

# Tentamen spraakherkenning en -synthese (1)

19 oktober 2009

Zet op iedere pagina je naam, je studentnummer en het volgnummer per pagina. Gebruik voor elke opgave (1-4) een apart vel. Als je voor 17.00 uur klaar bent, lever dan je tentamen in bij de surveillant en verlaat rustig de zaal.

Het cijfer voor dit schriftelijk tentamen bepaalt de helft van je eindcijfer voor dit college. Dat eindcijfer wordt echter pas aan het OWI doorgegeven als je alle practicumverslagen voor dit onderdeel hebt voltooid en ze allemaal zijn goedgekeurd. Nadere informatie hierover vind je op de Blackboard site.

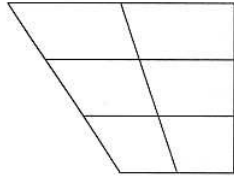
Beantwoord onderstaande vragen zo nauwkeurig, compleet en gedetailleerd mogelijk, zonder in herhalingen te vervallen. De vragen zijn zo gesteld dat een kort antwoord volstaat. Een erg lang antwoord is bijna zeker een verspilling van tijd. Het kan geen kwaad relevante ervaringen vanuit het hoorcollege en de practicumopdrachten in je antwoorden te verwerken. Bij de beoordeling van dit tentamen wordt meer gelet op begrip dan op feitenkennis. Besteed niet te veel tijd aan afzonderlijke opgaven. Als een opgave te veel tijd vergt, probeer dan eerst een andere opgave.

Veel succes!

David Weenink

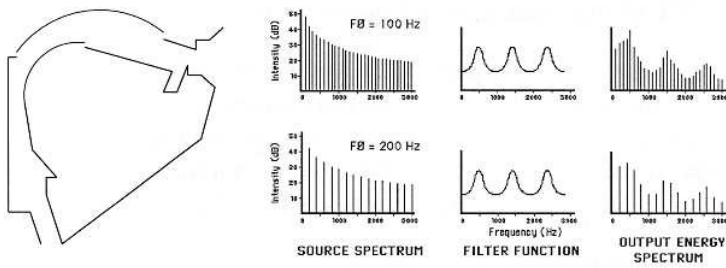
# 1 Spraakklanken

1. Schrijf de twaalf Nederlandse klinkers en de drie Nederlandse tweeklanken (diftongen) op. Je mag ze in geven in een woord van één lettergreep.
2. Zet de twaalf Nederlandse klinkers in een schematische klinkerruimte zoals onderstaande figuur. In deze figuur is de dimensie hoog/laag verticaal en de dimensie voor/achter horizontaal uitgezet.



3. Hoeveel fonemen heeft het woord *viering*? Schrijf ze op.
4. Rubriceer de fonemen van *viering* op manier en plaats van articulatie.

## 2 Spraaksignaalanalyse en corpora

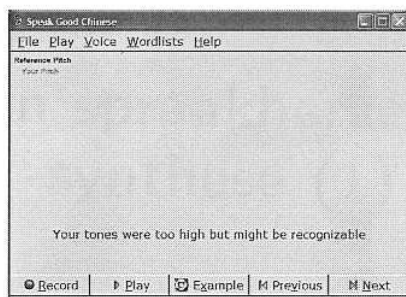


1. Leg uit wat bedoeld wordt met het bron-filter model aan de hand van bovenstaande figuren. Hou het kort.
2. Hoe worden formanten gevormd tijdens het spreken en wat is hun rol in de spraak?
3. Waarom zijn de formanten van mannenstemmen gemakkelijker te meten dan die van vrouwen of kinderen?
4. De toonhoogtemeter in Praat werkt via periodiciteitsdetectie. Leg kort hoe deze toonhoogtemeter ongeveer werkt.
5. Wat is het verschil tussen een oscillogram, een spectrum en een spectrogram.

### 3 Spraaksynthese en corpora

1. Normaliseer het volgende zinnetje:  
*Het tentamen SHS is op 19/10 en duurt van 14-17u.*
2. Beschrijf in het kort wat er zoal nodig is om een gedrukte-tekst-op-papier om te zetten in verstaanbare synthetische spraak. Welke stappen kunnen we hierin onderscheiden? Zeg tenminste iets over: text preprocessing; grafeem-foneemconversie via uitspraaklexicon en/of regels; prosodiecontrole (spreektempo, woordklemtoon, zinsaccent, intonatie, frasering), en feitelijke signaalgeneratie.
3. Geef aan waar spraakcorpora voor gebruikt kunnen worden in spraaktechnologie(onderzoek).
4. Het Corpus Gesproken Nederlands bevat ongeveer 1 miljoen woorden waarvan de automatische fonetische transcriptie en de woordgrenzen handmatig gecontroleerd zijn. Er is ook een automatische oplijning van de fonemen gemaakt. Bij de aanmaak van het Corpus is in een vroeg stadium besloten om de oplijning van de fonemen niet handmatig te controleren. Beargumenteer waarom het handmatig controleren van foneemgrenzen zo arbeidsintensief is. Betrek daarbij de betrouwbaarheid van de oplijning tussen de uitvoerders.
5. Het Corpus Gesproken Nederlands bevat ongeveer 8 miljoen *gesproken* woorden in verschillende spreekstijlen. Er zijn *tekst*corpora van meer dan een miljard woorden verdeeld over verschillende tekststijlen. Geef de redenen waarom verschillende spreek- en tekststijlen gebruikt worden. Geef voorbeelden van spreek- en tekststijlen.

## 4 SpeakGoodChinese



Het programma SpeakGoodChinese helpt studenten met het oefenen van de uitspraak van het Mandarijn Chinees.

1. Het MandarijnChinees is een toontaal. Wat is een toontaal en hoe verschilt een toontaal van het Nederlands, Engels of Duits?  
Waarom hebben Nederlandse studenten zoveel moeite om Chinees te leren?
2. Beschrijf in het kort de tonen van het Mandarijn Chinees. Gebruik zo mogelijk figuren.
3. Waarom denk je dat het meten van de toonhoogte zo moeilijk is voor toontalen?
4. SpeakGoodChinese heeft de beschikking over een verzameling van audio-opnamen van alle 1200 verschillende lettergrepen van het Mandarijn Chinees, inclusief alle voorkomende tonen. Alle woorden zijn door dezelfde spreker op een consistente manier ingesproken.
  - a) Stel je zou een difoonsynthese willen maken zodat je ook meerlettergrepige woorden kunt laten horen. Waarom kun je deze audiobestanden daar niet voor gebruiken?
  - b) Een alternatieve manier van synthese zou zijn om de audiofiles van de afzonderlijke lettergrepen achter elkaar te plakken. Wat voor problemen kun je hierbij verwachten?