

A Synchronous Corpus-Based Study on the Usage and Perception of Judgement Terms in the Pan-Chinese Context

Oi Yee Kwong* and Benjamin K. Tsou*

Abstract

This paper reports on a synchronous corpus-based study of the *everyday* usage of a set of Chinese judgement terms. An earlier study on Hong Kong data found that these terms were more polysemous than their English counterparts within the legal domain, and were even more fuzzily used in general news reportage. The current study further compares their usage in general texts from other Chinese speech communities (Beijing, Taiwan, and Singapore) to explore the regional differences in lexicalisation and perception of the relevant legal concepts. Corpus data revealed the distinctiveness of the Singapore data, and that the contrasting frequency distributions of the terms and senses could be a result of the varied focus in reportage or the use of alternative expressions for the same concepts in individual communities. The analysis will contribute to the construction and enrichment of Pan-Chinese lexico-semantic resources, which will be useful for many natural language processing applications, such as machine translation.

Keywords: Synchronous Corpus-Based Study, Legal Concept and Terminology, Regional Differences, Pan-Chinese Lexico-Semantic Resources

1. Introduction

In this paper, we report on a synchronous corpus-based study on a set of semantically related legal terms, and propose a Pan-Chinese lexico-semantic resource for the legal domain, such as one in the form of a thesaurus, to differentiate the usage and perception of closely related legal concepts and terminology across various Chinese speech communities.

The situation with language and law is a very interesting one in Hong Kong. As pointed out by Tsou and Kwong [2003], the legal system in Hong Kong has operated through English

* Language Information Sciences Research Centre, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong
E-mail: {rlolivia, rlbtsou}@cityu.edu.hk

solely for more than 150 years. Following the implementation of legal bilingualism in the 90's, Hong Kong became the first community that follows the Common Law system which at the same time allows the use of both English and Chinese in court proceedings. In contrast to the many precisely lexicalised legal concepts in English, given its long and established tradition in Common Law, the use of their Chinese equivalents is apparently more fuzzy, as is evident from the lack of one-to-one correspondence of legal terms between English and Chinese. This difference in the cross-lingual lexicalisation of legal concepts between English and Chinese has thus become a substantial linguistic hurdle in the implementation of legal bilingualism, and is directly related to whether both languages could eventually be used in balanced and proper ways to the same effect in the legal domain. The apparent fuzziness with Chinese legal terms is nevertheless peculiar in the Hong Kong context, but not in other places where Chinese is also used as the official language in the legal domain, such as in Mainland China. A possible reason is that preciseness is somehow diluted upon translation. In English, for instance, a "verdict" and a "sentence" are sufficiently distinguished, despite their semantic relatedness (as both are related to the results of a trial). However, when expressed in, or more often translated into, Chinese, the preciseness is somehow weakened. On the one hand, the expression or translation in Chinese might have to take into account the corresponding syntactic constraints and stylistic differences in the two languages in order to sound natural, hence the variation in expressing the same concept in different contexts. On the other hand, the translation could sometimes be affected by usages in other places. For example, although "contract" is mostly expressed as "合約" in Hong Kong, it is sometimes expressed as "合同", which is the term used in Mainland China for this concept.

A set of semantically related legal terms was studied by Tsou and Kwong [2003] with respect to their usage in the legal domain and in general texts. The terms are "裁定" (hold, convicted), "裁決" (determine, verdict), "判決" (judgement, conviction), "裁斷" (find, finding), and "裁判" (Magistracy)¹, all of which have one or more senses referring to some aspects of "judgement". Their uses were studied via a corpus of bilingual court judgments² and a general corpus of news articles from Hong Kong. From the corpus of bilingual judgments, it was observed that these Chinese terms were considerably polysemous, such that most of them were found as the renditions for multiple English terms, which are distinct though closely related in meaning. Even more varied usages were found for the Chinese terms

¹ The English terms are the more common translations of the corresponding Chinese terms as observed from Hong Kong court judgments. They are included here for reference only and are not necessarily the absolute or correct translations *per se*.

² An explicit distinction between the use of "judgment" and "judgement" is drawn here, as the inclusion or omission of the "e" is not arbitrary. "Judgment" refers specifically to the concluding writing for a court trial, while "judgement" refers to the action of judging in general.

in the general corpus, and the sense boundaries appeared to be more ambiguous. Although the confusion might be insignificant to the perception of the general readers, the subtlety therein could bear significant conceptual difference in the stricter legal domain, which is far less tolerable of impreciseness and ambiguity. Attempts were made to identify near-synonymous senses from the corpus examples, and to arrange them in terms of their semantic relatedness into a verb hierarchy and a noun hierarchy, in a way similar to WordNet [Miller *et al.* 1990].

The current work aims at expanding on the above study, to further explore and analyse the different usages and finely grained senses of the aforementioned Chinese legal terms, among news texts from various Chinese speech communities including Hong Kong, Beijing, Taiwan, and Singapore. Based on the usages of these terms in the corpus texts from different communities, we look for local differences in the lexicalisation, if any, and thus the perception of the corresponding legal concepts, which might be a result of the differences in social structures or legal systems. On the one hand, the different societies and legal systems might share similar legal concepts which are expressed in the same ways. On the other hand, in the case where they do not share similar lexical items, it is important to see what alternatives are used in different places to express related concepts. This is a necessary though preliminary step in the development of a Pan-Chinese lexico-semantic resource for legal terminology.

Efforts have been made by researchers in lexical semantics on the study of semantic relations among Chinese lexical items, with a view toward organising the lexical items into semantic networks. Gao [2001], for example, proposed a quantitative measure for the closeness and differentiation of near-synonyms among verbs denoting physical actions from a range of lexical semantic features. Cheng [2001] discussed the differentiation of related words from their individual focus and orientation. Nevertheless, work on lexical semantics and corpus-based lexicography often only drew reference from one particular corpus. Huang *et al.* [2000] worked on verbal semantics and near-synonyms of Mandarin Chinese as used in Taiwan. Tongyici Cilin [梅等 1984] is based exclusively on Chinese as used in post-1949 Mainland. However, linguistic variation is significant and especially salient for Chinese language used in different communities [Tsou *et al.* 2004]. Our corpus-based, Pan-Chinese approach, beginning with a set of domain-specific lexical items, thus has additional advantages for its indigenoussness, portability, and versatility. Such a Pan-Chinese lexical resource, when done in large scale, could contribute to natural language processing applications like machine translation and serve as a rich reference for legal and paralegal professionals. More importantly, the resource would capture the linguistic norms from more than one Chinese speech community.

In Section 2, we first briefly review the polysemy of legal terms and the complexity of translating legal terms from English to Chinese. Then in Section 3, we present the details of

the corpus analysis done in the current study. Results are discussed in Section 4, and we will conclude with future directions in Section 5.

2. Polysemy of Legal Terms

Tsou and Kwong [2003] started with a set of Chinese legal terms, all with senses related to “judgement” or “the action of judging” in the legal context, to study how the preciseness of legal concepts lexicalised in English is captured in their Chinese translations as shown in bilingual court judgments in Hong Kong, and how the preciseness of the latter is in turn preserved in a general corpus. The set of “judgement” terms includes “裁定” (hold, convicted), “裁決” (determine, verdict), “判決” (judgement, conviction), “裁斷” (find, finding), and “裁判” (Magistracy). They observed that in Hong Kong, despite the implementation of legal bilingualism for several years, legal concepts are not as precisely lexicalised in Chinese as in English. Thus while a particular legal concept could be relatively unambiguously expressed by one term in English, the same concept may not have a direct one-to-one corresponding term in Chinese. The fuzziness is carried over from legal contexts as in court judgments to informal contexts as in news reports. For instance, “裁決” has been identified as the translation equivalent for “decision”, “verdict”, and “award” in the bilingual corpus of court judgments. Similarly, the word “decision” has been rendered as “決定”, “裁決”, and “判決”, among many possibilities. Such a complex correspondence (as further illustrated in Table 1 and Figure 1) between English and Chinese legal terms can be explained by the fact that the use of English is much more mature in the Common Law system in Hong Kong. Many legal concepts are thus lexicalised and can be precisely expressed in English, whereas this preciseness is greatly weakened when terms are translated into Chinese.

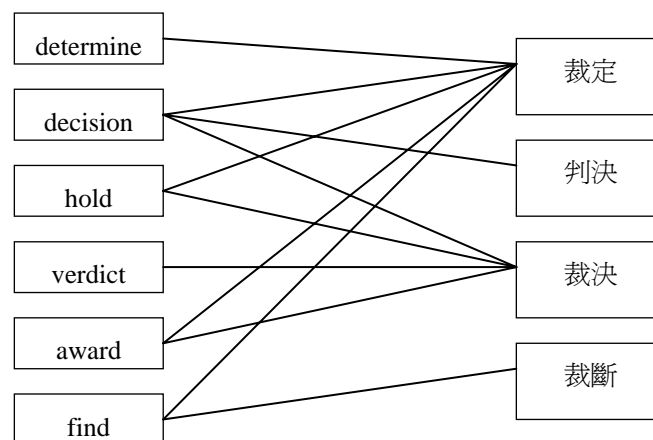


Figure 1. Interwoven mapping between English and Chinese legal terms

Table 1. Example of multiple renditions between English and Chinese legal terms

English	Chinese	Examples	
Decision	裁決	In <i>Mayson v. Clouet</i> [1924] AC980 the Privy Council approved the decision in <i>Howe v. Smith</i> (1884).	樞密院在 <i>Mayson v. Clouet</i> ([1924]AC980) 一案中對 <i>Howe v. Smith</i> ([1884]27Ch.D.89) 一案的裁決表示贊同。
Verdict	裁決	The jury returned their verdicts on 2 September 1997.	1997年9月2日，陪審團作出裁決。
Award	裁決	This was followed by a request that the tribunal should postpone making an award for two months.	這一段確認了賣方跟著請求仲裁庭延遲兩個月作出裁決。
Decision	決定	None of these decisions assist the appellant.	這些決定之中沒有一宗對上訴人有所幫助。
Decision	判決	The Court of Appeal, by a majority (Nazareth V-P and Liu JA, Rogers JA dissenting) reversed her decision, dismissing the claim to specific performance and holding that Douglas was entitled to forfeit the deposit.	上訴法庭以大比數（上訴庭副庭長黎守律和上訴庭法官廖子明同意，上訴庭法官羅傑志持異議）推翻她的判決，撤銷強制履行的申索和裁定第二與訟方有權沒收訂金。

Despite the complexity of multiple renditions, the morphemic structure of the individual Chinese terms might nevertheless indicate a core sense of the term, and thus suggest the focus of the relevant concept. For example, while the words “裁定”，“裁決” and “裁斷” share an identical morpheme “裁” (to judge), they could be differentiated by their second morphemes, which focus on “conclusion”，“decision”，and “inference” respectively. This distinction is similar to Cheng’s [2001] discussion of word families where similar and related words could be differentiated by their individual focus and orientation, or meaning facets.

In the current study, we are interested to see if the same kind of polysemy appears in the usage of the same legal terms in different Chinese speech communities. It is hypothesised that we may not find exactly the same usage of the same terms in the various communities, as their different social structures and legal systems might lead to different perception of the corresponding legal concepts, and the same concepts may not be equally salient for people in different communities. We will probe, using authentic corpus data, the perception of the various legal concepts in different communities, and see how the salience of “judgement” is reflected in the language used in different places; and if they do not use the terms in the same way, what alternative words are used to express similar concepts.

3. A Synchronous Corpus-Based Study

3.1 Materials

In this study, we further analyse three terms (called “target words” hereafter) which Tsou and Kwong [2003] studied, namely “裁定” (hold, convicted), “裁決” (determine, verdict), and “判決” (judgement, conviction). We leave out “裁斷” (find, finding), as it was only found in the bilingual court judgment corpus but not at all in the general corpus (LIVAC, as introduced below), and “裁判” (Magistracy), as it was found to mostly refer to the sense of “umpire” or “adjudication in a contest” when used in the general corpus.

Sentential contexts for the target words were extracted from a subset of the LIVAC corpus [Tsou *et al.* 2000]. LIVAC (<http://www.livac.org>) is a synchronous corpus developed by the Language Information Sciences Research Centre of the City University of Hong Kong. The corpus contains newspaper articles collected synchronously and regularly from six Chinese speech communities. The subset we used in the current study consists of texts from Hong Kong (HK), Beijing (BJ), Taiwan (TW) and Singapore (SG), covering local news, international news, sports news, entertainment news, and financial news, collected over the *same* period of time (for two years, 1997-98 and 2002-03). Each sub-corpus, that is, texts from each of the four places, contains about 5M Chinese characters, which yields about 3M words upon segmentation.

3.2 The Analysis

For each target word, 30 samples of their sentential contexts (where there was sufficient data) were randomly selected from each sub-corpus, and assigned a sense from the sense inventory as in Tsou and Kwong [2003] where appropriate. New senses were recorded when found. Upon sense tagging, the samples from the various sub-corpora were further analysed with respect to the sense distribution of each word in each community, and the similarities and differences of such distributions across the various communities. As in the previous study, the assignment of senses took into account the collocation patterns of the different senses and subcategorisation patterns of the verbal usages where applicable. In addition, bi-syllabic words containing the morphemes “裁”, “定”, “決” or “判” were retrieved from the sub-corpora. The retrieved words and their relative frequencies were studied, to disclose any perceptual difference of the related legal concepts and any alternative expressions of such concepts in the different communities.

4. Results and Discussion

4.1 Relative Frequency Distribution

The frequency of the target words from the various sub-corpora is shown in Table 2. The relatively low frequency of all the target words in BJ data is most notable. The small numbers readily indicate that court news does not receive as much attention in Beijing newspapers as in other places.

Table 2. Frequency of the target words in the sub-corpora

Word	HK	BJ	TW	SG
裁定	122	38	80	19
裁決	142	32	54	139
判決	160	66	210	341

Just comparing the absolute frequencies, “判決” ranked highest across the board. Its relative frequency is especially high for TW and SG. In the most dramatic case of SG data, there are 341 occurrences of “判決” but only 19 occurrences of “裁定”. This is very different from, for example, HK data, where the relative frequency of “判決” and “裁定” differs by less than 10%. To a certain extent, this difference in relative frequency suggests a variation of focus in news reportage in the two communities, assuming that the use of these words in them is not arbitrary. According to many legal dictionaries [e.g. 《法學詞典》編輯委員會 1985; 劉清景 2001], “判決” and “裁定” refer to different aspects of the ruling of a court. In particular, “判決” is often associated with the final determination on the main issue in a trial, whereas “裁定” usually refers to the conclusions to other factual disputes during a trial.

4.2 Sense Distribution

Sample sentences of the target words from BJ, TW and SG data were examined and each occurrence of the words was assigned a sense with reference to the sense set defined by Tsou and Kwong [2003]. The sense distributions were compared to those reported for Hong Kong data in the same study. The results are tabulated in Tables 3 to 5 for “裁定”, “裁決” and “判決” respectively (the number in brackets below each place refers to the number of samples checked and all figures reported are percentages). The second and third columns refer to data gathered from a bilingual corpus of Hong Kong court judgments and a subset of the current HK data from LIVAC respectively as reported in the earlier study. In this subsection we focus on the sense distribution with respect to individual words, and in the next we will further explore the regional differences observed.

4.2.1 裁定

“裁定” was earlier distinguished into four verb senses and two noun senses. No new sense was found from the data in the current study. The dominance of its verb usages is observed in all places except BJ. Over 70% of the samples for BJ were assigned sense 6 (as in “作出終審裁定”, “阻礙判決、裁定的執行”, “維持原判的裁定”, etc.). This contrasts enormously not only with other places but also with data from the legal domain. It also contrasts dramatically with SG data where no nominal usages were found at all for “裁定”. This may be a consequence of the small number of SG samples but is more likely a genuine difference in the usage of the word, as we will further discuss below. Another interesting observation is that in the legal corpus, “裁定” is seldom used to state the order given by the court (sense 3) and the BJ data are more or less in line with this. However, over 25% of the samples from HK, TW and SG fall under sense 3. This thus raises interesting questions regarding the saliency of the concepts in individual places.

Table 3. Sense distribution of “裁定”

Sense and Example	Legal (30)	HK (30)	BJ (30)	TW (30)	SG (19)
1. [v.] the court decides on the outcome of a case e.g. 法庭裁定...罪名成立。	43.33	43.33	3.33	16.67	31.60
2. [v.] the court resolves an issue in a case e.g. 法官裁定所提出的要求沒有得到滿意答覆。	36.67	20.00	6.67	30.00	26.30
3. [v.] the court gives an order e.g. 法官裁定港府要即時釋放他們。	3.33	26.67	6.67	26.67	42.10
4. [v.] to judge on some issue to resolve dispute e.g. 法庭需要裁定臨立會的合法性。	0.00	6.67	0.00	3.33	0.00
5. [n.] the resolution of an issue in dispute e.g. ...裁定回覆不能令人滿意。本席認為上述裁定正確無誤。	10.00	0.00	6.67	6.67	0.00
6. [n.] the decision on the outcome of a case e.g. 我認為暫委法官的裁定是正確的。	6.67	3.33	76.67	16.67	0.00

4.2.2 裁決

Sense tagging is notorious for its difficulty as the meaning in the new occurrence of a word is not always so clear-cut that a pre-defined sense could be unambiguously assigned to it. In this regard, the tagging for “裁決” was most difficult and confusing. The difficulty may be largely attributed to its relatively general meaning. For example, according to the hierarchies suggested by Tsou and Kwong [2003], the verb sense 裁決/1 and the noun sense 裁決/5 are

the top nodes³ in their respective hierarchies. Having the most general sense amongst others, it means that the word can be used in a relatively wide variety of contexts. When it comes to general news reports, they are not necessarily the correct and legitimate contexts. This is evident in two respects from Table 4.

Table 4. Sense distribution of “裁決”

Sense and Example	Legal (30)	HK (30)	BJ (25)	TW (30)	SG (30)
1. [v.] the court makes a decision based on evidence e.g. 若有法律觀點分歧，最終交由法庭裁決。	16.67	6.67	0.00	23.33 [†]	33.33
2. [v.] the court decides on the outcome/sentence/etc. e.g. 法官裁決...把謀殺罪減為誤殺罪...	0.00	3.33	0.00	23.33	20.00
3. [n.] the court's decision on the outcome of a case e.g. 陪審團達至誤殺的裁決。	30.00	6.67	20.00	13.33	0.00
4. [n.] the court's decision on monetary compensation e.g. ...拒絕執行公約裁決...	23.33	0.00	0.00	0.00	3.33
5. [n.] the court's decision on a case and orders e.g. ...會就法院上月底裁定港府要釋放十名越南人的裁決上訴。	16.67	53.33	40.00 [†]	30.00	23.33
6. [n.] the resolution of an issue e.g. ...關於證據接納性的裁決...	13.33	20.00	40.00 [†]	0.00	0.00
7. [n.] religious orders, etc. e.g. ...聽命於這類宗教裁決...	0.00	10.00	0.00	0.00	0.00
*8. [v.] the court decides on the outcome of a case e.g. 法官路易斯迪蘇沙裁決他的罪名成立。	--	--	--	3.33	10.00
*9. [v.] the court resolves an issue in a case e.g. ...裁決搭客沒有起訴保險公司的權利。	--	--	--	--	6.67
*10. [v.] to judge on some issue to resolve dispute e.g. 三司必須裁決的只是關係到公眾利益的法律問題。	--	--	--	--	3.33
*11. [v.] the action of judging by referee in sports events e.g. ...但二壘審竟判定外野手是接殺，然後訴請裁決...	--	--	--	6.67	--

³ In the verb hierarchy, {裁決/1} subsumes {裁定/1, 判決/1}, {裁定/2, 判決/2}, {裁定/3, 裁決/2, 判決/3}, while {裁定/2, 判決/2} further subsumes {裁定/4}. In the noun hierarchy, {裁決/5, 判決/4} subsumes {裁定/6, 裁決/3}, {裁定/5, 裁決/6}, {裁決/4}, while {裁定/6, 裁決/3} further subsumes {判決/5}. Senses within curly brackets belong to the same synonym set [Tsou and Kwong 2003].

First, the use of “裁決” is almost abused in the SG data. Senses 8 to 10 in Table 4 were unexpectedly found from SG. They are essentially equivalent to individual senses identified earlier for “裁定” and “判決”. In particular, 裁決/8 is apparently synonymous to {裁定/1, 判決/1}, 裁決/9 to {裁定/2, 判決/2}, and 裁決/10 to {裁定/4}. However, the low relative frequency for these senses of “裁決” and their absence from other regions suggest that these senses might more appropriately and specifically be replaced by the relevant senses of “裁定” and “判決” instead, as the examples for these senses in Table 4 sound slightly unnatural.

Second, the loose restriction on “裁決” is also reflected in the BJ and TW data (marked with † in Table 4) as the word is often used to refer to the decisions (or the action of making decisions) made by non-judiciary units (e.g. “完全由行政機關裁決” – TW, “世貿組織曾做出裁決” – BJ, etc.). Apart from this, an additional sense for “裁決” related to a referee’s judgement (sense 11) was found from the TW data; while sense 4 (decision on monetary compensation) is so domain-specific and technical that it is rarely found outside legal documents.

4.2.3 判決

As seen from Table 5, all regions show a similar sense distribution for “判決”, where senses 4 and 1 are the major senses. Sense 5 (i.e. conviction) is specific enough to appear only in the legal texts. Additional uses referring to a referee’s judgement in a sports event were observed from TW and SG data.

Table 5. Sense distribution of “判決”

Sense and Example	Legal (30)	HK (30)	BJ (30)	TW (30)	SG (30)
1. [v.] the court decides on the outcome of a case e.g. 本席判決上訴得直。	13.33	20.00	20.00	36.67	16.67
2. [v.] the court resolves an issue in a case e.g. 上訴法院判決受託人有權提出呈請。	3.33	6.67	0.00	0.00	3.33
3. [v.] the court gives an order or sentence e.g. 國際法庭未被授權判決罪犯死刑。	0.00	6.67	3.33	3.33	0.00
4. [n.] the decisions made by court, and related orders e.g. ...一宗上訴案的判決...	56.67	66.67	76.67	53.33	70.00
5. [n.] conviction, the judgment of being guilty e.g. 以上是二項罪名定罪判決的立場。	26.67	0.00	0.00	3.33	0.00
*6. [v.] the action of judging by referee in sports events e.g. 眾多裁判判決爭議的確曾阻擾比賽。	--	--	--	3.33	--
*7. [n.] the judgement of referee in sports events e.g. ...因爲不滿裁判的一個判決而摔球拍...	--	--	--	--	10.00

4.3 Regional Variation

An obvious difference among the three target words is that verb uses tend to dominate for “裁定”⁴ whereas “裁決” and “判決” are used more often as nouns. However, from Table 3, it can be seen that BJ has a lot more nominal usages (sense 6) of “裁定”. This observation is nevertheless in line with the findings of Kwong and Tsou [2003] on verb-noun categorial fluidity in Chinese, where in texts from BJ, about 18% of the verbs were found to undergo the verb-noun shift, compared to about 15% of the verbs in the TW and HK data. Hence this linguistic phenomenon might account for the dominance of sense 6 of “裁定” in BJ data.

There remain some interesting questions regarding the differences in sense distribution across the various regions:

1. Table 3 shows that “裁定” is mostly used in SG to state the order given by the court, that is, sense 3. Does this tell us anything about the salience of court orders in SG reportage?
2. Table 3 also shows that “裁定” is least used in sense 3 in BJ, compared to other places. The frequency of its synonymous senses 裁決/2 and 判決/3 is also extremely low in BJ data. So is the concept missing or expressed in another manner?
3. 裁決/3 is absent from SG data, and so is its synonym 裁定/6. Do SG news reports pay little attention to verdicts? Otherwise where has the concept been absorbed into other expressions?

The difference in sense distribution across various regions is on the one hand a result of the different linguistic norms and styles of language use, as exhibited by the dominance of nominal usages in BJ. Hence even though BJ does not use 裁定/3 or its synonyms, the relevant concepts might have been expressed via nominal uses such as 判決/4. On the other hand, it could reflect the varied approaches and perception in different communities regarding the concepts of judgement. For instance, these concepts are apparently less salient in BJ contexts given the relatively low frequency of the target words in BJ data. Moreover, since SG is found to use 裁定/3 heavily but not 裁決/3, it suggests that SG news tends to treat the conclusion (verdict) and the consequence (sentence and order) as a whole.

To further understand this Pan-Chinese variation, we have simultaneously, though only briefly, surveyed the use of more similar and related lexical items from various resources. For instance, among the words mined from the seed morpheme “判” from the corpus materials used in this study, words like “判刑”, “判監”, “判罰”, “判囚” and “判處” (all related to sentencing) are relatively more abundant in HK and TW than in SG, suggesting that HK and TW tend to distinguish between the verdict and the sentence more clearly. On the other hand,

⁴ In particular, sense 1 and sense 6 of “裁定” both refer to the same meaning but differ in syntactic category.

a preliminary survey through some internet resources on legal documents (e.g. websites with PRC judgments⁵ and Taiwan judgments⁶) shows that terms related to “decision” and “sentencing” like “判定”, “判處”, “判令”, and “判決” are shared by judgments produced in Hong Kong, Mainland China, and Taiwan region. However, a related but probably more specific term, “判命” (which appears to combine decision, sentencing and order), is found only in Taiwan judgments. The variation and the subtle difference among such related terms would be useful and should be captured in a comprehensive Pan-Chinese lexico-semantic resource for the legal domain. Thus the use of these words, their relations with the target words, and their variation in the Pan-Chinese context all require further investigation.

5. Conclusion

Thus in this study we have further analysed the usage and sense distribution of a set of closely related legal terms pertaining to judgement in the Pan-Chinese context. Linguistic data reveal variation in the salience of these concepts in various Chinese speech communities and the distinctiveness of the SG data. Based on the subtleties among various uses, we have further probed the salience of these concepts in the various communities, and the differences therein might be a result of the difference in legal systems. For instance, the use of the target words in HK and SG might be more influenced by translation from English legal terms than in BJ and TW. Alternatives for expressing similar and related legal concepts should be further explored and the study should be expanded with other sets of closely related legal terms.

As mentioned in the beginning of this paper, the subtle differences among the target words may be insignificant to the general readers. However, when it comes to high quality translation, especially translations which bear legal implications, the preciseness therein will definitely be indispensable. Hence our analysis of the use of legal terms in various Chinese speech communities will provide useful information for the construction of a Pan-Chinese legal term lexical resource as we witness the growing maturity of the Chinese language in the legal domain. Such an enriched lexical resource would be useful to legal and paralegal professionals, and for legal document translation between English and Chinese, by machines or by humans, as well as for many other natural language processing tasks.

⁵ <http://www.chinaiprlaw.com/wsjsx/wsjsx.htm>

⁶ 法源法律網 <http://db.lawbank.com.tw/FJUD/FJUDQRY01-1.asp>

Acknowledgements

An earlier version of this paper was presented in the Fifth Chinese Lexical Semantics Workshop, Singapore, June 2004. This work was supported by Competitive Earmarked Research Grants (CERG) of the Research Grants Council of Hong Kong under grant nos. CityU1233/01H and CityU1317/03H.

References

- Cheng, C.C., “論同義詞的語意輕重與側重”, Presented at the *Second Workshop on Chinese Lexical Semantics*, 2001, Peking University, Beijing, China.
- Gao, H., “A Specification System for Measuring Relationship among Near-synonyms of Physical Action Verbs,” Presented at the *Second Workshop on Chinese Lexical Semantics*, 2001, Peking University, Beijing, China.
- Huang, C-R., K. Ahrens, L-L. Chang, K-J. Chen, M-C. Liu, and M-C. Tsai, “The Module-Attribute Representation of Verbal Semantics: From Semantics to Argument Structure,” *Computational Linguistics and Chinese Language Processing*, 5(1), 2000, pp.19-46.
- Kwong, O.Y., and B.K. Tsou, “A Synchronous Corpus-Based Study of Verb-Noun Fluidity in Chinese,” In *Proceedings of the 17th Pacific Asia Conference on Language, Information and Computation (PACLIC 17)*, 2003, Singapore, pp. 194-203.
- Miller, G.A., R. Beckwith, C. Fellbaum, D. Gross, and K.J. Miller, “Introduction to WordNet: An on-line lexical database,” *International Journal of Lexicography*, 3(4), 1990, pp. 235-244.
- Tsou, B.K., and O.Y. Kwong, “When Laws Get Common: Comparing the Use of Legal Terms in Two Corpora,” *Language and Linguistics*, 4(3), 2003, pp. 609-629.
- Tsou, B.K., T.B.Y. Lai, and K. Chow, “Comparing Entropies within the Chinese Language,” In K.Y. Su, J. Tsujii, J.H. Lee and O.Y. Kwong (Eds), *Natural Language Processing – IJCNLP 2004*, LNAI Vol. 3248, Springer-Verlag, pp.466-475.
- Tsou, B.K., W.F. Tsoi, T.B.Y. Lai, J. Hu, and S.W.K. Chan, “LIVAC, A Chinese Synchronous Corpus, and Some Applications,” In *Proceedings of the ICCLC International Conference on Chinese Language Computing*, 2000, Chicago, pp. 233-238.
- 《法學詞典》編輯委員會, 《法學詞典(增訂版)》, 上海辭書出版社, 1985.
- 劉清景, 《新編法律大詞典》, 學知出版事業股份有限公司, 2001.
- 梅家駒、竺一鳴、高蘊琦、殷鴻翔, 《同義詞詞林》, 商務印書館/上海辭書出版社, 1984.

