

Adobe InDesign CS/CS2 に翻訳メモリツールを適用するための フィルターの開発について

2006年6月12日

株式会社アイデア・インスティテュート¹

研究開発本部 本部長 天野 薫

従来、Adobe FrameMaker で制作し翻訳展開していたようなマニュアル類まで、最近では Adobe InDesign を用いて制作することも多くなってきました。FrameMaker の場合には SDL Trados のような翻訳メモリ（TM）ツールを適用してレイアウトをよく保持したまま能率よく翻訳できましたが、InDesign ではどうでしょうか。FrameMaker と InDesign の特性の違いもありますが、その答えは、InDesign データから翻訳対象テキストを抽出し、そして逆に翻訳したテキストを InDesign データに戻すための機構（フィルター）に左右されます。ここでは、Trados を中心に標準的な InDesign 向けフィルターの現状を簡単にご報告するとともに、InDesign 文書のエクスポート形式のひとつである InDesign Interchange 形式（INX）に注目し、これを翻訳する立場から解説し、またそのためのフィルター開発の実際をご紹介します。

要約

- 現在では InDesign 最新版 CS2 も含め、各 TM ツールにフィルターが揃っている。Trados の場合、InDesign CS でのインライン枠の処理などやや不便。
- INX は InDesign 文書データを XML に書き出したもの。InDesign で開きほぼ同じ文書を復元することが可能。
- 原理的には素朴な INX 翻訳でインライン枠を含めレイアウトをよく保持することが可能。アイデア・インスティテュートで開発したフィルターで Trados 適用可能。

1 InDesign の交換形式 INX

1.1. InDesign Interchange 形式（INX）とは²

- InDesign 文書（INDD）を再構成するのに必要な情報を文字で書き出したもの。同じ InDesign で開く ▶ レイアウトも再現される。Word の RTF、FrameMaker の MIF に相当すべきもの。
- XML。一定の文書定義などはない。
- 用途：
データ下位互換性（CS2）（3.0.1 へのダウングレード）、
翻訳³、
その他。

¹ 翻訳会社。http://www.ideainstitut.co.jp

² 文献[1]。

³ INX を翻訳対象としているツールの例: Heartsome Translation Suite、SDL Trados 7.1、7.5=2006。

- 応用技術：
InCopy とのデータ交換（CS2）（INCX 形式）、
Page 部品の書き出し・再利用（CS2）（スニペット）

1.2. テキスト周辺

- テキストは文字属性が均一になるまでは分割される。
- 通常は接頭辞 c_ で始まる文字列。
例外的に、文字が e_s... の形（e_ssrq など）や処理命令（PI。<?ACE 18?>など）で表されることも。
- インライン枠等の挿入は PI で。

（例）単純なテキストとインライン枠の挿入

<pre><txsr prst="o_u62" crst="o_u63"> <pcnt> c_An inline image <?aid Self="rc_u9bcins41"?> . </pcnt> </txsr></pre>	<p>An inline image </p>
--	--

（例）文字スタイルの変化：txsr は入れ子にならない。

<pre><txsr prst="o_uda" crst="o_u63" > <pcnt>c_The basic idea is to translate</pcnt> </txsr> <txsr prst="o_uda" crst="o_ue7"> <pcnt>c_INX</pcnt> </txsr> <txsr prst="o_uda" crst="o_u63"> <pcnt>c_documents.</pcnt> </txsr></pre>	<p>The basic idea is to translate <u>INX</u> documents.</p>
---	---

（例）処理命令で表された文字

<pre><txsr prst="o_u62" crst="o_u63"> <pcnt> c_ <?ACE d800?> <?ACE dc00?> </pcnt> </txsr></pre>

2 TM ツールを InDesign に適用するためのフィルター（現状）

2.1. 各ツールの InDesign 対応⁴

InDesign CS (2003/10 3.0)、InDesign CS2 (2005/5 4.0) に対して：

- Star Transit XV
InDesign CS フィルター (2004/5 SP11)、
InDesign CS2 サポート (2006/3 SP16)
- SDL SDLX
InDesign CS フィルター (2004/10 2004Bonus)、
InDesign CS2 対応 (2005/12 2005SP1)

⁴ InDesign CS 対応に関し Transit、SDLX、Trados などの比較については文献[2]参照。

- Trados
Story Collector for InDesign CS（2004/8 6.5.5）、
InDesign CS2 対応の新フィルター（2005/12 SDL Trados 7.1）

2.2. SDL Trados の場合

SDLX のフィルターを利用することもできるが標準では、InDesign バージョンにより Story Collector か新フィルターのいずれかを用いる。

（それぞれ性質が全く異なる。）

2.2.1. Story Collector

- InDesign CS 対応（2.0 用のものもあるが、CS2 には対応しない）
- InDesign の標準機能（Export と Place）を利用するもの（プラグイン）。
 - 抽出形式 ISC は InDesign Tagged Text に基づく。
 - TOC、Index などがあるとクラッシュのおそれ⁵。
- 翻訳用に抽出するストーリーの順序を選択できる。
- インライン枠はストーリー先頭に寄せられてしまう。
（文中の挿入位置は保持されない。）

2.2.2. SDL Trados フィルター

- InDesign CS2 対応（CS 以下には対応しない）。
- INX を解析・翻訳する。
- フォント・マッピング等機能あり。

3 アイデア・インスティテュートで開発したフィルター

3.1. 要請

- InDesign CS と CS2 の両方がターゲット。
- 翻訳用に抽出するストーリーの順序を選択できる。
- 文中のインライン枠の挿入位置などを含めレイアウトを忠実に保持する。
- Trados での翻訳に適する。
- できるだけ短期に、動くものを実現する。

3.2. 実装方針

- INX ベース。
INX をテキスト本位の翻訳しやすい形式に変換する。
- INX の難しい解釈は避ける。
 - テキスト枠のページ上の位置情報などは InDesign に問い合わせる。
 - 文中に挿入された PI はそのままにする。
- ストーリーのテキストのみ翻訳対象とする。

⁵ Story Collector の問題については SDL Trados 製品リリースノートに記載あり [3]。

3.3. 中間形式

3.3.1. JOY（翻訳用 XML）

INX では属性別に切り分けられたテキスト断片を翻訳用に再構成したもの。

文中に現れるものは、PI のほかは次の要素のタグ

PS：段落区切り、

LS：任意改行、

c：文字スタイル、

o：その他文字属性、

e：特殊文字

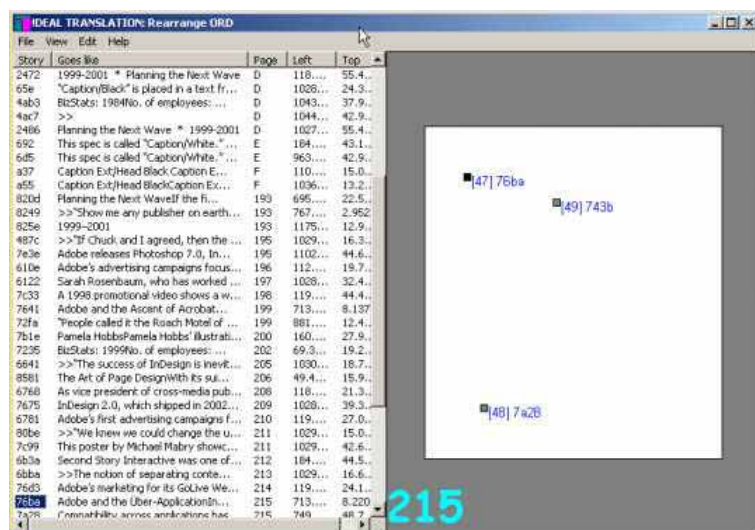
```
<pgf_style nam="Normal" i="1">
  The basic idea is to translate
  <c i="1">INX</c>
  documents. ...
  <PS />
  ...
</pgf_style>
```

3.3.2. ORD（テキスト抽出順序設定用 XML）

InDesign にストーリーの位置情報を問い合わせることで作成。

ORD でのストーリーの順序 ▶ JOY でのストーリーの順序

3.4. ORD 編集

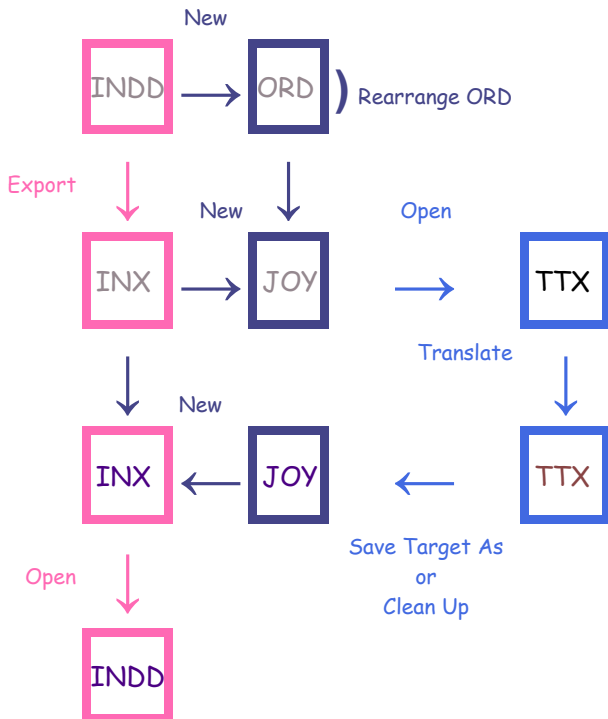


3.5. ワークフロー

InDesign、当該フィルター（Tradobeindex）、Trados を使用。

1. INX を用意する
2. Options 設定
3. ORD を生成
4. ORD を編集（ストーリーの並べ替え、削除）
5. JOY を生成
6. JOY を解析・翻訳

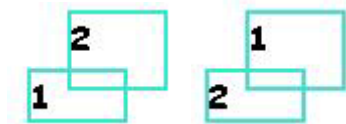
7. JOY のテキストを INX に埋め込む
8. INX を InDesign で Open、調整、保存



3.5.1. 設定事項

ORD オプション

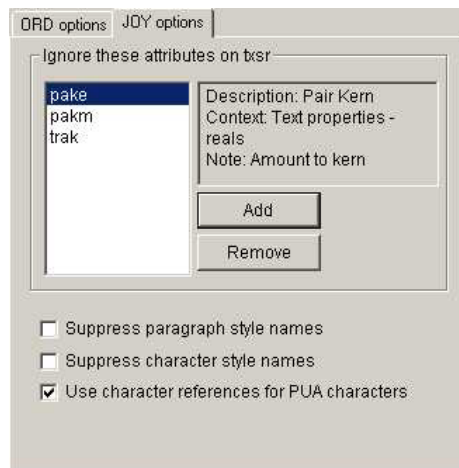
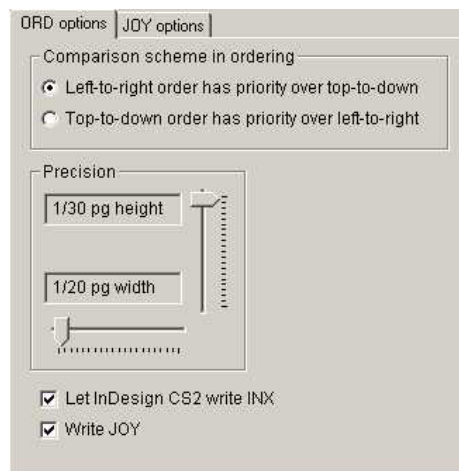
- Comparison scheme
左右、上下のどちらが優先か。

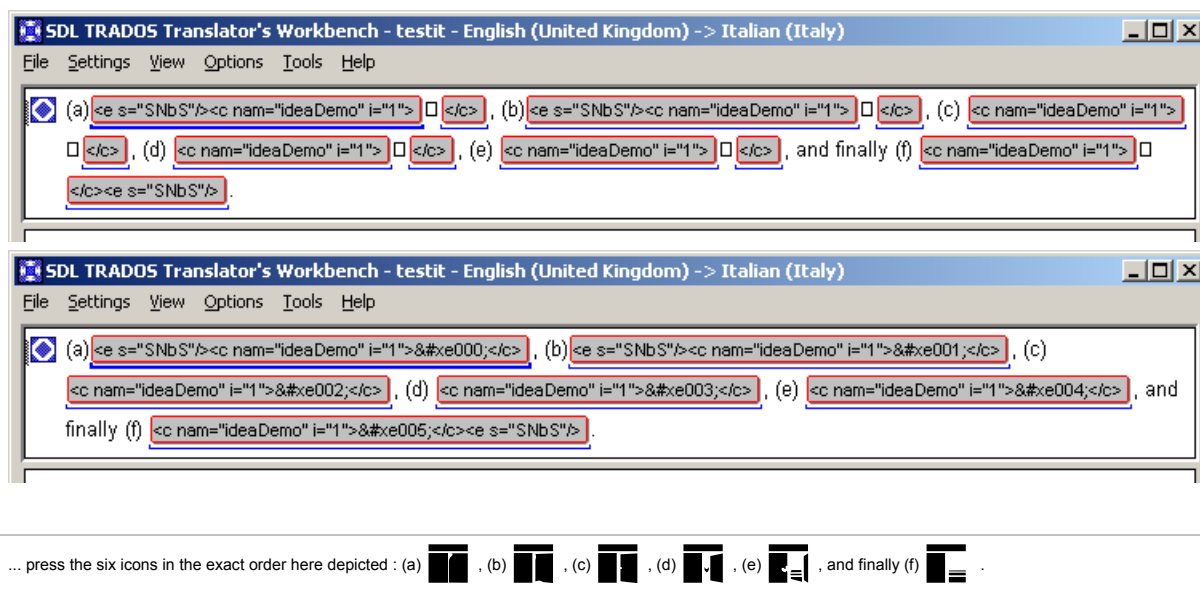


- Precision
どのくらい小さな差異を認めるか。
- InDesign CS2 に INX を書かせる
- JOY も生成する

JOY オプション

- Ignore these attributes
テキスト属性のうち JOY 生成の際に捨て去るもの。
- 段落スタイル名を省略
- 文字スタイル名を省略
- Use character reference for PUA characters
その意義：Unicode 私用領域の文字の区別





4 今後の課題と予定

現状、INX 翻訳の立場のフィルタとして期待できる基本性能は確保。

4.1. 改善したほうがよい点

- テキスト位置の情報取得と抽出順の設定
- ワークフローが適用か事前にデータ検証するための補助機能 (ダミー翻訳など)
- その他

4.2. 今後の課題

- 使用経験の蓄積▶問題点、適用範囲の見極め。
一般の方からもフィードバック頂けるよう成果フィルターを公開する。(準備中)
- ストーリー・テキスト以外の翻訳対象への対応
- 部分訳、多言語併記文書制作への応用 (準備中)
- 翻訳展開のための InDesign の機能拡張 (FrameMaker 流の相互参照機能など) とフィルター側の対応 (準備中)

参考文献

- [1] Adobe Systems Incorporated, *Working with INX file format*, Solution Document #10091 (2005/11/8) および *InCopy Interchange format*, Technical Note #10098 (2005/11/4); とともに InDesign CS2 4.0.2 Products SDK の一部として Adobe Systems 社から配布されています。
- [2] Uwe Scholz and Roman Kotzsch, *Localizing InDesign documents*, MultiLingual Computing & Technology #76 Vol 16 Issue 8 (2005), MultiLingual Computing, Inc.
- [3] SDL International, *SDL Trados 2006 Release Notes* (2006/3)または *SDL Trados 7.1 Release Notes* (2005/12)