

OLD ENGLISH UNIVERSAL DEPENDENCIES: CATEGORIES, FUNCTIONS AND SPECIFIC FIELDS

Javier Martín Arista
Universidad de La Rioja

ICAART2022

UNIVERSAL DEPENDENCIES

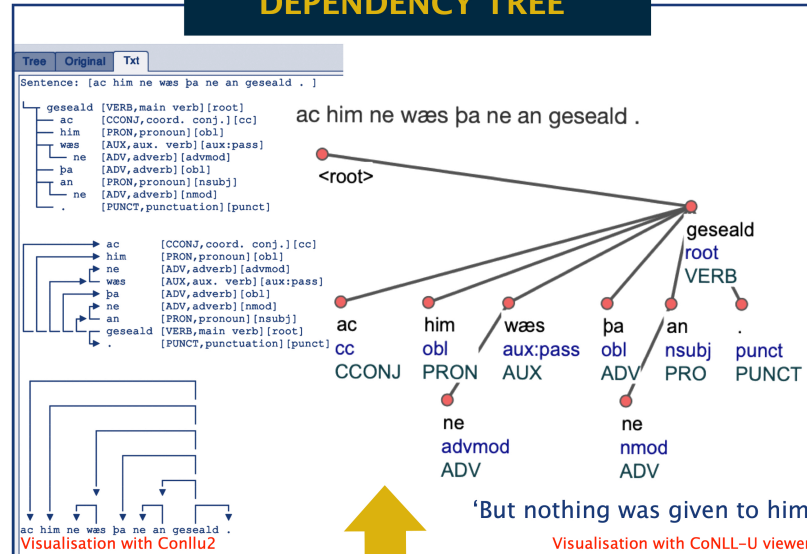
- Universal Dependencies (de Marneffe et al., 2021) is a model of morphological and syntactic annotation devised for the compilation of computerised data sets that facilitate cross-linguistic comparison aimed to natural language processing and to areas of linguistics like language acquisition and translation.
- The annotation includes UPOS (universal part-of-speech tags), XPOS (language-specific part-of-speech tags), Feats (universal morphological features), lemmas, and dependency heads and labels.
- The 2021 release of the UD comprises 183 treebanks over 104 languages.

de Marneffe, M., Manning, C., Nivre, J., & Zeman, D. (2021). Universal Dependencies. Computational Linguistics, 47 (2), 255-308.

OLD ENGLISH

- Old English is the historical stage of the English language spoken in England between approximately the fifth and the eleventh centuries (CE).
- Written records, which can be traced back to the seventh century, comprise approximately 3 million words in around 3,000 texts.
- Old English belongs to the West-Germanic Group of the Indo-European family of languages and is characterised by its explicit generalised morphological inflection and its consistently Germanic lexicon.

DEPENDENCY TREE



CATEGORIES

- Old English exponents of universal lexical categories: common noun, proper noun, main verb, auxiliary verb, adjective, demonstrative-article, numeral (card. and ord.), adverb, pronoun, adposition, coordinating conjunction, subordinating conjunction, interjection, foreign word, punctuation
- Old English values of universal morphological features: PronType (Dem-Art Ind Int Prs Rel), NumType (Card Ord), Poss (Yes), Foreign (Yes), Abbr (Yes), Typo (Yes), Gender (Fem Masc Neut), Number (Dual Plur Sing), Case (Acc Nom Dat Gen Ins), Reflex (Yes), Degree (Cmp Pos Sup), Person (1 2 3), VerbForm (Fin Inf Part), Mood (Imp Ind Sub), Tense (Past Pres), Voice (Act Mid Pass), Polarity (Neg), Person (1 2 3).

FUNCTIONS

- Syntactic annotation with Universal Dependencies comprises the identification of the dependent, of the head of dependency and of the dependency relation holding. It is based on the following principles:
 - Primacy of content words over function words.
 - Function words do not have dependents of their own.
 - Core arguments are compulsory and morphologically unmarked whereas oblique modifiers are optional and typically governed by preposition.

CONLL-U PLUS FORMAT

ID	FORM	LEMMA	UPOSTAG	XPOSTAG	FEATS	HEAD	DEPREL	DEPS	MISC	GLOSS	MORHPHREL
1	ac	ac	CCONJ	PRO	Uninflected=Yes Case=Dat	6	cc	-		but	-
2.	him	hē	PRO	pronoun	Gender=Masc Number=Sing Person=3 PronType=Prs	6	obl	-	him	him	-
3	næs	-	ADV	adverb	Polarity=Neg Uninflected=Yes	3	advmod	-	SpaceAfter=No	not	ø/BÉON
3.1	ne	ne	ADV	adverb	Polarity=Neg Uninflected=Yes	3	advmod	-	SpaceAfter=No	not	-
3.2	wæs	béon	AUX	aux. verb	Mood=Ind Number=Plur Tense=Pres VerbForm=Fin Uninflected=Yes	6	aux:pass	-		be	-
4	þa	þā	ADV	adverb	Uninflected=Yes Case=Nom	6	obl	-		then	-
5	nan	-	ADV	adverb	Polarity=Neg Uninflected=Yes	5	nmod	-	SpaceAfter=No	not	-
5.1	ne	ne	ADV	adverb	Polarity=Neg Uninflected=Yes	5	nmod	-	SpaceAfter=No	not	-
5.2.	an	ān	PRO	pronoun	Case=Nom Gender=Masc Number=Sing Person=3 PronType=Ind	6	nsubj	-		one	-
6	geseald	gesellan	VERB	main verb	Case=Nom Number=Sing Tense=Past VerbForm=Part Uninflected=Yes	0	root	-		given	sellan/SELLAN
7	.	-	PUNCT	punctuation	Uninflected=Yes	6	punct	-		-	-

