

/instituut voor
de Nederlandse
taal/

Detecting neologisms using Neoloog

IDS, Mannheim,
4 juni 2019

Vivien Waszink
vivien.waszink@ivdnt.org

Vivien?

- linguist/researcher at Instituut voor Nederlandse Taal (INT), Leiden, The Netherlands (Dutch Language Institute)
- lexicographer *Algemeen Nederlands Woordenboek* (ANW) and *Neologismenwoordenboek* (NW)
- ... today: Neoloog and NW (dictionary of neologisms)

New?

‘What we designate as “new words” are rarely in fact new in the strictest sense New words are pieced together from the fragments or the whole parts of old ones’ (Orin Hargraves)

Neologisms in Dutch

- mostly compounds: *toetsenbordkrijger* ‘desktop warrior’ and *batterijstress* ‘stress about the battery of an electric vehicle being low’
- sometimes derivations: *ontspullen*, *ontrommelen* ‘to declutter’
- sometimes multi-word units: *topless vergaderen* ‘to be in a meeting without using laptops, desktops, etc.’

Special kind of compounds: portmanteau words

glamping ‘glamorous camping’

brotox ‘botox for “bros”, for men’

hangry ‘hungry and angry’

infobesitas ‘overwhelming amount of information, especially on social media’

cyberchonder ‘someone who is excessively worried about having a serious illness and who is looking for symptoms online’

Words with new meanings and new loanwords: neologisms as well

- already existing words with a new meaning: *slim* ‘smart’; *eerlijk* ‘fair’

- new loanwords:

plogging: portmanteau-word of Swedish *plocka* ‘to pick’ and English *jogging*
frosecco ‘prosecco mixed with ice cubes; frozen prosecco’

thighbrow ‘skin roll created at the top and front of the thigh while someone is bending forward, sitting or kneeling’.

(The pose in which this ‘thighbrow’ is created is a popular pose in pictures on social media)

... but also more familiar words like *app*, *selfie*, etc.



Algemeen Nederlands Woordenboek (ANW)

- online, corpus-based dictionary
- describes contemporary Dutch from 1970 onwards
- describes Dutch as it is used in the Netherlands and in Vlaanderen ‘Flanders’, the Dutch-speaking region in Belgium
- the chronological successor of the *Woordenboek der Nederlandsche Taal* (WNT)
- lots of **neologisms** in the ANW
- ANW: a word is considered a neologism if it was coined in or after the year 2000
- a neologism is only added to the ANW if it is rooted in the lexicon, not if it is ephemeral

... but then you'll miss ...

- interesting morphological information and procedures of word coinage: *-gate*, *-syndroom/-complex*
ont-, *hypes/crazes* ending in *-ing*, etc.
- new composite forms coined in analogy with older ones:
hangoudere (*hangjongere*); *bore-out* (burn-out),
nobudgetfilm (lowbudgetfilm)
- neologisms often represent a certain era: economic crisis; the formation of a new government; the debate on racism; the debate on climate-change etc
- therefore: a special dictionary of neologisms:
Neologismenwoordenboek (NW)
- (... and also because visitors of ivdnt.org and users of the ANW seem to be interested in neologisms in particular)

Neologismenwoordenboek (NW)

- dictionary of neologisms
- both rooted neologisms and ephemeral ones
- detecting neologisms in newspapers using Neoloog
- test version available

Front page

- A
- B
- C
- D
- E
- F
- G
- H
- I
- J
- K
- L
- M
- N
- O
- P
- Q
- R
- S
- T
- U
- V
- W
- X
- Y
- Z

Info

Informatie over een woord Woorden met bepaalde kenmerken Beschrijf een woord

Zoek een woord of uitdrukking

Bijvoorbeeld: paard of heilige koe. Jokertekens werken ook: zoek *achtig om alle woorden te vinden die eindigen op 'achtig'. [Meer informatie »](#)

plasles

Elke week een beschrijving van een nieuw Nederlands woord in de webrubriek *Neologisme van de week*.

broodtekst

Antwoorden op vragen over woorden in de webrubriek *Woordbaak*.

oktember

Woorden of betekenissen die ooit in het Nederlands bestonden maar nu verdwenen zijn in de webrubriek *Terug in de taal*.

nieuwsbrief

Alles weten over het Nederlands? Schrijf je in voor de gratis digitale *nieuwsbrief* van het Instituut voor de Nederlandse Taal.

Design and presentation

- alphabetically ordered lemma list
- search box
- advanced search: words with specific properties/labels
- describe a word
- design in line with the design of ivdnt.org

Words with specific properties/labels

Informatie over een woord

Woorden met bepaalde kenmerken

Beschrijf een woord

Zoek woorden met bepaalde kenmerken

Ik zoek woorden die...

... vooral gebruikt worden in ✕

... zijn ✕

+ Nog een kenmerk

🔍 Zoek

describe a word

Informatie over een woord

Woorden met bepaalde kenmerken

Beschrijf een woord

Zoek een woord aan de hand van steekwoorden

klein huis duurzaam



1 resultaat gevonden.

Sorteer op: relevantie ▼

Toon: 20 resultaten ▼

Opslaan als...

tiny house

1.0 klein duurzaam huis

is vaak zelfvoorzienend; is een huis; klein woonhuis, soms op wielen, dat gebouwd ...

story

- each lemma starts with a short story about the neologism concerned
- in this story the word is defined, its etymology and morphology are elaborated, etc.
(comparable to Neologisme van de week (ivdnt.org))

Neologism of the week

Neologisme van de week

xtc-parfum

Door Vivien Waszink

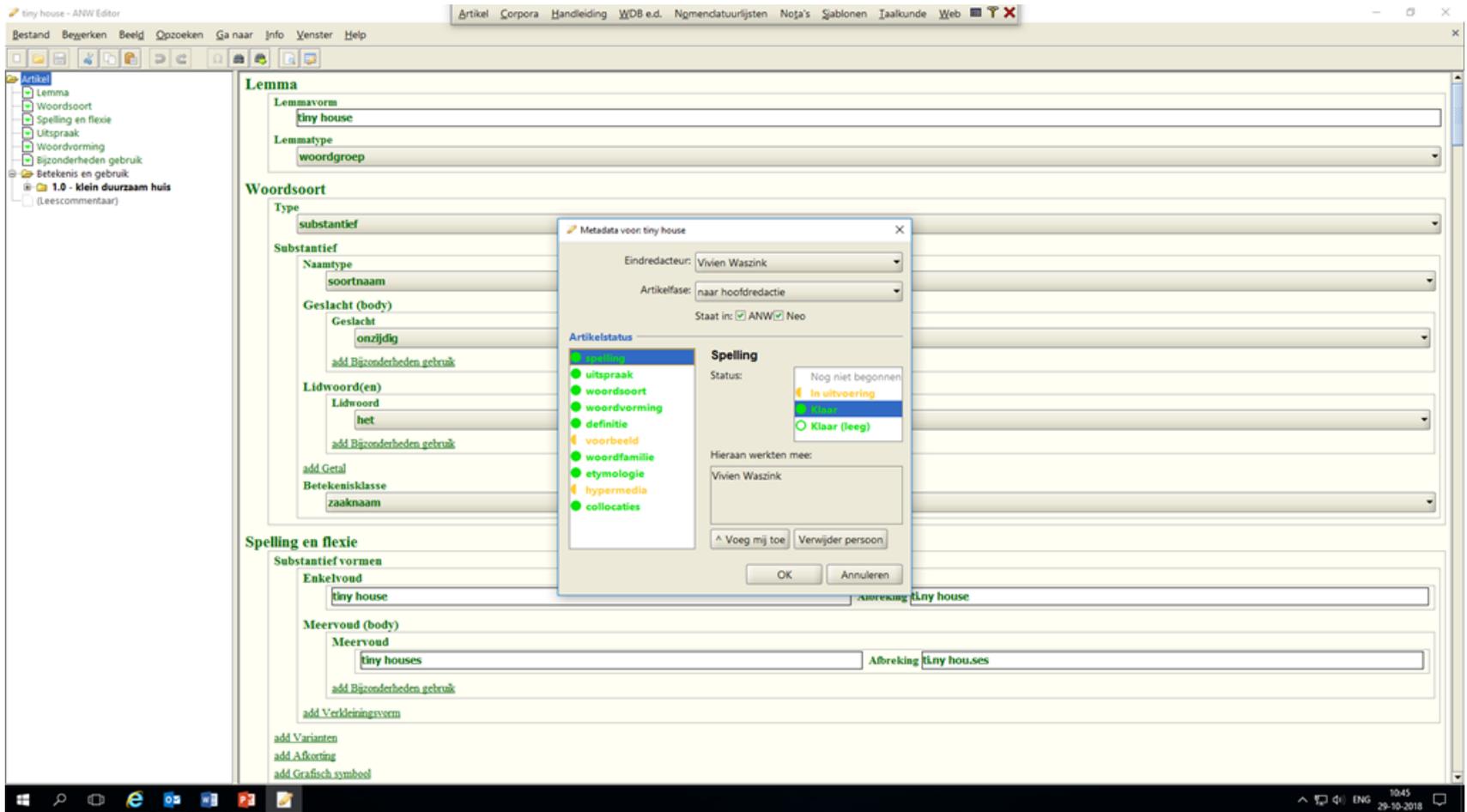
Speciaal parfum dat naar xtc ruikt en dat ervoor moet zorgen dat mensen bekend worden met de geur van xtc, zodat ze het in de gaten hebben als er een drugslab in hun buurt is waar xtc ontwikkeld wordt. Het lanceren van dit speciale parfum, genaamd XTACY, is onderdeel van een bewustwordingscampagne die is opgezet door twee officieren van justitie uit Zwolle. De officieren ontwikkelden samen met een reclamebureau deze zogeheten 'geurcampagne' in de hoop sneller drugslabs op te sporen. "Onder het motto: eens geroken wordt later herkend", legt officier van justitie Lars Stempher uit aan de NOS.

tiny house

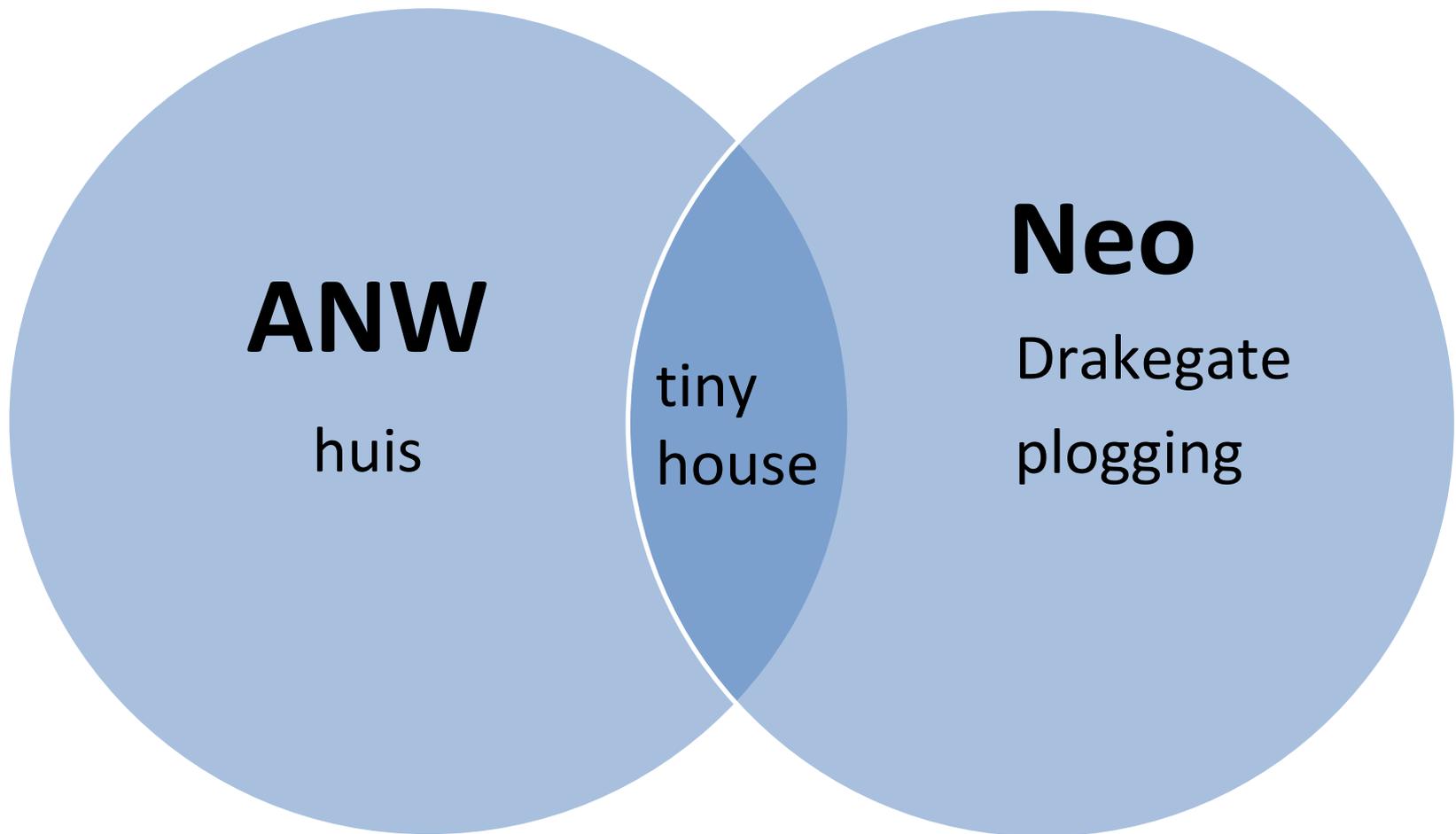
tiny house 1.0

Een tiny house is een klein huisje. Maar dan een huisje dat bewust klein is gehouden, speciaal voor mensen die genoeg willen nemen met minder leefruimte en in dat tiny house vaak ook zo bewust en duurzaam mogelijk proberen te leven. Met veel minder ruimte én veel minder spullen. In Amerika is de zogeheten 'tiny house movement' al wat langer actief: door de instortende economie en rampen als orkaan Katrina raakten veel mensen hun huizen kwijt en starters konden nog moeilijker aan een woning komen. De behoefte aan betaalbare en flexibele woonvormen ontstond en zo gingen mensen gewoon hun eigen huisjes bouwen.

Two dictionaries, one editor



So ...



Detecting neologisms through corpus analysis

- ANW-corpus
- CHN-corpus: Corpus Hedendaags Nederlands (until 2013) (contains newspaper corpus (NRC and The Standaard); using Neoloog to detect new words
- web as corpus

ANW-corpus

- over 102,5 million words
- consists of:
 - corpus of literary works
 - newspaper corpus
 - ‘domeinencorpus’: texts within specific domains of interest, for example concerning law, medicine, etc.
 - **corpus of neologisms (detected from around 20 newspapers and newsites)**

CHN-corpus

- Corpus Hedendaags Nederlands
- more than 800.000 texts
- consists of the corpora *5, 27 en 38 Miljoen Woorden Corpus* en het *PAROLE Corpus* (designed at the INT earlier)
- **recent newspaper texts (De Standaard and NRC) added: newspaper corpus**

Newspaper corpus: 1. De Standaard

- Flemish daily newspaper
- published in Belgium
- traditionally a christian-democratic newspaper (but not anymore)
- chief-editor: Karel Verhoeven

Newspaper corpus 2: NRC

- Dutch daily newspaper
- published in the Netherlands
- labeled left-liberal
- chief-editor: Peter Vandermeersch (until September 2019)

Neoloog

- computer tool to detect neologisms automatically
- designed at the INT by computer programmers Mathieu Fannee, Jan Niestadt and Rob van Strien
- for internal use only
- a corpus of Dutch newspapers is uploaded in Neoloog (*NRC* and *de Standaard*)
- updated every month
- ‘new words’ in this newspapers are detected by the computer as having appeared for the first time in and are presented in Neoloog in the form of a list
- every word can be viewed in its context

Neoloog

Neoloog

- 2016 11
- 2016 12
- 2017 01
- 2017 02
- 2017 03
- 2017 04
- 2017 05
- 2017 06
- 2017 07
- 2017 08
- 2017 09
- 2017 10
- 2017 11
- 2017 12
- 2018 01
- 2018 02
- 2018 03
- 2018 04
- 2018 05
- 2018 06
- 2018 07
- 2018 08
- 2018 09
- 2018 10
- 2018 11
- 2018 12
- 2019 01
- 2019 02
- 2019 03
- 2019 04

it ±193.562 rijen)
en

Kies een datum in menu boven links

<input type="text"/>	<input type="text"/>	Kiezen	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
lemma	freq	anw_status	neo	anw	opn	niet_neo

Geen resultaten. Probeer een ander zoekwoord.

Neoloog

2019 01

neologismen: Bestand in bewerking

4.007 rij(en) gevonden (uit =189.790 rijen)

Toon 10 rijen

ZOEK in gehele tabel:

ANW voorbereiding [UIT] Ik ben klaar! Maak ANW artikelen aan

Eerste Vorige 1 ... 21 22 23 ... 401 Volgende Laatste

woord	lemma	freq	anw_status	neo	anw	opn	niet_neo	twijfel	knop	comment
antiterrorismeketen	antiterrorismeket	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
anti-terrorismeoperaties	anti-terrorismeoperatie	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
anti-terrorisme-operaties	anti-terrorisme-operatie	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
antivleesdiscours	antivleesdiscours	8		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
appeljas	appeljas	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
app-vaardig	app-vaardig	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
april-grap	april-grap	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
april-november	april-november	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
arbeidersclubje	arbeidersclubje	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	
archeologieraadse	archeologieraadse	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden	

Neoloog

woord	lemma	freq	anw_status	neo	anw	opn	niet_neo
beachracen	beachrac	20		<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
beachracers	beachracer	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
beachraceseizoen	beachraceseizoen	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
beautyadventkalenders	beautyadventkalender	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
beautyadventskalender	beautyadventskalender	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
beauty liefhebber	beauty liefhebber	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
be-bop	be-bop	4		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
bebossingsactie	bebossingsactie	1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
bedankbier	bedankbier	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
bedankzin	bedankzin	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Neoloog

woord	lemma	freq	anw_status	neo	anw	opn	niet_neo	twijfel	knop
beachracen	beachrac	20		<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	Voorbeelden
beachracers	beachracer	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden
beachraceizoen	beachraceizoen	2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Voorbeelden

Neoloog

Beachrace De wind vol in het gezicht, hopend op wat grip in het mulle zand: beachracen is afzien. Zo ook in Scheveningen tijdens het EK. Lars Boom deed mee, als voorbereiding [bron]

volgend jaar in Frankrijk wordt gehouden, doen vooral veel Nederlanders mee. Nergens is het beachracen zo groot, de Nederlandse kustlijn is ideaal. Van Hoek van Holland naar Den Helder [bron]

en een stuk noordelijker bij IJmuiden. Maar ook in Frankrijk, België en Duitsland komt het beachracen op. De band van een beachracefiets is het meest zichtbare verschil met een normale [bron]

in 1993 al verreden, van Hoek van Holland naar Den Helder. Sinds 2013 is het beachracen een officiële discipline binnen de bond en inmiddels is er een Topcompetitie. In acht [bron]

. Met 625 euro en een jaar lang de kampioenstui als prijs. Nergens is het beachracen zo groot als in Nederland, waar de kustlijn ideaal is. Maar ook in de buurlanden [bron]

##beachracen##Beachrace De wind vol in het gezicht, hopend op wat grip in het mulle zand: beachracen is afzien. Zo ook in Scheveningen tijdens het EK. Lars Boom deed mee, als voorbereiding [bron]

##beachracen##volgend jaar in Frankrijk wordt gehouden, doen vooral veel Nederlanders mee. Nergens is het beachracen zo groot, de Nederlandse kustlijn is ideaal. Van Hoek van Holland naar Den Helder [bron]

##beachracen##en een stuk noordelijker bij IJmuiden. Maar ook in Frankrijk, België en Duitsland komt het beachracen op. De band van een beachracefiets is het meest zichtbare verschil met een normale [bron]

##beachracen##in 1993 al verreden, van Hoek van Holland naar Den Helder. Sinds 2013 is het beachracen een officiële discipline binnen de bond en inmiddels is er een Topcompetitie. In acht [bron]

##beachracen##. Met 625 euro en een jaar lang de kampioenstui als prijs. Nergens is het beachracen zo groot als in Nederland, waar de kustlijn ideaal is. Maar ook in de buurlanden [bron]

Beachrace De wind vol in het gezicht, hopend op wat grip in het mulle zand: beachracen is afzien. Zo ook in Scheveningen tijdens het EK. Lars Boom deed mee, als voorbereiding [bron]

volgend jaar in Frankrijk wordt gehouden, doen vooral veel Nederlanders mee. Nergens is het beachracen zo groot, de Nederlandse kustlijn is ideaal. Van Hoek van Holland naar Den Helder [bron]

en een stuk noordelijker bij IJmuiden. Maar ook in Frankrijk, België en Duitsland komt het beachracen op. De band van een beachracefiets is het meest zichtbare verschil met een normale [bron]

in 1993 al verreden, van Hoek van Holland naar Den Helder. Sinds 2013 is het beachracen een officiële discipline binnen de bond en inmiddels is er een Topcompetitie. In acht [bron]

. Met 625 euro en een jaar lang de kampioenstui als prijs. Nergens is het beachracen zo groot als in Nederland, waar de kustlijn ideaal is. Maar ook in de buurlanden [bron]

##beachracen##Beachrace De wind vol in het gezicht, hopend op wat grip in het mulle zand: beachracen is afzien. Zo ook in Scheveningen tijdens het EK. Lars Boom deed mee, als voorbereiding [bron]

##beachracen##volgend jaar in Frankrijk wordt gehouden, doen vooral veel Nederlanders mee. Nergens is het beachracen zo groot, de Nederlandse kustlijn is ideaal. Van Hoek van Holland naar Den Helder [bron]

Neoloog

neologismen: Bestand is KLAAR ZOEK in gehele tabel: beachracen

1 rij(en) gevonden (uit ±193.562 rijen)

Toon 10 rijen

ANW voorbereiding [AAN] Ik ben klaar! Maak ANW artikelen aan

Eerste Vorige 1 Volgende Laatste

woord	lemma	freq	anw_status	neo	anw	opn	niet_neo	twijfel	knop	comment	woordsoort	minidefinitie
beachracen	beachracen	20	Door Neoloog aangemaakt	<input type="checkbox"/>	Voorbeelden		werkwoord					

Neoloog

ANW voorbereiding [AAN] Ik ben klaar! Maak ANW artikelen aan

Eerste Vorige 1 Volgende Laatste

Kiezen

twijfel knop comment woordsoort minidefinitie

Voorbeelden werkwoord

Eerste Vorige 1 Volgende Laatste

Neoloog

neologismen: Bestand is KLAAR ZOEK in gehele tabel: beachracen

1 rij(en) gevonden (uit ±193.562 rijen)

Toon 10 rijen

ANW voorbereiding [AAN] Ik ben klaar! Maak ANW artikelen aan

Eerste Vorige 1 Volgende Laatste

woord	lemma	freq	anw_status	neo	anw	opn	niet_neo	twijfel	knop	comment	woordsoort
beachracen	beachracen	20	Door Neoloog aangemaakt	<input type="checkbox"/>	Voorbeelden		werkwoord				

Neoloog

neologismen: Bestand is KLAAR

ZOEK in gehele tabel: beachracen

1 rij(en) gevonden (uit ±193.562 rijen)

Toon 10 rijen

ANW voorbereiding [AAN] Ik ben klaar! Maak ANW artikelen aan

Eerste Vorige 1 Volgende Laatste

woord	lemma	freq	anw_status	neo	anw	opn	niet_neo	twijfel	knop	comment	woordsoort
beachracen	beachracen	20	Door Neoloog aangemaakt	<input type="checkbox"/>	Voorbeelden		werkwoord				

Neoloog

anw_status 

Door Neoloog aangemaakt

Door Neoloog aangemaakt

Door Neoloog aangemaakt

disadvantages

- list contains a lot of nonsense words
- list contains a lot of occasional compounds
- newspaper texts only

- ... what about neologisms in literature, vlogs, blogs, on social media, in youth language, etc.?

links

- ivdnt.org
- Neologism of the week:
<https://ivdnt.org/onderzoek-a-onderwijs/lexicologie-a-lexicografie/neologisme-van-de-week-archief>
- ANW: <http://anw.ivdnt.org/search>
- NW: