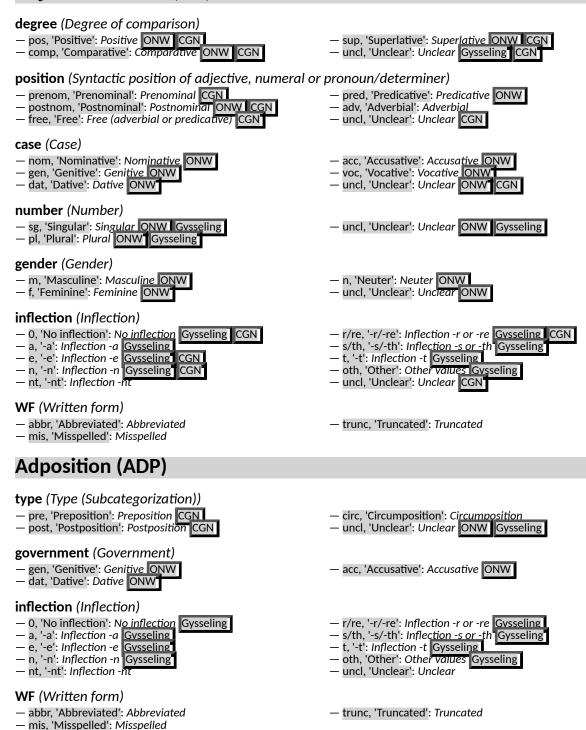
Bijlage 3: formele beschrijving tagset

In deze bijlage geven we een compact overzicht van de kenmerken van de tagset, waarbij we de corpora waarin kenmerken voorkomen achter het kenmerk zetten.

Het onderdeel 'declarations' geeft per corpus aan welke kenmerken, afhankelijk van de woordsoort of de waarden van andere kenmerken, verplicht zijn.

Adjective-Adverb (AA)



Adverb (ADV)

type (Type (Subcategorization)) reg, 'Regular': Regular ONW Gysseling CGN pron, 'Pronominal': Pronominal (adverb) ONW Gysseling uncl, 'Unclear': Unclear ONW CGN **subtype** (Subtype) dem. 'Demonstrative': Demonstrative adverb ONW Gysseling - prodrop, 'Pronoun drop': Pronominal adverb dropping the pronominal part - gen, 'General': General adverb Gysseling - indef, 'Indefinite': Indefinite adverb Gysseling - inter, 'Interrogative': Interrogative adverb ONW Gysseling - neg, 'Negative': Negative adverb ONW Gysseling - pers, 'Personal': Personal pronominal adverb (er) rel, 'Relative': Relative adverb ONW Gysseling resump, 'Resumptive': Resumptive adverb (mainly 'zo') Gysseling uncl, 'Unclear': Unclear ONW Gysseling inflection (Inflection) - r/re, '-r/-re': Inflection -r or -re Gysseling - s/th, '-s/-th': Inflection -s or -th' Gysseling - t, '-t': Inflection -t Gysseling - oth, 'Other': Other Values Gysseling — 0, 'No inflection': No inflection Gysseling – a, '-a': Inflection -a– e, '-e': Inflection -e Gysseling n, '-n': Inflection -n Gysseling nt, '-nt': Inflection -nt - uncl, 'Unclear': Unclear WF (Written form) abbr, 'Abbreviated': Abbreviatedmis, 'Misspelled': Misspelled - trunc, 'Truncated': Truncated Conjunction (CONJ) type (Type (Subcategorization)) - coor. 'Coordinating': Coordinating conjunction ONW Gysseling CGN - comp, 'Comparing': Comparing conjunction (als, dan, gelijkdat, alsof, ...) ONW Gysseling - qual, 'Qualitative': Oualitative conjunction 'als' in the sense 'in the quality of' Gysseling - expl, 'Explicative': Explicative conjunction Gysseling neg, 'Negative': Negative conjunction Gysseling rel, 'Relative': Relative Gysseling sub. 'Subordinating': Subordinating conjunction ONW Gysseling CGN uncl, 'Unclear': Unclear ONW Gysseling **inflection** (Inflection) O, 'No inflection': No inflection a, '-a': Inflection -a e, '-e': Inflection -e n, '-n': Inflection -n Gysseling Gysseling — r/re, '-r/-re': Inflection -r or -re Gysseling - s/th, '-s/-th': Inflection -s or -th Gysseling - t, '-t': Inflection -t Gysseling - oth, 'Other': Other values Gysseling — nt, '-nt': Inflection -ht - uncl, 'Unclear': Unclear **WF** (Written form) - abbr, 'Abbreviated': Abbreviated - trunc, 'Truncated': Truncated - mis, 'Misspelled': Misspelled Interjection (INT) inflection (Inflection) — 0, 'No inflection': No inflection Gysseling r/re, '-r/-re': Inflection -r or -re s/th, '-s/-th': Inflection -s or -th t, '-t': Inflection -t — a, '-a': Inflection -a — e, '-e': Inflection -e Gvsseling - n, '-n': Inflection -n Gysseling - nt, '-nt': Inflection -nt — oth, 'Other': Other values Gysseling - uncl, 'Unclear': Unclear **WF** (Written form) — abbr, 'Abbreviated': Abbreviated - trunc, 'Truncated': Truncated - mis, 'Misspelled': Misspelled Noun Common (NOU-C) **number** (Number) – sg, 'Singular': Singular ONW Gysseling CGN– pl, 'Plural': Plural ONW Gysseling CGN uncl, 'Unclear': Unclear ONW Gysseling CGN case (Case) nom, 'Nominative': Nominative ONW gen, 'Genitive': Genitive ONW CGN dat, 'Dative': Dative ONW CGN acc, 'Accusative': Accusative voc, 'Vocative': Vocative uncl, 'Unclear': Unclear ONW CGN

gender (Gender) n, 'Neuter': Neuter ONW CGNuncl, 'Unclear': Unclear ONW CGN - m, 'Masculine': Masculine ONW - f, 'Feminine': Feminine ONW inflection (Inflection) - 0, 'No inflection': No inflection - a, '-a': Inflection -a - e, '-e': Inflection -e - n, '-n': Inflection -n - nt, '-nt': Inflection -nt - CGN Gysseling Gysseling Gysseling CGN CGN - r/re, '-r/-re': Inflection -r or -re Gysseling - s/th, '-s/-th': Inflection -s or -th Gysseling CGN - t, '-t': Inflection -t Gysseling - oth, 'Other': Other values Gysseling - uncl, 'Unclear': Unclear CGN WF (Written form) — abbr, 'Abbreviated': Abbreviated - trunc, 'Truncated': Truncated - mis, 'Misspelled': Misspelled Noun Proper (NOU-P) number (Number) sg, 'Singular': Singular ONW CGNpl, 'Plural': Plural CGN uncl, 'Unclear': Unclear ONW CGN case (Case) nom, 'Nominative': Nominative gen, 'Genitive': Genitive CGN dat, 'Dative': Dative ONW CGN - acc, 'Accusative': Accusative voc, 'Vocative': Vocative uncl, 'Unclear': Unclear ONW CGN gender (Gender) m, 'Masculine': Masculinef, 'Feminine': FeminineONW n, 'Neuter': Neuter CGNuncl, 'Unclear': Unclear ONW CGN **inflection** (Inflection) - 0, 'No inflection': No inflection - a, '-a': Inflection - a Gysseling CGN - r/re, '-r/-re': Inflection -r or -re - s/th, '-s/-th': Inflection -s or -th - t, '-t': Inflection -t - oth, 'Other': Other values Gysseling - uncl, 'Unclear': Unclear CGN - a, '-a': Inflection -a Gysseling - e, '-e': Inflection -e Gysseling CGN - n, '-n': Inflection -n Gysseling CGN - nt, '-nt': Inflection -nt **WF** (Written form) - abbr, 'Abbreviated': Abbreviated - trunc, 'Truncated': Truncated - mis, 'Misspelled': Misspelled Numeral (NUM) type (Type (Subcategorization)) — card, 'Cardinal': Cardinal ONW Gysseling CGN — uncl, 'Unclear': Unclear Gysseling ord, 'Ordinal': Ordinal ONW Gysseling CGN **position** (Syntactic position of adjective, numeral or pronoun/determiner) free, 'Free': Free (adverbial or predicative) uncl, 'Unclear': Unclear ONW Gysseling CGN - prenom, 'Prenominal': Prenominal CGN postnom, 'Postnominal': Postnomina representation (Number representation) - dig, 'Digit': Digit mix-rom, 'Mix-roman': Combination of roman numeral with - let, 'Alphabetic': Alphabetic letters - rom, 'Roman': Roman uncl, 'Unclear': Unclear ONW Gysseling CGN — mix-dig, 'Mix-digit': Combination of digits and letters case (Case) nom, 'Nominative': Nominative gen, 'Genitive': Genitive ONW dat, 'Dative': Dative ONW acc, 'Accusative': Accusative ONW — uncl, 'Unclear': Unclear ONW CGN **number** (Number) — sg, 'Singular': Singular ONW — pl, 'Plural': Plural ONW uncl, 'Unclear': Unclear ONW **gender** (Gender) m, 'Masculine': Masculine ONWf, 'Feminine': Feminine ONW — n, 'Neuter': Neuter ONW uncl, 'Unclear': Unclear ONW

inflection (Inflection) — 0, 'No inflection': No inflection Gysseling CGN - r/re, '-r/-re': Inflection -r or -re Gysseling CGN - a, '-a': Inflection -a - e, '-e': Inflection -e - n, '-n': Inflection -n Gysseling CGN - s/th, '-s/-th': Inflection -s or -th Gysseling - t, '-t': Inflection -t Gysseling - oth, 'Other': Other values Gysseling - nt, '-nt': Inflection -nt uncl, 'Unclear': Unclear CGN **WF** (Written form) abbr, 'Abbreviated': Abbreviatedmis, 'Misspelled': Misspelled - trunc, 'Truncated': Truncated Pronoun-Determiner (PD) **type** (Type (Subcategorization)) poss, 'Possessive': Possessive recip, 'Reciprocal': Reciprocal ONW CGN refl, 'Reflexive': Reflexive ONW CGN rel, 'Relative': Relative ONW Gysseling CGN w-p, 'W-pronoun': Interrogative/relative pronoun of the 'w'- d-p, 'D-pronoun': Demonstrative/article/relative pronoun of the 'd'-series dem. Demonstrative': Demonstrative ONW Gysseling CGN excl, 'Exclamative': Exclamative CGN excl, Exclamative : Exclamative Continuous indef, 'Indefinite': Indefinite ONW Gysseling CGN on the continuous inter, 'Interrogative': Interrogative ONW Gysseling CGN on the continuous inter, 'Interrogative' ONW Gysseling CGN on the continuous series uncl, 'Unclear': Unclear Gysseling CGN **subtype** (Subtype) art, 'Article': Article ONW Gysseling CGNoth, 'Other': Other ONW Gysseling CGN uncl, 'Unclear': Unclear Gysseling CGN **position** (Syntactic position of adjective, numeral or pronoun/determiner) prenom, 'Prenominal': Prenominal ONW Gysseling CGN postnom, 'Postnominal': Postnominal ONW free. 'Free': Free (adverbial or predicative) ONW Gysseling CGN uncl, 'Unclear': Unclear ONW Gysseling CGN case (Case) nom, 'Nominative': Nominative ONW gen, 'Genitive': Genitive ONW dat, 'Dative': Dative ONW — acc, 'Accusative': Accusative ONW - uncl, 'Unclear': Unclear ONW CGN LN (Number ('inherent' Lemma number)) sg, 'Singular': Singular ONW Gysseling CGNpl, 'Plural': Plural ONW Gysseling CGN uncl, 'Unclear': Unclear ONW Gysseling CGN **LP** (Inherent person (pronouns)) 3, '3': Third person ONW Gysseling CGN uncl, 'Unclear': Unclear ONW Gysseling CGN 1, '1': First person ONW Gysseling CGN 2, '2': Second person ONW Gysseling CGN **LG** ('Inherent' lemma gender) m, 'Masculine': Masculine CGN m/f, 'Masculine or feminine': Masculine or feminine f, 'Feminine': Feminine ONW CGN - n, 'Neuter': Neuter CGN — m/n, 'Masculine or neuter': Masculine or neuter — uncl, 'Unclear': Unclear ONW CGN f/n, 'Feminine or neuter': Feminine or neuter **number** (Number) sg, 'Singular': Singular ONW CGNpl, 'Plural': Plural ONW CGN uncl, 'Unclear': Unclear ONW CGN gender (Gender) m, 'Masculine': Masculine ONW CGNf, 'Feminine': Feminine ONW CGN — n, 'Neuter': Neuter ONW CGN — uncl, 'Unclear': Unclear ONW CGN inflection (Inflection) – 0, 'No inflection': No inflection Gysseling CGN — r/re, '-r/-re': Inflection -r or -re Gysseling CGN - s/th, '-s/-th': Inflection -s or -th Gysseling - t, '-t': Inflection -t Gysseling - oth, 'Other': Other values Gysseling - uncl, 'Unclear': Unclear CGN - a, '-a': Inflection -a Gysseling - e, '-e': Inflection -e Gysseling CGN - n, '-n': Inflection -n Gysseling CGN - nt, '-nt': Inflection -nt **WF** (Written form) - abbr, 'Abbreviated': Abbreviated - trunc, 'Truncated': Truncated - mis, 'Misspelled': Misspelled

Residual (RES)

type (Type (Subcategorization)) - aff, 'Affix': Affix — meta, 'Meta': Meta CGN - symb, 'Symbol': Symbol CGN - oth, 'Other': Other Gysseling CGN - uncl, 'Unclear': Unclear ONW' Gysseling CGN - bre, 'Bounded Root Element': Bounded Root Element — for, 'Foreign': Foreign word CGN — form, 'Formula': Formula inflection (Inflection) — 0, 'No inflection': No inflection Gysseling r/re, '-r/-re': Inflection -r or -res/th, '-s/-th': Inflection -s or -th Gysseling – a, '-a': Inflection -a– e, '-e': Inflection -eGysseling - t, '-t': Inflection -t n, '-n': Inflection -n Gysseling nt, '-nt': Inflection -nt oth, 'Other': Other values Gysseling uncl, 'Unclear': Unclear Gysseling **WF** (Written form) abbr, 'Abbreviated': Abbreviatedmis, 'Misspelled': Misspelled — trunc, 'Truncated': Truncated CGN Verb (VRB) **type** (Type (Subcategorization)) mai, 'Main': Main Gysselingcop, 'Copula': Copula ONW – aux, 'Auxiliary': Auxiliary ONW– uncl, 'Unclear': Unclear ONW Gysseling finiteness (Verb finiteness) fin, 'Finite': Finite forms inf, 'Infinitive': Infinitive prespart, 'Present participle': Present participle ONW CGN pastpart, 'Past participle': Past participle ONW CGN uncl, 'Unclear': Unclear ONW Gysseling CGN **tense** (Verb tense) pres, 'Present': present tense ONW Gysseling CGN past, 'Past': past tense ONW Gysseling CGN — uncl, 'Unclear': Unclear ONW Gysseling CGN **mood** (Mood (subjunctives and imperatives are not in the molex database)) ind, 'Indicative': Indicative ONW CGN conj, 'Conjunctive': Conjunctive ONW CGN imp, 'Imperative': Imperative ONW Gysselinguncl, 'Unclear': Unclear ONW Gysseling CGN **number** (Number) - sg, 'Singular': Singular ONW CGN uncl, 'Unclear': Unclear ONW CGN — pl, 'Plural': Plural ONW CGN' **person** (Person (verb)) 1, '1': First person ONW2, '2': Second person ONW 3, '3': Third person ONWuncl, 'Unclear': Unclear ONW valency (Valency) trans, 'Transitive': Transitive ONWintr, 'Intransitive': Intransitive ONW impers, 'Impersonal': Impersonaluncl, 'Unclear': Unclear ONW - refl, 'Reflexive': Reflexive ONW conjugation (Conjugation class) pretpres, 'Preterite present': Preterite present uncl, 'Unclear': Unclear ONW - strong, 'Strong': Strong ONW - weak, 'Weak': Weak ONW - irreg, 'Irregular': Irregular ONW verbclass (Verb class) 1, '1': 1 (blijven,stijgen, ...) 2, '2': 2 (vliegen, liegen, bieden, bedriegen) 3a, '3a': 3a (vinden, winnen) 3b, '3b': 3b (worden, erven) ONW 5, '5': 5 (wezen, bidden, geven, genezen) 6, '6': 6 (varen, dragen, slaan) 7, '7': 7 (vangen, hangen) uncl, 'Unclear': Unclear ONW — 4, '4': 4 (spreken, komen) ONW inflection (Inflection) r/re, '-r/-re': Inflection -r or -re Gysseling s/th, '-s/-th': Inflection -s or -th Gysseling t, '-t': Inflection -t Gysseling CGN oth, 'Other': Other values Gysseling - 0, 'No inflection': No inflection Gysseling CGN - a, '-a': Inflection -a Gysseling CGN - e, '-e': Inflection -e Gysseling CGN - n, '-n': Inflection -n Gysseling CGN - nt, '-nt': Inflection -nt Gysseling - uncl, 'Unclear': Unclear CGN **WF** (Written form) abbr, 'Abbreviated': Abbreviated mis, 'Misspelled': Misspelled

Declarations

ONW

pos=AA => position pos=AA => degree,position,case,number,gender pos=NOU-C => case,number,gender pos=NOU-P => case,number,gender pos=VRB,mood=imp => finiteness pos=VRB,mood=imp => tense pos=VRB => type,finiteness,valency,conjugation pos=VRB,conjugation=strong => verbclass pos=VRB,conjugation=uncl => verbclass pos=VRB,finiteness=fin => mood,tense,number pos=VRB,finiteness=uncl => mood,tense,number pos=VRB,finiteness=fin,mood=ind => person pos=VRB,finiteness=fin,mood=conj => person pos=VRB,finiteness=fin,mood=uncl => person pos=VRB,finiteness=uncl,mood=uncl => person pos=PD,subtype=art => position pos=PD,type=pers => position pos=PD,type=refl => position pos=PD,type=recip => position pos=PD,type=rel => position pos=PD => type,case,position pos=PD,type=uncl => number,gender pos=PD,type=dem => number,gender pos=PD,type=inter=> number,gender pos=PD,type=rel=> number,gender pos=PD,type=indef=> number,gender pos=PD,type=recip=> LN,LG pos=PD,type=pers => LP,LN pos=PD,type=refl => LP,LN pos=PD,type=poss => LP, number, gender pos=PD, type=dem,position=uncl => subtype pos=PD, type=indef,position=uncl => subtype pos=PD, type=dem,position=prenom => subtype pos=PD, type=indef,position=prenom => subtype pos=PD, type=uncl,position=prenom => subtype pos=PD, type=uncl,position=uncl => subtype pos=PD, type=pers, LP=3, LN=sg => LG pos=PD, type=refl, LP=3, LN=sg => LG pos=PD, LP=uncl, LN=uncl => LG pos=PD, LP=uncl, LN=sg => LG pos=CONJ => type pos=ADP => type, government pos=ADV => type, subtype pos=NUM => type,position,case,number,gender pos=NUM => representation pos=RES => type

Gysseling

pos=NOU-C => number, inflection
pos=NOU-P => inflection
pos=AA => degree, position, number, inflection
pos=VRB => type, finiteness, inflection
pos=VRB,finiteness=fin => tense,mood
pos=VRB,finiteness=uncl => tense,mood
pos=NUM => type,position,representation,inflection
pos=PD => type,position
pos=PD, type=pers => LP, LN
pos=PD, type=refl|recip => LP, LN
pos=PD, type=uncl => LP, LN, inflection
pos=PD, type=dem => subtype

pos=PD, type=uncl => subtype pos=PD, type=dem => inflection pos=PD, type=indef => subtype pos=PD, type=indef => inflection pos=PD, type=rel => inflection pos=PD, type=inter => inflection pos=PD, type=jos= > inflection pos=ADV => type,subtype,inflection pos=ADV => type,inflection pos=CONJ => type,inflection pos=INT => inflection pos=RES => type,inflection

CGN

pos=NOU-C => number, inflection pos=NOU-C, number=sg => gender,case pos=NOU-C, number=uncl => gender,case pos=NOU-P => number, inflection pos=NOU-P, number=sg => gender,case pos=NOU-P, number=uncl => gender,case pos=AA =>degree, position, case, inflection pos=VRB => finiteness, inflection pos=VRB,finiteness=fin => tense,mood,number pos=VRB,finiteness=uncl => tense,mood,number pos=NUM => type,position,representation,inflection pos=NUM, position=prenom => case pos=NUM, position=uncl => case pos=PD => type,position,case,inflection pos=PD, type=pers => LP, LN pos=PD, type=pers, LP=3, LN=sg => LG pos=PD, type=pers, LP=3, LN=uncl => LG pos=PD, type=pers, LP=uncl, LN=uncl => LG pos=PD, type=pers | refl => LP, LN pos=PD, type=pers refl, LP=3, LN=sg => LG pos=PD, type=refl => LP, LN pos=PD, type=refl, LP=3, LN=sg => LG pos=PD, type=refl => LP, LN pos=PD, type=refl, LP=3, LN=uncl => LG pos=PD, type=refl, LP=uncl, LN=uncl => LG pos=PD, type=recip => LN, LP pos=PD, type=poss => LN, LP, number, gender pos=PD, type=poss, LP=3, LN=sg => LG pos=PD, type=poss, LP=uncl, LN=uncl => LG pos=PD, type=rel => number,gender pos=PD, type=rel|inter, position=free => number, gender pos=PD, type=rel inter, position=prenom => number,gender pos=PD, type=inter, position=free => number, gender pos=PD, type=inter, position=prenom => number, gender pos=PD, type=inter, position=uncl => number, gender pos=PD,type=excl=> number,gender pos=PD, type=dem, position=free => number, gender pos=PD, type=dem, position=prenom => number, gender pos=PD, type=dem, position=uncl => number, gender pos=PD, type=dem, position=prenom => subtype pos=PD, type=dem, position=uncl => subtype pos=PD, type=indef, position=free => number,gender pos=PD, type=indef, position=prenom => number,gender pos=PD, type=indef, position=uncl => number, gender pos=PD, type=indef, position=prenom => subtype pos=PD, type=indef, position=uncl => subtype pos=ADP => type pos=ADV => type pos=CONJ => type pos=RES => type