

2015年度 管理・情報専門委員会 成果報告

『中韓文献翻訳・検索システムに関する調査研究』



2016年3月10日

中国IPG 管理・情報専門委員会

中村崇諒 (コニカミノルタ(中国) 投資有限公司)

0. 委員会メンバー

以下11社、計18名（※途中参加、途中退会を含む）

IP FORWARD法律特許事務所
映橋知識産権諮詢（上海）有限公司
奥林巴斯（中国）有限公司
コニカミノルタ（中国）投資有限公司
ソニー（中国）有限公司
日本技術貿易株式会社
兄弟（中国）商業有限公司
パナソニックES社
パナソニックR&Dセンター（中国）有限公司
三菱電機（中国）有限公司
村田（中国）投資有限公司

宮川 幸子、本橋 たえ子、周 婷
高野 博成
蔡 博宇、陸 悅
中村 崇諒
石丸 幹朗、袁 改花
吳 礼
張 明慧、張 cong
宮 艷文
黃 善愛、任 之光
孟 翠敏
船木 由里、代 玲

本成果は、2014年度 特別テーマ調査委員会（中韓文献翻訳・検索システムに関する調査研究）を引き継いでおります。当該特別テーマ調査委員会には以下の皆様にもご協力を頂いております。

エプソン（中国）有限公司
キヤノン（中国）有限公司
兄弟（中国）商業有限公司
三菱化学控股管理（北京）有限公司

小林 利彦
浜田 祐一
奥田 聖二郎
杜 忻

1. 現状認識と課題

- 現在中国では2014年の特許出願件数が92.8万件（15年は110.2万件）に達し、実案も含めると膨大な数の専利文献が毎年出願されている。そのため、これらの専利文献を以下に効率よくクリアランスするかが問題となっているところである。そのような中、2015年1月5日に日本特許庁が「中韓文献 翻訳・検索システム」を無料公開した。
- 当該システムでは、「公報テキスト検索」と「公報番号索引照会」の2つが用意されているが、**翻訳精度と検索精度**がどの程度のものであるか多くの企業がその確認をとれていない。このような状況下、各社が本システムをバラバラに評価していれば時間・コストが膨大にされ、効率的ではないといった問題がある。

2. 調査目的

中韓文献翻訳・検索システムに関する調査研究について、前ページのとおり**翻訳精度と検索精度**の妥当性を評価することを本調査の目的とする。

The screenshot shows the search interface for patent publications. It includes fields for 'Publication Number' (公報番号), 'Search Project' (検索項目), 'Search Keyword' (検索キーワード), and 'Search Conditions' (検索条件). Below these are sections for 'Advanced Search' (近傍検索), 'NOT Search' (拡張NOT検索), and 'Search Result Summary' (検索結果総合式). At the bottom, there are buttons for 'Search' (検索), 'Screen Cleaning' (スクリーニング), 'Search Result List' (検索結果一覧), 'Reuse Search Conditions' (検索条件再利用), and 'Report Unmatched Words' (ヒットしないキーワードの報告).

項目番号	項目間接続	検索項目	検索キーワード	項目内接続
1	-	要約+請求の範囲	画像形成装置	OR
2	AND	出願人／権利者	柯尼卡美能达办公	OR
3	AND	国際特許分類		OR
4	AND	出願人／権利者		OR

検索結果

項目番号	ヒット件数	発行種別	文献種別	公知日	出願日
1	18	中国	韓国	From - To -	From - To -

図1 公報テキスト検索の画面

The screenshot shows the search interface for publication numbers. It includes fields for 'Publication Number' (公報番号) and 'Search Conditions' (検索条件). Below these are sections for 'Advanced Search' (近傍検索), 'NOT Search' (拡張NOT検索), and 'Search Result Summary' (検索結果総合式). At the bottom, there is a table for 'Input Format' (入力番号フォーマット) showing examples for China (CNA-nnnnnnnn, CNB-nnnnnnnn, CNC-nnnnnnnn, CNU-2nnnnnnnn, CNY-nnnnnnnn) and Korea (KRA-yyyynnnnnn, KRB1-10nnnnnnn, KRU-yyyynnnnnn, KRY1-20nnnnnnn).

公報番号	有無	発行国	公知日	出願日
CNA-103466357	有	中国	20131225	20130918

図2 公報番号索引照会の画面

【参考】日本特許庁 中韓文献翻訳・検索システム
URL : <http://www.ckgs.jpo.go.jp/>

3. 調査研究方法

(1) 翻訳精度

■評価定義 :

中国語ができる日本人が独立請求項の日本語翻訳を読んだ際に、発明の内容をどの程度理解できるか、A、B、Cの3段階で評価する。要約や説明書・図面の参照は可とする。

【A評価】軽微な誤訳がある程度で、B評価のものよりも正確に発明の内容を理解することができる。

【B評価】誤訳はあるが、単語がある程度正確に翻訳されており、大まかに発明の内容を理解することができる。

【C評価】誤訳が多く、発明の内容をほとんど理解できない。

■評価対象 : 独立請求項 1つ

■評価分野 : プロジェクター、プリンタ、ネットワーク、リチウム電池、情報処理、医療機器、エレベーター、複合機、自動車、光学ガラス、バイオ、照明、スピーカー、コンテンツサ、

■評価件数 : 70件 (5件/社)

■評価案件 : 特許公開、特許登録及び実案登録の何れか

3. 調査研究方法

評価例

評価

メインフレームの発明の内容をどの程度理解できるか、A、B、Cの3段階で評価
(要約や説明書・図面の参照可)

自由記載1（上側） メインフレームの評価 に対するコメントを記載

評価場所	分野	JPO翻訳	中文原文	評価	自由記載1(評価)	自由記載2(誤訳)
メイン	複合機	【請求項0001】1. 給紙機構、その特徴は、 大型用紙 になって 複数枚の小サイズ用紙 の枚で 寄せ集める ことができる 後処理装置 に向かうことに用いることは 用紙を搬送して *1、前記給紙機構が2個を含んで互いに狭圧する給紙ローラは、前記給紙ローラは軸と固定した軸の少なくとも1個の加圧ローラを含み、前記加圧ローラは給紙ローラ軸線方向の長さに沿い、用紙幅比を搬送されて設定値とすること、前記設定値は30%以上である。	1. 一种供纸机构，其特征在于，用于向可将多张小尺寸纸张拼合成大尺寸纸张的后处理装置搬送纸张*1，所述供纸机构包括两个相互夹压的供纸辊，所述供纸辊包括轴和固定在轴上的至少一个压辊。所述压辊沿供纸辊轴线方向的长度，与被搬送纸张宽度之比为设定值，所述设定值大于等于30%。	B	請求項中に「係り受け」の誤訳があり、当該部分の意味が理解し難い。しかし、説明書全体を読むことで内容を概ね理解することができる。	*1 係り受け ・複数枚の用紙を繋ぎ合わせて大サイズ用紙にすることができる後処理装置へ搬送することに用いられ、「とすべきところ、係り受けがおかしくなっている。 ・原文の「拼合…」にある「成(結果補語)」の訳がおかしい。
要約	複合機	給紙機構、後処理装置、画像形成システム、この給紙機構は 大型用紙 になって 複数枚の小サイズ用紙 の枚で 寄せ集める ことができる 後処理装置 に向かうことに用いて 用紙を搬送して *1、この給紙機構が2個を含んで互いに狭圧する給紙ローラは、前記給紙ローラは軸と固定した軸の少なくとも1個の加圧ローラを含み、前記加圧ローラは給紙ローラ軸線方向の長さに沿い、用紙幅比を搬送されて設定値とすること、前記設定値は30%以上である。この給紙機構は從来給紙機構を解決して用紙を搬送する時用紙可能性がありオフセット傾斜*2、しわの変形及び横湾曲変形の問題が発生する可能性がある*3。	一种供纸机构、后处理装置、图像形成系统，该供纸机构用于向可将多张小尺寸纸张拼合成大尺寸纸张的后处理装置搬送纸张*1，该供纸机构包括两个相互夹压的供纸辊，所述供纸辊包括轴和固定在轴上的至少一个压辊。所述压辊沿供纸辊轴线方向的长度，与被搬送纸张宽度之比为设定值，所述设定值大于等于30%。该供纸机构解决了现有供纸机构在搬送纸张时纸张可能会发生偏移倾斜*2、褶皱变形及横向弯曲变形的问题*3。	-	請求項1と同様。 一番最後の箇所は、日本人では意味が理解できない、又は誤解が生じる。	*2 単語の正確性 オフセット傾斜が誤訳・不明。ただし、中文「偏移傾斜」とあり、単純に「傾斜」となっていれば発生しなかったであろう誤訳。 ※3 係り受け 原文の「解决」と「問題」の係り受けがおかしい。正しくは、「この給紙機構は、従来の給紙機構が用紙を搬送する際にレジスト傾斜やしわによる変形、横曲変形が生じるといった問題を解決した」。

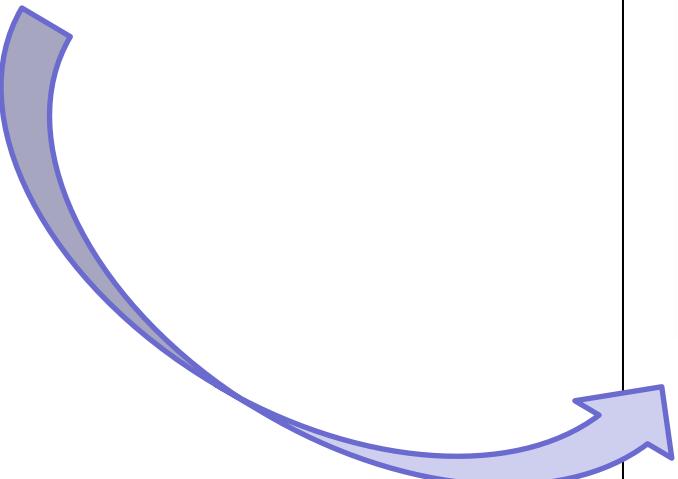
自由記載2

メインフレームと要約に関する誤訳を、「係り受け（名詞・動詞の誤り、修飾関係の誤り）」、「特許表現」、「単語の正確性」及び「訳抜け」の4つに分類し、コメントを記載

3. 調査研究方法

評価例

A、B、C評価についてイメージし難いと思いますので、評価例を添付資料に載せております。御参考ください。



評価項目	評価基準	評価結果	評価理由
1 要約	201310400400A CNA-10244827 総合評議会	【評価基準】評議会の運営が適切で、議論が活発かつ建設的であることを評価する。 評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。 評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。 評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。	評議会中に十分な議論が行われた。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。
2 要約	201310883705A CNA-10214469 総合評議会	【評価基準】評議会の運営が適切で、議論が活発かつ建設的であることを評価する。 評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。	評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。
3 要約	201310881123J CNA-10214460 総合評議会	【評価基準】評議会の運営が適切で、議論が活発かつ建設的であることを評価する。 評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。	評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。
4 要約	201310881123J CNA-10214460 総合評議会	【評価基準】評議会の運営が適切で、議論が活発かつ建設的であることを評価する。 評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。	評議会は、議題を明確に定め、議論を構造化して進行する。議論は、意見交換や議論を通じて、問題解決や意思決定に向けた具体的な行動計画を策定する。

3. 調査研究方法

(2) 検索精度

- 調査期間：2003年1月1日～2013年12月31日（検索式にて公報発効日を指定）
- 調査内容：当該システムによる日本語キーワードでの検索と、CNIPRによる中国語キーワードの検索との結果を比較し、どの程度検索結果が一致するか調査した。
- 調査項目：

◇網羅率（当該システムヒット件数／CNIPRヒット件数）

評価対象：「要約」、「請求の範囲」及び「全文」

当該システム（日本語）
の検索式

CNIPR（中国語）の検索式

分野	検索日	検索式		結果(ヒット件数)		網羅率
		公報テキスト検索（日本語検索）	CNIPR（中国語検索）	公報 テキス ト検索	CNIPR	
自動車部品	2015年7月6日	要約:パワーステアリング 電動操舵 操舵アシスト 要約:制御 コントロール 登録日:20030101:20131231	摘要=(动力转向 or 动力转向装置 or 动力操舵装置 or 动力方向盘 or 动力转向器 or 液压转向装置 or 电动助力转向) and 摘要=(控制 or 调节 or 管制 or 抑制 or 控制 or 控件 or 控) and 公开(公告)日=(20030101 to 20131231)	709	778	91%
		請求の範囲:パワーステアリング 電動操舵 操舵アシスト 請求の範囲:制御 コントロール 登録日:20030101:20131231	CLM=(动力转向 or 动力转向装置 or 动力操舵装置 or 动力方向盘 or 动力转向器 or 液压转向装置 or 电动助力转向) and CLM=(控制 or 调节 or 管制 or 抑制 or 控制 or 控件 or 控) and 公开(公告)日=(20030101 to 20131231)	898	1051	85%
		公報全文(書誌を除く):パワーステアリング 電動操舵 操舵アシスト 公報全文(書誌を除く):制御 コントロール 登録日:20030101:20131231	FT=(动力转向 or 动力转向装置 or 动力操舵装置 or 动力方向盘 or 动力转向器 or 液压转向装置 or 电动助力转向) and FT=(控制 or 调节 or 管制 or 抑制 or 控制 or 控件 or 控) and 公开(公告)日=(20030101 to 20131231)	3463	4077	85%

⇒要約

⇒請求の範囲

⇒全文

◇一致率（当該システムとCNIPRで一致する文献数／CNIPRヒット件数）

評価対象：「要約」（理由：評価件数調整のため）

評価件数：最大200件程度（理由：当該システムが千件以上の閲覧不可、案件番号確認で手作発生のため）

検索式の条件：IPCの利用可（理由：当該システムとCNIPRとでIPC検索のヒット件数が略一致するため）

4. 研究に基づく成果

(1) 翻訳精度

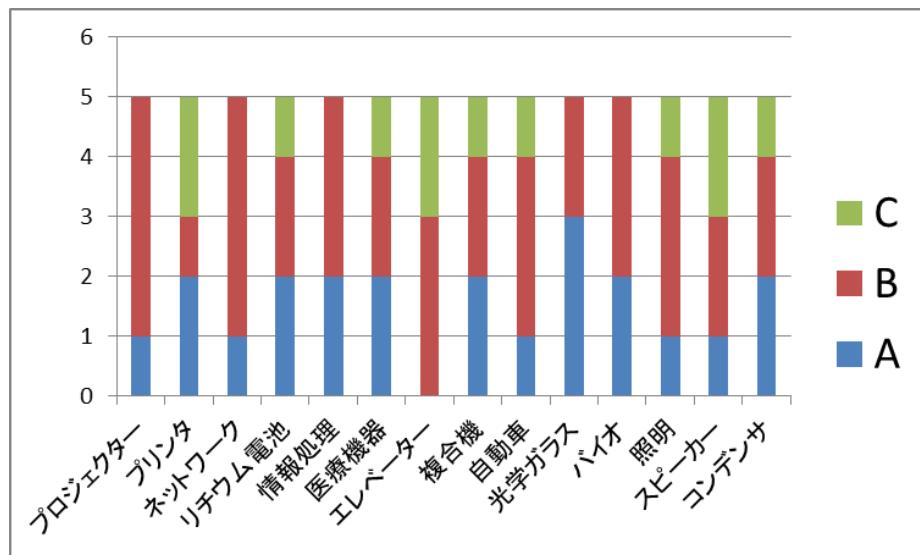


図3 翻訳精度の評価結果（分野毎）

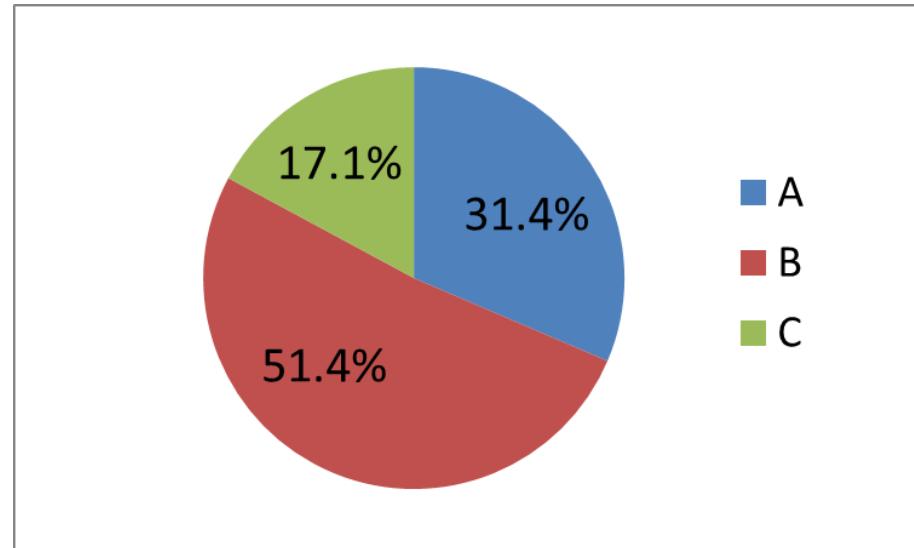


図4 翻訳精度の評価結果（総合）

■ 評価結果

- ・B評価以上が約83%あり、大まかに発明の内容を理解できる案件が多かった
- ・一方、C評価も約17%あり、発明を把握できない案件が介在した
- ・分野による翻訳精度の明確な差異を確認することはできなかった

■ 誤訳の傾向

- ・「係り受け」の誤訳が断トツで多かった

4. 研究に基づく成果

(1) 翻訳精度

<ハンドブックの作成>

中国語ができる日本人が本システムを利用するなどを前提とし、翻訳内に誤訳があったとしても、発明の内容を理解する上で手助けとなるハンドブックを作成した。

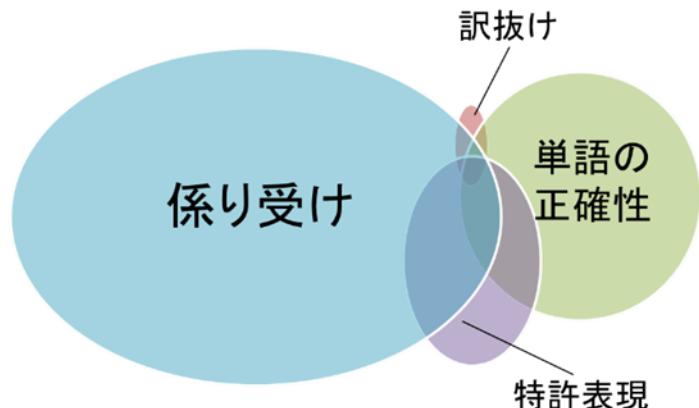


図5 誤訳の分類とその割合イメージ

ハンドブックを添付資料に載せております。御参考ください。

目次	
1. ハンドブックのねらい	1
2. ハンドブックの構成	3
3. ハンドブックで利用する符号の説明	3
4. ハンドブック	4
4.1 「～を含み（備え）」、「～から成る」又は「～組成する」に関する誤訳	4
4.2 「時」「情況」に関する誤訳	7
4.2.1 「～になつて～とき（時）」、「～とき（時）になつて」	7
4.2.2 「～になつて～前／～前になつて」又は「～になつて～後／～後になつて」	9
4.2.3 「～の過程（中）…」	10
4.2.4 「～場合（下限）」	11
4.3 訳されない方が意味を理解し易い誤訳	12
4.3.1 「その」	12
4.3.2 「その中」	13
4.3.3 「（に）用い／用いる／用いて」	14
4.3.4 「それによつて」	16
4.4 「～に基づいて等の表現を導く語句に関する誤訳	17
4.4.1 「～による／～によつて」	17
4.4.2 「～に使ひ」	19
4.5 請求項に比較的多く利用される語句に関する誤訳	20
4.5.1 「～になつて／～に致る／～になる」	20
4.5.2 「～に對／～に対する」	21
4.5.3 「～ごときでない」	22
4.5.4 「～から」	23
4.5.5 「～のため／あるいは／～のためにより」	24
4.5.6 「～が含まれた～を含む」	25

4. 研究に基づく成果

(2) 検索精度

- ・IPCの利用により一致率を上げられる
- ・キーワード利用時では、技術分野に対応して中国語と日本語が一対一で正確に翻訳されていれば、ここまでばらつき幅が広くなるとは考えにくい。よって、本システムとCNIPRとで検索アルゴリズムが異なることも考えられる（推測）。

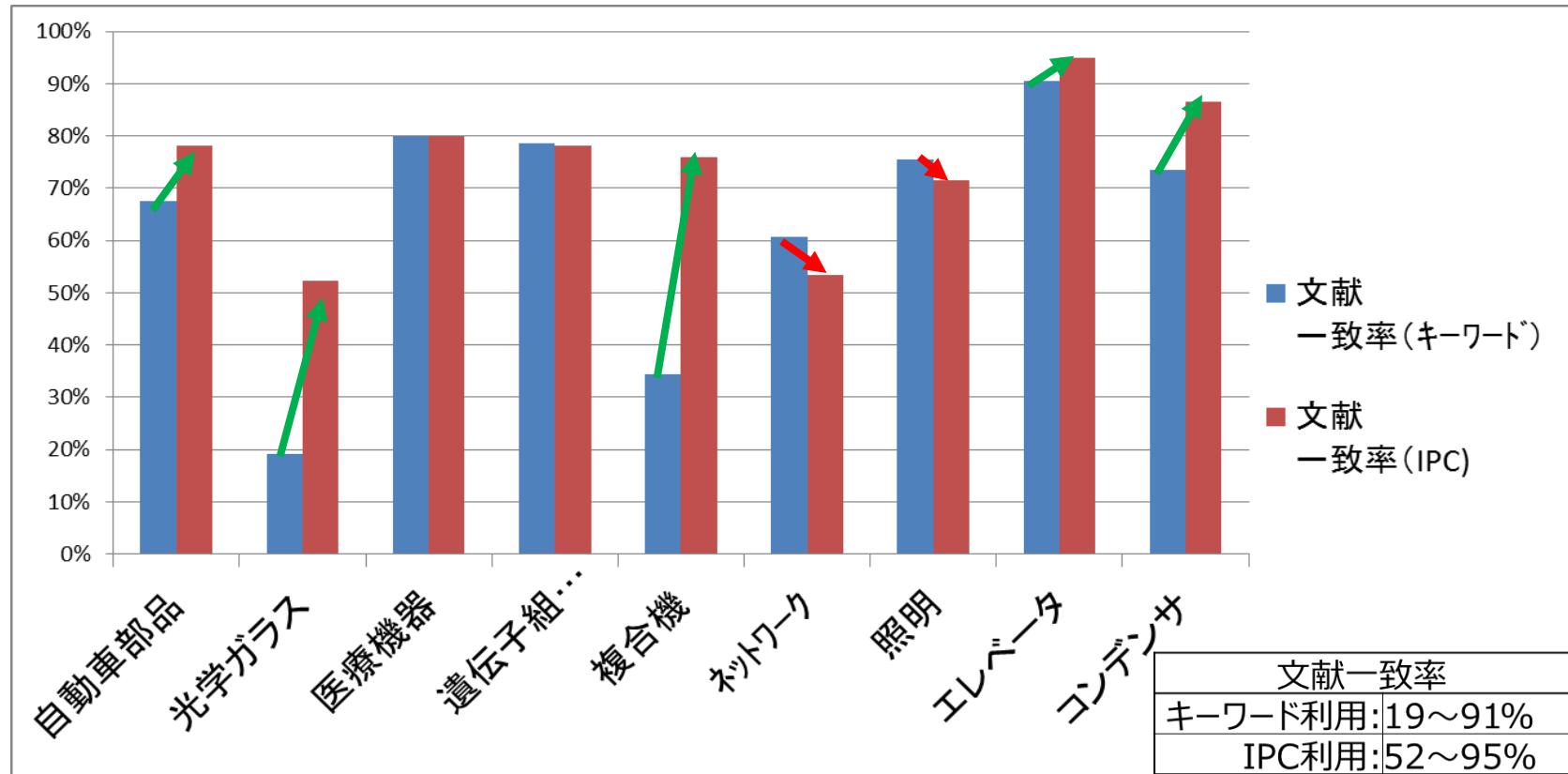


図6 誤訳の分類とその割合イメージ

■ 中国語の先行文献調査に活用

発明・実用新案の出願前先行文献調査を行う分には、中国語の解らない発明者・知財スタッフであっても概ね発明の内容を理解できる可能性がある。

■ 他社特許クリアランス時のノイズ取りに活用

例えば、母集団が1000件ある中からノイズを取って数百件に件数を絞ることは可能であると考える。ただし、発明の内容をほとんど理解できない翻訳文が含まれるため、完全にノイズを除去するのは困難である。この場合、より詳細な検討を行う場合には、別途人の手による翻訳等を検討する必要があると考える。

■ 日本特許庁への翻訳精度向上をアピール

JETRO様よろしくお願ひいたします。

なお、単語登録を充実させるだけでなく、中国語独特の文法表現による問題を解決できるよう（理由：係り受けの誤訳が多発のため）、改善を希望いたします。

7. 参考資料

■ 文献蓄積率

【評価内容】

CNIPRの蓄積数に対する、本システムの文献蓄積率をIPC毎に評価した。

【検索日】

特別テーマ調査委員会（2015年3月24日以前）及び本専門委員会（2015年5月20日以前）にて実施した。

【結果】

- (1) 2014年の蓄積率が低いが、他の年代は99%に近い蓄積があった。
- (2) 分野によってバラツキや100%を超える蓄積率が発生していた。

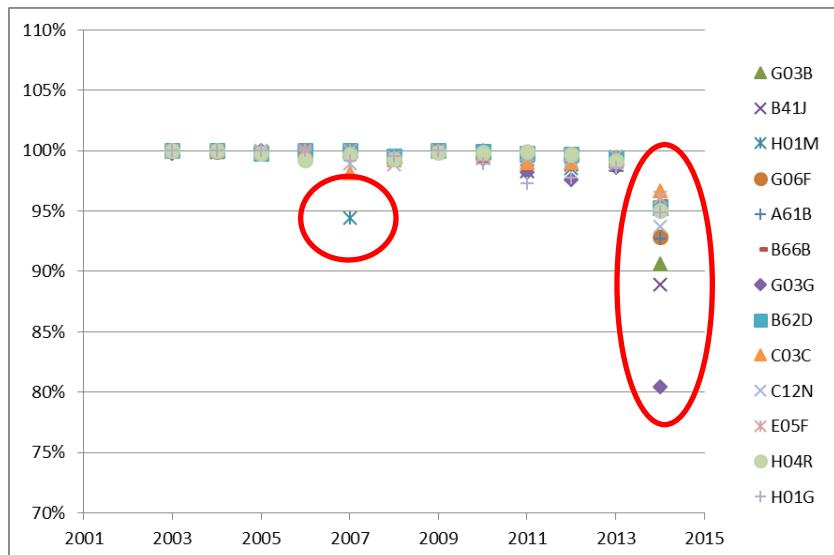


図7 特許公開

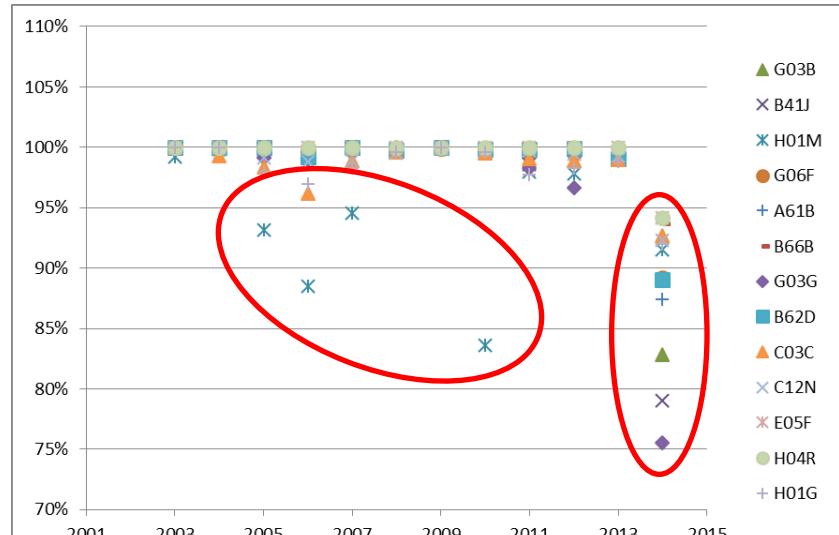


図7 特許公開

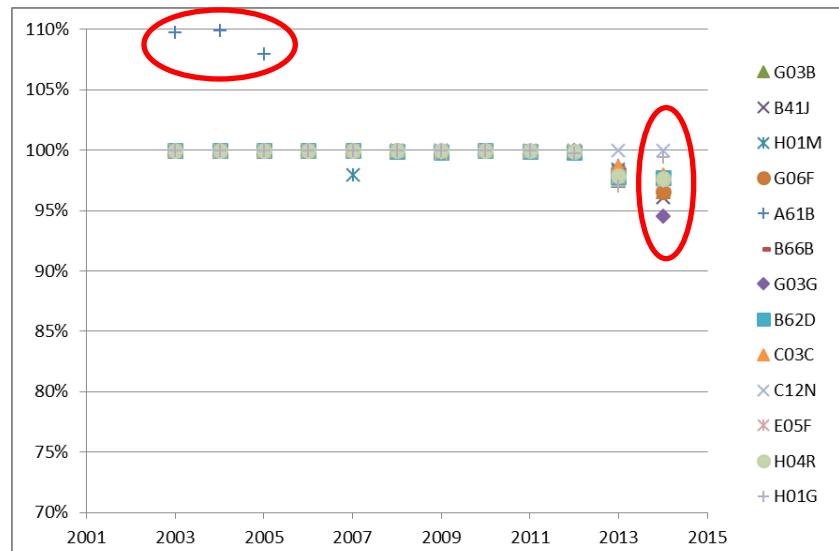


図9 実案登録