

統語解析情報付きコーパスを用いたテキストジャンル間の比較 三好伸芳

本研究の目的は、統語解析情報付きコーパス（ツリーバンク）である NPCMJ を用いて、文学的文章と説明的文章における主述関係の分布の差異を調査し、各テキストジャンルの特性の一端を構文的・文法的観点から明らかにすることである。BCCWJ をはじめとする従来の形態情報のみが付与されたコーパスでは、線状の関係を検索することは可能でも、次の下線部のような階層的関係（係り受け関係）を捉えることは困難である。

- (i) 老人は額を、元の窓の所へ立てかけて、席につくと、私にもその向う側へ坐る様に、手真似をしながら、私の顔を見つめて、こんなことを云った。 132_aozora_Edogawa-1929; JP

NPCMJ を用いた検索では、上述のような例文（本調査で実際に得られた文）も容易に取り出すことが可能である。

名詞や動詞がどのようなテキストジャンルに現れやすいかといった点については、樺島忠夫（1955）「類別した品詞の比率に見られる規則性」（『国語国文』24-6）などに指摘があるが、「どのようなテキストジャンルにおいて、どのような係り受け関係が生じやすいのか」という問題は、技術的な問題からこれまで扱われてこなかった。そこで本研究では、NPCMJ を用いて「どのタイプの名詞句が、どのタイプの主節述語と結びつくのか」を量的に調査した。

調査の結果、文学的文章に見られる顕著な特徴として、次のような代名詞を項とする名詞述語文が相対的に多く生起することが明らかになった。

- (ii) それは何か食べものの匂いだった。 469_aozora_Harada-1960; JP

当該構文では、代名詞の現場指示的性質と叙想的テンスにより、発話者の視点からの探索と場面転換が効果的に表現されている。以上の観察から、本調査では、「主語＝指示代名詞／名詞述語＝タ形」という構文が文学的文章に特徴的に表れており、テキストジャンル間の差異が階層的関係（係り受け関係）からも示せることを明らかにした。