

# 多品詞語カギリのアノテーションとガイドライン

宮岡 大<sup>1</sup> 上山あゆみ<sup>2</sup>

<sup>1</sup>九州大学文学部 1LT15145P@s.kyushu-u.ac.jp

<sup>2</sup>九州大学大学院人文科学研究院 ayumi.ueyama@kyudai.jp

## 1. はじめに

カギリという語は、名詞や接続助詞もしくは動詞など、いくつかの品詞にまたがる、さまざまな用法を持っている。コーパスがより有効性を持つためには、これらの用法の違いがアノテーションによって適切に区別されていることが望ましい。そのためには、川添他(2011)・田中他(2012)・宇津木他(2014)などで提唱されている方式に従い、必要に応じて言語学的テストに基づいて分類のガイドラインを作成することが有用である。本発表では、カギリのさまざまな用法について、分類ガイドラインを作成するにあたって、どのような困難な点があったかを明らかにした上で、実際にアノテーションを試み、その結果を述べる。

## 2. カギリの分類ガイドライン

まず、ここで提案するガイドラインを示す。それぞれの分類について、代表的な例文と分類の問題となりうる点を、そのあとで順に述べる。

A	時・条件	カギリを「～時」「～間は」「～ならば」「～以上」「～からには」に置き換えられる。
B	限定	「今日限りの特売」「その場限り」「1回限り」など、「期間の始点と終点+カギリ」の形式のもの。 カギリをダケに置き換えられる。
C	期限	「今日限りで打ち切り」「今季限りでの引退」など、「期間の終点+カギリ」の形式のもの。 カギリをマデに置き換えられる。
D	極限	「嬉しい限り」「うらやましい限り」「愚かしい限り」など、「感

		情・知覚した印象を表す語+カギリ」の形式のもの。
E	範囲内	「見た限り」「知る限り」「参照する限り」など、主に、「知覚動詞+カギリ」の形式のもの。
F	限度いっぱい	「できる限り」等、「可能表現+カギリ」の形式のもの。また、「声の限り」「見渡す限り」なども含む。
G	制限	該当するのは、「この限り」「その限り」のみ。
H	慣用表現	「限り(助詞)ある」「限り(助詞)ない」(下線部は活用する)。

## 3. 代表的な例文と問題となりうる点

### 3.1. 「A：時・条件」について

「A：時・条件」の代表的な例文は、次に示す通りである。

- (1) a. 余程の確証がない限り、慎重に構える
- b. 今後戦争が起こる限り、核兵器が使われるだろう

仁田(2004)では、この分類Aに当てはまるカギリを、「～時」「～間は」に置き換えることができる時文・条件文と、「～からには」に置き換えることができる原因・理由文とに分けている。私たちが当初は、その考え方にならって、2つを区別していたが、実際にアノテーションをしてみると、カギリを「～時」「～間は」と「～からには」のどちらにも置き換えられることが非常に多く、その2つを区別することが困難であった。

- (2) a. 余程の確証がない時は、慎重に構える
- b. 余程の確証がない間は、慎重に構える
- c. 余程の確証がないからには、慎重に構える

- (3) a. 今後戦争が起こる時は、核兵器が使われるだろう  
 b. ??今後戦争が起こる間は、核兵器が使われるだろう  
 c. 今後戦争が起こるからには、核兵器が使われるだろう

従って、「時・条件」を表すカギリと「原因・理由」を表すカギリを1つのグループとして区別をしないことにした。

### 3.2. 「B：限定」について

「B：限定」の代表的な例文は、次に示す通りである。

- (4) a. 順延は1回限りとされていた  
 b. その場限りで解決してはいけない  
 この分類についての説明は次の「C：期限」の節で行う。

### 3.3. 「C：期限」について

「C：期限」の代表的な例文は、次に示す通りである。

- (5) a. 11月限りで休部となった  
 b. 今季限りでの引退を表明している  
 「B：限定」と「C：期限」には、「数量の限定」という点では共通性がある。しかし、次のような例文が同じ分類になるのは不適切ではないかという疑問が生じた。

- (6) a. 今週限りの大安売り  
 b. 今週限りで引退する  
 そこで、「期間+カギリ」という表現において、その期間が始点と終点を示すような場合は、「B：限定」とし、その期間が終点のみを示すような場合は、「C：期限」とすることに結論づけた。

### 3.4. 「D：極限」について

「D：極限」の代表的な例文は、次に示す通りである。

- (7) a. 私にはうらやましい限りである  
 b. あまりにも愚かしい限りと言うほかない  
 「D：極限」については、当初、「嬉しい限り」「うらやましい限り」などの例文を見て、「感情を表す形容詞+カギリ」というラベルを付けていた。しかし、「愚かしい限り」などの、感情とは

少し異なる例も見られ、また、「光栄の限り」など、形容詞ではないものがカギリに前接する例もあったため、「感情・知覚した印象を表す語+カギリ」の形式のもの、という注記を加えた。

### 3.5. 「E：範囲内」について

「E：範囲内」の代表的な例文は、次に示す通りである。

- (8) a. 見た限りでは、うまく行っていない  
 b. 筆者の知る限り、そのような結果はない  
 「E：範囲内」は、典型例としては、「見た限り」「知る限り」があげられ、当初はそのラベルを「知覚動詞+カギリ」としていた。しかし、同様の用法で、「図を参照する限り」「文献をひもとく限り」など、知覚動詞が前接しない例も見られたため、このグループは知覚動詞に限るべきではないという結論になった。

また、これらの例は、多く次のような文脈であらわれている。

- (9) a. この数字を見た限り、彼の主張は正しい  
 b. 私が知る限り、それはあり得ない  
 そこで、このグループのラベルとして「範囲内」という名称を用いることにした。

### 3.6. 「F：限度いっぱい」について

「F：限度いっぱい」の代表的な例文は、次に示す通りである。

- (10) a. 可能な限り速度を上げる  
 b. 命の限りたたえ続けていく  
 c. 振れる限りに大きく首を振った  
 次のような例文の場合、「E：範囲内」にあたるという意見もあったが、私たちとしては、これらは「F：限度いっぱい」であるとみなしている。  
 (11) a. 見渡す限り大空が広がっている  
 b. 知る限りの関西弁を並べる

「E：範囲内」の場合、たとえば(9a)においては、「私が見ていない範囲では違うかもしれないが、私が見た範囲内では」という意味であり、「私の見た範囲」の内と外を区別し、その内側におけることであることを表している。そのため、カギリの後に対比を表す助詞「は」を入れても、元の意味として通じるのである。

これに対して、「F：限度いっぱい」の場合、たとえば(11a)では、「私が見渡すところ全て」と

いう意味であり、Eのような、「私が見渡している範囲」の内と外を区別する意味は無く、見渡している範囲が全てであることを表している。そのため、カギリの後に「全て」という言葉を入れても、元の意味として通じるものが多い。

このような考察に基づいて、(11)は「F：限度いっぱい」の例であると考えている。

### 3.7. 「G：制限」について

「G：制限」の代表的な例文は、次に示す通りである。

- (12)a. その他の利益付与は、この限りではない  
b. その限りでは、認められている

コーパスの中では、「この限りではない」「その限りにおいて当てはまる」など、「この限り」「その限り」という表現が多く見つかる。

当初、これらの例は、「B：限定」か、「F：限度いっぱい」のどちらかに分類しようと考えていたが、そのどちらとするべきかは、決定できないものが多かった。そもそも、これらの例におけるカギリの意味は「その範囲に該当する」ということであり、従って、BでもFでもないと考えるのが妥当ということになり、別のグループとして設定するにいたった。

### 3.8. 「H：慣用表現」について

「H：慣用表現」の代表的な例文は、次に示す通りである。

- (13)a. 言葉というものにも限りがある  
b. 限りもない哀愁が漂う

コーパスの中には、「限りある資源」「限りの無い命」など、「限り（助詞）ある」「限り（助詞）ない」という形式やそのバリエーションが多くみられる。当初、これらの例は、「G：制限」もしくは「B：限定」のどちらかに分類しようと考えていたが、分類の基準がたてにくく、また、ダケに置き換えられないという意味で「B：限定」とは考えられないということになり、最終的には、「H：慣用表現」として別のグループとみなすことにした。

## 4. アノテーション結果

『現代日本語書き言葉均衡コーパス』(BCCWJ)

からカギリを含んだ例文を抽出し、上記の分類ガイドラインにしたがって、実際にアノテーションを行った。（内訳は、PBから300、PMから61、LBaから57、PNから52、OTから30である。）ガイドライン設計者1人と九州大学文学部の学生1人で500件のアノテーションを行ない、それぞれのアノテーション結果からカップ値を計算したところ、0.798となった。ある程度、信頼性のあるガイドラインが作成できたことになる。今後、ずれの見られるところを中心に、検討を重ねていきたい。

## 謝辞

本ガイドラインの原型は、九州大学文学部での授業の中で、第一著者がメンバーであるグループで作成されたものである。同じグループのメンバーである、吉武柚里氏、香月望美氏、井上郁菜氏に感謝する。また、本研究は科研費基盤研究(C) No.16K02631 の助成を受けたものである。

## 参考文献

- 宇津木舞香, 佐藤未歩, 青木花純, 田中リベカ, 川添愛, 戸次大介 (2014) 「MCN コーパスにおける形式名詞「はず」「わけ」「つもり」のアノテーション」, 言語処理学会第 20 回年次大会発表論文集, pp.1067-1070.
- 川添愛, 斎藤学, 片岡喜代子, 戸次大介 (2011) 「言語情報の確実性に影響する表現およびそのスコープのためのアノテーションガイドライン Ver.2.4」, Technical Report of Department of Information Science, Ochanomizu University, OCHA-IS 10-4.
- 田中リベカ, 小池恵里子, 戸次大介, 川添愛 (2012) 「言語学テストに基づく意味アノテーションのガイドライン設計—確実性判断に関わる表現を中心に」, 言語処理学会第 18 回年次大会発表論文集, pp.401-404.
- 仁田円(2004)「条件文の周辺形式「場合(には)」と「かぎり(は)」について—時間を表す文との関連を中心に—」, 『大阪大学留学生センター研究論集 多文化社会と留学生交流』8号, pp. 37-53.