Towards a principled approach to sense clustering – a case study of wordnet and dictionary senses in Danish

Bolette S. Pedersen[1], Manex Agirrezabal[2], Sanni Nimb[3], Sussi Olsen[4], Ida Rørmann[5]
University of Copenhagen [1,2,4,5] & The Danish Society for Language and Literature[3]
Njalsgade 136, DK-2300 Copenhagen S[1, 2, 4,5,] Christians Brygge 1, DK-1219[3]
bspedersen@hum.ku.dk, manex.aguirrezabal@hum.ku.dk, sn@dsl.dk, saolsen@hum.ku.dk, idaroermannolsen@gmail.com

Abstract

Our aim is to develop principled methods for sense clustering which can make existing lexical resources practically useful in NLP – not too fine-grained to be operational and yet fine-grained enough to be worth the trouble. Where traditional dictionaries have a highly structured sense inventory typically describing the vocabulary by means of main- and subsenses, wordnets are generally fine-grained and *unstructured*. We present a series of clustering and annotation experiments with 10 of the most polysemous nouns in Danish. We combine the structured information of a traditional Danish dictionary with the ontological types found in the Danish wordnet, DanNet. This constellation enables us to automatically cluster senses in a principled way and improve inter-annotator agreement and wsd performance.

1 Lexical resources and word sense disambiguation (WSD)

Dealing with finegrained lexical sense inventories in NLP is a challenging task. Selecting the correct sense in a specific context is incredibly hard when word meaning is richly described with subtle and detailed sense distinctions as found in most wordnets and lexica.

To this end, coarse-grained word-sense disambiguation has become a well-established discipline over the years. One way to obtain a coarse-grained sense inventory is to cluster existing inventories either manually or automatically (Peters el al. 1998, Lapata & Brew 2004, Alvez et al. 2008, Izquierdo et al. 2009, McCarthy et al. 2016).

In recent years, also so-called supersense tagging has become popular where WordNet's *first beginners*[1] are applied as a cross-lingual sense inventory. In recent experiments on Danish corpora we achieved state of the art results in both annotator agreement and automatic supersense tagging (Alonso et al. 2015 and 2015b, Pedersen et al. 2016). Nevertheless, our experiments also demonstrated that the inventory was not particularly well suited for our purpose. First of all, the inventory proved *too* coarse in a considerable number of cases (see Alonso et al. 2016 for a discussion), and secondly, the set did not facilitate annotations across part-of-speech as in the case of de-verbal nouns resulting in unbalanced annotations between nouns and verbs.

In the present work, we pursue a slightly different path by returning to the monolingually and corpus-defined sense inventory of our monolingual lexical resources, the Danish wordnet, DanNet, and The Danish Dictionary (Den Danske Ordbog, DDO) on the basis of which DanNet was originally compiled (Pedersen et al. 2009). Our aim is to further examine the potential of a principled method for sense clustering to be performed automatically on existing fully-fledged sense inventories. The basic idea is to combine the structured information of a traditional Danish dictionary with the ontological types found in the Danish wordnet, DanNet, and to develop clustering methods on this basis.

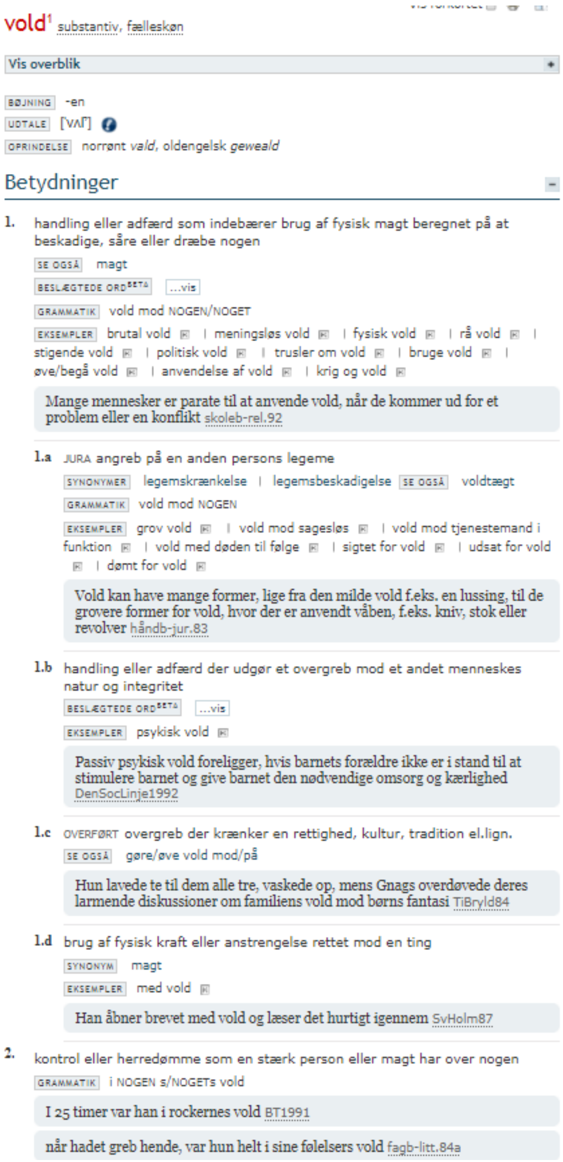
For our lexical sample study, we select 10 of the most polysemous nouns in Danish; we study how the senses are organized in DDO and DanNet and how they can be automatically clustered following two different principles: one allowing for clusters only within the same main sense, and one where also clustering of main senses are allowed except for the cases of homographs. For both sense inventories we perform manual annotation and word sense disambiguation using the LibLINEAR package and compare the results.

[1] Cf. https://wordnet.princeton.edu/man/lexnames.5WN.html

2 Sense organization in DDO and DanNet

2.1 Senses in DDO

Senses in DDO are according to normal convention organized in main- and subsenses as depicted in figure 1 for the lemma *vold* ('violence'):



vold[1] substantiv, fælleskøn

Vis overblik

BØJNING -en

UDTALE [ˈvʌlˀ]

OPRINDELSE norrønt *vald*, oldengelsk *geweald*

Betydninger

1. handling eller adfærd som indebærer brug af fysisk magt beregnet på at beskadige, såre eller dræbe nogen

SE OGSÅ magt

BESLÆGTEDE ORD BETA ...vis

GRAMMATIK vold mod NOGEN/NOGET

EKSEMPLER brutal vold | meningsløs vold | fysisk vold | rå vold | stigende vold | politisk vold | trusler om vold | bruge vold | øve/begå vold | anvendelse af vold | krig og vold

Mange mennesker er parate til at anvende vold, når de kommer ud for et problem eller en konflikt skoleb-rel.92

1.a JURA angreb på en anden persons legeme

SYNONYMER legemskrænkelse | legemsbeskadigelse SE OGSÅ voldtægt

GRAMMATIK vold mod NOGEN

EKSEMPLER grov vold | vold mod sagesløs | vold mod tjenestemand i funktion | vold med døden til følge | sigtet for vold | udsat for vold | dømt for vold

Vold kan have mange former, lige fra den milde vold f.eks. en lussing, til de grovere former for vold, hvor der er anvendt våben, f.eks. kniv, stok eller revolver håndb-jur.83

1.b handling eller adfærd der udgør et overgreb mod et andet menneskes natur og integritet

BESLÆGTEDE ORD BETA ...vis

EKSEMPLER psykisk vold

Passiv psykisk vold foreligger, hvis barnets forældre ikke er i stand til at stimulere barnet og give barnet den nødvendige omsorg og kærlighed DenSocLinje1992

1.c OVERFØRT overgreb der krænker en rettighed, kultur, tradition el.lign.

SE OGSÅ gøre/øve vold mod/på

Hun lavede te til dem alle tre, vaskede op, mens Gnags overdøvede deres larmende diskussioner om familiens vold mod børns fantasi TiBryld84

1.d brug af fysisk kraft eller anstrengelse rettet mod en ting

SYNONYM magt

EKSEMPLER med vold

Han åbner brevet med vold og læser det hurtigt igennem SvHolm87

2. kontrol eller herredømme som en stærk person eller magt har over nogen

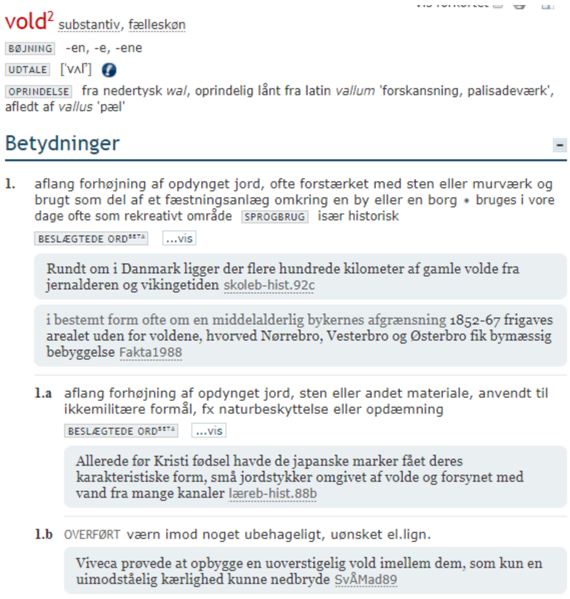
GRAMMATIK i NOGEN s/NOGETs vold

I 25 timer var han i rockernes vold BT1991

når hadet greb hende, var hun helt i sine følelsers vold fagb-litt.84a

Figure 1: Main- and subsenses in DDO of *vold* (violence, rampart, bank ..) in its violence sense.

In cases of homography where two lemmas take the same form without sharing etymology, two separate entries are established; in this case also an entry for the lemma *vold* in the sense of 'rampart' (Figure 2).



vold[2] substantiv, fælleskøn

BØJNING -en, -e, -ene

UDTALE [ˈvʌlˀ]

OPRINDELSE fra nedertysk *wal*, oprindelig lånt fra latin *vallum* 'forskansning, palisadeværk', afledt af *vallus* 'pæl'

Betydninger

1. aflang forhøjning af opdynget jord, ofte forstærket med sten eller murværk og brugt som del af et fæstningsanlæg omkring en by eller en borg • bruges i vore dage ofte som rekreativt område SPROGBRUG især historisk

BESLÆGTEDE ORD BETA ...vis

Rundt om i Danmark ligger der flere hundrede kilometer af gamle volde fra jernalderen og vikingetiden skoleb-hist.92c

i bestemt form ofte om en middelalderlig bykernes afgrænsning 1852-67 frigaves arealet uden for voldene, hvorved Nørrebro, Vesterbro og Østerbro fik bymæssig bebyggelse Fakta1988

1.a aflang forhøjning af opdynget jord, sten eller andet materiale, anvendt til ikkemilitære formål, fx naturbeskyttelse eller opdæmning

BESLÆGTEDE ORD BETA ...vis

Allerede før Kristi fødsel havde de japanske marker fået deres karakteristiske form, små jordstykker omgivet af volde og forsynet med vand fra mange kanaler læreb-hist.88b

1.b OVERFØRT værn imod noget ubehageligt, uønsket el.lign.

Viveca prøvede at opbygge en uoverstigelig vold imellem dem, som kun en uimodståelig kærlighed kunne nedbryde SvÅMad89

Figure 2: Main- and subsenses in DDO of *vold* (violence, rampart, bank ..) in its 'rampart' sense.

The overall principle for organizing senses within the same lemma follows Cruse (2000) by identifying different kinds of relations between main and subsenses:

- *Auto-hyponymy*: narrowed meaning with same hypernym, as in *to drink alcohol* as a subsense to *to drink*
- *Auto-superordination*: extended meaning with same hypernym as in *man* (male) vs *man* (person)
- *Auto-meronymy*: a part instead of the whole as in *door* meaning a piece of wood, metal or the like in contrast to *door* in the broader opening sense (as in *the door was made of wood* vs. *he closed the door*).
- *Auto-holonymy*: a whole instead of the part as in *body* meaning the whole body in contrast to *body* in the sense of the torso only.
- *Figurative*: sense where only part of the meaning (often its function) is derived from the core sense but used in a figurative/metaphorical context as in *window* in the sense *a window to the world*.

However, also the frequency of the senses (annotated in a set of randomly selected concordance lines (100-200 examples) from a balanced corpus of 40 mill. tokens (DDO Corpus (Norling-Christensen & Asmussen 1998)) was taken into consideration, as well as the communicative effect of the structure. The overall goal was to compile an 'easy to read' printed dictionary, especially by avoiding very deep sense structures. These two aspects considered, the relational principles defining subsenses to a particular main sense were not always followed. While figurative senses are typically described as subsenses to their main sense, frequent subsenses with a *non-figurative* relation (i.e. one the 4 'auto'-relations above) to the main sense were in fact in several cases described as an additional main sense instead of a subsense.

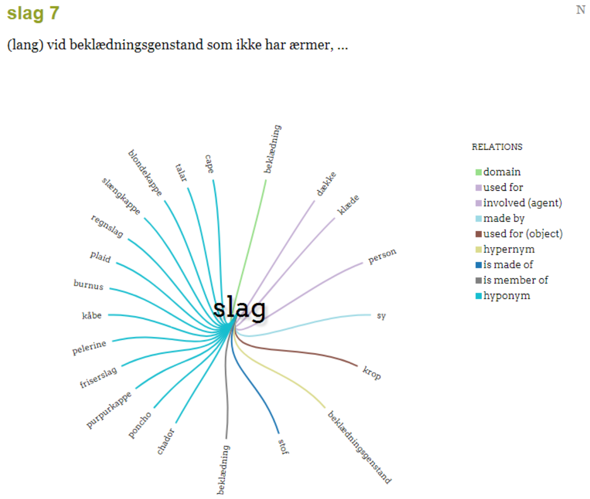
One example is the verb *æde* of which the first main sense describes the eating act of animals, whereas the second describes the eating act of humans, although the second is semantically derived from the first and therefore ought to be described as a subsense.

In other words, the semantic relatedness between word senses which we are looking for in order to be able to cluster senses in a principled way, is not always completely well reflected in the structure of the DDO entry. This inconsistency in structure – which is well-argued and also to our knowledge normal practice in lexicography – indicates why reuse of existing lexical resources in NLP is not just a straight-forward task. It also indicates that more than one experiment should preferably be performed; one where clusters are only established within main senses, and one where clustering also takes place across main senses (see Section 3).

2.2 Senses in DanNet

Senses in DanNet are organized in terms of synsets as in standard in wordnets (Fellbaum 1998). Each synset is assigned an ontological type based on EuroWordNets' top ontology, cf. Vossen 1999).

In contrast to the structure of a conventional dictionary where senses are typically organized in main and subsenses, the synsets that constitute the wordnet all have equal status. Further, each synset is inter-related to other synsets via semantic relations as shown in Figure 3.



slag 7

(lang) vid beklædningsgenstand som ikke har ærmer, ...

Figure 3: *Slag* in DanNet in its 'cape' sense and corresponding semantic relations

All synsets in DanNet are further assigned a complex ontological type following The EuroWordNet top-ontology (Vossen 1999) as depicted below in Figure 4 and 5.

- Origin
  - Natural
    - Living
      - Plant
      - Human
      - Creature
      - Animal
  - Artefact
- Form
  - Substance
    - Solid
    - Liquid
    - Gas
  - Object
- Composition
  - Part
  - Group
- Function
  - Vehicle
  - Representation
    - MoneyRepresentation
    - LanguageRepresentation
    - ImageRepresentation
  - Software
  - Place
  - Occupation
  - Instrument
  - Garment
  - Furniture
  - Covering
  - Container
  - Comestible
  - Building

Fig. 4: Ontological assignments to 1st Order Entities (cf. Vossen 1999:139)

- **SituationType**
    - Dynamic
        - BoundedEvent
        - UnboundedEvent
    - Static
        - Property
        - Relation
- **SituationComponent**
    - Cause
        - Agentive
        - Phenomenal
        - Stimulating
    - Communication
    - Condition
    - Existence
    - Experience
    - Location
    - Manner
    - Mental
    - Modal
    - Physical
    - Possession
    - Purpose
    - Quantity
    - Social
    - Time
    - Usage

Fig. 5: The EuroWordNet Top Ontology for 2nd and 3rd Order Entities cf. (Vossen et al. 1999:139)

Since our aim is to establish principled methods for sense clustering, it should be noted that the distinction between word senses is in several cases more fine-grained in DDO than the distinction between synsets in DanNet. This means that sometimes senses of the same word in DDO are in fact already members of the same synset in DanNet. These clusters were based on an idiosyncratic lexicographic judgment at the time of compilation of each synset but goes well in line with the more principled approach to sense clustering established here.

3 Establishment of clusters

Following the line of the discussion in Section 2, it does not seem appropriate just to collapse all DDO subsenses with its main sense; this would leave all metaphorical senses (which are indeed very frequent in our corpus) very poorly represented. We combine the information types from both resources: The DDO and DanNet and to this end, we perform three annotation experiments:

- Experiment 1 ('regular') where all main and subsenses are maintained.
- Experiment 2 ('clustered') where subsenses are clustered if they are of the same ontological type, and
- Experiment 3 ('clustered reduced') where also main senses are clustered if they are of the same ontological type.

Even if the ontology enables groupings of synsets which are ontologically similar (for instance artifact/part of artifact artifact/group of artifacts, person/groups of persons), we have in these experiments adopted a rather conservative approach and only clustered senses with the exact same ontological type.

Often a narrowed or an extended sense will have the same ontological type, in other cases a similar one. In contrast, figurative senses are typically of a completely different ontological type and are preserved with this method.

	Ex. 1 regular	**Ex. 2 clustered**	**Ex. 3 clustered reduced**
Selskab (company, party, association)	10	6	5
Plads (room, space, square, post)	13	9	6
Slag (battle, stroke, cape)	17	11	10
Skud (shot, shoot, dosis)	12	12	11
Skade (harm. injury, magpie, skate)	6	5	4
Kort (card, map)	10	4	3
Vold (violence, bank)	9	7	5
Hul (hole, gap)	14	11	8
Blik (look, glace, tin)	7	6	4
Model (model, pattern, design)	8	7	6

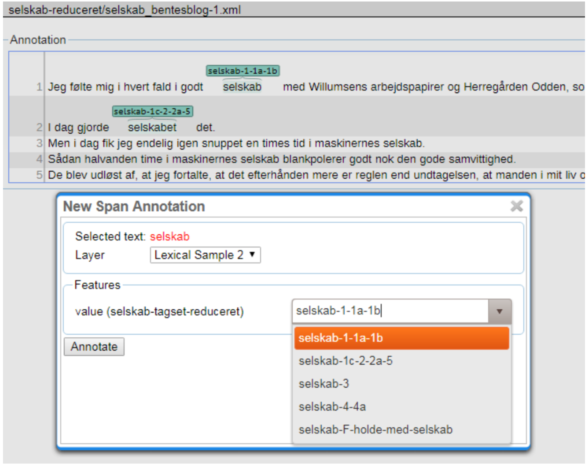
Table 1: Number of sense clusters in ex. 1- 3 excluding idiomatic expressions which do not cluster

4 Corpus and annotation

The texts selected for annotation have been extracted from the 45 million words CLARIN Reference Corpus (Asmussen 2012). This corpus comprises a wide variety of text types and domains: blog, chat, forum, magazine, Parliament debates (written down by professionals), and newswire, of which the latter constitutes 48 % of the entire corpus. In line with the Senseval approach (www.senseval.org), the number of annotated sentences for each noun varies according to the number of DDO senses of the noun (100 + 15*no. of senses), resulting in 177 to 535 sentences per noun.

It turned out that the otherwise very frequent nouns that we selected are not very frequent in social media texts, and since it is important for the project to have all text types including social media represented in the annotated data, all sentences from this text type that contained the noun in question were extracted from the corpus. Still to reach the specified number of sentences for each noun, we ended up with a majority of sentences from newswire texts.

For the annotation task we used the tool WebAnno (Yimam et al., 2013), which facilitates calculation of the inter-annotator quality and adjudication of the annotated files. For each occurrence of the word to be annotated, the annotators select a sense from the list of clustered senses in a drop down menu, see fig. 6.



selskab-reduceret/selskab_bentesblog-1.xml

Annotation

selskab-1-1a-1b

1 Jeg følte mig i hvert fald i godt selskab med Willumsens arbejdspapirer og Herregården Odden, so

selskab-1c-2-2a-5

2 I dag gjorde selskabet det.

3 Men i dag fik jeg endelig igen snuppet en times tid i maskinernes selskab.

4 Sådan halvanden time i maskinernes selskab blankpolerer godt nok den gode samvittighed.

5 De blev udløst af, at jeg fortalte, at det efterhånden mere er reglen end undtagelsen, at manden i mit liv o

New Span Annotation

Selected text: selskab

Layer: Lexical Sample 2

Features

value (selskab-tagset-reduceret): selskab-1-1a-1b

Fig 6: WebAnno annotation of *selskab* (company, party, association ..).

4.1 Annotation results

All sentences have been doubly annotated by advanced students and researchers and around 2% of the examples have been curated. The results from the three annotation experiments can be seen in Figure 7.

We apply Krippendorffs α (cf. Krippendorffs 2011) which calculates chance corrected agreement coefficients, i.e. sets off the fact that it is easier to agree on few tags than on many. Values range from 0 to 1, where 0 is perfect disagreement and 1 is perfect agreement. It is customary to require α ≥ .80 in most annotations tasks, however, for sense annotation where more tentative conclusions are still acceptable, we consider α ≥ .67 reasonable and useful. With this measure, as can be seen, only experiment 3 achieves 'acceptable' intercoder agreement for all words[2].

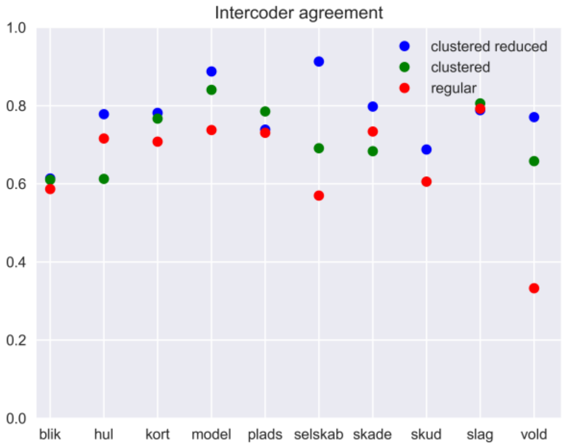


Fig. 7: Intercoder agreement (IA) (Krippendorffs α) in experiment 1-3

When curating 2% of the annotated material, we observed three kinds of discrepancies among annotators:

- *Plain errors*: Diverging annotations due to wrong pos tags or because the annotator had erroneaously skipped a word, for instance in cases with more than one lexical occurrence per sentence.
- *Incomplete or unclear tag set:* Diverging annotations in cases where a new/unconventional sense of the word was not covered by the tag set, or where the lexical description of a tag was unclear or blurred.
- *Underspecified examples*: Diverging annotations where the precise word sense

[2] It should be noted that we are here dealing with some of the most complex and polysemous words in Danish; i.e. agreement measures will most presumably differ for the rest of vocabulary.

could not be deduced from the isolated example (most divergences).

The annotators report that the annotations tasks are generally hard and that they are often in doubt, in particular when annotating with the full sense inventory where the distinctions are often very subtle. In contrast, they report that the generated clusters are somewhat more intuitive for them to work with, a fact which is reflected in an increased annotator agreement for the clustered senses, and also an increased agreement from experiment 2 to experiment 3.

One example is *selskab* (company, association, party) where groups of people doing things together can be more or less temporary resulting in different senses in the fine-grained experiment – but in only one cluster in the cluster experiments; a fact which increased agreement quite a lot. Further, where some clusters at first sight seem awkward, they often prove to ease annotation substantially. An example is *plads* which with its 'space' sense as a physical space/room/area is clustered with the 'square' sense as an urban, open area, square or field. Even though there are slightly different associations with these two senses it proves quite convenient to think of them as part of the same 'physical' cluster. Another noteworthy issue is the associations that we make regarding the digital universe, as in *plads på harddisken* (disc space) or *plads på skrivebordet* (space on the (computer) desktop). Are these examples abstract or concrete? Intercoder disagreement proves that annotators are in doubt.

In some cases, annotators report that clusters are really too coarse in experiment three, as exemplified with *kort* (card, map ..) where two very different kinds of artifacts are clustered (playing cards and maps) because they are of the same ontological type: Image Representation.

In a few cases, however, the ontologically based cluster separations seem to play a minor role. The ontological types of *fysisk skade* (physical injury/damage*)* and *psykisk skade* (psychological injury/damage) differ, where a psychological injury is more abstract and non-physical. But is this distinction really crucial? One can argue that the association of being injured, in either one of these ways, is more relevant to the context than whether the damage is physical or not, a fact which is demonstrated by quite a lot of underspecified corpus examples leading to disagreement among annotators because they had to choose one or the other.

Finally, the annotators meet a dilemma when dealing with metaphors. In the metaphor '*et skud i bøssen'* (one shot left), expressing one's only chance, the word *skud* is not the actual bullet, but rather the figurative sense of a chance. It is important to have a consensus of whether to stay inside the metaphorical picture and annotate within it, or whether to annotate with the actual intention. We chose consensus regarding the former solution, but still these cases lead to disagreement a number of times.

5 Word sense disambiguation using the LibLINEAR package

We also perform an experiment to see how empirical methods can perform in such hard tasks. The task is to disambiguate some specific words in a sentence (lexical sample task), and to see if there is any significant improvement of the prediction accuracies, when using clustered word senses.

The features that we use include a bag of lemmas of the whole sentence. We also include the next and previous four lemmas. These last elements are devised to disambiguate idiomatic expressions whose structure is mostly fixed.

As currently the data includes information from several annotators, training and evaluating Machine Learning classifiers is not straightforward. The main problem is the evaluation of a model. If two or more annotators have tagged a word in a sentence with diverging sense cluster tags, we consider it correct if an ML classifier classifies that instance as one of those sense clusters (either of them). This corresponds well to the fact that most divergences are caused by underspecified corpus examples. For learning, if two different annotators have tagged an instance, we consider it to be two different instances, resulting in some cases where we can have two instances with the same attributes, but with different outputs.

As the amount of data is limited, we decided to perform a 5-Fold Cross-Validation to check if the classifiers work sufficiently. We train a Linear Support Vector Machine for its robustness when used with a high number of features.

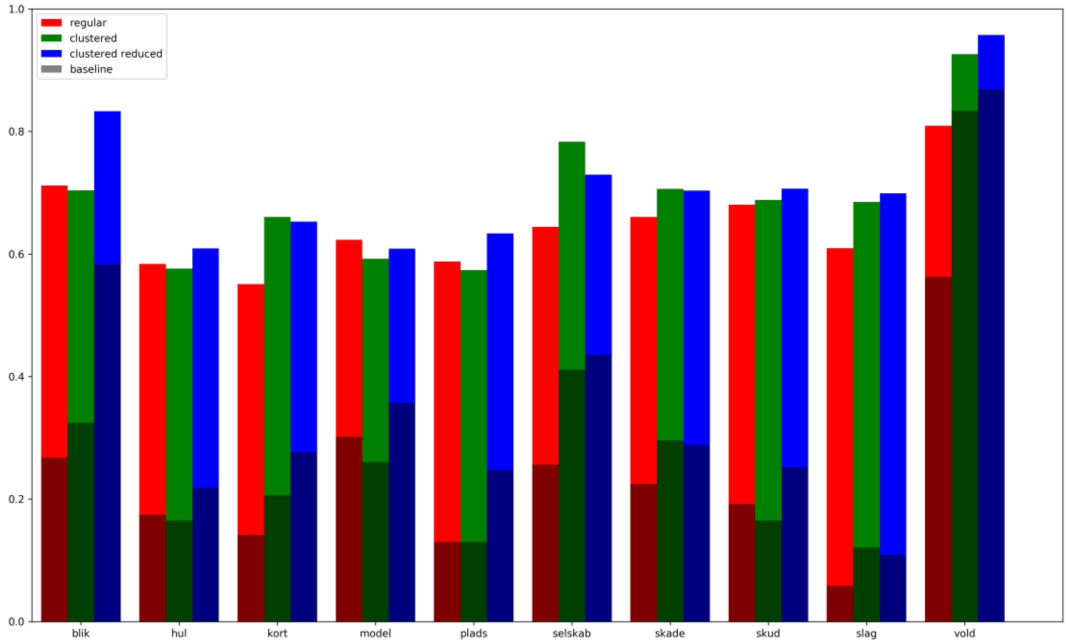


Fig. 8: Accuracies of the three experiments (regular, clustered, reduced clusters) compared to a baseline.

The toolkit that we employ is the well-known LibLINEAR package[3] (Fan et al. 2008), included in the module *scikit-learn* (Pedregosa et al. 2011) from Python.

Accuracies of the word disambiguation tasks with the three types of sense inventories compared to a baseline are provided in Figure 8. On average, reduced clusters can be seen to outperform the experiments with the more fine-grained sense inventories.

6 Concluding Remarks

In this paper we have examined how we can cluster noun senses in a principled way based on dictionary and wordnet information in combination (main and sub-senses versus ontological typ ing). We have dealt with some of the hardest and most polysemous nouns in Danish. We have further examined how systematically clustered noun senses influence inter-annotator agreement and automatic word sense disambiguation in a positive way, resulting in our last experiment (reduced clusters) in a sense inventory which seems actually manageable and well-functioning for both the annotators and the automatic disambiguation system. How our method will apply to verbs and adjectives is still an open question; for these word classes other information types than ontological typing may be more crucial.

It would also be interesting in future work to study how principled clustered based on lexicons and wordnets as presented in this paper compare to the word profiles that appear with word embeddings and sense induction methods.

Finally, only little space has however been left to discuss to which extent the meaning distinctions that are established by our clustering methods are actually relevant. Relevance depends on our purpose and on the kind of language technology service we are aiming at, where translation generally demands a high degree of detail, information search quite less, and question answering maybe something in between. In future work we would like to include relevance criteria as a more dominant feature encompassing also elements such as sense frequency and predominance information of senses; information which is however not directly accessible for Danish at the current stage.

Ackowledgements

Thanks to Anna Braasch, Sara Lee Naldal, and Ida Hauerberg Wolthers for assisting with the sense annotations.

[3] https://www.csie.ntu.edu.tw/~cjlin/liblinear/

References

Alvez, Javier, Jordi Atserias, Jordi Carrera, Salvodaor Climent, Egoitz Laparra, Antoni Oliver, German Rigau (2008). Complete and consistent annotation of wordnet using the top concept ontology. *LREC Proceedings* 2008.

Asmussen, J. (2012). CLARIN-Referencekorpus. Sprogteknologisk Workshop October 31, 2012, University of Copenhagen. http://cst.ku.dk/Workshop311012/sprogtekno2012.pdf

Cruse, D.A (2000). *Meaning in Language*. Oxford: Oxford University Press.

DDO = *Den Danske Ordbog*. (E. Hjorth et al). 2003-2005. Det Danske Sprog- og Litteraturselskab & Gyldendal, Copenhagen.

Fan, R. E., Chang, K. W., Hsieh, C. J., Wang, X. R., & Lin, C. J. (2008). LIBLINEAR: A library for large linear classification. *Journal of machine learning research*, 9(Aug), 1871-1874.

Fellbaum, Christiane (ed.). (1998). *WordNet: An Electronic Lexical Database*. MIT press.

Izquirdo, Rubén, Armando Suárez, and German Rigau. (2009). An empirical study on class-based word sense disambiguation. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics pp 389-397*. The Association for Computational Linguistics.

Krippendorff , K. (2011). Agreement and Information in the Reliasbility of Coding. In: *Communication Methods and Measures 5 (2)* pp: 93-112.

Lapata, Mirella and Chris Brew (2004). Verb Class Disambiguation Using Informative Priors. *Computational Linguistics*, 30(1): 45-73.

Martínez Alonso, Héctor; Anders Johannsen; Sussi Olsen; Sanni Nimb; Nicolai Hartvig Sørensen; Anna Braasch; Anders Søgaard; Bolette Sandford Pedersen. (2015). Supersense tagging for Danish. In *Proceedings of the 20th Nordic Conference of Computational Linguistics NODALIDA 2015, Linköping Electronic Conference Proceedings #109*, ACL Anthology, Linköping University Electronic Press, Sweden.

Martínez Alonso, Héctor; Barbara Plank; Anders Johannsen; Anders Søgaard. 2015b. Active learning for sense annotation. In *Proceedings of the 20th Nordic Conference of Computational Linguistics NODALIDA 2015, Linköping Electronic Conference Proceedings #109*, ACL Anthology, Linköping University Electronic Press, Sweden.

Martínez Alonso, Héctor; Anders Johannsen; Sanni Nimb; Sussi Olsen; Bolette Sandford Pedersen. 2016. An empirically grounded expansion of the supersense inventory. In *Proceedings of Global Wordnet Conference 2016*.

McCarthy, Diana, Marianna Apidianaki & Katrin Erk (2016). Word Sense Clustering and Clusterability. In: *Computational Linguistics, Vol. 42, no. 2.*

Norling-Christensen, Ole & Jørg Asmussen: The Corpus of The Danish Dictionary. *Lexikos. Afrilex Series, 8* 1998, 223–242

Pedersen, B.S, S. Nimb, J. Asmussen, N. Sørensen, L. Trap-Jensen, H. Lorentzen. 2009. DanNet – the challenge of compiling a WordNet for Danish by reusing a monolingual dictionary. *Language Resources and Evaluation, Computational Linguistics Series, pp.269-299.*

Pedersen, Bolette Sandford; Braasch, Anna; Johannsen, Anders Trærup; Martínez Alonso, Héctor; Nimb, Sanni; Olsen, Sussi; Søgaard, Anders; Sørensen, Nicolai. 2016. The SemDaX Corpus - sense annotations with scalable sense inventories. In *Proceedings of the 10th edition of the Language Resources and Evaluation Conference*. Portorož, Slovenia.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Vanderplas, J. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12(Oct), 2825-2830.
Peters, Wim, Yvonne Peters & Piek Vossen (1998.). Automatic sense clustering in EuroWordNet. In: *First International Conference on Language Resources & Evaluation 1998*, Granada, Spain.

Vossen, P (ed). 1999. *EuroWordNet, A Multilingual Database with Lexical Semantic Networks*. Kluwer Academic Publishers, The Netherlands.

Yimam, S.M., Gurevych, I., Eckart de Castilho, R., & Biemann, C. (2013). WebAnno: A Flexible, Web-based and Visually Supported System for Distributed Annotations. *Proceedings of ACL-2013*, demo session, Sofia, Bulgaria.