

Spelspiek



*Michel Boekestein (INL)
boekestein@inl.nl*

Spelspiek

RU | STEVIN

Spelspiek is...

Een spellingshulpmiddel, beschikbaar via:

- MSN



info@spelspiek.nl

- SMS



SPELLING woord 3669

- Webbrowser



www.spelspiek.nl

Spelspiek

- Demonstratieproject STEVIN
(www.stevin-tst.org)

- Gemaakt door



MSN



Spelspiek

ru | STEVIN

Browser

Spelspiek Uitleg - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.spelspiek.nl/ Google

Spelspiek

Zoek hier naar de **spelling** van een bepaald woord.

[Uitleg](#) [Blog](#)

hoe schrijf je bjoetiekees

gezochte woord: bjoetiekees

Je schrijft niet 'bjoetiekees', maar 'beautycase'!

taagebruiker neert wel eens zijn twijreis over de manier waarop je een bepaaid woord moet schrijven.

Er bestaan al verschillende kanalen via welke je de spelling van woorden kunt opzoeken, of de officiële regels van de spelling van de Nederlandse taal kunt

SMS

MSN

(INL) INSTITUUT VOOR NEDERLANDSE LINGUISTIEK

Polderland LANGUAGE AND SPEECH TECHNOLOGY

van Dale

elitech

rû | STEVIN

Spelspiek

rû | STEVIN

Spelspiek is niet ...

- Een on-line spellingscorrectieprogramma
- Een volledig interactief dialoogsysteem
- Een grammatica-correctieprogramma

Voordelen van Spelspiek

- Laagdrempelige toegang via MSN, sms, browser
- Begrijpt spellingvragen in natuurlijke taal
- Biedt intelligente herkenning van spelfouten
- Herkent ambigue woorden
- Geeft betekenisinformatie waar nodig
- Gebruikers kunnen feedback geven (en we doen er iets mee)
- Bij onbekende woorden en onbekende spelfouten wordt deskundige menselijke hulp ingeschakeld
- ... en daardoor een continu verbeterende kwaliteit en dekkinggraad

Spelspiek is bedoeld voor ...

Iedereen die geïnteresseerd is in / moeite heeft met de officiële spelling. Met name:

- Opgroeiende kinderen en jongeren
- Professionele taalgebruikers, zoals tekstschrijvers en vertalers

Spellingcorrectie

Vergelijk woord met ijklijst. Nadelen?

- dekking (te lage dekking nadelig, maar te hoge ook!):
 - OOV's vs. verwarring (homofonen);
- het Nederlands telt "oneindig" veel woorden: eenpersoonssportvliegtuigje
- "*wild u effen wachtte op de doctor*";
heeftten; wort; gelt

Spellingcorrectie

Dus: OOV-probleem kun je niet oplossen. Wat voor *mechanismen* kan men inzetten om:

- Goed gespelde woorden goed te rekenen, en
- incorrect gespelde woorden fout te rekenen?

Maar ook:

De precisie daalt met deze mechanismen

- Samenstellingsanalyse
- Spelfoutenvoorspeller (bv. regelsysteem)
- Leren vanuit de praktijk

Samenstellingsanalyse

recht[s] + bijstand + verzekering

Let op de precisie: *hoogtenmeter* is niet *hoogten + meter!*

- oplossingen:
 - kijken naar andere, wel bekende samenstellingen met *hoogte*
 - beredeneren: *hoogte* heeft twee meervouden, dus geen tussen-n.

doorde = waarschijnlijk *door de*

- oplossing:
 - kijk naar woordklasse van de samenstellende delen

Grens bereikt: *blindedarm* vs. *blindenbegeleider*

informatie aan de gebruiker. Bv. met voorbeelden:

- "kies uit het type
 - 'blinde + rest': *blindedarmontsteking*, ... of:
 - 'blinden + rest': *blindenstok*, *blindenschool*, *blindenschrift*, ..."

Leren uit de praktijk

- Beruchte fouten opnemen (*onmiddelijk, hondehok, ballustrade*)
- Veranderingen t.o.v. spelling '95 opnemen: (*aëro-... => aero...; aids-... => aids...; bohémien => bohemien; skister => skiester; slangekruid => slangenkruid*)
- Laat de spellingcorrector fouten (+ de correcties) opnemen die hij in de praktijk tegenkomt. Handwerk.
- Automatische leeralgoritmen

Regels

- *pyjama* $p[yj|i|iei|iey|ii|iy|y|ij][a|aa]m[a|aa]$

Geeft 32 pyjama's:

*pyjama, pyjamaa, pyjaama, pyjaamaa, piama,
piamaa, piaama, piaamaa, pieiama, pieiamaa,
pieiaama, pieiaamaa, pieyama, pieyamaa,
pieyaama, pieyaamaa, piama, piamaa,
piaanama, piaanamaa, piyama, piyamaa, piyaama,
piyaamaa, pyama, pyamaa, pyaama, pyaamaa,
pijama, pijamaa, pijaama, pijaamaa*

Informatie naar gebruiker

- homofonen (wort, word, wordt) + info:

flushes

flushes /meervoud van flush ; pokerterm#

flushjes /meervoudig verkleinwoord van
flush ; pokerterm#

flusjes /straks#

flusjes /zo-even

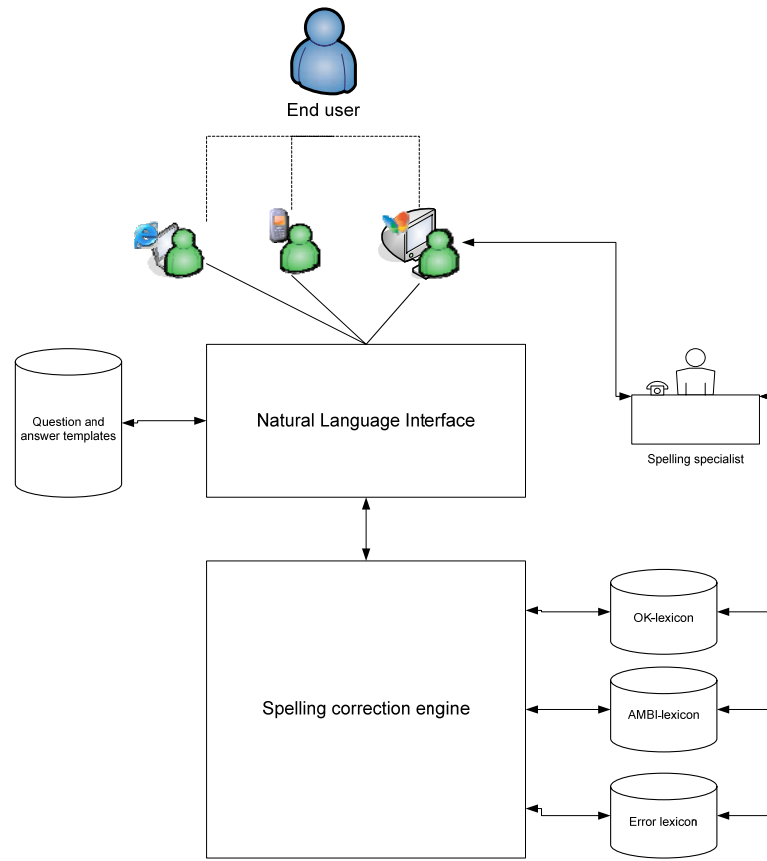
- OK, onbekend, bekende fout

Grenzen/moeilijkheden

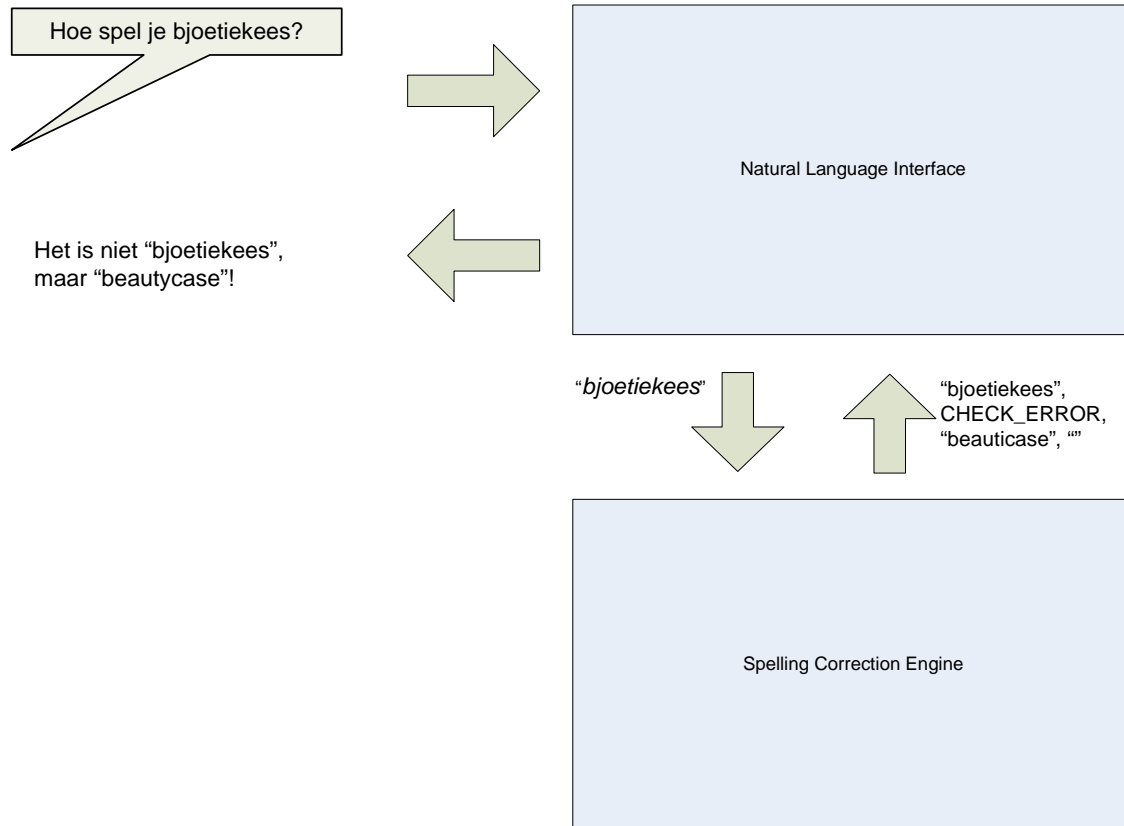
- *Meerwoordige uitdrukkingen*
- *grammatica (herkend/herkent)*

- *Verbeteringen aan herkenning*
- *Vervolledigen informatie aan gebruiker*
- *Toekomst*

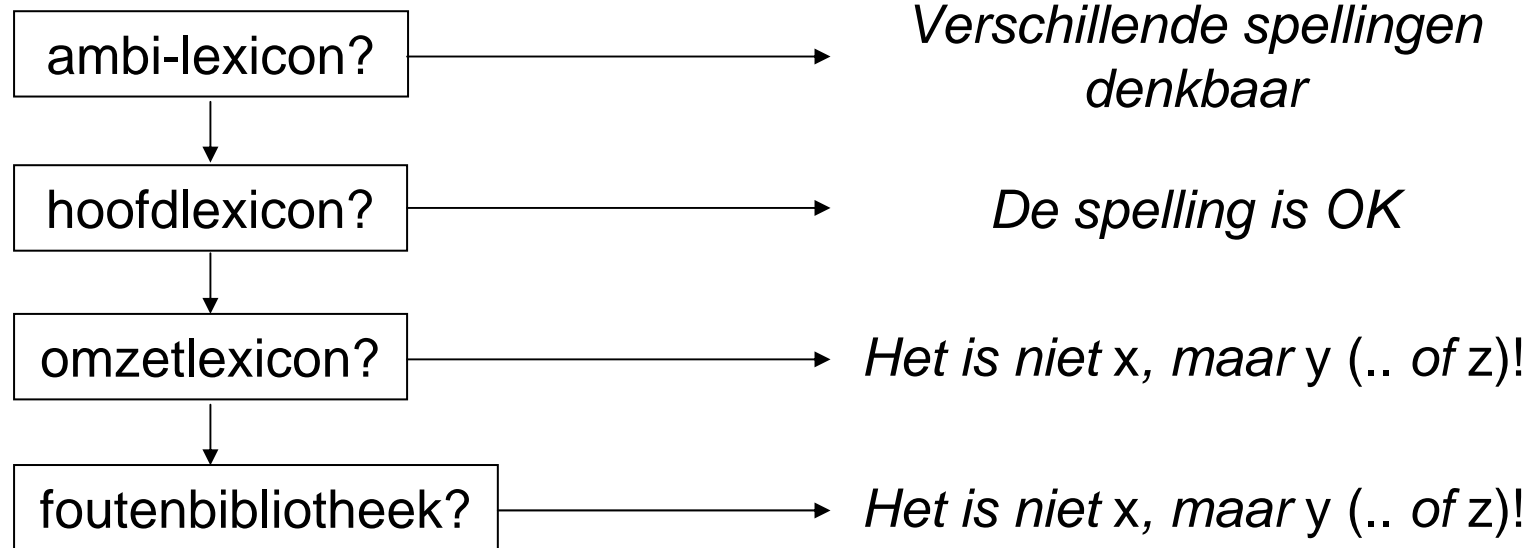
Hoe werkt Spelspiek?



De natuurlijke taal - interface



Volgorde omspeller



Uit de praktijk

- Ongeveer 12000 *verschillende* woorden gevraagd sinds juli 2007
- Top 10: pannenkoek, paddestoel, paddenstoel, interessant, paardebloem, onmiddelijk, email, online, produkt, liefde
- Ongeveer 300 mails in de laatste 3 maanden
- Waarnaartoe, motormanagement, moeraswolfsspinn, Dode Zeerollen, geluidsniveaueis

Probeer het zelf!

- MSN: add user *info@spelspiek.nl*
- SMS: *3669 SPELLING <word>*
- Web: Visit *www.spelspiek.nl*