

# 翻訳は語順が大切

～日本語風の英語を介した英日・日英統計翻訳～

## どんな研究？

対訳例文を大量に使って翻訳の規則を確率的にモデル化し、機械翻訳を実現する「統計的機械翻訳」に関する研究です。これまで英語と日本語の間の翻訳では語順の違いが大きく、うまく翻訳できないことがありました。この展示では、英語と日本語の語順の違いを克服する技術を紹介します。

## どこが凄い？

日本語の特徴に着目し、英語を日本語に近い語順に並べ替えた「日本語風の」英語を介した翻訳技術を考案しました。特許文書の英日翻訳では、翻訳精度を従来高性能の翻訳ソフトを上回る水準に向上させました。日英翻訳でも統計的機械翻訳の最先端の方法より約2倍高速な翻訳を可能にしました。

## どんな風に役立つ？

統計的機械翻訳は翻訳対象の対訳例文があれば学習できるので、例えば社内文書のような特定の分野の文をこの技術によって精度よく翻訳することができるようになり、言語の壁を越えたコミュニケーションを円滑に行えるようになります。将来的には中国語など他の言語への展開も目指しています。

## 英日翻訳

One terminal electrode 20 is positioned on the chip element body 10.

従来の統計翻訳

一方の端子電極 20 は、チップ素子上に位置している本体 10。

単語の並べ替えに失敗した誤訳

英語の並べ替え

構文に基づく簡単な規則

One terminal electrode 20 ga chip element body 10 on positioned is.

日本語語順の英語

逐語訳なので簡単に訳せる!

統計翻訳

一方の端子電極 20 は、チップ素子本体 10 上に位置している。

語順も正しい正確な訳!!

## 日英翻訳

One terminal electrode 20 is positioned on the chip element body 10.

正しい英語への並べ替え

単語を翻訳しない統計翻訳!

日本語語順の英語

One terminal electrode 20 ga chip element body 10 on positioned is.

逐語訳なので簡単に訳せる!

統計翻訳

一方の端子電極 20 は、チップ素子本体 10 上に位置している。

### 日本語の特徴「主辞後置性」

(主辞≒係り受け関係の「係り先」)

「日本語は原則左から右に係る」

彼は 彼女を 好き  
主辞

一方...

「英語は係る方向が混在」

He likes her  
主辞

### 「主辞後置並べ替え」

「英語でも主辞(係り先)を後ろに並べ替えれば日本語風の語順になる」

He ga her o likes  
助詞の「が」「を」を補充

## 関連文献

- [1] H. Isozaki, K. Sudoh, H. Tsukada, M. Nagata, "Head Finalization: A Simple Reordering Rule for SOV Languages," Proc. of WMT-MetricsMATR. pp.244-251, 2010  
[2] K. Sudoh, X. Wu, K. Duh, H. Tsukada, M. Nagata, "Post-ordering in Statistical Machine Translation," Proc. of MT Summit XIII. pp.316-323, 2011

## 連絡先

須藤克仁 (Katsuhito Sudoh) 協創情報研究部 言語知能研究グループ  
E-mail : sudoh.katsuhito[at]lab.ntt.co.jp ( {at} の部分を @ に置き換えてください)