

TR-IT-0190

日韓間変換主導翻訳の中間時評価

Report on Mid-term performance evaluation of Transfer-Driven
Machine Translation between Japanese and Korean

山本 和英 古瀬 蔵
Kazuhide YAMAMOTO and Osamu FURUSE

1996 年 9 月

概 要

変換主導型翻訳 (TDMT) の日韓、韓日翻訳の中間時 (1996 年 8 月時点) 評価結果について報告する。評価は翻訳品質、原言語構造、翻訳時間の三項目で行なった。評価の結果、日韓翻訳は 71% の翻訳率、韓日翻訳は 53% の翻訳率となり、両翻訳共に誤った構造解析を行なっても出力さえすれば概ね高品質の翻訳結果となることが明らかになった。翻訳時間は平均的な文長で日韓 0.4 秒 (9.46 形態素 / 文)、韓日 0.1 秒 (8.70 形態素 / 文) となり、実時間処理の要求に十分対応できることが確認された。

目次

1	はじめに	1
1.1	評価対象文	2
1.2	評価項目	2
1.2.1	翻訳品質評価	2
1.2.2	原言語構造評価	3
1.2.3	翻訳時間	3
1.3	用語の定義	4
2	日韓翻訳	5
2.1	概要	5
2.2	構造評価との関連	7
2.3	翻訳処理時間	8
2.4	誤りの分析	9
3	韓日翻訳	10
3.1	概要	10
3.2	構造評価との関連	12
3.3	翻訳処理時間	13
3.4	誤りの分析	14
3.5	Rank-D の分類	14
4	まとめ	17
	参考文献	19
	(付録)Rank-D 文一覧	20

第 1 章

はじめに

本報告書では、1996年8月時点における TDMT[Yam96b] の日韓翻訳、韓日翻訳の性能評価結果について述べる。以下、本章で評価作業の概要を示した後、次章では日韓翻訳、第3章では韓日翻訳について述べ、最終章でまとめを行なう。また今後取り組むべき問題を提示する目的で、日韓翻訳、韓日翻訳それぞれの品質評価で Rank-D¹ となった文の一覧を付録に示す。

今回の中間時評価で行なった評価項目は以下の三点である。

1. 翻訳品質
2. 原言語構造
3. 翻訳処理時間

このうち、翻訳品質と原言語構造の評価については翻訳対象文と学習文が異なるオープンテストで、翻訳処理時間の評価については変換知識学習に使用した文の一部(日韓 779 文、韓日 132 文)を対象にしたクローズドテストで行なった。また、翻訳処理時におけるパターンマッチングのビーム幅(任意の部分文字列に対し、保持する構造数)は、予備実験で有意な差が得られなかったため、1のみとした。

なお、本報告書で述べる翻訳評価では、翻訳入力を文字列ではなく、形態素タギングデータとしており、形態素解析失敗による翻訳への影響及び形態素処理時間の影響を評価の対象外としている。TDMTにおける形態素解析の性能については Technical Report[Yam96a] を参照のこと。

以下では、日韓、韓日翻訳に関係する部分について評価項目の概略を述べる(詳細は文献[Mim96]を参照のこと)。

¹Rank(翻訳品質評価)については1.2.1節で述べる。

1.1 評価対象文

日韓翻訳は、形態素情報が付与された日英バイリンガル会話から 69 会話を無作為に選定し、そのうち通訳者の発話を含まない日本語話者の発話のみを翻訳評価対象とした。ただし、選定した 69 会話が高頻度語、多会話に出現する語をなるべく多く含むことを確認した。評価文を日英会話から選択した理由は、日英翻訳、日独翻訳と同一の文を対象とし、比較調査も行なうためである。また、オープンテストを行なうため、いずれも翻訳未訓練会話から選定した。

韓日翻訳は、形態素情報が付与された日韓バイリンガル会話より 92 会話を無作為に選定し、韓国語話者の会話部分を対象とした。これらも日韓翻訳と同様に翻訳未訓練会話である。

それぞれの会話には発話者(旅行者/乗務員など)、会話場面(空港/ホテルなど)、会話形態(対面/電話)などの会話状況を与える。

1.2 評価項目

1.2.1 翻訳品質評価

翻訳品質評価は、目的言語を母語とし、原言語にも堪能である評価者(日韓翻訳、韓日翻訳いずれも二名)に依頼した。

評価基準は以下の 5 段階とした²。

Rank-A No problem: 問題なし

Rank-B Fair: ほぼ大丈夫、容易に内容が理解できる

Rank-C Acceptable: 何とか内容が理解できる

Rank-D Nonsense, Wrong sense: 内容が理解できない、または全く誤った内容である

未出力 (翻訳失敗: 翻訳結果を出力しない)

これらのうち、翻訳結果を出力しない未出力文については、出力文の評価が不可能であるので評価は行なっていない。また、Rank-B/Rank-C/Rank-D については、翻訳結果が不自然な理由を以下の問題点の中から選択する作業(複数指摘を認める)、および翻訳結果についてのコメントも同時に依頼した。

[語選択] Word Selection: 語の選択が不適切

[活用] Conjugation: 出現語の形態、活用

²ただし未出力は自動的に決まるので、実際に評価者が評価するのは Rank-A から Rank-D の 4 段階である。

[構文] Word Order: 語順の誤り、主格や目的格の欠落による構文誤り

[語用] Sentence Style: 文体

ごちない、あるいは native は使わない、など

[情報量] Information: 情報が不適切

きれいな文でも文脈上不適切な場合も該当する

[限定詞] Determiner: 冠詞や限定詞などの誤り

また、今回の評価に関し、以下の注意を評価者に与えた。

- 対象はあくまで話し言葉である
- 出力結果は文字でなく音で聞いたと仮定する
- 大文字、小文字、ハイフン、句読点の問題は無視する
- 同じ出力結果でも、文脈に応じて評価結果が変わることもあり得る
- (英語における)Mr./Ms. Suzuki など、文脈でも決められない場合は誤りとしてしない

1.2.2 原言語構造評価

原言語構造評価は TDMT の構造出力を理解できる評価者(日韓翻訳、韓日翻訳いずれも一名³)に依頼した。評価は、翻訳品質評価と同一の評価文を対象に、それぞれの原言語構造に対して正解もしくは誤りの評価を与える方法で行なった。つまり、構造に何らかの誤りがある場合は誤りとし、すべて正しい場合に限り正解の評価を与えた。

1.2.3 翻訳時間

翻訳時間評価は、変換知識学習に使用した文(日韓 779 文、韓日 132 文)を対象にして計測した。計測した時間には、形態素解析に要する時間、および Garbage Collection に要する時間を含まない。

³日韓翻訳と、日英、日独翻訳は同一の日本語に対して評価するので、原言語構造はすべて同一である。このため、原言語構造評価は同一の評価者が行なった。

1.3 用語の定義

以下、本報告書で使用する語句の定義を行なう。

- 出力率
全翻訳評価文(のべ)に対する、何らかの翻訳結果を出力した文(のべ)の割合。
- 翻訳精度
翻訳結果を出力した全文に対する、Rank-C以上の文(のべ)の割合。
- 翻訳率(= 出力率 × 翻訳精度)
全翻訳評価文(のべ)に対する、Rank-C以上の文(のべ)の割合。

第 2 章

日韓翻訳

本章では日韓翻訳の評価結果について述べる。なお、評価を行なった時点での日韓翻訳の訓練文数はのべ 1041 文、異なり 943 文であり、(“X の Y” などの) 変換パターン規則数は 680、変換パターン用例数は 3605 である。

2.1 概要

表 2.1 に日韓翻訳における翻訳品質の評価結果の概要を示す。なお、表中の「割合」はのべで計算した割合¹を、「文長」は一文あたりの平均形態素数を、「意味距離」は一文平均の意味距離を意味する。

表 2.1: 日韓翻訳評価: 概要

	#1 ²	#2	割合	文長	意味距離
Rank-A	537	531	42.8%	5.08	0.403
Rank-B	225	297	20.9%	11.01	1.770
Rank-C	103	66	6.8%	12.88	2.536
Rank-D	36	7	1.7%	11.79	2.201
未出力	346	346	27.7%	14.06	—
計	1247	1247	100%	9.46	0.753

表 2.1 から、日韓翻訳での出力率、翻訳精度、翻訳率を計算すると、以下のようになる。

- 出力率: 72.3%
- 翻訳精度: 97.6%

¹例えば Rank-A では $(537 + 531)/(1247 \times 2) \approx 0.428$ と計算した。

²以下、#1、#2 はそれぞれの翻訳評価者を表す。

- 翻訳率 (= 出力率 × 翻訳精度): 70.5%

文長と翻訳品質の関係では、Rank-A の文長が他と比較して極端に短い。また意味距離との関係でも、Rank-A だけは極端に意味距離が小さい。このことは、会話に頻出するあいさつなど、パターンマッチによって意味距離計算することなしに翻訳されている定型の短い表現の多くが Rank-A と判定されていることを示している。

表 2.2: 日韓翻訳での評価者間の相関

#1\#2	Rank-A	Rank-B	Rank-C	Rank-D	計
Rank-A	464	69	4	0	537
Rank-B	60	148	16	1	225
Rank-C	7	67	28	1	103
Rank-D	0	13	18	5	36
計	531	297	66	7	901

表 2.2に、2名の評価者間の評価の相関を示す。この表より、両評価者が一致した評価をした割合が645文(未出力を除く評価文901文に対して71.6%)、Rank-AとRank-Bなど、評価が一段階異なるものもこれに含めると876文(同97.2%)であり、評価者の個人差はほぼないと考えることができる。

全評価文(のべ1247文)のうち、意味距離が0であった文数は232文(18.6%)であった。表 2.3に、これら意味距離が0であった232文の評価結果を示す。

表 2.3: 意味距離0の文の評価

	#1	#2	計
Rank-A	220	212	432
Rank-B	10	17	27
Rank-C	1	3	4
Rank-D	1	0	1
計	232	232	464

表 2.3より、意味距離が0の文のほとんど(93.1% : 432文 ÷ 464文)が Rank-A となっていることがわかるが、その一方で Rank-C が4文、Rank-D が1文存在した。以下に、これらに該当する Rank-C、Rank-D の5文を示す。

表 2.4: 意味距離 0 で Rank-C または Rank-D の文

日本語入力	韓国語翻訳結果	備考
そうですね	그렇죠?	(頻度 2)
よろしいですか	좋습니까?	
確認いたします	확인해 드리겠습니다.	
承りました	받았습니다.	

以上示した5文は、いずれも単一の文の翻訳結果として見た場合には問題ないことから、これらの文の正確な翻訳のためには、前文以前の会話状況の反映が必要である。しかし、全評価文に対してこれらの比率は非常に低いため、日韓翻訳に関しては日英翻訳などより文脈処理を行なう必要性は低いと結論づけることができる。

2.2 構造評価との関連

今回の中間時評価では、翻訳品質評価と同一の文に対して、その原言語構造が正しいかどうかの評価も行なった。その結果、および翻訳品質評価との関係を以下に示す。

表 2.5: 構造評価との関連

	正しい構造	誤った構造	(未出力)	正しい構造の割合
Rank-A	1008	60	0	94.4%
Rank-B	389	133	0	74.5%
Rank-C	102	67	0	60.4%
Rank-D	31	12	0	72.1%
未出力	16	6	670	2.3%
計	1546	278	670	62.0%

この表より、構造を誤った139文(全評価文の11.1%)のうちの21.6%(両評価者で計60文)がRank-A、また93.5%(同260文)がRank-C以上であった。原言語構造が誤ったまま変換処理を行なった場合、翻訳結果の質に決定的な影響を与えずRank-Dが少ないのは、日本語と韓国語の構造が類似しているためである。

また同表より、何らかの構造を出力したにもかかわらず翻訳結果を出さない文が11文存在することが明らかになった。これは生成処理のバグであると考えられ、解決は容易である。

2.3 翻訳処理時間

図 2.1に、入力形態素数と翻訳時間の関係を示す。

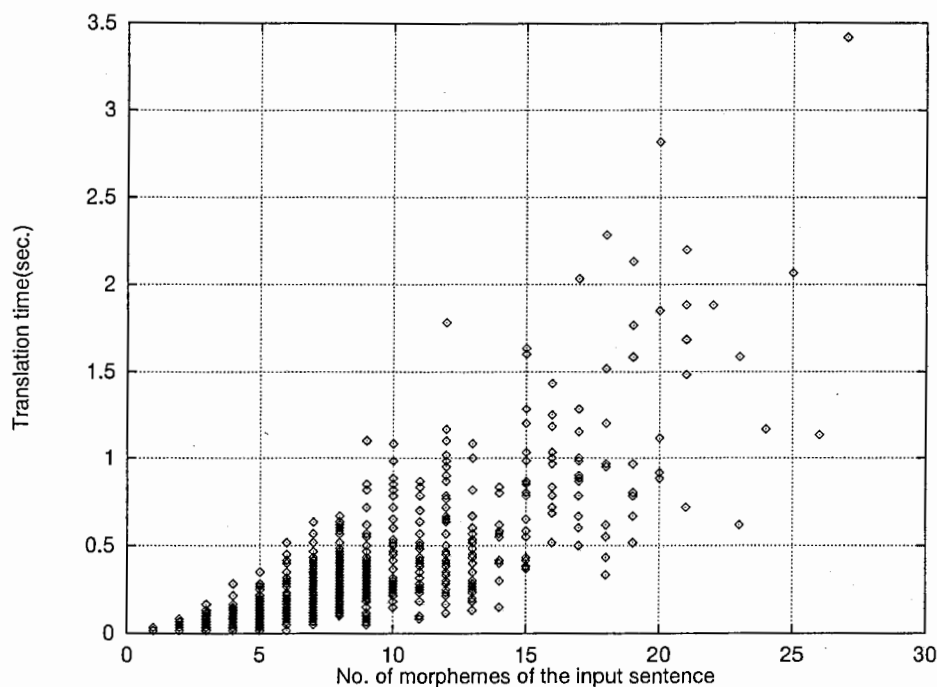


図 2.1: 入力形態素数と翻訳時間の関係

計測の結果、1文が9形態素の文の平均処理時間は0.35秒、10形態素では0.42秒であった。本評価実験の平均形態素数が、9.46(形態素/文)であることを考えると、平均処理時間はほぼ0.4秒程度と考えてよい。ただし、評価における入力タグデータであり、形態素解析に要する時間は含んでいない。

この図より、形態素数の増加に伴い翻訳時間が増加することがわかるが、両者の関係が線形であるかどうかは確認できない。同一の形態素数においても翻訳時間にかなりの幅が見られることから、形態素数以外の要因も翻訳時間に影響していることがこの図から推測される。

2.4 誤りの分析

表 2.6に、Rank-B、Rank-C、Rank-D の評価文に対して評価者が指摘した原因を集計したものを示す。

表 2.6: 誤りの原因分類

	語選択	活用	構文	語用	情報量	限定詞
評価者 #1	157	117	16	183	59	6
評価者 #2	274	105	28	242	2	0
Rank-B(522 文)	254	80	6	353	38	3
Rank-C(169 文)	137	123	31	57	14	2
Rank-D(43 文)	40	19	7	15	9	1
計	431	222	44	425	61	6

評価者との関係では、「情報量」の項目以外は大きな差はない。ただし、全体的には評価者 #2 のほうがやや多く指摘する傾向にある。「情報量」の項目は、指摘する基準の解釈が評価者によって異なっている差が大きい。なお、「限定詞」の項目は、韓国語に相当する現象がほぼないと考えられたため、ほとんど指摘されなかった。

また表 2.6より、「構文」誤りが他の誤りに比較して少ないことがわかる。また「構文」誤りは日英/日独翻訳の評価結果³と比較しても少ないことから、この傾向は韓国語翻訳の一つの特徴であると考えられる。

一方、最も多い「語選択」は Rank-D のほとんど(93%)で指摘されていることから、最も重要な問題であると考えられる。総数では「語用」の誤りも「語選択」と同様に多いが、軽微な誤り(Rank-Bにより集中している)なので、さほど深刻ではない。

³詳細はそれぞれの Technical Report[Mim96][Pau96]を参照のこと。

第 3 章

韓日翻訳

本章では韓日翻訳の評価結果について述べる。評価を行なった時点での訓練文数はのべ 506 文、異なり 493 文であり、変換パターン規則数は 311、変換パターン用例数は 1701 である。

3.1 概要

表 3.1 に韓日翻訳における翻訳品質評価結果を示す。

表 3.1: 韓日翻訳評価: 概要

	#1	#2	割合	文長	意味距離
Rank-A	344	344	29.4%	4.97	0.489
Rank-B	123	188	13.3%	6.44	1.052
Rank-C	164	72	10.1%	8.82	1.836
Rank-D	62	89	6.5%	8.42	2.004
未出力	476	476	40.7%	12.16	—
計	1169	1169	100%	8.70	0.599

表 3.1 から、出力率、翻訳精度、翻訳率を算出すると、以下のようになる。

- 出力率:59.3%
- 翻訳精度:89.1%
- 翻訳率 (= 出力率 × 翻訳精度):52.8%

日韓翻訳と比較して、変換パターン数、用例数共に約半分、もしくはそれ以下であることを考慮すると、出力率が低いのはやむを得ないと考えられる。今後、パターン/用例を蓄積していくことによって出力率は向上することが予想できる。

翻訳精度は、前述したように学習が不十分であるにもかかわらず高い値となった。これは、韓国語が日本語と構文の面で非常に類似しているためと考えられる。

文長と Rank との関係では、未出力文に対する平均文長が Rank-C や Rank-D と比較して極端に長いことがわかる。日韓翻訳でも同様に未出力文の平均文長が最も長い、その差は小さい。このことから、日韓翻訳と比較して韓日翻訳では学習量の少なさに起因する長文に対する弱さをはっきりと傾向に出ている。

表 3.2: 韓日翻訳での評価者間の相関

#1\#2	Rank-A	Rank-B	Rank-C	Rank-D	計
Rank-A	310	26	1	7	344
Rank-B	25	76	11	11	123
Rank-C	9	83	45	27	164
Rank-D	0	3	15	44	62
計	344	188	72	89	693

表 3.2 に、評価者間の評価の相関を示す。この表より、両評価者が一致した評価をした割合が 475 文 (未出力を除く評価文の 68.5%)、評価が一段階異なるものも含めると 662 文 (未出力を除く評価文の 95.5%) であり、日韓翻訳の評価者と同様に、韓日翻訳でも評価者の個人差はあまりないと考えることができる。

次に、表 3.3 に、評価文のうち、意味距離が 0 である文の評価結果を示す。

表 3.3: 意味距離 0 の文の評価

	#1	#2	計
Rank-A	141	135	276
Rank-B	36	35	71
Rank-C	1	1	2
Rank-D	0	7	7
計	178	178	356

韓日翻訳では、意味距離が 0 で Rank-C または Rank-D である文が合計のべ 9 文あった。それらを表 3.4 に示す。以下の表からわかるように、のべでは韓日のほうが多いが異なりは日韓、韓日共に 4 文で同じである。また、実際は「こちらでございます」「こちらです」で三分の二 (6 文) を示すことから、Rank-D になった原因を日韓翻訳の時よりも絞りこむことができ、対策は立てやすい¹。なお、「わかりそう

¹この文の直訳は「こと(に)あります」である。韓国語では買い物客が店員にお金を手渡す時、および店員が買い物客に品物を手渡す時にこの表現を使う。

ます」は日本語生成における「です/ます」の選択誤り、「乗られます」は訳語選択の誤りと平叙文/命令文の選択誤りが重なったものである²。

表 3.4: 意味距離 0 で Rank-C または Rank-D の文

韓国語入力	日本語翻訳結果	(模範訳)
여기 있습니다.(頻度 5)	こちらでございます。	(どうぞ。)
여기 있어요.	こちらです。	(どうぞ。)
알 수 있을 것 같습니다.	分かりそうです。	分かりそうです。
타세요.	乗られます。	お滑りください。

3.2 構造評価との関連

表 3.5 に、構造評価結果と、翻訳品質評価の関係を示す。

表 3.5: 構造評価との関連

評価	正しい構造	誤った構造	(未出力)	正しい構造の割合
Rank-A	638	50	0	92.7%
Rank-B	248	63	0	79.7%
Rank-C	152	84	0	64.4%
Rank-D	86	65	0	57.0%
未出力	0	0	952	0%
計	1124	262	952	48.1%

この表より、全文に対して正しい原言語構造を出す割合が、日韓翻訳よりも低いことがわかる(日韓は 62%、韓日は 48%)。未出力文も誤った構造とみなすと、半分以上の文が誤った構造に解析していることになる。

また Rank との関係でも、構造が正しいにもかかわらず Rank-D となる率も韓日の方が高いことがわかる。日韓の構造誤りは Rank-B を最頻値とした評価なのに対し、韓日は Rank-C の値が最大である。このことから、韓日の方が深刻な誤りの可能性があり、学習量の増加、日本語生成の整備が必要である。

² 「滑って下さい」が該当文脈において最も近い直訳である。韓国語では「乗る」と「滑る」は同一の語(타다)によって表記する。

3.3 翻訳処理時間

図 3.1に、入力形態素数と翻訳時間の関係を示す。

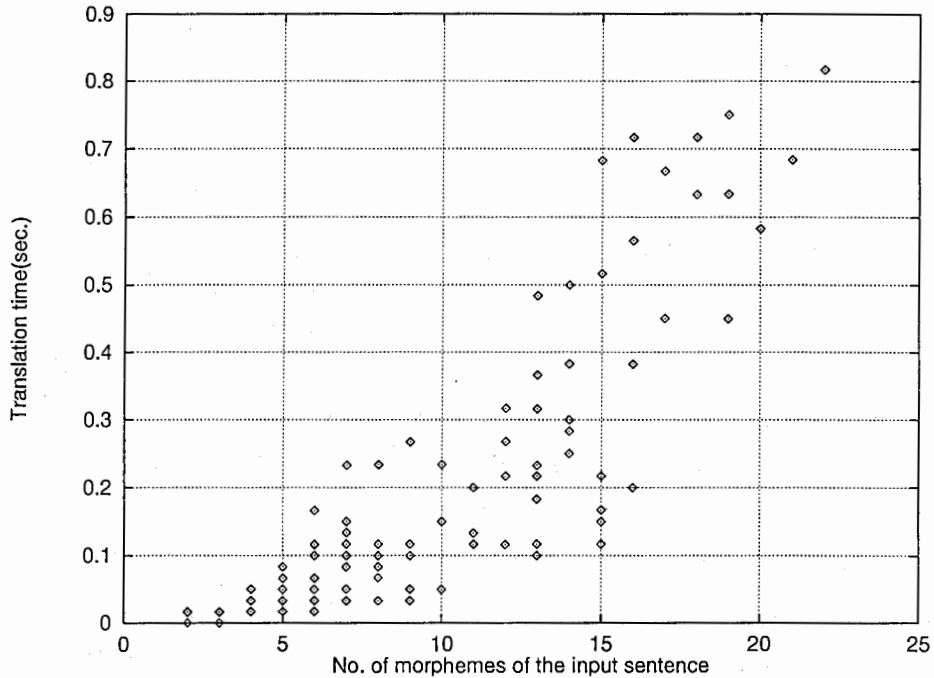


図 3.1: 入力形態素数と翻訳時間の関係

計測の結果、8形態素の入力文の翻訳処理時間の平均が 0.098 秒、9 形態素で 0.100 秒であった。韓日翻訳の評価に使用した全文の平均文長が 8.70 形態素であることを考慮すると、韓日翻訳の平均的な翻訳処理時間は約 0.1 秒であるとみなせる。

韓日翻訳の場合、日韓翻訳よりデータ数が少ないが、翻訳時間はすべての形態素数にわたって明らかに速い結果となっている。韓日翻訳においては、日韓翻訳よりも学習された変換ルール用例の数が大幅に少ないことが翻訳時間の差に影響していると考えられる。

3.4 誤りの分析

表 3.6: 誤りの原因分類

	語選択	活用	構文	語用	情報量	限定詞
評価者 #1	271	18	31	209	146	1
評価者 #2	196	59	3	157	55	0
Rank-B(311 文)	134	25	2	97	148	0
Rank-C(236 文)	197	26	17	201	21	1
Rank-D(151 文)	136	26	15	68	32	0
計	467	77	34	366	201	1

表より、韓日でも日韓と同様に(出力結果の)「構文」誤りは少ないことがわかる。Rank-D の多く(90%)で指摘されていることから、最も対策が必要なのは日韓翻訳と同様、「語選択」であると考えられる。

また、韓日翻訳の場合、日韓の原因分布と異なるのは「語用」の誤りが比較的深刻なことである。これは日韓翻訳で「語用」に問題がある文の多くが Rank-B であったのに対し、韓日翻訳では「語用」の多くが Rank-C であることからわかる。

3.5 Rank-D の分類

韓日における Rank-D の総数は、のべ 151 文、異なり 107 文である。以下ではこの原因を分類する。なお、以下に示す例はすべて、評価者のどちらか、あるいは両方が Rank-D とした文である。また、文数は異なりで示す。

1. 日本語生成の失敗：23 文

「早くご覧ください。」(빨리 봐 주세요.)

「御飯を出るのないますか。」(밥 나오는 것 없어요?)

「お客様がこちらに荷物を上げておかれてください。」

(손님 여기 짐을 올려 놓으십시오.)

2. 訳語選択の失敗：59 文

(a) 「なる」に関する問題：13 文

「お客様のお名前は どうやって なる のでしょうか。」

(손님 성함은 어떻게 되시죠?)

「酒二瓶まで なります。」(술 두 병까지 됩니다.)

- (b) 「分」と「人」³: 8文
 「何分ました。」(몇 분이세요?)
 「一分あたり四万ウォンです。」(한 분당 사만 원입니다.)
- (c) 「される」に関する問題: 5文
 「どのようなことにされますか。」(어떤 걸로 하시겠습니까?)
 「はい今のホテルの中にあるレストランで食事されるのです。」
 (네 지금 호텔 안에 있는 레스토랑에서 식사하실 수 있습니다.)
- (d) 「使う」と「書く」⁴: 4文
 「はい書かれうるのです。」(네 쓰실 수 있습니다.)
 「はい考えより早く着いてそうするのですが今部屋書けましようか。」
 (네 생각보다 빨리 도착해서 그런데 지금 방 쓸 수 있을까요?)
- (e) その他: 29文
 「英語ではケイ子供のエムワイオーユー円以来のエッチイーイーです。」⁵
 (영어로는 케이 아이 엠 와이 오 유 엔 지 에이치 이 이.)
 「青い色の表紙をなっている水道の蛇口をロックしておきますし赤いのだ
 けつけられましたか。」⁶
 (파란 색 표지가 되어 있는 수도꼭지를 잠가 놓고 빨간 것만 트셨습니까?)
3. 平叙、疑問、命令文の選択失敗⁷: 5文
 「私どものホテルの宿泊客ました。」(저희 호텔 투숙객이세요?)
 「人參茶だけ召しあがります。」(인삼차만 드세요.)
4. 一見正しい翻訳結果(要文脈処理): 13文
 「こちらでございます。」(여기 있습니다.)
 「売って行ってください。」⁸(많이 파세요.)

³韓国語では同一の語(분)で表記する。

⁴韓国語では同一の語(쓰다)で表記する。

⁵正しくは「英語ではケイアイエムワイオーユーエヌジーエッチイーイー。」

⁶正しくは「青色の表示の蛇口を縮めておいて赤いだけまわされましたか。」

⁷韓国語の非格式体では叙述形、疑問形、命令形、勧誘形が同一形態になる。文末の記号(?)はTDMTの入力には含まれないので、このような選択誤りが起こる。

⁸直訳は「たくさん売ってください」である。この表現は、買い物客が商品購入後、分かれ際に店員に対して使われる。

5. 分類不能、原因不明など：7文

「あげられますか。」(한국내 연락처 좀 주시겠습니까?)

「これは新たに出るのありませんし大分できたモデル人ですがなぜこっち高いですか。」⁹(이건 새로 나온 것도 아니고 오래 된 모델인데 왜 이리 비싸요?)

⁹正しくは「これは新しく出たわけでもなく古いモデルなのになぜこらも高いんですか。」

第 4 章

まとめ

変換主導型翻訳 (TDMT) における日韓 / 韓日翻訳の中間時評価結果について述べた。本報告書の内容をまとめると以下のようになる。

- 日韓は翻訳率 71%、出力さえすれば概ね高品質の翻訳結果を出す。最大の課題は入力パターンを効率的に学習して未出力を減らすことであり、訳語選択も重要な課題である。
- 韓日は翻訳率 53%、4 割以上が未出力の状況は日韓よりも深刻である。
- 日本語 / 韓国語共に、生成部の整理 / 小改良で出力文の精度向上は可能。
- (前文以前を参照する) 文脈処理は韓国語についてはあまり大きな問題ではない。一文内だけで解決できる処理がまだ多く残っている。

今回の中間時評価の主な諸元と評価結果を表 4.1 にまとめる。

表 4.1: 日韓 / 韓日翻訳: まとめ

	日韓翻訳	韓日翻訳
訓練文数 (のべ)	1041	506
訓練文数 (異なり)	943	493
変換パターン規則数	680	311
変換パターン用例数	3605	1701
出力率	72.3%	59.3%
翻訳精度	97.6%	89.1%
翻訳率	70.5%	52.8%

日韓翻訳は言語の類似性により日英翻訳や日独翻訳よりも容易であると直観的に考えられてきた。また韓日翻訳についても英日翻訳などよりも容易であると言われてきた。今回の TDMT 中間時評価はそれを定量的なデータで裏付けたことになる。

しかしながら、韓国語処理の容易性はあくまでも相対的なものであって、日韓翻訳、韓日翻訳に課題がないことを意味するわけではない。現在の日韓/韓日 TDMT の最大の問題点は未出力文、つまり翻訳結果を出力しない文がまだ非常に多い点である。これに関しては、特に韓日翻訳が用例蓄積が少ないのでその意味では出力率が低いのは当然の結果であり、学習をしていくに伴って出力率は増加することは予想できる。日韓翻訳、韓日翻訳共に、用例蓄積に際してさらに効率的な変換パタンの学習が望まれる。

参考文献

- [Mim96] 美馬秀樹, 古瀬蔵: 日英間変換主導翻訳の中間時評価, Technical Report TR-IT-0189, ATR 音声翻訳通信研究所 (1996).
- [Pau96] PAUL, M., 傅康晴, 古瀬蔵: 日独変換主導翻訳の中間時評価, Technical Report TR-IT-0191, ATR 音声翻訳通信研究所 (1996).
- [Yam96a] 山本和英, 隅田英一郎: TDMT 用形態素解析プログラム (日本語 / 英語 / 韓国語) — 性能評価報告 —, Technical Report TR-IT-0192, ATR 音声翻訳通信研究所 (1996).
- [Yam96b] 山本和英, 古瀬蔵, 飯田仁: 用例に基づく日韓の対話翻訳処理機構, 全国大会講演論文集, 第 53 回, 4L-10, pp. 2/71-72, 情報処理学会 (1996).

(付録) Rank-D 文一覧

以下に、今回の評価で二名の評価者のうちどちらかが Rank-D と判断した文の一覧を示す。TDMT の翻訳性能改善のためには、定量的な評価と同時に個々の出力、特に誤翻訳の出力を分析、検討する必要がある。このため、以下に挙げる一覧は今後の翻訳性能の向上のための資料として価値が高いと判断し、ここに全文を掲載する。

様式は以下の通り：

入力: < 入力文 > (会話 ID) 出力: < 翻訳結果 > (意味距離) < 構造評価結果 > / 評価者 #1 / 評価者 #2
--

ここで、構造評価結果は { 正, 誤 } のいずれか、評価者による評価は

{Rank- A/B/C/D}[語選択 - 活用 - 構文 - 語用 - 情報量 - 限定詞]

であり、[...] 内の問題点は評価者が指摘した場合に 1、そうでない場合に 0 としている。

例：

入力: はいそのとおりでございます (AT130082) 出力: 예 그 길.(1.00) 正 /D[100000]/C[111000]
--

会話 ID:AT130082 中の一入力文「はいそのとおりでございます」の翻訳結果は「예 그 길.」、翻訳の際に原言語構造も正しく構成され、意味距離は 1.00 であった。この翻訳結果に対して評価者 #1 は Rank-D と、評価者 #2 は Rank-C と評価した。また、翻訳文の問題点として、評価者 #1 は「語選択」を、評価者 #2 は「語選択」「活用」「構文」の 3 点を指摘した。

付録 A: 日韓翻訳

入力: はいそのとおりでございます (AT130082)

出力: 예 그 길.(1.00)

正 /D[100000]/C[111000]

入力: そうですね (BT320011)

出力: 그렇죠?(0.00)

正 /D[100000]/B[010100]

入力: なるほど結構たくさんあるんですね (BT320011)

出力: 어 괜찮히 많이 있군요.(1.67)

誤 /D[100000]/C[111100]

入力: そのダブルベッドが二つある部屋ということですね (BT320011)

出力: 그 더블베드가 두 개 있는 방으로하고요?(0.67)

正 /D[110100]/C[110100]

入力: いかがなさいましたか (CT130051)

出力: 어때요했어요?(1.78)

正 /D[100000]/B[110000]

入力: おみやげ用の表のことでございますね (CT130051)

出力: 선물용 앞면 말이시군요.(2.00)

誤 /D[100000]/C[110100]

入力: 予備の毛布でしたらお部屋のクローゼットに置いてあると思いますがお探しいただけた
でしょうか (CT330072)

出力: 예비가 모포면 방의 옷장에 놓아 있다고 생각하지만 찾아 주신 것이었습니까?(4.06)

誤 /D[110100]/C[110000]

入力: 御一緒にお持ちいたします (CT330072)

出力: 같이 있어 드립니다.(0.67)

正 /D[110000]/D[111000]

入力: それでしたら一日当たりが一万千九百円となります (KT330011)

出力: 그러면 하루 당침이 만천구백 엔으로 됩니다.(1.13)

正 /D[110000]/B[100000]

入力: 七月四日あたりで開いている部屋があるかどうかお聞きしたいんですが (OT320031)

出力: 칠 월 사 일당에서 여는 방이 있는지 듣고 싶은데요.(3.42)

誤 /D[100100]/D[111100]

入力: 湯涌温泉の中の純和風旅館で白雲楼というのがお勧めかと思えます (OT330021)

出力: 유와쿠 온천의 중 준 일본식 여관에서 하쿠운로라는 것이 권을 것 같습니다.(3.78)

誤 /D[110110]/C[111100]

入力: お客様金沢のことはよくご存じですか (OT330021)

出力: 손님 가나자와 말은 잘 아십니까?(0.83)

誤 /D[100000]/C[110100]

入力: 山鉾巡行というのはアメリカのパレードのようなもので山鉾がたくさんの人に引かれて
街に出ます (OT330071)

出力: 꽃수레 순행이는 것은 미국 행진 같은 것에 꽃수레가 많이 사람에게 빼시고 가에 나갑니다.(5.87)

誤 /D[110100]/C[110000]

入力: 二十九日の方も同じ部屋でお取りできます (TAS13006)

出力: 이십구 일에 있는도 같은 방으로 찾되겠습니다.(2.42)

正 /D[111000]/B[100000]

入力: 当日はなにでこちらへいらっしゃいますか (TAS13016)

出力: 당일은 무엇서 이쪽으로 계십니까?(1.09)

正 /D[100100]/C[111000]

入力: できたりんごか何かをむいて持って来てほしいんですけど (TCC22073)

出力: 될 수 있으면 사과 꺾을 향해서 갖고 원하는데요.(2.98)

正 /D[100010]/C[110000]

入力: これからホテル内のフラワーショップのほうに電話をしてみますので折り返しという形になるかと思いますが (TCC23054)

出力: 이것부터 호텔내 꽃가게의 쪽에 전화를 해 봐 드릴 테니까 다시라는 모양이 될 거는데요.(4.97)

誤 /D[110100]/C[111000]

入力: ではそのバックをお願いしたいと思います (TCS12023)

出力: 그럼 그 단체관광을 부탁하고 싶다고 생각합니다.(1.93)

正 /D[100100]/D[100010]

入力: いかがなさいましたか (TCS13005)

出力: 어때요했어요?(1.78)

正 /D[110100]/B[100100]

入力: いらっしゃる前に一応電話をなさってください (TCS13005)

出力: 계시는 앞으로 일단 전화를 하셔 주십시오.(2.72)

正 /D[100100]/C[110100]

入力: お手数をおかけします (TCS32012)

出力: 수수를 걸어 드립니다.(1.78)

正 /D[000100]/D[111000]

入力: で懐石などのお食事は御不要でございますか (TCS33035)

出力: 그래서 그러세요.(2.17)

正 /D[100010]/D[000010]

入力: はい窓際のお席は承れます (TCS33050)

出力: 예 창가의 자리는 삼가드립니다.(0.90)

正 /D[100000]/C[110000]

入力: お忘れ物を横浜のホテルでされたということなんですけどどのようなものなんでしょうか (TCS33061)

出力: 분실물을 요코하마의 호텔에서 소매치기 당했다는 것인데 어떤 것인가요?(3.31)

正 /D[100010]/B[110000]

入力: 宅配便はお一つですよ (TCS33061)

出力: 택배편은 하나죠?(1.08)

正 /D[100000]/B[000100]

入力: 実は支店長が転勤で帰国しますので心に残るパーティーにしようと思っています (TDS32005)

出力: 실은 지점장이 전근에서 귀국할 테니까 마음을 남는 파티로 하려고 하다고 생각하고 있습니다.(4.36)

誤 /D[110100]/B[100100]

入力: ですからお話がはずむことと存じます (TDS33008)

出力: 그래서 말씀이 들뜨는 로 압니다.(1.81)

正 /D[100000]/B[100000]

入力: 前の便ですか (TGS13002)

出力: 앞 편하십니까?(0.68)

正 /B[010000]/D[110000]

入力: ええそういうことです (THS33003)

出力: 예 그런 이죠.(0.60)

正 /D[000010]/B[100100]

入力: その時何かいるものはありますか (TKS32002)

出力: 그 시간 무엇입니까? 있는 것은 있어요?(1.65)

正 /D[101110]/C[110000]

入力: それから乗り捨てはどこの営業所でもできるんですか (TKS32002)

出力: 그리고 내림은 어디 영업소에서도 되나요?(1.99)

正 /D[100000]/B[100100]

入力: はい承りますフィリップス様 (TKS33003)

出力: 예 받는 필립스씨.(0.83)

誤 /C[010000]/D[111000]

入力: はいお取りいたしますのでしばらくお待ちください (TOS23001)

出力: 예 찾아 드릴 테니까 잠깐 기다려 주십시오.(0.31)

正 /D[100000]/C[110100]

入力: じゃあついでに聞きますがゴルフ道具のレンタルとかもできるんでしょうか (TOS32009)

出力: 그럼 김에 들겠는데요 골프 도구 렌탈이라든가도 될까요?(4.45)

正 /D[110010]/B[100000]

入力: 奈良へは以前にいらっしゃったことはありますか (TOS33005)

出力: 나라에는 이전에 계시는 일은 있어요?(1.34)

正 /D[100000]/B[100100]

入力: いいえそのご予算でも松花堂というすてきなお弁当を召し上がっていただけますよ (TSS33001)

出力: 아뇨 그 예산이나 쇼카이도라는 멋있단 도시락을 들어 주시겠는데요.(4.00)

正 /D[110100]/C[111000]

入力: できればもう一品なにか追加で御注文くださいませ (TTS33002)

出力: 될 수 있으면 다시 일 산품 뭔가 추천데요 주문이여 주십시오.(4.10)

誤 /D[111000]/C[110000]

入力: 三つのお肉を一つのお皿の方に盛り合わせにいたしましてお部屋の方に順にお持ちいたします (TTS33002)

出力: 세반의 고기를 하나 접시로 모리아와세로 하고 방에 순서에 있어 드립니다.(4.53)

正 /D[110001]/C[111000]

付録 B: 韓日翻訳

入力:그 요금은 어떻게 되죠? (AF350031)
出力: その料金はどうやってなるのでしょうか。 (1.74)
正 /C[000100]/D[100110]

入力:여기 있습니다. (AF350031)
出力:こちらでございます。 (0.00)
正 /A[000000]/D[000010]

入力:알아보고 오겠습니다. (AF440012)
出力:しらべてみますし来ます。 (1.19)
正 /D[110100]/C[010100]

入力:네 지금 호텔 안에 있는 레스토랑에서 식사하실 수 있습니다. (AF440012)
出力: はい今のホテルの中にあるレストランで食事されるのです。 (2.81)
誤 /D[101100]/B[000100]

入力:네 생각보다 빨리 도착해서 그런데 지금 방 쓸 수 있을까요? (AF450011)
出力: はい考えより早く着いてそうするのですが今部屋書けましょうか。 (3.93)
誤 /D[100100]/D[100110]

入力:그럼 짐만 좀 맡아 주시겠어요? (AF450011)
出力: では荷物だけ預て頂けますか。 (1.63)
正 /D[100010]/B[010000]

入力:네 여기 있습니다. (AF450011)
出力: はいこちらでございます。 (0.00)
正 /B[000010]/D[100010]

入力:여기 있어요. (AF450012)
出力:こちらです。 (0.00)
正 /B[000010]/D[100010]

入力:그럼 지금 방으로 못 들어가요? (AF450012)
出力: では今部屋にません入りますか。 (1.62)
誤 /D[001100]/D[011100]

入力:말씀 좀 여쭙어 보겠습니다. (AF450013)
出力: お話ちょっとお聞きしてみます。 (2.54)
正 /D[100010]/C[000110]

入力:손님 성함은 어떻게 되시죠? (AT340043)
出力: お客様のお名前は どうやってなれるのでしょうか。 (1.38)
正 /C[100100]/D[100110]

入力:걱정 마시고 오십시오. (AT340043)
出力: 心配語ですし来られてください。 (2.29)
誤 /D[100100]/D[110100]

入力:그렇게 하십시오. (CF340053)
出力: そのようにされてください。 (1.06)
正 /C[100100]/D[000100]

入力:파란 색 표지가 되어 있는 수도꼭지를 잠가 놓고 빨간 것만 트셨습니까? (CT340042)
出力: 青い色の表紙をなっている水道の蛇口をロックしておきますし赤いのだけつけられましたか。 (7.60)
誤 /D[100100]/D[110100]

入力:손님 계란은 계란 프라이하고 스크램블하고 오무렛이 있습니다. (CT340062)

出力: お客様の卵は目玉焼きとスクランブルとオムレツはあります。(2.50)
誤 /D[101100]/B[100000]

入力: 어떤 걸로 하시겠습니까? (CT340062)
出力: どのようなことにされますか。(1.00)
正 /D[100100]/C[100100]

入力: 손님 주스는 어떤 걸로 하시겠습니까? (CT340062)
出力: お客様のジュースはどのようなことにされますか。(2.17)
誤 /D[101100]/C[100100]

入力: 네 준비되어 있습니다. (CT340062)
出力: はい手配れています。(2.21)
正 /D[100100]/C[110100]

入力: 지금 곧 올라가겠습니다. (CT340072)
出力: 今すぐに登ります。(1.51)
正 /B[100010]/D[100000]

入力: 전화를 끊고 기다리십시오. (CT340082)
出力: 電話を買って待たれてください。(2.48)
正 /D[100100]/C[100110]

入力: 연결됐습니다. (CT340082)
出力: 連絡できておりました。(1.43)
正 /B[100010]/D[100000]

入力: 일본 연결됐습니다. (CT340083)
出力: 日本の連絡できておりました。(1.43)
誤 /C[101010]/D[100000]

入力: 네 다른 건 됐어요. (CT350043)
出力: はいほかのものはなりました。(0.82)
正 /C[100100]/D[100100]

入力: 밥 나오는 것 없어요? (CT450013)
出力: 御飯を出るのないますか。(2.86)
誤 /C[100100]/D[110000]

入力: 그럼 됐어요. (CT450013)
出力: ではなりました。(0.40)
正 /D[100100]/D[100000]

入力: 대한 항공 분이에요? (FF450012)
出力: 大韓航空分ですか。(1.17)
正 /C[100100]/D[100000]

入力: 제 짐이 안 나오는 거 있죠. (FF450012)
出力: 私に荷物は出ませんのあるです。(2.38)
誤 /D[110100]/C[101100]

入力: 영어로는 케이 아이 엠 와이 오 유 엔 지 에이치 이 이. (FF450012)
出力: 英語ではケイ子供のエムワイオーユー円以来のエッチイーイーです。(5.86)
誤 /D[101100]/D[100010]

入力: 빨리 봐 주세요. (FF450012)
出力: 早くご覧てください。(0.99)
正 /D[110100]/D[100000]

入力: 손님 여기 짐을 올려 놓으십시오. (GF340022)

出力: お客様がこちらに荷物を上げておかしてください。(3.68)
誤 /D[111100]/C[100000]

入力:됐습니다. (GF340023)
出力: になりました。(0.40)
正 /B[100010]/D[100000]

入力:한국내 연락처 좀 주시겠습니까? (GT340011)
出力: あげられますか。(2.55)
誤 /D[101010]/D[000010]

入力:대구는 옛날 역이구 동대구는 새로 생긴 큰 역입니다. (HF340011)
出力: テグは昔駅でトンテグは新たにできる大きい駅です。(2.51)
誤 /D[110100]/C[110100]

入力:각각 얼마나 됩니까? (HF350011)
出力: それぞれどのくらい良いのですか。(0.97)
正 /D[100010]/D[100000]

入力:여기 있습니다. (HF350011)
出力: こちらでございます。(0.00)
正 /B[000010]/D[100000]

入力:이걸로 잡아 드릴까요? (HF440012)
出力: これに捕まえらせて頂きましょうか。(2.03)
正 /D[110100]/D[110000]

入力:가격 차이가 좀 나요. (HF440012)
出力: 価格違いにちょっと出ます。(2.13)
正 /D[101100]/D[000100]

入力:네 다음 분. (HF440013)
出力: はい次の分です。(0.50)
正 /D[100010]/D[100010]

入力:예 잠깐 보겠습니다. (HF440013)
出力: はいしばらくご覧します。(1.38)
正 /D[100100]/D[100010]

入力:네 전부 일행은 몇 분이세요? (IF440013)
出力: はい全部の一行は何分しました。(1.67)
誤 /D[101100]/D[110000]

入力:그러니까 여권을 가지고 열 시하고 두 시에 가면 아니죠? (OF350021)
出力: ですからパスポートを持って十時と二時に行けばありませんですね。(3.73)
誤 /D[100100]/D[100000]

入力:이 호텔에서 얼마 정도 걸립니까? (OF350023)
出力: このホテルでいくらかかりますか。(1.17)
正 /B[100010]/D[100000]

入力:마지막으로 한 가지만 더 묻겠어요. (OF350023)
出力: 最後に一種類だけもう聞きます。(3.32)
誤 /C[101100]/D[100100]

入力:여기 안내소죠? (OF450012)
出力: じゃ案内所ですね。(1.22)
正 /D[100010]/D[100000]

入力:배는 다시 선착장으로 돌아가나요? (OF450012)

出力:おなかはまだ船着き場に帰りますか。(2.69)
正 /D[100100]/D[100100]

入力:그리고 지금 같으면 몇 시 차 탈 수 있어요? (OF450012)
出力:それから今同じたら何時の車乗れますか。(3.46)
誤 /D[100100]/C[100100]

入力:한 분당 사만 원입니다. (OT340011)
出力:一分あたり四万ウォンです。(1.50)
誤 /D[100010]/D[100000]

入力:네 분이죠? (OT340011)
出力:四分ですね。(0.44)
正 /D[100010]/D[100000]

入力:한 분당 사만 원입니다. (OT340012)
出力:一分あたり四万ウォンです。(1.50)
誤 /D[100010]/D[100000]

入力:한 분만 말씀하세요. (OT340012)
出力:一分だけおっしゃってください。(0.89)
正 /D[100010]/D[100000]

入力:조선 호텔로 모시려 가겠습니다. (OT340012)
出力:チョソンホテルにご案内致しに行きます。(0.72)
正 /D[100100]/C[100000]

入力:몇 분이세요? (PF440012)
出力:何分ました。(0.50)
誤 /D[100100]/D[110010]

入力:들어가서 구경만 하실 거면 입장권이 있구요. (PF440013)
出力:入って見物だけされると思えば入場券はあるです。(3.28)
誤 /D[100100]/C[110100]

入力:사이즈가 어떻게 되시죠? (PF440022)
出力:サイズにどうやってなられるのでしょうか。(0.88)
正 /C[100100]/D[110000]

入力:네 됩니다. (PF440022)
出力:はいなります。(0.00)
正 /D[100010]/D[100000]

入力:하루 타시겠습니까? (PF440022)
出力:一日乗られますか。(1.56)
正 /B[100010]/D[100000]

入力:타세요. (PF440022)
出力:乗られます。(0.00)
誤 /B[100010]/D[100000]

入力:네 얼마나 쓰시겠습니까? (PF440023)
出力:はいいくらくらい書かれますか。(1.00)
正 /D[100010]/D[100000]

入力:양말을 두꺼운 걸 신으면 그렇게 잘 안 맞을 경우도 있습니다. (PF440023)
出力:靴下を厚いのを履けばそのようによく合う場合もありません。(4.61)
誤 /D[001100]/D[110100]

入力:네 잘 받았습니다. (PF440023)

出力: はいよく貰っておりました。(0.77)
正 /D[100010]/C[000100]

入力:가서 타 보세요. (PF440023)
出力: 行って乗ってみられます。(0.67)
正 /B[100010]/D[100000]

入力:여기 있습니다. (RF340013)
出力: こちらでございます。(0.00)
正 /A[000000]/D[100000]

入力:그럼 내일 오전 길로 두 장 주시겠습니까? (RF350012)
出力: では明日午前ことに二枚あげられますか。(2.20)
誤 /D[100100]/D[100100]

入力:지불은 뭐로 하시겠습니까? (SF440011)
出力: 支払は何にされますか。(0.42)
正 /A[000000]/D[100000]

入力:네 쓰실 수 있습니다. (SF440011)
出力: はい書かれうるのです。(0.73)
正 /D[100100]/D[100000]

入力:저희 호텔 투숙객이세요? (SF440011)
出力: 私どものホテルの宿泊客ました。(1.50)
誤 /D[100100]/C[110010]

入力:방 번호가 어떻게 되시죠? (SF440011)
出力: 部屋番号をどうやってなられるのでしょうか。(0.88)
正 /C[100100]/D[100000]

入力:네 물론 쓸 수 있습니다. (SF440012)
出力: はいもちろん書けるのです。(1.07)
正 /D[100100]/D[100000]

入力:카드 받았습시다. (SF440012)
出力: 카드貰っておりました。(1.21)
誤 /D[100010]/D[000100]

入力:네 부가 가치세 십 프로 포함되어 있습니다. (SF440013)
出力: はい付加価値税十人プロ含まれています。(1.82)
誤 /C[100100]/D[100000]

入力:처음이세요? (TF440011)
出力: 初めてました。(1.00)
誤 /D[100100]/D[100000]

入力:한방 약차인데요. (TF440011)
出力: 漢方の煎じ薬ですが。(1.29)
正 /A[000000]/D[100000]

入力:네 저희는 메뉴는 따로 없구요. (TF440013)
出力: はい私どもはメニューは別送ないのです。(2.57)
正 /D[100100]/D[100010]

入力:유명한 한국 인삼을 분말로 한 차로 건강에도 좋고 미용에도 아주 좋은 겁니다. (TF440013)
出力: 有名韓国人参を粉末で車で健康もいいですし美容もとてもいいのです。(9.65)
誤 /D[101100]/D[110000]

入力:인삼차 드실 때 커피 드시면 별로 안 좋습니다. (TF440013)

出力: 人参茶召しあがる時のコーヒー召しあがれば特にないいです。(5.47)
誤 /D[101100]/D[110000]

入力:인삼차만 드세요. (TF440013)
出力: 人参茶だけ召しあがります。(1.44)
正 /B[100010]/D[100000]

入力:친구 분도 인삼차 한번 들어 보세요. (TF440013)
出力: 友人の分も人参茶を一度入ってみてください。(3.41)
正 /D[100100]/D[100000]

入力:안 먹어 보고 어떻게 맛을 아세요? (TF440013)
出力: 食べてみませんでしてどうの味を知りました。(4.06)
誤 /D[101100]/D[110000]

入力:커피랑 먹어서 꼭 나쁜 건 없어요. (TF440013)
出力: コーヒーと食べてぜひ悪いのはありません。(3.16)
誤 /D[100100]/D[110000]

入力:커피는 안 드세요? (TF440013)
出力: コーヒーは召しあがりませんか。(1.50)
正 /C[100100]/D[000010]

入力:예 전부 나왔습니다. (UF340011)
出力: はい出ませんでした。(1.17)
正 /D[101100]/D[000010]

入力:귀중품은 없었습니까? (UF340011)
出力: 貴重品はなくていましたか。(1.33)
正 /C[100100]/D[010000]

入力:방금 이 버스에서 내린 사람인데요. (UF450011)
出力: 今このバスで降りた人ですが。(3.16)
正 /C[100100]/D[100100]

入力:그런데 없어요. (UF450012)
出力: とろろでないます。(0.61)
正 /C[100000]/D[100000]

入力:예 네모 난 거요. (UF450013)
出力: はい四角を出るのです。(0.98)
正 /C[100100]/D[100000]

入力:예 그리고 비행기 표도 들어 있어요. (UF450013)
出力: はいそれから飛行機切符もかかっています。(1.43)
正 /C[100100]/D[100000]

入力:아니 없어요. (UF450023)
出力: ないないます。(1.40)
誤 /C[100100]/D[110000]

入力:머리만 아프십니까? (VF340021)
出力: 頭だけ痛いれるですか。(2.26)
正 /C[100100]/D[010000]

入力:예 여기 있습니다. (VF340021)
出力: はいこちらでございます。(0.00)
正 /A[000000]/D[100100]

入力:그리고 나머지 두 알은 식후 삼십 분에 하나씩 드십시오. (VF340022)

出力:それから残り二たまでは食後三十分に一召しあがってください。(4.21)
誤 /C[100110]/D[100100]

入力:그냥 머리만 아프세요? (VF340023)
出力:そのままの頭だけ痛かったです。(2.67)
誤 /C[101100]/D[110000]

入力:여기 있습니다. (VF340023)
出力:こちらでございます。(0.00)
正 /A[000000]/D[000100]

入力:여기 있습니다. (VF350012)
出力:こちらでございます。(0.00)
正 /A[000000]/D[000100]

入力:인삼 제품은 많이 있는데요 뭘 원하세요? (VF440012)
出力:人參の製品はかなりあるのですが何を望みました。(3.85)
誤 /C[100100]/D[110000]

入力:술 두 병까지 됩니다. (VF440012)
出力:酒二瓶までなります。(1.50)
誤 /C[100100]/D[100000]

入力:네 출국장 나가서 바로 앞에 있어요. (VF440012)
出力:はい出国者名簿出てすぐ前にあります。(0.91)
誤 /D[100100]/D[100000]

入力:그런 일은 없어요. (VF440012)
出力:その日はありません。(2.11)
正 /D[100100]/D[110000]

入力:그럼 됐어요. (VF450012)
出力:ではなりました。(0.40)
正 /D[100100]/D[100000]

入力:많이 파세요. (VF450012)
出力:売って行ってください。(1.10)
正 /D[100010]/D[100010]

入力:그럼 그걸로 주시구요. (VF450021)
出力:ではそれにあげられるです。(1.26)
正 /D[100100]/D[100000]

入力:원이나 달러도 쓸 수 있습니까? (VF450021)
出力:ウォンくらいドルも利用できるのですか。(1.32)
誤 /D[100010]/D[100000]

入力:말씀 좀 여쭙어 보겠습니다. (VF450023)
出力:お話ちょっとお聞きしてみます。(2.54)
正 /D[100010]/C[000100]

入力:우리 집엔 코끼리가 없으니까 괜찮아요. (VF450023)
出力:私達の店には象がないのでいいです。(1.36)
正 /C[100100]/D[100000]

入力:이건 새로 나온 것도 아니고 오래 된 모델인데 왜 이리 비싸요? (VF450023)
出力:これは新たに出るものもありませんし大分できたモデル人ですがなぜこっちは高いですか。
(5.69)
誤 /D[100100]/D[110100]

入力:됐습니다. (WF340011)

出力:なりました。(0.40)

正 /C[100100]/D[000100]

入力:거기서 영수증도 발급 받을 수 있습니까? (WF450011)

出力:そこから領収証も発給を貰えますか。(3.02)

誤 /C[100100]/D[100100]

入力:제가 전무님 부인의 친구 되는 사람인데요. (ZT350011)

出力:私が専務さんの御夫人の友人なる人ですが。(5.29)

誤 /C[100100]/D[100100]