

A Collaborative Interlingual Index for interoperable wordnets

Piek Vossen, VU University Amsterdam, Netherlands

Francis Bond, Nanyang Technological University, Singapore

John McCrae, University of Bielefeld, Germany

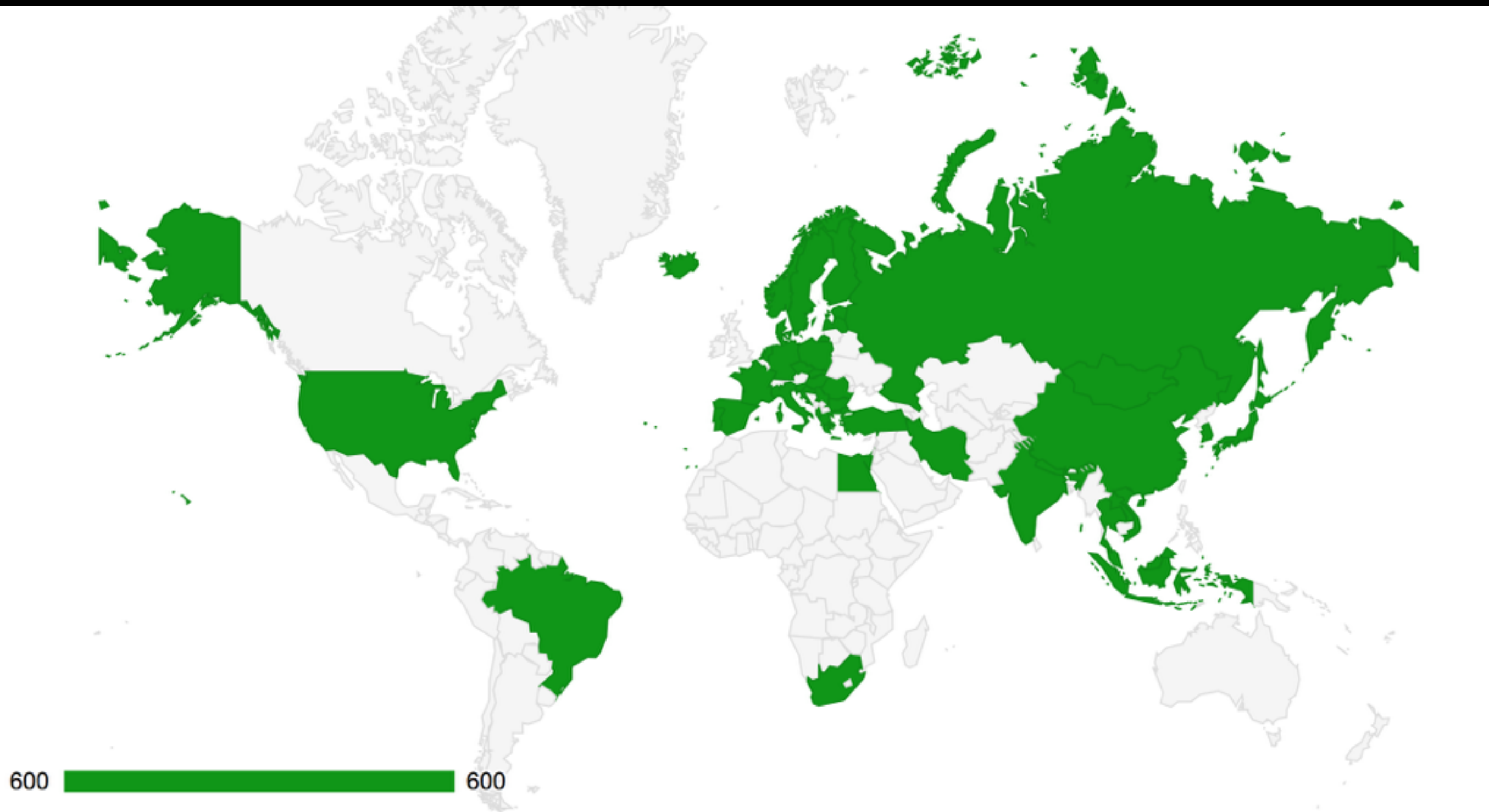
WordNet

- Relational model of meaning
- Synonyms represent a single concept
- Concepts are related through semantic relations
- Glosses support relations

25 years of wordnets

- Princeton WordNet since 1990 (1.5, 1.6, 2.0, 3.0)
- EuroWordNet
- Multiwordnet, BalkaNet
- Indowordnet, Asian wordnet, African wordnet
- Open Multilingual Wordnet
- Babelnet, Wiktionary

Global Wordnet Map



Bond & Paik (2012)

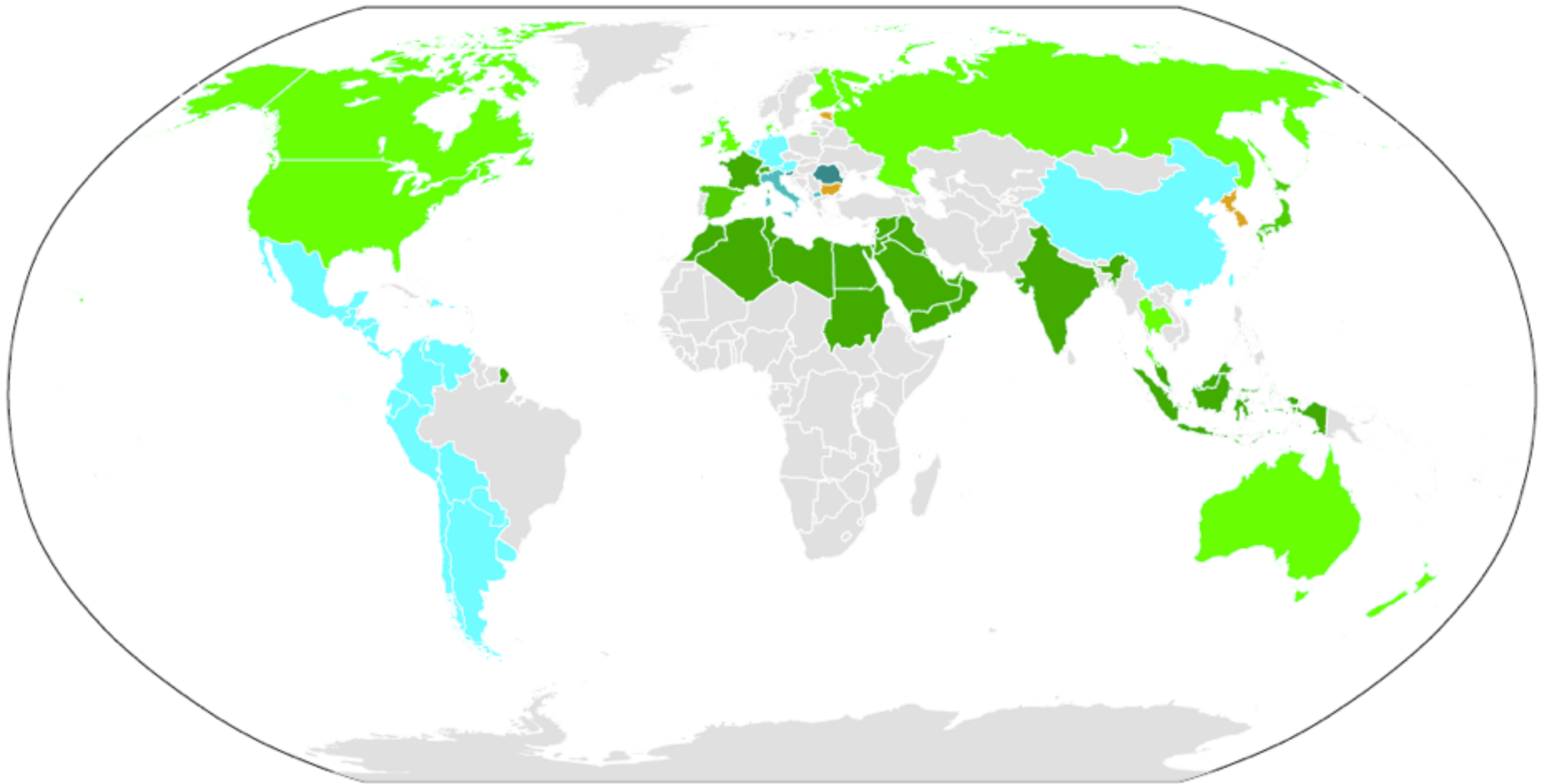


Figure 1: Map of Countries showing WordNet availability

Countries with **open source wordnets** are in green; **free for research wordnets** are in blue; **non free wordnets** are in brown. The lighter the color, the more synsets.

Citation counts from Google Scholar (accessed on 2011-09-23)

Francis Bond and
Kyonghee Paik (2012)
A survey of wordnets and
their licenses. GWC 2012.
Matsue. 64–71

Name	Language	# Synsets	Release	License	Citation	Count
Open Source						
Princeton WN ^{*€}	English	155,000	1991	WordNet	Fellbaum (1998)	6,821
FinnWordNet	Finnish	117,700	2010	WordNet	Lindén and Carlson. (2010)	0
Russian WN	Russian	117,000	2004	Wordnet	Balkova et al. (2008)	15
Thai Wordnet	Thai	73,593	2007	WordNet	Thoongsup et al. (2009)	4
DanNet [*]	Danish	65,000	2008	WordNet	Pedersen et al. (2009)	8
Japanese WN [*]	Japanese	57,000	2009	WordNet	Isahara et al. (2008)	24
Catalan WN [*]	Catalan	42,000	1999	GPL	Benítez et al. (1998)	17
LSG	Irish Gaelic	32,742	?	GNU FDL	http://borel.slu.edu/lsg/	—
Hindi WN	Hindi	28,687	?	GNU FDL	Jha et al. (2001)	10
WOLF	French	22,000	2009	Cecill-C [†]	Sagot and Fišer (2008)	22
Wordnet Bahasa [*]	Malay, Indonesian	20,000	2011	MIT	Nurril Hirfana et al. (2011)	—
Spanish WN ^{*⊙€}	Spanish	15,556	2006	LGPL	Farreres et al. (1998)	65
Catalan WN ^{*⊙€}	Catalan	15,556	2006	LGPL	Benítez et al. (1998)	17
Arabic WN [*]	Arabic	11,269	2008	CC BY SA	Black et al. (2006)	28
Hebrew WN [*]	Hebrew	5000	2006	GPL	Ordan and Wintner (2007)	0
Free for Research						
Chinese WN [*]	Chinese	115,424	2008	res/com	Xu et al. (2008)	0
KorLex ^{*⊙}	Korean	90,000	2007	res/com	Yoon et al. (2009) (nouns)	—
Spanish WN ^{*€}	Spanish	62,000	1999	res/com	Farreres et al. (1998)	65
Cornetto ^{*€}	Dutch	70,371	2009	res/com	Vossen et al. (2008)	19
GermaNet ^{*€}	German	69,594	2011	res/com	Kunze and Lemnitzer (2002)	52
MultiWN ^{*€}	Italian	38,877	2008	res/com	Pianta et al. (2002)	143
MWN [*]	Macedonian	33,276	2010	CC BY NC	Saveski and Trajkovski (2010)	0
Ro-WordNet [*]	Romanian	30,000	soon	no-deriv.	Tufiş et al. (2008)	9
Czech WN ^{*€}	Czech	29,000	1999	res/com	Pala and Smrř (2004)	34
SloWnet [*]	Slovene	20,000	2010	CC BY NC SA	Fišer and Sagot (2008)	13
Non Free (Available for Research)						
KorLex [*]	Korean	130,878	2007	res/com	Yoon et al. (2009)	5
Estonian ^{*€}	Estonian	47,000	—	ELRA	Kerner et al. (2010)	0
EuroWordNet					Vossen (1998)	728
Dutch	Dutch	44015	1999	ELRA	ELRA-M0016	
Spanish	Spanish	23370	1999	ELRA	ELRA-M0017	
Italian	Italian	48529	1999	ELRA	ELRA-M0018	
German	German	15,132	1999	ELRA	ELRA-M0019	
French	French	22,745	1999	ELRA	ELRA-M0020	
Czech	Czech	22,745	1999	ELRA	ELRA-M0021	
Estonian	Estonian	9,317	1999	ELRA	ELRA-M0022	
ItalWordNet	Italian	49,360	1999	ELRA	ELRA-M0018	
BasqWN	Basque	30,281	?	ELRA	Pociello et al. (2011)	0
BulNet ^{*⊙}	Bulgarian	23,715	2004	ELRA	ELRA-M0041 (Koeva, 2008)	3

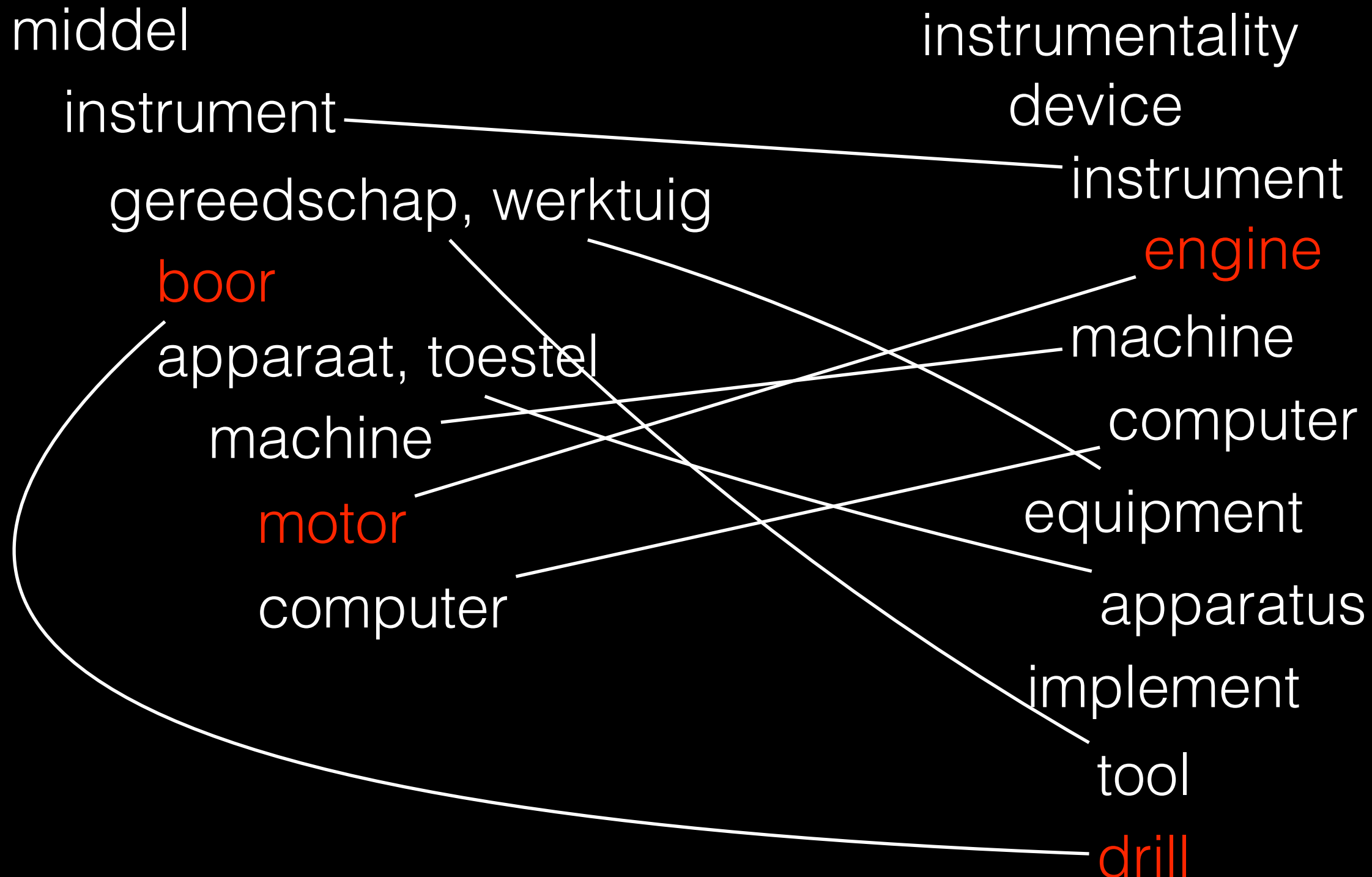
MERGE

Table 1: Catalog of WordNets

EuroWordNet design

- Each wordnet is structured according to the Princeton model: synsets & semantic relations between them, some with native glosses
- Synsets have equivalence relations to the Inter-Lingual-Index (ILI) = fund of concepts provided by Princeton (IndoWordnet uses Hindi as ILI)
- Different equivalence relations are allowed; mimic the wordnet-relations
- No semantic relations imposed on the ILI:
 - HUB for sameAs relations
 - relations are expressed in each wordnet

Merging wordnets



Status of the global net

- Wordnets built using different methods: merge or expand, manual or (semi-)automatic
- Different sets of relations were used
- Different interpretations of relations (with the same name)
- Different ways of defining synonyms (strict, loose)
- Different degrees of polysemy
- Differences in coverage

Status of the global net

- Linked to different versions of the English WordNet
- Released in different formats
- Using different license schemes
- Fixed Anglo-Saxon ILI: changes through English WordNet
- No central hosting of the ILI

From a global net
to
the Global Wordnet Grid
a platform for conceptual
interoperability

Global Wordnet Grid

- All wordnets linked to a single fund of concepts.
- Merge of concepts in all languages, not depending on English WordNet.
- Available as LLOD, one license: CC-BY-SA3.0, CC-BY-SA4.0.
- Adaptable by the wordnet-language community.
- As many wordnets as possible linked to the Grid and available as LLOD through the same license.

Motivation for GWG

- Platform for achieving linguistic & conceptual interoperability:
 - **Lexical** level:
 - Universals and idiosyncrasies in lexicalisation across languages
 - What is a word and what is a concept?
 - **Textual** level:
 - From Text to RDF: populating the LOD from textual data.
 - Understanding **any** language by machines in a similar way
 - Example NewsReader project: www.newsreader-project.eu

Wikinews corpus:
120 articles fully annotated

From text to RDF across mentions across language

NLP interpretation in NAF

RDF representation in SEM

A380 makes maiden **flight** to US. March 19, 2007.
The Airbus A380, the world's largest passenger plane, was set to **land** in the United States of America on Monday after a test **flight**. One of the A380s is **flying** from Frankfurt to Chicago via New York; the airplane will be **carrying** about 500 people.

El A380 hace su vuelo inaugural a los EEUU. 19 de marzo del 2007.
El Airbus A380, el mayor avión de pasajeros del mundo, aterrizó el lunes en los Estados Unidos de América, tras un vuelo de prueba. Uno de los A380s volará de Frankfurt a Chicago pasando por Nueva York; el avión llevará unas 500 personas.

Eerste vlucht van A380 naar V.S. 19-Mar-07. De Airbus A380, het grootste passagiersvliegtuig ter wereld, maakte zich maandag op om na een testvlucht te landen in de Verenigde Staten van Amerika. Een van de A380-machines vliegt van Frankfurt naar Chicago via New York en vervoert ongeveer 500 mensen.

```
<entity id="e3" type="LOCATION"> <!--United States-->
  <span><target id="t28","t29"/></span>
  <eRef conf="0.94" ref="dtpb:United_States" reftype="en"/>
</entity>
<predicate id="pr5"> <!--flying-->
  <eRef ref="fn:Bringing", "fn:Motion", "fn:Operate_vehicle", "fn:Ride_vehicle",
    "fn:Self_motion", "wn:01451842-v",
    <eRef ref="wn:01847845-v", "wn:01840238-v", "wn:02140965-v"/>
  <span><target id="t44"/></span>
  <role id="rl14" semRole="A1"> <!--One of the A380s-->
    <eRef ref="fn:Bringing@Theme", "fn:Motion@Theme", "fn:Operate_vehicle@Vehicle",
      "fn:Ride_vehicle@Theme", "fn:Self_motion@Self_mover"/>
  <span><target id="t39","t40","t41","t42"/></span>
</role>
<role id="rl15" semRole="AM-DIR"> <!--from Frankfurt-->
  <span><target id="t45","t46"/></span> </role>
<role id="rl16" semRole="AM-DIR"> <!--to Chicago-->
  <span><target id="t47","t48"/></span> </role>
<role id="rl17" semRole="AM-MNR"> <!--via New York-->
  <span><target id="t49","t50","t51"/></span>
</role>
</predicate>
<timex3 id="tmx2" type="DATE" value="2007-03-19"> <!--Monday-->
  <span><target id="w33"/></span>
</timex3>
```

```
<entity id="e2" type="ORGANIZATION"> <!--EEUU-->
  <span><target id="t9"/> </span>
  <eRef conf="0.99" ref="dbp:Estados_Unidos" reftype="es"/>
  <eRef conf="0.99" ref="dbp:United_States" reftype="en"/>
</entity>
<predicate id="pr3"> <span> <target id="t49"/> </span> <!--volar-->
  <eRef ref="fn:Bringing,"fn:Motion", "fn:Operate_vehicle", "fn:Ride_vehicle",
    "fn:Self_motion"/>
  <eRef ref="wn:01451842-v", "wn:01847845-v", "wn:01840238-v", "wn:02140965-v"/>
  <role id="rl8" semRole="arg0"> <!--Uno de los A380s-->
    <span> <target id="t45","t46","t47","t48"/> </span>
    <eRef ref="fn:Bringing@Agent", "fn:Motion@Theme", "fn:Operate_vehicle@Driver",
      "fn:Ride_vehicle@Theme", etc./>
  </role></predicate>
<timex3 id="tx3" type="DATE" value="2007-03-19"> <!--el lunes-->
  <span><target id="w30"/><target id="w31"/> </span>
</timex3>
```

wn:01451842-v;wn:01847845-v;wn:01840238-vwn:
02140965-v;wn:01941093-v

a sem:Event, fn:Bringing, fn:Motion,
fn:Operate_vehicle, fn:Ride_vehicle, fn:Self_motion;
rdfs:label "volar", "fly", "verlopen", "vliegen" ;

gaf:denotedBy

wikinews:english_mention#char=202,208>,
wikinews:english_mention##char=577,580>,
wikinews:dutch_mention##char=1034,1042>,
wikinews:dutch_mention#char=643,650>,
wikinews:dutch_mention#char=499,505>,
wikinews:dutch_mention#char=224,230>,
wikinews:spanish_mention#char=218,224>,
wikinews:spanish_mention#char=577,583> ;

sem:hasTime nwrtime:20070391;

sem:hasPlace

dbp:Frankfurt_Airport, dbp:Chicago ,
dbp:Los_Angeles_International_Airport,
nwr:airbus/entities/Chicago_via_New_York;

sem:hasActor

dbp:Airbus_A380,
nwr:airbus/entities/Los_Angeles_LAX ,
dbp:Frankfurt, nwr:airbus/entities/A380-machines.

wn:01451842-v;wn:01847845-v;wn:01840238-v;
wn:30-02140965-v

a sem:Event, fn:Bringing, fn:Motion,fn:Operate_vehicle,
fn:Ride_vehicle, fn:Self_motion;

rdfs:label "flight" ;

gaf:denotedBy

wikinews:english_mention##char=19,25,
wikinews:english_mention##char=174,180,
wikinews:english_mention##char=566,572;

sem:hasTime nwrtime:20070391;

sem:hasActor dbp:United_States_dollar, dbp:Qantas .



Predicate Matrix

Cross-lingual wordnets

Second Workshop
on
Natural Language Processing and Linked Open Data
(NLP&LOD2)

Collocated with RANLP 2015

11 September 2015

Hissar, Bulgaria

(<http://www.bultreebank.org/NLP&LOD2/>)

The Workshop: Topics of interest (1)

- NLP processing for LOD: reasons for low precision and inconsistencies
- Enhancing NLP applications with LOD
- Information extraction from LOD using NLP techniques
- Manipulating LOD (cleaning, adding information, deleting information, reconstructing facts) with NLP techniques
- LOD as a corpus
- Mapping LOD to common sense ontologies and language data

The Workshop: Topics of interest (2)

- Storing LOD in RDF bases
- Methodological and theoretical approaches to LOD
- Handling polysemy and metonymy of entities in LOD
- Incompleteness of LOD data
- LOD as unbalanced data through countries, cultures and topics of interest
- Insufficient reasoning in NLP and LOD
- Dynamics of LOD & NLP: versioning, replication, provenance, etc.

Important dates

- Submission deadline: 5 July 2015
- Notification of acceptance: 7 August 2015
- Camera-ready copies due: 22 August 2015
- Workshop date: 11 September 2015

Why adapt the ILI?

- Better mapping across languages following a merge approach.
- Bypass English gaps to map languages.
- Share resources across languages: ontologies, domains, sense-tagged corpora.
- Harmonise the semantics of wordnets across languages:
 - definition of synonymy, relations
 - similarity & relatedness measures
- Study universals and idiosyncrasies across languages.

Not lexicalized in English

- Cultural specific concepts
 - klunen_{Dutch} = walk on skates over land from one frozen water to another
 - Udhiyah_{Arabic} = slaughtering of a lamb during the period of Eid-Adha
- Pragmatic concepts
 - Gender variants:
 - Lehrer_{German} and Lehrerin_{German} = teacher

Not lexicalized in English

- Aspect variants in Slavic languages:
 - vypít_{Czech} = to drink up,
 - pít_{Czech} = to be drinking
- Lexical inclusion:
 - hilamos_{Tagalog} = to wash one's face
 - alevín_{Spanish} = small fish
 - bemahlen_{German} = paint something (obligatory obj)
- Compounds:
 - kindermeel_{Dutch} = flour for children
 - tarwemeel_{Dutch} = flour made of oats

Should there be a limit?

- NO! we could even include adjective-noun or verb-object pairs
- What about productivity?
- What about duplicate concepts?
- Productivity and compositionality need to be observed.
- Cross-lingual lexicalisation determines value:
 - what is linked across languages works, what is not linked disappears eventually

Why not use an ontology?

- Lack of consensus on ontologies.
- Axiomizing concepts in an ontology is more complex.
- Coverage of ontologies is still too low.
- Ontologies can be linked to the ILI as well: we can do both!
- ILI can feed ontologies with new concepts acquired from languages

What defines a concept?

- Keep it as simple as possible:
 - A unique IRI
 - English gloss
 - **LINKED** to a Synset in at least one wordnet through a sameAs relation
 - **LINKED** to a hypernym synset in a wordnet or being a well-defined top : —> *no orphans!!*

Design properties

- Glosses in other languages optional, English gloss required!
- Concept identifiers are unique and never deleted or modified.
- No further semantics is imposed on the ILI.
- Concepts have NO PART OF SPEECH
- Linking: gwa:sameAs or gwa:similarTo, no other links are allowed

Protocol for changes

- New concepts can be proposed but need to be linked to at least one wordnet.
- The only way to change a concept is by changing its English gloss.
- Concepts can be voted for and commented on.
- Concepts never disappear but can be ignored
- Duplication check: plagiarism!

ONION MODEL

- Kernel of fund consists of concepts that:
 - shared by all associated wordnets
 - sufficiently voted for (defined sufficient: nr., global/cultural spread)
 - axiomized through an ontology
 - passed the consistency checking
- Outer layer contains the most recently proposed new concepts linked to a single language.
- In between layers link to more languages, are moderated

ONION MODEL

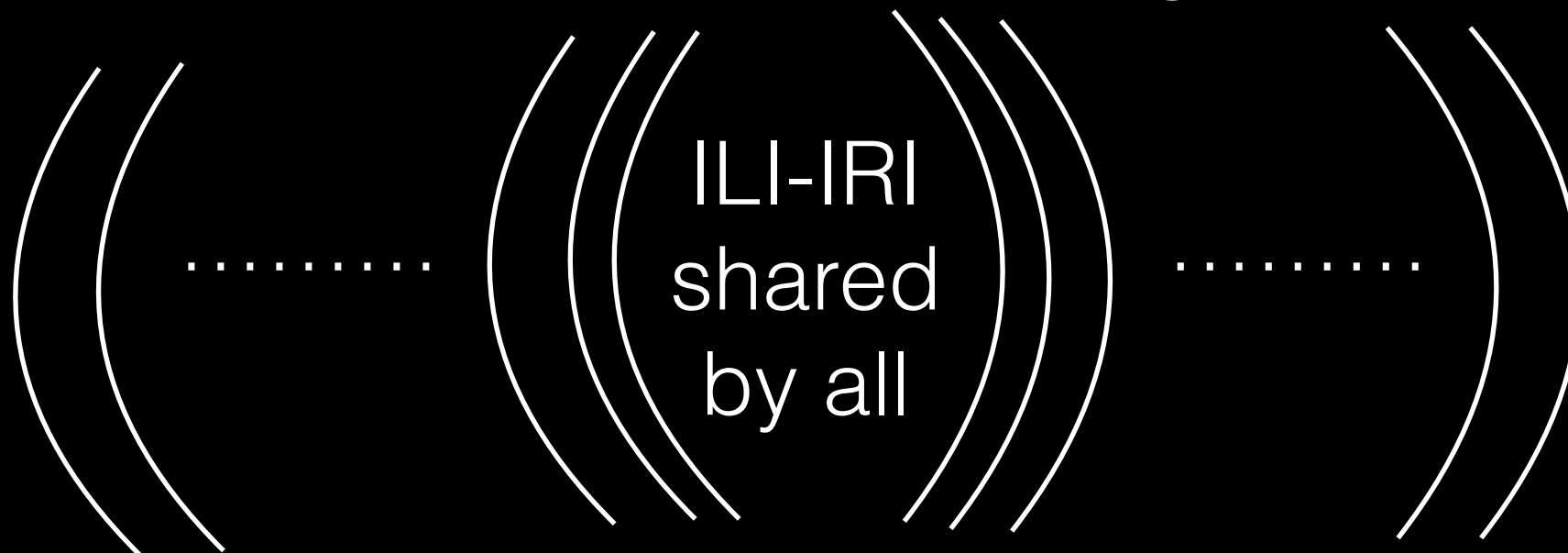


MODERATED



VALIDATED

VOTED



New
ILI-IRI
linked
to 1

ILI-IRI
shared
by 2

ONTOLOGIZED

ILI-IRI
shared
by 2

New
ILI-IRI
linked
to 1



R D F



ILI Community platform

- The community:
 - wordnet-members: wordnet builders and ontologizers
 - ili-moderator: moderator for the overall platform
 - wordnet-moderators: moderators for each language-wordnet community
- Every member belongs to a group associated with the wordnet of a language
- Add new members: ili moderator

ILI Community platform

- What can members do:
 - vote for concepts: the whole world?
 - comment on concepts: members
 - modify concepts: members
 - promote concepts to inner layers: language moderators

ILI Community platform

- Define your preference for alerts:
 - any modification
 - modification of concepts linked to your resource
 - modification of concepts related to concepts in your resource

Domain communities

- Specialists in domains should include their terminology for a languages and the corresponding concepts for the ILI
- Use the Collaborative environment for achieving cross-lingual interoperability in domains

Discrimination of concepts

- Gloss similarity can be used to find ILI concepts that are similar
- Semantic relations (of any linked wordnet) can be used to find glosses of siblings or co-hyponyms
- ILI groupings can be created for too fine-grained concepts

Technical implementation

- Specification of the data model to store concepts, editing history, provenance
- Github repository for hosting the ILI-concepts
- Social community software for voting and editing
- Export functions to WN-LMF, WN-RDF, LEMON,...
- Versioning
- Hosting

Consistency checking

- Check relations imposed on concepts from any linked wordnet or ontology
- How many hypernym matches across word nets?
- How consistent are antonymy relations?
- etc.....

Statistics

- Cross-wordnet sharing of concepts (gwa:sameAs & gwa:similarTo)
- Different parts-of-speech realisations
- Linkage in external wordnet:
 - subclass relation: top-leaf-middle, depth
 - other relations

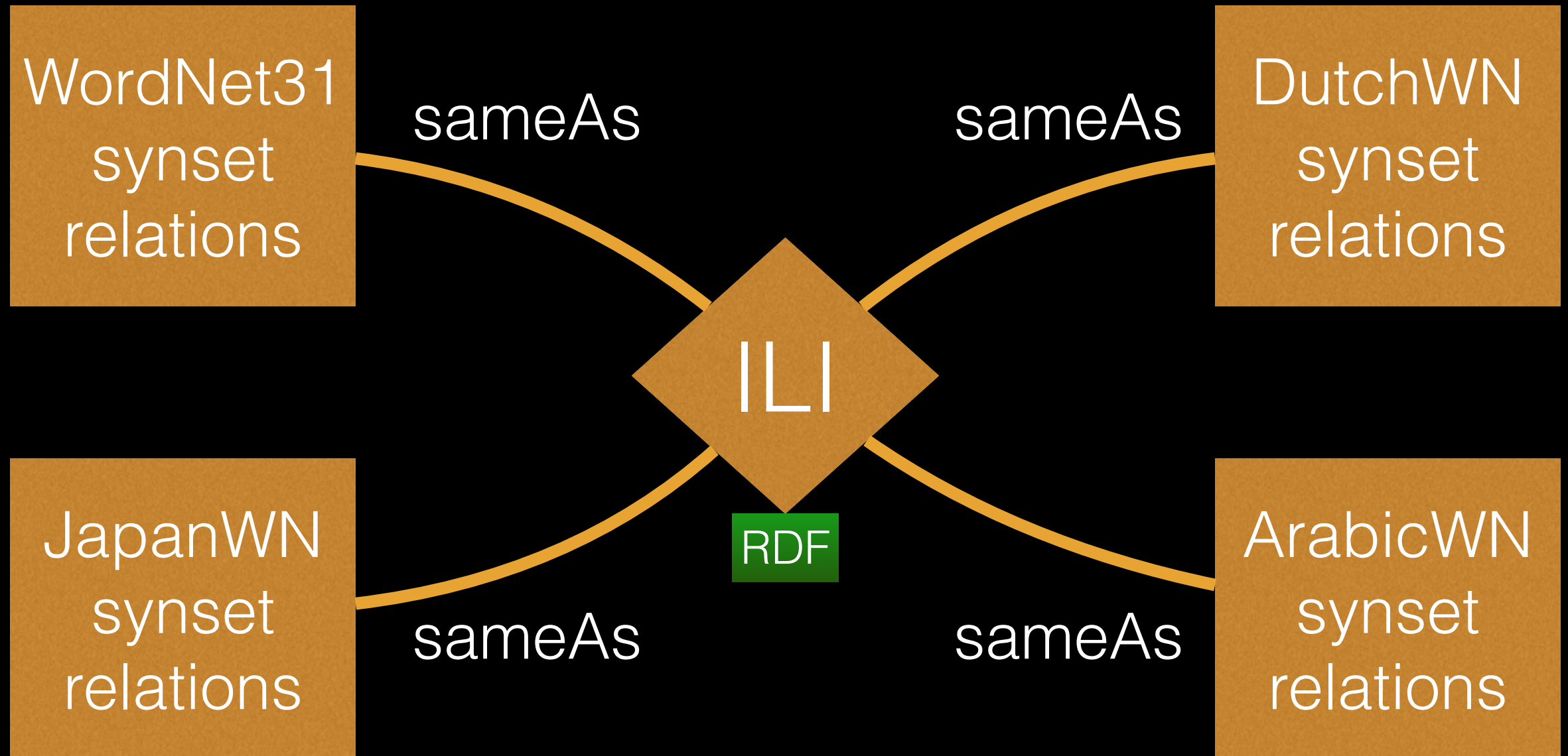
Bulk import of concepts

- BabelNet 2.5
 - Merge of WordNet, Open Multilingual WordNet, Wikipedia, OmegaWiki, Wikidata, Wiktionary
 - 9.3M synsets, 21,7M definitions
 - 7.7M images linked to synsets
 - creativecommons.org/licenses/by-nc-sa/3.0/

Project plan

- Design the data structure
 - Set up the repository with versioning and onion layers
 - Bulk import of ILI records
-
- Bulk import of linked wordnets
 - Develop tools for checking and gloss comparison/suggestion
 - Set up social community platform

Global Wordnet Grid



WN-LMF

RDF

LEMON

WN-LMF

RDF

LEMON

ILI as LLOD

- John McCrae
- <http://data.lider-project.eu/ili>

```

@prefix cc: <http://creativecommons.org/ns#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix vann: <http://purl.org/vocab/vann/> .
@prefix voaf: <http://purl.org/vocommons/voaf#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .

```

Wordnets

```

@prefix pwn30: <http://wordnet-rdf.princeton.edu/wn30/> .

```

```

@prefix ili: <http://globalwordnet.org/ili/> .
@base <http://globalwordnet.org/ili/ili.ttl>. #fixme

```

```

<> a voaf:Vocabulary ;
    vann:preferredNamespacePrefix "ili" ;
    vann:preferredNamespaceUri "http://globalwordnet.org/ili" ;
    dc:title "Global Wordnet ILI"@en ;
    dc:description "The shared Inter-Lingual Index for the global wordnets. It consists of
a list of concepts of instances with defintions, and their links to open wordnets."@en ;
    dc:issued "2014-07-30"^^xsd:date ;
    dc:modified "2014-07-30"^^xsd:date ;
    owl:versionInfo "2014 unstable release"@en ;
    dc:rights "Copyright Global Wordnet Association" ;
    cc:license <http://creativecommons.org/licenses/by/4.0> ;
    cc:attributionName "Global Wordnet Association";
    cc:attributionURL <http://www.gwa.org>;
    dc:contributor <http://john.mccr.ae> , <http://www3.ntu.edu.sg/home/fcbond/>,
<http://vossen.info> ;
    dc:publisher <http://www.gwa.org> .

```

```

<Concept> a owl:Class .

```

```

<Instance> a owl:Class .

```

```
<i1>    a    <Concept> ;
        owl:sameAs    pwn30:eng-00001740-a ; # able
        skos:definition "(usually followed by `to') having the necessary means or skill or
know-how or authority to do something"@en ;
        dc:source    pwn30:eng-00001740-a .

<i2>    a    <Concept> ;
        owl:sameAs    pwn30:eng-00002098-a ; # unable
        skos:definition "(usually followed by `to') not having the necessary means or skill or
know-how"@en ;
        dc:source    pwn30:eng-00002098-a .

<i3>    a    <Concept> ;
        owl:sameAs    pwn30:eng-00002312-a ; # abaxial, dorsal
        skos:definition "facing away from the axis of an organ or organism"@en ;
        dc:source    pwn30:eng-00002312-a .

<i4>    a    <Concept> ;
        owl:sameAs    pwn30:eng-00002527-a ; # adaxial, ventral
        skos:definition "nearest to or facing toward the axis of an organ or organism"@en ;
        dc:source    pwn30:eng-00002527-a .

<i5>    a    <Concept> ;
        owl:sameAs    pwn30:eng-00002730-a ; # acroscopic
        skos:definition "facing or on the side toward the apex"@en ;
        dc:source    pwn30:eng-00002730-a .

<i6>    a    <Concept> ;
        owl:sameAs    pwn30:eng-00002843-a ; # basiscopic
        skos:definition "facing or on the side toward the base"@en ;
        dc:source    pwn30:eng-00002843-a .
```


Open Dutch Wordnet-LMF

```
<LexicalEntry id="erker-n-1">
  <Lemma partOfSpeech="noun" writtenForm="erker"/>
  <Sense definition="naar buiten uitspringende raampartij" id="r_n-12592" provenance="cdb2.2_Manual+google_api" synset="eng-30-02813399-n"/>
</LexicalEntry>
<LexicalEntry id="ermitage-n-1">
  <Lemma partOfSpeech="noun" writtenForm="ermitage"/>
  <Sense id="o_n-104799465" provenance="cdb2.2_Auto" synset="eng-30-03515934-n"/>
</LexicalEntry>
<LexicalEntry id="ernst-n-11">
  <Lemma partOfSpeech="noun" writtenForm="ernst"/>
  <Sense id="t_n-219537486" provenance="google_api+wiktionary" synset="eng-30-07512315-n"/>
</LexicalEntry>
```

```
<Synset id="odwn-10-105981777-v">
  <Definition gloss="tot iets noodzaken" language="nl" provenance="odwn"/>
  <Definition gloss="something to compel" language="en" provenance="google-translate"/>
  <SynsetRelation provenance="pwn" relType="has_hyperonym" target="eng-30-02506546-v"/>
</Synset>
<Synset id="odwn-10-107584917-v">
  <Definition gloss="vernietigen" language="nl" provenance="odwn"/>
  <Definition gloss="destroy" language="en" provenance="google-translate"/>
  <SynsetRelation provenance="odwn" relType="has_hyperonym" target="eng-30-01645601-v"/>
</Synset>
```

ODWN synsets

```
<Synset id="eng-30-02813399-n" ili="i50597">
  <Definition gloss="naar buiten uitspringende raampartij" language="nl" provenance="odwn"/>
  <Definition gloss="a window that sticks out from the outside wall of a house" language="en" provenance="pwn"/>
  <Definition gloss="outwardly projecting windows" language="en" provenance="google-translate"/>
  <SynsetRelation provenance="pwn" relType="has_hyperonym" target="eng-30-04587648-n"/>
  <SynsetRelation provenance="pwn" relType="has_hyponym" target="odwn-10-106040031-n"/>
  <SynsetRelation provenance="pwn" relType="has_hyponym" target="eng-30-03855333-n"/>
</Synset>
```

PWN synsets

```
<Synset id="eng-30-06863751-n" ili="i72552">
  <Definition gloss="toonsoort" language="nl" provenance="odwn"/>
  <Definition gloss="any of 24 major or minor diatonic scales that provide the tonal framework for a piece of music" language="en" provenance="pwn"/>
  <Definition gloss="key" language="en" provenance="google-translate"/>
  <SynsetRelation provenance="pwn" relType="has_hyperonym" target="eng-30-06814870-n"/>
  <SynsetRelation provenance="odwn" relType="has_hyperonym" target="eng-30-04731497-n"/>
  <SynsetRelation provenance="pwn" relType="has_hyponym" target="eng-30-06864308-n"/>
  <SynsetRelation provenance="pwn" relType="has_hyponym" target="eng-30-06864192-n"/>
  <SynsetRelation provenance="pwn" relType="has_hyponym" target="eng-30-06864407-n"/>
</Synset>
```

ODWN-1.0

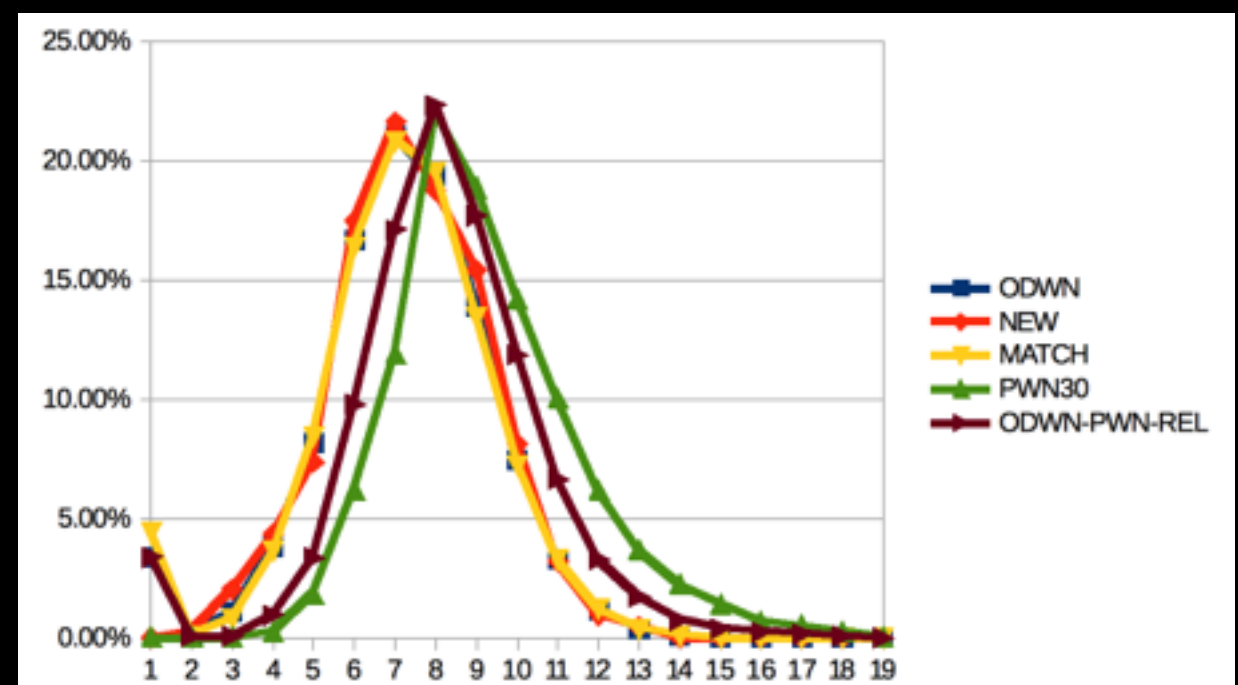
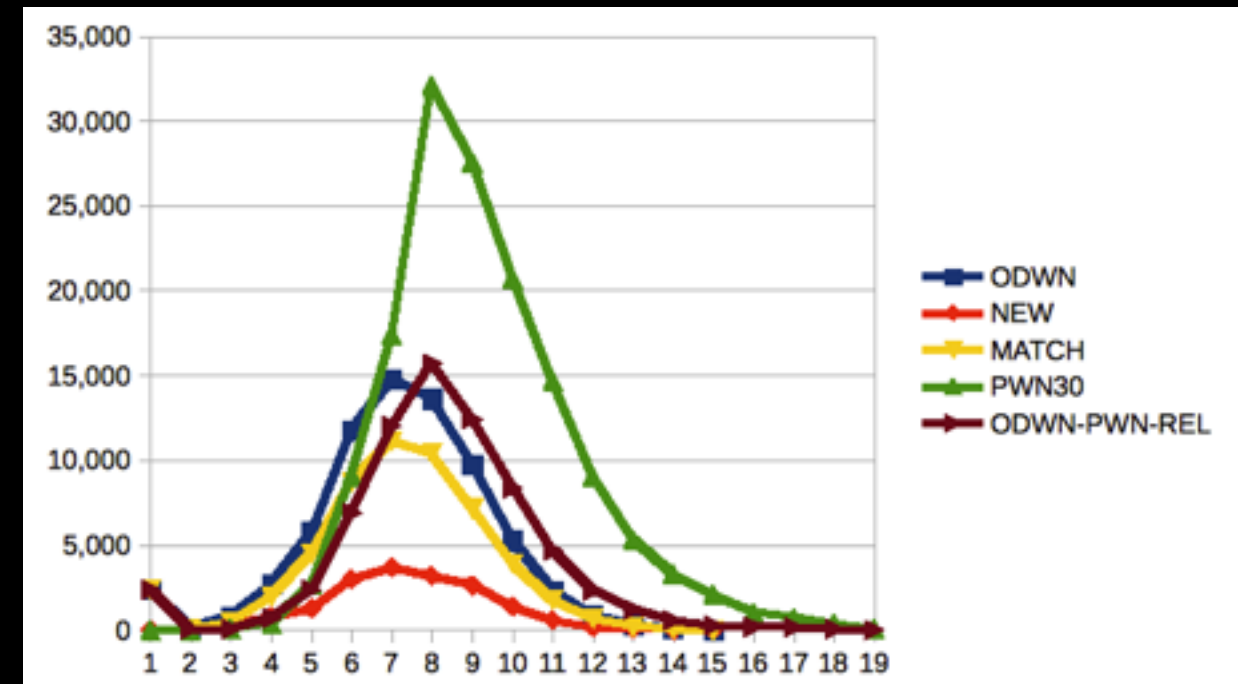
Open Dutch Wordnet 1.0	TOTAL	NOUNS	VERBS
Word forms	57,602	50,255	7,347
Lexical Units	94,140	78,612	15,528
Odwn Synsets with Dutch synonyms	21,636	15,992	5,644
Pwn Synsets with Dutch synonyms	19,980	15,706	4,274
Pwn Synsets without Dutch synonyms	75,376	66,409	8,967
Total	116,992	98,107	18,885

ODWN	DEFINITIONS
TOTAL	173,042
DUTCH	41,417
ENGLISH-PWN	95,190
ENGLISH-GoogleTrans.	36,435

Distribution over levels

NOUNS

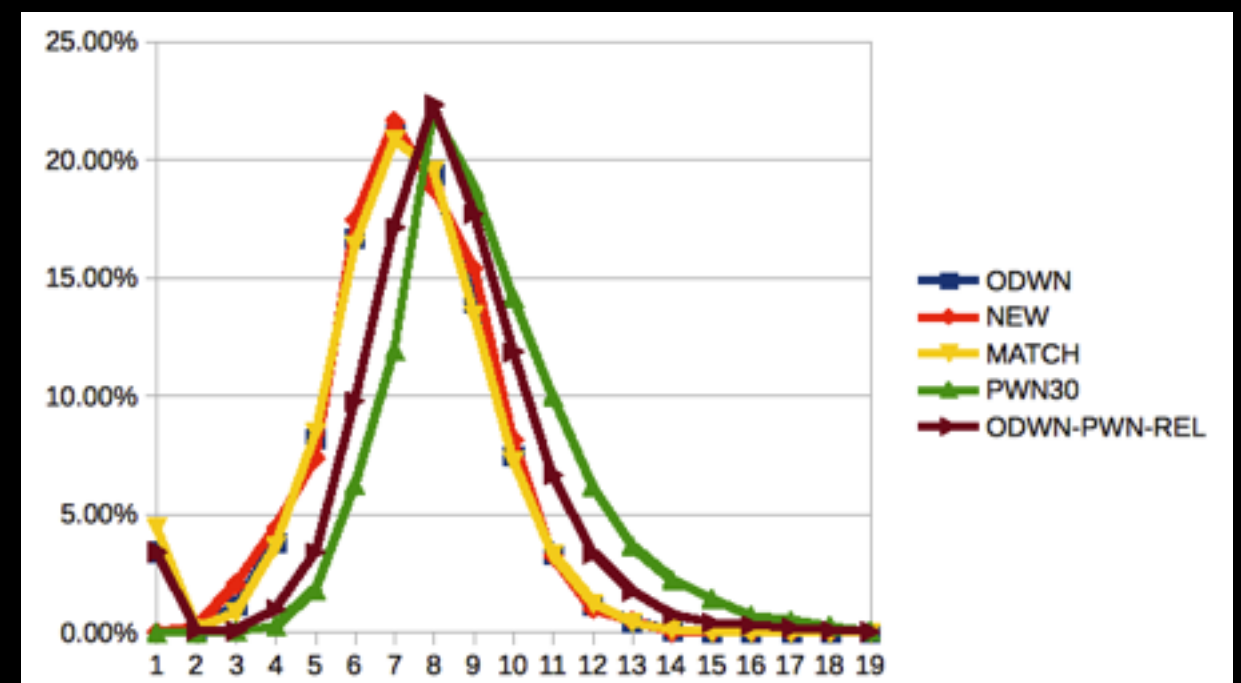
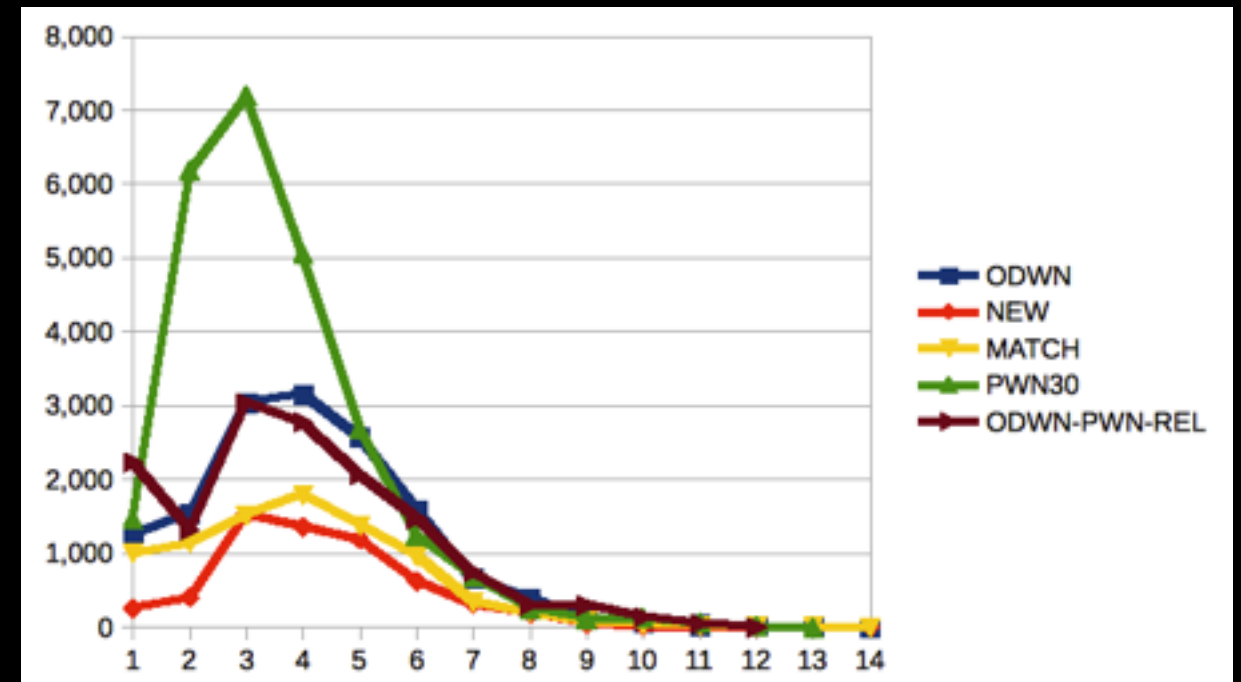
NOUNS	ODWN	NEW	MATCH	PWN30	ODWN-PWN-REL
1	2,385	1	2,384	1	2,385
2	110	44	66	4	58
3	786	343	443	32	52
4	2,684	739	1,945	379	691
5	5,759	1,248	4,511	2,658	2,368
6	11,702	2,953	8,749	9,075	6,876
7	14,775	3,655	11,120	17,393	12,036
8	13,611	3,176	10,435	32,124	15,682
9	9,775	2,604	7,171	27,564	12,441
10	5,251	1,373	3,878	20,681	8,333
11	2,311	545	1,766	14,624	4,680
12	820	166	654	9,006	2,318
13	288	78	210	5,320	1,202
14	61	3	58	3,273	531
15	1	0	1	2,041	272
16				991	192
17				720	134
18				367	55
19				59	13
TOTAL	70,319	16,928	53,391	146,312	70,319



Distribution over levels

VERBS

VERBS	ODWN	NEW	MATCH	PWN30	ODWN-PWN-REL
1	1,253	253	1,000	1,465	2,229
2	1,542	405	1,137	6,174	1,314
3	3,033	1,517	1,516	7,192	3,032
4	3,149	1,355	1,794	5,054	2,772
5	2,564	1,189	1,375	2,681	2,057
6	1,587	618	969	1,242	1,470
7	650	306	344	688	750
8	387	188	199	251	294
9	119	40	79	110	300
10	53	8	45	121	135
11	28	9	19	55	39
12	14	5	9	13	3
13	14	2	12	1	
14	2	1	1		
TOTAL	14,395	5,896	8,499	25,047	14,395



Matching strategies

- 21,636 ODWN synsets:
 - 7,489 match in another sense
 - 14,147 without match in any sense
- Google translate of lemmas
- Google translate of Dutch glosses

Complicated sense matches

```
<LexicalEntry id="ergernis-n-1">
  <Lemma partOfSpeech="noun" writtenForm="ergernis"/>
  <Sense definition="irritatie" id="r_n-12580" provenance="google_api+wiktionary" synset="eng-30-05832264-n"/>
</LexicalEntry>
<LexicalEntry id="ergernis-n-2">
  <Lemma partOfSpeech="noun" writtenForm="ergernis"/>
  <Sense id="r_n-12580_sub_1" provenance="cdb2.2_None" synset="odwn-10-104952541-n"/>
</LexicalEntry>
<LexicalEntry id="ergernis-n-3">
  <Lemma partOfSpeech="noun" writtenForm="ergernis"/>
  <Sense id="r_n-12580_sub_2" provenance="google_api+wiktionary" synset="eng-30-14406573-n"/>
</LexicalEntry>
<LexicalEntry id="ergernis-n-4">
  <Lemma partOfSpeech="noun" writtenForm="ergernis"/>
  <Sense id="r_n-12580_sub_3" provenance="google_api+wiktionary" synset="eng-30-07518261-n"/>
</LexicalEntry>
<LexicalEntry id="ergernis-n-5">
  <Lemma partOfSpeech="noun" writtenForm="ergernis"/>
  <Sense id="r_n-12580_sub_4" provenance="google_api+wiktionary" synset="eng-30-00418394-n"/>
</LexicalEntry>
<LexicalEntry id="ergernis-n-6">
  <Lemma partOfSpeech="noun" writtenForm="ergernis"/>
  <Sense id="r_n-12580_sub_5" provenance="google_api+wiktionary" synset="eng-30-07551418-n"/>
</LexicalEntry>
```

Translations of lemmas

1. Translate synonyms to English with Google-Translate
2. Look-up all translations in WordNet-PWN
 1. If monosemous or single synset take it
 2. else get *shared* synsets and hypernyms and check for matches with the ODWN hypernyms
 1. If match take it
 2. else get the synsets with highest similarity score (Leacock & Chodorow 1998) and above 2.0

NOT AN ENTRY IN WordNet-PWN

odwn-10-100326505-n=[zit:3] TRANS odwn-10-100326505-n=[is]
odwn-10-100324629-n=[overheidssector:1] TRANS odwn-10-100324629-n=[public sector]
odwn-10-100323103-n=[gloria:1] TRANS odwn-10-100323103-n=[Gloria]
odwn-10-100321930-v=[schurken:1] TRANS odwn-10-100321930-v=[thugs]
odwn-10-100319553-n=[tegenkracht:1] TRANS odwn-10-100319553-n=[to force]
odwn-10-100318571-n=[MBO:1] TRANS odwn-10-100318571-n=[MBO]
odwn-10-100318143-v=[heropenen:1] TRANS odwn-10-100318143-v=[reopening]
odwn-10-100317175-v=[omwaaien:1] TRANS odwn-10-100317175-v=[blown over]
odwn-10-100316580-n=[staalkabel:1] TRANS odwn-10-100316580-n=[SWR]
odwn-10-100310993-n=[waterkracht:1] TRANS odwn-10-100310993-n=[hydropower]
odwn-10-100308232-v=[verzoeken:2] TRANS odwn-10-100308232-v=[requests]
odwn-10-100305471-n=[bouwpremie:1] TRANS odwn-10-100305471-n=[building premium]
odwn-10-100305105-n=[duivelswerk:1] TRANS odwn-10-100305105-n=[devil's work]
odwn-10-100304724-n=[milieuramp:1] TRANS odwn-10-100304724-n=[environmental disaster]
odwn-10-100304639-n=[natuurpad:1] TRANS odwn-10-100304639-n=[nature trail]
odwn-10-100303840-n=[symptoombestrijding:1] TRANS odwn-10-100303840-n=[symptoms]
odwn-10-100303825-v=[set uppen:2,set-uppen:2] TRANS odwn-10-100303825-v=[set upping, set up data]
odwn-10-100303085-n=[wolindustrie:1] TRANS odwn-10-100303085-n=[wool industry]
odwn-10-100302950-n=[FM-zender:1] TRANS odwn-10-100302950-n=[FM Transmitter]
odwn-10-100302281-n=[eenakter:1] TRANS odwn-10-100302281-n=[acter]
odwn-10-100302142-n=[reünist:1] TRANS odwn-10-100302142-n=[reünist]
odwn-10-100299997-n=[gezinsverzorging:1] TRANS odwn-10-100299997-n=[family care]
odwn-10-100299543-n=[perscensuur:1] TRANS odwn-10-100299543-n=[press censorship]
odwn-10-100298134-n=[belle:1] TRANS odwn-10-100298134-n=[Belle]
odwn-10-100295695-n=[dressman:1] TRANS odwn-10-100295695-n=[dressman]
odwn-10-100295434-n=[chartermaatschappij:1] TRANS odwn-10-100295434-n=[charter company]
odwn-10-100293669-n=[harmoniemodel:1] TRANS odwn-10-100293669-n=[harmony model]
odwn-10-100291645-v=[potverteren:1] TRANS odwn-10-100291645-v=[Marken]
odwn-10-100290337-n=[hallenkerk:1] TRANS odwn-10-100290337-n=[hallchurch]
odwn-10-100289577-n=[badjuffrouw:1] TRANS odwn-10-100289577-n=[badjuffrouw]
odwn-10-100285322-n=[duivenvlucht:1] TRANS odwn-10-100285322-n=[pigeons flight]
odwn-10-100284834-n=[minimumsnelheid:1] TRANS odwn-10-100284834-n=[minimum rate]
odwn-10-100284712-n=[vertaalcomputer:1] TRANS odwn-10-100284712-n=[translation computer]

MONOSEMOUS ENTRY IN WordNet-PWN

odwn-10-100310851-n=[Noord-Ier:1] TRANS eng-30-08887841-n=Northern Ireland
odwn-10-100300801-n=[slangenbezweerder:1] TRANS eng-30-10615584-n=snake charmer
odwn-10-100300379-v=[wegsluipen:1] TRANS eng-30-01917123-v=slink
odwn-10-100296374-n=[pinkstergemeente:1] TRANS eng-30-10414865-n=Pentecostal
odwn-10-100291038-n=[vliegtuigbouwer:1] TRANS eng-30-02686568-n=aircraft
odwn-10-100285397-n=[wonderdokter:1] TRANS eng-30-10784544-n=witch doctor
odwn-10-100284846-n=[entreegeld:1] TRANS eng-30-13321495-n=entrance fee
odwn-10-100279661-v=[afwaaien:3] TRANS eng-30-01290945-v=blow off
odwn-10-100268396-n=[opsporingsregister:1] TRANS eng-30-06508816-n=data file
odwn-10-100260016-n=[kniebeschermer:1] TRANS eng-30-03623198-n=knee pad
odwn-10-100257258-n=[adel:4,adelendom:1,nobiliteit:1,noblesse:1] TRANS eng-30-14432032-n=nobility,
odwn-10-100255586-n=[geboortestreek:1] TRANS eng-30-08510169-n=homeland
odwn-10-100248287-n=[q:1] TRANS eng-30-06832896-n=q
odwn-10-100241160-n=[luilak:2] TRANS eng-30-10608385-n=slacker
odwn-10-100241028-n=[methadonbus:1] TRANS eng-30-03754014-n=methadone
odwn-10-100240814-v=[verknippen:1] TRANS eng-30-10463714-n=cut up
odwn-10-100236808-v=[afrijden:10] TRANS eng-30-02011302-v=ride off
odwn-10-100223757-n=[humaniora:1] TRANS eng-30-06153846-n=humanities
odwn-10-100210198-n=[zelfmoordpiloot:1] TRANS eng-30-10673776-n=suicide bomber
odwn-10-100208620-n=[F-16:1] TRANS eng-30-00041074-n=F-16

PARENT MATCH WordNet-PWN & ODWN

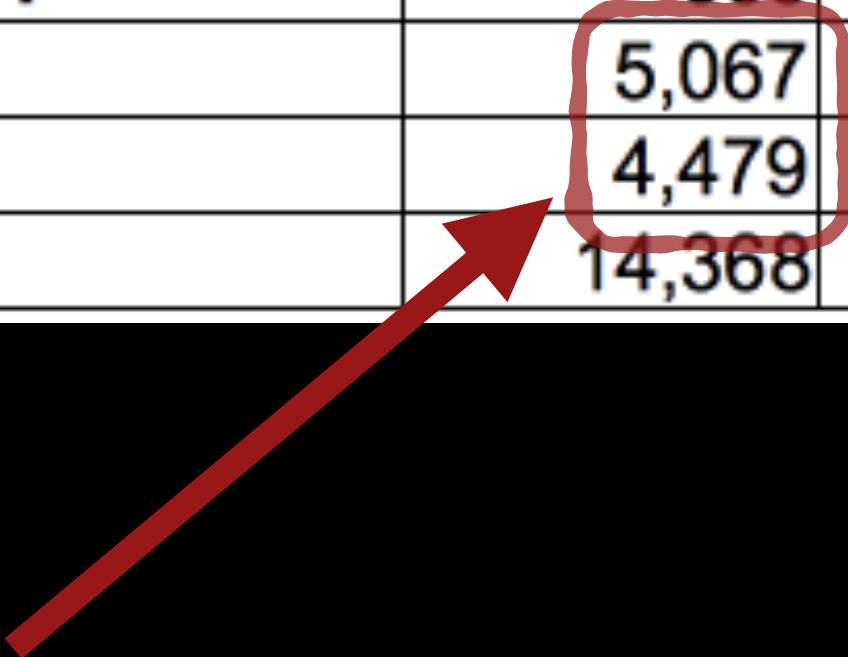
odwn-10-109232236-n=[topkwaliteit:1]	TRANS eng-30-04728068-n=quality
odwn-10-109229564-n=[studieboek:1]	TRANS eng-30-06410904-n=textbook
odwn-10-109226742-n=[ergotherapie:1]	TRANS eng-30-00658082-n=occupational therapy
odwn-10-109220717-n=[rookworst:1]	TRANS eng-30-07675627-n=sausage
odwn-10-109220717-n=[rookworst:1]	TRANS eng-30-00002684-n=sausage
odwn-10-109220254-n=[eindsprint:1]	TRANS eng-30-00294452-n=sprint
odwn-10-109210977-n=[radarinstallatie:1]	TRANS eng-30-03575240-n=radar
odwn-10-109210246-n=[grut:1,kinderzegen:1,kroost:1]	TRANS eng-30-10373998-n=fry, blessing of children, offspring
odwn-10-109208446-n=[seksist:1]	TRANS eng-30-00007846-n=sexist
odwn-10-109207525-n=[safe sex:1]	TRANS eng-30-00844254-n=safe sex
odwn-10-109201209-v=[calculeren:1]	TRANS eng-30-00637259-v=calculate
odwn-10-109195775-n=[pinacothek:1,schilderijenkabinet:1]	TRANS eng-30-03412058-n=art gallery, paintings cabinet
odwn-10-109186767-n=[samensteller:1]	TRANS eng-30-09614315-n=composer
odwn-10-109179831-n=[pro:1]	TRANS eng-30-06648724-n=pro
odwn-10-109169751-n=[concessie:3]	TRANS eng-30-00030358-n=concession
odwn-10-109164231-n=[Vlaming:1]	TRANS eng-30-09620078-n=Fleming

L&C SIM MATCH > 2.0 WordNet-PWN & ODWN

odwn-10-109951353-n=[molm:1] TRANS eng-30-14844693-n=mold
odwn-10-109860771-n=[bakkersknecht:1] TRANS eng-30-09632518-n=baker
odwn-10-109837643-n=[Israëliër:1] TRANS eng-30-00007846-n=Israeli
odwn-10-109768435-n=[hutkoffer:1] TRANS eng-30-02774630-n=trunk
odwn-10-109747342-n=[eindstation:1] TRANS eng-30-04412901-n=terminal
odwn-10-109739035-n=[boeteling:1] TRANS eng-30-00007846-n=penitent
odwn-10-109722611-n=[vrouwtje:3] TRANS eng-30-00007846-n=female
odwn-10-109694534-n=[tankschip:1] TRANS eng-30-04194289-n=tanker
odwn-10-109629155-n=[tekstbegrip:1] TRANS eng-30-05804793-n=comprehension
odwn-10-109617280-n=[baar:1] TRANS eng-30-03563967-n=bar
odwn-10-109597167-n=[kankergezwel:1] TRANS eng-30-14239918-n=cancer
odwn-10-109591471-n=[houtsnede:1] TRANS eng-30-04003597-n=woodcut
odwn-10-109567922-n=[reformator:1] TRANS eng-30-00007846-n=reformer
odwn-10-109523889-n=[afzetter:1] TRANS eng-30-09998101-n=cheat
odwn-10-109493447-n=[straatveger:1] TRANS eng-30-00007846-n=scavenger
odwn-10-109467078-n=[rekenaar:1] TRANS eng-30-09617867-n=calculator

Translations of lemmas

LEMMAMATCHING	TOTAL	NOUNS	VERBS
MONOSEMOUS	2,032	1,660	372
DIRECT SYNONYM OR HYPERNYM	2,455	1,939	516
Leacock-Chodorow SIMILARITY	335	334	1
NOT ENTRY PWN	5,067	4,236	831
ENTRY BUT NO MATCH	4,479	2,966	1,513
TOTAL	14,368	11,135	3,233



Try to match the translations of the glosses:

- Dice score content words (length $\geq$ 3)
- Compare co-hyponyms and co-hyponyms of hypernyms

GLOSS MATCHING

DICE MATCHING RESULTS NO ENTRY	TOTAL	NOUNS	VERBS
TOTAL	5,067	4,236	831
NO DEFINITION	377	309	68
ONE WORDS GLOSS	363	188	175
MORE THAN ONE WORD	3,206	1,603	1,603
NO MATCH RESULT	825	672	153
top (70-100)	37	34	3
middle (31-69)	1,725	1,518	207
bottom (1-30)	1,740	1,515	225

DICE MATCHING RESULTS NO MATCH	TOTAL	NOUNS	VERBS
TOTAL	4,479	2,966	1,513
NO DEFINITION	140	80	60
ONE WORDS GLOSS	708	300	408
MORE THAN ONE WORD	3,206	1,603	1,603
NO MATCH RESULT	705	467	238
top (70-100)	42	28	14
middle (31-69)	1,303	972	331
bottom (1-30)	1,581	1,119	462

GLOSS MATCHING

DICE MATCHING RESULTS NO ENTRY	TOTAL	NOUNS	VERBS
TOTAL	5,067	4,236	831
NO DEFINITION	377	309	68
ONE WORDS GLOSS	363	188	175
MORE THAN ONE WORD	3,206	1,603	1,603
NO MATCH RESULT	825	672	153
top (70-100)	37	34	3
middle (31-69)	1,725	1,518	207
bottom (1-30)	1,740	1,515	225

DICE MATCHING RESULTS NO M	TOTAL	NOUNS	VERBS
TOTAL	4,479	2,966	1,513
NO DEFINITION	140	80	60
ONE WORDS GLOSS	708	300	408
MORE THAN ONE WORD	3,206	1,603	1,603
NO MATCH RESULT	705	467	238
top (70-100)	42	28	14
middle (31-69)	1,303	972	331
bottom (1-30)	1,581	1,119	462

ILL candidates

ZERO DICE SCORE

odwn-10-109994970-n	[bierbuik][iemand met zo'n buik,someone with such a belly]
odwn-10-109967773-v	[instoppen][in bed leggen,lay in bed]
odwn-10-109936203-n	[kaalslag][het kappen van bomen,cutting down trees]
odwn-10-109930274-n	[anticlimax][tegenvaller bij hoge verwachtingen]
odwn-10-109918239-n	[vibrator][trillend toestel,vibrating device]
odwn-10-109913187-n	[zee-engte][vernaauwing in de zee door land]
odwn-10-109890697-n	[afweersysteem][afweer tegen ziektes,disease resistance]
odwn-10-109871362-n	[nieuwsvoorziening][verzorging v.h. nieuws]
odwn-10-109865276-n	[receptie][officiële ontvangst]
odwn-10-109854072-v	[bestellen][iemand laten komen,allow anyone]
odwn-10-109837238-n	[hamerstuk][vergaderpunt zonder discussie,meeting point without discussion]
odwn-10-109834436-n	[coupure][bedrag v.e. bankbiljet]
odwn-10-109803922-v	[ontlokken][bewering afdwingen,enforce allegation]
odwn-10-109797015-v	[afstrijken][strijkwerk afmaken,ironing finish]
odwn-10-109796989-v	[meedraaien][samen draaien,rotate together]
odwn-10-109795703-n	[imperialist][strever naar machtsuitbreiding,striver for power expansion]

LOW DICE < 30

20	odwn-10-109974807-n	[bootvluchteling][iemand per boot zijn land ontvlucht, someone boat fled country]
20	eng-30-10017890-n	[DP, displaced_person, stateless_person][a person forced to flee from home or country]
29	odwn-10-109969533-n	[schuimspaan][platte lepel met gaatjes, flat spoon with holes]
5	eng-30-03623556-n	[knife][gerei om mee te snijden, edge tool used as a cutting instrument; has a pointed blade with a sharp edge and a handle, paraphernalia]
19	eng-30-03361945-n	[flatware][tableware that is relatively flat and fashioned as a single piece]
17	eng-30-03025755-n	[chopstick][one of a pair of slender sticks used as oriental tableware to eat food with]
25	eng-30-04220945-n	[silver_plate][tableware that is plated with silver]
25	eng-30-03445472-n	[gold_plate][tableware that is plated with gold]
22	eng-30-02763714-n	[awl][scherpe staaf om gaatjes te prikken, a pointed tool for marking surfaces or for punching small holes, sharp stick to poke holes]
14	eng-30-03524574-n	[hoe][landbouwwerktuig met gekromd blad om de grond los te hakken, bepaald tuingereedschap, a tool with a flat blade attached at right angle]
8	eng-30-03239726-n	[drill][apparaat om gaten te maken, machinale boor, elektrische boormachine, a tool with a sharp point and cutting edges for making holes]
12	eng-30-03929202-n	[pickaxe, pickax, pick][scherp en puntig werktuig, houweel met punt en beitel, a heavy iron tool with a wooden handle and a curved head to dig]
16	eng-30-04023249-n	[puncher, punch][apparaat dat gaatjes maakt, apparaat dat kaarten ponst, a tool for making holes or indentations, device that makes holes]
6	eng-30-04051439-n	[rammer][schroefpers, a tool for driving something with force, screw press]
19	eng-30-02908217-n	[brush][borstel met handvat, an implement that has hairs or bristles firmly set into a handle, brush with handle]
5	eng-30-04050066-n	[rake][tuingereedschap, a long-handled tool with a row of teeth at its head; used to move leaves or loosen soil, garden tools]
18	eng-30-03403643-n	[funnel][voorwerp met wijde hals en smalle tuit, a conically shaped utensil having a narrow tube at the small end; used to channel the flow of liquids or granular materials]
16	eng-30-03454211-n	[grappler, grappling_iron, grapnel, grapple, grappling_hook][a tool consisting of several hooks for grasping and holding; often thrown with a rope]
29	eng-30-03919430-n	[muller, pestle, pounder][a heavy tool of stone or iron (usually with a flat base and a handle) that is used to grind and mix material]

30 < = DICE < 70

50	odwn-10-109994127-n	[flinter][zeer dun schijfje of stukje, very thin slice or shred]
50	eng-30-09428741-n	[section][a very thin slice (of tissue or mineral or other substance) for examination under a microscope]
41	odwn-10-109993315-v	[lachen][vrolijkheid uitdrukken, express mirth]
35	eng-30-01067194-v	[menace][express a threat either by an utterance or a gesture]
41	eng-30-00929839-v	[connote, imply][express or state indirectly]
35	eng-30-01067380-v	[beam][express with a beaming face or smile]
37	eng-30-00032778-v	[sneer][express through a scornful smile]
41	eng-30-01067512-v	[smile][express with a smile]
41	eng-30-00928959-v	[say][communicate or express nonverbally]
41	odwn-10-109978071-n	[varsity][roeiwedstrijd tussen studenten, rowing competition between students]
41	eng-30-07468534-n	[swimming_meet, swim_meet][a swimming competition between two or more teams]
34	odwn-10-109975023-n	[peillood][apparaat om diepte van water te meten, device to measure depth of water]
34	eng-30-04260192-n	[sonic_depth_finder, fathometer][depth finder for determining depth of water or a submerged object by means of sound]
37	odwn-10-109960175-n	[breedtecirkel][denkbeeldige cirkel over de aarde, imaginary circle the earth]
37	eng-30-13873849-n	[circlet][a small circle]
32	eng-30-05102889-n	[osculating_circle, circle_of_curvature][the circle that touches a curve (on the concave side) and whose radius is the radius of curvature]
37	odwn-10-109950849-v	[stationeren][een standplaats geven, give a stand]
37	eng-30-01517515-v	[upend][set, turn, or stand on end]
37	eng-30-01515196-v	[launch, set_in_motion][get going; give impetus to]
33	odwn-10-109949756-v	[meerijden][met anderen rijden, others ride]
33	eng-30-02376958-v	[interact][act together or towards others or with others]
36	odwn-10-109945969-n	[matje][lang nekhaar bij een verder kort kapsel, long neck hair with a further short haircut]
36	eng-30-05259109-n	[bob][a hair style for women and children; a short haircut all around]

DICE ≥ 70

83	odwn-10-109697118-n	[brandschade]	[schade door brand,damage by fire]
83	eng-30-00403911-n	[burn]	[damage inflicted by fire]
75	odwn-10-109616370-v	[opbieden]	[hoger bieden,bid higher]
75	eng-30-02299801-v	[outbid]	[bid higher than others]
71	odwn-10-109571923-n	[wikkel]	[papiertje waarin iets verpakt wordt,paper in which something is wrapped]
71	eng-30-04605726-n	[wrap,wrapping,wrapper]	[the covering (usually paper or cellophane) in which something is wrapped]
83	odwn-10-109187549-n	[schuimplastic]	[bepaalde kunststof]
83	eng-30-03979377-n	[polyester]	[bepaalde harde kunststof,any of a large class of synthetic fabrics,certain hard plastics]
73	odwn-10-109076330-n	[hit]	[klein soort paard,kind of small horse]
73	eng-30-02382437-n	[pony]	[klein soort paard,any of various breeds of small gentle horses usually less than five feet high]
83	odwn-10-108712975-n	[rammelaar]	[mannetje van het konijn,male rabbit]
83	eng-30-02324514-n	[lapin]	[castrated male rabbit]
83	odwn-10-108646433-n	[rotswand]	[steile zijde v.e. rots]
83	eng-30-09246464-n	[cliff,drop,drop-off]	[steile rots,a steep high face of rock,cliff]
76	odwn-10-108287212-v	[uitknippen]	[vormen met een schaar,shapes with scissors]
76	eng-30-01248950-v	[scissor]	[met een schaar,cut with or as if with scissors,with scissors]
72	odwn-10-107859308-n	[Maagd]	[één v.d. twaalf sterrenbeelden,one of the Twelve Constellations]
72	eng-30-09232687-n	[cancer]	[één v.d. twaalf sterrenbeelden,a small zodiacal constellation in the northern hemisphere; bet
77	odwn-10-107748764-v	[omklappen]	[van stand veranderen,of state change]
77	eng-30-01831531-v	[move]	[van stand veranderen,move so as to change position, perform a nontranslational motion,of state
75	odwn-10-107646993-v	[afgieten]	[kookwater weggieten,cooking water pour out]
75	eng-30-02070296-v	[pour_out,decant,pour]	[pour out]
75	eng-30-02069788-v	[pour_out,effuse]	[pour out]

Next steps

- fuzzy sense mappings: 7,489 synsets
- matching ODWN synsets need to be merged with the PWN synsets: about 5,000 synsets
- non-matching synsets:
 - no glosses or single word gloss
 - ILI-candidates: about 1,500 zero-dice and 3,000 low-dice scores

Future GWG

- Test procedure for adding wordnets and extending ILI
 - WordNet3.0 and WordNet3.1
 - Dutch Open WordNet
 - Multilingual Wordnet Database
- Launch the website with pointers to ILI repository, Github for releasing wordnets
- Install matching procedures
- Involve the community