

Chiara Zanchi

University of Pavia

chiara.zanchi01@unipv.it

Silvia Luraghi

University of Pavia

silvia.luraghi@unipv.it

Erica Biagetti

University of Pavia / Bergamo

erica.biagetti01@universitadipavia.it

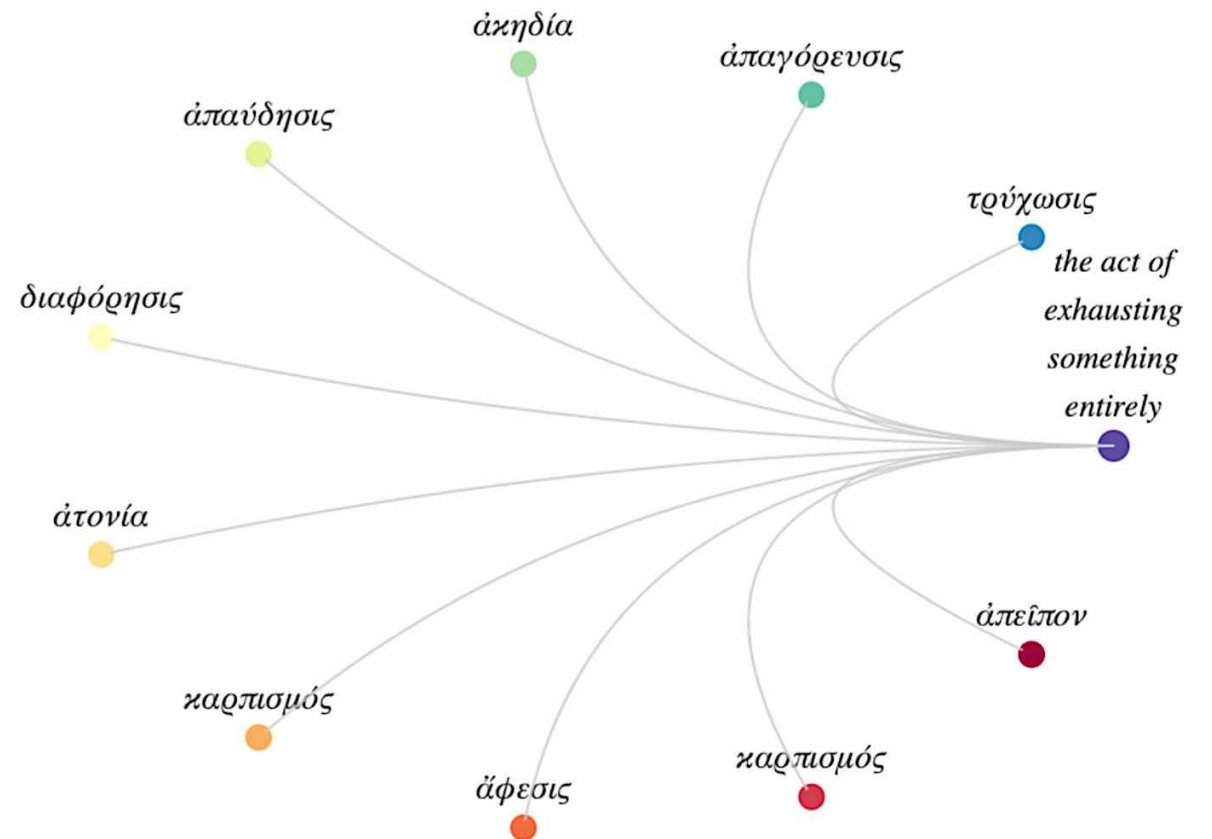
LINKING THE ANCIENT GREEK WORDNET TO THE HOMERIC DEPENDENCY LEXICON

INTRODUCTION

➤ **WordNets** (WNs): lexical databases in which meaning is stored in a relational way

- Lemmas and Synsets
- Lexical Relations
- Semantic Relations

➤ **Pilot study:** linking the Ancient Greek WordNet (AGWN) to the Homeric Dependency Lexicon (HoDeL)



STRUCTURE OF THIS PRESENTATION

1. The family of WordNets for ancient Indo-European languages
2. Enriching the Ancient Greek WordNet with sentence frames
3. Syntactic information about verbs' construction in WordNet
4. Verb frames in HoDEL
5. Discussion

THE FAMILY OF WORDNETS FOR ANCIENT INDO-EUROPEAN LANGUAGES

➤ **Ancient Greek WN:**

- Family of WNs for ancient Indo-European languages
- WNs for Sanskrit and Latin
- University of Pavia, University of Exeter, Center for Hellenic Studies (Harvard University)

➤ **Architecture**

- Integration with other resources for ancient languages → URIs
- Interoperability among WNs of this family
- Maximal usage of the synsets deriving from the Princeton WordNet

THE FAMILY OF WORDNETS FOR ANCIENT INDO-EUROPEAN LANGUAGES

- **Principled view of polysemy** → structured semantic network
- Diachronic and stylistic metadata associated with each synset

pléō ‘sail’

(1)	Literal senses ‘sail’, ‘float’ a. v#01260993 travel by boat b. v#01299504 float on the surface of water
(2)	Metonymic sense ‘depart’ v#01262245 move away from a place into another direction
(3)	Metaphoric sense ‘behave in a certain way’ v#00007023 behave in a certain manner

Sense	Period	Genre	Loci
(1)a	Archaic (8 th -6 th BCE)	poetry epic historiography theater ...	Il.3.444, Pi.P.4.69 X.Cyr.6.16 A.Ag.691
(2)	Archaic (8 th -6 th BCE)	epic	Od.12.5
(3)	Classical (5 th c.-323 BCE)	theater philosophy oratory	S.Ant.190 D.19.250 Pl.Lg.813d

THE FAMILY OF WORDNETS FOR ANCIENT INDO-EUROPEAN LANGUAGES

➤ Language-specific features and morphological annotation

Field	Subfield	Value
Etymology		PIE. *pleu- 'float'
Lemma		<i>pléō</i>
POS		Verb
Morpho		v1spia--1e
Uri		π07077
Morphology	Principal Parts	<i>pleúsomai épleusa pépleuka epléusthēn pleusthēsomai pèpleusmai</i>
	Prosody	<i>pléō</i>
Form Tokens	Form	v1sfia--1e
	Token	<i>plesoūmai</i> (FUT.IND.1SG)
	Alternative	✓ (late form of <i>pleúsomai</i>)

THE FAMILY OF WORDNETS FOR ANCIENT INDO-EUROPEAN LANGUAGES

➤ Extended set of lexical relations

Relation	Label and Example	Inverse
Derivation	<i>makró-tēs</i> ‘length’ IS DERIVED FROM <i>makrós</i> ‘long’	IS RELATED TO
Antonymy	<i>anomía</i> ‘lawlessness’ IS PRIVATIVE OF <i>nomía</i> ‘lawfulness’	HAS PRIVATIVE
Parasynthesis	<i>ánoos</i> ‘without understanding’ IS PARASYNTHETIC OF <i>nóos</i> ‘mind’	HAS PARASYNTHETON
Composition	<i>makropoiéō</i> ‘enlarge’ IS COMPOSED OF <i>makrós</i> ‘big’; <i>makropoiéō</i> ‘enlarge’ IS COMPOSED OF <i>poiéō</i> ‘make’	COMPOSES
Inclusion	<i>thalássia érga</i> ‘navigation’ INCLUDES <i>thalássios</i> ‘related to the sea’; <i>thalássia érga</i> ‘navigation’ INCLUDES <i>érgon</i> ‘work’	IS INCLUDED IN
Participle	<i>eikós</i> ‘seeming like’ IS PARTICIPLE OF <i>éoika</i> ‘to be like’	HAS PARTICIPLE

ENRICHING THE ANCIENT GREEK WORDNET WITH SENTENCE FRAMES

- **Semi-automatic two-step workflow**
- Pilot study of Homeric verbs with HoDeL (<https://hodel.unipv.it/hodel-res>)
- Ultimately, whole Ancient Greek and Latin Dependency Treebank (AGLDT)
- **HoDeL**: corpus-based lexicon of Homeric verbs induced from AGLDT 2.0 via a series of SQL queries whose results are stored in a relational database, which lies behind the online interface
 - It allows obtaining **corpus-based data concerning sentence frames**, their frequency, and instantiations
 - It gives information about the **syntactic and semantic features of verbal dependents**
 - It provides **aligned translation of Homeric passages**

ENRICHING THE ANCIENT GREEK WORDNET WITH SENTENCE FRAMES

➤ Mapping the Homeric verbs in HoDeL with the corresponding entries in AGWN

- Two data extractions of the lemmas having 'v' as a postag in HoDeL database and in AGWN Github repository (<https://github.com/greekwordnet>)
- HoDeL: 2,482 results; AGWN: 28,405 results → automatic comparison

Type of correspondence (HoDeL-to-Ancient Greek WordNet)	Frequency
one-to-one	2,256 (90.94%)
one-to-many (two)	41 (1,65%)
one-to-zero	185 (7.41%)
TOTAL	2,482 (100%)

ENRICHING THE ANCIENT GREEK WORDNET WITH SENTENCE FRAMES

One-to-many correspondences:

- homophonous verbs, represented by ambiguous character strings: “δέω” (BetaCode: de/w), verbs assigned URIs “δ01019” and “δ01020”, corresponding to the meanings ‘tie’ and ‘lack’
- verbs contained twice in AGWN due to importation errors from the electronic version of the reference dictionary of AGWN, the Liddell-Scott-Jones: “τρέφω” being assigned the URIs τ02408 and τ02407

One-to-zero correspondences:

- different lemmatizations
- different notations
- errors in AGWN

Mismatch type	HoDeL	AGWN
(a) different voice	μετατρέπομαι	μετατρέπω
(b) variants	διαπέταμαι	διαπέτομαι
(c) participles	θυμηγερέων (v)	θυμηγερέων (n)
(d) hiatus	δικνέομαι	διῦκνέομαι
(e) subscript iota	θνήσκω	θνήσκω
(f) missing verbs	καταβρόχω	καταβρόξειε (n)
(g) incorrect pos	μέμονα	μέμονα (n)

ENRICHING THE ANCIENT GREEK WORDNET WITH SENTENCE FRAMES

- Annotators asked to assign the extracted sentence frames to the correct synset → **manual sense disambiguation**
- Pros of this workflow:
 - HoDeL provides aligned English translations of the AG passages
 - More convenient to assign sentence frames to each lexicon entry while AGWN is under construction
- The added value of sentence frames extracted from corpora:

Frequency distribution of sentence frames, the most frequent filling words for each frame-slot and their associative connections with synsets (synsets make it possible to generalize these associative connections across verbal lexemes)

SYNTACTIC INFORMATION ABOUT VERBS' CONSTRUCTION IN WORDNET

- 35 sentence frames indicating “**the number of noun arguments that the verb subcategorizes for**”
- The noun slots in these syntactic frames are not linked to either thematic roles or semantic categories
- Sentence frames constructed by the WN annotators based on their knowledge of the language, and the examples are **not extracted from corpora**

- **S: (v) announce, denote** (make known; make an announcement) *"She denoted her feelings clearly"*
 - direct troponym / full troponym
 - direct hypernym / inherited hypernym / sister term
 - derivationally related form
 - **sentence frame**
 - Somebody —s something
 - Something —s somebody
 - Something —s something
 - Somebody —s that CLAUSE

SYNTACTIC INFORMATION ABOUT VERBS' CONSTRUCTION IN WORDNET

- Verbs' constructions are largely language specific → **sentence frames cannot be exported from the Princeton WN when creating a new WN for a different language**
- Es. **GermaNet** (<https://weblicht.sfs.uni-tuebingen.de/rover/>)
 - Syntactic categorization indicated and partly corpus-based
 - *verkünden* 'announce': three synsets. First synset 'öffentlich sagen' ('announce'): three verb frames with examples
 - NN.Dn.DS – *Sie verkündete, dass sie heiraten ihn würde.* “She announced that she was going to get married.”
 - NN.Dn.AN – *Der Sektenführer verkündete seinen Anhängern die frohe Botschaft.* “The leader of the sect announced his followers the good news.”
 - NN.AN – *Er verkündete die Fusion der beiden Unternehmen.* “He announced the merger of the two companies.”
 - Language specific information (case marking, verbal aspect) and language independent information (complements and adverbial, obligatory and optional; cf. locative PP with *wohnen* 'live (in a place)' BL is obligatory, directional PP with *segeln* 'sail' Bd is optional); no animacy feature of fillers in GermaNet

VERB FRAMES IN HODEL

- Case studies: *pléō* ‘sail, float’ (28 occurrences already validated) & *angéllō* ‘announce’ (27 occurrences not yet validated)
- GermanNet annotation + animacy information
- **Some relevant features of Ancient Greek:**
 - Null subject language, but the subject triggers verbal agreement → + NN for null subjects
 - Grammaticalized verbal aspect interacting with tense
 - Three aspectual stems: present/imperfective, aorist/perfective, perfect/resultative
 - Future tense not sensitive to aspect
 - Different aspectual stems of the same verb can show different values for voice (e.g. *gígnomai*:PRS.M/P ‘become’ vs. *gégona*.PF.ACT)

pléō: three synsets ‘float’, ‘sail’, ‘depart’ (metaphoric sense not relevant)

1 NN(-a/+a) ...impf

1ⁱ NN(-a/+a) ...fut

1 NN(-a/+a) ...impf

a) FLOAT

Od. 5.240 *tá* *hoi* *plóoien* *elaphrôs*
DEM.NOM.PL 3SG.DAT float.OPT.PRS.3SG lightly
‘They (=the trees) would float lightly for him.’

b) SAIL

Od. 10.80 *ennêmar* *mèn* *homôs* *pléomen*
for_nine_days PTC alike sail.IMPF.1PL
‘We sailed for nine days alike (night and day).’

c) DEPART

1ⁱ NN(-a/+a) ...fut

Od. 12.5 *háma* *d’* *ēoî* *phainoménēphi* *pleúsesth’*
together PTC dawn.DAT appearing.DAT sail.FUT.MID.2PL
‘At dawn you will depart.’

| *pléō*: three synsets ‘float’, ‘sail’, ‘depart’
(metaphoric sense not relevant)

2 NN(-a/+a) ...impf PP

2ⁱ Perlative PPs or adverbs

2ⁱⁱ Direction PP

2 NN(-a/+a) ...impf PP

b) SAIL

2ⁱ Perlative PPs or adverbs (*epì oínopa pónton* ‘over the wine-dark sea’, Il. 7.88)

2ⁱⁱ Direction PPs

c) DEPART

Source PPs or adverbs (*apò Krētēs* ‘from Crete’ Od. 14.253; *póthen* ‘where from?’ Od. 3.71)

aggéllō synset #00659537 | make known

1 NN(+a) ...ptcp.fut/aor Nd(+a)

1ⁱ NN(+a) ...fut

1 NN(+a) ...ptcp.fut/aor Nd(+a)

Od. 16.150

sú

2SG.NOM

g'

PTC

aggeilas

announce.PTCP.AOR.NOM

opísō

back

kie

come.IMP.PRS.2SG

‘After announcing, come back.’

1ⁱ NN(+a) ...fut (Il. 9.617)

aggéllō synset #00659537 | make known

2 NN(+a) ...impf/aor Na(-a) Nd(+a)

2ⁱ NN(+a) NA (+A)

2ⁱⁱ NN(-a) NA(-a)

2 NN(+a) ...impf/aor Na(-a) Nd(+a)

Il. 15.159 *Poseidáōni* *ánakti* *pánta* *tád'* *aggeîlai*
P.DAT king.DAT all.ACC.PL DEM.ACC.PL announce.INF.AOR
'(Go) to announce all these things to kingly Poseidon.'

2ⁱ NN(+a) NA (+a) 'bring information about someone' (Od. 14.122-123)

2ⁱⁱ NN(-a) NA(-a) (Od. 13.93-94 with an inanimate subject *astēr* 'star' and an inanimate object *pháos* *Ēoús* 'the light of dawn')

aggéllō synset #00659537 | make known

3 NN(+a) ...aor ND(+a) INFINITIVE

3 NN(+a) ...aor ND(+a) INFINITIVE

Od. 16.350 *keinois* *aggeilōsi* *thoōs* *oíkonde*
DEM.DAT.PL announce.SBJV.AOR.3PL quickly home

néesthai

travel.INF.PRS.MID

‘(That they) announce them to quickly travel home.’

aggéllō synset #00659537 | make known

4 NN(+a) ...impf/aor COMPL CLAUSE

4ⁱ NN(+a) AcI

4ⁱⁱ NN(+a) ND *hótti* clause

4 NN(+a) ...impf/aor COMPL CLAUSE

4ⁱ NN(+a) AcI

Il. 8.517-8.519 *kérukes ...* *aggellóntōn* *paídas ...* *gérontas*
herald.NOM.PL announce.IMP.PRS.3PL boy.ACC.PL old_man.ACC.PL
léxasthai
gather.INF.AOR.MID

‘Let the heralds announce that boys (and) old men gather.’

4ⁱⁱ NN(+a) ND *hótti* clause

Il. 22.438-439 *ou* *gár* *hoí ...* *éggeil'* *hótti* *rhá*
NEG PTC 3SG.DAT announce.AOR.3SG that PTC
hoi *pósis* *éktothi mímne* *puláōn*
3DAT.SG spouse.NOM outside abide.IMP.3SG gate.GEN.PL

‘No (messenger coming) had announced her that her husband abode outside the gates.’

SENTENCE FRAMES - INVENTORY

NN(-a/+a) ...impf

NN(-a/+a) ...fut

NN(+a) ...fut

NN(-a/+a) ...impf PP

NN(+a) ...ptcp.fut/aor Nd(+a)

NN(+a) ...impf/aor Na(-a) Nd(+a)

NN(+a) NA (+A)

NN(-a) NA(-a)

NN(+a) ...aor ND(+a) INFINITIVE

NN(+a) ...impf/aor COMPL CLAUSE

...to be continued!

DISCUSSION

- The issue of **obligatoriness** → We must rely on corpora, which may be partial in coverage
- The issue of **omitted OBJs** → should we accord them the same treatment as omitted SBJs?
- **Formulas** of the Homeric language and frequency counts
- Importance of verbal aspect/tense for sentence frames: different aspectual profile of the two verbs
 - *pléō* only features imperfective forms, while *aggéllō* can be used both in the imperfective and in the perfective aspect
 - sentence frame 1 of *aggéllō* is almost only limited to participles, with a preference for the future tense (7 occurrences out of 9, plus one occurrence of a finite form)

**THANK YOU
FOR YOUR ATTENTION**

REFERENCES

Biagetti E., Zanchi C., Short W. M. (2021), Toward the creation of WordNets for ancient Indo-European languages. Proceedings of the 11th Global WordNet Conference, S. Bosch, C. Fellbaum, M. Griesel, A. Rademaker, P. Vossen (eds), EAACL/GWC, Global WordNet Association, pp.258-266.

Baayen R H., Piepenbrock R., Gulikers L. (1995), CELEX2 LDC96L14. Web Download. Philadelphia: Linguistic Data Consortium.

Duhoux Y. (1992), The verb in Ancient Greek [Le verbe en grec ancien]. Peeters, Louvain-La-Neuve.

Fellbaum C. (ed.) (1998), WordNet: An electronic lexical database. MIT Press, Cambridge, MA.

Luraghi S. (2003), Definite referential null objects in Ancient Greek. Indogermanische Forschungen 108, pp. 169-196.

Shi L., Mihalcea R. (2005), Putting pieces together: Combining FrameNet, VerbNet and WordNet for robust semantic parsing. International conference on intelligent text processing and computational linguistics, pp. 100-111. Springer, Berlin.

Tyler A., Evans V. (2003), The Semantics of English Prepositions: Spatial Scenes, Embodied Meanings and Cognition. CUP, Cambridge, UK.

Zanchi C., Luraghi S. (2020), Presenting Hodel—A New Resource for Research On Homeric Greek Verbs. Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference “Dialog 2020” [Komp’yuternaya Lingvistika i Intellektual’nye Tekhnologii: Trudy Mezhdunarodnoy Konferentsii “Dialog 2020”], Moscow, Supplementary volume, pp.1188-1201.

Zanchi,C. (forthc.). The Homeric Dependency Lexicon: what it is and how to use it. Journal of Greek Linguistics.