

/instituut voor
de Nederlandse
taal/

**A good match:
a Dutch collocation, idiom and
pattern dictionary combined**

Woordcombinaties
Word Combinations

lut.colman@ivdnt.org
carole.tiberius@ivdnt.org

Word combination

Any meaningful (type of) combination of words with spaces: free combinations and multiword expressions, like collocations, fixed expressions and idioms, but also more abstract semantically motivated (valency) patterns and conversational routines

The landscape of phraseological resources for Dutch

- 1 idiom dictionary
 - Van Dale Idioomwoordenboek (1999)
- 1 collocation dictionary
 - Combinatiewoordenboek (de Kleijn 2003)
- A Contrastive Verb Valency Dictionary D-E-F (Contragram 2006)
- Some proverb dictionaries (old and modern)
- MWEs in lexical databases: DuELME, Cornetto
- Phraseological information scattered over layered microstructures in general language dictionaries
- Focus on idioms, less on collocations



A good match: a Dutch collocation, idiom and pattern dictionary combined

/instituut voor
de Nederlandse
taal/



Woordcombinaties

Voorbeeldzinnen

Combinatiemogelijkheden

Patronen

Formules

aanmoedigen

Zoeken

Deze taaltool ontsluit een grote verzameling taal materiaal met teksten uit boeken, kranten, tijdschriften en het web.

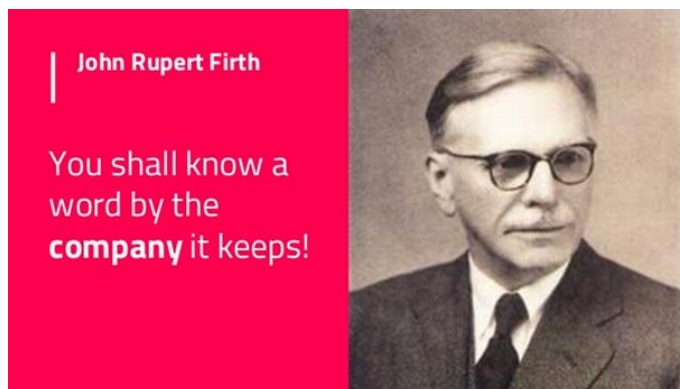
Je kan zien hoe woorden gebruikt worden in zinnen en welke woorden met elkaar gecombineerd worden:

- Voorbeeldzinnen: bekijk het trefwoord in representatieve voorbeeldzinnen.
- Combinatiemogelijkheden: met welke andere woorden of woordgroepen wordt het trefwoord gecombineerd?
 - Patronen: hoe bouw je zinnen met het trefwoord?
 - Formules: bekijk veelgebruikte formules met het trefwoord.

3-year pilot with verbs

- Database and web application
 - reference tool for comprehension and production
 - resource for CALL, CALT, researchers & developers of textbooks, teaching materials and test materials
- Stand-alone online dictionary & plug-in in language learning material and other dictionaries
- Target users: advanced L2 learners of Dutch in higher education
- Macrostructure: selection of frequent verbs
- Pilot corpus: \pm 300 million tokens
- Tools: Sketch Engine with GDEX and CPA, DWS (under construction)

Inspiration



- **SkE_{LL}**: Sketch Engine for Language Learning (Lexical Computing, Brno)
- **PDEV**: Pattern Dictionary of English Verbs (Patrick Hanks, Wolverhampton)
- **E-VALBU**: Das elektronische Valenzwörterbuch deutscher Verben (IDS, Mannheim)
- **StringNet Navigator 4.0** (David Wible, Taiwan)
- **CVVD**: Contrastive Verb Valency Dictionary (Contragram, Gent)
- **LEAD**: Louvain EAP dictionary
- **SweCcn**: Swedish Constructicon van Språkbanken (Gothenburg)
- **Combinatiewoordenboek** de Kleijn (Piet de Kleijn)
- Etc.

TNE & CPA (Patrick Hanks)

- TNE: Theory of Norms and Exploitations
 - Words have meaning potentials rather than meanings
 - Meanings are evoked by context
 - Norms are ‘normal’ phraseological patterns as main carriers of meaning
 - Exploitations are creative uses of normal patterns
- CPA: Corpus Pattern Analysis
 - Corpus-driven technique for mapping meaning onto words in context
 - Sketch Engine supports annotation of usage patterns in corpus samples
- Pattern: form-meaning pair
 - Slots in patterns populated by semantic types from an ontology:
 - [[Human]], [[Animal]], [[Furniture]],...
 - Semantic types are categories of lexical sets (collocates):
 - [[Furniture]]: table, chair, bed, ...

Option 1: examples with verb in context

Sometimes good example sentences suffice

- to work out the meaning of the verb
- to get a bird's-eye view of its usage patterns

GDEX: Good Dictionary EXamples

Customer data is **analysed** to answer certain key requirements.

The promising results are now being **analysed** .

The efficacy of rival cloud graphics is routinely **analysed** .

It was those data that were **analysed** here .

I am wary about **analysing** humour too much .

General estimating equation was applied to **analyse** data.


Many linguists **analyse** Japanese pitch accent somewhat differently.

These are then **analysed** by senior managers.

I **analysed** the situation as extraordinarily simple.


Two researchers reviewed and **analysed** this final study set.

<https://skell.sketchengine.co.uk>

 Voorbeeldzinnen

 Combinatiemogelijkheden

 Patronen

 Formules

analyseren

 Zoeken

analyseren

We hebben ondertussen al alle data **geanalyseerd** .

Hij **analyseert** de dagelijkse consumptie van beelden in kranten en weekbladen en ook in film .

De onderzoekers **analyseerden** voor het eerst het profiel van mensen tegen wie een vordering voor uithuiszetting loopt .

Elke zucht wordt uitvergroot , langs alle kanten bekeken en de impact ervan **geanalyseerd** .

Christ , als ze maar niet mijn kinderen gaat zitten **analyseren** .

Deze componenten worden dan gebruikt om de huidige structuur van de Belgische varkenshouderij te **analyseren** .

Integendeel , er ligt veel nadruk op **analyseren** en je goed kunnen uitdrukken .

De concrete invulling is echter nog niet helder ; daardoor kunnen ook de effecten nog niet worden **geanalyseerd** .

Option 2: word sketch

The screenshot shows the SKEL interface for the word 'analyse'. At the top, there is a search bar with 'analyse' and a magnifying glass icon. To the right of the search bar are two tabs: 'Examples' and 'Word sketch', with 'Word sketch' being the active tab. Below the search bar, the word 'analyse' is shown with a 'verb' tag and a 'Context' button. The main content area is divided into several sections, each with a title and a list of related words in a light blue box:

- subject of analyse**: researcher, scientist, paper, study, report, book, team
- object of analyse**: datum, sample, behaviour, DNA, situation, trend, result, pattern, structure, impact, content
- modifiers of analyse**: critically, separately, carefully, properly, fully, then, first, here, also, as, only
- words and**: interpret, evaluate, report, understand
- or analyse**: describe, collect, collate, gather, examine, record, monitor, identify

analyseren

subject

laboratorium onderzoeker expert instituut computer deskundige

object

monster gegeven **datum** situatie informatie tekst ontwikkeling staal resultaat wedstrijd probleem zaak beeld

bepaling

grondig verder goed nader kritisch nauwkeurig zorgvuldig snel automatisch

and/or

interpreteren verbeteren optimaliseren rapporteren beoordelen oplossen vergelijken evalueren verwerken bespreken beschrijven

subject Who analyses? In a passive sentence this may be expressed in a phrase introduced by *door* (*by*).

het **laboratorium** de onderzoeker de expert het **instituut** de computer de deskundige



Drie **laboratoria analyseerden** onafhankelijk het mitochondriaal DNA uit een kies .

More ...



Zijn **laboratorium analyseert** momenteel everzwijnvlees om de uitbraak met de parasiet in ons land op het spoor

De stalen worden **geanalyseerd** door **laboratoria** in Portugal en Zwitserland .

Drie **laboratoria analyseerden** onafhankelijk het mitochondriaal DNA uit een kies .

De stalen worden **geanalyseerd** door de **laboratoria** van het KIK .

De stalen werden **geanalyseerd** door het **Laboratorium** voor Biodiversiteit en Evolutionaire Genomica (KU Leuven

Option 3: patterns & meanings

PDEV: analyse analyze

[Access full data](#)

Displayed here are **All patterns** .
Other options:

1 Pattern: **Human** or **Computer** or **Process** analyses **Information** or **Anything**

Implicature: **Human** or **Computer** or **Process** examines **Information** or **Anything** methodically and in detail, in order to understand and explain it

Example: The researchers will read and analyze the written material and will follow up some responses with face-to-face interviews.

E-VALBU

1 analyseren

Strukturbeispiel:	jemand analysiert etwas mit irgendetwas
Im Sinne von:	jemand untersucht systematisch die einzelnen Faktoren von etwas mittels irgendetwas
Satzbauplan:	K_{sub} , K_{akk} , (K_{adv})
Beispiele:	(2) Mit einem Supercomputer analysierten sie die Signale der Synchrotron-Strahlung. (3) Die Regierung hat die Ursachen der Unzufriedenheit falsch analysiert. (Zeit, 08.02.1985, S. 15) (4) [Valenzreduktion]Die Untersuchung analysiert die Lage auf dem Arbeitsmarkt.

Pattern 1: iemand | iets analyseert iets

someone | something analyses something

Meaning: iemand | iets onderzoekt iets door het methodisch te ontleden

Example: De onderzoekers analyseerden gegevens over 6870 mensen van 16 jaar en ouder. [More examples](#) 

Pattern 1: iemand | iets analyseert iets



laboratorium
onderzoeker
expert
instituut
computer
deskundige

STs: [[Human]], [[Institution]], [[Computer]]

More features on demand: verb forms, structure and optionality, passivization, pragmatics

Idioms & proverbs

- Patterns with very limited or unique lexical preferences and/or specific phrase structures in some or all of the slots
 - Pattern: (uitdrukking) **Iemand** begraaft de strijdbijl
(idiom) **Someone** buries the hatchet *the spade
 - Pattern: (spreekwoord) De appel valt niet ver van de boom | niet ver van de stam
(proverb) The apple *the pear never falls far from the tree
- Using patterns as canonical forms allows for STs and lexical sets made explicit in the variable slots: **someone** = [[Human or Institution]]
- Also thematic searches and reverse searches on our wish list
 - From theme to pattern: e.g. actions & behaviour → **Someone** buries the hatchet
 - From meaning to pattern: e.g. make peace → **Someone** buries the hatchet



Option 4: conversational routines (formules)

- “highly conventionalized prepatterned expressions whose occurrence is tied to more or less standardized communication situations” (Coulmas 1981)
- Aijmer (2014: 2) distinguishes three major classes:
 - Formulaic speech acts, such as apologizing, thanking and greeting
 - *I'm sorry, but ...* or *Thank you!*
 - Discourse markers
 - *As I say ...*
 - Attitudinal routines expressing the speaker's attitude or emotion
 - *Go to hell!*
- Routines come natural to native speakers, but they are difficult to master for non-native language learners
- Pragmatic function predominates over the referential meaning (Aijmer 2014: 11) → not practicable to list them with the semantically motivated pattern-meaning pairs

Pragmatic meaning → pragmatic search


- Access through predefined lists of speech acts and/or situations

I want to 

ask information

give information

apologise

express emotion 

...

anger

joy

doubt

surprise

in this situation 

public transport

work

shopping

conversation

...

Hoe laat vertrekt/gaat de trein naar x? (What time does the train for x depart?)

Coming soon on www.ivdnt.org !

/instituut voor
de Nederlandse
taal/



Woordcombinaties

Voorbeeldzinnen

Combinatiemogelijkheden

Patronen

Formules

aanmoedigen

Zoeken

Deze taaltool ontsluit een grote verzameling taalmateriaal met teksten uit boeken, kranten, tijdschriften en het web.

Je kan zien hoe woorden gebruikt worden in zinnen en welke woorden met elkaar gecombineerd worden:

- Voorbeeldzinnen: bekijk het trefwoord in representatieve voorbeeldzinnen.
- Combinatiemogelijkheden: met welke andere woorden of woordgroepen wordt het trefwoord gecombineerd?
 - Patronen: hoe bouw je zinnen met het trefwoord?
 - Formules: bekijk veelgebruikte formules met het trefwoord.

