

# **Are groupings for word sense definitions that boost inter-annotator agreement the same as groupings based on language learner ability?**

Takahiko Aiba

Hirosaki University

## **Abstract**

This thesis work investigates whether or not word sense groupings are the same in two very different types of lexical resources: (i) OntoNotes, with groupings to assist preparing word sense training data for Natural Language Processing (NLP); and (ii) the English Vocabulary Profile, with groupings based on differing abilities of second language learners. To answer this question, 47 verbs were investigated by counting the correspondence of the word senses in the two resources. The investigation shows 81.6% of groupings of word senses in the two resources overlap. This degree of similarity suggests both resources should be of applicable value for their two distinct target domains: NLP and language teaching.

## **1. Introduction**

The goal of this investigation is to question whether there is overlap between groupings of word senses based on: (i) an attempt to achieve high inter-annotator agreements regarding sense selections, and (ii) observations about which senses of words language learners are able to use at different levels of ability. To reach this goal, groupings of word senses from two different resources are compared: OntoNotes and the English Vocabulary Profile.

This overview paper is organized as follows. Section 2 introduces the OntoNotes resource. Section 3 introduces the English Vocabulary Profile resource. Section 4 reports results from an investigation into the degree of alignment between the groupings of word senses from the two resources. Section 5 is an overall summary.

## **2. What is OntoNotes?**

Word sense ambiguity gives the Natural Language Processing (NLP) problem of needing to extract the appropriate word sense for a given instance of word use. The aim of OntoNotes (Weischedel et al. 2012) is to provide word sense disambiguation information through the annotation of a large corpus which is composed of many genres such as news, broadcast, talk shows, weblogs, usenet newsgroups, and conversational telephone speech.

The word senses in OntoNotes are based on WordNet word senses. WordNet is an online lexical database which is based on psycholinguistic theories of human lexical memories according to Miller et al. (1993). WordNet is an important resource for word sense disambiguation because of its wide lexical coverage, but the rate of inter-annotator agreement is below 70% for word senses

(Weischedel et al., 2017), so OntoNotes tried to create groupings of word senses which could improve inter-annotator agreement.

In order to explain how to group word senses in WordNet into word senses in OntoNotes, consider the verb “agree”. In WordNet there are seven word senses, while OntoNotes has four word senses. The word sense “agree1” in OntoNotes (“concur, be in accord”) is the grouping of two word senses from WordNet: (i) be in accord, be in agreement (WN=agree%2:32::00::), and (ii) achieve harmony of opinion, feeling, or purpose (WN=agree%2:32:04). This grouping of the sense definitions from WordNet leads to course-grained word sense definitions that improve annotator productivity and achieve inter-annotator agreements to the target of 90% on average.

### **3. What is the English Vocabulary Profile?**

Available on the Internet from <https://www.englishprofile.org/>, the English Vocabulary profile (Capel 2010) is a description from corpus evidence of attainment abilities for English based on the Common European Framework of Reference (CEFR). CEFR is a series of ‘can-do’ statements about language abilities for different levels, from beginner levels through to expert levels. Such information is useful for teachers, designers of course syllabi, publishers making textbook materials, examination boards preparing exams, etc.

As of 2012, the English Vocabulary Profile has just under 7,000 headwords. The task of creating the English Vocabulary Profile was to provide the CEFR level (A1, A2, B1, B2, C1, C2) for the individual sense meanings of words, fixed phrases, and idioms. The sense groupings to CEFR levels of words are supported by corpus evidence from the Cambridge International Corpus, the Cambridge Learner Corpus and the Cambridge English Corpus. Each of these corpora have sizes of over a billion words of naturally occurring data.

### **4. Investigation**

This section reports information about the degree to which the groupings of word senses obtained from OntoNotes align with the groupings of word senses obtained from the English Vocabulary Profile.

#### **4.1 The selection of data**

47 verbs were chosen from a list of verbs at random that contained at least two groupings of CEFR level senses in the English Vocabulary Profile dataset. The 47 verbs give a lot of word senses (polysemy), and some of them have entries for all CEFR levels (A, B, C) of the English Vocabulary Profile. This gives a dataset where there is a lot of opportunity for mismatches in the groupings of the senses between the two resources to become apparent, and therefore provides a strong test as to the degree of correspondence.

This gave verb data with 296 OntoNotes word senses. The task for finding the degree of correspondence became one of counting which word sense definitions in OntoNotes are corresponding with the CEFR levels of the English Vocabulary Profile.

#### 4.2 An example with perfect alignment

First of all, consider “agree” as an example. When it comes to “agree”, it is an example of perfect alignment. “agree” has 4 word senses in OntoNotes from 7 WordNet sense definitions. Taking a look at Table 1, two “agree” word sense definitions with WordNet sense keys were clustered into the single OntoNotes “agree1”. Also, three “agree” word sense definitions with WordNet sense keys were clustered into the single OntoNotes “agree2”. One “agree” word sense definition with a WordNet sense key was clustered into the OntoNotes “agree3”. One “agree” word sense definition with a WordNet sense key was clustered into the OntoNotes “agree4”. Table 2 shows which word sense definitions of OntoNotes correspond with which CEFR levels in the English Vocabulary Profile. Agree1 is 100% A2 level, Agree2 is 100% C1 level, agree3 is 100% B1 level and agree4 is 100% N/A. Agree4 is “be agreeable or suitable”, but the English Vocabulary Profile doesn’t have that word sense definition, so it is “N/A”. Even though it has N/A, all word sense definitions are just applied to one CEFR level of the English Vocabulary Profile. This means that, for this example, the clustering of word sense definitions in OntoNotes is the same as the clustering of word sense definitions in the English Vocabulary Profile.

sense key	definition	CEFR	Ontonotes
agree%2:32:00::	be in accord; be in agreement	A2	agree1
agree%2:32:01::	consent or assent to a condition, or agree to do som	B1	agree3
agree%2:42:00::	be compatible, similar, or consistent; coincide in th	C1	agree2
agree%2:42:01::	go together	C1	agree2
agree%2:42:03::	show grammatical agreement	C1	agree2
agree%2:40:02::	be agreeable or suitable	-	agree4
agree%2:32:04::	achieve harmony of opinion, feeling, or purpose	A2	agree1

Table 1.

Verb	Ontonotes	A1	A2	B1	B2	C1	C2	N/A
agree	agree1		2/2(100%)					
	agree2					3/3(100%)		
	agree3			1/1(100%)				
	agree4							1/1(100%)

Table 2.

### 4.3 An example without perfect alignment

So, do all word sense definitions in OntoNotes correspond to just one CEFR level of the English Vocabulary Profile? Focusing on “look”, it has 6 word sense groupings in OntoNotes (Table 3). However, at “look1”, 2/15(13.3%) is A1, 2/15(13.3%) is B1, 4/15(26.7%) is B2, 2/15(13.3%) is C1, and 5/15(33.3%) is N/A. This means, for this example, the word sense groupings in OntoNotes are not the same as the word sense groupings in the English Vocabulary Profile. Like this, 47 verbs were examined, with 296 word sense definitions in OntoNotes, and manual counts were made concerning which word sense definitions in OntoNotes applied to CEFR levels of the English Vocabulary Profile.

Verb	OntoNote:	A1	A2	B1	B2	C1	C2	N/A
look	look1	2/15(13.3%)		2/15(13.3%)	4/15(26.7%)	2/15(13.3%)		5/15(33.3%)
	look2		4/4(100%)					
	look3				1/3(33.3%)			2/3(66.7%)
	look4		1/3((33.3%)					2/3(66.7%)
	look5			1/4(25%)	1/4(25%)			2/4(50%)
	look6				1/2(50%)			1/2(50%)

Table 3

### 4.4 Overall results

47 verbs (296 word sense definitions in OntoNotes) were examined, and the results are interesting. First of all, the ratio of perfect results like “agree” was counted. As shown in section 4.3, “agree1” is 100% A2, “agree2” is 100% C1, “agree3” is 100% B1 and “agree4” is 100% N/A. 4/47 (8.5%) of verbs show perfect results (agree, ask, save, use). Of the 296 word sense definitions in OntoNotes, the ratio of 100% corresponding to a CEFR level of the English Vocabulary Profile is 83/296 (28%). N/A results are not included here. Only 28% is corresponding perfectly to a CEFR level of the English Vocabulary Profile, so does this mean that groupings for word sense definitions that boost inter-annotator agreement are unlike groupings based on language learner ability?

Looking at the word sense definition “know1” in OntoNotes, 1/3 (33.3%) is A1 and 2/3 (66.7%) is A2 (Table 4). Even though “know1” is separated into two levels, this separation is not a big difference because the corresponding CEFR levels are adjoining. So, let’s focus on this adjacency of groupings. Also, looking at the word sense definition “serve3” (Table 5), even though it is separated into two levels, 1/4(25%) is B1 and 3/4(75%) is C1, “serve3” is almost C1 level because it is above 50%. So, let’s also focus on whether or not each word sense definition in OntoNotes can get above the level of 50% corresponding with a CEFR level of the English Vocabulary Profile.

Verb	OntoNote:	A1	A2	B1	B2	C1	C2	N/A
know	know1	1/3(33.3%)	2/3(66.7%)					
	know2			4/5(80%)				1/5(20%)
	know3		1/1(100%)					
	know4			1/1(100%)				
	know5							1/1(100%)

Table 4

Verb	OntoNote:	A1	A2	B1	B2	C1	C2	N/A
serve	serve1					2/3(66.7%)	1/3(33.3%)	
	serve2					2/3(66.7%)		1/3(33.3%)
	serve3			1/4(25%)		3/4(75%)		
	serve4		2/4(50%)					2/4(50%)
	serve5						1/1(100%)	
	serve6							1/1(100%)
	serve7							1/1(100%)

Table 5

The ratio of above 50% corresponding word sense definitions with a CEFR level of the English Vocabulary Profile is 159/296 (53.7%). Looking at “catch5” of Table 6, even though “catch5” is separated into three levels, A2, B1, and N/A, the ratio of A2 and B1 in total is 66.6%. This takes the overall result of correspondence above 50%. 27 word sense definitions in OntoNotes don’t have a ratio of correspondence that is above 50%. However, in these 27 word sense definitions, 14 word sense definitions exceeded a ratio of correspondence that is above 50% if focusing on adjacent results and counting their sum. Given this results, 173/296 (58.4%) word sense definitions show a correspondence ratio that is above 50%. However, these “296” include N/A parts. Looking at the Table 2, “agree4” is 100% N/A. This means no data can be extracted, so if 100% N/A are removed from consideration, the number is 212 (84 word sense definitions are 100% N/A. From this, 173/212 (81.6%) word sense definitions are found to be above 50%. In response to this result, we can say that 81.6% of groupings for word sense definitions that boost inter-annotator agreement match groupings based on language learner ability.

Verb	OntoNote:	A1	A2	B1	B2	C1	C2	N/A
catch	catch1	4/8(50%)					2/8(25%)	2/8(25%)
	catch2			2/8(25%)	4/8(50%)			2/8(25%)
	catch3						2/6(33.3%)	4/6(66.7%)
	catch4	1/6(16.7%)			2/6(33.3%)	1/6(16.7%)		2/6(33.3%)
	catch5		2/6(33.3%)	2/6(33.3%)				2/6(33.3%)
	catch6							1/1(100%)
	catch7			1/3(33.3%)	1/3(33.3%)	1/3(33.3%)		
	catch8				1/3(33.3%)			2/3(66.7%)
	catch9							1/1(100%)

Table 6

**4.5 Summary of all results**

Table 7 gives a summary of all results.

Percentage of corresponding	296 word sense definitions	212 word sense definitions (not included N/A)
100%	83/296(28%)	83/212(39.2%)
above 50%	159/296(53.7%)	159/212(75%)
above 50% (included adjacent)	173/296(58.4%)	173/212(81.6%)

Table 7. Summary of all results

From the summary of all results in Table 7, we are able to conclude that 81.6% of groupings of word senses based on an attempt to achieve high inter-annotator agreement overlap with groupings made from observations about what language learners are able to use at differing levels of ability when it comes to focusing on above 50% (including adjacent results) corresponding with CEFR levels of the English Vocabulary Profile without N/A.

This data doesn't have a lot of cases that show 100% correspondence with CEFR levels of the English Vocabulary, but focusing on cases that reach a correspondence ratio that is above 50%, we find that 75% of groupings of word senses in OntoNotes are the same as groupings of CEFR levels for word senses in the English Vocabulary Profile even when adjacent results are not included. Overall, it seems reasonable to conclude that the data was able to show a high percentage of corresponding groupings of word senses in OntoNotes with groupings of word senses for CEFR levels of the English Vocabulary Profile.

## 5. Conclusion

This thesis work has investigated the correlation between groupings for word senses in OntoNotes and the English Vocabulary Profile. This was achieved by counting which word sense definitions in OntoNotes have corresponding CEFR based groupings in the English Vocabulary Profile. It was found that 81.6% of groupings for word sense definitions that boost inter-annotator agreement overlap with groupings based on language learner ability. This degree of similarity suggests both resources should be of applicable value for their two very distinct target domains: Natural Language Processing and language teaching. However, there are differences in the data content of the resources. This is because word senses in OntoNotes are based on word senses from WordNet which offer a very broad coverage, while word senses in the English Vocabulary Profile are based on what English learners were observed to know and actively use. Results regarding what the resources contain can vary greatly from word to word. As future work, it will be necessary to investigate English nouns in the same way as verbs in order to further consolidate the results of this survey.

## References

- CAMBRIDGE UNIVERSITY PRESS (2015). English Profile – The CEFR for English. Homepage: <https://www.englishprofile.org/> (accessed 2022/02/24).
- CAMBRIDGE UNIVERSITY PRESS. (2013). Introductory Guide to the Common European Framework of Reference (CEFR) for English Language Teachers. <https://www.englishprofile.org/images/pdf/GuideToCEFR.pdf>.
- Capel, Annette. (2010). A1-B2 vocabulary: Insights and issues arising from the English Profile Wordlists project. English Profile Journal, Volume 1, Cambridge University Press.
- Capel, Annette. (2012). Completing the English Vocabulary Profile: C1 and C2 Vocabulary. English Profile Journal, Volume 3, Cambridge University Press.
- Miller, G. (1993). Nouns in WordNet: A Lexical Inheritance System. In G. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. Miller. Introduction to Wordnet: An On-Line Lexical Database. Princeton University.
- Yu, Liang-Chih, Chung-Hsien Wu, and Eduard Hovy. (2008). OntoNotes: Corpus Cleanup of Mistaken Agreement Using Word Sense Disambiguation. *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, pages 1057–1064, Manchester, August 2008.
- Weischedel, Ralph, Eduard Hovy, Marcus Mitchell, Martha Palmer, Robert Belvin, Sameer Pradhan, Lance Ramshaw, and Nianwen Xue. (2017). OntoNotes: A Large Training Corpus for Enhanced Processing. In Joseph Olive, Caitlin Christianson, and John McCary. Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Language Exploitation, Springer.

Weischedel, Ralph, Sameer Pradhan, Lance Ramshaw, Jeff Kaufman, Michelle Franchini, Mohammed El-Bachouti, Nianwen Xue, Ann Houston. (2012). OntoNotes Release 5.0 with OntoNotes DB Tool v0.999 beta. <http://www.bbn.com/NLP/OntoNotes>.