

■ TAALWETENSCHAP

Op zoek naar de oertaal

Met technieken uit de biologie reconstrueren taalkundigen taalstambomen. Met grammatica kunnen ze verder in het verleden kijken dan met woordenschat.

door bruno van wayenburg

HET zal de eerste en vermoedelijk ook de allerlaatste keer zijn dat er een woord in de taal Lavukaleve in het wetenschapsvakblad *Science* staat. 'Tau vegome', ofwel 'hand', is dat woord. Dat lijkt overigens helemaal niet op 'obi', hetzelfde woord in het Touo of 'kakau' in het Savosa-vo.

Toch zijn de drie talen verwant, samen met nog twaalf zogeheten Papoeatalen gesproken op eilanden ten oosten van Papoea-Nieuw-Guinea. Dat ontdekten Michael Dunn van het Max Planckinstituut in Nijmegen en collega's uit Nijmegen en Cambridge. Opmerkelijk aan de vondst is vooral dat de ontdekking gedaan werd met technieken die normaal het domein zijn van DNA- en evolutiebiologen.

De Papoeatalen, er zijn er in totaal 23, golden voorheen als taalkundig geïsoleerd, net zoals het Baskisch in Europa. Dunn liet zien dat ze in een ver verleden vermoedelijk toch teruggaan op één oertaal, hoewel de splitsing zo lang geleden is dat gemeenschappelijke woorden óf onherkenbaar verbasterd, of helemaal verdwenen zijn.

Opmerkelijk, want normaal gesproken baseren taalkundigen hun oordeel over de familiegeschiedenis van talen juist op overeenkomstige, verwante woorden, zoals het Nederlandse 'vader' en het Duitse 'Vater', die terug te voeren zijn op een gemeenschappelijke oervorm (zie *Europese eenwording begon in oertijd*).

Ondanks ontegenzeggelijke successen zijn er grenzen aan die aanpak. Talen slijten. Klanken veranderen of vallen weg, woorden worden vervangen door leenwoorden van naburige talen, of ze veranderen onherkenbaar van betekenis. Sporen van verwantschap vervagen daarom op de lange duur. Voorbij zo'n 10.000 jaar geleden zijn de paar overblijvende verwante woorden niet meer te onderscheiden van toevallige overeenkomsten, en is er weinig zinnigs te zeggen over taalverwantschap, vermoeden de meeste taalkundigen.

Pogingen om dat toch te doen, bijvoorbeeld de stelling van Merrit Ruhlen en Joseph Greenberg dat ze woorden uit de allereerste oertaal



Melanesische zeilkano. Een nieuwe stamboom van de Papoeatalen die in dit gebied ten oosten van Papoea-Nieuw-Guinea worden gesproken, komt op z'n hoofdtakken netjes overeen met de verschillende eilandengroepen. © science

hadden gereconstrueerd ('tik' zou 'vinger' betekenen en 'akwa' water), worden door collega's weinig serieus genomen.

Maar een taal is meer dan woorden alleen. Ze bestaat ook uit grammaticaregels, en er zijn aanwijzingen dat die beter tegen erosie bestand zijn. „De Papoeatalen laten opvallende structurele overeenkomst zien”, zegt Dunn. Zo hebben de meeste het werkwoord meestal aan het eind van de zin, in tegenstelling tot de omliggende Austronesische talen. Ook maken ze vaak geen onderscheid tussen de klank 'l' en 'r', en hebben ze vaak grammaticale geslachten, zoals het Frans of ook het Nederlands. „Dat laatste komt niet zo vaak voor dat dat een toevallige overeenkomst zou zijn”, zegt Dunn.

Al eerder zijn er daarom verwantschappen tussen Papoeatalen voorgesteld, al waren die zonder bewijs in de vorm van verwante woorden moeilijk hard te maken. Op aanwijzing van zijn Britse collega's besloot Dunn om het daarom over een andere boeg te gooien, en de vermoedens te toetsen met trucs van evolutiebiologen. Als hun technieken aan de hand van gemuteerde DNA-codes een stamboom van levende wezens kunnen opstellen, was het idee, dan kunnen wij taalkundigen aan de hand van muterende grammaticale eigenschappen best een taalstamboom maken.

Van vijftien Papoeatalen codeerden Dunn en collega's 125 grammaticale eigenschappen, van het type 'de taal heeft een speciale meervoudsvorm voor twee stuks

(ja/nee)', of 'de taal heeft voorzetsels (ja/nee). De hele exercitie leverde een lange lijst van 1.875 ja/nee-antwoorden op. „Bijna elk antwoord in die lijst kostte ons twintig minuten hard nadenken”, zegt Dunn, die zich baseerde op gepubliceerde taalbeschrijvingen en gegevens uit het veldwerk van zijn eigen onderzoeksgroep.

Na dit monnikenwerk zetten de onderzoekers de computer aan, met software geleend van hun biologische collega's. Die zoekt tussen de duizenden mogelijke stambomen naar de 'zuinigste' stamboom, ofwel de stamboom met zo'n structuur dat het aantal veranderingen van een grammaticale eigenschap tijdens de evolutie minimaal is.

Het idee is dat de eenvoudigste verklaring voor de gegevens, dus die met het minste aantal 'mutaties', ook de meest plausible is.

Getest op de zestien ook gecoedeerde talen van de Austronesische taalfamilie leverde deze methode inderdaad een stamboom die heel aardig klopte met de laatste taalkundige inzichten.

Toen volgde de proef op de som: het zoeken van een zuinige stamboom voor de Papoeatalen. De vraag was van tevoren hoe je het resultaat zou moeten interpreteren. Zelfs uit een aanwezigheidslijst van de lagere school valt immers een stamboom te destilleren, maar dat wil nog niet zeggen dat die ook wat zinnigs meedeelt.

De stamboomreconstructiesoftware kent weliswaar statistische trucs om te schatten hoe sterk het

'stamboosignaal' in de gegevens is, zegt Dunn. „Maar uiteindelijk belangrijker voor de interpretatie was de observatie dat de hoofdtakken van de stamboom overeenkwamen met de verschillende eilandengroepen.” Zo werden de talen van één tak gesproken op de Solomoneilanden, die van een andere tak op de Bismarck-archipel, die van een volgende tak op Bougainville, en die van nog een uitloper op de Louisiade-archipel.

Ofwel: de stamboom komt overeen met het beeld van de sprekers van één oertaal die zich verspreiden, waarna de talen per regio verder evolueerden en splitsen. Net zoals in Europa Germaanse talen in het Noorden worden gesproken, Romaanse in het Zuiden en Slavische in het Oosten. „De boom was ook geografisch plausibel”, zegt Dunn. De takken pasten als puzzelstukjes op de kaart.

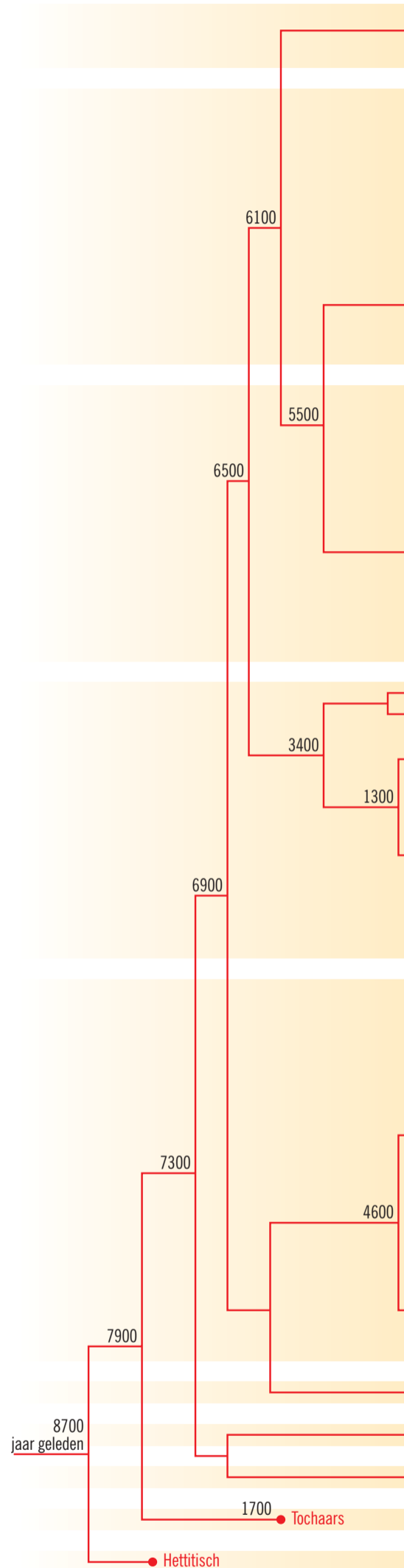
DAARMEE lijkt het erop dat onderzoekers voor het eerst een taalverwantschap voorbij de woord-erosiegrens van 10.000 jaar hebben laten zien. En dat klopt ook met wat er archeologisch bekend is over het gebied. Vanaf minstens 35.000 jaar geleden is Zuid-Oost-Azië en Australië gekoloniseerd door de mens, terwijl 4.000 jaar geleden een nieuwe immigratiegolf volgde van de sprekers van Austronesische talen, die nu de Papoeatalen omringen. Vermoedelijk zijn Papoeatalen een taalkundig overblijfsel van de eerste invasie, die ver vóór de tweede invasie al opgesplitst was in verschillende talen.

Béatrice Lamiroy, hoogleraar vergelijkende taalkunde aan de KU Leuven, vindt het onderzoek van Dunn veelbelovend, en de uitkomst 'plausibel', al vindt ze dat de methode ook op andere, liefst grotere, taalfamilies getest moet worden voor de uiteindelijke bruikbaarheid, „maar het idee erachter lijkt me te kloppen.” Interessant vindt Lamiroy ook dat Dunn en collega's grammaticale kenmerken van de taal gebruiken, en niet de sneller veranderende woordenschat. „Grammatica is het skelet van de taal, die blijft langer stabiel.”

Wel is de vraag hoe ver de blik nu verder reikt. Zicht op de bijna mythische oertaal aller oertalen biedt de methode ongetwijfeld niet. „We weten niet óf, en zo ja wanneer, die ooit gesproken werd”, zegt Dunn, „maar ook het signaal van deze methoden verdwijnt hoogstwaarschijnlijk lang voor die tijd.”

TAALPUZZELSTUKJES VALLEN OP HUN PLAATS

INDO-EUROPEES DEEL VAN DE T

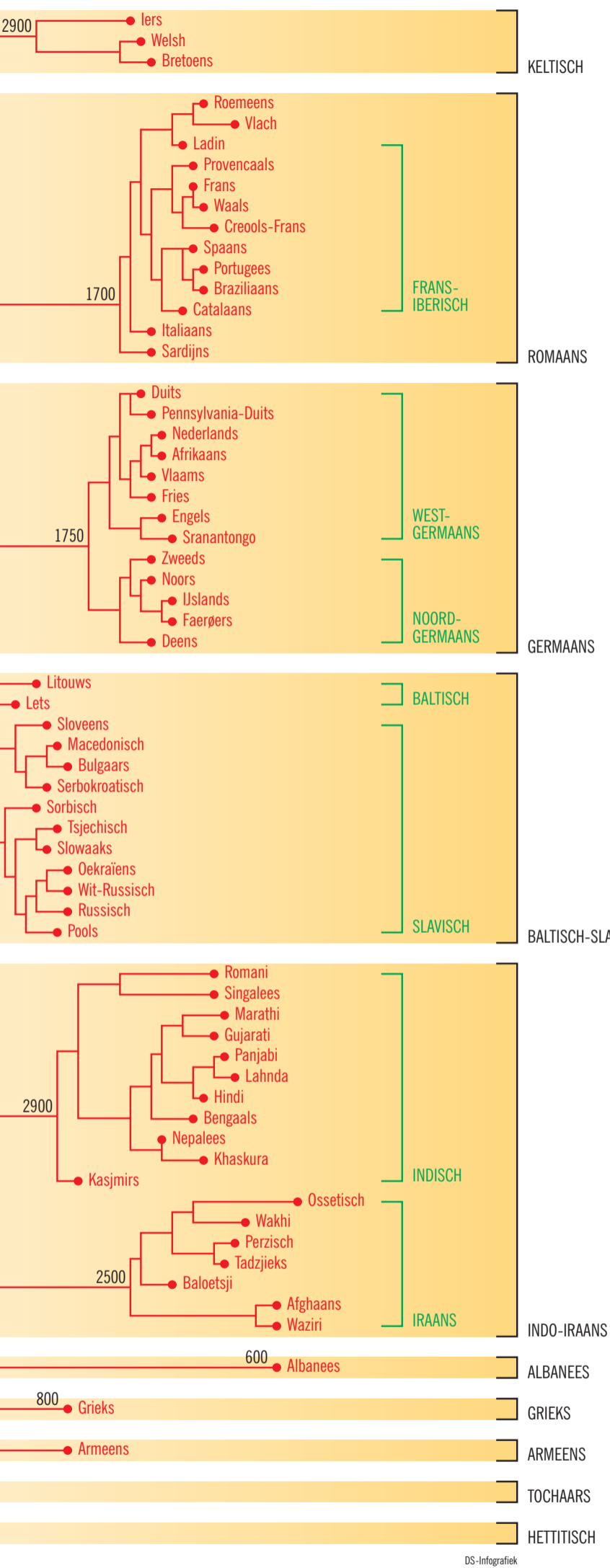


www.relatiegids.be

het klikt meteen



TALBOOM



DS-Infografiek

Taalkunde en biologie: een huwelijk met aanlooppromen

„Ik moet zeggen dat ik ook extreem sceptisch was”, zegt taalkundige Michael Dunn over evolutiebiologische methoden voor taalkunde, die de afgelopen jaren veel controverse hebben opgeroepen onder taalkundigen. ‘Junk science’, waanwetenschap, was aanvankelijk hun oordeel, dat nu langzaam wat milder lijkt te worden.

De overeenkomsten lijken zo voor de hand te liggen: talen zijn net biologische soorten. Ze kunnen levend zijn en sterven, ze kunnen over de aarde zwerven, uitwisselingen aangaan met andere talen, zelfs kinderen krijgen en hele families vormen. De parallel is zelfs nog gedetailleerder: woorden *muteren*, net als stukjes DNA dat doen. Talen kunnen woorden uit andere talen overnemen, zoals bij seksuele voortplanting genen gemengd worden.

Al jaren ontwikkelen evolutiebiologen computermethoden om uit de DNA-gegevens van verschillende soorten weer de oorspronkelijke evolutiestambomen te herleiden. *Cladistische* methoden, heten zulke rekenmethoden, een bloeiende tak van de bio-informatica, met grote bijdragen aan het begrip van de evolutie.

Russell Gray, een Nieuw-Zeelands bioloog die onderzoek deed aan de evolutie van pinguïns, liefhebberde ook in het Maori, de oorspronkelijke taal van zijn land en lid van de Austronesische taalfamilie. Gray raakte gefascineerd door de overeenkomsten tussen de disciplines en besloot zijn arsenaal ook op talen uit te proberen. „Daar was al langer over gesproken, maar het was lastig om het ook echt te doen.”

In 2000 kwam hij met een eerste concreet resultaat: een stamboom van de Austronesische talen, geproduceerd met een biologische ‘zuinigste’ boommethode.

In tegenstelling tot Dunn baseerde Gray zich op verwantschappen tussen woorden in Austronesische talen, zoals beoordeeld door taalkundig historici. In Grays optimale, zuinigste stambomen, was het aantal keren dat een woord vervangen werd door een niet-verwant woord minimaal. Behalve op details klopte

deze stamboom heel aardig met die waar de Austronesisch-deskundigen het al over eens waren, dus inhoudelijk was de uitkomst niet revolutionair. Ook een cladistische stamboom van de Bantoe-talen, door de Britse Clare Holden, schopte geen heilige huisjes om.

Meer stof deed een publicatie in 2003 over de Keltische talen van de bioloog Peter Forster en Alfred Toth uit Cambridge in *The Proceedings of the National Academy of Sciences* opwaaien. Forster maakte geen gebruik van taalkundige input, gebruikte zijn eigen lekenoordeel over de verwantschap van woorden, en maakte veel fouten bij de behandeling van zijn basisgegevens. Taalkundigen, onder wie de Amerikaanse Pennsylvania University, kraakten zijn artikel. „Waardeeloos”, was daarin een van de vriendelijker termen.

„Taalkundigen kunnen nogal fel zijn”, lacht Dunn. Bovendien helpt het niet dat het gebied nogal gevoelig is voor wilde theorieën van amateurtaalkundigen. Tot verdere irritatie pochten Forster en Toth in een krantenartikel dat taalkundigen „niets van hun methode begrepen.”

Onder dit ongunstig gesternte verscheen in *Nature* een half jaar later Grays volgende analyse, over de Indo-Europese taalfamilie, met gebruikmaking van de nieuwste cladistische technieken van de evolutionair biologen. Het resultaat was een hele set stambomen, allemaal taalkundig plausibel (*een voorbeeld ziet u hiernaast*).

De controversiële angel was deze keer Grays nieuwe datering, gebruikmakend van ijkpunten langs de boom, zoals de opsplitsing van het Germaans (de eerste duidelijk Noord-Germaanse inscripties dateren van het Scandinavië van de derde eeuw) of de afsplitsing van het oud-Grieks (de eerste Myceense teksten zijn van 1.500 voor Christus). Door die data op de stamboom te passen kwam Gray uit op een ouderdom van 8.700 jaar voor het Proto-Indo-Europees, in plaats van de geaccepteerde 6.000 jaar.

Die laatste datering klopt met de

geaccepteerde hypothese dat de originele Indo-Europeanen dezelfde zijn als het Koergan-ruitervolk van de steppen boven de Zwarte Zee, die met de nieuwe technologie van paard en wagen kans zagen hun taal uiteindelijk over heel Europa te verspreiden. Maar 8.700 jaar klopt veel beter met een concurrerende theorie, die de Indo-Europese oorsprong plaatst in het huidige Anatolië in Zuid-Turkije, dicht bij de ‘vruchtbare halve maan’, het gebied waar ongeveer in die tijd de landbouw uitgevonden werd. Dan zou het Indo-Europees zijn wijde verspreiding uiteindelijk te danken hebben aan boeren, net als overigens ook de Austronesische en de Afrikaanse Bantoe-talen.

Opnieuw waren de taalkundevakbladen te klein. „Volstrekt belachelijk”, vindt Indo-Europeanist Aleksander Lubotsky van de Leidse universiteit de datering — en de methode — nog altijd. Taalverandering, in wezen een sociaal proces, is volgens Lubotsky niet te vergelijken met mutaties in DNA, een chemisch-fysisch proces. Volgens hem zijn de ijkpunten veel te ruim, en daarnaast — saillant detail — kende de Indo-Europese oertaal een woord voor ‘wiel’, dat pas rond 5.500 jaar geleden is uitgevonden. Volgens Gray is het echter best denkbaar dat het woord voor wiel zich samen met de technologie verspreid heeft, of is afgeleid van een oudere stam die ‘draaien’ betekent.

Maar ondanks de dateringskwestie lijkt de stemming onder taalkundigen aan het omslaan, merkt Gray, „aanvankelijk was communiceren hierover echt een nachtmerrie. Nu luisteren mensen tenminste, en komen ze met argumenten”.

Béatrice Lamiroy, taalkundige aan de KU Leuven, vindt de kruisbestuiving tussen taalkunde en biologie waardevol. „De parallel tussen taalevolutie en biologische evolutie is trouwens al opgemerkt door Charles Darwin. Heel onorthodox vind ik het dan ook niet, wel veelbelovend, al moeten we zien wat we er uiteindelijk aan hebben”, zegt Lamiroy. „Ik denk dat het ijs nu wel gebroken is”, zegt Dunn. (bvw)

Europese eenwording begon in oertijd

NIET alleen ‘vader’, ‘vater’ en ‘father’ lijken op elkaar, ook het Griekse ‘pèter’, het Sanskriet ‘pita’ en het Ierse ‘athir’ hebben er wel wat van. En inderdaad gaan deze woorden terug op één oervorm, en horen deze talen bij één familie, de Indo-Europese taalfamilie, met nog tientallen leden van Russisch tot Portugees, en van IJslands tot Afghaans en Bengaals. Al voor de geschreven geschiedenis hebben deze talen zich verspreid over Europa en grote delen van Zuid-West-Azië.

Met veel moeite hebben historisch taalkundigen door woordvormen en klanken te vergelijken de talloze klankveranderingen in kaart gebracht die tot deze diversiteit geleid hebben. Er is zelfs een reconstructie van de hypothetische oertaal gemaakt, het ‘Proto-Indo-Europees’ (PIE). Het woord voor is vader in het PIE is *ph2ter (het sterretje geeft aan dat het om een gereconstrueerde vorm gaat, de h2 is een bepaalde keelklank).

Ook tientallen andere taalfamilies

hebben inmiddels geheimen prijsgegeven aan de comparatieve methode, zoals de Austronesische talen (Indonesisch, Hawaïaans, Maori en honderden andere talen van Madagaskar tot de Stille Oceaan), de Afro-Aziatische (met Arabisch, Hebreeuws en Berber), de Oeraalse (Fins, Hongaars, Ests en vele talen in de voormalige Sovjet-Unie) of de Turkse (Turks, Oezbeeks). Sommige ‘geïsoleerde’ talen, zoals Baschisch of Japans, hebben helemaal geen familie. (bvw)

Morgen in
Jobat
Carrière-maker.

**HOE HOU JE
HET VOL AAN
DE TOP?**

Professor
Marc Buelens (UG)
over DE NIEUWE LEIDER

Ken jij de impact
van je daden?

Ontdek hoe het
kan in de reeks:

**DE MANAGER
EN ZIJN COACH**

**Focus op ICT &
ENGINEERING**

Computer
zoekt vrouw

