

# Activiteit 3

---

## Zeg het opnieuw! —Tekst compressie

### Samenvatting

Omdat computers maar een beperkte ruimte hebben om informatie op te slaan, moeten we deze informatie zo efficiënt mogelijk beschrijven. Dit heet compressie. Door de data te coderen voor het opslaan en weer te decoderen wanneer het weer wordt opgehaald kan de computer meer data opslaan en het sneller over het internet versturen.

### Kerdoelen

- Taal: het herkennen van patronen in woorden en tekst.
- Techniek: Technische kennis en begrip. Hoe werken computers.

### Vaardigheden

- Kopieren van geschreven tekst

### Leeftijd

- 9 jaar en ouder

### Materialen

- *Voorbeeld voor in de klas: Zeg het opnieuw!*

Iedere leerling heeft nodig:

- *Werkblad: Zeg het opnieuw!*

Optionele uitbreidingsopdrachten:

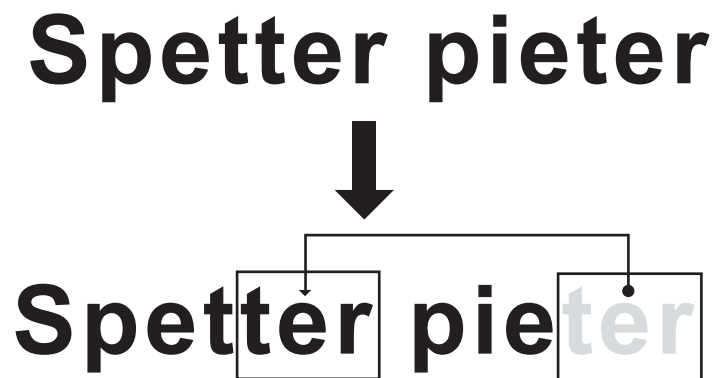
- *Werkblad: Extra voor xperts*
- *Werkblad: Kort en krachtig*

## Zeg het opnieuw!

---

### Introductie

Computers moeten veel data opslaan en doorsturen en wel zo dat de data niet te veel ruimte innemen of dat het heel lang duurt om de informatie door te sturen. Ze verkleinen de tekst op ongeveer deze manier.



### Demonstratie en discussie

Laat het hele refrein van *Spetter pieter pater* zien in de klas (zie 'Voorbeeld voor in de klas, pagina 29). Zoek patronen in de letters van dit gedicht. Vind je groepen van 2 of meer letters die herhaald worden of misschien zelfs hele woorden of zinsdelen? (Vervang deze met rechthoeken net zoals hierboven.)

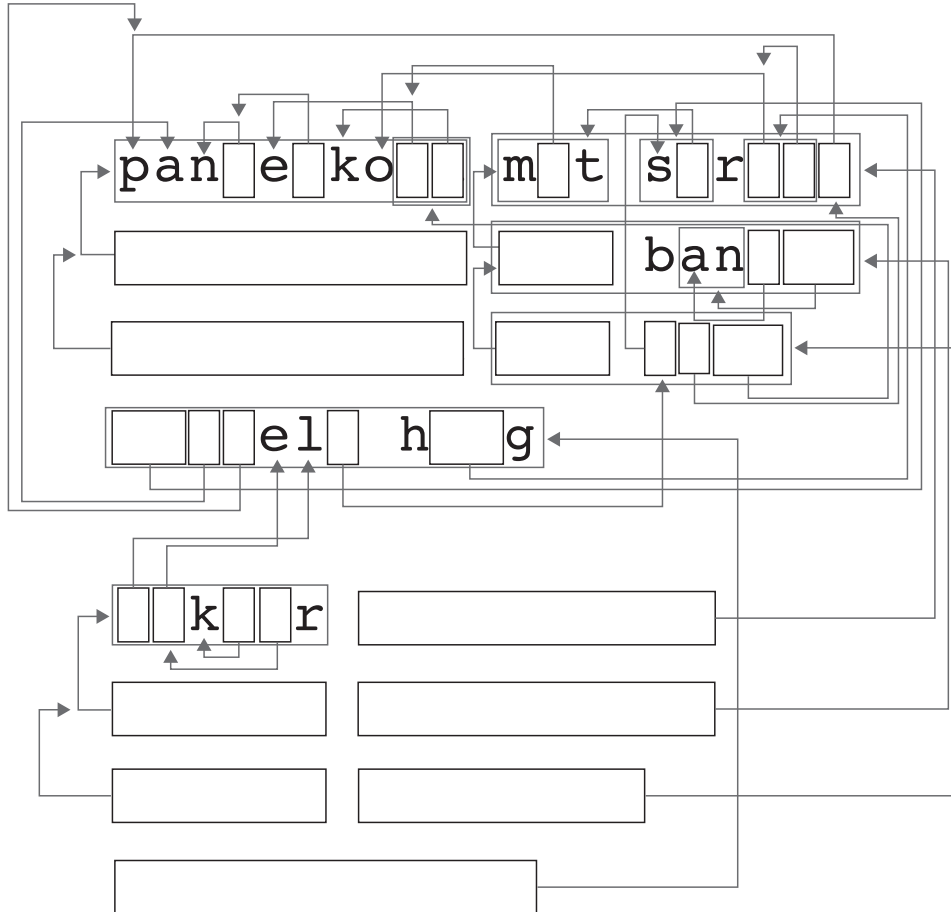
Voorbeeld voor in de klas: Zeg het opnieuw!

**Spetter pieter pater  
lekker in het water  
ga maar vast naar huis  
ik kom een druppel later**

# Werkblad: Zeg het opnieuw!

---

Veel woorden en letters missen in dit gedicht. Kan je de missende letters en woorden weer vinden? Je vindt deze in de rechthoeken waar met de pijlen naar verwezen wordt.



Kies nu een simpel gedicht of kinderversje en maak je eigen puzzel. Zorg ervoor dat de pijlen altijd naar een eerder stuk in de tekst verwijzen. Je puzzel moet ontcijferd kunnen worden door van links naar rechts te gaan en van boven naar beneden, net zoals we lezen.

**Uitdaging:** Probeer zo min mogelijk van de oorspronkelijke tekst te behouden!

Een aantal suggesties: *Slaap kindje slaap, Lang zal ze leven, Drie kleine kleutertjes* — of probeer een stuk uit *Kat met de hoed!*

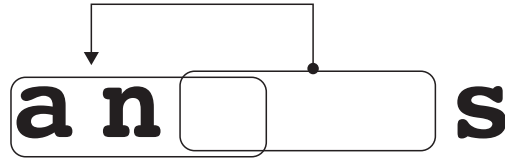
**Tip:** Om een wirwar van pijlen te vermijden: Schrijf ruim, laat veel ruimte over tussen de woorden en regels als je ze opschrijft, zodat je ruimte hebt voor rechthoeken in andere rechthoeken en pijlen die naar de rechthoeken verwijzen.

Het is ook makkelijker om je puzzel te maken door eerst het hele gedicht op te schrijven en daarna pas te beslissen waar de rechthoeken moeten komen.

## Werkblad: Extra voor Experts

---

Hoe zou je deze puzzel oplossen?



Missende tekstdelen kunnen ook naar zichzelf wijzen. In dit geval kan de het woord prima ontcijferd worden als de letters van links naar rechts worden gekopieerd. Iedere letter is bruikbaar voor de kopie waarvoor hij nodig is. Dit is bruikbaar voor computers als er veel bepaalde letters achter elkaar voor komen of sprake is van een patroon van letters.

Bij computers worden de rechthoeken en pijlen weer gegeven door getallen. Bijvoorbeeld, **ananas** kan geschreven worden als **an (2, 3) s**. “2” betekent tel twee letters terug om het beginpunt voor het kopiëren te vinden,

**an\_\_s**

en “3” betekent kopieer de 3 volgende letters:

**ana\_\_s**

**anan\_s**

**ananas**



Omdat er twee getallen zijn gebruikt om dit woord te coderen zijn normaal gesproken alleen groepen van twee of meer letters de moeite waard om te verkleinen, anders wordt er geen ruimte bespaard. Sterker nog: de grote van het bestand zou zelfs kunnen groeien als er twee getallen worden gebruikt om een letter te coderen.

Verzin nu wat woorden of een zin voor jezelf en codeer ze zoals een computer zou doen als ze verkleind zouden worden. Kan een klasgenoot ze ontcijferen?

## Werkblad: Kort en krachtig

---

Hoeveel woorden heb je hier echt nodig?

Stel je voor dat je een computer bent en zo veel mogelijk op je harde schijf wilt opslaan. Streep alle groepen van 2 of meer letters door die al een keer gebruikt zijn. Deze zijn niet meer nodig omdat ze vervangen kunnen worden door een verwijzing. Probeer zo veel mogelijk letters door te strepen.

*Ik ken een dame die een vogel doorslikte,  
je schrikt hé, een vogel doorslikte.  
Ze slikte de vogel om de spin op te vangen  
die kriebelend friemelend in haar slokdarm bleef  
hangen.  
Ze slikte de spin, om de vlieg weg te jagen  
En waarom de vlieg?  
Ik hoor het je vragen  
Dat weet alleen zij, tot het einde der dagen.*

## Waar gaat dit eigenlijk over?

---

De opslagcapaciteit van computers neemt ongelooflijk snel toe, in de laatste 25 jaar is de opslagruimte van een 'gewone' computer ongeveer een miljoen keer groter geworden. Maar we vinden ook steeds meer dingen die we op onze computer op willen slaan. Hele boeken, zelfs hele bibliotheken, muziekcollecties en films, maar die kosten allemaal wel opslagruimte. En dat is niet het enige probleem; we willen het ook allemaal delen en grote bestanden zijn een probleem op het internet, dan duurt het te lang om te downloaden. En we maken computers ook steeds kleiner, zelfs een telefoon of een horloge moet tegenwoordig veel op kunnen slaan.

Gelukkig is er een oplossing voor dit probleem. In plaats van alleen maar meer opslagruimte bij te kopen of een snellere internetverbinding te nemen, kunnen we data verkleinen door een compressie toe te passen zodat de data minder ruimte inneemt. Dit proces van verkleinen en weer uitpakken van data gebeurt gewoonlijk automatisch door de computer. Wij merken alleen dat de harde schijf meer op kan slaan en websites sneller zichtbaar zijn, maar de computer doet al het harde werk.

Er zijn veel methodes van compressie uitgevonden. De methode in deze opdrachten, met het verwijzen naar eerder gebruikte stukken tekst, heet 'Ziv-Lempel encoding' of 'LZ encoding', uitgevonden door 2 Israëlische professoren in de jaren zeventig. Het kan in iedere taal gebruikt worden en kan makkelijk de data tot de helft verkleinen. Je kent het waarschijnlijk wel onder de naam 'zip' bestanden en het wordt ook gebruikt voor 'GIF' afbeeldingen.

Andere methodes van tekstcompressie zijn gebaseerd op het idee dat letters die vaak gebruikt worden kortere codes moeten hebben dan andere, net als de Morsecode.

## Oplossingen en hints

---

