

曖昧性を考慮した料理名のコンピュータ処理

東海大福岡短大 ○八尋剛規 九州女子大附属高校 近葉美希
福岡教育大 澤田吉苗

目的 コンピュータ上で料理名を文字列として文字列比較・検索を行うと、実際に同一の料理であっても、その記述の違いにより、比較や検索が正確に行えない。これは日本語の特徴である助詞や修飾語等による曖昧性を考慮していないためである。そこで、この問題を解決するため、料理名の特徴を調査し、曖昧性を排除した表現におきかえる手法を考察した。

方法 まず、料理名の文法的特徴を調査した。調査対象は料理雑誌などに掲載された日本語表記の料理名とし、フランス料理、中華料理などは対象外とした。次に、最長単語一致法、料理名の文法的特徴を考慮したいくつかの例外処理、食品成分表に記載されている食品名を主に登録している単語辞書を用いて単語分解を行った。最後に、単語分解された料理名から、材料名や調理名などを抽出し、それぞれの材料名や調理名に割り当てた記号や数値を使って一定の書式に置き換えた。

結果 料理名の文法的特徴として、1) 料理名は主に4つの構文をもち (A:材料名、調理名のみ (料理名全体に占める割合は20.8%)、B:材料名+調理名 (同 11.6%)、C:材料名の調理名 (同 44.9%)、D:材料名と材料名の調理名 (同 21.8%)、E:その他 (同 1.0%))、2) 料理名に使用される助詞は格助詞 (連体修飾) の「の」 (料理名全体の「の」の出現頻度に占める割合は92.2%) 及び格助詞 (並列) の「と」 (同 90.5%) であった。料理名の単語分解の成功率は95%程度であった。料理名を記号化することによって、曖昧性を含まない表現で処理することができるようになった。