Wow! What a useful extension!
Introducing Non-Referential Concepts to Wordnet

Luis Morgado Da Costa and Francis Bond

Linguistics and Multilingual Studies,
Nanyang Technological University, Singapore
luis.passos.morgado@gmail.com, bond@ieee.com

Abstract

In this paper we present the ongoing efforts to expand the depth and breath of the Open Multilingual Wordnet coverage by introducing two new classes of non-referential concepts to wordnet hierarchies: interjections and numeral classifiers. The lexical semantic hierarchy pioneered by Princeton Wordnet has traditionally restricted its coverage to referential and contentful classes of words: such as nouns, verbs, adjectives and adverbs. Previous efforts have been employed to enrich wordnet resources including, for example, the inclusion of pronouns, determiners and quantifiers within their hierarchies. Following similar efforts, and motivated by the ongoing semantic annotation of the NTU-Multilingual Corpus, we decided that the four traditional classes of words present in wordnets were too restrictive. Though non-referential, interjections and classifiers possess interesting semantics features that can be well captured by lexical resources like wordnets. In this paper, we will further motivate our decision to include non-referential concepts in wordnets and give an account of the current state of this expansion.

1. Introduction

In this paper we motivate and describe our ongoing efforts to expand the Open Multilingual Wordnet (OMW)'s coverage by introducing non-referential concepts to wordnet hierarchies (Bond and Foster, 2013). Namely, the introduction of interjections as a language independent class of concepts, and the introduction of numeral classifiers as a language specific class of concepts.

This expansion has mainly been motivated by and done in parallel with the ongoing semantic annotation of the NTU-Multilingual Corpus: NTU-MC (Tan and Bond, 2011). The NTU-MC is a multilingual, multi-genre parallel corpus that is currently being tagged with the Princetong Wordnet, PWN (Fellbaum, 1998), the Japanese Wordnet (Isahara et al., 2008), the Chinese Open Wordnet (Wang and Bond, 2013) and the Wordnet Bahasa (Nurril Hirfana, Suerya and Bond, 2011) through the OMW.

The original design of PWN, which the other three wordnet projects are based on, includes only contentful/referential open class words: nouns, verbs, adjectives and adverbs. These classes of concepts are related by a coherent set of semantic relations (**hypernmy, antonymy, meronymy**, . . .) within this lexical resource, and can be used to sense tag the most part of any given corpus. Most wordnet projects follow PWN, hence restricting the classes of words available for annotation tasks using these resources.

During the annotation the NTU-MC, we found several classes of words that we wanted to annotate, but that were not included in the original set of concepts that were described and linked by wordnets. The first class we added was pronouns, including demonstratives, interrogative, personal and indefinite pronouns, and their associated modifiers (Seah and Bond, 2014). These were added primarily to allow cross-lingual linking of concepts, since it was often the case that a noun in one language would be linked to a pronoun in another. In this paper we will, similarly, motivate and describe our current efforts on further expanding wordets to include interjections/exclamatives and numeral classifiers.

2. Interjections

Interjections are generally perceived as words or phrases that constitute a whole linguistic act. That is they generally do not combine with other words in integrated syntactic constructions. They do not refer to events of individuals, but instead carry expressive meaning (Huddleston and Pullum, 2002). Following Jovanović (2004) and Ameka (1999), we use the term broadly, covering plain interjections, greetings and many more uses, described below.

While generally absent from semantic ontologies like the Princeton Wordnet (Fellbaum, 1998), interjections are abundant in many different corpora. In particular, any corpus that includes direct discourse is likely to contain interjections. Consider the following snippet from *The Adventure of the Speckled Band* (Conan Doyle, 1892), a short story currently being annotated as part of the NTU-MC:

(1) a. "Ah! That is suggestive. Now, on the other side of this narrow wing runs the corridor from which these three rooms open. There are windows in it, of course?"

    b. "Yes, but very small ones. Too narrow for anyone to pass through."

    c. "Thank you. That is quite settled" said he, rising and putting his lens in his pocket.

    d. "Hullo! Here is something interesting"

When we look at the currently available sense inventory for the highlighted lemmas (Table 1), we observe that none of the available senses are adequate to tag the highlighted interjections. While some of these concepts may actually

synset	lemmas	definition
00049758-r	**now**	indicates a change of subject or activity
15119919-n	**now**	the momentary present
00049220-r	**now**, at present	at the present moment
00049102-r	**now**	used to preface a command or reproof or request
00048475-r	**now**, today, nowadays	in these times
00048739-r	immediately, at once, right away, **now**, (...)	without delay or hesitation; with no time intervening
00049685-r	**now**	in the immediate past
00049433-r	**now**	in the historical present; at this point in the narration of a series of past events
07203900-n	**yes**	an affirmative
07229245-n	**thank you**	a conversational expression of gratitude
06632511-n	hello, hi, **hullo**, howdy, how-do-you-do	an expression of greeting

Table 1: Partial Sense Inventory (PWN3.0)

seem viable, for example 07203900-n to tag *yes* or 07229245-n to tag *thank you*, looking closely at their part-of-speech and examples shows that none of these senses are suitable. In this case, both are nominal concepts, so the synset 07203900-n (*yes*) could be used to tag, for example – *I was hoping for a yes* – , and 07229245-n to tag – *Few job candidates send thank yous* – where both underlined words are behaving as nouns.

According to their definitions, there are at least two interjections in the PWN (Table 2). These concepts received adverb as their part-of-speech. We can also find cases where it's slightly confusing which were the lexicographer's intentions. Table 3 shows three hyponyms of the concept 06629392-n, defined as *acknowledgment or expression of goodwill at parting*. Though not entirely clear, the nominal part-of-speech and the absence of other similar senses leads us to believe that these senses were also introduced in PWN for situations similar to the examples above. In this case, these nominal senses would be used to tag the following example: *Her good-mornings/good-afternoons/good-nights were hasty and mumbled.*

Therefore, currently, interjections are insufficiently represented in PWN. Interjections should, at least, be treated in a way they can be easily distinguished from other non-interjective senses.

In our efforts to enrich Wordnet with a new class of interjective senses, we followed the broad sense of interjection as defined, for example, by Ameka (1999). This broader sense captures many kinds of words and expressions. For example, those used:

- to express emotions such as surprise, disgust, etc. (e.g. *wow, ugh, yuk, gosh, ...*)

- in greetings, leave-taking, thanking, apologizing, etc. (e.g. *hello, thank you, goodbye, ...*)

- for swearing (e.g. *damn, shit, bite me, ...*)

- for responding (e.g. *yes, no, OK, yeah, you bet, ...*)

- onomatopoeically (e.g. *hush, boo, meow, oink, ...*)

To add these to the wordnet, we introduce a hierarchy inspired by that proposed by Jovanović (2004), where interjections are grouped according to their predominant pragmatic features. Jovanović (2004) proposes a set of 21 classes of expressive interjections, grouped by their emotional expressive potential along with an indefinite number of non-expressive interjections. The 21 expressive classes are: anger, annoyance, approval, contempt, delight, disgust, enthusiasm, fear, impatience, indignation, irritation, joy, pain, pity, pleasure, relief, sorrow, surprise, sympathy, triumph and wonder.

These classes are far from being exhaustive, and Jovanović (2004) recognizes the fact that many non-expressive interjections have a limited pragmatic range, and are difficult to group into bigger sets. Between these non-expressive interjections, we can find many interjections that concern greetings, leave-taking, thanking, apologizing, an extremely large range of onomatopoeia, and many others.

In the interest of the semantic annotation task, since we don't believe that some ambiguous lemmas could be reliably distinguished between some of these classes, we have chosen to merge some of the classes proposed by Jovanović (2004). We put forward a tentative adaptation, that may be subject to future review. Table 4 shows the concepts and current number of senses added for each interjection type we have created so far. We have reduced Jovanovic's 21 classes of expressive interjections to 11 classes, and are still going through the remaining data for non-expressive interjections, organizing it in different concepts. When compared to other concepts, some of these new interjective concepts have many members, as we are only enforcing near synonymy between senses.

All the concepts in Table 4 have been added to the OMW hierarchy, with senses for English and Mandarin Chinese. A new part-of-speech for non referential concepts was created with the tag 'x', which should be valid for a wide range of non semantically empty non-referential concepts such as interjections, classifiers, bound morphemes, particles, determiners and others.

We have also chose to standardize definitions for new interjective concepts, which take the form: *'an expression that is uttered ...'*.

4324

synset	lemmas	definition
00150351-r	right, right on	an interjection expressing agreement
00049889-r	now now	interjection of rebuke

Table 2: Existing Interjections (PWN3.0)

synset	lemmas	definition
06632671-n	morning, good morning	a conventional expression of greeting or farewell, used to wish someone a good morning
06632807-n	afternoon, good afternoon	a conventional expression of greeting or farewell, used to wish someone a good afternoon
06632947-n	good night	a conventional expression of farewell

Table 3: Ambiguous Concepts (PWN3.0)

We give interjections a flat hierarchy, and they are all examples of **utterance**, 07109847-n "the use of uttered sounds for auditory communication." and so linked with **exemplifies**. Because they do not denote any referent, they are not linked to other words by the hypernymy relation. This was inspired by Seah and Bond (2014), where pronouns were linked to the **pronoun** synset using **domain usage**, we renamed the relation to make it clear it has a different meaning from the other domain relations.

Whenever possible, we enrich this flatter hierarchy with links to other existing concepts (e.g. *good night* is linked to 15167474-n, night, *the period spent sleeping*) with **see also**. We give three examples here:

```
80000000-x (pain)
eng-lemmas    ah, oh, ouch, ow, wow, yipe, yow
cmn-lemmas    哎呀, 啊呀, 哎哟, 啊哟, 哎哟喂, 疼
definition    an expression that is uttered to show
              physical hurt or pain
exemplifies   15167474-n (utterance)
see also      14322699-n (pain, hurting)
```

```
80000001-x (general greeting)
eng-lemmas    aloha, ciao, g'day, good day, hallo, hal-
              loa, halloo, hallow, hello, hi, howdy,
              hullo, 'sup
cmn-lemmas    你好, 您好, 喂, 哎, 你吃了吗, 嗨, 咳
definition    an expression that is uttered as a general
              greeting, regardless of the time of day
exemplifies   15167474-n (utterance)
see also      06630017-n (greeting)
```

```
80000002-x (checkmate)
eng-lemmas    checkmate, mate
cmn-lemmas    将死
definition    an expression that is uttered during a
              game of chess to declare that the final
              winning move has taken place
exemplifies   15167474-n (utterance)
see also      00167764-n (checkmate)
```

We have worked through roughly half of the data pro-

vided by Jovanović (2004), but there is a long tail of non-expressive very specific synsets (like **checkmate** or **tally-ho**). We're still working through the big class of onomatopoeias and other specific interjections, but they were left out of this first batch of additions because they require some work to create individual concepts with detailed definitions and relations.

Although this abstract presents examples for only English and Chinese, we hope to soon expand these entries to include interjective senses in the Japanese Wordnet (Isahara et al., 2008) and the Wordnet Bahasa (Nurril Hirfana, Suerya and Bond, 2011). To accomplish this more effectively, we are looking into processing data from multilingual sources like omniglot.com and Wiktionary.

2.1. Exclamatory Pronouns

The terms exclamation or exclamative can be used for emphatic or expressive utterances in general. One of the formal features that characterize exclamatory constructions is the frequent co-occurrence of exclamatory pronouns and interjections, typically with interjections that convey surprise (Michaelis, 2001). Exclamatory pronouns (2) had been excluded by Seah and Bond (2014), following the policy of only including words with propositional content.

(2) a. Wow, <u>what</u> a beautiful house!

b. It's amazing <u>how</u> much this costs!

c. This is <u>such</u> an amazing view!

Out of these three, **such** appears in the Wordnet as both an adjective (01554230-a) and as an adverb (00147386-r). We added (or refined the existing entries for) the exclamative pronouns **what** and **how**. At the same time we added the adverbs **so** and **such** that are often used in exclamations (Huddleston and Pullum, 2002, pp 918–924), and the demonstrative degree specifiers **this** and **that**.

(3) a. <u>How</u> good was that cake!

b. The cake was <u>so</u> good!

c. The cake was <u>that</u> good!

Concept	Eng Senses	Total	Cmn Senses	Total
Surprise, Wonder	ah, blimey, . . .	58	啊, 哎哟, . . .	30
Pity, Sorrow	alas, dear, . . .	19	哎呀, 哦, . . .	6
Joy, Pleasure, Delight	boy, hurrah, . . .	27	哈哈, 嘻, . . .	9
Anger, Annoyance, Irritation, Indignation	damn, rats, . . .	42	啊呀, 行啦, . . .	10
Approval, Triumph, Enthusiasm	aha, hubba-hubba, . . .	10	行呀, ok, . . .	7
Contempt, Disgust, Impatience	bah, humph, . . .	59	呸, 唏, . . .	3
Pain	ouch, ow, . . .	7	啊呀, 啊哟, . . .	5
Sympathy	now, tsk, . . .	2	好啦, 唉, . . .	3
Fear	eeeek, oh, . . .	3	啊, 哎呀, . . .	3
Relief	whew, whoof	2	啊, 哎	2
Encouragement	attaboy, heigh, . . .	16	加油, 乖乖, . . .	5
Attention-Seeking	ahem, hoy, . . .	36	嗨, 嘘, . . .	5
Toasting	prosit, salut, . . .	10	干杯, 嘿, . . .	4
General Greetings	howdy, g'day, . . .	13	你好, 喂, . . .	7
Morning Greetings	morning, good morning	2	早上好	1
Afternoon Greetings	good afternoon, afternoon	2	下午好	1
Night Greetings	bonsoir, good evening	2	晚上好	1
General Farewells	ciao, bye-bye, . . .	21	再见, 拜拜, . . .	4
Night Farewells	good night, sweet dreams, . . .	5	晚安	1
Checkmate	mate, checkmate	2	将死	1
Total Number of Senses		338		108

Table 4: Interjective Senses

3. Classifiers

Most languages make use of classifiers (CL). For example, most languages have measure classifiers, similar to *a kilo of tea*, or group CLs, as in *a school of fish*. These two kinds of classifiers (i.e. measure and group classifiers) are sub-classes of nouns in many languages, and have been considered as nouns in PWN since its beginning. But some languages (e.g. Chinese, Japanese, Thai, etc.) have a much richer variety of classifiers including, for example, sortal classifiers: where a number of semantic features (e.g. physical, functional, etc.) must be shared between CLs and nouns they can quantify to license their usage. Consider the following examples from Chinese:

(4) 两　只　狗
　　 liǎng zhǐ gǒu
　　 2　　CL dog

　　 "two dogs"

(5) 三　台　电脑
　　 sān tái diànnǎo
　　 3　 CL computer

　　 "three computers"

(6) *三　只　电脑
　　 sān zhǐ diànnǎo
　　 3　 CL computer

　　 "three computers"

Examples (4), (5) and (6) show how the simple act of counting in Mandarin Chinese involves pairing up nouns with specific classifiers, if incompatible nouns and classifiers are put together then the noun phrase is infelicitous. In some languages (e.g. Chinese or Japanese), classifiers can be used as anaphoric references to elided nouns. This means that in (4) and (5), the words for *dog* and *computer* can be elided if the context and the semantic features of the CLs are enough to resolve the anaphoric reference. Therefore, though often redundant, CLs add to the semantic content of noun-phrases they are quantifying.

Different scholars treat this class of words very differently. Chao (1965), the traditional and authoritative native Chinese grammar, splits CLs into nine different classes. Cheng and Sybesma (1998) draw a binary distinction between **count-classifiers** and **massifiers**. Erbaugh (2002) splits CLs into three categories (**measure**, **collective** and **sortal classifiers**). Measure classifiers describe quantities (e.g. 'a bottle of', 'a mouthful of'), collective classifiers describe arrangement of objects ('a row of', 'a bunch of'), and sortal classifiers refer to a particular noun category (which can be defined, for example, by shape).

In our analysis, we follow the taxonomy previously proposed in Bond and Paik (2000), distinguishing between five major classes of classifiers: **sortal** (which classify the kind of the noun phrase they quantify); **event** (which are used to quantify events); **mensural** (which are used to measure the amount of some property); **group** (which refer to a collection of members); and **taxonomic** (which force the noun phrase to be interpreted as a generic kind).

In the Chinese portion of the NTU-MC corpus, classifiers make up roughly 2.5% of the words, with 260 types for 3,775 instances. These numbers are similar in the Japanese portion of the NTU-MC, and a bit lower in the Indonesian portion, mainly due to the fact that the usage of classifiers is optional in modern Indonesian. Motivated by their implicit semantics and relations with nouns, as well as a considerable presence of classifiers in corpora, we decided to extend

the wordnets for Chinese, Japanese and Indonesian with a new set of sortal classifier concepts.

Following the explanation given in Section 1, classifiers also receive the part-of-speech 'x' to mark their non-referentiality. And again, similar to what happened with interjections, classifiers were also given a standardized definition with the form " *a . . . classifier used . . . , such as . . .* ", where the kind of classifier, the general class of nouns they are used with, and one or more examples must be provided. All classifiers link to **classifier**, 06308436-n "a word or morpheme used in some languages in certain contexts (such as counting) to indicate the semantic class to which the counted item belongs." through **exemplify**, also in a flat hierarchy. They are not linked to other words by the hypernymy relation.

Instead, with the introduction of numeral classifiers, two new concept relations were also introduced: **classifies** and its reverse **classified_by**. These new relations are used to link classifiers with other concepts, i.e. nouns or verbs, showing that the use of a specific classifier has been licensed for that concept. A confidence score can be used to weight these relations making it, therefore, possible to choose a preferred classifier out of many classifiers linked to a single concept.
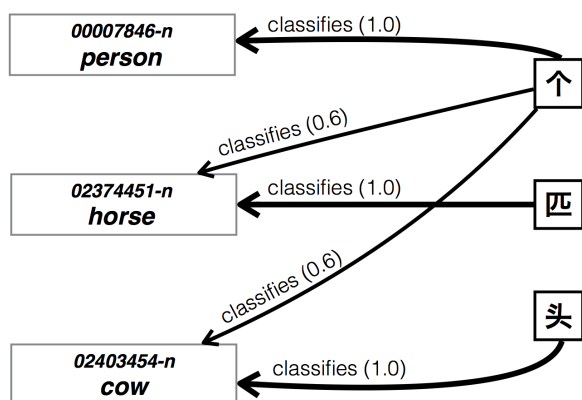

Figure 1: Example of classifier network

We plan to adapt the work presented in Morgado da Costa et al. (2016) to provide an automatic mapping between classifiers and other wordnet concepts. There is existing data that can be easily imported to Chinese Open Wordnet, and we will be looking into repeating the method for the other languages (i.e. Japanese and Indonesian).

These links can be immensely useful for second language learning, in tasks like machine translation, or any other domain where the knowledge of which classifier can be used with a specific noun or verb can be of value.

So far we have focused specifically on sortal classifiers. For Chinese, we started from the taxonomy provided by Gao (2010) to select sortal classifiers. For Japanese and Indonesian we worked from existing lists collated from corpora (Mok et al., 2012). Our initial efforts include the inclusion of 71 Chinese sortal classifiers in the Chinese Open Wordnet, 47 Japanese sortal classifiers in the Japanese Wordnet, and 30 Indonesian sortal classifiers in Wordnet Bahasa.

Here are the lists of classifiers collated for each language:

- Chinese: 埯, 把, 本, 部, 册, 层, 出, 道, 顶, 栋, 堵, 朵, 幅, 个, 根, 股, 管, 级, 家, 件, 间, 节, 具, 卷, 棵, 颗, 口, 粒, 辆, 绺, 枚, 门, 面, 名, 盘, 匹, 片, 篇, 撇, 期, 圈, 扇, 首, 艘, 所, 条, 贴, 挺, 头, 味, 尾, 位, 眼, 员, 盏, 张, 只, 支, 枝, 柱, 株, 幢, 尊, 座, 床, 轮, 份, 杆, 瓣, 块, 台

- Japanese: 振, 丁, 両. 人, 代, 件, 位, 体, 作, 個, 冊, 冠, 匹, 台, 合, 問, 回, 基, 床, 戸, 手, 把, 曲, 本, 束, 条, 枚, 校, 株, 棟, 機, 点, 発, 社, 粒, 組, 羽, 色, 行, 話, 軒, 通, 部, 階, 隻, 面, 頭

- Indonesian: orang, buah, ekor, biji, batang, lembar, pucuk, bilah, bidang, bentuk, bulir, utas, kuntum, butir, patah, pintu, batu, bengkawan, carik, eksemplar, helai, kaki, keping, labuh, laras, lonjor, siung, unit, untai, urat

Also, despite the obvious lexical overlap between languages like Chinese and Japanese, we found that the range of things that can be classified with a single classifier is very different across languages. For this reason, we gave up the expectation of being able to share sortal classifiers across languages. For the time being, regardless of the similarity of their coverage, each classifier concept is strictly monolingual.

We now present one example for each language:

80000003-x	
cmn-lemmas	把 *bǎ*
definition	a sortal classifier used with tools and objects with a handle, such as hammers, brooms, guitars or teapots
exemplifies	06308436-n (classifier)
classifies	03481172-n (hammer)
classifies	02906734-n (broom)
classifies	04398044-n (teapot)

80000004-x	
jpn-lemmas	振 *furi*
definition	a sortal classifier used for weapons with a blade, such as knifes, swords or daggers
exemplifies	06308436-n (classifier)
classifies	03624134-n (knife)
classifies	04373894-n (sword)
classifies	03158885-n (dagger)

80000005-x	
ind-lemmas	utas
definition	a sortal classifier used with threadlike objects, such as threads, ropes or wires
exemplifies	06308436-n (classifier)
classifies	04426788-n (thread)
classifies	04108268-n (rope)
classifies	04594218-n (wire)

4. Conclusions and Future Work

In this paper we have motivated and described the introduction of two non-propositional classes of words into wordnet. We introduced a new part-of-speech 'x' to be used by both interjections and classifiers, and for any non-referential concepts to be added in the future. We have added over 300 new interjective senses to English and over 100 to Chinese. As well as 71 sortal classifier senses to Chinese, 47 to Japanese and 30 to Indonesian.

Our future plans include not only expanding the coverage of interjections for English and Chinese, but also to expand the language coverage of these resources. We will be looking into automating mappings and processing data from multilingual sources like `omniglot.com` and Wiktionary.

The work presented here focused on sortal classifiers, in future work we will expand to cover other types of classifiers, such as group and event classifiers. We will also be looking into automatically populating the new classifiers with links to nouns and verbs, making this resource more valuable for multiple natural language processing tasks.

In the more distant future, we would also like to continue our effort to expand the coverage of classes of words to wordnets. Our next target will be prepositions. English prepositions are often translated as nouns in Chinese and Japanese: for example *between* is translated as *aida* "space or region between" in Japanese. We hope to build on existing semantic taxonomies for prepositions such as (Schneider et al., 2015).

We commit to release the data described in this work, by releasing them in the Open Multilingual Wordnet[1] and by attempting to merge them with the upstream wordnet projects. We hope to inspire other projects to proceed with similar extensions for different languages.

5. Acknowledgements

6. Bibliographical References

Ameka, F. K. (1999). Interjections. *Concise encyclopedia of grammatical categories*, pages 213–216.

Bond, F. and Foster, R. (2013). Linking and extending an open multilingual wordnet. In *51st Annual Meeting of the Association for Computational Linguistics: ACL-2013*, pages 1352–1362, Sofia.

Bond, F. and Paik, K. (2000). Re-using an ontology to generate numeral classifiers. In *18th International Conference on Computational Linguistics: COLING-2000*, pages 90–96, Saarbrücken.

Chao, Y. (1965). *A Grammar of Spoken Chinese*. University of California Press.

Cheng, L. L.-S., Sybesma, R., et al. (1998). Yi-wan tang, yi-ge tang: Classifiers and massifiers. *Tsing Hua journal of Chinese studies*, 28(3):385–412.

Conan Doyle, A. (1892). *The Adventures of Sherlock Homes*. George Newnes, London.

Erbaugh, M. S. (2002). Classifiers are for specification: Complementary functions for sortal and general classifiers in Cantonese and Mandarin. *Cahiers de linguistique-Asie orientale*, 31(1):33–69.

Fellbaum, C. (1998). A semantic network of English verbs. In Christine Fellbaum, editor, *WordNet: An Electronic Lexical Database*, chapter 3, pages 70–104. MIT Press.

Gao, H. (2010). Computational lexicography: A feature-based approach in designing an e-dictionary of Chinese classifiers. In *Proceedings of the 2nd Workshop on Cognitive Aspects of the Lexicon*, pages 56–65. Coling 2010.

Huddleston, R. and Pullum, G. K. (2002). *The Cambridge Grammar of English*. Cambridge University Press, Cambridge. (principal authors).

Isahara, H., Bond, F., Uchimoto, K., Utiyama, M., and Kanzaki, K. (2008). Development of the Japanese WordNet. In *Sixth International conference on Language Resources and Evaluation (LREC 2008)*, Marrakech.

Jovanović, V. Ž. (2004). The form, position and meaning of interjections in english. *FACTA UNIVERSITATIS-Linguistics and Literature*, (Vol. 3/11):17–28.

Michaelis, L. A. (2001). Exclamative constructions. *Language typology and language universals*, 2:1038–1050.

Mohamed Noor, N., Sapuan, S., and Bond, F. (2011). Creating the open Wordnet Bahasa. In *Proceedings of the 25th Pacific Asia Conference on Language, Information and Computation (PACLIC 25)*, pages 258–267, Singapore.

Mok, H., Gao, E., and Bond, F. (2012). Generating numeral classifiers in Chinese and Japanese. In *Proceedings of the 6th Global WordNet Conference (GWC 2012)*, Matsue. 211-218.

Morgado da Costa, L., Bond, F., and Gao, H. (2016). Mapping and generating classifiers using an open chinese ontology. *Proceedings of the 8th Global WordNet Conference (GWC2016)*.

Schneider, N., Srikumar, V., Hwang, J. D., and Palmer, M. (2015). A hierarchy with, of, and for preposition supersenses. In *Proceedings of The 9th Linguistic Annotation Workshop*, pages 112–123, June.

Seah, Y. J. and Bond, F. (2014). Annotation of pronouns in a multilingual corpus of Mandarin Chinese, English and Japanese. In *10th Joint ACL - ISO Workshop on Interoperable Semantic Annotation*, Reykjavik.

Tan, L. and Bond, F. (2011). Building and annotating the linguistically diverse NTU-MC (NTU-multilingual corpus). In *Proceedings of the 25th Pacific Asia Conference on Language, Information and Computation (PACLIC 25)*, pages 367–376, Singapore.

Wang, S. and Bond, F. (2013). Building the Chinese Open Wordnet (COW): Starting from core synsets. In *Proceedings of the 11th Workshop on Asian Language Resources, a Workshop at IJCNLP-2013*, pages 10–18, Nagoya.

---

[1]`http://compling.hss.ntu.edu.sg/omw/`