

groting van de taak: *horizontale taakverruiming*, z  
*taal*; *verticale taakverruiming*, zie bij *verticaal*<sup>1</sup>;  
**ving** (v.), structurele wijziging in taken of de  
binnen een bedrijf, instelling of afdeling; ...  
(v.); ...**waardering** (v.); ...**werk** (o.), werk dat  
richt wordt.

**taal** (v.(m.); -tje) [Mnl. *tale* (hetgeen iem. zegt, ge  
Limburg ook: getal), ~ *tellen*, de Idg. verbinding  
ker], **1** (g.mv.) het op basis van een universeel  
vermogen gevormde geheel van tekens en regel  
betekenisdragende elementen te combineren  
mens gebruik maakt, om zijn gedachten te art  
wereld te ordenen en te communiceren, syn.  
*taal is een evoluerend, sociaal symbolencomplex,*  
*gearticuleerde intrinsiek-variabele, arbitraire eenheid*  
*waarop de mens in en aan de periferie van sociaal*

# Drie landen, één taalbeleid

Blijft het Nederlands overeind in het geweld van tegen elkaar opboksende wereldtalen zoals het Engels, het Arabisch en het Chinees? En wat kan de overheid doen om onze kleine taal in leven te houden? Veel meer dan u zou denken.

Eerst even een misverstand rechtzetten: het Nederlands is geen kleine taal. Het is de moedertaal van ruim 15 miljoen Nederlanders en 6 miljoen Vlamingen; en nog eens 400.000 Surinamers maken er dagelijks gebruik van. Er zijn ruim 6.000 talen in de wereld en op de ranglijst staat het Nederlands tussen de 45<sup>ste</sup> en 50<sup>ste</sup> plaats. Dat is hoger dan bijvoorbeeld het Grieks of het Zweeds. Sommige onderzoekers houden ook rekening met het belang van een taal, bijvoorbeeld in publicaties, bij internationale contacten of op het internet. Zij plaatsen het Nederlands zelfs in de top twintig.

### Een gezonde taal

Het Nederlands verkeert dus in stralende gezondheid. Toch is er zorg nodig als we dit zo willen houden, want talen kunnen snel terrein verliezen. We zagen dit in de voorbije decennia gebeuren met onze dialecten. Waar ouders hun streektaal niet meer doorgaven aan hun kinderen, verdween deze heel snel en definitief.

Essentieel voor het voortbestaan van een taal is dat ze beschikt over alle middelen om zich te ontwikkelen en aan te passen aan de voortdurend veranderende behoeften. Een taal die men niet kan schrijven, verdwijnt in tijden waarin schrijven belangrijk is. Daarom is een vaste spelling essentieel. Een taal die men niet kan leren in een school voor vertalers en tolken, verdwijnt uit het assortiment waarmee internationale contacten worden onderhouden. Daarom zijn goede vertaalinstellingen en goede opleidingen Nederlands in het buitenland noodzakelijk. Een taal die niet meedoet met de ontwikkelingen op het internet, wordt zwakker naarmate we meer informatie halen van het web. Daarom zijn voorzieningen zoals woordenlijsten en grote hoeveelheden tekstmateriaal nodig, die ontwikkelaars van digitaal materiaal en software kunnen gebruiken voor hun producten.

### Meer middelen voor taal

Het Nederlands is dan wel geen kleine taal, maar de overheid die een taalbeleid kan voeren om het gezond te houden, is verspreid over drie landen. Los van elkaar zijn hun mogelijkheden erg beperkt. Samen werken ze efficiënter en kunnen ze meer middelen inzetten. Daarvoor hebben ze de Taalunie.

De Nederlandse Taalunie is een intergouvernementele organisatie die werd

opgericht in 1980 door een verdrag tussen Nederland en België. Sinds 2004 is Suriname geassocieerd lid. Aanvankelijk was het doel van de Taalunie 'integratie van Nederland en de Vlaamse gemeenschap in België op het gebied van de Nederlandse taal en letteren'. Tegenwoordig wordt de nadruk minder gelegd op integratie dan op samenwerking ten behoeve van de taalgebruiker, en op behoud van het taalkundige en literaire erfgoed.

De belangrijkste werkerterreinen van de Taalunie zijn: de taal zelf, het Nederlands in digitale toepassingen, het onderwijs in en van het Nederlands, literatuur en leesbevordering, en de positie van het Nederlands in Europa en in de wereld.

### Samenwerken helpt

In al die gevallen wordt gezocht naar samenwerking tussen de deelnemende landen. Zo is de Taalunie alleen bevoegd voor onderwijs waar het over gezamenlijke projecten gaat, naast het beleid van de Nederlandse, Vlaamse en Surinaamse minister van Onderwijs. Ook op het werkerterrein van de literatuur en leesbevordering is dat goed zichtbaar. Er zijn bijvoorbeeld fondsen voor de letteren in Nederland en Vlaanderen; de taak van de Taalunie is die fondsen te laten samenwerken in projecten die de landsgrenzen overschrijden. Een derde sprekend voorbeeld is te vinden in het Instituut voor Nederlandse Lexicologie, waar in opdracht van de Nederlandse Taalunie bestanden worden aangelegd die zowel bruikbaar zijn voor de productie van woordenboeken als voor ICT-toepassingen. Geen van de drie landen zou zulke belangrijke, maar kostbare voorzieningen zelfstandig kunnen financieren. Samen lukt het wel. Hetzelfde principe geldt bij het STEVIN-programma<sup>103</sup>. Hier heeft de Taalunie een groot deel van de beleidsvoorbereiding voor haar rekening genomen en de financierende overheden bij elkaar gebracht.

*Linde van den Bosch,  
Algemeen secretaris van de  
Nederlandse Taalunie*

## Nederlandse Taalunie

### Wat is de Taalunie?

De Nederlandse Taalunie is een organisatie waarin Nederland, Vlaanderen en Suriname de krachten bundelen in dienst van de Nederlandse taal, het taalonderwijs en de letteren. Ministers bepalen het beleid en parlementsleden zien erop toe. Een raad van deskundigen adviseert. De voorbereiding en uitvoering van het beleid zijn in handen van een algemeen secretariaat, dat samenwerkt met commissies en werkgroepen.

Het doel van de Taalunie is de gebruiker van het Nederlands te ondersteunen. Daarmee blijft het Nederlands een levenskrachtige taal. De strategische uitgangspunten van de Nederlandse Taalunie kunnen we samenvatten in vier slogans:

- Het Nederlands past overal
- Het Nederlands klinkt in vele stemmen
- Het Nederlands tussen andere Europese talen
- Het Nederlands is ons culturele huis

Lange Voorhout 19,  
Postbus 10595  
2501 HN Den Haag, Nederland

T: + 31 70/346 95 48  
F: + 31 70/365 98 18  
E: info@taalunie.org

Meer informatie over de werkzaamheden van de Taalunie vindt u op [taalunieversum.org](http://taalunieversum.org). Ook de publicaties van de Taalunie zijn via de website te bestellen.

<sup>103</sup> Zie ook elders in dit nummer: p. 30.

# Spraakmakende en betekenisvolle technologie

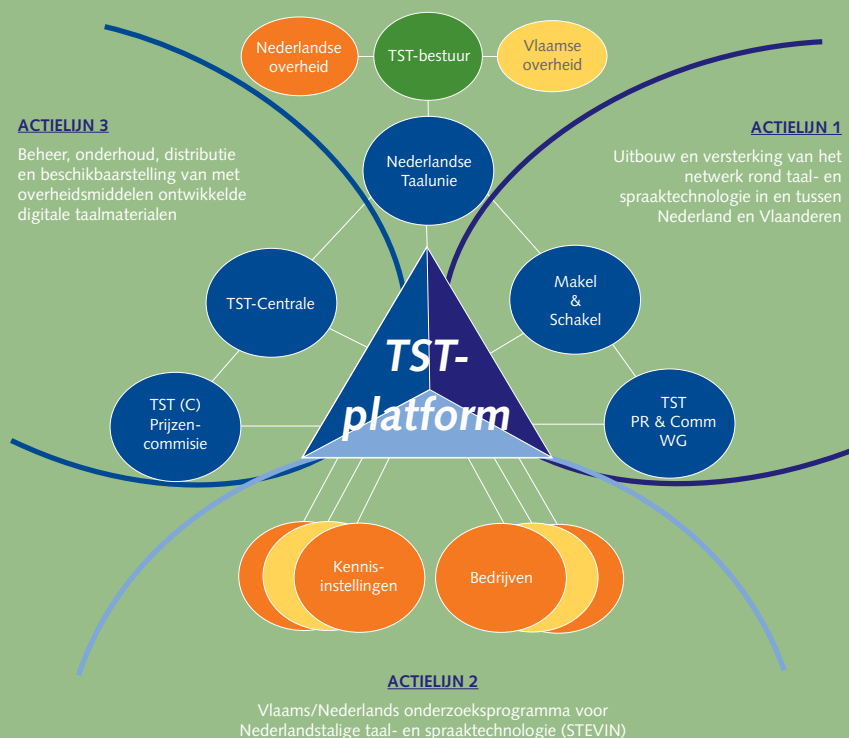
Ondanks alle (kleine) verschillen en varianten, gebruiken Vlamingen en Nederlanders dezelfde taal. Samenwerking tussen Vlaanderen en Nederland op het vlak van taal- en spraaktechnologie (TST) lijkt dan ook evident. Toch heeft het tijd gevergd. Het spraakmakende debacle van het spraakbedrijf Lernout & Hauspie Speech Products in 2001 heeft, zeker in Vlaanderen, voor een echte schokgolf gezorgd. Niet alleen bij onderzoekers, industriëlen en het gerecht, maar ook bij de vele modale burgers die hoopten een graantje mee te pikken. Nog steeds wekt TST wantrouwen op. Ten onrechte.

## Vanuit een diepe leperse vallei ...

In 1999 richtten Vlaanderen en Nederland het TST-platform op, om de activiteiten rond taal- en spraaktechnologie te stroomlijnen en coördineren. Drie actielijnen stonden centraal (zie figuur 2): netwerken; inventariseren en prioriteren van onderzoek en ontwikkeling; onderhoud van materialen.

2004 was het startjaar van een effectieve en geïntegreerde samenwerking op het vlak van taal- en spraaktechnologie voor het Nederlands. Deze samenwerking lag in het verlengde van de drie eerder uitgestippelde actielijnen. Een TST-Centrale<sup>104</sup> (actielijn 3) werd boven de doopvont gehouden, een STEVIN-onderzoeks- en stimuleringsprogramma<sup>105</sup> (actielijn 2) ging van start en een makel- en schakel-functie (actielijn 1) werd actief. STEVIN<sup>106</sup> (Spraak- en Taaltechnologische Essentiële Voorzieningen In het Nederlands) staat in voor de creatie van digitale basisvoorzieningen onder de vorm van onderzoeksresultaten (datacollecties of corpora verrijkt met bijkomende informatie, software, demonstratoren, algoritmes, ...). Het makelen en schakelen zorgt ervoor dat de spelers in kaart en met elkaar in contact gebracht worden. De TST-Centrale dient om de digitale TST-voorzieningen te onderhouden en distribueren.

Figuur 2: De drie TST-actielijnen gecoördineerd door de Taalunie





### Naar een piek in de polders

*Een TST-relatiebureau: makelen en schakelen*

De makel- en schakelfunctie heeft als doel Vlaamse en Nederlandse spelers in het TST-veld – de kenniscentra, technologieaanbieders, softwareproducenten, belangenorganisaties en overheden – te identificeren en bij elkaar te brengen. Het verspreiden van een TST-nieuwsbrief<sup>107</sup> en het organiseren van de TST-beurs Taal in Bedrijf<sup>108</sup> zijn belangrijke activiteiten in dit kader. Het makelen en schakelen wordt op continue basis uitgevoerd door de Nederlandse Taalunie<sup>109</sup>, die hiervoor onder meer een centrale Infodesk<sup>110</sup> beheert en onderhoudt. Een specifiek onderdeel ervan is de aanwezigheid in Europese netwerken die TST voor het Nederlands ten goede komen.

*Een TST-fabriekje: het STEVIN-programma*

Het bedrijfsleven is niet geneigd te investeren in de opbouw van een digitale basisinfrastructuur voor het Nederlands gezien de beperkte afzetmarkt, bijvoorbeeld in vergelijking met het Engels. Dergelijk marktfalen dwingt overheden ertoe om zelf de investeringen te dragen. Enkele doelstellingen staan hierbij voorop. Een

adequate digitale taalinfrastructuur van essentiële data en tools voor het Nederlands moet zowel voor de kennisinstellingen als het bedrijfsleven beschikbaar worden (doelstelling 1). De kennisinstellingen hebben de data nodig als grondstof en de tools als hulpmiddel bij het uitbouwen van hun wetenschappelijke activiteiten (doelstelling 2a); bijvoorbeeld nieuwe methodes om informatie uit documenten te filteren. Voor het bedrijfsleven zijn ze onmisbaar om complexe applicaties te ontwikkelen voor een aantal talen, maar primair voor het Nederlands (doelstelling 2b); bijvoorbeeld voor een navigatiesysteem dat straatnamen in het Nederlands correct uitspreekt. De economische en culturele positie van het Nederlands in de wereld van informatietechnologie moet immers behouden blijven (doelstelling 3).

Bijkomende flankerende activiteiten bevorderen de kennisoverdracht en vraagstimulering (doelstelling 4). Demonstratieprojecten moeten aantonen dat TST voor het Nederlands een rijpe technologie is om nuttige en rendabele toepassingen te maken (zie ook verder). Dit om de vraag naar dergelijke producten aan te wakkeren. Netwerksubsidies worden verleend om workshops of andere evenementen in te richten om kennis te

verspreiden. Enkele educatieve projecten hebben specifiek als doel om TST in de kijker te plaatsen bij studenten en schoolgaande jeugd. Zo ondersteunt STEVIN mee de opbouw van een TST-sectie binnen de taalwetenschappenvakpagina van de populaire Nederlandse site 'Kennislink'<sup>111</sup>. Beslissers en investeringsverantwoordelijken binnen een organisatie kunnen hun kennis over TST bijspijkeren dankzij TST-masterclasses.

Het adequaat regelen van intellectuele eigendomsrechten van de materialen en softwaretools wordt in nauwe samenwerking met de TST-Centrale gerealiseerd. In principe wordt de Taalunie eigenaar van alle projectresultaten, wat niet gebruikelijk is bij onderzoeksprojecten. Dit opent evenwel de mogelijkheid om een beleid van open innovatie<sup>112</sup> op te zetten via de TST-Centrale. Bedrijven kunnen zo eigen specifieke producten ontwikkelen op basis van gemeenschappelijke basiscomponenten (doelstelling 5).

### Wat rolt er van de band?

STEVIN levert een waaier aan resultaten op<sup>114</sup>. Een onderscheid kan gemaakt worden tussen onderzoeksprojecten en demonstratieprojecten:

- Zo leveren verschillende onderzoeksprojecten (€ 400-500.000 subsidie) een gegevensverzameling af (corpus) van teksten die verrijkt werden met allerhande taalkundige informatie, bv. woordsoort, noemvorm, .... Ook hulpmiddelen en procedures om verrijkingen toe te voegen aan een corpus worden opgeleverd. Dit type van projecten richt zich specifiek op de digitale basisinfrastructuur. Andere betreffen eerder onderzoek en ontwikkeling, bv. het verbeteren van een spraakherkenner en onderdelen die er nuttig voor zijn, zoals methodes om beter met ruis en achtergrondlawaai om te gaan. Ten slotte zijn er onderzoeksprojecten die toepassingsgericht zijn. Daarbij wordt bijvoorbeeld software ontwikkeld die teksten automatisch samenvat en vertaalt. Of die informatie distilleert uit teksten. Een ander project is dan weer maatschappelijk relevant omdat het methodes ontwikkelt die het alfabetiseringsproces van allochtonen digitaal ondersteunen. Hoofdzakelijk kennisinstellingen nemen deel aan onderzoeksprojecten. In totaal werden negentien dergelijke projecten goedgekeurd.
- De demonstratieprojecten (€ 50-100.000) beslaan een breed gamma aan toepassingen. Van het ondertitelen van tv-programma's tot de automatische productie van een gesproken krant voor blinden over spellingscorrectoren voor dyslectici, jargonbewakers, automatische antwoordlijnen en het slim opzoeken van informatie in webportalen. Dergelijke projecten zijn in hoofdzaak bestemd voor kmo's. Deze passen kant-en-klare technologie toe in nieuwe producten of diensten voor 'echte' klanten.

Een mooi voorbeeld van de samenwerking tussen bedrijfsleven en kennisinstellingen, Vlaamse en Nederlandse partners, en de aanwending van basisinfrastructuur in een toegepast project? Autonomata en Automata Too. Hetzelfde consortium schuilt achter beide projecten: TeleAtlas<sup>114</sup>, Nuance Communications International<sup>115</sup> en de Universiteit Gent<sup>116</sup> aan Vlaamse zijde en de universiteiten van Utrecht en Nijmegen aan Nederlandse zijde. Het eerste Autonomataproject leverde onder meer een specifieke softwarecomponent op, waarmee eigennamen beter uitgesproken en herkend worden. Deze component wordt meegeleverd met de commerciële Nuancesoftware. Vooral voor autonavigatiesystemen is een dergelijke verbetering van groot belang; vandaar de betrokkenheid van TeleAtlas. Het Autonomata Tooproject bouwt hierop verder. Het doel is om een demonstrator te bouwen, die informatie verschaft over zogenaamde 'points of interest' (POIs). Dit zijn lokale bezienswaardigheden, officiële instanties – toerismekantoor, ziekenhuis, politiekantoor, ... – openbaar transport, hotels, banken enz. De namen van deze POIs kunnen op vele manieren worden uitgesproken. Soms al in het Nederlands, maar zeker als een eigenaam of vreemd woord deel uitmaakt van de benaming van het POI. Het uitspreken en herkennen van dergelijke namen kan erg lastig zijn voor een computer. Het probleem wordt des te prangender gezien het belang van deze POIs, ook voor buitenlanders, wat de spraakvariatie alleen maar vergroot. Autonomata Too probeert hiervoor een oplossing te vinden.

### En wat daarna?

EWI startte een TST forecastoefening<sup>117</sup> op. Uit enkele voorbereidende brainstormen en rapporten distilleerden we 45 stellingen, die door een 70-tal experts beoordeeld en becommentarieerd werden. Deze resultaten zijn voorgesteld tijdens Taal in Bedrijf 2008<sup>118</sup>. Ook de Nederlandse overheid en enkele belangorganisaties brachten hun wensen naar voor. Op basis van dit alles streven we in de komende periode naar een synthese waar alle financierende overheden kunnen achter staan, wat hopelijk resulteert in een opvolgingsprogramma voor STEVIN.

*Peter Spyns,  
Afdeling Strategie en Coördinatie en  
STEVIN-programmacoördinator voor de  
Nederlandse Taalunie*



## nū | STEVIN

### Het budget van STEVIN

Het totaalbudget voor STEVIN bedraagt € 11,4 miljoen. De Vlaamse overheid, vertegenwoordigd door het departement EWI, het IWT en het FWO, staat in voor 1/3<sup>de</sup> (€ 3,8 miljoen). De overige 2/3<sup>de</sup> (€ 7,6 miljoen) komt van de Nederlandse ministeries van Economische Zaken en van Onderwijs, Cultuur en Wetenschap, en de Nederlandse Organisatie voor Wetenschappelijk Onderzoek. De Nederlandse Taalunie<sup>119</sup>, als intergouvernementele organisatie, treedt op als het financiële verdeelpunt, overziet en coördineert de activiteiten. STEVIN loopt tot eind 2011.

104 Zie elders in dit nummer: p. 33.

105 <http://www.stevin-tst.org>.

106 Zie [http://www.stevin-tst.org/documenten/stevin\\_mjp\\_150904.pdf](http://www.stevin-tst.org/documenten/stevin_mjp_150904.pdf) voor een uitvoerige beschrijving van het STEVIN-meerjarenplan.

107 <http://taalunieversum.org/taal/technologie/nieuwsbrief>.

108 <http://www.taalinbedrijf.org>.

109 Zie ook elders in dit nummer: p. 28.

110 <http://taalunieversum.org/taal/technologie>.

111 <http://www.kennislink.nl/web/show?id=129397>.

112 EWI-Review (2) 2: 46 – 49.

113 We verwijzen de geïnteresseerde lezer naar twee thematische uitgaven van het tijdschrift DIXIT (van de Nederlandse organisatie voor TAAI- en Spraaktechnologie – NoTAS) over STEVIN. Elk project beschrijft zichzelf in één pagina: [http://www.stevin-tst.org/pers/#in\\_media](http://www.stevin-tst.org/pers/#in_media).

114 <http://www.teleatlas.com/index.htm>.

115 <http://netherlands.nuance.com>.

116 <http://www.elis.ugent.be>.

117 <http://ewi-vlaanderen.be/ewi/index.php?id=485>.

118 <http://www.taalinbedrijf.org>.

119 Zie elders in dit nummer: p. 28.

# De TST-Centrale: voor het behoud van digitaal Nederlandstalig onderzoeksmateriaal

Voor onderzoek en ontwikkeling in de Nederlandstalige taal- en spraaktechnologie is er voldoende digitaal bronmateriaal van het Nederlands nodig. Vaak wordt dit materiaal in onderzoeksinstituten samengesteld met financiële steun van de overheid. Om kapitaalvernietiging tegen te gaan, wordt beheer en hergebruik van de materialen gestimuleerd door de materialen te centraliseren in een Nederlands-Vlaams portaal: de Centrale voor Taal- en Spraaktechnologie of TST-Centrale.



De materialen die bij de TST-Centrale beschikbaar zijn, bestaan uit verzamelingen teksten of spraakfragmenten (corpora), geordende lijsten van verschillende (delen van) woorden (lexica), elektronische woordenboeken en taal- en spraaktechnologische software (tools; bijvoorbeeld een spellingchecker). De verzamelingen tekst en spraak zijn in de meeste gevallen verrijkt met taalkundige informatie over vorm, uitspraak, woordsoort, betekenis en onderlinge relaties in een zin. Deze corpora, lexica, woordenboeken en tools zijn vaak het resultaat van een door de overheid gesubsidieerd project. Projectsubsidies dekken over het algemeen alleen de kosten voor de ontwikkeling, maar niet voor het beheer en de distributie van de uiteindelijke resultaten. Het belangrijkste doel van de TST-Centrale is daarom de materialen na afloop van het project in beheer te nemen en beschikbaar te stellen voor nieuwe projecten op het gebied van onderwijs, onderzoek en ontwikkeling.

## STEVIN draagt bij

Een belangrijke leverancier van taalmaterialen is het Vlaams-Nederlandse subsidieprogramma STEVIN (Spraak- en Taaltechnologische Essentiële Voorzieningen In het Nederlands)<sup>120</sup>. De resultaten van de

STEVIN-projecten worden overgedragen aan de Nederlandse Taalunie, waarna de TST-Centrale voor het verdere beheer en de distributie zorgt. Bijzonder is dat de TST-Centrale nauw betrokken is bij bepaalde aspecten van de STEVIN-projectprocedure: eigendomsrecht kwesties werden samen met een gespecialiseerde STEVIN-commissie en juristen geklaard; er werd een beleid opgesteld voor het opleveren van resultaten als open source; en de TST-Centrale verzorgde enkele publieke themadagen over de lopende projecten. Inmiddels is een eerste reeks STEVIN-projectresultaten beschikbaar via de TST-Centrale.

Beschikbaarstelling van de materialen verloopt via de website van het Instituut voor Nederlandse Lexicologie (het instituut waarbij de TST-Centrale ondergebracht is<sup>121</sup>). In de online productcatalogus wordt van alle beschikbare materialen een afzonderlijke productbeschrijving gegeven. Vele ervan zijn online te doorzoeken en direct te bestellen bij de TST-Centrale. Na het ondertekenen (of online accepteren) van een gebruikersovereenkomst kunnen vrijwel alle producten gratis gebruikt worden voor onderzoekdoeleinden. Voor commerciële partijen geldt een marktconforme prijs.

## Over toegevoegde waarde

De medewerkers van de TST-Centrale hebben, naast praktische kennis over beheer en distributie, ook inhoudelijke kennis opgebouwd over de verschillende materialen. De servicedesk helpt gebruikers met specifieke vragen of problemen. Indien nodig kan er contact opgenomen worden met de ontwikkelaars van het materiaal. Bij de servicedesk kan men ook workshops of gastcolleges over bepaalde materialen aanvragen. Op die manier zorgt de TST-Centrale ervoor dat de waarde van digitale onderzoeksmaterialen voor het Nederlands en de kennis ervan niet verloren gaat.

*Laura van Eerten,  
TST-Centrale*

## Meer?

In 2004 is de TST-Centrale opgericht op initiatief van de Nederlandse Taalunie. De TST-Centrale wordt gefinancierd door de Nederlandse Taalunie en is als project ondergebracht bij het Instituut voor Nederlandse Lexicologie, met vestigingen in Leiden en Antwerpen.  
Website: [www.inl.nl/tst-centrale](http://www.inl.nl/tst-centrale) e-mail: [servicedesk@inl.nl](mailto:servicedesk@inl.nl)

<sup>120</sup> Zie ook elders in dit nummer: p. 30.

<sup>121</sup> <http://www.inl.nl>.