoo  
‘shampoo with conditioner’

With the expression *rinse in shampoo*, Japanese speakers refer to such a bottle as (2), that is, what English speakers would call “Shampoo with Conditioner” or “Shampoo and conditioner in one.”

(2)



(Shimada and Nagano (2014))

A similar example is given in (3).

- (3) *chiizu in hanbaagu*  
cheese with hamburger  
‘hamburger with cheese into it’

It refers to such a dish as shown in (4).

(4)



(<https://cookpad.com/recipe/3524376>)

The English preposition *in* appears in (3), connecting *chiizu* ‘cheese’ and *hanbaagu* ‘hamburger.’

In this paper, through corpus-based search, we first argue that these types of expressions are formed conforming to a grammar of a particular language and should be taken as grammatical ones. It is thus to be covered by the linguistic research. Second, we point out that they can be ambiguous in interpretation and pose a potential challenge to natural language processing.

Our corpus is Japan’s biggest recipe site called *CookPad*, where visitors can search through user-created recipes. Almost 80 to 90

percent of all Japanese women in their 20s and 30s are said to be using this site. Due to the vast quantity and the freshness of the data, this site can be used as a corpus of ordinary speakers' spontaneous linguistic expressions. Another advantage is that pictures and instructions of dishes help us check and identify the intended interpretation and the grammatical structure of a given recipe name.

## 2. Functions of borrowed prepositions

The grammatical element *-in* is not an original Japanese morpheme. So one might wonder whether the expression in (1) is what is wrongly produced by non-native speakers of English as an incorrect English expression. Namiki (2005) suggests that, as a correct English phrase, (1) should have the structure in which *in* and *shampoo* 'shampoo' form a constituent, modifying *rinse* 'conditioner.' However, as the English translation shows, *-in* in (1) actually takes the preceding noun *rinse*, rather than the following noun *shampoo*. The sequence *rinse in*, but not *in shampoo*, makes a constituent, and behaves as a modifier of *shampoo*. This usage is never possible for English prepositions.

Nevertheless, Namiki (2005) regards (1) as not a destructive and hopeless expression, but a well-formed compound in Japanese. His insight is that the English preposition *in* is borrowed in Japanese to derive expressions like (1) and (3). Also, the borrowing of *in* is assumed to be motivated by the preexistence of the native deverbal noun *iri* 'with, possessing', which is similar to the English preposition *in* in sound and meaning. (5) and

(6) refer to the same dish that (1) and (3) do, respectively.

- (5) *rinse iri shampoo*  
rinse entering shampoo  
'shampoo with conditioner'
- (6) *chiizu iri hanbaagu*  
cheese entering hamburger  
'hamburger with cheese into it'

Namiki's explanation immediately faces a serious problem, however. Other prepositions than *in* are also allowed to produce a similar expression, though we have no native deverbal nouns similar to these prepositions in sound and meaning. *On*, for example, often appears as a linker in a recipe name.

- (7) *kuriimuchiizu on chikin*  
cream cheese on chicken  
'chicken with cream cheese sause'



(<https://cookpad.com/recipe/305114>)

- (8) *tomato on karee doria*  
tomato on curry doria  
'a curry casserole with sliced tomatoes'



(<https://cookpad.com/recipe/849251>)

As in the case of *in*, the interpretations suggest that *on* and the preceding nouns make a unit, modifying the following nouns in (7) and (8).

According to Namiki’s idea, *in* is still prepositional in function in (1) and (3), though the noun it takes lies leftward, not rightward. The function of borrowed *in* is not changed in nature. We propose here that English preposition change into a kind of lexemic predicates through borrowing.

*Rinse-in* (1) and *hanbaagu-in* (3) mean ‘containing rinse’ and ‘containing cheese,’ respectively. *Kuriimuchiizu-in* in (7) and *tomato-on* in (8) mean ‘topped with cream cheese sause’ and ‘topped with tomatoes,’ respectively. Remember that *in* in (1) and (3) can be paraphrased into the native deverbal noun *iri*. It should be noted here that this form is simultaneously for the adverbial form of Japanese verbs. Likewise, *on* in (7) and that in (8) can be paraphrased into the adverbial form of the Japanese verb *kake* ‘put, pour’ and *nose* ‘top,’ respectively. It can be said that the borrowed *-in* and *-on* function as the predicates like *contain* and *top* and do not correspond to functional elements like prepositions.

### 3. On ambiguity

Quite interestingly, we have found through a Google search not only instances of *rinse in shampoo* attested by Namiki, but also instances of *shampoo in rinse*. What is important is that both expressions are used in the same meaning. That is, both (1) and (9) mean “shampoo with conditioner.”

- (9) *shanpuu in rinse*  
shampoo in rinse  
‘shampoo with conditioner’

(Shimada and Nagano (2014))

As in the case of (1), to guarantee the interpretation ‘shampoo with conditioner’ in (9), it is necessary to make a constituent consisting of *in* and *rinse* which modifies *shanpuu*. In (9), however, *rinse* comes rightward, so *in* takes the noun following it. This is the same combining pattern observed in English, where prepositions always takes nouns in the right side position.

*On* in recipe names can also combine with the noun in the right position. Consider (10).

- (10) *tomato on siokonbu*  
tomato on salty seaweed  
‘tomatoes with salty seaweed’



(<https://cookpad.com/recipe/1401320>)

(10) refers to a dish with tomatoes topped with a kind of seaweed. This means that *on* and *siokonbu* ‘salty seaweed’ make a constituent, modifying the preceding noun *tomato*. (10) contrasts with (8) in that the leftmost noun *tomato* combines with *on* in (8), while it does not in (10). It is thus a part of modifier in some cases and a modifiee in other cases. In other words, a partner of *in* or

*on* can be changeable.

In sum, Japanese *X-in/on-Y* sequences is two-way ambiguous in interpretation. *X* can be either a part of modifier, making a unit with *in/on*, or a modifiee. *In/on* can combine with either *X* or *Y*.

#### 4. Concluding remarks

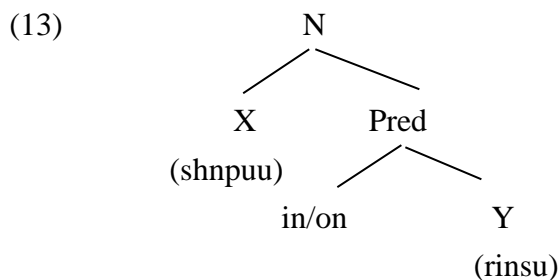
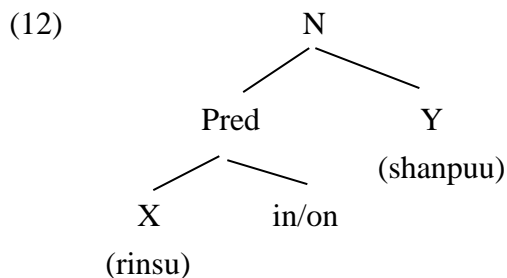
In this paper, we have observed that:

- (11) a. English prepositions are borrowed into Japanese as a kind of predicates. Japanese *X-in/on-Y* sequences standing for brand names conform to a Japanese grammar.
- b. Japanese *X-in/on-Y* sequences are two-way ambiguous in that *in/on* can directly combine with either *X* or *Y*.

As concluding remarks, some comments on (10a) and (11b) are in order.

It follows from (10a) that borrowed *in* and *on* are not functional prepositions in spite of taking the same form as English prepositions. Rather, they are lexical predicates. This means that the form and the function are independent in a grammatical system.

Turning to (10b), linguistically speaking, the ambiguity of *X-in/on-Y* sequences indicates that human language is interpreted depending on a hierarchical structure, not a visible, linear, or sequential pattern. The followings are two potential hierarchical structures for Japanese *X-in/on-Y* sequences.



(12) and (13) are, for example, structures for (1) and (9), respectively. In addition to the category status of *in/on*, NLP must somehow recognize the two hierarchical possibilities in (12) and (13) and properly differentiate between them to successfully interpret expressions involving borrowed *in/on*.

#### References

- Namiki, T. (2015) Nihongo-no atarashii taipu-no fukugougo: rinsu-in shanpuu-to rinsu-iri shanpuu. In T. Oishi, T. Nishihara & K. Toyoshima, eds. *Gendaikeigairon-no chouryu*. Tokyo: Kurosio, 1-19.
- Shimada, M. and A. Nagano (2014) Borrowing of English Adpositions in Japanese, Paper read at the annual conference of LAGB.

\* We used recipe data provided by Cookpad Inc. and the National Institute of Informatics. Also, the sample pictures are all cited from the Cookpad site (<https://cookpad.com/>). We gratefully acknowledge Cookpad Inc. for the permission to use materials from their site. This research is financially supported by JSPS KAKENHI Grant Number 16K13234.