

Spraaktechnologie

Gerrit Bloothoof

Gerrit.Bloothoof@let.uu.nl

Communicatie

taal was er eerst als spraak en gebaar

pas veel later als schrift

- en lezen en schrijven
(eerste taaltechnologie)

Communicatie

- Spraak
 - akoestische representatie van boodschap, van taal
 - nadruk, emotie (supra-linguistisch)
 - overdracht
 - snel
 - draagt ver
 - alle richtingen (geen oogcontact nodig)

Communicatie

- Gebaar
 - mimiek, lichaamstaal, gebarentaal
 - aanwijzen
 - lipbewegingen
 - ondersteunend (doven)

soms ook ondersteunend in spraaktechnologie

Schrift

relatief recente, technologische ontwikkeling

tekst

symbool
letter
eenduidig
sprekeronafhankelijk
nadruk expliciet /niet
emotie expliciet /niet
langzaam
één richting
vastleggen
doorzoekbaar

spraak

geluidsignaal
klank
niet-eenduidig
sprekerafhankelijk
nadruk impliciet
emotie impliciet
snel
alle richtingen
opslaan (zeer recent)
doorzoekbaar (zeer, zeer recent)

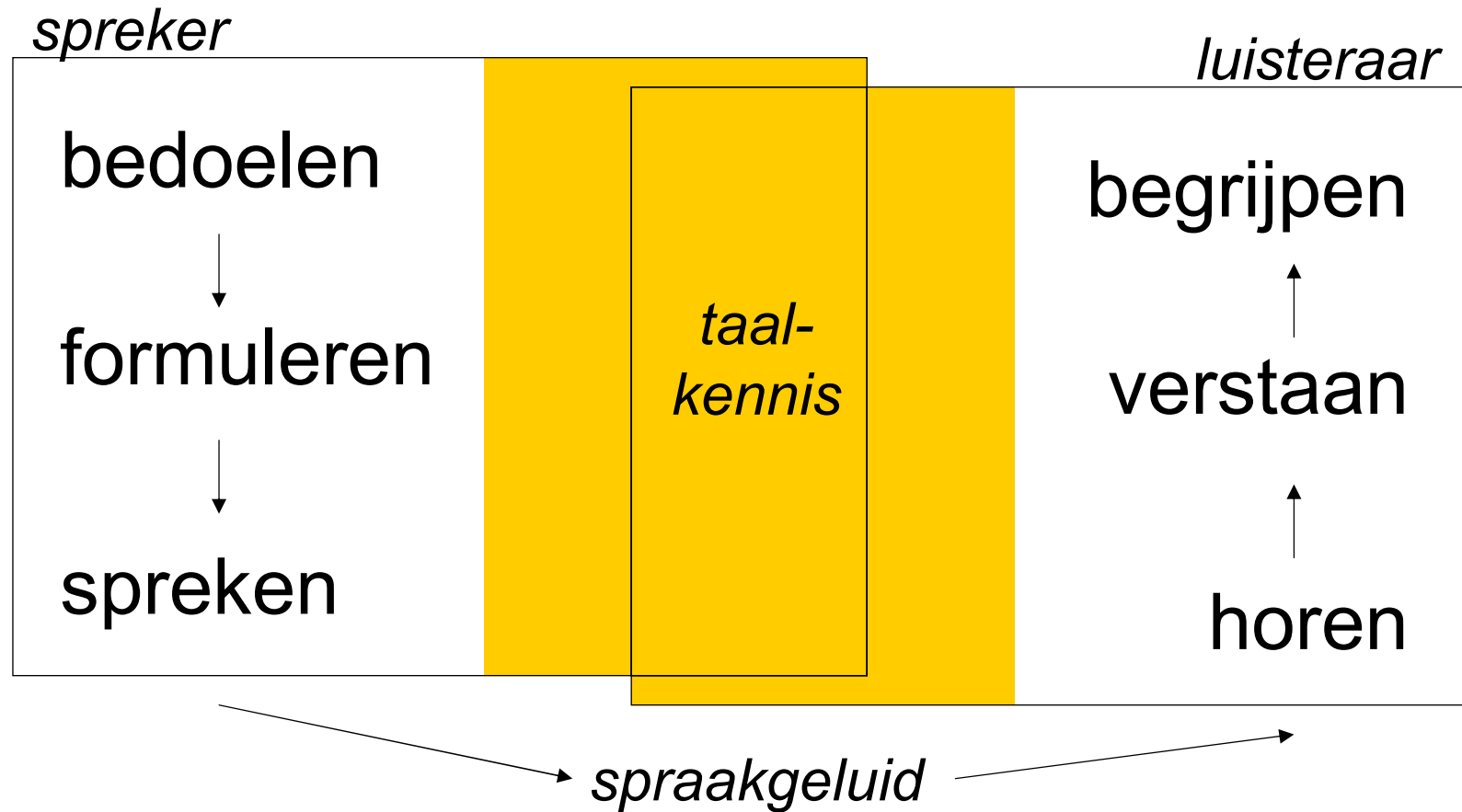
Wetenschap en Technologie

- Begrijpen van fundamentele processen bij de mens
 - fonetiek
 - psycho/neurolinguïstiek
 - taalwetenschap
- Communicatieprocessen met machines
 - luisteren naar machines
 - spreken tot machines
 - dialoog met machines

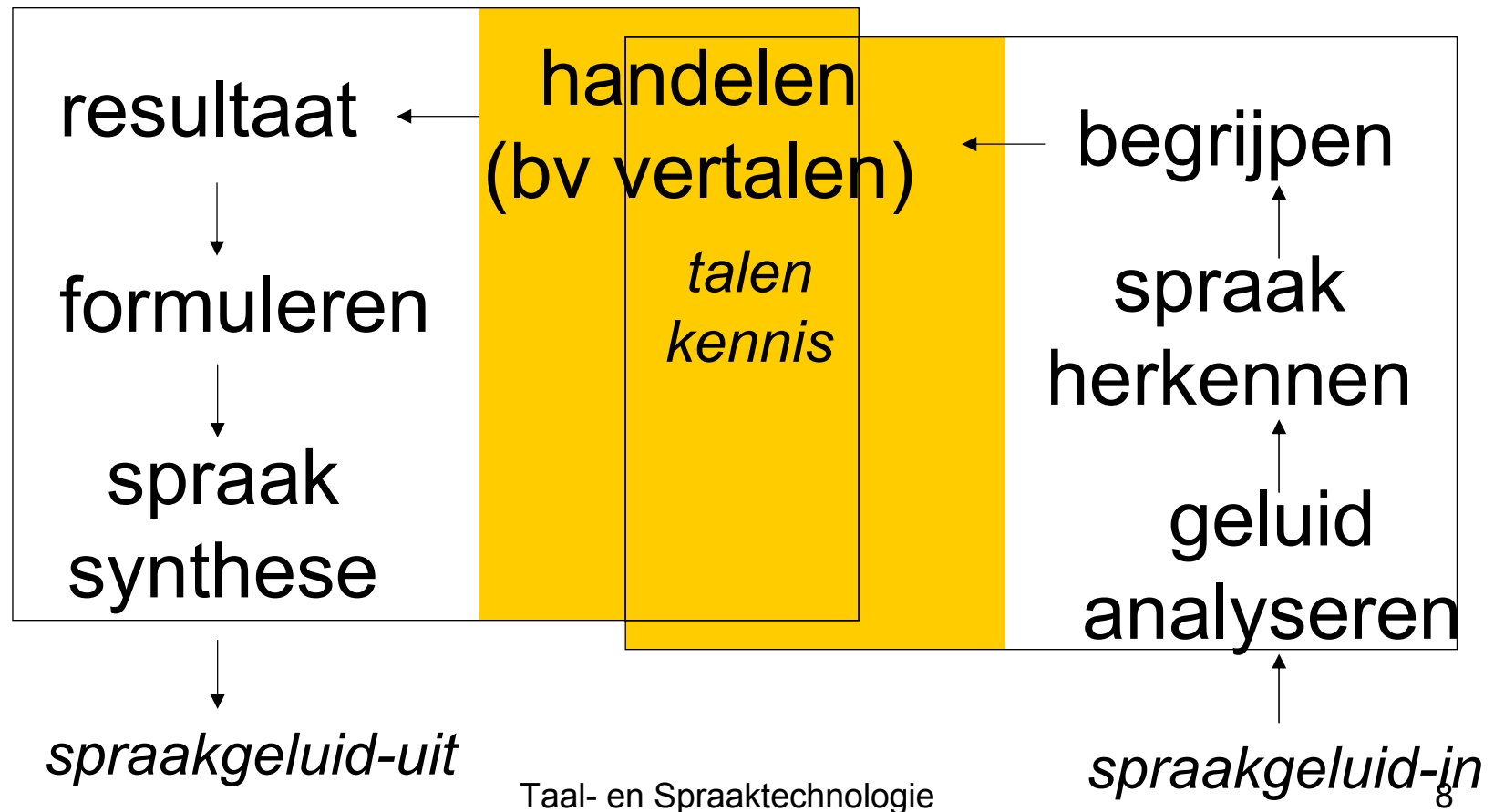
Modellen nodig van deze processen

Sterke wisselwerking tussen wetenschap en technologie

de spraakketen



de spraakketen (computer)



Hoofdpijnen in spraaktechnologie

- **Spraaksynthese**
 - audiovisuele spraak
 - spraakgeneratie 
- **Spraakherkenning**
 - spraakbegrijpen
- **Dialogsystemen**
 - multimodale systemen
- **Spraak-naar-spraakvertalen**
- **Sprekerherkenning**
 - verificatie
 - identificatie
- **Taalherkenning**
- **Spraakcodering**
 - 64 Kb/s naar 2 Kb/s

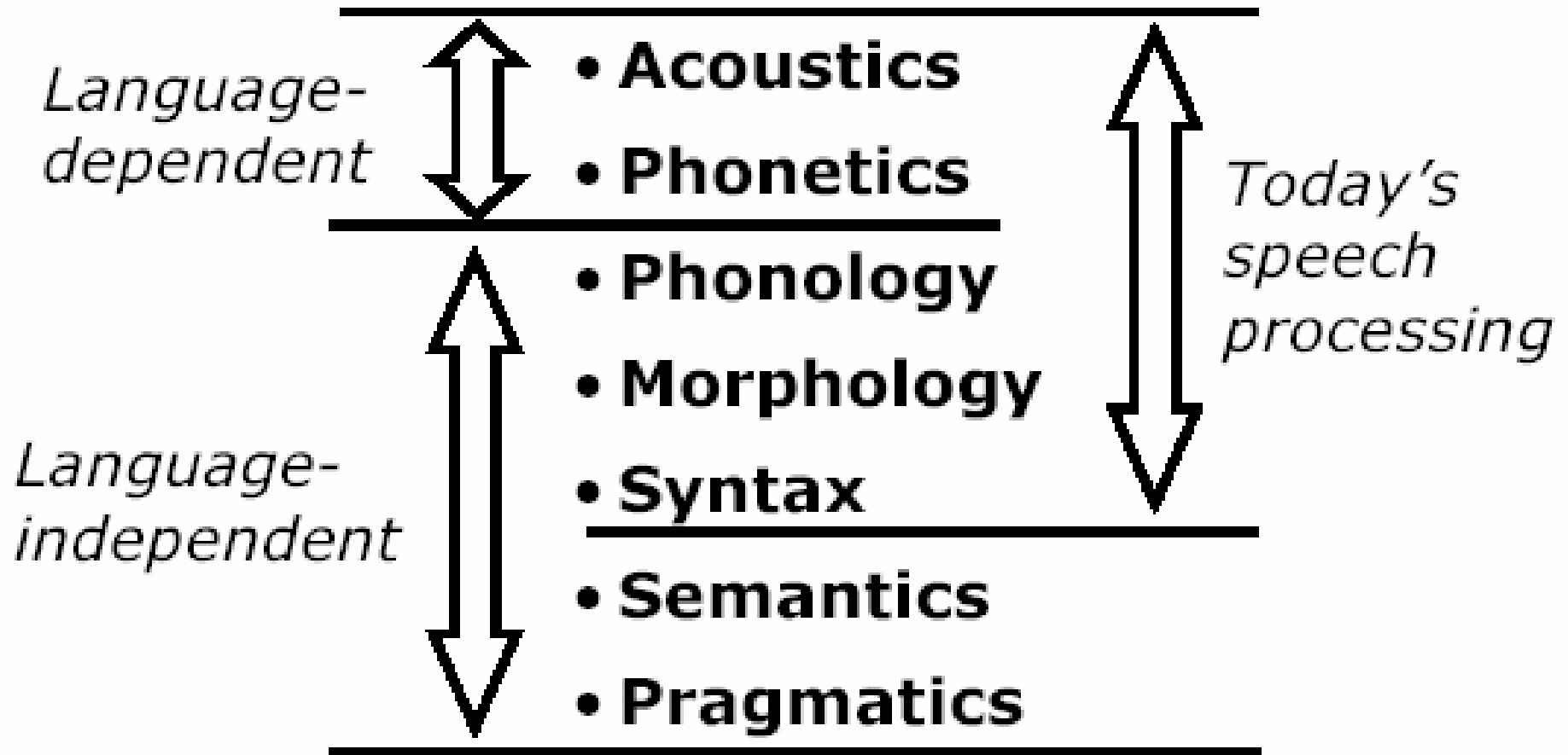
Toepassingen

- Voorleesmachines
 - voor blinden
 - E-mail reader, SMS-reader, Webreader
 - informatie (stations, vliegvelden)
- Hands-free bediening
 - gehandicapten
 - ingewikkelde besturingen, consumentenelectronica
- Spreken met een computer
 - Naturally Speaking,..(RSI preventie)
- Dialoogsystemen (multi-modaal, multi-linguaal)
 - interactieve informatie- en besteldiensten (OV, girofoon, koerslijn)
 - Computer-Aided (Language) Learning
- Zoeksystemen (audio-archieven)
- Toegangs- en transactiebeveiligingen

Spraak- en taaltechnologie

- **Spraaksynthese**
 - Letter-klanksymbool omzetting
 - Morfologische analyse voor klemtoontoekeening
 - Zinsontleding voor ritme en melodie
 - Semantische ontleding voor herkennen van oude en nieuwe informatie met effect op ritme en melodie
- **Spraakherkenning**
 - Uitspraakmodellen
 - Taalmodel
- **Dialogsystemen**
 - Dialooganalyse
 - Begrijpen en beslissen
- **Spraak-naar-spraak vertaling**
 - Automatisch vertalen

Zeven lagen



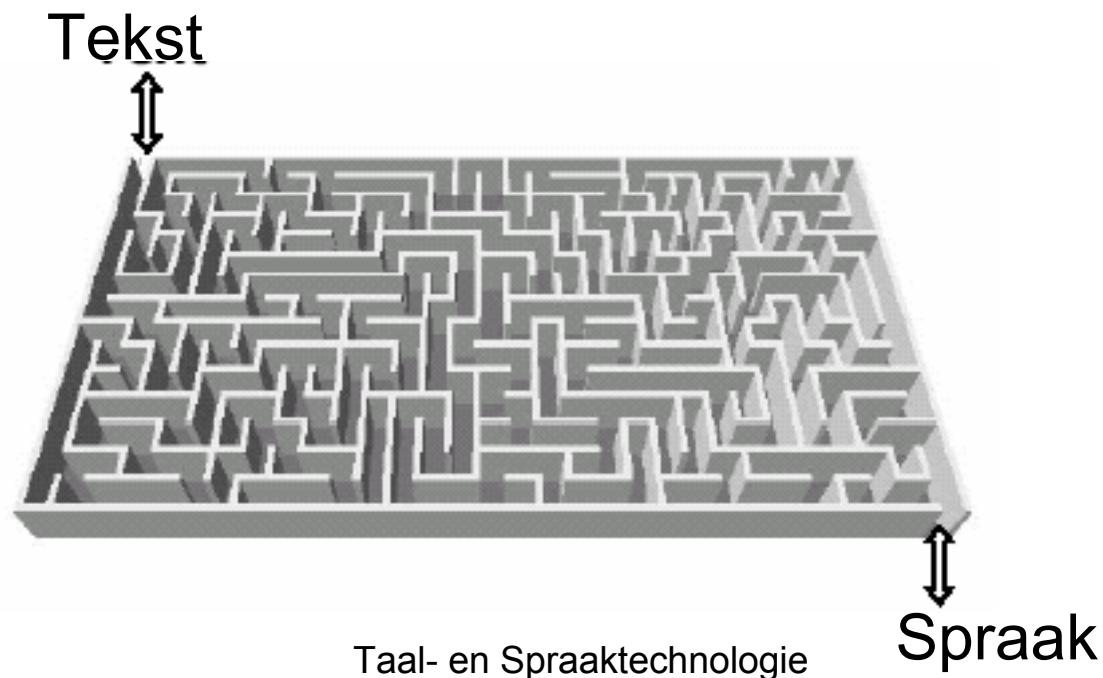
Tekst en spraak

- Tekst is informatie in de vorm van een verzameling symbolen
- Spraak is een geluidsignaal dat (dezelfde) informatie bevat

of niet helemaal...

Spraaktechnologie

- Symbool naar signaal conversie
- Signaal naar symbool conversie



Problemen

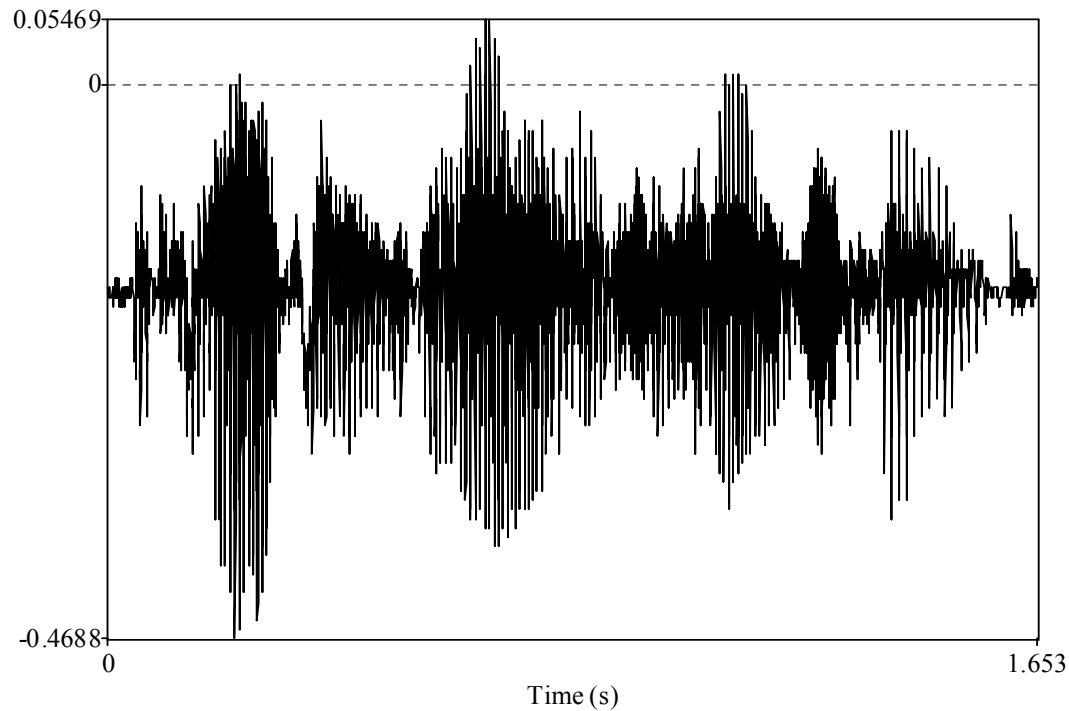
- taal moet zo *geproduceerd* worden dat ze *begrepen* kan worden.
- voor spraak:
 - zo weinig mogelijk energie gebruiken in productie
 - zuinig articuleren (spreek niet alles (precies) uit)
 - spraak verre van perfect en eenduidig
 - gebruik kennis van de taal bij luisteraar
 - gebruik context, maar hoe
 - mensen verschillen
 - variatie in klank en taal
 - waar zit invariantie (als die er al is)? (als bij symbolen)

Communicatievoorwaarden

spreker en luisteraar

- hebben toegang tot elkaar
 - lucht
 - lucht + telefoon (elektromagnetische golven | analoog of digitaal)
 - **verstoringen (akoestiek ruimte, achtergrondlawaai, lijnvervorming)**
- kennen elkaars taalcode
 - drukken boodschap op dezelfde manier uit in codering / decodering

Het leven is mooi als de zon schijnt



geen pauzes, slordigheden,
sprekereigenschappen,
verstoringen

Modelkeuze

Niet noodzakelijk de modellen uit psycholinguïstiek en fonetiek

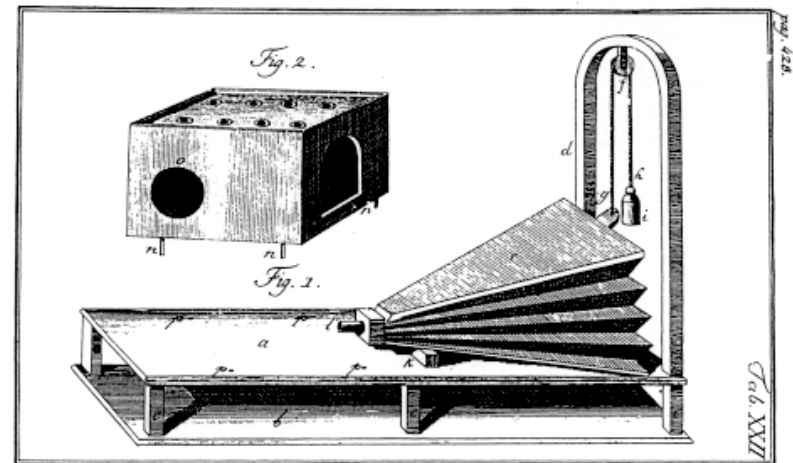
- Copiëren

- spraaksynthese (18e eeuw)

- Imiteren

Zelfde functie, maar (gedeeltelijk) via andere processen

- een vliegtuig klapwiek niet
- spraaksynthese [voorbeelden](#)



Regels en kansen

Regels

- eenduidig (consistent), als ... dan ...
- expliciete kennis

Beperkingen

- variabiliteit in akoestische realisatie
- ambiguiteiten

Beschrijf verschijnselen als *waarschijnlijkheden* die door training geleerd kunnen worden

Spraak in deze cursus

- Basis spraakanalyse (fonetiek)
- Spraaksynthese (difoonsynthese)
 - Theorie
 - Practicum (in teams)
 - Grafeem-foneem regels
 - (Regels voor toonhoogte en duur)
 - (Gezichtsuitdrukking)
- Spraakherkenning
 - Theorie