

2021年5月17日

中国特許情報の異表記・誤訳 用語編(その4)

アジア特許情報研究会:伊藤徹男

1. はじめに

ここでは既に紹介した誤訳の問題ではなく、どちらかという異表記として検索式にも反映させるべき中国語「文字化け文字」と「異体字」を紹介します。

1) 文字化け文字

「文字化け文字」とは、文字コードが異なるために日本で使われている通常の PC では(Google Chrome や IE などブラウザの問題でもなく)表示されない文字のことを指します。文字化け文字は台湾特許に多く見られますが、中国大陸特許にも存在します。

英語情報から洩れる「文字化け文字」も極めて少ないので、調査という観点からは無視してもよいものです。したがって、このような現象もある、という情報として受け止めていただければ結構です。

2015年までにかかなり多くの化合物の文字化け文字を抽出して検索式に異表記として用いていました。1つの化合物に8種もの文字化け文字がある化合物なども存在していました。この文字化け文字も2015年の台湾特許庁訪問時に改善を要望しましたが、これについては「改善する予定はない」とのことでしたが、昨年(2020年8月)確認したところでは、抽出した文字化け文字では検索できないものも現れ、逆に困惑しています。用語辞書の修正が必要だと思っておりましたが、修正に取り掛かれずにいました。一部の文字化け文字を多く含む用語(多くは化合物)について検証した結果を改めて紹介します。

2) 異体字

異体字という聞きなれない用語ですが Web には以下のようにも紹介されています。

『異体字(variant)とは、文字としては異なる見た目を持つものの、意味や発音が同じであり、歴史的・文化的な背景などから、ある文化圏や言語圏では同一の文字として見なされる文字のことを指します。

例えば「国」と「國」、「沢」と「澤」は、日本では新字と旧字として区別されるなど、意味的には同じであっても一般的には異なる文字として扱われています。

日本では明らかに意味が違う文字なのに、他の国では異体字と見なされ同一視される例もあります。例えば「機」と「机」、「葉」と「叶」などは、中国語圏では片方はもう片方の異体字として扱われます。「千葉」と「千叶」が同一視されるとは日本語の感覚では想像しづらく、「機上」と「机上」に至っては、日本語では意味が異なります。しかし、中国語圏ではこれらは同じ意味・文字となります。』

<https://www.nic.ad.jp/ja/basics/terms/variant.html>

様々な漢字異体字を紹介した情報も存在します。

「異体字 一覧」

https://www.google.com/search?lr=&as_qdr=all&sxsrf=ALeKk01hUqhj7FhgKhCU6Uz2ZrZi4IOcOg:1621243990336&source=univ&tbm=isch&q=%E7%95%B0%E4%BD%93%E5%AD%97+%E4%B8%80%E8%A6%A7&hl=ja&sa=X&ved=2ahUKEwiuiY3BtNDwAhVPxIsBHbW1AFMQjJkEegQIDxAB

日本でも戦前の旧字体と戦後の新字体が異体字だとされていますが、ここで紹介する異体字は、上記のようなものではなく、特許情報中に存在する「よく目を凝らしてみても同じように見える表記の中国語」という、さらに特殊な文字です。「歴史的には昔使われたもの」とも定義されています。

2. 文字化け文字

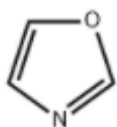
化合物としてはN原子やP原子を含むものに文字化け文字が多いのでその数例を例示します。

a)NITRILE

もっとも単純な化合物の例として「DINITRILE」の台湾特許を挙げます。DINITRILE は「二腈」と表記されますが、「二・」と「二・」もありますので、「二腈 or 二・ or 二・」を異表記として検索します。ACETONITRILE の場合も「乙腈 or 乙・ or 乙・」となります。

また、不飽和 (unsaturated nitrile)、丙烯 (acrylonitrile)、芳族 (aromatic nitriles)などと半角スペース部分が nitrile と英訳されているので、その部分をコピーして(*nitrile* and ())@ti と検索しましたが、nitrile としては何も検索できませんでした。

b)OXAZOLE



左に示すように環形化合物に N 原子や O 原子などが挿入された OXAZOLE などの複素環化合物にも多く見られます (図1)。

磺胺甲噁唑[◆]噻[◆]唑[◆]生物[◆]的[◆]制[◆]法[◆]和[◆]用[◆]途

Preparation method and use of sulfalene oxazole derivative

用于治疗心血管及有关疾病的噻唑或[◆]噻[◆]唑[◆]生物[◆]衍[◆]生[◆]物

Thiazole or oxazole derivatives which are useful in treatment of cardiovascular diseases

用作糖尿病治疗中的PPAR-α和PPAR-γ激活剂的羧酸取代的[◆]噻[◆]唑[◆]生物[◆]衍[◆]生[◆]物

Carboxylic and substituted oxazole derivatives for use as PPAR-α and PPAR-γ in the treatment of diabetes

噻唑和[◆]噻[◆]唑[◆]生物[◆]衍[◆]生[◆]物

Thiazole or oxazole derivatives

図1. OXAZOLE の文字化け文字(CN 公開特許)

また、台湾特許庁データベース GPSS では、(*oxazol*)@ti や(*quinolin*)@ti として、後方一致検索や部分一致検索で benzoxazole(苯并・唑)、isoxazolidine(异・唑)などの中国大陸特許や ALKYLPIPERAZINE(烷基哌・)、ANILINOPIPERAZINE(苯胺・)などの台湾特許も英語で検索が可能ですが、後方一致検索や部分一致検索ができないデータベースでは、やむなく文字化け文字での検索も動員せざるを得ません。

3. 漢字の異体字

よく目を凝らしてみても同じように見える表記の中国語があります。これも「異体字」というらしい。この異体字は文字コードが違うために検索では差別されて抽出されますが、中国大陸特許だけでなく、2015年以前は台湾特許にも存在しました。2015年の台湾特許庁への要望で(簡体字で繁体字も検索できるように、という)台湾特許にも存在する異体字は正規表記の中国語(繁体字でも簡体字でも)で検索するとその区別はなくなり、解消されました。

しかし、CNIPR などの中国特許データベースでの大陸特許検索では、現在でもなお、正規表記中国語と異体字は区別されて抽出されるので注意が必要です。異体字の存在数は極めてわずかなので無視してもいいかとは思いますが、権利判断調査では検索式に異体字も追加する必要があります。

また、このような(よく目を凝らしてみても同じように見える表記の中国語)異体字の存在はほとんど知られていませんので異体字で出願すれば検索から逃れることができる、とも言えます。見た目の公報表記は変わりませんので権利行使上も問題ないものと思われます。かつては日本特許でも意識的に検索逃れをしたのではないかというような例をセミナーで紹介したことがあります。

どんな文字が異体字として使われているか網羅した訳ではありませんが、以下のような例を挙げることができます。CNIPR 1990～2020 年公開特許の出現数。

表1. 中国語異体字の出現数

正規表記	TI	AB	CL	異体字	TI	AB	CL
内视镜	129	182	274	内视镜	2	1	1
减少	14026	788624	189836	减少	1	70	36
同步	47902	197831	378876	同步	3	6	51
研究	4432	98778	33950	研究	0	0	0
卞	3713	17985	89405	卞	3	46	266

表1の異体字「研究」については、発明の名称～請求の範囲では確認できませんが、出願人名では正規表記「研究院」で 333953 件、異体字「研 究院」では 53 件を確認できます。

また、benzyl の正規表記「苳」と異体字「苳」以外に似たような中国語として「卞」が benzyl と英訳されており、異体字かどうか不明ですが、請求の範囲で 1060 件存在します(卞胺(benzylamine)とし

て132件)。権利侵害調査ではこれも異表記として検索式に加えています。

公報やディスプレイ上では判別できず、検索してはじめて確認できる、というのは困ったものです。調査対象範囲の他の用語についても早急に確認してみたいと思います。

4. 中国語のスペルミス(誤記)

中国語が英訳されたもので特にスペルの長い英語には必ずと言っていいほどスペルミスが存在することについては、これまで何度か紹介しましたが、中国語にもスペルミスと思われる誤記があります。

「請求の範囲」を翻訳した英語情報は権利範囲を示したのですが、権利を確定する「請求の範囲」そのものではないので権利解釈上は問題になりません。中国語の誤記など誤記と判断されない場合、あるいは出願人などからの英語情報を元に中国語に翻訳された出願書類の誤訳が元で要求する権利内容とは異なる内容のものとなって無効審判などが起こされたりすることがあります。

ここでは異体字ではないそのような誤記について紹介したいと思います。数からすれば圧倒的に異体字より多いです。

中国特許における中国語のスペルミスについては、INFOPRO 2014 で「タイプミスを含む中国公報に関する留意点」として渋谷氏が中国特許情報中の中国語表記について示唆に富んだ報告をしています。

予稿:http://patentsearch.punyu.jp/asia/INFOPRO2014type_miss_pre.pdf

発表資料:http://patentsearch.punyu.jp/asia/INFOPRO2014type_miss.pdf

ここでは、己烷(hexane)の己が(化学分野では六を示す hexa として使われる)、「巳烷」「巳烷」と誤記されたものが存在し、それぞれ hexane と英訳されており、中でも「巳烷」の誤記が多く、さらには「环巳烷」が「ring alkane」と誤訳されている公報も散見され、これは巳が無視されて「环烷」として英訳されたと考えられる、と考察しています。己烷に対する巳烷は概念の異なる用語に翻訳されたものではないと思われるので、誤訳ではなく誤記として扱います。

このように「よく目を凝らしてみても同じように見える」異体字とは違って、「見た目が似た漢字(中国語)」が同一の英訳として表記されるものがいくつかあります。過去にも挙げました熱触媒としての azobisisobutyronitrile (AIBN)がその1つです。英語のスペルも長いので英語スペルミスもかなり多いのですが、中国語について 2000-2018 年の公開特許における出現数を表2のように示し、それぞれの誤記についても注釈しています。

表2. azobisisobutyronitrile (AIBN)の中国語誤記

	AB	CL	FULL
① 偶氮二异丁腈	1,864	12,270	31,469
② 偶氮二异丁氰	125	811	1,338
③ 偶氮二异庚腈	99	2,838	3,722
④ 偶氮异丁腈	14	153	1,245
⑤ 偶氮二异丁晴	13	136	259
⑥ 偶氮二异丁晴	7	128	224
⑦ 偶氮二乙丁腈	8	43	76
⑧ 偶氮二己丁腈	8	5	15
⑨ 偶氮异二丁腈	3	11	25
⑩ 偶氮二异丁氢	3	4	5
⑪ 腈偶氮二异丁	1	4	7
⑫ 偶氮二异丁基氰	1	2	4

- ③ 「庚(hepta(七))」は「丁(butyl)」の誤記
 ④ 「二」(bis)が欠落したもの
 ⑤⑥「晴」「晴」は「腈(nitrile)」の誤記と思われる。
 ⑦「乙」はethylであり、「异(iso)」の誤記
 ⑧「己(hexyl)」は「异(iso)」の誤記
 ⑩「氢(hydrogen)」は「腈(nitrile)」の誤記

そしてさらに AIBN の「nitrile」としての出現数を表3のように示しています。

表3. Nitrile に該当する中国語の存在数

	TI	AB	CL
① nitrile and 腈	1364	6750	22829
② nitrile and 氰	23	784	6243
⑤ nitrile and 晴	0	66	105
⑥ nitrile and 晴	17	535	1619

しかし、その後の検証で、表2中の「⑤⑥「晴」「晴」は「腈(nitrile)」の誤記と思われる。」については、以下のように誤記ではなく使われている例もありますので、訂正したいと思います。

腈

丙腈(propanenitrile)

己二腈(adiponitrile)

氰(主として cyano として)

氢氰(hydrocyanation)

还原氰基(reduce cyano groups)

氰基苯并(cyanobenzo)

氰酸酯(cyanate ester)

晴

乙晴(acetonitrile)

丁晴(acrylonitrile)

丙烯腈(acrylonitrile)

丙二腈(malononitrile)

晴

丁腈(nitrile butadiene) 同一要約中に丁腈もあり

丙烯腈(acrylonitrile)

丁腈橡胶(nitrile rubber or nitrile butadiene rubber)

同一要約中に丙烯腈(acrylonitrile)もあり。

その他、Polyethylene terephthalate(PET)の中国語誤記については、「中国特許情報の異表記・誤訳 用語編(その1)」で中国語の異表記としても紹介していますので割愛しますが、以下のような中国語の誤記例が数多くあります。ここでは一例を挙げます。

「中国特許情報の異表記・誤訳 用語編(その3)」では誤訳として紹介した羟(hydroxy)と羧(carboxy)の問題も中国語表記が似ているが故だと言えます。

・吡啶(pyridine)と吡啶

中国語「吡啶」は pyridine と英訳されますが、見た目にも同じような「吡啶」が Pyridine と英訳されている例があります。pyridinone も「吡啶酮」が一般的で「吡啶酮」はほとんど使われませんし、pyrazolyl も「吡啶基」が使われ、「吡啶基」は極めてわずかです。しかし、中国大陸特許や台湾特許いずれでも両方が使われています。異体字ではなく文字が酷似している、という問題です。

文字化け文字や異体字などは極めてわずかなので無視してもいいと思いますが、中国大陸特許では英語のスペルミス、誤訳だけでなく、中国語のスペルミス(誤記)なども存在するので要注意です。特に、他社特許のウォッチだけでなく、自社の出願特許が権利行使できる内容になっているかも見直してみるのも重要です。

以上