

## 1.14. Dialectometrische indeling van de Nederlandse dialecten

De dialectindeling in 1.12 is gebaseerd op klankgrenzen en die in 1.13 op subjectieve verwantschapsoordelen. Sindsdien heeft men geprobeerd ook een indeling te vervaardigen zuiver gebaseerd op taalkundige metingen, via de zogenoemde dialectometrische analyse. Onderzoekers Wilbert Heeringa, Peter Kleiweg, John Nerbonne en Martijn Wieling hebben de fonetische transcripties van 562 woorden zoals deze in 613 verschillende Nederlandse en Vlaamse plaatsen worden uitgesproken met elkaar vergeleken. Ze hebben gebruik gemaakt van materiaal dat binnen het GTRP-project (zie de Algemene inleiding) werd verzameld.

De aanpak is altijd gebaseerd op de vergelijking tussen twee uitspraken van hetzelfde woord, waar men het verschil in uitspraak met de zogenoemde Levenshtein-afstand meet. Als de reeksen dezelfde lengte hebben, dan kan men ze boven elkaar zetten en simpelweg tellen op hoeveel plaatsen ze verschillen:

Gronings	[son] en
Haarlems	[zon] ...verschillen op één plek en hebben dus een afstand van één.

De Levenshtein-afstand houdt verder rekening met de mogelijkheid dat geluiden kunnen worden toegevoegd dan wel verloren kunnen gaan. Dat ziet men het eenvoudigst aan de hand van een voorbeeld van twee uitspraken van het woord *melk*:

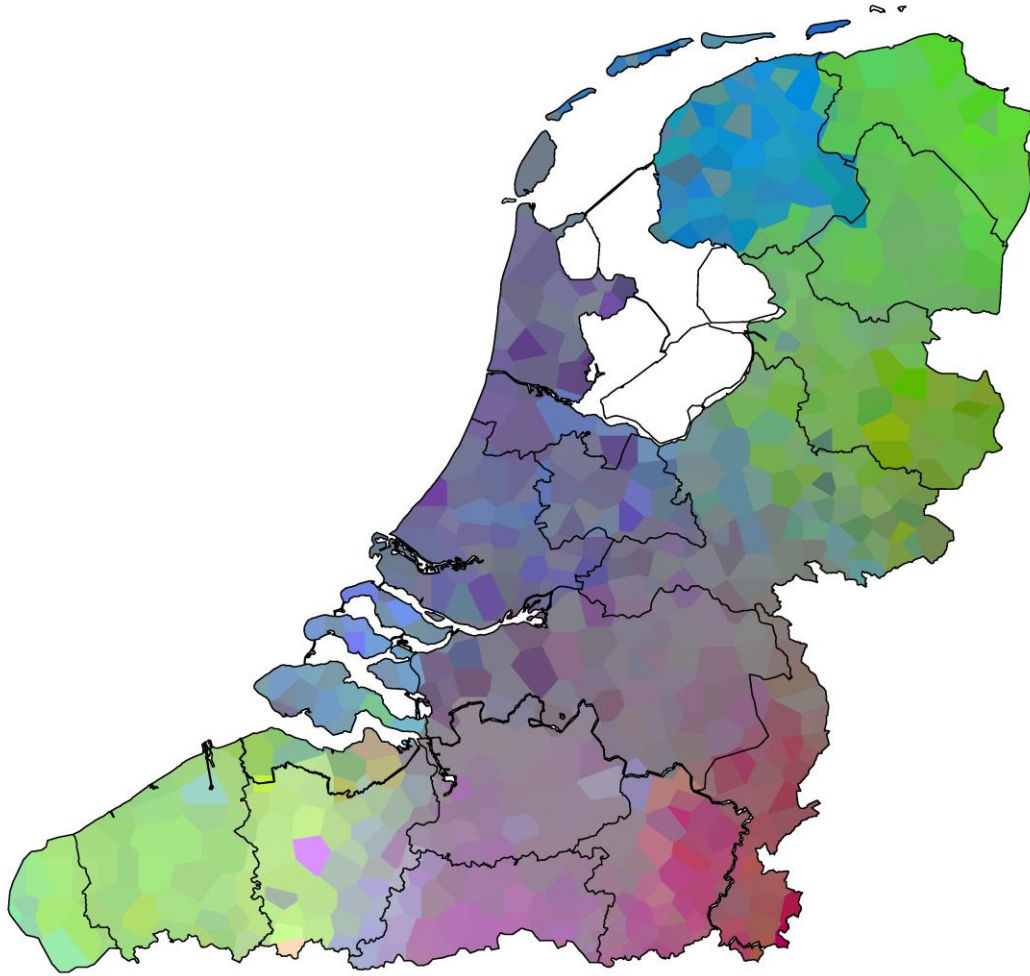
Fries (Grouw)	[mòlkə], met ə voor de stomme e in bv. <i>de</i>
Haarlem	[mèlək]

De Levenshtein-afstand vergelijkt deze uitspraken door de vervanging van [ò] door [è], het verlies van de [ə] aan het einde van [mòlkə] en de toevoeging van de [ə] tussen [l] en [k] alle als één verschil te zien. Dit geeft een totale afstand van drie en staat gelijk aan het aantal handelingen dat nodig is om de Friese uitspraak te veranderen in de Haarlemse uitspraak.

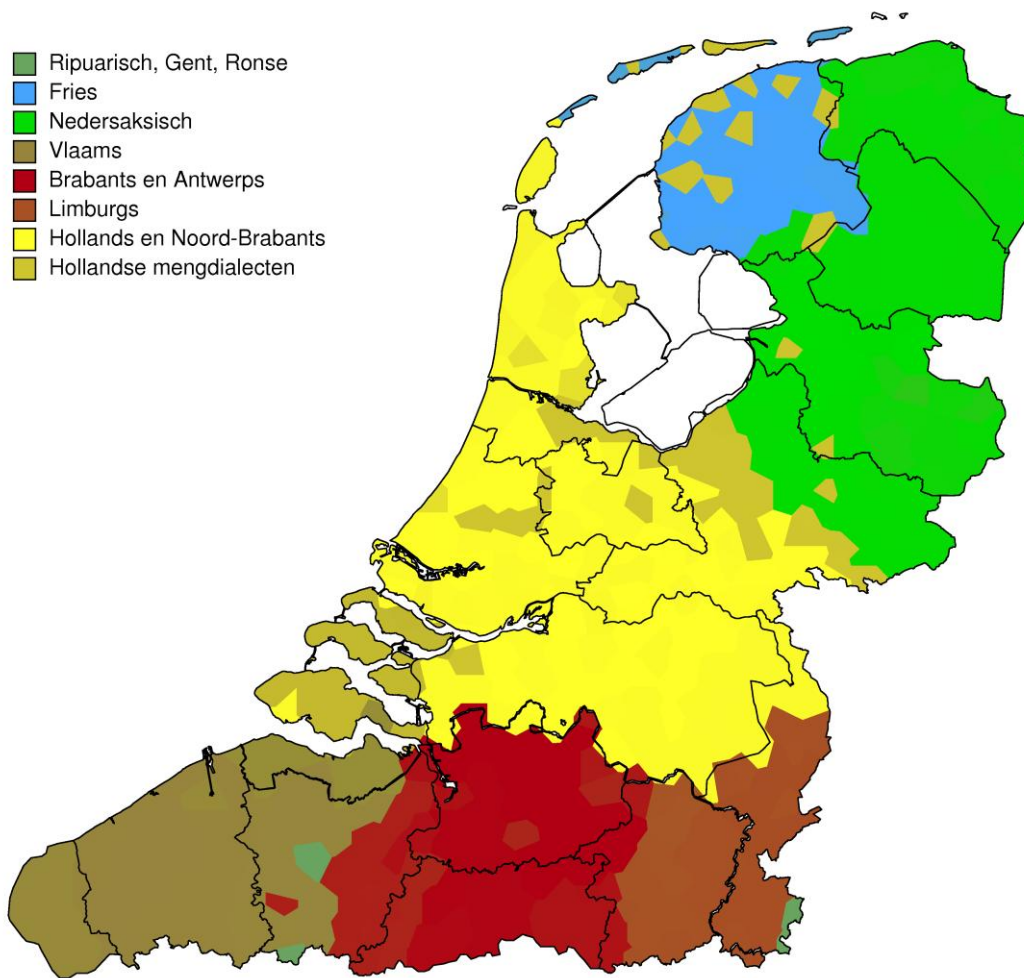
Dezelfde procedure wordt voor alle 562 woordparen uitgevoerd om op die manier voor ieder woordpaar de afstand te bepalen. Hiervan wordt vervolgens het gemiddelde berekend om zo het verschil tussen het Fries in Grouw en het Nederlands in Haarlem te vinden. Er zijn ook verfijningen van de Levenshtein-afstand waar veel voorkomende vervangingen minder zwaar wegen, maar daar gaan we hier verder niet op in. Op basis van de gemiddelde Levenshtein-afstanden tussen de 613 dialecten (in totaal 187.578 plaatsparen, en dus ruim 100 miljoen woordvergelijkingen!) verkrijgt men een vrij veelomvattend beeld van de dialect-verhoudingen.

Wanneer men nu via een statistische methode op basis van de onderlinge afstanden de dialecten in drie dimensies plaatst, kan men kleuren toewijzen op basis van deze drie coördinaten. Taalkundig vergelijkbare plaatsen krijgen dan vergelijkbare kleuren. Het resultaat is te zien in kaart 1.14a. Omdat de meeste dialecten maar weinig verschillen van de omliggende dialecten, verschillen de kleuren in een klein gebied weinig van elkaar. De kaart doet recht aan de gedachte dat sprake is van een *dialectcontinuüm*.

Het is ook mogelijk om groepen of ‘clusters’ vergelijkbare dialecten te zoeken op basis van de onderlinge dialectafstanden. Dialecten met onderling kleine afstanden worden in dezelfde groep geplaatst en dialecten met onderling grote afstanden in een verschillende groep. De groepen worden in kaart 1.14b getoond. De kaart laat ruwweg een verdeling in acht groepen zien, waar kleine kleurverschillen binnen een groep laten zien dat verdere opdeling mogelijk is.



Kaart 1.14.a. Visualisatie van het Nederlandse dialectlandschap als dialectcontinuüm op basis van variatie in de klankcomponenten.



Kaart 1.14b. Indeling van het Nederlandse dialectlandschap in acht groepen.