

# 副詞の統語的特徴に着目した見かけの複合語の構造解析法

7B-1

太田悟

宮崎正弘

新潟大学大学院自然科学研究科

## 1 はじめに

日本語処理において、名詞、および名詞相当の接辞が結合して構成される複合語は、難しい処理対象の一つである。従来、複合語を自動分割し構造を解析する方法 [1][2] や曖昧さを絞り込む方法 [3] に関しては様々な提案がなされてきた。しかし複合語の構造化に重点が置かれてきたため、「昨日山」のようにその内部に文節境界を含む見かけの複合語に関しては、十分な対処がなされていなかった。そのため見かけの複合語の原因となる見かけの要素が、後方に続く複合語と一緒に構造化されるといった誤った解析が行われてきた。

本稿では見かけの要素を持つ複合語を対象として、用例を基に見かけの要素である副詞性名詞を、複合語の構成要素となるか否かの観点から三種類に分類し、文節境界の有無を決定する方法を提案する。

## 2 見かけの複合語とは

見かけの複合語とは、本来複合語とはいえない「見かけの要素+名詞（または名詞相当の複合語）」の構造を持つ形態素列のことである。見かけの要素とは、複合語の構成要素とはなりえないにも関わらず、見かけ上複合語を構成するように見える形態素のことをいう。図1の「早速個人登録」は、見た目にはこれらすべての形態素で複合語を構成するかのように見えるが、実際には「早速」は統語的に離れた関係である見かけの要素であり、形態素列「個人登録」が真の複合語となる。この見かけの要素は副詞性名

詞であり、直後の名詞との間には図1のように文節境界が入る。

例：早速個人登録をします。

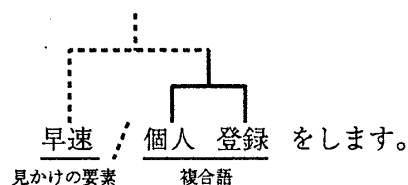


図1. 見かけの複合語

## 3 副詞性名詞の分類

時間や数量などを表す名詞は、格助詞を付与することなしに副詞的に使われる。このような名詞を副詞性名詞と呼び、表1のようなものがある。この副詞性名詞が見かけの複合語を作り出す原因となっている。

表1. 副詞性名詞の例

	例
時間関連	一晩中、従来、同時
人関連	一人一人、全員、単独
その他	事実、事前、早速、一部、多数

ここでは、副詞性名詞を複合語の構成要素となるか否かの観点から、表2のように分離型・接続型・状況依存型の三種類に分類した。以下に各分類の特徴と分類基準を示す。

### 1. 分離型

基準…直後に名詞が続くと独立してしまい、必ず文節境界が入る形態素

特徴…分離型の副詞性名詞は単語の独立性が高いという性質があり、それ自身で文節境界を必要とする形態素である。

Structural Analysis of Japanese Pseudo Compound Noun based on syntactic feature of adverb

Satoru Ohta, Masahiro Miyazaki  
Niigata University

2. 接続型

基準…直後に固有名詞以外の名詞が続くと、必ず一緒に複合語として構造化される形態素

特徴…接続型の副詞性名詞は単語の従属性が強いという性質があるため単独では用いられず、後方の名詞や複合語と一緒に構造化される。

3. 状況依存型

基準…直後に名詞が続くと、その名詞と一緒に複合語として構造化される場合とされない場合の二つの状況が考えられる形態素

特徴…状況依存型に属する副詞性名詞は、それ自身の形態素だけで文節境界の有無を判定することは困難であるため、直後に続く単語の種類についてもチェックする。

事実調査 をすることが必要だ。

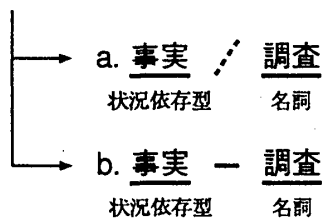


図 2. 状況依存型の例

状況依存型の副詞性名詞は文節境界の有無の判定が難しいが、その直後に続く名詞はある程度決まった字面である場合が多い。これを字面による共起関係と言い、状況依存型の直後に続く名詞と共起関係が成立する場合に、文節境界は入らず一緒に構造化される。

表 2. 副詞性名詞の分類

	例
分離型	一晩中、一人一人、早速
接続型	同時、単独、事前
状況依存型	従来、全員、事実、一部、多数

本分類は新聞記事から抽出した複合語 5943 個のうち、副詞性名詞を含んだ複合語 566 個の使用例の分析を基に行なった結果である。

4 ルールと制御

4.1 各分類型のルール

分類したそれぞれの型の副詞性名詞に対応するルールを表 3 に示す。本ルールにより文節境界の有無を決定する。

表 3. ルール

	接辞	固有名詞	副詞性名詞	サ変名詞	一般名詞
分離型	×	○	○	○	○
接続型	×	○	×	×	×
状況依存型	×	○	×	○	×

○ 文節境界有り  
× 無し

4.2 ルールの制御

形態素解析により単語分割された結果が入力されると、分割された単語のうち副詞性名詞に対しては、分類表を参照してルールの適用を行なう。これにより、文節境界が入るものに対しては、副詞性名詞とその後方に続く単語との間に文節境界を設定し、複合語の構造解析を行なう。

5 おわりに

見かけの複合語の構造を解析するため、複合語の構成要素となるか否かの観点から、副詞性名詞を 3 つの型に分類し、ルール化とその制御法について提案した。これにより見かけの要素を解消することが可能となった。今後は状況依存型副詞性名詞や、構造化された副詞性名詞への対処が必要である。

参考文献

[1] 宮崎、池原、横尾：複合語の構造化に基づく対訳辞書の単語結合型辞書引き、情報処理学会論文誌、Vol.34、No.4、pp.743-754(1993)  
 [2] 佐野、宮崎：拡張 CYK 法による日本語複合名詞の構造解析法、信学会秋季大会、No.D-51(1992)  
 [3] 前川、宮崎：日本語複合名詞の構造的曖昧さの絞り込み法とその評価、情報処理学会第 49 回全国大会、No.1G-5(1994)