

/instituut voor de
Nederlandse taal/

jaarverslag

2025

verslag aangaande de werkzaamheden,
gebeurtenissen enz. in het afgelopen jaar;
jaarlijks verslag

(Woordenboek der Nederlandsche Taal)

Jaarverslag 2025
Instituut voor de Nederlandse Taal

Inhoudsopgave

1. Inleiding: het jaar 2025	6
2. Language Resources Repository en CLARIN-centrum.....	8
Language Resources Repository	8
Geografische verspreiding van gebruikers.....	8
Het INT als CLARIN-centrum	9
CLARIN-België.....	9
CLARIN-Nederland	9
3. Corpusinfrastructuur.....	9
Monitorcorpus	10
Historische corpora.....	10
Verrijking.....	10
Verrijking met woordsoort en lemma	11
Syntactische annotatie.....	11
Metadata.....	11
Informatie-extractie	11
4. Beschrijving van de woordenschat door de eeuwen heen.....	12
Datamodel voor de centrale kennisbank van de woordenschat.....	12
Het centrale lexicon (GiGaNT)	12
Betekenisregister	13
Lexicografische producten, gebruikersapplicaties, API's en datasets	13
Woordenlijst.org	13
Algemeen Nederlands Woordenboek (ANW).....	14
Woordenboek van Nieuwe Woorden (WNW).....	14
Woordcombinaties	15
Historische woordenboeken	15
Vertaalwoordenschat.....	16
Woordpeiler	16
API's en datasets.....	17
5. Beschrijving van de Nederlandse dialecten	17
Elektronische Woordenbank van de Nederlandse dialecten (eWND).....	17
Database van de Zuidelijk-Nederlandse Dialecten (DSDD).....	17
Digitale infrastructuur voor het Bildts en het Overijssels.....	17

Overig.....	17
6. Expertisecentrum voor Nederlandstalige Terminologie.....	18
Veldondersteuning.....	18
Termenlijsten	18
Nederlands als wetenschapstaal.....	18
JuriTerm	19
HOTNeV.....	19
Terminologierubrieken op het INT/ENT-web	20
TermWerk	20
Terminologisch netwerk	20
7. Grammatica.....	21
Grammaticaportaal.....	21
Taaladvies	21
e-ANS.....	21
Taalportaal	22
8. Nationale en internationale samenwerkingsverbanden	22
Netwerken	22
IMPACT Centre of Competence	22
European Language Data Space.....	22
ELEXIS Association.....	23
Netwerkprojecten.....	23
Universality, diversity and idiosyncrasy in language technology (UniDive, 2022-2026)	23
European Network on Lexical Innovation (ENEOLI, 2023-2027)	23
Onderzoeks- en infrastructuurprojecten	24
CLARIAH-Vlaanderen (2021-2025).....	24
CLARIAH-VL+ (2025-2028).....	24
SSHOC-NL (Social Science and Humanities Open Cloud for the Netherlands, 2024-2029)	24
SABeD (2021-2024)	25
PressMint (1 juni 2025 tot en met 31 mei 2027)	25
Geparset Corpus van de Zuidelijk-Nederlandse Dialecten (2024-2028).....	25
Pilotproject Duidelijke Taal (2023-2024)	25
Spread the News (2020-2025)	26
Buitenlandse Aardrijkskundige Namen – fase 1	26
Namenloket (1-03-2025 – 28-02-2027)	26
Macroscopic (2026-2030).....	27
Digitalisering BasiScript (2025)	27

Overige infrastructurele dienstverlening.....	27
Infrastructuur Papiamento/u.....	27
Etymologiebank	28
GLAD.....	28
DAGENTA	28
Pallas	28
9. Systeembeheer en ICT-capaciteit	28
10. Disseminatie.....	29
Doelgroepenbeleid	29
Onderzoekers en taalkundigen	29
Onderwijs.....	29
Algemeen publiek	29
Communicatiemiddelen.....	30
Website	30
Huisstijl.....	33
Nieuwsbrieven & persberichten	34
Sociale media	34
Podcasts	34
Andere populairwetenschappelijke activiteiten.....	35
Bijlage 1: Raad van Toezicht en Raad van Advies	37
Bijlage 2: Medewerkers	37
Bijlage 3: publicaties, lezingen, media, prijzen etc.	39
Publicaties	39
Lezingen en presentaties	45
Actieve deelname en organisatie van congressen en workshops	49
Onderwijs.....	50
In de media	51
Diversen	52
Bijlage 4: Taalmaterialen.....	54
B4.1 Downloads van taalmaterialen.....	54
B4.2 Toevoegingen en updates aan Taalmaterialen.....	57

1. Inleiding: het jaar 2025

Dit jaarverslag geeft een overzicht van de werkzaamheden en resultaten in 2025 van het Instituut voor de Nederlandse Taal, hét kennis-, infrastructuur- en onderzoeksinstituut voor het Nederlands in het Nederlandse taalgebied, dat wil zeggen Nederland, Vlaanderen, Suriname en de Cariben. Deze werkzaamheden werden uitgevoerd in het kader van het [Meerjarenbeleidsplan](#) 2023-2027.

Een groot deel van de werkzaamheden vloeit voort uit een belangrijk zwaartepunt van dit plan, de integratie van alle componenten van de woordenschatbeschrijving in één centrale, modulair georganiseerde, relationele kennisbank van de Nederlandse woordenschat door de eeuwen heen. Vandaaruit worden bestaande en nieuwe lexicografische producten verder ontwikkeld. Er werd gewerkt aan een versterking van de corpusworkflow en het corpusgebaseerde lexicografische proces, resulterend in een verbinding tussen primaire corpusdata en afgeleide lexicale data. Hoofdstuk 4 geeft een uitgebreide beschrijving van deze werkzaamheden. De andere hoofdstukken laten zien dat er daarnaast nog tal van andere projecten worden uitgevoerd op het gebied van grammatica, dialecten, terminologie.

Large Language Models (LLM's) spelen een steeds grotere rol in het werk van het INT. In 2025 werd onderzocht hoe LLM's kunnen worden ingezet in het lexicografische werk. Dit zal voortgezet worden in het Horizon Europe-project ELEXAI, dat in 2025 werd aangevraagd door een Europees consortium van taalinstellingen en inmiddels is toegewezen. INT is ook partner in het in 2025 aanvrage en gehonoreerde infrastructuurproject Cultur-AI. In 2025 werd verder bekend dat het NWO Large Scale Research Infrastructure-project Macroscopic gehonoreerd is. Het INT zal hierin een bijdrage leveren aan het bouwen van een mediacorpus.

Andere projectaanvragen met AI waarbij het INT betrokken is zijn FIAT (inclusieve automatische transcriptie) en CLARIAH Prime, en er zijn nog meer projectvoorstellen in voorbereiding. Wat hierbij helpt is dat het INT in december 2025 erkend is als onderzoeksinstituut door NWO, hetgeen de betrokkenheid bij aanvragen vergemakkelijkt. Het recente advies van de Raad voor de Nederlandse Taal en Cultuur voorziet een centrale rol voor het INT in de ontwikkeling van infrastructuur voor LLM's, benchmarking van LLM's, en kennisdeling over LLM's. Twee medewerkers van het INT, Vincent Vandeghinste en Bram Vanroy, zijn lid van de Taaluniecommissie die dit advies moet uitwerken.

Het INT heeft in 2025 zijn rol in de Europese infrastructuurnetwerken verder versterkt. Vincent Vandeghinste is Nationaal Coördinator voor CLARIN en vertegenwoordigt België in het National Coordinators Forum. Carole Tiberius heeft dezelfde rol voor Nederland. Vincent Vandeghinste is daarnaast lid van het Knowledge Infrastructure Committee van CLARIN-ERIC, en lid van de CLARIN Board of Directors. Het INT werkt verder binnen de European Language Data Space, waar het de verantwoordelijkheid voor Nederland neemt.

De onderzoekers van het INT deelden hun kennis in 2025 in wetenschappelijke publicaties, lezingen, workshops en webinars. Het INT heeft ook taalprofessionals en het bredere publiek als doelgroep. Het besteedde in 2025 veel aandacht aan het bereiken van deze doelgroepen, onder andere met populairwetenschappelijke publicaties (waaronder het zeer gunstig ontvangen eindejaarsgeschenk *Taaltaart*), zes nieuwsbrieven, negen podcasts, en volop aanwezigheid in de media. De diagrammen in hoofdstuk 9 geven een goed beeld van de mate waarin het INT zijn publiek bereikt.

Het INT speelt een centrale rol in onderzoek en infrastructuur voor het Nederlands, is een gewilde partner in projectaanvragen, en is succesvol in het verwerven van subsidies. De financiële situatie blijft echter een belangrijk punt van zorg. De jaar in jaar uit teruglopende structurele financiering is een serieus gevaar voor de ontwikkeling van het instituut. Gezien de veelheid aan taken is de belastbaarheid van medewerkers een voortdurend punt van aandacht.

In de personeelsformatie waren er in 2025 enkele belangrijke wijzigingen. Op 1 september begon Sjef Barbiers als nieuwe directeur van het INT. Hij is de opvolger van Frieda Steurs, die op 1 oktober na negen jaar met pensioen ging. Frieda kreeg bij haar afscheid de erepenning van de Taalunie en de Matthias de Vriespenning vanwege haar grote verdiensten voor de Nederlandse taal. Zij blijft als gastonderzoeker verbonden aan het instituut.

Bij de aanvang van 2025 werkten er 33 personeelsleden (27,2 fte). Eind december 2025 waren dat 37 personeelsleden (29,4 fte). Per 1 oktober begonnen twee jonge taalkundige onderzoekers die als taak hebben om bij te dragen aan de uitbouw van de centrale kennisbank voor de Nederlandse woordenschat.

Sinds december 2025 huurt het INT het pand aan het Rapenburg 61 niet meer van de Universiteit Leiden maar rechtstreeks van de eigenaar. Als gevolg daarvan moesten ook veel van de diensten van de universiteit herzien worden. Het INT is in gesprek met de Universiteit Leiden over de huisvesting na 2028.

Februari 2026

prof. dr. Sjef Barbiers

wetenschappelijk directeur-bestuurder

2. Language Resources Repository en CLARIN-centrum

Language Resources Repository

Via de website taalmaterialen.ivdnt.org stelt het Instituut voor de Nederlandse Taal (INT) bronnen, data en tools beschikbaar voor taalkundig onderzoek en taal- en spraaktechnologie (TST), hoofdzakelijk gericht op het Nederlandse taalgebied. De Taalmaterialenwebsite kreeg 32.718 pageviews van 14.651 verschillende gebruikers. Dat is voor beide getallen een stijging van meer dan 20%.

Een aantal van die tools en resources zijn ook beschikbaar vanuit het INT CLARIN-portal, op <https://portal.clarin.ivdnt.org>. Deze site kreeg 16.143 pageviews (+210%) van 3100 (+20%) verschillende bezoekers.

Het totaal aantal downloads voor zowel commerciële als niet-commerciële producten is licht gedaald ten opzichte van vorig jaar, maar ligt wel hoger dan de voorgaande jaren:

- Niet-commerciële downloads: 1015 (1091 in 2024, 816 in 2023, 522 in 2022)
- Commerciële downloads: 35 (44 in 2024, 20 in 2023, 5 in 2022)

Een gedetailleerde lijst over welke taalmaterialen hoe vaak zijn gedownload is te vinden in Bijlage B4.1. Welke nieuwe toevoegingen en updates er in 2025 geweest zijn, is te vinden in Bijlage B4.2.

Om taalmaterialen en het CLARIN-portal te ontdubbelen werd er onderzoek verricht naar welke repository-omgevingen geschikt zijn voor het INT. Hiervoor werd aanvankelijk Dataverse getest, om uiteindelijk voor DSpace 8 te kiezen. Daarvan werd een eerste testversie opgezet, die in 2026 verder aangevuld en getest wordt.

Er kwamen 149 helpdeskvragen en -verzoeken binnen in 2025. Tot midden maart vallen daar enkel de vragen en verzoeken i.v.m. de taalmaterialen onder. Vanaf midden maart vallen daar ook de mails van servicedesk@ivdnt.org onder (en vanaf dat moment is het taalmaterialen-adres van de website en uit de licenties verwijderd), en worden deze vragen intern bijgehouden in een databank.

Geografische verspreiding van gebruikers

Gebruikers dienen zich te registreren alvorens ze producten kunnen downloaden. Daarbij wordt ook hun e-mailadres vastgelegd. Aan de hand van de extensie van die adressen kan een beeld worden verkregen vanuit welke landen belangstelling bestaat voor onze taalmaterialen. Hieronder een tabel met alle extensies die vaker dan 4 keer voorkomen.

Extensie	Aantal
com	448
nl	245
be	97
de	20
org	12
net	8
eu	7
ai	5

Het INT als CLARIN-centrum

CLARIN-België

Voor 2025 vervulde Vincent Vandeghinste voor het INT de rol van Nationaal Coördinator voor CLARIN op Europees niveau, als vertegenwoordiger voor België in het National Coordinators Forum (NCF). Vincent is ook lid van het Knowledge Infrastructure Committee van CLARIN-ERIC. Vanaf 1 januari 2025 trad Vincent ook toe tot de CLARIN Board of Directors, met een mandaat voor 2 jaar. Daarnaast nam Jesse de Does vanuit het INT de vertegenwoordiging van België op zich in het Standing Committee on CLARIN Technical Centres. Het INT vervulde een actieve rol als liaison tussen Belgische onderzoekers en de Europese CLARIN-infrastructuur en tussen de Vlaamse onderzoekers in CLARIAH-VL+ en de CLARIN-infrastructuur van het INT als CLARIN-B-centrum voor België. Vincent was lid van het Programme Committee voor de CLARIN-conferentie in 2025 te Wenen. K-Dutch, het CLARIN Knowledge Centre for Dutch, werd verder uitgebreid, waarbij de website verder werd uitgebreid en verfijnd. K-Dutch kreeg 546.592 unieke bezoekers (stijging van meer dan 30.000%) maar aangezien meer dan 542.000 hiervan uit China of Singapore kwamen, is het allicht interessanter om te weten dat er 954 bezoekers uit Nederland of België kwamen.

CLARIN-Nederland

Sinds 1 juli 2025 vervult Carole Tiberius de rol van CLARIN Nationaal Coördinator voor CLARIN op Europees niveau, als vertegenwoordiger voor Nederland in het National Coordinators Forum (NCF).

3. Corpusinfrastructuur

Zorgvuldig samengestelde en wetenschappelijk onderbouwde corpora vormen een essentieel onderdeel van de taalinfrastructuur voor het Nederlands. Op basis van deze primaire taaldata kan de Nederlandse taal worden gedocumenteerd en kunnen taalapplicaties ontwikkeld worden. De corpusinfrastructuur van het INT bevat naast corpora ook tools voor dataprocessing en ontsluiting.

Een belangrijk deel van de werkzaamheden aan de corpusinfrastructuur wordt uitgevoerd voor de verdere uitbouw van de centrale kennisbank voor de Nederlandse woordenschat, maar resulteert ook in corpusinfrastructuur voor de brede onderzoeksgemeenschap. Daarnaast is het INT betrokken in diverse projecten waarin corpora worden gebouwd, waarbij het INT, naast expertise ook infrastructurele ondersteuning biedt voor het bouwen, gebruiken of ter beschikking stellen van het corpusmateriaal.

Monitorcorpus

Het Corpus Hedendaags Nederlands (CHN) bevat al het hedendaags Nederlandse corpusmateriaal dat bij het INT beschikbaar is. Het corpus is in 2025 continu uitgebreid met krantenmateriaal uit Nederland, Vlaanderen, Suriname en de Caribische rijkdelen. Voor de interne versie wordt deze update wekelijks uitgevoerd, voor de externe versie elke maand. Het interne corpus bevat momenteel ruim 10,75 miljoen documenten en bijna 4,5 miljard woorden. Verder is er gewerkt aan een fundamentele herimplementatie van de corpusprocessing-workflow, het corpusmanagementsysteem en de daaraan gekoppelde automatische indexeringscomponent. Deze herimplementatie houdt onder andere in het afzonderlijk toegankelijk maken van de corpusmetadata voor handmatige nabewerking in een daarvoor ontwikkelde omgeving, het verder uitbreiden van de metadata conform het CMDI-metadatamodel en het implementeren van een incrementeel indexeringsysteem voor het CHN. De verbeterde corpusworkflow wordt in 2026 in gebruik genomen.

Historische corpora

In 2025 is een nieuwe versie van het *Couranten Corpus* online gepubliceerd, met verbeterde taalkundige verrijking conform de [TDN](#)-kerntagset. Daarnaast is er gewerkt aan de uitbreiding van het krantenmateriaal van dit corpus. Het afgelopen jaar heeft een groep van ongeveer 20 enthousiaste vrijwilligers onder leiding van Nicoline van der Sijs en met ondersteuning van het INT ervoor gezorgd dat alle 2733 afleveringen van de *Oprechte Haerlemsche Courant* zijn getranscribeerd en gecorrigeerd. Daarna zijn deze vrijwilligers hetzelfde gaan doen voor enkele kleinere 17e-eeuwse kranten. De data zullen uiteindelijk onderdeel gaan vormen van een groot diachroon corpus van kranten. Daarnaast is er verder gewerkt aan de uitbreiding van het trainings- en evaluatiemateriaal voor de verrijking met woordsoort en lemma. Verder is het *Historical Corpus of Dutch*, een gebalanceerd corpus dat ontwikkeld is door onderzoekers van de Vrije Universiteit Brussel en de Universiteit Leiden, online gepubliceerd. Het *Language of Leiden Corpus* is voorbereid voor online publicatie in 2026. Tot slot is er gewerkt aan een nieuwe release van alle online historische corpora met de nieuwe versie van de software waarmee deze corpora toegankelijk worden gemaakt (zie onder Informatie-extractie), waarbij ook alle manuals zijn bijgewerkt. De technische update zal plaatsvinden in het voorjaar van 2026.

Verrijking

De INT-corpora worden samengesteld uit al bestaand digitaal materiaal of, waar nodig, uit door digitalisering verkregen materiaal. De brondata worden geconverteerd naar eenzelfde XML-standaard (TEI) en zorgvuldig van metadata voorzien en daarna automatisch taalkundig verrijkt.

Metadata en taalkundige verrijking bieden een nadrukkelijke meerwaarde om zinvolle informatie uit de corpora te kunnen extraheren. De werkzaamheden van het afgelopen beleidsjaar worden hieronder gespecificeerd.

Verrijking met woordsoort en lemma

Voor de verrijking van historisch Nederlands is verder gewerkt aan de ontwikkeling van goudstandaarddatasets voor de verrijking met woordsoort en lemma. Inmiddels zijn ruim 760.000 woorden (863.543 tokens), gespreid over de 13e tot 19e eeuw, volledig nagekeken conform de [lemmatiseerprincipes](#) en de [TDN](#)-kerntagset.

Syntactische annotatie

Op grond van reeds in 2023 verricht voorwerk is besloten de *Universal Dependencies*-richtlijnen aan te houden voor syntactische annotaties.

In 2025 is verder gewerkt aan de annotatieprincipes. Als gevolg van uitgebreid overleg met Gertjan van Noord en Gosse Bouma over de implementatie van Universal Dependencies, heeft Gosse Bouma een aantal aanpassingen en verbeteringen geïmplementeerd in het conversietraject van Lassy-annotatie naar UD-annotatie. Daarnaast is de implementatie van UD-parsing in GaLAHaD doorontwikkeld, en is er een prototype van een query-builder voor niet-technische gebruikers ontwikkeld.

Metadata

De corpora hebben een gemeenschappelijk metadataformaat, met ruimte voor subcorpus-specifieke metadata. Ten behoeve van verdere uniformering en systematisering is gewerkt aan een CMDI-tekstprofiel dat toepasbaar is voor alle corpusdata van het INT. In 2025 is gewerkt aan de implementatie van het CMDI-tekstprofiel voor het CHN (zie boven) .

Informatie-extractie

Het corpusmateriaal wordt toegankelijk gemaakt via een applicatie waarmee in de corpora gezocht kan worden. Wanneer de IPR (Intellectual Property Rights) het toelaten, wordt het corpusmateriaal ook als dataset beschikbaar gesteld in de language resource repository. De software is open-source beschikbaar.

Voor de ontwikkeling van de corpusapplicatie, die bestaat uit de corpuszoekmachine [BlackLab](#) en de [corpusfrontend](#), ligt de prioriteit bij de ondersteuning van de diverse INT-taken. Deze werkzaamheden worden ondersteund door samenwerking met diverse partijen die de software gebruiken, en door de mogelijkheden die externe projecten bieden voor verdere ontwikkeling van de software.

In 2025 is er gewerkt aan het wegwerken van wat in de software-engineering “technical debt” genoemd wordt, om ervoor te zorgen dat de ontwikkelde software onderhoudbaar en toekomstbestendig is. Hieronder vallen aanpassingen aan nieuwe versies van softwaretools als Java,

Vue en Tomcat en verbetering van codekwaliteit en documentatie. Voor BlackLab is daarnaast onder andere gewerkt aan optimalisatie, incrementeel indexeren en het ondersteunen van externe metadata, en uitgebreide mogelijkheden voor het zoeken op delen van documenten en relaties. Voor het BlackLab-frontend is er – naast het ondersteunen van deze uitgebreide zoekmogelijkheden, het wegwerken van technical debt en refactoring – gewerkt aan de standaardisering van de corpusconfiguraties en het opzetten van de documentatiewebsite. Ook is het indexeren en publiceren van corpora significant vereenvoudigd door nieuwe indexeerscripts. Tot slot is er gewerkt aan een *prototype query builder* voor syntactisch zoeken.

4. Beschrijving van de woordenschat door de eeuwen heen

Zoals uiteengezet in het Meerjarenbeleidsplan 2023-2027 is de koppeling van verschillende lexicografische databanken op zowel vorm- als betekenisniveau de belangrijkste innovatie en de grootste uitdaging voor de woordenschatbeschrijving in de huidige beleidsperiode.

In wat volgt beschrijven we de activiteiten die in 2025 voor de verschillende onderdelen van de centrale kennisbank en de daaraan gekoppelde lexicografische producten zijn uitgevoerd.

Datamodel voor de centrale kennisbank van de woordenschat

Voor de uitbreiding van het datamodel werd voortgebouwd op de al bestaande onderdelen in de huidige infrastructuur, met name het centrale lexicon GiGaNT (cf. infra), de koppeling op lemma-niveau met de lexicografische databanken en het Diachroon seMantisch lexicon van de Nederlandse Taal (DiaMaNT). Er werd in kaart gebracht welke bijkomende datacategorieën nodig zijn, welke koppelingen binnen de kennisbank wenselijk zijn en welke daarbuiten (bv. met corpusdata). In 2025 zijn de datamodellen van GiGaNT geëvalueerd en aangepast en is begonnen met de implementatie van die aanpassingen. Daarnaast is er een datamodel uitgewerkt, voor de koppeling tussen GiGaNT en de corpusdata uit het CHN (cf. infra).

Het centrale lexicon (GiGaNT)

In 2025 zijn de reguliere werkzaamheden aan het centrale lexicon voortgezet. Deze houden in: het onderhoud aan zowel de historische component (GiGaNT-Hilex) als de moderne component (GiGaNT-Molex), het verder werken aan de integratie van deze componenten en de uitbreiding van GiGaNT-Molex ten behoeve van de diverse lexicografische eindproducten die een koppeling met GiGaNT hebben (cf. infra). Afgezien van de toevoeging van nieuwe lexemen aan GiGaNT-Molex, is er ook verder gewerkt aan het toevoegen van uitspraakinformatie, met aandacht voor de uitspraak in zowel Nederland als Vlaanderen. Inmiddels zijn van de in totaal ruim 300.400 lexemen in Molex er ruim 299.980 van handmatig geverifieerde uitspraakinformatie voorzien, goed voor 378.648 verschillende uitspraken. Tot slot is er gewerkt aan de implementatie van het gereviseerde datamodel van GiGaNT.

De onderliggende databank van de in 2024 gelanceerde applicatie Woordpeiler (cf. infra) vormde in 2025 het uitgangspunt voor de ontwikkeling van een frequentiemodule (Frequency Lexicon Service, FreqLeS) die frequentiegegevens uit het CHN dynamisch aan GiGaNT-Molex koppelt en zo aan de hele centrale woordenschat-infrastructuur. De frequentiemodule wordt ook gebruikt voor twee nieuwe databanken: In ExoLex worden ontbrekende kandidaat-lemmata voor GiGaNT-Molex gerangschikt naar opnameprioriteit op basis van verschillende types frequentiegegevens uit het CHN. In ChloroLex worden die frequentiegegevens samengebracht met gegevens uit GiGaNT-Molex om een wetenschappelijk onderbouwde selectie te maken voor het Groene Boekje dat in 2026 verschijnt. De frequentiemodule is bovendien generiek gemaakt zodat ze ook voor historische corpora en GiGaNT-Hilex ingezet kan worden.

Tot slot is er ook gewerkt aan een technische upgrade van de infrastructuur voor de bewerking van structurele data, Lex'it. Dat heeft onder andere geleid tot een publieke [release](#) van dit Rapid Application Developmentplatform.

Betekenisregister

In 2025 werd de proefversie van het betekenisregister verder uitgewerkt. Voor de hedendaagse woordenschat werd een semi-automatische selectie gemaakt van ruim 2500 lemmata met een betekenisbeschrijving in een van de lexicografische databanken van het INT (ANW, Woordcombinaties en RBN). Vervolgens werd er een manueel gevalideerde koppeling gemaakt tussen het betekenisregister enerzijds en de lexicografische databanken anderzijds. Daarna werden aangepaste definities toegevoegd voor gebruik in de specifieke eindproducten zoals het ANW. Het datamodel werd verder verfijnd met het oog op de ontwikkeling van een eigen editeeromgeving voor het betekenisregister in 2026. In het kader van de aanvraag van het Horizon Europe InfraDev-project ELEXAI werd bovendien in kaart gebracht hoe het betekenisregister voor het Nederlands verder ontwikkeld kan worden binnen de Europese lexicografische infrastructuur. Deze aanvraag werd recent goedgekeurd en het ELEXAI-project start in de 2de helft van 2026 (cf. infra). Ten slotte werden op basis van ANW en WNW datasets aangemaakt voor het finetunen en testen van Large Language Models (LLMs) voor betekenisbeschrijvingstaken die in de toekomst vanuit het betekenisregister ondersteund zullen worden. Zo werden de datasets intern gebruikt voor experimenten met automatische definitie-generatie van neologismen en in samenwerking met de Master in AI aan de KU Leuven voor experimenten met het expanderen van glossen naar definities en met het vereenvoudigen van definities voor tweedetaalleiders.

Lexicografische producten, gebruikersapplicaties, API's en datasets

Woordenlijst.org

Ook dit jaar is er weer gewerkt aan de uitbreiding van Woordenlijst.org. Dat resulteerde in een update in november 2025. Dat betekent dat er nu 233.258 lemmata en 845.573 paradigmavormen online staan. Een volgende technische update is ook voorbereid, die in het voorjaar van 2026 zal worden uitgevoerd. Daarnaast is er in het najaar gestart met de voorbereidende werkzaamheden voor het nieuwe Groene Boekje. Er is gewerkt aan de criteria voor selectie, een workflow en tools om de selectie te kunnen doen en een omgeving om mogelijke selecties te kunnen evalueren. Tot

slot heeft het INT ook dit jaar weer ondersteuning geboden aan de Commissie Spelling bij haar werkzaamheden.

Algemeen Nederlands Woordenboek (ANW)

Het *Algemeen Nederlands Woordenboek* (ANW) is een corpusgebaseerd, digitaal, multimediaal woordenboek van het eigentijdse Nederlands in Nederland en de Caribische rijkdelen, Vlaanderen en Suriname. De taalperiode die het ANW bestrijkt, loopt van 1970 tot heden. Naast de woorden uit de kernwoordenschat van de genoemde taalgebieden wordt in het ANW ook een selectie van door ons verzamelde neologismen (nieuwe woorden, nieuwe verbindingen, nieuwe uitdrukkingen, nieuwe betekenissen van al bestaande woorden) opgenomen.

Door de redactie zijn in 2025 568 volledig bewerkte, nieuwe trefwoorden ('artikelen') online gezet. Daarvan waren 385 trefwoorden neologismen, die ook in het *Woordenboek van Nieuwe Woorden* (WNW) terechtkwamen. Eén trefwoord kan meerdere gelijkvormige woorden beschrijven. Het gaat dan bijvoorbeeld om woorden waarvan de woordsoort of etymologische herkomst verschilt. In de huidige versie van het ANW zijn in totaal ruim 300.000 woorden opgenomen, die behandeld zijn in ruim 87.000 trefwoorden. De trefwoorden bevatten in totaal ruim 47.000 betekenissen.

Alle nieuwe woorden zijn door de redactie ook (toegevoegd en) gelinkt aan GiGaNT-Molex, het grote computationele lexicon van het hedendaags Nederlands dat ook de basis vormt van de Woordenlijst Nederlandse Taal en het Groene Boekje.

De redactie is verder gevorderd met het selecteren van betekenissen uit Woordcombinaties, het Referentiebestand Nederlands (RBN) en de Vertaalwoordenschat om die toe te voegen aan ANW-lemmata die nog geen definitie hebben.

De redactie schrijft mee aan de webrubriek 'Nieuw woord van de week', waarin elke week een actueel neologisme wordt besproken.

Woordenboek van Nieuwe Woorden (WNW)

Het INT onderzoekt de vorming van nieuwe woorden en houdt die woorden onder andere bij in woordenboeken. Sinds 2018 wordt er op het INT gewerkt aan het *Woordenboek van Nieuwe Woorden* (WNW), een onlinewoordenboek waarin woorden die vanaf het jaar 2000 zijn ontstaan, worden beschreven. Het WNW neemt niet alleen de beklijvende nieuwe woorden op (bijvoorbeeld *app*, *selfie*, *keuzestress* en *ontspullen*), maar juist ook veel nieuwe woorden die kort bestaan (zoals *appaire* en *stookschaamte*). Ook in 2025 is er verder gewerkt aan dit woordenboek. Het WNW bevatte eind 2025 bijna 24.000 woorden. Het WNW wordt dagelijks geüpdatet.

Alle nieuwe woorden zijn door de redactie ook (toegevoegd en) gelinkt aan GiGaNT-Molex, het grote computationele lexicon van het hedendaags Nederlands dat ook de basis vormt van de Woordenlijst Nederlandse Taal en het Groene Boekje.

Ook afgelopen jaar is er iedere week een neologisme behandeld in de rubriek 'Nieuw woord van de week'.

Woordcombinaties

Woordcombinaties is een online taaltool waarin het gebruik van woorden in context wordt beschreven en beschikbaar gesteld aan taalgebruikers en leerders van het Nederlands. Het project integreert de beschrijving van collocaties, idiomen en semantisch gemotiveerde syntactische patronen en behandelt de combinatoriek van werkwoorden en substantieven. Het project is corpusgebaseerd en gebruikt de Sketch Engine als corpus-querytool. Woordcombinaties toont hoe woorden gebruikt worden in goede voorbeeldzinnen, welke woorden met elkaar gecombineerd worden (collocaties of combinatiemogelijkheden) en hoe valentiepatronen en andere patronen samen met collocaties gebruikt worden voor het bouwen van zinnen. In Woordcombinaties worden ook idiomen en conversatieformules geïnventariseerd en systematisch bewerkt.

In het verslagjaar zijn patronen bewerkt van werkwoorden uit de schooltaalwoordenschat en het *Frequentiewoordenboek Nederlands* (Tiberius en Schoonheim 2014). Daarnaast zijn er ook voorbeeldzinnen en combinaties bewerkt van werkwoorden en substantieven. De redactie is in september uitgebreid. In het laatste kwartaal is daarom tijd besteed aan het opleiden van de nieuwe krachten.

Op computationeel gebied is onderzoek gedaan naar de mogelijkheden van AI voor het ondersteunen van de manuele annotatie en identificatie van de patronen binnen Woordcombinaties. Zo is onderzocht of LLM's ingezet kunnen worden voor het automatisch toekennen van semantische types aan argumenten in patronen. De resultaten zijn door Carole Tiberius en Jesse de Does gepresenteerd in een keynote op de eLex 2025-conferentie in Slovenië.

Om Woordcombinaties in het onderwijs onder de aandacht te brengen, zijn toepassingen beschreven in het online Handboek Didactiek Nederlands.

Historische woordenboeken

De beschrijving van de historische woordenschat van het Nederlands is te vinden in de verschillende historische woordenboeken van het INT. Deze woordenboeken zijn online beschikbaar in het historische-woordenboekenportaal (gtb.ivdnt.org): het *Oudnederlands Woordenboek* (ONW), het *Vroegmiddelnederlands Woordenboek* (VMNW), het *Middelnederlandsch Woordenboek* (MNW) en het *Woordenboek der Nederlandsche Taal* (WNT).

Om de rijkdom van deze historische woordenboeken en van de webapplicatie waarin ze te raadplegen zijn nog meer onder de aandacht te brengen, werd net als in voorgaande jaren in de webrubriek 'Terug in de taal' een twintigtal bijdragen rond aansprekende historische woorden en begrippen gepubliceerd.

Op verschillende manieren zijn de Historische Woordenboeken – en andere historische taalmaterialen van het INT – in de schijnwerpers geplaatst. Zo is er voor erfgoedmedewerkers in Vlaanderen een webinar over het gebruik van de historische-woordenboekenapplicatie georganiseerd, dat door deelnemers in een enquête positief werd beoordeeld. Daarnaast zijn tijdens presentaties de woordenrijkdom van deze woordenboeken en de onderzoeksmogelijkheden ervan voor het voetlicht gebracht. Uit de log-gegevens én uit de reacties en vragen die het INT regelmatig zowel van specialistische als niet-specialistische gebruikers ontvangt, blijkt dat de webapplicatie ondertussen een veelgebruikt werkinstrument geworden is.

De woordenboekdata vormen de kern van de historische-lexiconmodule van GiGaNT (GiGaNT-Hilex). GiGaNT-Hilex bevat de volgende woordenboekelementen: historisch lemma, modern Nederlands equivalent, woordsoort van een lemma, geattesteerde woordvormen van het lemma in de citaten en de citaten zelf. De werkzaamheden die ook in 2025 in dit kader zijn uitgevoerd, leveren correcties op die voor de online woordenboeken eveneens relevant zijn.

Citatenmateriaal van zowel MNW en WNT is gebruikt als trainings- en evaluatiemateriaal voor het verrijken van historisch corpusmateriaal (zie boven).

Vertaalwoordenschat

De *Vertaalwoordenschat* is een online-applicatie voor tweetalige woordenboeken. Via deze online-applicatie worden de tweetalige bestanden ontsloten die, onder meer in opdracht van de Commissie Lexicologische Vertaalvoorzieningen (CLVV, 1993-2003), zijn ontwikkeld voor taalparen die op de commerciële markt niet spontaan aan bod kwamen.

In september 2017 heeft het INT de *Vertaalwoordenschat*-applicatie gelanceerd. Inmiddels staan het Nederlands-Nieuwgrieks / Nieuwgrieks-Nederlands, het Nederlands-Portugees / Portugees-Nederlands, het Nederlands-Estisch, het Nederlands-Fins / Fins-Nederlands, het Nederlands-Deens / Deens-Nederlands, en het Nederlands-Arabisch / Arabisch-Nederlands online. In 2025 is het Nederlands-Indonesisch toegevoegd aan de *Vertaalwoordenschat*. De digitale versie van het Nederlands-Indonesisch ondersteunt zowel taalproductie als taalreceptie. Bij verschillende trefwoorden is extra informatie beschikbaar voor de Nederlandstalige of de Indonesischtalige gebruiker. Op deze manier zijn de twee woordenboeken Nederlands-Indonesisch, één voor Nederlandstaligen en één voor Indonesischtaligen, die eerder gedrukt zijn verschenen op basis van het bestand, in één digitale versie verenigd. Verder is in 2025 door een externe redacteur gewerkt aan het updaten van de bestanden Nederlands-Arabisch/Arabisch-Nederlands. De woordenboeken zijn gratis raadpleegbaar via de website van de *Vertaalwoordenschat*.

Woordpeiler

De applicatie *Woordpeiler* maakt het mogelijk om voor elk woord te visualiseren hoe het gebruik in de verschillende taalvariëteiten van het Nederlands door de tijd heen geëvolueerd is. De grafieken zijn gebaseerd op het krantenmateriaal van het Corpus Hedendaags Nederlands en onderzoekers met een CLARIN-login kunnen vanuit *Woordpeiler* ook de onderliggende gegevens raadplegen. Daarnaast is er voor intern gebruik een module voor trendanalyses ontwikkeld die toelaat om nieuwe of kenmerkende woorden in kaart te brengen voor een willekeurig tijdsinterval uit de periode 2000 tot nu. In 2025 werd de applicatie verder ontwikkeld met een uitbreiding naar n-grammen en een nieuwe zoekinterface voor geavanceerd zoeken. Binnen het instituut is *Woordpeiler* geïntegreerd in de eigen woordenschatbeschrijving. Naast de directe toepassing binnen de neologismen-detectie vormt de onderliggende databank (FreqIDb) ook de basis voor het koppelen van frequentiegegevens aan GiGaNT-MoLex (cf. supra). Daarnaast werd een instantie voor de historische kranten corpora (Courantenpeiler) ontwikkeld. De toepassing wordt daarnaast ook gebruikt voor wetenschapscommunicatie over trends in de woordenschat naar het brede publiek (rubriek *Woordpeiling*). *Woordpeiler* werd ook voorgesteld op congressen (CLIN, eLex) en gepromoot via een webinar tijdens de Week van het Nederlands.

API's en datasets

Ook in 2025 werd maandelijks de lexiconservice geüpdatet waarmee GiGaNT-Hilex toegankelijk wordt gemaakt voor gebruik in bijvoorbeeld corpusapplicaties als Delpher, het Corpus Middelnederlands of de Gekaapte Brieven.

5. Beschrijving van de Nederlandse dialecten

Elektronische Woordenbank van de Nederlandse dialecten (eWND)

In 2025 zijn met behulp van vrijwilligers enkele nieuwe woordenboeken toegevoegd aan de elektronische Woordenbank van de Nederlandse Dialecten (eWND).

Database van de Zuidelijk-Nederlandse Dialecten (DSDD)

In 2025 zijn in de Database van de Zuidelijk-Nederlandse Dialecten (DSDD) bestaande concepten verder nagekeken op onregelmatigheden. Op basis daarvan zijn de nodige correcties aangebracht, of zijn dubbele concepten opgespoord, verwijderd of ont dubbeld. Het *Woordenboek der Zeeuwse dialecten* is verder bewerkt voor opname in de DSDD. Eind 2025 bevatte de DSDD ongeveer 31.000 concepten (met ongeveer 600.000 trefwoorden).

Digitale infrastructuur voor het Bildts en het Overijssels

In 2025 hebben de vrijwilligers van het Bildts Aigene materiaal voor het Bildts toegevoegd en verder bewerkt. In 2025 werd gestart met een vergelijkbare infrastructuur voor het Overijssels. Het doel is op basis van bestaande gekoppelde woordenboeken een zakwoordenboek te maken voor het West-Overijssels. In het kader hiervan werden drie bestaande Overijsselse amateurdialectwoordenboeken gedigitaliseerd en in Lex'it verder bewerkt. De gegevens werden gekoppeld aan de bestaande Nederlandse lijst van het Bildtse project en uitgebreid met nieuwe begrippen uit de drie bestaande Overijsselse woordenboeken. Voor beide projecten is een zoekomgeving gebouwd waarin de data gepubliceerd zullen worden.

Overig

In 2025 verscheen tweewekelijks 'Uit de Streek', een rubriek over dialectwoorden uit de DSDD en de woordenbanken. Een deel van die columns werd herwerkt en opgenomen in het boekje *Taaltaart*, het nieuwjaarsboekje dat in 2025 door het INT werd uitgegeven (zie verder). De projectleider werkte ook mee aan de tweemaandelijks rubriek over streektaalen in het tijdschrift *Onze Taal*. In drie tijdschriften worden regelmatig artikelen van de projectleider opgenomen (*Nehalennia*, *De diejalektgazette* en *Zeeuws Erfgoed*). Het INT heeft in 2025 ook weer Limburgs cursusmateriaal toegevoegd, aangepast en beschikbaar gesteld via de website Limburgsluisteren.ivdnt.org op verzoek van Levende Talen Limburgs.

6. Expertisecentrum voor Nederlandstalige Terminologie

Het Expertisecentrum Nederlandstalige Terminologie (ENT) richt zich primair op veldondersteuning en op positionering in een terminologisch netwerk. Voor beide kerntaken is in 2025 het volgende gerealiseerd.

Veldondersteuning

Termenlijsten

Termenlijsten zijn een essentieel onderdeel van het lexicon van een taal. Ze dienen concrete doelen waar ze voorzien in de omschrijving en afbakening van begrippen voor de beoefening van een vak, in onderwijsverband bij de instructie van leerstof of waar ze in meertalig verband vertaalsuggesties aanleveren voor beoefenaren van taalberoepen. Op een algemener en abstracter niveau vormen ze ook een graadmeter die uitdrukt in hoeverre een taal beschikt over expressiemiddelen buiten de algemene communicatiebehoeften. De ENT-webpagina voorziet daarom in een rubriek die het mogelijk moet maken om Nederlandstalige eentalige of meertalige termenlijsten binnen een enkele verzamelplaats te raadplegen. De ordening gebeurt volgens de onderwerpencatalogus van de Library of Congress, een internationaal systeem dat ook veel Nederlandse universiteitsbibliotheken hanteren om hun collectie te classificeren. De uitbreiding gebeurt gestaag. In 2025 werden ruim 100 nieuwe termenlijsten toegevoegd.

In 2025 is de ENT-webpagina ook uitgebreid met een nieuwe rubriek voor historische termenlijsten. Deze opdeling is ingegeven vanuit een streven naar toegankelijke en overzichtelijke termenverzamelingen. Taalprofessionals die bezig zijn met actuele onderwerpen zullen veelal geen behoefte hebben om geregeld lijsten tegen te komen met vakjargon uit vroegere tijden. Daarentegen wekken ze des te meer de belangstelling van onderzoekers van historisch tekstmateriaal. Deze termenlijsten verdienen daarom een aparte rubriek die momenteel ruim 30 links naar online beschikbare termenverzamelingen telt. Ook hier wordt geclassificeerd volgens voornoemde onderwerpencatalogus van de Library of Congress.

Nederlands als wetenschapstaal

Naast deze collectie van termenlijsten die bestaat uit een groeiende verzameling links, host het ENT ook extern ontwikkelde termenbestanden zoals de termenlijsten in het kader van het project 'Nederlands als wetenschapstaal'. Onderzoek naar onderwijs en woordenschat heeft duidelijk gemaakt dat de opstap naar het eerste bachelorjaar een betere begeleiding wenselijk maakt voor zowel geschreven vaktaal als gesproken academisch Nederlands. Aansluitend bij deze behoefte werd een samenwerkingsverband opgestart tussen de Stichting Nederlands/Vlaams Platform Taalbeleid Hoger Onderwijs, de Taalunie, KU Leuven, UGent en het INT. Het proefproject 'Nederlands als wetenschapstaal - van corpora naar terminologielijsten' heeft als doel om taalhulpmiddelen te ontwikkelen die deze overstap van het middelbaar naar het hoger onderwijs makkelijker maken. Dit project leverde drie grote termenverzamelingen op voor scheikunde, wiskunde en natuurkunde. Ook

hier zorgde het INT voor terminologische en computationele begeleiding van de publicatie. Het [scheikundebestand](#) staat online en is zoekbaar. In 2024 is een script ontwikkeld om de vele interne verwijzingen tussen ingangen met een semi-automatische aanpak te optimaliseren. De toepassing daarvan begon eind 2024 en liep door in 2025. Scheikunde is inmiddels afgewerkt. Voor het wiskundebestand is de semi-automatische aanpak aangevat. Het is momenteel beschikbaar als pdf maar zal, evenals het natuurkundebestand, als termenbank worden gepubliceerd naar het voorbeeld van het online scheikundebestand.

JuriTerm

Het pilootproject JuriTerm NL-VL: Rechtsvergelijkende terminologiedatabank voor het Nederlands (01/02/2025–31/01/2026) heeft tot doel een wetenschappelijk onderbouwde en digitaal ondersteunde infrastructuur uit te bouwen voor rechtsvergelijkende terminologiebeschrijving tussen Nederland en Vlaanderen. Het project, uitgevoerd door het INT in samenwerking met de Taalunie, en met ondersteuning van de Vlaamse Juristenvereniging, het Tijdschrift voor Privaatrecht en de Stichting Ridder René Victor, beoogt een databank waarin juridische begrippen niet alleen worden gedefinieerd, maar ook systematisch in hun onderlinge relaties en in hun overeenkomsten en verschillen tussen beide rechtssystemen worden beschreven. Daarmee levert JuriTerm een instrument voor juristen, rechtslinguïsten en vertalers en sluit het aan bij de harmonisatie-doelstellingen van de Taalunie inzake rechts- en overheidsterminologie.

In de eerste projectfase werd een uitgebreid literatuuronderzoek uitgevoerd, resulterend in een interdisciplinaire methodologie die lokaal-ontologisch, empirisch en rechtsvergelijkend van aard is. Als uitgangspunt werd de Bolzano-methode gekozen, aangepast aan de Nederlandse en Vlaamse context en afgestemd op bestaande lexicografische bronnen. De methodologie werd besproken en goedgekeurd door een interdisciplinaire begeleidingscommissie van juristen en terminologen uit Nederland en Vlaanderen.

Parallel werd de digitale werkomgeving ontwikkeld. TermWerk werd uitgebreid tot JuriTermWerk, met bijkomende data-categorieën op begrips-, (rechts)systeem- en termniveau, waaronder begripsstatus, rechtsvergelijkende situering en opmerkingen over rechtsgevolgen. Daarnaast werd een eigen corpus Juridisch Nederlands samengesteld met tijdschriften uit beide rechtsgebieden, als basis voor empirische onderbouwing van definities en contexten. Voor de gebruikersinterface werden mock-ups ontwikkeld; een eerste versie van de online omgeving is voorzien begin 2026.

Inhoudelijk werden 232 termfiches volledig uitgewerkt en circa 100 bijkomende fiches voorbereid, binnen de domeinen verbintenissenrecht, arbeidsrecht en sociaalzekerheidsrecht.

Tot slot werd ingezet op disseminatie via lezingen, publicaties en een projectpagina op de INT-website. In 2026 wordt het project voortgezet en een verdere uitbreiding van de databank en versterking van de coöperatieve dimensie met rechtswetenschappelijke partners voorzien.

HOTNeV

Ten slotte was het ENT in 2025 ook zelf actief in het beschrijven van termen. In het HOTNeV-project (Hoger Onderwijs Terminologie in Nederland en Vlaanderen) worden onderwijstermen geanalyseerd en beschreven volgens de principes zoals die zijn vastgelegd in de internationale terminologie-

standaard ISO 704. De resultaten van het onderzoek van de termen, hun vertaalequivalenten en beschrijving worden opgeslagen in een database. Het onderzoek komt tot stand in het kader van stagebegeleiding en de resultaten worden opgeslagen in een tweetalige termenbank Nederlands en Brits-Engels met de terminologietool QTerm. Een export met verdere uitbreiding van de data was ook in 2025 beschikbaar via de INT-website Taalmaterialen.

Terminologierubrieken op het INT/ENT-web

Opnieuw zijn in 2025 vier uitgebreide nieuwsbrieven verstuurd die vervolgens op de website ter beschikking kwamen en daarna werden gearchiveerd. Ze informeren over terminologie in vaste rubrieken zoals een terugblik naar voorbije evenementen, nieuws over ontwikkelingen en publicaties, tools en bronnen, berichten van NL-Term en een agenda. In de loop van het jaar is ook de ENT-webpagina met de evenementenrubriek over Europese terminologiecongressen, -symposia en -workshops bijgehouden.

TermWerk

Op het vlak van terminologische applicaties heeft het INT in 2025 samen met de Taalunie een nieuwe tool onder de naam TermWerk gelanceerd. Het gaat hierbij om de integratie van de vroegere termextractor TermTreffer en de termbank-editor TermBeheerder, zodat één tool ontstaat die geschikt is voor de invoer, de verwerking en het beheer van termmateriaal. Het resultaat is een applicatie die in een breed, generiek verband en afgestemd op het Nederlands en zijn taaleigenschappen, de termextractiemodule en de editeerfunctionaliteiten van beide systemen combineert. Deze is door het INT zelf volledig ontwikkeld, rekening houdend met de vereisten van de gebruikers. De tool werd getest door studenten van de KU Leuven en van ITV-hogeschool en voorgesteld aan de leden van NL-TERM en van de Termraad. De functies van TermWerk worden doorlopend verder verbeterd. Zo werden binnen het TermWerk-plus-project in 2025 functies voor import en export van termbanken en voor de vergelijking van termlijsten met een termenbank toegevoegd. Ook beschikken gebruikers nu over meer flexibiliteit in de keuze en configuratie van beschrijvingsvelden in de termenbank.

Terminologisch netwerk

Om zich te positioneren in het terminologische netwerk nam het ENT in 2025 deel aan verschillende evenementen en samenwerkingsvormen. Zo participeerde het aan de driemaandelijke bijeenkomsten van de Termraad Nederlands. Op dit overlegplatform voor terminologische afstemming binnen het Nederlandse taalgebied werken EU-terminologen samen met Belgische en Nederlandse partners uit overheidsdiensten, terminologieverenigingen en vertaalopleidingen. Het doel van de bijeenkomsten is terminologische ondersteuning voor deze instellingen door de verzameling, beschrijving en uniformering van terminologie in specifieke vakgebieden. Op de Termraadvergadering van 26 september in Luxemburg gaf het ENT een presentatie over termextractie en termbeheer, enerzijds op semi-automatische wijze met Word en Excel, en anderzijds op automatische wijze met de nieuw ontwikkelde tool TermWerk. Opnieuw was er samenwerking tussen het INT en NL-Term op de jaarlijkse TiNT-dag (Terminologie in het Nederlandse Taalgebied) die plaatsvond in Antwerpen op 14 november. Het ENT belichtte er zijn

werkzaamheden vanuit het kader van de ISO-standaard 29383 over terminologiebeleid. Van 27 tot 29 november nam het ENT deel aan het tweejaarlijkse internationale congres van de European Association for Terminology. Deze editie vond plaats in Innsbruck, en het ENT participeerde in de panellezingen. Ook droeg het INT centraal bij aan de organisatie en de uitgave van het vierde deel in de reeks *Handbook of Terminology* dat op deze conferentie werd voorgesteld. 'Terminology planning in Europe' biedt per taalgebied een overzicht van de terminologische stand van zaken. Aan dit boek droeg het ENT voor het Nederlands bij met een hoofdstuk onder de titel 'Efforts and challenges in translating concept to reality: The Dutch case'. Ten slotte vond stagebegeleiding ditmaal plaats voor twee masterstudenten van de Universiteit Leiden.

7. Grammatica

Grammaticaportaal

De testversie van het grammaticaportaal, de hub voor de grammaticaonderdelen van het INT, is verrijkt met extra koppelingen met Taaladvies.net en de e-ANS. Daarnaast is er in een pilot onderzocht hoe de site mede toegankelijk zou kunnen worden gemaakt met behulp van een RAG-chatbot die gevoerd is met de beschikbare grammaticadata.

Taaladvies

Aan Taaladvies.net, een zeer drukbezochte site die onder beheer is bij het INT, is afgelopen jaar regulier onderhoud gepleegd. Er is onderling contact tussen de redacties van Taaladvies.net en de ANS om de kruisverwijzingen bij herziene hoofdstukken te corrigeren.

e-ANS

Voor al achter de schermen is er dit jaar veel werk verzet aan de herziening van de e-ANS. Er is gewerkt aan de inhoudelijke revisie van de hoofdstukken 22 (Passieve zinnen), 23 (Soorten zinnen naar communicatieve functie), 24-26 (Nevenschikking) en 27 (Samentrekking). Daarnaast is er gewerkt aan de redactie van hoofdstukken 5 (Het voornaamwoord), 18 (De werkwoordelijke constituent) en 20 (De zinsdelen, waarvan 20.3, 20.6-7). Deze hoofdstukken worden inhoudelijk in overeenstemming gebracht met de huidige inzichten in de betreffende onderdelen van de Nederlandse grammatica. Daarnaast worden ze op verschillende manieren verrijkt, volgens de doelstellingen van de herziening van de e-ANS, met actuele literatuurverwijzingen, koppelingen naar taaladviesbronnen, en voorbeelden uit het Corpus Hedendaags Nederlands en het Corpus Gesproken Nederlands.

Voor de hoofdstukken 5, 18 en 20 zijn daarnaast onderwijsmodules gemaakt, die tegelijk met de hoofdstukken zullen worden gepubliceerd (verwachting: 2026). Deze modules bevatten samenvattingen, termenlijsten, veelgestelde vragen en lesideeën. Ze zijn speciaal ontwikkeld voor docenten en leerders van de Nederlandse grammatica, voor wie het niveau van de ANS vaak net te hoog ligt. Ook verscheen in 2025 de online rubriek Grammaticasafari, waarin op luchtige toon

onderwerpen uit de ANS voor een breed publiek uitgelicht worden. Een aantal van deze columns zijn gebundeld in het boek *Taaltaart*, dat met enkele INT-collega's is geschreven.

Er is regulier onderhoud gepleegd en de toegankelijkheid van de website is verbeterd door toevoeging van een onderwerpenregister.

Taalportaal

Het onderdeel Nederlandse syntaxis van Taalportaal bestaat uit het werk *Syntax of Dutch*. Van dit werk is een nieuwe editie in de maak, die ook op Taalportaal zal gaan verschijnen. Van deze nieuwe versie zijn, in samenwerking met de auteur, Hans Broekhuis, de bronbestanden in Word omgezet naar een testversie in Taalportaal-xml. Daarnaast is dr. Camiel Hamans in overleg begonnen aan een uitbreiding van de Nederlandse morfologie.

8. Nationale en internationale samenwerkingsverbanden

Netwerken

IMPACT Centre of Competence

Het INT is voorzitter van het IMPACT Centre of Competence (www.digitisation.eu). Dit is een non-profitorganisatie bestaande uit publieke en commerciële organisaties met als doel de digitalisering van historisch materiaal “beter, sneller, en goedkoper” te maken. Het centrum voorziet in data, tools, services en expertise op het gebied van document imaging, taaltechnologie en het verwerken van historisch tekstmateriaal. Het IMPACT Centre of Competence is sedert 2019 ook CLARIN Knowledge Centre.

European Language Data Space

De European Language Data Space (LDS) is door de EU officieel aangesteld om een Europees dataplatform en Europese datamarkt te ontwikkelen voor het verzamelen, creëren, uitwisselen en (her)gebruiken van meertalige en multimodale taaldata. De basis hiervoor is gelegd in eerdere initiatieven zoals het European Language Resource Coordination-initiatief (ELRC) en de European Language Grid (ELG). Ook binnen de European Language Data Space is het INT het nationale aanspreekpunt. Op 23 januari 2025 vond de gezamenlijke LDS Country Workshop voor België en Nederland plaats in Breda. Er waren 52 deelnemers uit België en Nederland (vertegenwoordigers van bedrijven, overheidsinstanties en onderzoeksinstituten). Vooral de panelsessies, die tot veel discussie leidden, werden gewaardeerd.

ELEXIS Association

Het INT was partner in ELEXIS, een Europees Horizon 2018-2022-project, waarbinnen een infrastructuur voor e-lexicografie is opgezet. Om de duurzaamheid van de infrastructuur te waarborgen is een nieuwe associatie in het leven geroepen, de ELEXIS Association. Doel van de associatie is het organiseren en coördineren van lexicografisch gerelateerde activiteiten en NLP-activiteiten voor zover die relevant zijn voor de lexicografie. Het INT is lid geworden van de associatie. Carole Tiberius is Deputy van de President van de ELEXIS Association. In 2025 is gewerkt aan een herziening van de Europese projectaanvraag omtrent LLM's en lexicografie (ELEXAI), die in september 2025 is ingediend en in februari 2026 is goedgekeurd. Het INT leidt de werkpakketten die LLM's integreren in de lexicografische workflow.

Netwerkprojecten

Universality, diversity and idiosyncrasy in language technology (UniDive, 2022-2026)

Het INT neemt deel aan het Europese onderzoeksnetwerk UniDive (Universality, diversity and idiosyncrasy in language technology). Het doel van deze COST-actie is om te onderzoeken hoe taaltechnologie verbeterd kan worden door te focussen op universaliteit, maar tegelijkertijd ook de diversiteit van talen in het oog te houden. Dit om te voorkomen dat talen met weinig middelen en bedreigde talen buitengesloten worden in de digitale wereld. Het INT is actief binnen twee werkgroepen (Corpusannotatie en Lexicon-Corpus Interface). In 2025 is een eerste versie van het Nederlandse PARSEME-corpus gereleased (als onderdeel van PARSEME corpus 2.0).

European Network on Lexical Innovation (ENEOLI, 2023-2027)

Het INT neemt deel aan het Europese COST-onderzoeksnetwerk European Network On Lexical Innovation (ENEOLI). De belangrijkste doelstellingen van het netwerk kunnen als volgt worden samengevat: 1) een gemeenschappelijke, meertalige kernterminologie voor lexicale innovatie definiëren; 2) digitale methodes en instrumenten in kaart brengen die gebruikt worden om lexicale innovaties in Europese talen te identificeren en te verklaren; 3) vergelijkende studies uit te voeren naar lexicale innovaties in Europese talen, met speciale aandacht voor ontleningen en hun equivalenten; 4) specifieke training in neologie aan te bieden aan vertalers, redacteurs, journalisten, technisch schrijvers en docenten. Het INT heeft in 2025 vooral bijgedragen aan het in kaart brengen van internationaal gangbare lexicografische selectiecriteria voor neologismen met het oog op het verbeteren van de eigen neologismenworkflow. Daarnaast was het INT ook betrokken bij de 1ste ENEOLI Training School in Lissabon in juni.

Onderzoeks- en infrastructuurprojecten

CLARIAH-Vlaanderen (2021-2025)

Binnen het FWO-IRI-project 'CLARIAH-VL: Advancing the open humanities service infrastructure' was de hoofdtak van het INT het voorzien van de benodigde infrastructuur voor het opzetten van de Digital Text Analysis Dashboard & Pipeline. In samenwerking met KU Leuven werd hiervoor de automatische corpusannotatieservice [TextLens](#) gebouwd.

CLARIAH-VL+ (2025-2028)

De vervolgaanvraag *CLARIAH-VL+ op weg naar de 'SSH Open Science Cloud' in Vlaanderen*, werd eind december 2024 gehonoreerd door het FWO. Het INT behoudt hierin min of meer dezelfde rol, nl. het voorzien van de benodigde infrastructuur voor het uitbouwen van TextLens tot een nieuwe tool voor treebank-querying in universal dependency treebanks. De taalkundige kant van het project werd gepresenteerd op de CLARIN-conferentie in Wenen.¹

SSHOC-NL (Social Science and Humanities Open Cloud for the Netherlands, 2024-2029)

Dit vervolproject van CLARIAH Plus beoogt te komen tot een consortium van onderzoeksinfrastructuren, gericht op het creëren van een ecosysteem van diensten, gegevens en instrumenten voor de sociale wetenschappen en menswetenschappen. Het consortium wordt geleid door ODISSEI, de Nederlandse nationale infrastructuur voor sociale wetenschappen, en CLARIAH, de Nederlandse nationale infrastructuur voor geesteswetenschappen. Het INT werkt samen met de groep van Antske Fokkens van de VU aan infrastructuur voor het methodologisch verantwoord inzetten van machinelearning en AI voor dataverrijking.

Als vervolgstap in de gebruikerstudie naar het potentieel van automatische dataverrijkingstools voor de SSH-onderzoeksgemeenschap hebben we, voortbouwend op de eerdere enquête, gerichte een-op-eeninterviews gehouden om bestaande kennislacunes te aan te vullen. Parallel hieraan is verder gewerkt aan ondersteuning voor methodologisch verantwoorde dataverrijking. Het GaLAHaD-platform is uitgebreid met een toolgerichte vergelijking van named-entity-annotaties en bijbehorende geaggregeerde statistieken (geïnspireerd op eerder werk aan de VU), en met exportmogelijkheden voor named entities, afhankelijkheidsrelaties en alinea- en zinsstructuur naar diverse bestandsformaten.

Het INT heeft een hands-on workshop over GaLAHaD en LAnCeLoT georganiseerd, met deelnemers van verschillende partijen uit Nederland, België en Duitsland. In het kader van evaluatie is gewerkt aan de ontwikkeling van benchmarks voor de beoordeling van LLM-prestaties op uiteenlopende voor het lexicografisch proces relevante taken, waaronder het genereren van woordenboekdefinities op basis van voorbeeldzinnen, word sense disambiguation en word sense induction, en diverse taken

¹ Vincent Vandeghinste, Walter Daelemans, Luna De Bruyne, Tim Van de Cruys and Els Lefever (2025). CLARIN in CLARIAH-VL+: Plans for 2025-2028. In [Proceedings of the CLARIN Annual Conference](#), Vienna. pp. 118-122.

gerelateerd aan Corpus Pattern Analysis. De resultaten worden in 2026 gepubliceerd. Daarnaast is er een [projectwebsite](#) ingericht als centrale hub met informatie en verwijzingen naar lopende werkzaamheden en evenementen.

In het kader van training is er in april een [workshop](#) over de platforms GaLAHaD en LAnCeLoT georganiseerd, waarin deze door het Instituut voor de Nederlandse Taal ontwikkelde tools voor automatische verrijking en handmatige verfijning van linguïstische annotaties in historische corpora zijn geïntroduceerd. Tevens is in november een [workshop](#) over methodologische reflectie en toepassingen van LLM's binnen de Sociale en Geesteswetenschappen verzorgd, met een theoretische ochtendsessie en een praktische middagsessie.

SABeD (2021-2024)

Van het Spoken Academic Belgian Dutch Corpus van de KU Leuven werd in september 2025 een online doorzoekbare [versie](#) beschikbaar gesteld. Deze werd 691 keer geraadpleegd in 2025.

PressMint (1 juni 2025 tot en met 31 mei 2027)

Het project PressMint: Interoperable Corpora of Historical Newspapers is een CLARIN-flagship-project dat tot doel heeft een meertalige, vergelijkbare, geannoteerde, vertaalde en interoperabele verzameling corpora van Europese historische kranten uit het begin van de 20^e eeuw samen te stellen. Het INT draagt hieraan bij door de conversie en taalkundige verrijking van onder meer het Couranten Corpus.

Geparset Corpus van de Zuidelijk-Nederlandse Dialecten (2024-2028)

Dit project is een vervolg op het project Gesproken Corpus van de Zuidelijke Nederlandse Dialecten (GCND) en voorziet in een aanzienlijke geografische uitbreiding waardoor het corpus het hele Europees-Nederlandse gebied zal bevatten. Het project gebruikt de bestaande audio-gealigneerde transcripties en annotaties uit de eerste fase van het GCND om NLP-tools te hertrainen en zo de syntactische annotatie van de nieuwe data te versnellen. De rol van het INT is om de voor het GCND gebouwde infrastructuur om de data te ontsluiten hierop aan te passen en uit te breiden.

In 2025 is verder gewerkt aan de workflow voor data-processing, aan het implementeren van de resultaten van Universal Dependency parsing in het corpus en de corpusapplicatie, en aan het implementeren van zoekmogelijkheden naar sprekerfragmenten, zodat zoekresultaten op een correcte wijze gefilterd kunnen worden op eigenschappen van de sprekers.

Pilootproject Duidelijke Taal (2023-2024)

De resultaten van het pilootproject Duidelijke Taal werden verder geanalyseerd en in 2025 gepresenteerd op de CLARIN-conferentie² in Wenen en de Plain Language Experience in Brussel. Een uitgebreide publicatie is ingediend en onder review.

² **Vincent Vandeghinste, Bram Vanroy, Job van Doeselaar (2025).** Human Evaluation of Automated Text Simplification through Crowdsourcing. In [Proceedings of the CLARIN Annual Conference](#). Wenen, pp.143-147.

Spread the News (2020-2025)

2025 is het laatste jaar van het project ‘Spread the new(s). Understanding standardization of Dutch through 17th-century newspapers’, nadat de einddatum is verlengd, onder andere vanwege zwangerschaps- en ouderschapsverlof van promovenda Machteld de Vos, die gedeeltelijk op het INT en gedeeltelijk bij de Radboud Universiteit werkte. In 2025 zijn de laatste artikelen, de inleiding en de conclusie ingeleverd bij de promotores Nicoline van der Sijs en Helen de Hoop, en in januari 2026 is de tekst van het proefschrift naar de manuscriptcommissie gestuurd.

Buitenlandse Aardrijkskundige Namen – fase 1

Het is de bedoeling om de website <https://namen.taalunie.org> te vernieuwen. Op deze website zijn alle buitenlandse aardrijkskundige namen waarvoor een Nederlandse vorm bestaat, te vinden. De schrijfwijze en het gebruik van buitenlandse aardrijkskundige namen in het Nederlands is gestandaardiseerd en de Commissie Aardrijkskundige Namen van de Nederlandse Taalunie beheert en actualiseert de inhoud van de website. In fase 1 van het project wordt de data van de BAN-lijst in een database-omgeving ondergebracht zodat de BAN-lijst gemakkelijk onderhouden en geüpdatet kan worden. In 2025 is het datamodel voor de database verder verfijnd en is er begonnen met het onderbrengen van de gegevens in de nieuwe verbeterde structuur. Via een daarvoor ontwikkelde applicatie wordt de data verder geoptimaliseerd.

Namenloket (1-03-2025 – 28-02-2027)

Het project Namenloket werd in 2024 voorbereid door Variaties vzw - Koepelorganisatie voor Dialecten en Oraal Erfgoed in Vlaanderen, de vakgroep Taalkunde van de UGent en het INT, en het werd goedgekeurd in januari 2025. Het project wil een digitaal interactief informatieportaal bouwen over het Vlaams namen-erfgoed, de eigennamen die Vlamingen vanouds aan zichzelf, maar ook o.a. aan plaatsen en dieren geven. Bezoekers zullen er een antwoord vinden op vaak gestelde vragen over eigennamen en zullen ook vragen kunnen stellen. Via een dataplatform zullen erfgoedvrijwilligers en -onderzoekers hun eigen namenverzamelingen digitaal kunnen ontsluiten. Het portaal zal ook didactisch materiaal bevatten. Partners in het project naast de hogergenoemde initiatiefnemers (Variaties vzw, INT en vakgroep Taalkunde UGent) zijn Histories vzw, de Koninklijke Commissie voor Toponymie en Dialectologie, Didactica (Antwerp School of Education, Universiteit Antwerpen), Meemoo (Vlaams instituut voor het archief) en de Universiteit Keulen. Het INT is naast de bouw van de website verantwoordelijk voor het opzetten van een dataplatform met zoekmodule waarop erfgoedvrijwilligers en onderzoekers hun eigen dataverzamelingen (bv. namenwoordenboeken en -glossaria) kunnen uploaden. Op die manier kunnen ze hun werk digitaliseren, verduurzamen, beter doorzoekbaar maken en voor een breder publiek beschikbaar stellen. Dit project ontwikkelt de digitale infrastructuur en protocollen nodig voor die integratie en maakt bij wijze van aanzet één dataverzameling – het monumentale *Woordenboek van de familienamen in België en Noord-Frankrijk* van Frans Debrabandere – voor het eerst volledig digitaal beschikbaar en doorzoekbaar voor het brede publiek. In 2025 is de website gebouwd. Er is gewerkt aan de databewerking van het *Woordenboek van de familienamen in België en Noord-Frankrijk* en aan een omgeving om automatisch bewerkte data manueel te verifiëren.

Macroscopic (2026-2030)

Het INT neemt deel aan het NWO Large-Scale Research Infrastructure-project *Macroscopic*, dat beoogt een onderzoeksinfrastructuur op te zetten die ons in staat kan stellen om complexe sociale en culturele processen te onderzoeken. Het project is een initiatief van twee grote Nederlandse onderzoeksinfrastructuren: ODISSEI, gericht op sociale en economische data, en CLARIAH, dat culturele en linguïstische archieven beheert.

Macroscopic stelt onderzoekers in staat om op een veilige manier omvangrijke datasets met elkaar te combineren en te analyseren, verspreid over sociale, culturele en digitale domeinen van de gehele Nederlandse bevolking. Het INT is betrokken bij de ontwikkeling van een omvattend Netherlands Media Corpus en is daarnaast verantwoordelijk voor een gebruikersvriendelijke ontsluiting daarvan door middel van een gebruikersvriendelijke annotatie- en analyse-infrastructuur (de Tekstsuite) die – in samenwerking met LUCL – mede door de inzet van AI-technieken tot nieuwe onderzoeksmogelijkheden zal leiden.

Digitalisering BasiScript (2025)

Eind 2024 werd door het archief van de Radboud Universiteit (RU) aan het INT gevraagd of er interesse was om het papieren BasiScript-archief (door kinderen geschreven tekstjes) over te nemen, gezien het verlopen van de standaard bewaartermijn van 10 jaar bij de RU. Hoewel een deel van deze gegevens al eerder getranscribeerd en voor onderzoek beschikbaar gesteld werden op Taalmaterialen, bleek een groot deel nooit getranscribeerd te zijn. Dit deel van de dataverzameling dreigde verloren te gaan. Het INT heeft van de Taalunie extra financiering gekregen om deze 120.000 pagina's door een externe commerciële partner te laten scannen. De resulterende pdf's zijn nu beschikbaar bij het INT, waarbij zal worden gekeken naar automatische transcriptie en anonimisering. In alle geval is de data behoeft voor vernietiging.

Overige infrastructurele dienstverlening

Infrastructuur Papiamentu/u

Op 1 februari 2024, de Internationale Dag van de Moedertaal, hebben het INT en het Nationaal Taalinstituut Curaçao (NTIC) een samenwerkingsverband gesloten dat vastgelegd is in een Memorandum of Understanding. Het doel is om samen te werken aan de uitbouw van de digitale lexicale infrastructuur voor het Papiaments. Het INT brengt zijn kennis en expertise in op het gebied van (lexicale) databases en bouwt en host de infrastructuur.

In 2026 start een vervolgproject, dat zich met name richt op de ontwikkeling van een Schooltaalwoordenlijst in Papiamentu als Cultureel Erfgoed voor het Onderwijs op Curaçao. Ook hier bestaat de rol van het INT uit technische en lexicaal-infrastructurele ondersteuning.

Etymologiebank

Sinds 2020 zijn de [etymologiebank](#) en de [uitleenwoordenbank](#) in beheer van het INT. In 2025 hebben drie stagiairs een bijdrage geleverd aan de verrijking van de etymologiebank. Verder is in 2025 een projectplan opgesteld om de infrastructuur rond de etymologiebank te vernieuwen, dat in 2026 zal worden uitgevoerd. Dit is noodzakelijk om continuïteit, integratie met de lexicale infrastructuur van het INT en doorontwikkeling van de publieksapplicatie mogelijk te maken, aangezien zowel de huidige publieksinterface, de onderliggende software, als het databewerkingsproces verouderd en kwetsbaar zijn.

GLAD

Verder is het INT verantwoordelijk voor het hosten van de data van het Global Anglicism Database Network (GLAD); hiervoor is een online bewerkingsapplicatie met behulp van het Lex'it-platform ontwikkeld. De publieksversie met Engelse leenwoorden in 17 talen is [hier](#) te vinden. De Nederlandse gegevens van GLAD zijn in 2025 uitgebreid tot 10.430 trefwoorden. Drie stagiairs, waaronder een Russische, hebben aan GLAD meegewerkt.

DAGENTA

Het INT host de Database Geschiedenis Nederlandse Taalkunde ([DAGENTA](#)). Een Nijmeegse stagiaire heeft links gelegd tussen enkele grammatica's uit DAGENTA en de e-ANS.

Pallas

De [Digitale Pallas](#), een 18e-eeuws Russisch meertalig woordenboek, is door het INT in 2025 ondergebracht in een speciale Wordpress-site met extra informatie voor onderzoekers.

9. Systeembeheer en ICT-capaciteit

Er is in 2025 flink gesleuteld aan de IT-infrastructuur van het INT. In de backend zijn verschillende opslagsystemen vervangen en efficiënter gemaakt waardoor we netto meer opslagruimte tot onze beschikking hebben. De servicedesk en de webredactie zijn opgenomen in het centrale servicemanagementsysteem. Er is flinke voortgang gemaakt met het vervangen van commerciële software door opensourcevarianten. Er is ook gewerkt aan het vervangen van de taalmaterialenwebsite door een DSpace-instantie.

Er zijn dit jaar geen nieuwe applicaties in beheer genomen; het totaal aantal online aangeboden services is nu 114. Beveiligingsincidenten zijn dit jaar beperkt gebleven tot 2 spoed-updates van software-onderdelen. De gemiddelde uptime van de extern aangeboden applicaties bedroeg 99,907%.

10. Disseminatie

Doelgroepenbeleid

Onderzoekers en taalkundigen

Het INT richt zich traditioneel op onderzoekers en taalkundigen. Bestaande contacten met onderzoekers, al dan niet verbonden aan wetenschappelijke instituten en universiteiten, werden onderhouden en waar mogelijk geïntensiveerd en uitgebreid. Taalkundigen van het INT verzorgden aan de KU Leuven het vak *Computationele Lexicologie* in de Master Taalkunde en Taal- en Letterkunde, en aan de Universiteit Leiden de collegereeks *Corpus Lexicography*. Ook werden er geregeld webinars, presentaties en lezingen gehouden op wetenschappelijke bijeenkomsten en congressen, werd er een workshop LLM's georganiseerd voor onderzoekers en sponsorde het INT de Taaldag van de Belgische Kring voor Linguïstiek (BKL).

Onderwijs

Een tweede belangrijke doelgroep is het onderwijs: leerlingen, studenten en docenten. Informatie voor deze doelgroep wordt bijeengebracht op de onderwijspagina's op de website van het INT. Daarnaast zijn de educatieve posters (Posters in de klas) geüpdatet en onder de aandacht gebracht bij docenten. Onder andere in de nieuwsbrief, op conferenties (HSN-conferentie) en festivals (Week van het Nederlands, Het Taalfeest). Er werden presentaties gehouden op het Colloquium Neerlandicum en het INT heeft opnieuw de Nederlandse opgave verzorgd voor de jaarlijkse Taalkunde-olympiade van de Universiteit Leiden. Daarnaast trad het INT als sponsor op voor de TABU Dag in Groningen, voor en door studenten taalwetenschap.

Algemeen publiek

Een derde doelgroep is het algemene publiek, de taalgeïnteresseerden en taalliefhebbers. Die doelgroep wordt op verschillende manieren bediend. Onder andere via de wekelijkse populairwetenschappelijke rubrieken op de website over grammatica, dialectwoorden, nieuwe woorden en historische woorden. Het INT is actief op verschillende sociale media en maakt op regelmatige basis nieuwe afleveringen van de podcastseries *Over taal gesproken* en *Waar komt pindakaas vandaan?* (die ook gebruikt worden als lesmateriaal). Het INT verstuurt tien keer per jaar twee soorten nieuwsbrieven aan abonnees. In de Week van het Nederlands organiseerde het INT een webinar over Woordpeiler. Daarnaast zijn in samenwerking met de Stichting Nederlandse Dialecten en de Taalunie twee dialect-webinars georganiseerd gericht op een breed publiek. Verder hielden medewerkers lezingen voor een divers publiek, waren zij te horen in radioprogramma's en droegen zij bij aan publicaties voor een algemeen publiek dat belangstelling heeft voor taal in het algemeen en het Nederlands in het bijzonder. Zie daarvoor ook het uitgebreide overzicht in bijlage 3.

Communicatiemiddelen

Website

De website ivdnt.org is vaak een van de eerste dingen die het publiek ziet van het INT, en dat maakt het een belangrijk communicatiemiddel. Traditioneel waren veel communicatie-uitingen erop gericht om bezoekers naar de website te trekken, maar online gedrag beweegt zich steeds meer naar 'zero-click': gebruikers krijgen hun antwoord op een zoekvraag direct in Google, in AI-tools of op sociale media zonder ook maar één keer door te klikken. Dat vraagt om een andere aanpak: minder doorverwijzen, de doelgroep direct bedienen op het platform zelf en zorgen dat INT-informatie goed gevonden wordt in zoekmachines, door AI-tools en op de sociale media.

Sinds 2023 worden de webstatistieken bijgehouden met Plausible. In 2025 telde de website 587.083 paginaweergaven (392.819 unieke bezoekers), een daling van zo'n 7% ten opzichte van het voorgaande jaar (626.482 paginaweergaven en 423.944 unieke bezoekers in 2024).



Webstatistieken uit Plausible

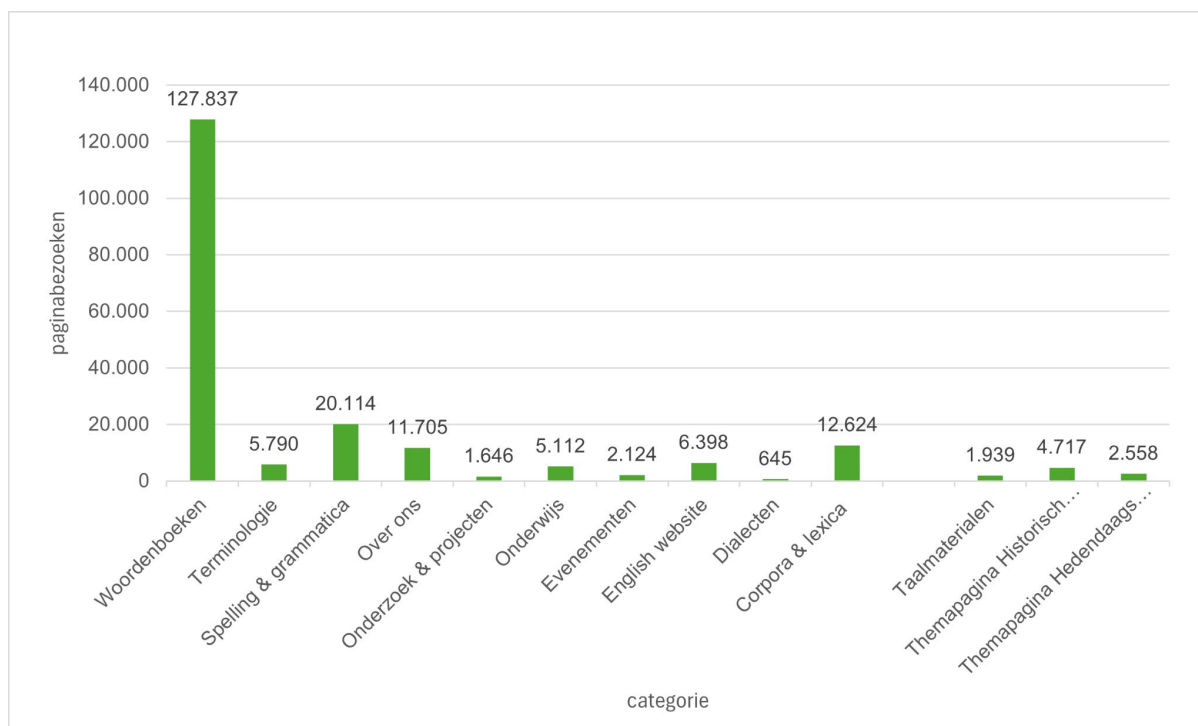
Het grootste deel van de bezoekers (195.000) komt via Google op de website terecht. 86.000 bezoekers tikken ivdnt.org direct in de browser in (tegenover 78.700 in 2024), 67.000 bezoekers (tegenover 57.100 in 2024) komen via diverse nieuwsbrieven (eigen nieuwsbrief 2.800 (t.o.v. 1.900 in 2024), Taalpost, Taalunienieuwsbrief) en zo'n 6.000 via de sociale media. De meesten daarvan vinden de website via LinkedIn, gevolgd door Facebook, Reddit, Instagram, Bluesky en Mastodon. Sinds 2025 is het INT niet meer actief op X (voorheen Twitter). 1.338 bezoekers kwamen op ivdnt.org terecht vanuit ChatGPT.

Van de statische content wordt al jaren achtereen de informatiepagina over het *Woordenboek der Nederlandsche Taal* (WNT) het meest bezocht, met 28.996 paginabezoeken (23.378 unieke bezoekers) in 2025 (de statistieken van de historische woordenboeken zelf zijn te vinden in

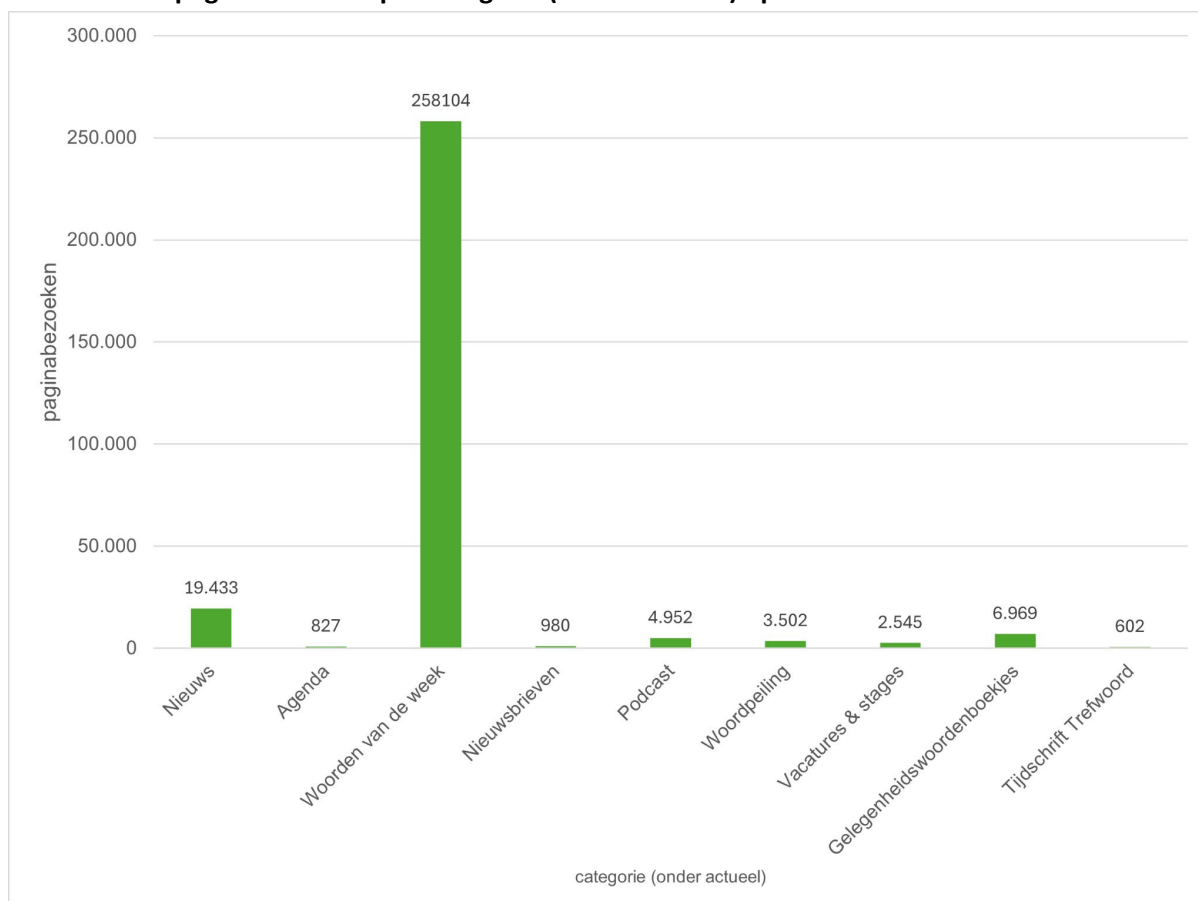
hoofdstuk 3). De woordenboekzoekfunctie is de tweede meest bezochte pagina, met 23.843 paginabezoeken. Er is een nieuwe overzichtspagina voor Dialecten toegevoegd.

Al jaren is 'Nieuw woord van de week' de best bezochte actuele rubriek, met in totaal 136.628 bezoeken (99.915 unieke bezoekers) in 2025. Gemiddeld blijven de bezoekers zo'n 4,5 minuten op de pagina. Ook de (sinds 2020 sporadisch bijgehouden) rubriek 'Woordbaak' wordt nog regelmatig geraadpleegd met in 2025 75.602 bezoeken in totaal (67.238 unieke bezoekers). Bezoekers blijven gemiddeld zo'n 21 minuten op de pagina. In de rubriek 'Grammaticasafari' zijn minder afleveringen dan gebruikelijk verschenen in verband met de voorbereidingen voor het boek *Taaltaart*.

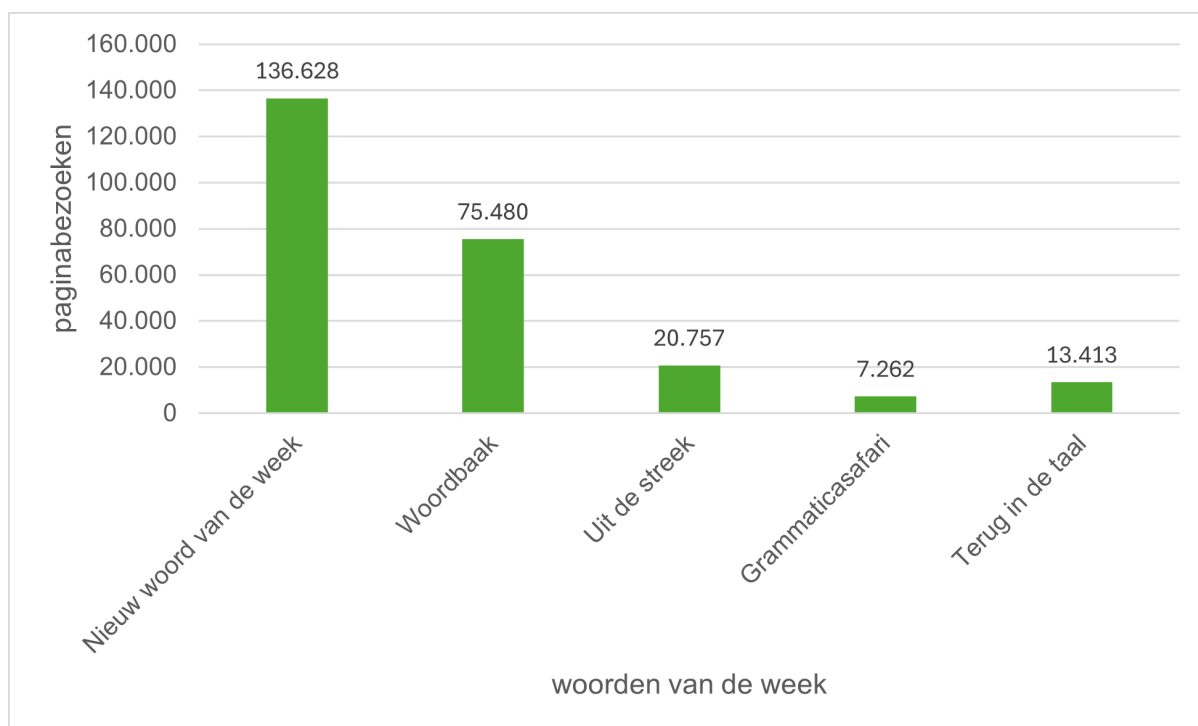
Totaal aantal paginabezoeken per categorie op de website



Totaal aantal paginabezoeken per categorie (onder Actueel) op de website



Totaal aantal paginabezoeken Woorden van de week



43% van de bezoekers raadpleegt ivdnt.org op een smartphone, 53% maakt gebruik van een desktop en 3% van een tablet. 8% van de bezoekers scrolt door tot het einde van de homepage.

De webredactie en de webmaster werken continu aan de verbetering van de website. Zo worden fouten hersteld, kapotte links verwijderd of vervangen, wordt content geüpdatet, de veiligheid verbeterd en wordt op verschillende manieren gewerkt aan het versnellen van de laadtijd en aan zoekmachineoptimalisatie (SEO). Wanneer je via een incognitovenster in Google zoekt op termen als 'nieuwe woorden' (eerste resultaat), 'hedendaags Nederlands' (eerste resultaat), 'historisch Nederlands' (eerste resultaat) of 'terminologie' (eerste resultaat), verschijnt de INT-website steeds hoog in de zoekresultaten.

Op het gebied van veiligheid en gebruikerservaring zijn er bijvoorbeeld foto's toegevoegd van de contactpersonen per project, en is bij terminologie de plug-in van Google Maps vervangen door Open Street Maps. Daarnaast is er gewerkt aan de redactieomgeving van de website, zodat het editen van de voorpagina gemakkelijker is geworden en ook de voorpagina van Woordpeiler eenvoudig geüpdatet kan worden.

Huisstijl

Een team bestaande uit de webmaster, softwareontwikkelaars en de communicatieadviseur heeft de standaardwebapplicatie doorontwikkeld die als vertrekpunt dient voor de ontwikkeling van nieuwe INT-applicaties. De gelijkvormigheid en herkenbare huisstijl zorgen ervoor dat alle door het INT ontwikkelde websites door de gebruikers ook als zodanig herkend worden. In 2025 is er een begin gemaakt aan een Universeel Woordenboek; een webapplicatie die standaard voor de (nieuwe)

woordenboeken gebruikt zal worden. Ook is de website Woordpeiler gelanceerd, waarbij er intensief samengewerkt is om de huisstijl goed naar voren te laten komen.

Nieuwsbrieven & persberichten

In 2025 zijn er zes algemene INT-nieuwsbrieven verschenen en vier nieuwsbrieven over terminologie. Het abonneebestand groeit elk jaar gestaag. De algemene nieuwsbrief telde eind 2025 4.708 abonnees, van wie per editie 35,7% de nieuwsbrieven opent; van de 3.592 geadresseerden van de nieuwsbrief terminologie doet 19,5% dat. In de algemene nieuwsbrieven zijn het vooral de Woorden van de week en de nieuwsberichten waarop doorgeklikt wordt naar ivdnt.org.

Er zijn vier persberichten verstuurd: Woordpeiler (23-01-2025), Sjem Barbiers nieuwe directeur-bestuurder Instituut voor de Nederlandse Taal (31-01-2025), Woordpeiling eerste kwartaal (09-04-2025) en Manosfeer en graaiflatie: verschijn- en verdwijnwoord van 2025 (20-12-2025). Ook zijn er twee mailings verstuurd over de Week van het Nederlands (31,6% opens) en het nieuwe boek *Taaltaart* (32,9% opens).

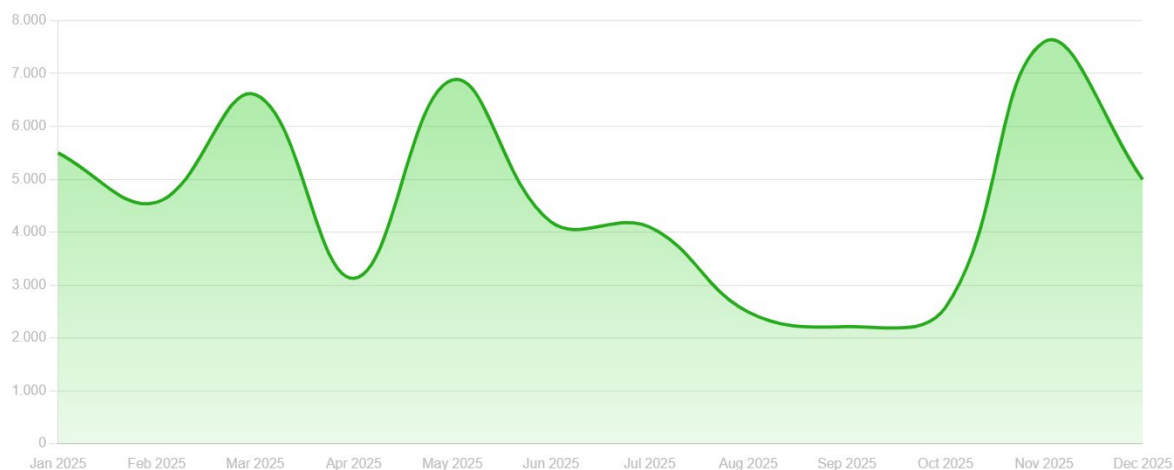
Sociale media

Alle berichten die verschijnen op ivdnt.org worden met behulp van Blog2Social doorgeplaatst op LinkedIn (4.953 volgers), Facebook (1.615 vind-ik-leuks), Mastodon (638 volgers) en Bluesky (538 volgers). Voor Instagram (2.711 volgers) wordt voor elke post een aparte afbeelding gemaakt in de huisstijl. Op LinkedIn zijn het vooral vacatures en nieuwsberichten die veel interactie opleveren. Op Facebook blijft de interactie achter. Op Instagram scoren de 'Woorden van de week' het beste, en dan met name 'Nieuw woord van de week'. Van de verschillende typen posts zijn het de reels die het grootste publiek bereiken. Het gaat dan in de meeste gevallen om fragmenten uit de podcast. LinkedIn en Instagram blijven de twee kanalen waar het meest op ingezet wordt en waar de meeste activiteit plaatsvindt.

Podcasts

Over taal gesproken

In 2025 zijn er negen afleveringen verschenen van de podcast 'Over taal gesproken', die het INT samen met Onze Taal maakt. Laura van Eerten (INT) en Raymond Noë (Onze Taal) spreken elke aflevering een deskundige over het Nederlands. In totaal zijn er nu 41 afleveringen verschenen en de twee best beluisterde in de eerste zeven dagen na publicatie (van 2025) zijn 'Wat taal verradt – met Freek Van de Velde' (2.472 downloads) en 'De taal van depressie – met Aafke Romeijn' (2.285 downloads). Over alle afleveringen verspreid waren er in 2025 in totaal zo'n 55.000 downloads van de podcast. 72% van de luisteraars komt uit Nederland en 12% uit België. 'Over taal gesproken' heeft eind 2025 7.593 abonnees op Spotify en 2.441 abonnees op Apple Podcasts. In 2025 zijn er voor het eerst twee liveopnames gemaakt tijdens het Taalfeest in TivoliVredenburg.



Totaal aantal downloads Over taal gesproken in 2025

Waar komt pindakaas vandaan?

Er is samen met Onze Taal een nieuw seizoen met drie afleveringen gemaakt van de podcast 'Waar komt pindakaas vandaan?' Dit keer in samenwerking met de Week van het Nederlands en ingesproken door de ambassadeurs, gepubliceerd tijdens de Week van het Nederlands. In 2025 zijn er in totaal 38.210 afleveringen gedownload. 79% van de luisteraars komt uit Nederland en 7% uit België. Ook uit deze podcast zijn verschillende fragmenten opgenomen in lesmethodes Nederlands.



Totaal aantal downloads Waar komt pindakaas vandaan in 2025

Andere populairwetenschappelijke activiteiten

De redactie van het *Woordenboek van Nieuwe Woorden* publiceerde enkele artikelen over nieuwe woorden, hield veel lezingen, webinars en interviews (op radio en tv) over dit onderwerp en begeleidde diverse profielwerkstukken over neologismen.

In samenwerking met de Taalunie heeft het INT drie webinars georganiseerd: een over Woordpeiler en twee over dialect. Twee van de drie webinars vonden plaats in de Week van het Nederlands.

Er is twee keer een Taalpubquiz georganiseerd (op Nederlands Centraal en in de Bibliotheek Zaandam), en voor de Taalkrant van de Week van het Nederlands is er een digitale Taalquiz gemaakt.

Aan het einde van het jaar bracht het INT weer een populairwetenschappelijke publicatie uit die als eindejaarsgeschenk verstuurd werd naar INT-relaties. Dit keer was dat het boek *Taaltaart. Het Nederlands in vier smaken*.

Het INT was al verantwoordelijk voor de techniek achter de website Neerlandistiek.nl en zal ook in 2026 de hosting en het onderhoud op zich nemen.

Bijlage 1: Raad van Toezicht en Raad van Advies

Raad van Toezicht

Mr. Mieke Zaanen (voorzitter)

Drs. Erik Boels

Mr. Frank Judo

Raad van Advies

Prof. dr. Hans Bennis (voorzitter)

Willem Bongers-Dek

Prof. dr. Lisa Cheng

Dr. Els Lefever

Dr. Eric Mijts

Dr. Dirk Pijpops

Dr. Freek Van de Velde

Bijlage 2: Medewerkers

Prof. dr. Frieda Steurs – directeur/bestuurder (in dienst tot 30 september 2025)

Prof. dr. Sjef Barbiers – directeur/bestuurder (in dienst sinds 1 september 2025)

Dr. Brenda Assendelft – onderzoeker/taalkundige (in dienst sinds 1 februari 2025)

Dr. Maaïke Beliën – onderzoeker/taalkundige

Dr. Roland de Bonth – onderzoeker/taalkundige

Dalí Chirino – onderzoeker/taalkundige (in dienst sinds 1 oktober 2025)

Lic. Lut Colman – onderzoeker/taalkundige

Lic. Griet Depoorter – onderzoeker/taalkundige

Lic. Katrien Depuydt – onderzoeker/taalkundige

Lic. Veronique De Tier – onderzoeker/taalkundige

Dr. Jesse de Does – computerlinguïst

Job van Doeselaar – webmaster

Laura van Eerten MA – communicatieadviseur

Drs. Mathieu Fannee – systeemontwikkelaar

Drs. Thomas Haga – taalkundig assistent

Lisa van den Hoven - taalkundig assistent (in dienst sinds 1 januari 2025)

Dr. Kris Heylen – onderzoeker/taalkundige

Ruud de Jong – systeemontwikkelaar

Dr. Dirk Kinable – onderzoeker/taalkundige

Uma Kraus – managementassistent

Dr. Martin Kroon – onderzoeker/taalkundige (in dienst sinds 1 oktober 2025)

Drs. Marco van der Laan – systeembeheerder

Dr. Frank Landsbergen – computerlinguïst

Dr. Vince Liégeois – onderzoeker/taalkundige (in dienst sinds 17 februari 2025)

Koen Mertens – systeemontwikkelaar

Drs. Jan Niestadt – systeemontwikkelaar

Vincent Prins, MSc - systeemontwikkelaar

Dr. Kathy Rys – onderzoeker/taalkundige

Prof. dr. Carole Tiberius – computerlinguïst

Dr. Vincent Vandeghinste – computerlinguïst en CLARIN-coördinator

Dr. Bram Vanroy – onderzoeker/computerlinguïst

Lic. Katrien Van pellicom – onderzoeker/taalkundige

Drs. Boukje Verheij – onderzoeker/taalkundige

Rayvano van Vliet – systeembeheerder

Drs. Vivien Waszink – onderzoeker/taalkundige

Karin van Weerlee – managementassistent

Jeroen Willems – taaldata-engineer (in dienst sinds 1 januari 2025)

Laetitia de Winter – administrateur

Bijlage 3: publicaties, lezingen, media, prijzen etc.

Publicaties

Barbiers, S. (2025). Micro-comparative syntax, dialectology, and sociolinguistics. In: S. Barbiers, M. Polinsky, & N. Corver (Eds.), *The Cambridge handbook of comparative syntax* (pp. 203–227). Cambridge University Press.

Barbiers, S., Polinsky, M., & Corver, N. (2025). The comparative syntactic enterprise: An introduction. In: S. Barbiers, M. Polinsky, & N. Corver (Eds.), *The Cambridge handbook of comparative syntax* (pp. 1–19). Cambridge University Press.

Barbiers, S., Polinsky, M., & Corver, N. (Eds.). (2025). *The Cambridge handbook of comparative syntax*. Cambridge University Press.

Barbiers, S., Vanden Wyngaerd, G., & Van der Wal, J. (2025). Theoretical approaches to comparative syntax. In: S. Barbiers, M. Polinsky, & N. Corver (Eds.), *The Cambridge handbook of comparative syntax* (pp. 23–74). Cambridge University Press.

Beliën, M. (2025). Tussenwerpsels in het volle zonlicht. Op: *Neerlandistiek*, 22 maart 2025

Bonth, R. de. (2025). 21 bijdragen aan de webrubriek *Terug in de taal*. <https://ivdnt.org>

Bonth, R. de. (2025). ‘4 T’s en 2 C’s: Taalnorm, taalwerkelijkheid, taalgevoel, taalverandering én het Couranten Corpus’. In: J. van Rijt & M. van Oostendorp (Eds.), *Zodat taal boeit: Inspirerende taalkunde voor in de klas* (pp. 174–181). Pica.

Bonth, R. de. (2025). ‘8 à 9 literatuurgeschiedenissen’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘Alsof het gedrukt staat: Copyrette & printerette’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘Babette moet hangen!’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘Bakkers klaar? Puzzelen maar!’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘Croquetterette’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘De vorstin der dranken: Ronette’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘En we noemen haar *ette’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘Er staat een centaurette op de gang’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘Ettelijke vrouwennamen’. <https://www.neerlandistiek.nl>

Bonth, R. de. (2025). ‘Etymologica: Toetje’. <https://www.neerlandistiek.nl>

- Bonth, R. de. (2025). 'Godmiljaar, Ben Salemans is er niet meer!'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Groslijst: Peter Drehmanns, *Sidderingen. De glorie van het verval en het verval van de glorie*'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Mijn papier is op': Brieven van Antonie Cornelis Oudemans aan Matthias de Vries. In: 2025.
- Bonth, R. de. (2025). 'Onder de wol met Betje en Aagje'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Rekenen in, op en met de computerette'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Schoon met stoom: De stomerette'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Speciale speciaalzaken'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Supplement op het Middelnederlandsch handwoordenboek'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Teweerstellen met tankette'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Vacuette: Voor stof- en bloedzuigers'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Wat stond daar? Het woordenboek als leeshulp'. <https://www.neerlandistiek.nl>
- Bonth, R. de. (2025). 'Woorden van (1) de (2) jaar'. <https://www.neerlandistiek.nl>
- Bonth, R. de., De Tier, V., Landsbergen, F., Verheij, B., & Waszink, V. (2025). *Taaltaart: Het Nederlands in vier smaken*. Scriptum.
- Bouillon, P., et al. (2025). *Proceedings of Machine Translation Summit XX: Volume 2*. European Association for Machine Translation.
- De Tier, V. (2025). Dialect wordt een hip museumstuk. In: *Jouw dialect leeft nog even. De Lage Landen*, mei 2025, pp. 15–23
- De Tier, V. (2025). Onder de Zeeuwse oogleden: dialect over strontjes, slapertjes en meer. In: *Nehalennia* aflevering 227, najaar 2025, 19–21.
- De Tier, V. (2025). 'Begrijpt de computer binnenkort ook Zeeuws-Vlaamse dialecten?'. In: *Zeeuws Erfgoed*, jg. 24, 1, pp. 30–31.
- De Tier, V. (2025). 'Berewolven en biete bouwen: namen voor de kinderschrik in West-Vlaanderen, In: *Dielektgazette Bachtn de Kuupe*, jg. 20, nr. 3, pp. 7–9.
- De Tier, V. (2025). 'De impact van (streek)geschiedenis op de Zuid-Bevelandse dialecten'. In: *Alle geschiedenis is lokale geschiedenis. Vijftig jaar Heemkundige Kring De Bevelanden*. 1975–2025, pp. 33–45.

De Tier, V. (2025). 'En de boer ... hij ploegde voort. Landbouwwoordenschat in de Zeeuwse dialecten'. In: *Zeeuws Erfgoed*, jg. 24, 3, pp. 28–29.

De Tier, V. (2025). 'Hennestrontjes schieten: knikkeren in West-Vlaanderen', In: *Diejalektgazette Bachtn de Kuupe*, jg. 20, nr. 3, pp. 7–9.

De Tier, V. (2025). 'Katje mie, katje were. Over West-Vlaamse namen voor kinderspelletjes'. In: *Diejalektgazette Bachtn de Kuupe*, jg. 20, nr. 1, pp. 7–9.

De Tier, V. (2025). 'Van West-Vlaamse troepen, bendes en kooien. Hoe we hier en elders spreken over groepen vogels, dieren en ... druiven', In: *Diejalektgazette Bachtn de Kuupe*, jg. 20, nr. 4, pp. 7–9

De Tier, V. (2025). 'West-Vlaamse stemmen uit het verleden, nu met ondertitels!', In: *Diejalektgazette Bachtn de Kuupe*, jg. 20, nr. 2, pp. 7–9.

De Tier, V. (2025). 'Zeeuwse stemmen uit het verleden met ondertitels. In: *Nehalennia*, 225, voorjaar 2025, pp. 16–17

De Tier, V. & Oostendorp, M. van (2025). 'Kindergriezels'. In: *Onze Taal*, nr. 6, pp. 16–17.

De Tier, V. & Oostendorp, M. van (2025). 'Op kraamvisite'. In: *Onze Taal*, nr. 4, pp. 16–17.

De Tier, V. & Oostendorp, M. van, (2025). 'De grote voorjaars schoonmaak. In: *Onze Taal*, nr. 2, pp. 16–17

Doumen, J., Prins, V., & Vandeghinste, V. (2025). Textlens: An online dashboard for digital text analysis. In: *Digital Humanities Benelux 2025 Conference*. <https://doi.org/10.5281/zenodo.15425643>

Eerten, L. van (2025). Bijdragen Taalkalender over dierenwoorden, *Onze Taal*.

Fannee, M. (2025a). *Atlas van laatmiddeleeuws Warmond: Reconstructie van het dorp en zijn omgeving (1250–1574)* (4e herz. ed.). Historisch Genootschap Warmelda.

Fannee, M. (2025b). De (toch niet helemaal) verloren gegane middeleeuwse archieven van de kerk van Warmond. In: *De Hekkensluiter*, 22(2), pp. 8–25.

Fannee, M. (2025c). *Warmonds toponymisch woordenboek: Middeleeuwen & vroegmoderne tijd* (4e herz. ed.). Historisch Genootschap Warmelda.

Federico M. et al. (onder wie Tiberius, C.) (2025). DiBiMT: A Gold Evaluation Benchmark for Studying Lexical Ambiguity in Machine Translation. In: *Computational Linguistics 2025*; 51 (2): pp. 343–413.

Heylen, K., Prins, V., Depuydt, K., De Does, J., Van Eerten, L., & Haga, T. (2025). Woordpeiler: A new tool for visualizing and analyzing lexical trends in contemporary Dutch. In: *eLex 2025 book of abstracts* (p. 83).

Joby, C., & Van der Sijs, N. (2025). *Dutch and contact linguistics*. John Benjamins.

Kallas, J., Koppel, K., Heylen, K., Kernerman, I., Ostroški Anić, A., Vezzani, F., & Arhar Holdt, Š. (2025). Neology in practice: Lexicographic and terminological approaches to lexical innovation. In: I. Kosem

et al. (Eds.), *Electronic lexicography in the 21st century (eLex 2025): Intelligent lexicography*, pp. 652–672.

Karakanta, A., Tiberius, C., Wijnholds, G., Prokić, J., Westera, M., & Verberne, S. (Eds.). (2025). *Computational linguistics in the Netherlands journal* (Vol. 14).

Klosa-Kückelhaus, A., & Tiberius, C. (2025). The lexicographic process revisited. In: *International Journal of Lexicography*, 38(1), pp. 1–12.

Landsbergen, F. (2025). De ANS in de klas. In: *Handboek didactiek Nederlands*. Levende Talen.

Liégeois, V. (2025a). Die deutschsprachige Gemeinschaft und die (Kon)föderalisierung Belgiens: Die Abschaffung des Senats als kritische Fallstudie. In: *Europäisches Journal für Minderheitenfragen*, 13(3-4), pp. 415–441. <https://doi.org/10.35998/ejm-2025-0020>

Liégeois, V. (2025b). Naar een kortere zinsbouw in wetgeving: Veel voordelen, vier criteria en één hindernis. In: *Tijdschrift voor Wetgeving*, 26(2), pp. 168–176.

Liégeois, V., & Bernaerts, J. (2025). Die deutschsprachige Rechtsterminologie Belgiens: Eine metalinguistische Studie zu Übersetzungsfragen des Terminologieausschusses. *Zeitschrift für Angewandte Linguistik*, 83, 272–302. <https://doi.org/10.1515/zfal-2025-2014>

Liégeois, V., & Gautier, L. (2025a). Domain-sensitive weather phraseology from a constructionist perspective: A case study based on Dutch dialects in Flanders. *Dutch Journal of Applied Linguistics*, 14.

Liégeois, V., & Gautier, L. (2025b). Existential structures as specialised constructions: Evidence from a corpus of Flemish weather reports. *Études de linguistique et d'analyse des discours*, 11. <https://doi.org/10.35562/elad-silda.1673>

Martelli, F., Perrella, S., Campolungo, N., Munda, T., Koeva, S., Tiberius, C., & Navigli, R. (2025). DiBiMT: A gold evaluation benchmark for studying lexical ambiguity in machine translation. *Computational Linguistics*, 51(2), pp. 343–413. https://doi.org/10.1162/coli_a_00541

Morozova, I., & Barbiers, S. (2025). On the category and morphosyntax of numerals. *Linguistic Variation*. [Advanced online publication](#).

Oostendorp, M. van & De Tier, V. (2025). ‘Chris en Robin kussen elkaar’. In: *Onze Taal*, nr. 1, pp 16–17.

Oostendorp, M. van & De Tier, V. (2025). ‘Hoe meer kinders, hoe meer vreugd’. In: *Onze Taal*, nr. 5, pp 16–17.

Oostendorp, M. van & De Tier, V. (2025). ‘werkwoorden ... op het einde. In: *Onze Taal*, nr. 3, pp 16–17.

Resi, R. & Steurs, F. (Eds), (2025). *Handbook of terminology* (Vol. 4, 799 p.).

Resi, R. & Steurs, F., (2025). Concluding Remarks. Dimensions of terminology across Europe. In: R. Resi & F. Steurs (Eds.), *Handbook of terminology* (Vol. 4, pp. 763–777).

Shterionov, D., et al. (2025). *Proceedings of the third international workshop on automatic translation for signed and spoken languages (AT4SSL)*. European Association for Machine Translation.

Sijs, N. van der (2025). *Dijt onze woordenschat alsmaar uit? Over opbouw en geschiedenis van de Nederlandse woordenschat*. Radboud University Press Nijmegen

Sijs, N. van der (2025). 'Als een uitdijend heelal? Over de Nederlandse woordenschat, die alsmaar lijkt te groeien', in: *Onze Taal* 3, pp. 58–61

Sijs, N. van der (2025). 'Chaper 17. "O this is Eden!". Dutch overseas animal names', in: *Dutch and Contact Linguistics: The Dutch language outside the Low Countries*, ed. by Christopher Joby and Nicoline van der Sijs, John Benjamins, pp. 537–558.

Sijs, N. van der (2025). 'Chapter 1. The role of Dutch in the circulation of loanwords', in: *Dutch and Contact Linguistics: The Dutch language outside the Low Countries*, ed. by Christopher Joby and Nicoline van der Sijs, John Benjamins, pp. 12–45.

Sijs, N. van der (2025). 'De lotgevallen van de Burghers op Sri Lanka', in: *Internationale Neerlandistiek*, themanummer Emigratie uit de Lage Landen, 63, 3, pp. 230–251.

Sijs, N. van der (2025). 'De middelste zin van... Twee minuten stilte'. Op: *Neerlandistiek*

Sijs, N. van der (2025). 'De rol van de mens bij kunstmatig-intelligent machinelezen'. Op: *Neerlandistiek*.

Sijs, N. van der (2025). 'Dutch in contact with other languages', in: Gijsbert Rutten & Pierre Swiggers (eds), *The Dutch Language, 1500-1800: New Perspectives*.

Sijs, N. van der (2025). 'Een verrassende taalpeiling bij Nederlanders, Vlamingen en Walen', in *Neerlandia* 4, 11.

Sijs, N. van der (2025). 'Etymologica: (De) Oekraïne'. Op: *Neerlandistiek*.

Sijs, N. van der (2025). 'Etymologica: cyperse kat'. Op: *Neerlandistiek*.

Sijs, N. van der (2025). 'Etymologica: eskimo-ijsjes'. Op: *Neerlandistiek*.

Sijs, N. van der (2025). 'Etymologica: huisvlijt'. Op: *Neerlandistiek*

Sijs, N. van der (2025). 'Etymologica: lapjeskat en schildpadkat'. Op: *Neerlandistiek*.

Sijs, N. van der (2025). 'Etymologica: sibbediksie'. Op: *Neerlandistiek*

Sijs, N. van der (2025). 'Etymologica: sterkhouder'. Op: *Neerlandistiek*

Sijs, N. van der (2025). 'Etymologica: vingerbob'. Op: *Neerlandistiek*

Sijs, N. van der (2025). 'Etymologica: zekerfeest'. Op: *Neerlandistiek*

Sijs, N. van der (2025). 'Geef mij maar Nederlands (of Frans)!', in: *De Lage Landen*

- Sijs, N. van der (2025). 'Had de Statenvertaling invloed op de taal van de 17e-eeuwse kranten?'. Op: *Neerlandistiek*
- Sijs, N. van der (2025). 'Het taalregister van de oudste Nederlandstalige kranten'. Op: *Neerlandistiek*.
- Sijs, N. van der (2025). 'Hoe goed ken jij je Jiddisch, goochemerd?', in: *750xAmsterdam*.
- Sijs, N. van der (2025). 'Kaiserbroodje', in: *Onze Taal* 2, 34.
- Sijs, N. van der (2025). 'Killivisjes', in: *Onze Taal* 4, 36.
- Sijs, N. van der (2025). 'Leve de historische taalkunde!'. Op: *Neerlandistiek*
- Sijs, N. van der (2025). 'Met het oog op morgen'. Op: *Neerlandistiek*.
- Sijs, N. van der (2025). 'Mugshot', in: *Onze Taal* 6, 54.
- Sijs, N. van der (2025). 'Stemvee', in: *Onze Taal* 3, 64.
- Sijs, N. van der (2025). 'Suppie, verschuiven en deepy: wie kent ze nog?', in *Neerlandia* 3, pp. 30–33.
- Sijs, N. van der (2025). 'Van engelenbak naar skybox', in: *Onze Taal* 5, p. 36.
- Sijs, N. van der (2025). 'Verrijkend taalonderwijs dankzij etymologische kennis', in: *Zodat taal boeit, Inspirerende taalkunde voor in de klas*, red. Jimmy van Rijt en Marc van Oostendorp, Uitgeverij Pica, pp. 157–165.
- Sijs, N. van der (2025). 'Verschenen: Onderzoeksrapport "Taal in de Lage Landen"'. Op: *Neerlandistiek*
- Sijs, N. van der (2025). 'Woordenboekgebruik'. Op: *Neerlandistiek*
- Sijs, N. van der (2025). 'Woordsalade', in: *Onze Taal* 1, 40.
- Sijs, N. van der & Coornstra, L. (2025). 'Bedankt voor het lenen!', in: *Onze Taal* 2, pp. 32–33
- Sijs, N. van der et al. (2025). *Onderzoeksrapport Talen in de Lage Landen. Taalhouding, taalkennis en taalgebruik van studenten en docenten in het Nederlandse en Belgische hoger onderwijs*. INT en Fryske Akademy, 120 pagina's.
- Sijs, N. van der et al. (2025). 'A cross-linguistic study on pseudo-Anglicisms in six languages', in: *Anglicisms Around the Globe: Cross-Linguistic Studies on the Impact of English*, ed. Henrik Gottlieb et al., Routledge, 235–266.
- Snel, B., Van der Meulen, M., & Landsbergen, F. (2025). Grammaticales van de grootste expert: De ANS. *Levende Talen Magazine*, 2025(1), 16–19.
- Soete, F., & Vandeghinste, V. (2025). Optimising controllable sentence simplification for Dutch. *Computational Linguistics in the Netherlands Journal*, 14, pp. 21–41.

Steurs, F. (2025). Meeting the societal and technological challenges to linguistic diversity: The case of Dutch, Frisian, Papiamentu and Afrikaans. In: S. Kirchmeier & T. Váradi (Eds.), *Linguistic diversity and language planning*. ELTE Hungarian Research Centre for Linguistics.

Steurs, F., & Kinable, D. (2025). Efforts and challenges in translating concept to reality. In: R. Resi & F. Steurs (Eds.), *Handbook of terminology* (Vol. 4, pp. 377–391).

Steurs, F., Kockaert, H., & Alswlaiman, A. A. (Eds.). (2025). *Al-Ġāmi‘ fī ‘ilm al-muṣṭalaḥ = Handbook of terminology*. King Salman Global Academy for Arabic Language.

Steurs, F., Kinable, D., & Heylen, K. (2025). Het INT: Schatkamer van de Nederlandse taal – het ENT, het expertisecentrum Nederlandstalige terminologie. In: *TiNT 2024* (pp. 66–71). NL-Term.

Tallarico, G., Costa, R., Salgado, A., Freixa, J., Heylen, K., Humbley, J., Kernerman, I., Klosa-Kückelhaus, A., Ostroški Anić, A., & Vezzani, F. (2025). Lexicography and neology: Building bridges through a multilingual approach with ENEOLI. *Lexikos*, 35, pp. 294–317.

Vandeghinste, V., & Kontino, T. (Eds.). (2025). *Selected papers from the CLARIN Annual Conference 2024*. Linköping University Electronic Press.

Vandeghinste, V., Vanroy, B., & Van Doeselaar, J. (2025). Human evaluation of automated text simplification through crowdsourcing. In: *Proceedings of the CLARIN Annual Conference*, pp. 143–147.

Vandeghinste, V., Daelemans, W., De Bruyne, L., Van de Cruys, T., & Lefever, E. (2025). CLARIN in CLARIAH-VL+: Plans for 2025–2028. In: *Proceedings of the CLARIN Annual Conference*, pp. 118–122.

Vonk, F., De Bonth, R., & Rutten, G. (2025). Els Ruijsendaal (1942–2025) en het talenonderwijs: Een in memoriam. <https://www.neerlandistiek.nl>

Waszink, V. (2025a). ‘Doe noncha, bruh. Waarom spreken jongeren jongerentaal?’. In: *Neerlandia*, 2025(4).

Waszink, V. (2025b). ‘Fawaka? Bun! Jongerentaal speelt leentjebuurt bij Sranantongo’. Op: *De Lage Landen*.

Waszink, V., & Van Eerten, L. (2025). Frisse taalquiz. Op: *Taalkrant*.

Lezingen en presentaties

Barbiers, S. Wat is het verschil tussen een groot formaat geleerde en een geleerde van groot formaat? 5de LUCL Retired & Kicking Symposium, Universiteit Leiden, 24 oktober 2025.

Beliën, M. Op grammaticasafari! Een reis doorheen het Nederlands. Nacht van het Nederlands, Gent, 3 oktober 2025.

Beliën, M. Paal en perk aan partikels? Een perspectief vanuit de ANS. Dag van de Nederlandse partikels, Amsterdam, 12 september 2025.

Bonth, R. de. Presentatie Couranten Corpus voor vrijwilligers krantenproject. Leiden, Instituut voor de Nederlandse Taal, 2 mei 2025.

Bonth, R. de. Presentatie over INT voor vrijwilligers die preken HaverSchmidt digitaliseren. Leiden, Instituut voor de Nederlandse Taal, 29 september 2025.

Bonth, R. de. Presentatie over Woordenboek der Nederlandsche Taal voor ITV Hogeschool. Leiden, 8 april 2025.

Colman, L., & Landsbergen, F. Grammatica – Woordcombinaties en de ANS voor vertalers. Presentatie voor de ITV Hogeschool voor Tolken en Vertalen, Leiden, 8 april 2025.

De Tier, V. De impact van streekgeschiedenis op de Zuid-Bevelandse dialecten. Goes, 23 mei 2025.

De Tier, V. De Oost-Vlaamse dialecten in de DSDD. Heusden, 6 november 2025.

De Tier, V. De West-Zeeuws-Vlaamse dialecten. Lodewijk College, Terneuzen, 1 december 2025.

De Tier, V. De Zeeuwse dialecten. De Vroone, Kapelle, 9 december 2025.

De Tier, V. De Zeeuwse dialecten: geschiedenis en kenmerken. Goes, 10 april 2025.

De Tier, V. Dialecten uit Wetteren en omgeving. Wetteren, 27 februari 2025.

De Tier, V. Nederlandse dialecten. Zomercursus Taalunie, 11 augustus 2025.

De Tier, V. West-Vlaamse dialecten in de DSDD. Roeselare, 15 januari 2025.

De Tier, V. Zeeuwse kaartjes in de DSDD. Biervliet, 15 mei 2025.

De Tier, V., Depuydt, K., & Fannee, M. Een infrastructuur voor het Bildts: De ontwikkeling van een digitale dialectinfrastructuur voor dialectliefhebbers. Posterpresentatie, Sociolinguistics Circle, Leiden.

Depuydt, K., & De Does, J. An infrastructure for historical Dutch corpus development. CLARIN Conference 2025, Wenen, 2 oktober 2026.

Doumen, J., Prins, V., & Vandeghinste, V. Textlens: An online dashboard for digital text analysis. Digital Humanities Benelux 2025 Conference, Amsterdam.

Eerten, L. van, samen met V. Waszink. Frisse Taalquiz. Bibliotheek Zaandam, 6 maart 2025.

Eerten, L. van, samen met V. Waszink. Frisse Taalquiz. Nederlands Centraal, Taalunie, 26 mei 2025

Eerten, L. van. Waar komt pindakaas vandaan? Presentatie tijdens de Nacht van het Nederlands, Gent, 3 oktober 2025.

Heylen, K., De Jong, R., De Does, J., & Fannee, M. TermWerk: An integrated web application for Dutch terminology extraction and management. Posterpresentatie, CLIN35-congres, Leuven, 12 september 2025.

Heylen, K., Depuydt, K., De Does, J., Prins, V., Van Eerten, L., & Haga, T. Woordpeiler: A new tool for visualizing and analyzing lexical trends in contemporary Dutch. Presentatie, eLex 2025, Bled, 20 november 2025.

Heylen, K., Depuydt, K., De Does, J., Prins, V., Van Eerten, L., & Haga, T. Woordpeiler: Visualizing and analyzing lexical trends in contemporary Dutch. Posterpresentatie, CLIN35-congres, Leuven, 12 september 2025.

Kallas, J., Koppel, K., & Heylen, K. Neology in practice: Lexicographic and terminological approaches to lexical innovation. Presentatie, eLex 2025, Bled, 18 november 2025.

Kinable, D. ENT: Jaarlijkse rapportering – Beschouwingen vanuit een ISO-standaard. Presentatie, TiNT-dag (NL-Term), Antwerpen, 14 november 2025.

Kinable, D. Focus on ISO standard 29383: Dissemination by the Centre of Expertise for Dutch Terminology. Panellezing, Terminology Planning in Europe (European Association for Terminology), Innsbruck, 27–29 november 2025.

Kinable, D. Word, Excel en TermWerk voor extractie en beheer van Nederlandstalige terminologie. Presentatie, Termraad Nederlands, Luxemburg, 26 september 2025.

Landsbergen, F. Onderwijsmaterialen van het Instituut voor de Nederlandse Taal. Posterpresentatie, Neerlandistiekdagen, Nijmegen, 4 april 2025.

Landsbergen, F. Standaardtaal vs. variatie in de Algemene Nederlandse Spraakkunst (ANS). Presentatie, 22e Colloquium Neerlandicum (IVN), Brussel, 27 augustus 2025.

Liégeois, V. A comparative law terminological database for Dutch legal language: A functional introduction based on employment contract law. Presentatie, YLRC25 – Young Legal Researchers Conference, Hasselt, 5 december 2025.

Liégeois, V. Conceptuele rechtsvergelijking: Open normen in het Nederlandse en Belgische aansprakelijkheidsrecht. Lezing, Colloquium van het Instituut voor Filosofische en Sociaalwetenschappelijke Educatie vzw, Gent, 25 oktober 2025.

Steurs, F. Openingslezing op de Workshop The language data space workshop, Breda, 23 januari 2025.

Steurs, F. Lezing op de Unesco Language Technology days, Parijs, 24-26 februari 2025.

Steurs, F. Lezing voor de CLARIN strategy days: The socio-economic impact of language technology. Invited lecture. Leuven, 5 februari 2025.

Steurs, F. Lezing Prince-academie (online) Taal in transformatie. Technologie, economie en de kracht van generatieve AI. 25 maart 2025.

Steurs, F. Lezing De schatkamer van het Nederlands bij de ontvangst Secretarissen KNAW (Mieke Zaanen), 1 april 2025.

Steurs, F. Lezing De schatkamer van het Nederlands bij de ontvangst van de studenten Vakgroep Taalkunde UGent, 4 april 2025.

Steurs, F. Lezing De schatkamer van het Nederlands bij de ontvangst van de studenten ITV, 8 april 2025.

Steurs, F. lezing afdeling OvdP Nijmegen over EFNIL: Het meertalige Europa. Kansen en uitdagingen in tijden van AI, 24 april 2025.

Steurs, F. Nederlands Centraal: deelname aan paneldiscussie (moderator: Frits Spitz), 26 mei 2025.

Steurs, F. Colloquium Neerlandicum: panel gespecialiseerd vertalen. Openingslezing : Vertalen in tijden van transformatie, 27 augustus 2025.

Steurs, F. KANTL: ondertekening samenwerkingsakkoord INT-KANTL, Gent, 28 augustus 2025

Steurs, F. Terminology work: new methods in times of AI, Workshop Pretoria, 1-2 oktober 2025.

Steurs, F. Intelligent Language Technology: AI and the Impact on Languages and Language Policy, Workshop Pretoria, 1-2 oktober 2025.

Steurs, F. Towards a Comprehensive Digital Language Infrastructure with AI, Workshop Pretoria, 1-2 oktober 2025.

Steurs, F. lezing en debat: Next gen terminology work with AI (Workshop), Stellenbosch, 6-7 oktober 2025.

Steurs, F. Lezing en debatleider over Terminology Teaching. EAFT workshop. Innsbruck, 27-28 november 2025.

Steurs, F. Lezing voor de leerstoel Nederlands in Essen, Duisburg, 4 december 2025.

Tiberius, C., Colman, L., & Janssen, M. Constructing a PARSEME-NL corpus: Part I. Posterpresentatie, UniDive 3rd General Meeting, Boedapest, 29–30 januari 2025.

Tiberius, C., & De Does, J. LLMs and lexicography at the Dutch Language Institute. Keynote, eLex 2025 – Intelligent Lexicography, Bled, Slovenië, 19 november 2025.

Vandeghinste, V., Daelemans, W., De Bruyne, L., Van de Cruys, T., & Lefever, E. CLARIN in CLARIAH-VL+: Plans for 2025–2028. Poster, CLARIN Annual Conference, Vienna.

Vandeghinste, V., Vanroy, B., & Van Doeselaar, J. Human evaluation of automated text simplification through crowdsourcing. Poster, Brussels Plain Language Experience 2025.

Vandeghinste, V., Vanroy, B., & Van Doeselaar, J. Human evaluation of automated text simplification through crowdsourcing. Poster, CLARIN Annual Conference, Vienna.

Waszink, V. Gesproken column over inclusieve taal. Nacht van het Nederlands, Gent, 3 oktober 2025.

Waszink, V. 'Twijfelmoed, brotox en vertiktokken: Nieuwe woorden en het WNW'. Lezing, Hogeschool Tolken en Vertalen ITV, 8 april 2025.

Actieve deelname en organisatie van congressen en workshops

Roland de Bonth:

- Workshop over Couranten Corpus tijdens Neerlandistiekdagen (3 en 4 april 2025, Nijmegen, Radboud Universiteit)
- Webinar *Werken met online historische woordenboeken*, voor Histories – organisatie voor erfgoedvrijwilligers in Vlaanderen en Brussel (7 april 2025, online)
- Workshop over GaLAHaD & LAnCeLot (9 april 2025, Leiden, Universiteit Leiden)

Katrien Depuydt:

- Workshop over GaLAHaD & LAnCeLot (9 april 2025, Leiden, Universiteit Leiden)

Jesse de Does:

- Workshop over GaLAHaD & LAnCeLot (9 april 2025, Leiden, Universiteit Leiden)
- LLM workshop, Amsterdam, VU, 27 November 2025

Kris Heylen:

- Tallarico, Giovanni, Rute Costa, Ana Salgado, Ana Ostroški Anić, , Federica Vezzani, John Humbley, Judit Freixa, Kris Heylen, Petar Božović, Špela Arhar Holdt, Vincent Balnat & Weronika Woźniak. Organisatie van de Second General Meeting of the European Network On Lexical Innovation (ENEOLI). Budapest, 10-11 april 2025.
- Heylen, Kris en Veronique De Tier. Organisatie van het symposium "Een Lans voor het Nederlands" ter gelegenheid van het afscheid van Frieda Steurs als directeur van het Instituut voor de Nederlandse Taal. Leiden, 16 september 2025.
- Heylen, Kris. Deelname als vertegenwoordiger van het INT aan de Internationale Gespreksdag van de Marnixkring, Gent, 15 november 2025.
- Kernerman, Ilan en Kris Heylen. Organisatie van de 7th Globalex Workshop on Lexicography and Neology (GWLN-8). Bled, 20 november 2025

Vincent Prins:

- Workshop over GaLAHaD & LAnCeLot (9 april 2025, Leiden, Universiteit Leiden)
- LLM workshop, Amsterdam, VU, 27 November 2025

Frieda Steurs:

- samen met Carole Tiberius & Veronique Hoste. Organisatie LDS Country Workshop voor België en Nederland, 23 januari 2025. Breda.
- Workshop WoLaLa: large language models. Lid van het wetenschappelijk comité. Georganiseerd door het centrum voor de Hongaarse taal (ELTE Research Centre for Linguistics), Budapest 20-21 november 2025
- EAFT workshop in Innsbruck: organisatie van het congres, openingslezing, debatleider over Terminology Teaching, 27-28 november 2025

Carole Tiberius:

- samen met Veronique Hoste & Frieda Steurs. Organisatie LDS Country Workshop voor België en Nederland, 23 januari 2025. Breda.
- Deelname aan UniDive WG2 meeting in Budapest, 28 januari 2025.
- Deelname aan UniDive 3rd general meeting in Budapest, 29-30 januari 2025.
- Deelname aan ALT-EDIC Meeting NL Partners in Den Haag als vertegenwoordiger van CLARIN, 12 maart 2025.
- Deelname aan AI-Verordening in de praktijk georganiseerd door Teekens Karstens in Leiden, 17 april 2025.
- Deelname aan CLARIN conferentie in Wenen, 27-30 september 2025.
- Deelname aan [first CLARIAH Community Connect 2025](#) in Utrecht, 3 december 2025.

Veronique De Tier:

- Sociolinguistics Circle, Leiden, 4 april
- Streektaalconferentie Streektaal op het Podium. Stichting Nederlandse Dialecten. Emmen, 2 oktober 2025

Vincent Vandeghinste:

- Organisatiecomité Computational Linguistics in the Netherlands 2025 te Leuven.
- LREC 2026 Senior Area Chair
- LREC 2026 Workshop ParlaClarin V

Bram Vanroy:

- Workshop over GaLAHaD & LANCeLot, 9 april 2025, Leiden, Universiteit Leiden
- LLM workshop, Amsterdam, VU, 27 November 2025

Onderwijs

Heylen, Kris en Uri Horesh. Gastseminarie Neology in Cultural and Social Contexts. 1st ENEOLI Training School on Lexical Innovation, Faculdade de Ciências Sociais e Humanas da Universidade NOVA de Lisboa, 12 Juni 2025.

Heylen, Kris en Vincent Vandeghinste. Collegereeks Computatieve Lexicologie (B-KUL-F0VG6A) binnen de Master Taalkunde aan de KU Leuven: september-december 2025.

Steurs, F., Bedrijven en de ROI van terminologiewerk. Gastcollege KU Leuven, master in het vertalen, 6 maart 2025

Tiberius, Carole. MA-collegereeks *Corpus Lexicography*. Universiteit Leiden, februari-mei 2025.

Tiberius, Carole. BA-college *Taal Mentaal 2: Semantiek van woorden en zinnen*. Universiteit Leiden, 20 oktober 2025.

Vandeghinste, Vincent. 'Text-to-pictograph conversion and vice versa.' Language Engineering Applications. Master of Artificial Intelligence. KU Leuven. 3 maart 2025.

Vandeghinste, Vincent. 'Language Technology Infrastructures.' Language Engineering Applications. Master of Artificial Intelligence. KU Leuven. 12 mei 2025.

Waszink, Vivien. Module inclusieve taal voor de Hogeschool Tolken en Vertalen ITV

In de media

Barbiers, Sjef. Interview over nieuwe functie als directeur bij het INT. *De Taalstaat* NPO Radio 1, 8 februari 2025

De Tier, Veronique. De boesjkammeree. op: Nostalgie Plus, 29 september 2025

De Tier, Veronique. Maalplak. op: Nostalgie Plus, 6 oktober 2025

De Tier, Veronique. Klakkaards. op: Nostalgie Plus, 13 oktober 2025

De Tier, Veronique. Puppegolle. op: Nostalgie Plus, 15 december 2025

Eerten, Laura van. Over wat een boer (oprisping) te maken heeft met een boer (beroep). In de podcast *Alledaagse Vragen* van BNNVARA, 19 februari 2025

Eerten, Laura van, samen met Vivien Waszink. Online Taalquiz in de Taalkrant van de Week van het Nederlands, oktober 2025

Eerten, Laura van. Twee live opnames van de podcast *Over taal gesproken*. Taalfeest, 11 oktober 2025

Eerten, Laura van. Bekendmaking verschijn- en verdwijnwoord. In *De Taalstaat* NPO Radio 1, 20 december 2025

Van pellicom, Katrien. Interview over het allerlaatste Groene Boekje op papier. In: *Onze Taal 5 - 2025*, pp. 26-27.

Verheij, Boukje. Interview over het woord *eftelingsk* in [Woordenlijst.org](https://woordenlijst.org). In *De Nacht Van* op NPO Radio 1, 14 april 2025

Verheij, Boukje. Over of er woorden uit het woordenboek verdwijnen. In de podcast *Alledaagse Vragen* van BNNVARA, 28 april 2025

Waszink, Vivien. [Interview](#) over nieuwe woorden. *NRC*, 1 januari 2025

Waszink, Vivien. Interview over inclusieve taal. *Daphne op Donderdag*, podcast van Daphne Deckers, 30 januari 2025

Waszink, Vivien. [Interview](#) over jongerentaal. *Vrij Nederland*, 26 januari 2025

Waszink, Vivien met Laura van Eerten. Frisse Taalquiz bibliotheek Zaandam, 6 maart 2025

Waszink, Vivien. [Interview](#) over jongerentaal. Podcast *Dat smaakt naar Meertens*, 6 maart 2025

Waszink, Vivien. [Interview](#) over nieuwe woorden. *Quest*, 9 maart 2025

Waszink, Vivien. [Interview](#) over nieuwe woorden. Grote Taalshow Studium Generale Utrecht, 12 maart 2025

Waszink, Vivien. [Interview](#) over de taal van De Jeugd van Tegenwoordig. *De Lage Landen*, 14 maart 2025

Waszink, Vivien. [Interview](#) over het woord 'gezellig'. Podcast *Alledaagse Vragen (BNNVARA)* op het NPO Luister Podcastfestival, 23 maart 2025

Waszink, Vivien. Interview over jongerentaal. Podcast *Spreek de Streek*, 27 maart 2025

Waszink, Vivien. [Interview](#) over datingtaal. *De Telegraaf*, 10 april 2025

Waszink, Vivien. [Interview](#) over 'moedertaal' en 'vaderland'. Podcast *Alledaagse Vragen (BNNVARA)*, 20 mei 2025

Waszink, Vivien met Laura van Eerten, taalquiz. Nederlands Centraal, Taalunie, 26 mei 2025

Waszink, Vivien. Interview over inclusieve taal. *Plusmagazine*, juni 2025.

Waszink, Vivien. [Interview](#) over 'stief-'. Podcast *Alledaagse Vragen (BNNVARA)*, 18 augustus 2025

Waszink, Vivien. Interview over therapietaal. *Onze Taal*, september 2025.

Waszink, Vivien. [Interview](#) over de taal van De Jeugd van Tegenwoordig. NPO radio 1 *Nieuws en Co*, 3 oktober 2025

Waszink, Vivien. [Interview](#) over jongerentaal. In: *Linda*, 23 november 2025

Waszink, Vivien. [Interview](#) over 6-7 en kinderwoord van het jaar. *NOS Journaal*, 16 december 2025

Waszink, Vivien. [Interview](#) over woorden van het jaar. Podcast *Over Taal gesproken*, 16 december 2025

Diversen

Roland de Bonth

- Auteur en jurylid van het Groot Leiderdorps Dictee (5 november 2025, Leiderdorp)
- Coördinator en jurylid Profielwerkstukwedstrijdprijs van Levende Talen Nederlands 2025
- Redacteur *Trefwoord*, het digitale tijdschrift voor lexicografie en lexicologie.
- Redacteur Taalbeschouwing van het Handboek Didactiek Nederlands.
- Kolomleider literatuuronderwijs HSN 37, 28 & 29 november 2025, Tilburg University
- Auteur opdracht IV. Naamvallen in het Nederlands voor de Taalkundeolympiade 2025 van de Universiteit Leiden

Jesse de Does

- Lid van de Advisory Board van het GLOBALISE-project van het Huygensinstituut.
- Vertegenwoordigt België in de Technical Centres Committee (TCC) van CLARIN.

Katrien Depuydt

- Vertegenwoordiger van de voorzitter van de executive board in de managementboard van het IMPACT Centre of Competence.
- Lid van de Advisory Board van het GLOBALISE-project van het Huygensinstituut.

Frank Landsbergen

- Review voor online handboek Levende Talen Nederlands.

Vincent Vandeghinste

- [Nederlands Centraal](#), (2025) Panel discussion "Ontwikkeling AI voor het Nederlands", Taalunie, Rotterdam
- Nationaal coördinator CLARIN-België, lid van National Coordinators Forum CLARIN, lid van Board of Directors CLARIN, lid van Knowledge Infrastructure Committee CLARIN, Coördinator van K-Dutch, CLARIN knowledge centre for Dutch
- Lid van Taalunie Werkgroep AI voor het Nederlands (2025-2026)
- Redactielid *Computational Linguistics in the Netherlands Journal*.

Nicoline van der Sijs

- Lid adviesgroep Taalnormen en identiteitsvorming bij regionale en minderheidstalen in Europa van Hans Van de Velde
- Samenwerking met Horst Simon (Freie Universität Berlin), prof. dr. Natalia Filatkina (Universität Hamburg) en prof. dr. Andrea Rapp (Technische Universität Darmstadt), "Historische Fremdsprachenlehrwerke digital. Sprachgeschichte, Sprachvorstellungen und Alltagskommunikation im Kontext der Mehrsprachigkeit im Europa der Frühen Neuzeit" (DAGENTA en taalgidsjes)
- Samenwerking met Gina Kavaliūnaitė-Holvoet, Litouwen, in "The Chylinski Bible Database", gebaseerd op de Statenvertaling, en subsidieaanvraag ingediend voor dit project
- Samenwerking met Hanna Fischer, Andreas Bieberstedt (University of Rostock) (Alfred Lameli, Deutscher Sprachatlas in Marburg): "Lexical Information System for Low German (LIND)" (eWND)
- Externe begeleider postdoc project Anna Bosman en Robert Chamalaun over hoe spelvaardigheid verbeterd kan worden via etymologie
- Consultant project Ute Boonen 'Das Niederländische als Mittlersprache für Übersetzungen ins Deutsche im 17. und 18. Jahrhundert'
- Beschermvrouwe van project 'Nederlands ter Plekke' door Károli Universiteit Boedapest o.l.v. Julia Albert
- in 2025 Everwinus Wassenbergh Penning gekregen: <https://neerlandistiek.nl/2025/02/everwinus-wassenbergh-penning-voor-nicoline-van-der-sijs/> - in 2025 door de lezers van Onze Taal verkozen tot een van de drie invloedrijkste taalgebruikers van de eeuw
- mentor for the University of New Europe (UNE)
- lid van het Taalstaatpanel voor de verkiezing van Taalstaatmeester
- voorzitter van het bestuur van de Stichting Jiddische Lexicografie Amsterdam
- redacteur van *Neerlandistiek*

- redacteur van *Trefwoord*, tijdschrift op het gebied van lexicografie en lexicologie

Katrien Van pellicom

- Lidmaatschap Commissie Aardrijkskundige Namen (CAN) van de Taalunie.

Carole Tiberius

- Board of Trustees Adam Kilgariff Prize
- Nationale Coördinator CLARIN-NL
- National Anchor Point LDS (European Language Data Space)
- Deputy of the President of the ELEXIS Association
- Chair eLex Supervisory Board
- Secretary-Treasurer EURALEX (European Association of Lexicography)

Veronique De Tier

- Voorzitter Stichting Nederlandse Dialecten
- Bestuurslid Variaties vzw. Koepelorganisatie voor dialecten en oraal erfgoed.
- Bestuurslid Brusseleir!eind
- Redactielid Nehalennia
- Redactieraad De Lage Landen
- Redactieraad Taal en Tongval

Vivien Waszink

- Auteur lesmethode Nieuw Nederlands, Noordhoff, methodes vmbo 3 gt, 4 havo, 5 havo, 4 vwo, 5 vwo en 6 vwo.
- Lid Taalexplorium, een samenwerkingsverband van diverse organisaties die educatieve activiteiten rond taal organiseren in o.a. bibliotheken.
- Voorzitter ondernemingsraad

Bijlage 4: Taalmaterialen

B4.1 Downloads van taalmaterialen

Deze downloadstatistieken bevatten geen downloads door systeembeheer, webmaster en taalmaterialenmedewerkers. Dubbele downloads (van eenzelfde account) worden slechts eenmaal meegeteld. Downloads van bestelpakketten die enkel een licentie bevatten werden eveneens niet meegeteld.

Een aandachtspuntje is dat omwille van de GDPR een aantal spraakcorpora vanaf 20 maart niet meer rechtstreeks te downloaden waren: ook mensen die een niet-commerciële licentie wensten, moesten vanaf dan een getekende licentie opsturen. Die overgang is duidelijk in de cijfers te zien. Bijvoorbeeld van 1 januari tot 20 maart is de NC-versie van het CGN 46 keer gedownload. Vanaf 20 maart is het bestelpakket (met de licentie en bestelinstructies) van de NC-versie van het CGN 162 keer gedownload, maar er zijn slechts 38 getekende licenties binnengekomen.

Alle materialen waarvoor sinds 20 maart een getekende licentie nodig is:

- Corpus Gesproken Nederlands (CGN)
- Corpus Ondertitelde UvN-colleges (COUC)
- AUTONOMATA-POI-corpus
- AUTONOMATA-namencorpus
- IFA Dialogue Video Corpus
- IFA Corpus
- Bliss Spoken Dialogue Dataset

Niet-commerciële materialen	Aantal licenties
Corpus Gesproken Nederlands (CGN)	84
Frequentielijsten corpora	83
SoNaR-corpus	76
CELEX-2 Dutch	59
CGN-annotaties	46
CHN N-grams	40
Corpus Pathologische en Normale Spraak (COPAS)	32
Lassy Klein-corpus	31
GiGaNT-Molex	29
SoNaR Nieuwe Media Corpus	29
INT Historische Woordenlijst	28
Referentiebestand Nederlands (RBN)	25
CombiLex	24
AI-Trainingset - Tag de Tekst voor Named Entity Recognition (NER)	22
Dutch Parallel Corpus (DPC)	21
Medische Termen Belgisch-Nederlands (MedTermBN)	19
Spoken Academic Belgian Dutch Corpus (SABeD)	19
Cd-rom Middelnederlands	18
The LiLaH Emotion Lexicon of Greek, Kurdish, Turkish, Spanish, Farsi and Chinese	18
Lassy Groot-corpus	17
Corpus Gysseling (Data)	15
Eindhoven Corpus	15
PAROLE-lexicon	15
Dataset Synthetische Simplificatie	14
Corpus Middelnederlands (Data)	13
DuOMAn Subjectivity Lexicon	12
Dutch Idiom Database: Native Speakers (DID-NS)	12
Menselijke evaluatie van geautomatiseerde tekstvereenvoudiging: resultaten van crowdsourcing	12
RND Woordenlijsten	12
Wablief-corpus	12

e-Lex	9
NAMES Corpus	9
Referentiebestand Belgisch-Nederlands (RBBN)	9
SoNaR Character N-grams	9
Moroccorp	8
WAI-NOT Corpus	8
Belgian Covid Sign Language Corpus – BeCoS Corpus	6
BLISS Dialogue Summaries	6
Hoger Onderwijs Terminologie in Nederland en Vlaanderen- HOTNeV	6
Medische Pilot (MedPilot)	6
Brieven als Buit - Gouden Standaard	5
Corpus Ondertitelde UvN-colleges (COUC)	5
DuELME	5
Hotel Review Corpus in Nederlandse Gebarentaal- NGT_HoReCo	5
OMBI Nederlands-Arabisch	5
BLISS Spoken Dialogue Dataset	4
IFA Corpus	4
OMBI Arabisch-Nederlands	4
Annotated Corpora for Term Extraction Research (ACTER)	3
AUTONOMATA-transcriptietoolset	3
Boarnsterhim Corpus 2.0 (BHC 2.0)	3
Corpus Vlaamse Gebarentaal (Corpus VGT)	3
SumNL-samenvattingencorpus	3
AUTONOMATA-namencorpus	2
COREA-coreferentiecorpus	2
DAESO-corpus: parallelle Nederlandstalige monolinguale treebank	2
Dutch C-CLAMP	2
Gold Standard Parallel Corpus of Sign and spoken Language - GoSt-ParC-Sign	2
IFA Dialogue Video corpus	2
INT IMPACT NE-lexicon	2
Woordenboek Vlaamse Gebarentaal (Woordenboek VGT)	2
Afrikaans Genre Classification Corpus	1
Autshumato English-Afrikaans Parallel Corpora	1
Autshumato English-isiZulu Parallel Corpora	1
Dupira	1
Hotel Review Corpus in Spanish Sign Language - LSE_HoReCo	1
Hotel Review Corpus in Vlaamse Gebarentaal - VGT_HoReCo	1
isiXhosa Genre Classification Corpus	1
Lwazi Afrikaans Pronunciation Dictionary	1
Lwazi isiZulu ASR Corpus	1
Lwazi Sesotho Pronunciation Dictionary	1
Lwazi Siswati ASR Corpus	1

MuST-Cinema-PE: post-editing in automatic subtitling	1
OMBI Nederlands-Deens	1
OMBI Nederlands-Indonesisch	1
PaCo-MT Parallele Corpora	1
PAROLE-lexicon	1
Eindtotaal	1015

Commerciële materialen	Aantal licenties
Corpus Gesproken Nederlands (CGN) Commercieel	9
COPAS	8
JASMIN-spraakcorpus Commercieel	6
BasiLex-lexicon Commercieel	3
BasiLex-corpus Commercieel	2
BasiScript-corpus Commercieel	1
DuELME Commercieel	1
Dutch Idiom Database: Native Speakers (DID-NS) Commercieel	1
e-Lex Commercieel	1
Lassy Groot-corpus Commercieel	1
SoNaR Groot-corpus Commercieel	1
SumNL-samenvattingencorpus Commercieel	1
Eindtotaal	35

B4.2 Toevoegingen en updates aan Taalmaterialen

Boarnsterhim 2.0	Het Boarnsterhim Corpus bestaat uit 250 uur spraak in zowel West-Fries als Nederlands door dezelfde groep tweetalige sprekers. Het corpus bevat originele opnamen uit 1982-1984 en een replicerend onderzoek 35 jaar later.	Nieuwe versie
Corpus Hedendaags Nederlands - CHN (Online)	Het Corpus Hedendaags Nederlands (CHN) is een tekstverzameling met ongeveer 10,1 miljoen teksten uit boeken, blogs, kranten, tijdschriften en journaaluitzendingen uit Nederland, Vlaanderen, Suriname en de Nederlandse Antillen.	Nieuwe versie
Couranten Corpus	Het Couranten Corpus bevat dertien zeventiende-eeuwse Nederlandse kranten uit de periode 1619-1700 die momenteel op Delpher.nl beschikbaar zijn. Er zijn 109.532 artikelen in te vinden met een totaal van 18.926.425 woorden.	Nieuwe versie

Eindhoven Corpus	Het Eindhoven Corpus is een verzameling Nederlandstalige geschreven en getranscribeerde gesproken teksten uit de periode van 1960 tot 1976. Het corpus bevat ca. 768.000 tokens.	Nieuwe versie
Spoken Academic Belgian Dutch Corpus - SABeD	Het Spoken Academic Belgian Dutch Corpus (SABeD) bestaat uit 200 colleges gegeven op Vlaamse hogescholen en universiteiten.	Nieuwe versie
Corpus Oudfries (Online)	Het Corpus Oudfries bevat een grote aantal woorden van de Oudfriesse taal van ca. 1200-1550, die door Rita van de Poel zijn gelemmatiseerd en van PoS-markeringen voorzien als onderdeel van haar promotieonderzoek.	Nieuw taalmateriaal
Dutch C-CLAMP	The Dutch Corpus of Contemporary and late Modern Periodicals (Dutch C-CLAMP) is een corpus dat bestaat uit een verzameling artikelen uit 13 culturele of literaire tijdschriften die in Vlaanderen en Nederland zijn gepubliceerd. Het is een historisch corpus met teksten uit de periode tussen 1837 en 1999.	Nieuw taalmateriaal
Elektronische Woordenbank van de Nederlandse Dialecten - eWND (Online)	De elektronische Woordenbank van de Nederlandse Dialecten (eWND) bevat een alsmaar groeiend aantal oude en modernere Nederlandse dialectwoordenboeken.	Nieuw taalmateriaal
GaLAHaD (Online)	GaLAHaD (Generating Linguistic Annotations for Historical Dutch) is ontwikkeld binnen het CLARIAH-PLUS-project. De applicatie biedt een flexibele omgeving voor automatische verrijking van historische teksten.	Nieuw taalmateriaal
Gesproken Corpus van de zuidelijk-Nederlandse Dialecten - GCND (Online)	Het GCND-corpus (voluit Gesproken Corpus van de zuidelijk-Nederlandse Dialecten) is een taalkundig geannoteerd corpus van gesproken dialecten uit het zuidelijke Nederlandse taalgebied — met name dialecten uit België, het noorden van Frankrijk en het zuiden van Nederland.	Nieuw taalmateriaal
Historical Corpus of Dutch - HCD (Online)	Het Historisch Corpus van het Nederlands (HCD) is een diachronisch, regionaal gebalanceerd corpus van verschillende genres geschreven Nederlands.	Nieuw taalmateriaal
LAnCeLoT (Online)	LAnCeLoT (Linguistic Annotation Corpus Laundry Tool) is ontwikkeld binnen het CLARIAH-PLUS-project. De applicatie stelt onderzoekers in staat	Nieuw taalmateriaal

	om taalkundige verrijkingen handmatig te corrigeren en te verfijnen.	
Nederlands als Wetenschapstaal: Natuurkunde (Data)	Nederlands als wetenschapstaal: Natuurkunde (data) is een dataset met natuurkundige vaktermen in het Nederlands die is ontwikkeld als onderdeel van het project Nederlands als wetenschapstaal. Het materiaal bevat een lijst met fysische termen inclusief definities en Engelse equivalenten	Nieuw taalmateriaal
Nederlands als Wetenschapstaal: Scheikunde (Online)	Nederlands als wetenschapstaal: Scheikunde (online) is een online taalmateriaal dat deel uitmaakt van het project Nederlands als wetenschapstaal. Het biedt een zoekbare lijst met scheikundige terminologie in het Nederlands, waarin Nederlandse vaktermen uit de chemie zijn verzameld met Engelse equivalenten.	Nieuw taalmateriaal
SUBCAT-lexicon	Het SUBCAT-lexicon is een uitgebreid subcategorisatielexicon voor werkwoorden. Het bestaat uit ongeveer 12.012 werkwoordsingangen met abstracte subcategorisatieschemata.	Nieuw taalmateriaal
Termwerk (Online)	TermWerk is een webapplicatie voor termextractie, termbeschrijving en termbeheer in het Nederlands.	Nieuw taalmateriaal
Textlens (Online)	Textlens is een online tekstverwerkingsdashboard voor taken als automatische tokenisatie, lemmatisering, tagging, named entity recognition en afhankelijkheidsanalyse voor Nederlands, Engels, Frans en Duits.	Nieuw taalmateriaal
Woordpeiler (Online)	Met Woordpeiler kan men bekijken hoe vaak woorden door de tijd heen voorkomen in teksten uit Nederlandstalige kranten.	Nieuw taalmateriaal
Dataset containing hypothetical manner clauses in English and Dutch	Deze dataset bevat door komma's gescheiden bestanden (csv) met informatie over het gebruik van bijzinnen die worden ingeleid door het voegwoord 'as if' in de hedendaagse Britse spraak en van bijzinnen die worden ingeleid door het voegwoord 'alsof' in de hedendaagse Nederlandse spraak.	Nieuw taalmateriaal
Dutch Renaissance Poetry Corpus	Het Dutch Renaissance Poetry Corpus bevat alexandrijnen en jambische pentameters geschreven door een selectie van Nederlandse Renaissancedichters (eind 16de en 17de eeuw).	Nieuw taalmateriaal

Text2Picto (Online)	Text2Picto is een vertaaltool die bedoeld is om de communicatie voor mensen met een leesbeperking te verbeteren. De tool vertaalt zinnen naar pictogrammen, dat wil zeggen grafische symbolen die dienen als vervanging voor verbale communicatie.	Nieuwe versie
Global Anglicism Database - GLAD (Online)	De Global Anglicism Database bevat meer dan 80,000 Engelse leenwoorden die voorkomen in zeventien verschillende talen.	Nieuw taalmateriaal
Machine Translation Evaluation Online	MATEO is een online platform voor het evalueren van machinevertalingen via een gebruiksvriendelijke interface.	Nieuw taalmateriaal
The Digital Pallas (Online)	De Digitale Pallas is de digitale versie van de Comparative Dictionary of All Languages and Dialects (1790-1791), een Russisch woordenboek van Peter Simon Pallas.	Nieuw taalmateriaal
GCND GrETEL (Online)	GCND GrETEL is de zoek- en query-applicatie voor het Gesproken Corpus van de zuidelijk-Nederlandse Dialecten (GCND). Het maakt het mogelijk om de handmatig geverifieerde syntactische annotaties van het GCND-corpus te doorzoeken en te analyseren.	Nieuw taalmateriaal