

Panel on Lexical Units

Compounds as words

Piek Vossen, Irion Technologies & Vrije Universiteit
Amsterdam

Global Wordnet Conference
Szeged, Hungary, 22-25 Jan 2008

Are all compounds words?

Are all compounds types?

- water = H₂O and is a type of substance
 - {**thewater**}NL ('tea water' => for making tea)
 - {**koffiewater**}NL ('coffee water' => for making coffee)
 - {**bluswater**}NL (=> for extinguishing fire)
- Many compounds refer to sets of types in specific circumstances;
 - Do not represent new disjunct types
 - Are dependent on these types

Water used for a purpose

- theewater (*for making tea*)
- koffiewater (*for making coffee*)
- bluswater (*for extinguishing fire*)
- scheerwater (*for shaving*)
- afwaswater (*for cleaning dishes*)
- waswater (*for washing*)
- badwater (*for bathing*)
- koelwater (*for cooling*)
- spoelwater (*for flushing*)
- drinkwater (*for drinking*)

Water occurring somewhere or originating from

- putwater (*in a well*)
- slootwater (*in a ditch*)
- oppervlaktewater (*at the surface*)
- welwater (*out of a spring*)
- leidingwater, gemeentepils, kraanwater (*out of the tap*)
- gootwater (*in the kitchen sink or gutter*)
- grachtwater (*in a canal*)
- kwelwater (*coming from underneath a dike*)
- grondwater, grondwater (*in the ground*)
- buiswater (*on a ship*)

Being the result of a process

- pompwater (*being pumped away*)
- smeltwater, dooiwater (*melting snow and ice*)
- afvalwater (*waste water*)
- condens, condensatiewater, condenswater (*from condensation*)
- lekwater (*leaking water*)
- regenwater (*rain water*)
- spuiwater (*being drained for water maintenance*)

Availability & cliché

- Compositionality is not sufficient criteria
- All combinations are possible but not all combinations occur within language-culture
- We know (have statistical intuition) that a compound is available for use in a language community
- Availability 'hinders' (mildly blocks) other constructs:
 - not using the cliché confuses people
 - using the cliché confirms interpretation (releases the stereotypes and associations)
- Language generation point of view

Non-systematicity of compounds

muziek (music) (89,300,000)		muziekinstrument (musical instrument)	muzikant (musician)
klassieke muziek (classical music) (1,130,000)	<i>gitaarmuziek</i> (<i>guitar music</i>) (35,300)	drums	gitarist (guitarist) (1,400,000)
popmuziek (pop music) (515,000)	<i>pianomuziek</i> (<i>piano music</i>) (33,600)	piano	pianist (piano player)
jazzmuziek (jazz music) (33,200)	<i>trompetmuziek</i> (<i>trumpet music</i>) (1,200)	trompet (trumpet)	drummer
country muziek (country music)	<i>saxmuziek</i> (<i>sax music</i>) (28)	gitaar (guitar)	jazzgitarist (9,380) (jazz guitar player)
popgitaarmuziek (pop guitar music) (0-5-6,-0)	<i>drummuziek</i> (<i>drum music</i>) (394)	saxophone	acoustische gitarist (acoustic guitar player) (94-2)
jazzgitaarmuziek (jazz guitar music) (1-8-5-2)	gitaarpopmuziek (guitar pop music) (60-6-8-4)	countrygitaar (country guitar) (189-1,390)	electrische gitarist (electric guitar player) (154)
	countrygitaarmuziek (country guitar music) (0-0-2-0)	jazzgitaar (jazz guitar) (4,730-3,410)	jazzdrummer (jazz drummer) (9,820)
klassieke gitaarmuziek (classical guitar music) (271-364)	klassieke pianomuziek (classical piano music) (647)	klassieke gitaar (classical guitar) (20,700)	country giratist (0) (country guitar player)
		electrische gitaar (electric guitar) (28,800)	
		acoustische gitaar (acoustic guitar) (8,400)	