

辞書定義文と分散表現を用いた漢語の和語への言い換え

関沢祐樹 梶原智之 小町守(首都大学東京)

sekizawa-yuuki@ed.tmu.ac.jp

目的

机上 → 机の上
 気丈 → 気持ちが丈夫

同音異義語		難易度
漢語	多い	高い
和語	少ない	低い

漢語を和語へ言い換えて可聴性・可読性を向上させることを目的とした

先行研究

萩行正嗣, 黒橋禎夫. 辞書定義文を用いた
 二字漢語の言い換え表現の生成.
 言語処理学会第15回年次大会発表論文集, pp.256-259, 2009.

辞書定義文で漢語の語構成漢字に
 対応する部分とそれを漢語の言い換えとする

二字漢語: 水鳥

辞書定義文: 川や湖の水辺にすむ鳥

定義文解析: 川や
 湖の | 水辺に—すむ—鳥

対応部探索: 川や
 湖の | 水辺に—すむ—鳥

言い換え: 水辺にすむ鳥

問題点: 先行研究では、定義文に語構成漢字の
 対応部分がないと言い換えの生成が出来ない

実験結果

漢語に対して言い換えを生成した数

漢語数: 13,213	先行研究	提案手法
言い換えた漢語数	5,430 (41.1%)	13,008 (98.5%)

漢語に対して正しい言い換えを生成した数
 13,213語からランダムに500語を選択

漢語数: 500	先行研究	提案手法
正しい言い換え数	120 (24.0%)	124 (24.8%)

和語は漢語では修飾されない。例えば、「直射日光」と言うが、「直射日の光」とは言わない
 したがって、漢語連続の名詞句は和語の名詞句と文脈が異なる

提案手法

先行研究で言い換えを生成できなかった漢語を
 辞書定義文を使わずに、辞書定義文から抽出した
 言い換えパターンを用いて言い換えを生成



正しい言い換えを生成できた例

二字漢語: 初旬
 定義文: 月の初めの十日間
 言い換え: 初め

言い換え出来そうだけれど言い換え出来なかった例

二字漢語: 日光
 予想言い換え: 日の光
 言い換え: 日の出

言い換え前後の難易度の平均 (日本語教育語彙表, ver.0.1)

