

翻訳生産性向上のテクニック

ー PC-Transer 翻訳スタジオ活用法 第3回 文型翻訳メモリを使いこなす

小室誠一 (MT 研究会主宰 : <http://www.babel.co.jp/mtsg/>)

翻訳生産性を飛躍的に向上させるには、訳文の自動生成機能を活用することが不可欠です。従来の翻訳メモリソフトを利用した手作業による翻訳処理ではどんなに頑張っても 2 倍程度にしかありませんが、PC-Transer の「文型翻訳メモリ」を使えば生産性 10 倍も夢ではありません。

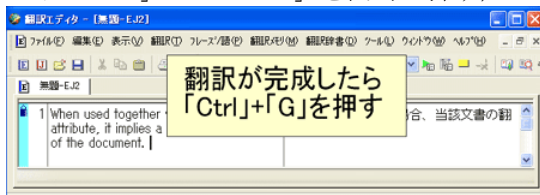
今回は前回に引き続いて、「文型一致文検索」を行うために必要な「文型翻訳メモリ (タグ付けした翻訳メモリをこう呼ぶことにします)」の作成方法を少し詳しく見てまいりましょう。

●簡単なタグの付け方

タグ付けというとなんだか面倒な感じがしますが、「翻訳メモリペイン」の「タグ指定」ボタンを使えば大変簡単です。

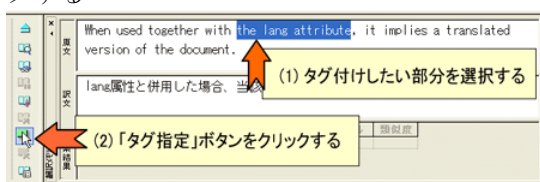
(1) 対訳エディタで翻訳を完成させる

(2) ショートカットキー「Ctrl+G」を押す (「コントロール」キーと「G」を同時に押す)



(3) 対訳文が「翻訳メモリペイン」にセットされる

(4) 「翻訳メモリペイン」で原文の変更可能な部分を選択して「タグ指定」ボタンをクリックする



(5) 同様に訳文の対応する部分を選択して「タグ指定」ボタンをクリックする

(6) 複数ある場合は(4)(5)を繰り返す

タグは1文中に最高10個まで付けることができますが、連続して（隣接して）記述することはできません。また、あまりたくさん付けてもうまく行かないことがあります。2～3 程度にしておくのが適切でしょう。

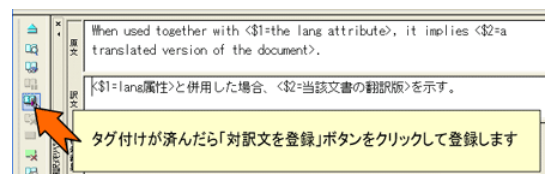
「翻訳メモリペイン」の「タグ指定」ボタンでタグ付けすると、以下のように自動的に数字が付きます。

原文 : When used together with <\$1=the lang attribute>, it implies <\$2=a translated version of the document>.

訳文 : <\$1=lang 属性>と併用した場合、<\$2=当該文書の翻訳版>を示す。

「対訳文を登録」ボタンをクリックすると「登録」ダイアログが開いて、事前に指定した翻訳メモリに登録できます。

この時、タグを省いた完全翻訳メモリとタグの付いた文型翻訳メモリが一度に登録されます。



●タグの詳細指定

タグは「<\$タグ名>」という記号でくくりますが、タグ名がただの数字だと、どのような意図でタグ付けしたか、後でわかりにくくなる場合があります。

There is illustrated in <\$1> a <\$2>.
<\$1>に<\$2>が示される。

タグ名には、英数字、ひらがな、カタカナ、漢字が使えます。「\$=」を含めて 64 バイト(全角 32 文字)まで使用できます。

そこで、以下のように、変更可能な部分にどのような語句が入るのがわかるようにタグ名を付けることができます。

There is illustrated in <\$図の番号> a <\$対象物>.
<\$図の番号>に<\$対象物>が示される。

さて、PC・Transer の文型翻訳メモリはこれだけではありません。実は何と訳文の生成情報も記述することができるのです！

生成情報はタグの後に「/」で区切って何個でも記述できます。

原文を名詞句として訳すには「NP」、動詞句として訳すには「VP」、形容詞句なら「AP」、副詞句なら「DP」、さらに文末がピリオドでない原文を文として訳すなら「SE」などが指定できます。指定しないときは名詞句になります。

試しに以下の文型翻訳メモリを登録します。

原文：It took me <\$1> to <\$2>.
訳文：<\$2>のに私は<\$1>かかった。

この翻訳メモリを使って次の文を訳します。

「It took me three days to clean the room.」

「clean the room」は動詞句ですが、<\$2>に生成情報の指定がないので名詞句扱いとなり、文型検索にマッチせずに機械翻訳されてしまいました。

「部屋をきれいにするために、私には 3 日か

かった。」

そこで、文型翻訳メモリの訳文側に動詞句の指定を追加します。VP は動詞句、S は終止形で訳出するという意味です。

原文：It took me <\$1> to <\$2>.
訳文：<\$2/VP:S>のに私は<\$1>かかった。

こんどはうまく文型検索にマッチして以下のような訳文が出力されました。

「部屋をきれいにするのに私は 3 日かかった。」

●「文型翻訳メモリ」使用上の注意

もう少し例を挙げてみましょう。

以下の文型翻訳メモリを登録します。

原文：It is dangerous to <\$1>.
訳文：<\$1/VP:S>と危険だ。

この翻訳メモリを使っていくつか例文を訳してみます。

It is dangerous to drink too much.
飲みすぎると危険だ。

It is dangerous to tease a bear.
クマをからかうと危険だ。

It is dangerous to wake up a sleepwalker.
夢遊病者を起こすと危険だ。

It is dangerous to put metal in the microwave.
金属を電子レンジに入れると危険だ。

It is dangerous to heat an organic compound in a distilling apparatus.
蒸留している装置で有機化合物を熱すると危険だ。

一定のパターンでうまく訳出されています。

ここで注意しなければならないのは、タグ付けした部分に機械翻訳した訳文が埋め込まれるということです。つまりタグ付けした部分が不自然な訳文になる可能性があるのです。

It is dangerous to hire somebody who has too much experience.

あまりにたくさんの経験をする誰かを雇うと危険だ。

この場合、「～と危険だ。」は問題ありませんが、「あまりにたくさんの経験をする誰か」は不自然です。「somebody who has too much experience」を「経験のありすぎる人」などとユーザー辞書に名詞で登録しておく必要があります。

ただし、簡易登録だと「経験のありすぎる人を借りると危険だ。」と、「hire」が「借りる」と訳出されてしまいます。

そこで、詳細登録で、「意味素性」を「人間」に指定します。すると、「経験のありすぎる人を雇うと危険だ。」とうまく訳し分けてくれます。

このように、文型翻訳メモリを活用するには、効果的なタグ付けとともにユーザー辞書の援護射撃も忘れてはなりません。常に、NP、VP、AP、DP など生成情報を念頭において、単語ではなく「フレーズ」を登録して行くことが大切です。

「文型翻訳メモリ」は大変有効な機能です。それほど手間がかからないので、翻訳の作業手順に含めれば無理なく蓄積できます。「文型翻訳メモリ」とフレーズ辞書が十分に蓄積されれば、驚くほど生産性が向上することは間違いありません。

