

Hoorzitting Spelling, Vlaams Parlement, maandag 20 februari 2006.

Stelling: De commotie rond de spellingaanpassing wordt gevoed door een fundamenteel ontbreken van een adequate spellingcontrole bij de Media.

De Media publiceren artikelen in hun kranten en op hun web sites waarin zeer frequent spellingfouten voorkomen. De meeste van deze fouten zijn niet gerelateerd aan de spellingaanpassing van oktober 2005. Een deel is terug te voeren op een verschil in denken over spellingregels tussen belangrijke categorieën (Van Dale versus het Groene Boekje, 1995).

We kunnen het de Media echter niet geheel kwalijk nemen dat er zoveel foutieve spellingwijzen in hun publicaties voorkomen. We moeten daarvoor analyseren op welke wijze men de spelling controleert.

Spellingcheckers die met de redactionele opmaaksystemen worden meegeleverd, zijn meestal niet in staat om op een juiste wijze spellingfouten vast te stellen. Deze spellingcheckers keuren een substantieel deel van de fouten in teksten goed. Naarmate de lengte van woorden toeneemt, neemt ook de kans op het accepteren van fouten toe.

Het is voor de meeste ontwikkelaars van spellingcheckers onmogelijk een lexicon te ontwikkelen dat het merendeel van de Nederlandse woordenschat dekt. In plaats daarvan stelt men een beperkt woordenboek samen met woordstammen (lemmata). Men creëert kunstmatig woorden door combinaties van deze stammen te maken. In de wiskunde noemt men deze combinaties permutaties. Veel van deze permutaties zijn in feite niet-bestaande of foutief gespelde woorden. De grote hoeveelheid combinaties, die de techniek van permutaties oplevert, maakt dat veel spellingfouten niet worden opgemerkt.

Uit eerdere — door ons uitgevoerde — testen bleek dat 40 tot 50 % van de woorden met verkeerde tussenletters of andere kleine vervormingen door de meeste van de spellingcheckers werden goedgekeurd. Voor het Nederlands valt te denken aan woorden als: *mistanden*, *bolletjeslikker*, *briefbus*, *straatgids*, *rechtzaak*, etc.

De spellingaanpassing van oktober 2005 heeft een nieuwe categorie problemen voor spellingcheckers veroorzaakt, zeker wat betreft de mate van voorkomen: de spelling is over de woordgrens gegaan. Deze categorie woorden bevat nu een spatie: *24 uursdoktersdienst*, *18 holesgolfbaan*, etc. Deze combinaties worden door spellingcheckers in stukken gehakt en niet in hun samenhang op spelling gecontroleerd. Vroeger kwam dit ook veel voor. Een fout als *gebruik maken* werd dan overgeslagen.

*TALO heeft een nieuwe generatie spellingalgoritmen en bijbehorende processen ontwikkeld om bovenstaande fouten wèl te kunnen detecteren. TALO is bezig met ontwikkelaars binnen de grafische sector de geschikte taaltechnologie op te zetten. Wij hebben echter vaak te maken met redactionele systemen die niet over die functies beschikken om volgens passende normen taal te behandelen. Op grond van geleverde methodieken van de fabrikanten is het onmogelijk meer dan één woord te selecteren. Ontwikkelaars moeten andere methoden vinden om een adequate spellingchecker toe te kunnen voegen aan zo'n redactioneel systeem. Het probleem ligt bij de aangeboden interfacetechniek van fabrikanten.

De Taalunie heeft er nooit eerder aan gedacht dit aspect van spellingcontrole bij haar oordeel te betrekken, maar uit gesprekken blijkt dat men zich bewust is geworden hoe belemmerend de functionaliteit van spellingcheckers kan zijn voor correct taalgebruik.

*TALO levert een bijdrage aan de spellingcontroletechnologie, mede gebaseerd op ervaringen met de Duitse spellinghervorming van 1996.

Dr. J.C. Woestenburg, *TALO b.v.