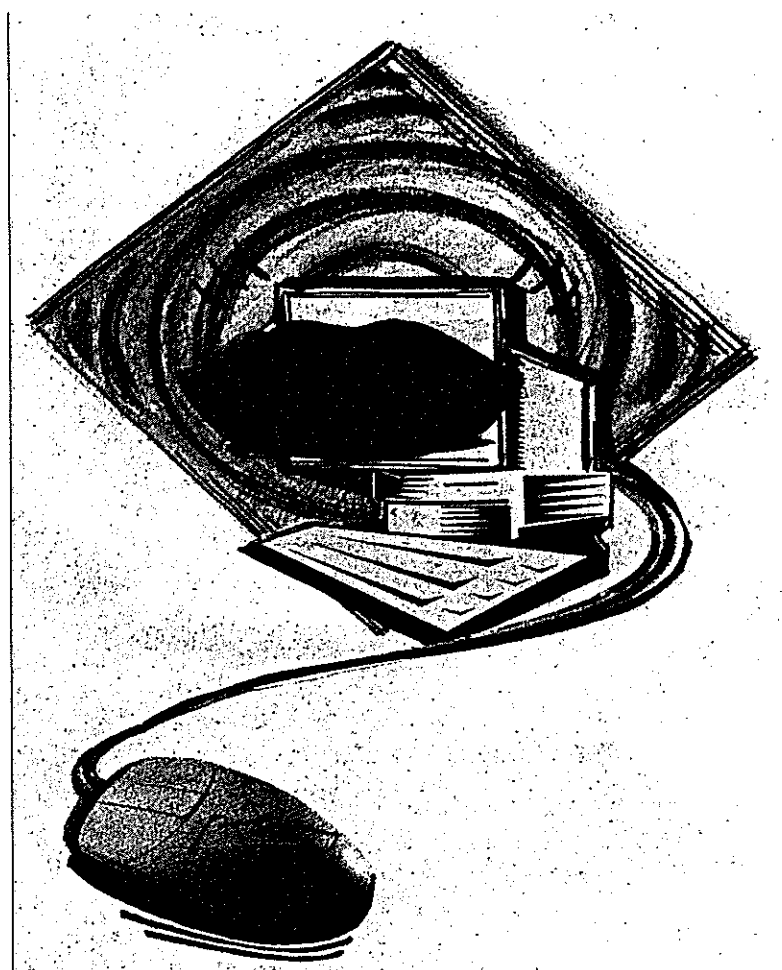


De Geïntegreerde proef : Spraaaktechnologie



Naam: Seys Michiel
& Simpelaere Dennis
Leraars: Mr. Verhaeghe
& Mevr. Degryse
Vak: De geïntegreerde proef

Schooljaar: 1998-1999
Klas: 614
Nrs: 10 & 11

spraaktechnologie

WOORD VOORAF

Wij hebben van school uit de opdracht gekregen om een geïntegreerde proef te maken. Hieruit hebben wij geopteerd voor het onderwerp ruimtevaart

Ruimtevaart omvat heel veel. Het is dus logisch dat wij maar een klein deel kunnen bespreken, zo zijn we terecht gekomen bij spraaktechnologie omdat dit in de toekomst zeker een belangrijke rol zal spelen in het uitvoeren van opdrachten in de spaceshuttle.

Voor dit werk hebben wij op de nodig steun kunnen rekenen van Mvr. Degryse en Mr. Verhaeghe. Voor de informatie zijn wij ten rade geweest in de bibliotheek te Brugge en te Torhout, bij Lernout & Hauspie en op het internet.

Hierbij danken wij iedereen die ons bij deze geïntegreerde proef heeft geholpen en in het bijzonder Mvr. Degryse en Mr. Verhaeghe voor de begeleiding.

Inhoudstabel

Hoofdstuk 1. Geluid

| | |
|--|---------|
| 1.1 Inleiding | p 1 |
| 1.1.1 Wat is geluid | p 2 |
| 1.2 Geluidsgolven | p 3-4 |
| 1.2.1 Resonantie | p 4-5 |
| 1.3 Geluidsgolven opvangen | p 5 |
| 1.3.1 Oren | p 6-7 |
| 1.3.2 Microfoons | p 7 |
| 1.4 Geluidsgolven voortbrengen | p 8 |
| 1.4.1 De stem | p 8 |
| 1.4.2 Luidsprekers | p 9 |
| 1.4.3 Muziekinstrumenten | p 9-10 |
| 1.4.3.1 Snaarinstrumenten | p 11 |
| 1.4.3.2 Blaasinstrumenten | p 12 |
| 1.4.3.3 Slaginstrumenten | p 12 |
| 1.5 Geluid opnemen en afspelen | p 13 |
| 1.5.1 Geluid op band | p 13 |
| 1.5.1.1 Elektriciteit en magnetisme | p 13-15 |
| 1.5.1.2 Geluid op band | p 16-17 |
| 1.5.1.3 Het magnetisatieproces | p 17-18 |
| 1.5.1.4 Vierkantsverhouding | p 19 |
| 1.5.1.5 Vormingsvrije opname | p 20 |
| 1.5.1.6 Geluidsband voor spoel en cassette | p 21 |
| 1.5.1.7 De opbouw van de band | p 21 |
| 1.5.1.8 De banddikte | p 22-23 |
| 1.5.2 Geluid op cd | p 24 |
| 1.5.2.1 Inleiding | p 24 |
| 1.5.2.2 Analoog en digitaal opnemen | p 24-25 |
| 1.5.2.3 Wat is het dat de naald van Edison's fonograaf | p 25-29 |
| 1.5.2.4 De cd verstaan | p 30 |
| 1.5.2.5 De cd-speler | p 31 |
| 1.5.2.6 DVD: de nieuwe generatie van dataopslag | p 31 |

Hoofdstuk 2. Spraaktechnologie

| | |
|--|-------------|
| 2.1 Principe | p 32 |
| 2.1.1 Waarom? | P 32 |
| 2.1.2 Het nut van spraaktechnologie | p 32 |
| 2.1.3 Oplijning | p 33 |
| 2.1.4 Hypothesen | p 34 |
| 2.1.5 Opel astra | p 35 |
| 2.1.6 Krant | p 35 |
| 2.1.7 PIN-code | p 36 |
| 2.1.8 Wat is spraaksynthese? | p 37 |
| 2.1.8.1 Allofoonspraak | p 37 |
| 2.1.8.2 Difoonspraak | p 37 |
| 2.1.8.3 Aaneengeschakelde zinnen | p 37 |
| 2.1.9 Verschillende types van spraakherkenning | p 38 |
| 2.2 Werking | p 38 |
| 2.2.1 Fonetische engines | p 38 |
| 2.2.2 Functionele eisen | p 39 |
| 2.2.2.1 Dialoogstructuur | p 39 |
| 2.2.2.2 Softwarestructuur | p 39 |
| 2.2.2.3 Afhandeling van fouten | p 40 |
| 2.2.3 De structuur van de stemcomputer... | p 40 |
| 2.2.3.1 Stemmodule | p 40 |
| 2.2.3.2 Dialoogmodule | p 40 |
| 2.2.3.3 Communicatiemodule | p 41 |
| 2.2.3.4 Dialoogconcept | p 41 |
| 2.2.4 Lexicon-, uitspraak- en ... | p 42 |
| 2.2.4.1 Het lexiconmappingbestand | p 43 |
| 2.2.5 Programma verloop | p 44 |
| 2.2.5.1 Programmafases | p 44 |
| 2.2.5.2 Programmaverloop | p 44 |
| 2.2.5.3 Communicatie tussen modules | p 45 |
| 2.2.5.4 Lexiconbestanden | p 45 |
| 2.2.6 Globale variabelen | p 47 |
| 2.2.6.1 Normale condities | p 47 |
| 2.2.6.2 Initialisatie | p 47 |
| 2.2.6.3 Dialoogafhandeling | p 47 |
| 2.2.7 Praktische kant | p 49 |
| 2.2.7.1 Instelling van de apparatuur | p 49 |
| 2.2.7.2 Trainingsfase | p 50 |

Hoofdstuk 3. Toepassingen en toekomst

| | |
|-------------------------------------|------|
| 3.1 Voxtron Flanders | p 51 |
| 3.2 SyVox | p 53 |
| 3.3 Softamend nv. | P 54 |
| 3.3.1 Inleiding | p 54 |
| 3.3.2 een korte voorstelling | p 55 |
| 3.3.2.1 Het actiedefinitie veld | p 55 |
| 3.3.2.2 De gebruikers defenitie | p 56 |
| 3.3.2.3 Alarm prioriteit | p 56 |
| 3.3.2.4 Flexibele systeem setup | p 57 |
| 3.3.2.5 Systeem configuratie | p 57 |
| 3.3.3 Logboek | p 58 |
| 3.3.4 Technische vereisten | p 58 |
| 3.4 Creative labs, Incorporated | p 59 |
| 3.5 Creator Ltd. | p 60 |
| 3.6 Probleemstellingen | p 61 |
| 3.7 Voordelen van spraaktechnologie | p 62 |
| Besluit | |

Hoofdstuk 1. Geluid

1.1 Inleiding

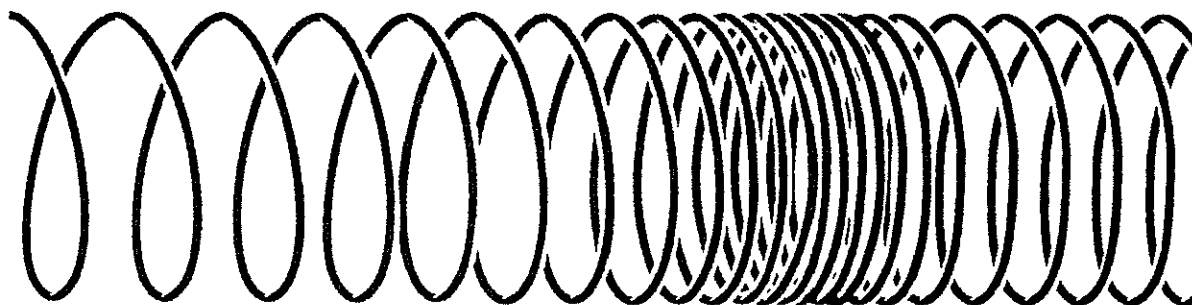
Geluid is zodanig verweven met ons dagelijkse leven dat we er niet meer bij stilstaan. Toch is het uiterst nuttig. We doen een beroep op geluid voor communicatie door middel van spraak en muziek. Onverwachte geluiden waarschuwen ons vaak voor gevaar. En zelfs zwakke achtergrondgeluiden kunnen ons heel wat vertellen over onze omgeving.

Ook dieren maken gebruik van geluid om te communiceren. Heel wat dieren brengen een specifiek geluid voort om andere dieren te waarschuwen of om de aandacht te trekken van een soortgenoot. Vleermuizen en dolfijnen maken gebruik van geluiden om zich te oriënteren. Er bestaat zelfs een soort garnaal die geluid gebruikt om zijn prooi te verdoven.

We vinden de wereld een erg lawaaiërig plaats, en toch hoort de mens slechts een klein gedeelte van alle geluiden, namelijk van 20 tot 20.000 Hz (zie later). Dus heel wat geluiden die dieren voortbrengen, vallen buiten ons gehoorvermogen. Het is evenwel ook zo dat niet alle dieren in staat zijn alle geluiden te horen die de mens maakt. Elk levend wezen heeft zijn eigen gehoorvermogen, aangepast aan zijn omgeving en zijn levenswijze.

1.1.1 Wat is geluid?

Geluid ontstaat als een voorwerp begint te trillen: wanneer een trommelstok het oppervlak van een trommel raakt, dan trilt het soepel bewegende vlies van de trommel op en neer waardoor de lucht erboven in beweging komt. Beweegt het vlies van de trommel naar omhoog, dan wordt de lucht erboven samengedrukt (of gecomprimeerd) en stijgt de luchtdruk. Beweegt het vlies naar beneden, dan zet de lucht uit en daalt de druk. De drukverschillen die zo in de lucht boven het vlies van de trommel ontstaan, planten zich in alle richtingen voort, zoals golven in een vijver: het zijn geluidsgolven.



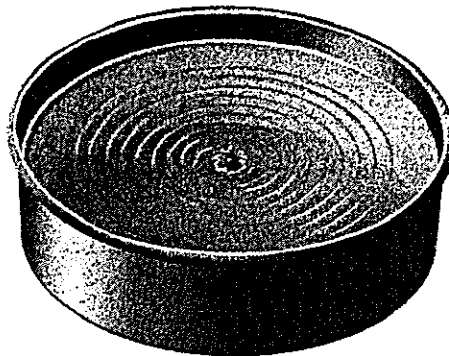
Geluid is een vorm van energie. Geluidsgolven brengen geluidsenergie van de ene plaats naar de andere, maar daarvoor hebben ze stofdeeltjes nodig, zoals atomen en moleculen. Die zorgen ervoor dat de energie zich kan verplaatsen. Geluidsgolven brengen stofdeeltjes immers aan het trillen zodat deze deeltjes tegen elkaar aanbotsen en zo geluidsenergie overdragen. Dat betekent dat geluid zich kan voortbewegen door lucht, vloeistoffen en vaste voorwerpen. Geluid kan zich evenwel niet voortbewegen door een luchtledige ruimte omdat een dergelijke ruimte geen deeltjes bevat (vb. zie het heelal = muisstil).

De snelheid waarmee geluid zich van de bron weg beweegt, noemt men de geluidssnelheid. Die hangt af van de materie waar het geluid doorheen beweegt en ook de temperatuur van die materie. Zo bedraagt ze 344 m/s in lucht van 20°C en 1437 m/s in water van 15°C. De snelheid ligt hoger in vloeistoffen omdat die meer deeltjes bevatten om te trillen. In vaste voorwerpen ligt de snelheid zelfs nog hoger.

Geluid kan niet met de ogen worden waargenomen, maar er bestaat een toestel waarmee men geluid zichtbaar kan maken nl. de oscilloscoop. Eerst vangt een microfoon (zie later) de geluidsgolven op en zet de trillingen ervan om in elektrische signalen. Op zijn beurt vangt de oscilloscoop deze elektrische signalen op en geeft ze op een scherm weer als een opeenvolging van golven. De vormen van de geluidsgolven kunnen dan worden gemeten en opgenomen. Met behulp van Fourieranalyse kan men dan de grondtonen scheiden van de boventonen (zie later).

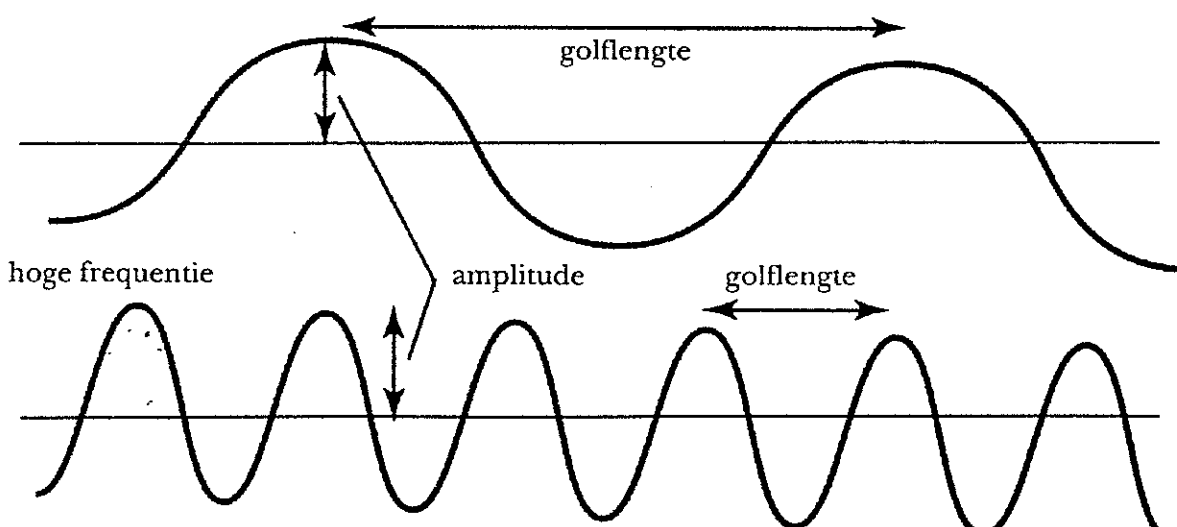
1.2 Geluidsgolven

Als je een steen in een vijver of een meer gooit, verschijnt er een kring kleine golven op het oppervlak van het water. De golven verplaatsen zich steeds verder weg van de plaats waar de steen in het water terechtkwam. Hoe groter de steen, hoe groter de golven. Geluidsgolven die van een geluidsbron weg bewegen, vertonen een soortgelijk patroon. Geluidsgolven kun je dan ook best vergelijken met die rimpelingen op het wateroppervlak.



Er zijn een aantal technische termen in verband met geluidsgolven. De golflengte is de afstand tussen het top- of dalpunt van opeenvolgende golven. De hoogte van de golf noemt men de amplitude. Geluidsgolven met een hogere amplitude klinken luider, aangezien ze meer geluidsenergie dragen.

De snelheid van het geluid wordt gemeten door het aantal golven te tellen die in een seconde voorbijkomen. Men noemt dit de frequentie (in golven per seconde). Hoe groter het aantal golven per seconde, hoe groter de frequentie. Kleinere voorwerpen kunnen sneller trillen dan grote en kunnen zo geluiden met een hogere frequentie voortbrengen. Geluiden met een hoge frequentie bezitten een hoge toonhoogte, geluiden met een lage frequentie bezitten een lage toonhoogte. De frequentie wordt gemeten in Hertz (Hz) wat het aantal geluidsgolven per seconde weergeeft.



De geluidssnelheid hangt af van het materiaal waar het geluid doorheen beweegt. Geluid beweegt zich sneller voort in vloeistoffen dan in lucht, omdat de deeltjes in een vloeistof talrijker zijn en dichter bij elkaar liggen. Dat betekent dat de geluidsenergie sneller kan worden verplaatst. Vaste stoffen bevatten zeer veel deeltjes die bovendien heel dicht tegen elkaar liggen. Bijgevolg kunnen de geluidsgolven zich nog sneller door vaste stoffen bewegen dan door vloeistoffen of lucht. De snelheid van het geluid door glas bedraagt bijvoorbeeld 5000 m/s De snelheid van lucht 344 m/s en die heeft een speciale naam: Mach 1.

We zullen niet verder ingaan op het begrip Mach omdat dit niet zo veel meer met spraaktechnologie te maken heeft.

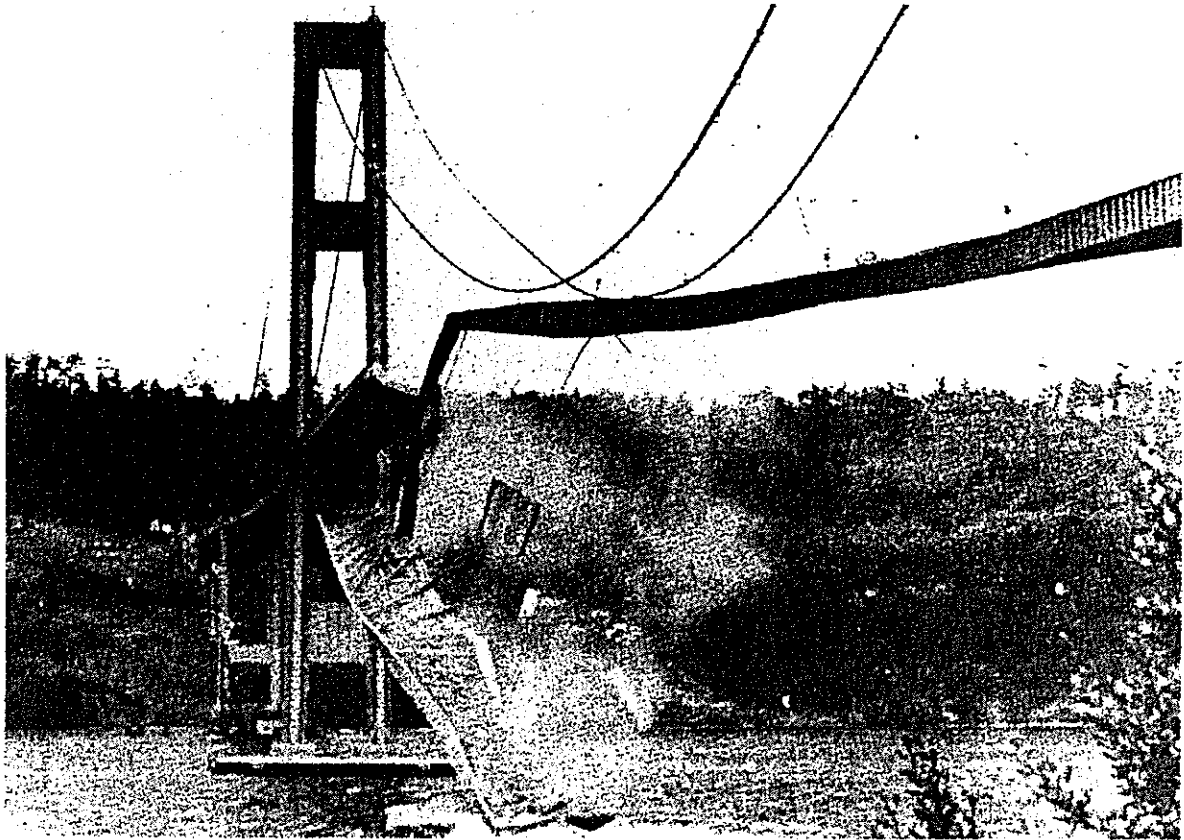
1.2.1 Resonantie

Wanneer voorwerpen trillen, dan is er één bepaalde frequentie waarbij dat het gemakkelijk gaat. Als een voorwerp vrij mag trillen, dan trilt het met zijn eigenfrequentie. Deze frequentie hangt af van de vorm en het materiaal waaruit het voorwerp is vervaardigd. Wanneer je bijvoorbeeld tegen een kleine bel tikt met haar eigenfrequentie, dan ontstaat er een galmend geluid, de grondtoon. Een grotere bel die uit hetzelfde materiaal is vervaardigd en dezelfde vorm heeft, heeft een lagere eigenfrequentie en bijgevolg een grondtoon met een lagere frequentie.

Resonantie komt tot stand wanneer een voorwerp energie opvangt met dezelfde frequentie als zijn eigenfrequentie. Je kunt dat best vergelijken met het duwen van iemand op een schommel. Om een goede schommelbeweging te krijgen moet diegene die duwt elke duw zorgvuldig timen. Als men niet op het juiste ogenblik duwt, ontstaat er geen goede oscillatie (schommeling). De persoon die duwt, beperkt dan enkel de schommelbeweging. Hij moet de frequentie van zijn duwen aanpassen aan de frequentie van de bewegende schommel.

Resonantie kan een verrassend krachtig effect hebben. Ze is heel belangrijk bij het bespelen van muziekinstrumenten. Als een muzikant een noot speelt met dezelfde grondtoon als dat van een glas op een tafel in de nabijheid, dan begint het glas te trillen doordat het de geluidsenergie van de omringende trillende lucht opvangt. Er ontstaat resonantie als het instrument met exact dezelfde frequentie trilt, waardoor de lucht rondom het glas met de eigenfrequentie van het glas begint te trillen.

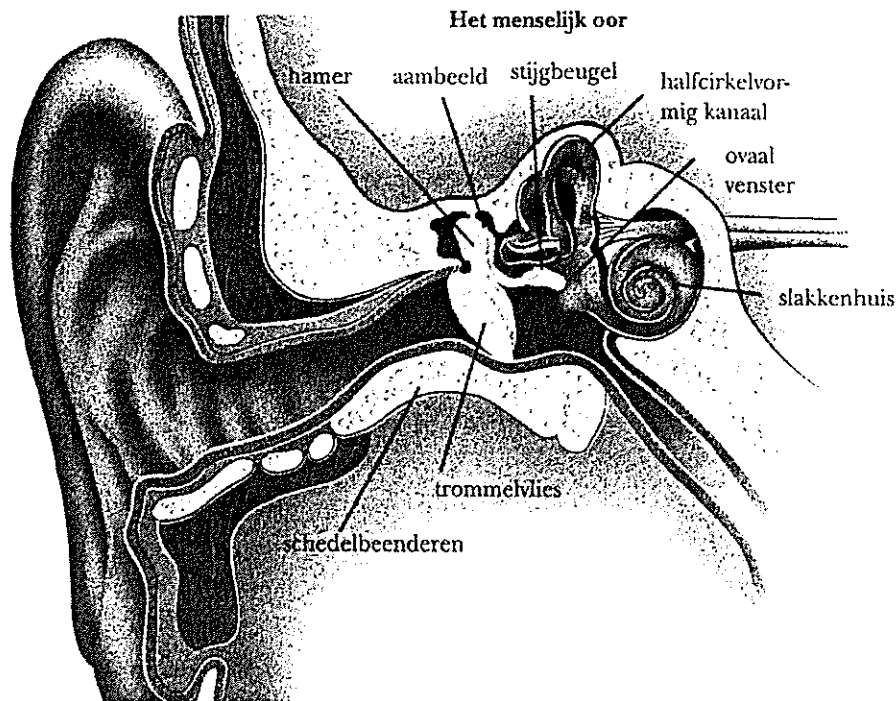
De meeste bouwwerken resoneren met één of meerdere frequenties (dat wil zeggen dat ze meer dan één frequentie hebben). Resonantie kan uitzonderlijk veel schade aanrichten in onvoorziene omstandigheden. Vaak krijgen bruggen een speciale aërodynamische vormgeving, waardoor energie geen luchtstromen in beweging kan brengen en er op het bruggendeck dus geen resonantie kan ontstaan. Het effect van de resonantie zou er immers kunnen toe leiden dat de brug begint te wiegen en uiteindelijk zelfs instort.



1.3 Geluidsgolven opvangen

Geluidsgolven dragen niet zoveel geluidsenergie mee en in lucht planten ze zich niet voort. Daarom ontwikkelden vele dieren gedurende miljoenen jaren speciale organen om hen te helpen zwakke geluiden te ontdekken en te versterken. Deze geluiden zetten ze dan om in signalen die naar de hersenen worden gestuurd. Het oor is één van de resultaten van deze langzame evolutie. De mens heeft ook elektronische toestellen ontwikkeld die geluidsgolven kunnen opvangen, versterken en overbrengen. Dankzij dergelijke uitvindingen kunnen wij op een snelle en duidelijk verstaanbare wijze met elkaar communiceren over verre afstanden.

1.3.1 Oren



De oren van zoogdieren zijn geëvolueerd tot uiterst gespecialiseerde organen die geluidsgolven opvangen en versterken vooraleer ze door te sturen naar de hersenen. Dergelijke oren hebben alle een zelfde bouw, hoewel er bepaalde aanpassingen zijn volgens de levenswijze van elk type zoogdier.

Een oor bestaat uit drie delen: buiten-, midden- en binnenoor. Het buitenoor is dat deel dat we aan de buitenkant van het hoofd kunnen zien (de oorschelp). De wetenschappelijke naam is 'pinna'. De functie van de pinna bestaat erin de geluidsgolven op te vangen en ze naar het middenoor te leiden. De vorm van de oorschelp verschilt van zoogdier tot zoogdier.

Omhullende beentjes beschermen het middenoor. Het buiten- en middenoor zijn van elkaar gescheiden door een membraan: het trommelvlies of tympanum. In het middenoor bevinden zich drie kleine beentjes: de hamer, het aambeeld en de stijgbeugel (kleinste beentje van het lichaam). Samen vormen ze de gehoorbeentjes. Deze beentjes hebben een heel precieze positie in het oor en worden door spieren op hun plaats gehouden. Wanneer geluidsgolven via de oorschelp het trommelvlies bereiken, begint het trommelvlies te trillen. Deze trilling brengt de hamer in beweging. Die duwt op zijn beurt tegen het aambeeld, dat dan ook begint te trillen. Ten slotte bereiken de trillingen de stijgbeugel. Alle gehoorbeentjes samen zorgen ervoor dat de trillingen ongeveer 22 keer worden versterkt. Er is geen geluidsverlies, aangezien de trillingen van beentje tot beentje worden doorgegeven. De trillende stijgbeugel raakt een membraan, het ovaal venster, dat de scheiding vormt tussen midden- en het binnenoor.

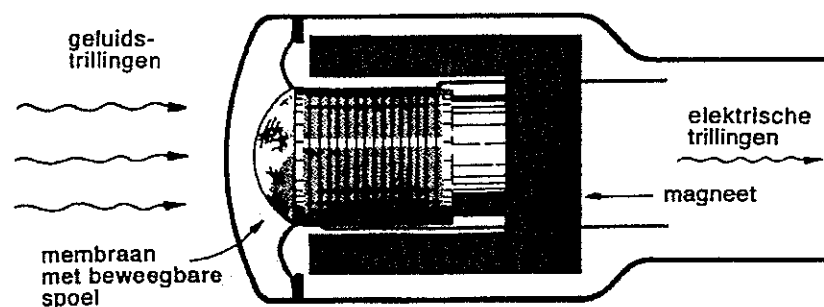
Het binnenoor verschilt duidelijk van het midden- en het buitenoor. Het is namelijk niet gevuld met lucht, maar met vloeistof. Het bevat twee organen: één voor het gehoor en één voor het evenwicht. Het orgaan voor het gehoor heet het 'slakkenhuis', een gekronkelde en met vloeistof gevulde structuur met receptoren (voelorgaantjes). Deze receptoren zijn verbonden met zenuwen. Wanneer het membraan van het ovaal venster begint te trillen, dan drukt het op de vloeistof in het slakkenhuis die ook begint te trillen. De vloeistof in het slakkenhuis drukt op de receptoren, die op die manier via het zenuwstelsel berichten doorgeven aan de hersenen.

De tweede structuur in het binnenoor is die van de halfcirkelvormige kanalen. Dat zijn uiterst gespecialiseerde sensoren die de hersenen informatie doorgeven over de houding en de beweging van het hoofd. Er zijn drie halfcirkelvormige kanalen die in een rechte hoek ten opzichte van elkaar staan zodat de hersenen bewegingen in alle richtingen kunnen onderscheiden. Beweegt het hoofd van links naar rechts, dan begint de vloeistof in één halfcirkelvormige kanaal te bewegen. De bewegende vloeistof drukt op tasharen die in het halfcirkelvormige kanaal liggen. Deze haren zenden informatie naar de hersenen over de houding van het hoofd. De hersenen gebruiken deze informatie om het evenwicht en de lichaamshouding te controleren.

1.3.2 Microfoons

Wetenschappers bestempelen microfoons als energieomvormers. Hiermee wordt bedoelt dat ze geluidsenergie omzetten in elektrische energie. Deze omzetting steunt op het principe van de inductie. Een elektrische stroom wordt in een stuk draad geïnduceerd (met andere woorden: er ontstaat een elektrische stroom in de draad) als dat stuk draad in en uit een magnetisch veld wordt bewogen. Hoe sneller de draad wordt bewogen, hoe groter de opgewekte stroom.

De geluidstrillingen brengen een membraan in trilling. Daar is een licht spoeltje van dun, geïsoleerd koperdraad aan bevestigd dat in hetzelfde ritme meebeweegt. Het spoeltje is in een magnetisch veld opgehangen. Daardoor wordt er een elektrisch kleine spanning in opgewekt waarvan de polariteit dus in sterkte evenredig is aan de bewegingen van het membraan en de geluidstrillingen. Zo wordt de kinetische energie (de beweging) omgezet in elektrische energie. De in de spoel opgewekte elektrische stroom wordt via verbindingsdraden naar de versterker gestuurd, en vervolgens naar de luidspreker.



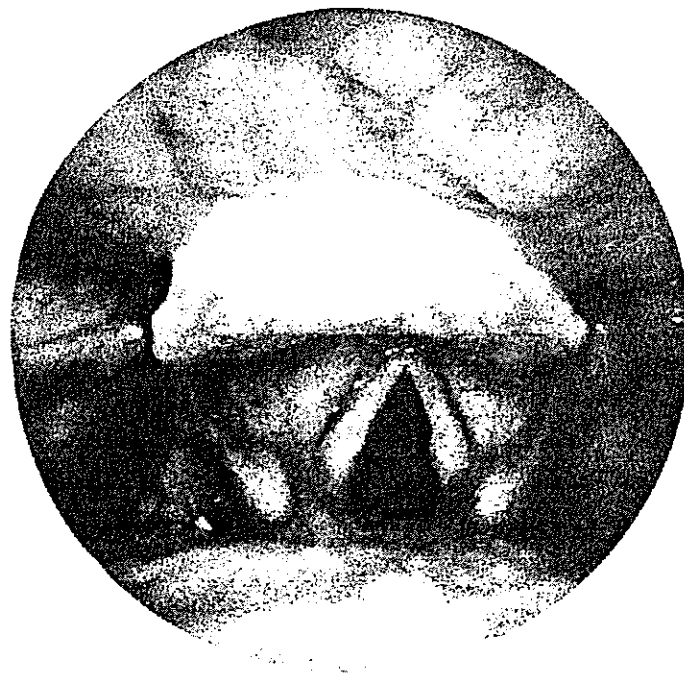
Het principe van de veelgebruikte magneto-dynamische microfoon.

1.4 Geluidsgolven voortbrengen

Veel dieren gebruiken geluiden om met elkaar te communiceren. Er zijn overigens niet veel verschillen tussen de manieren waarop de verschillende dieren geluiden maken. De mens van zijn kant heeft geleerd om kunstmatige geluiden te produceren door middel van een hele waaier instrumenten en toestellen. Niettemin ontstaan ook deze geluiden op een manier sterk die sterk lijkt op de methodes die men in de natuur aantreft.

1.4.1 De stem

De menselijke stem wordt voortgebracht door het strottenhoofd, dat is een op een doosje gelijkende holle structuur in de luchtpijp die je vooraan de hals kunt zien en voelen en die vaak de adamsappel wordt genoemd. De wetenschappelijke naam is 'larynx'. Het strottenhoofd bestaat uit twee stellen plooiën die trillen als er lucht doorstroomt. De trillingen brengen geluiden voort. Als je je strottenhoofd aanraakt en spreekt, voel je de trillingen. Het ene uiteinde van elke stemband is vooraan in het strottenhoofd vastgemaakt, terwijl het andere uiteinde is vastgemaakt aan één van de twee beweegbare kraakbeenderen in het strottenhoofd zelf. Als je inademt, worden de stembanden uit elkaar getrokken zodat de lucht naar de longen kan. Als je uitademt zonder te spreken, blijven de stembanden uit elkaar. Om te spreken of te zingen, worden de twee kraakbeenringen tijdens uitademen dicht bij elkaar gebracht. Daardoor wordt de luchtpijp gedeeltelijk versperd en moet de lucht tussen de stembanden worden geduwd, zodat die beginnen te trillen. De toonhoogte van de stem hangt af van de lengte en dikte van de stembanden en van de grootte van het strottenhoofd. Mannen hebben een diepere stem dan vrouwen omdat zij een groter strottenhoofd hebben met langere en dikkere stembanden. De nauwkeurige controle over onze stembanden stelt ons in staat heel wat geluiden te produceren. Als een kind leert spreken, leert het niet alleen een taal gebruiken, maar ook hoe het zijn stembanden moet gebruiken om het gewenste geluid te maken.



1.4.2 Luidsprekers

Een luidspreker doet eigenlijk het tegenovergestelde van een microfoon. Een elektrische stroom rond een draad in een magnetisch veld zorgt ervoor dat elektrische energie wordt omgezet in kinetische energie. De elektrische energie stroomt door een spoel die rond een buisje is gewikkeld. Dat buisje is vastgemaakt aan een papieren conus (trechtersvormig membraan) die in een magnetisch veld wordt gehouden. Elektrische energie die door een spoel in een magnetisch veld stroomt, wordt omgezet in kinetische energie, waardoor de spoel in beweging komt en de papieren conus voorwaarts en achterwaarts duwt. De bewegingen van de conus veroorzaken drukgolven in de lucht, die worden waargenomen als geluid. Het geluid is sterker, aangezien de versterker het signaal van de microfoon krachtiger heeft gemaakt. Hoe krachtiger het signaal, hoe meer de conus beweegt.

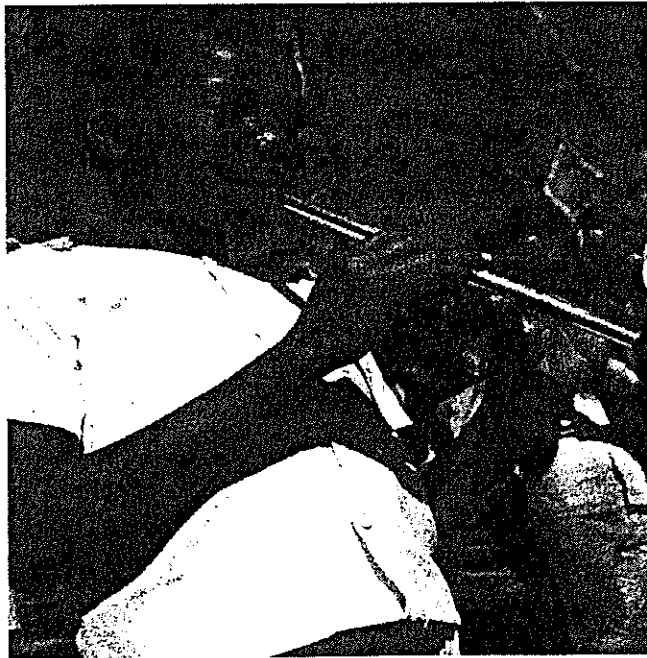
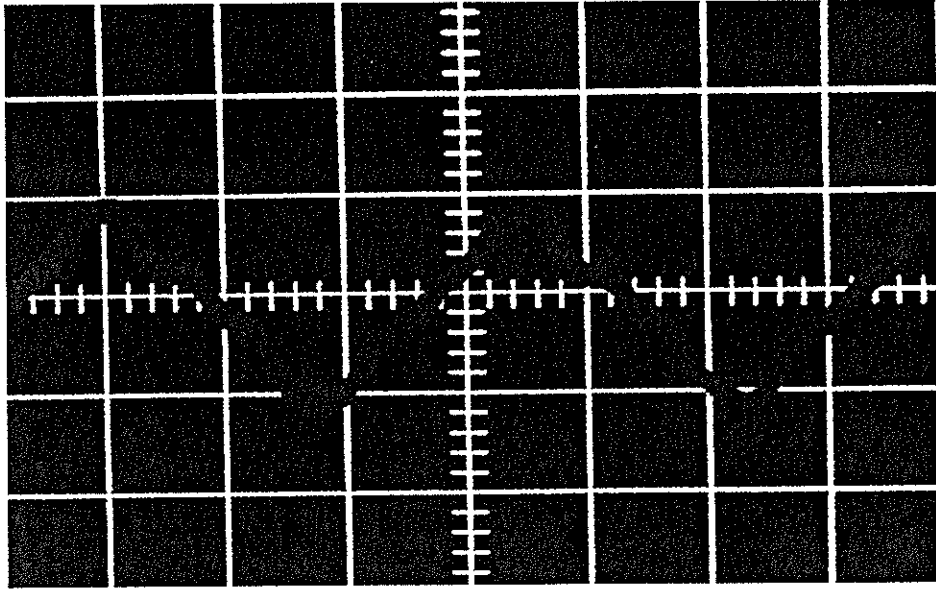
1.4.3 Muziekinstrumenten

Muzieknoten worden net als alle geluiden geproduceerd door drukgolven van trillende voorwerpen. Muziekinstrumenten kunnen vervaardigd zijn uit hout, metaal, snaren en pijpen die een trillende luchtkolom bevatten. Noten van muziekinstrumenten zijn gebaseerd op octaven. Telkens de toonhoogte van een geluid één octaaf verhoogt, verdubbelt de frequentie ervan. De meeste instrumenten vereisen ook een middel om de geproduceerde geluiden te versterken. Als je eenvoudigweg tokkelt op een vioolsnaar, maakt dat niet veel geluid. Is deze snaar evenwel over een holle doos gespannen, dan is het geluid veel sterker, want zowel de doos als de lucht binnenin de doos trillen in harmonie met de snaar. De resonantie versterkt het geluid.

Grote muziekinstrumenten produceren lage noten, net zoals grote dieren, lage luide geluiden voortbrengen. Zo kan een klein kind veel hogere noten zingen dan een volwassene. Een zelfde ontwikkeling stellen we vast van de viool via altviool naar de cello. Een blaasinstrument, zoals een dwarsfluit, heeft een aantal gaatjes over de lengt van de buis. Bij het bespelen van het instrument worden sommige gaatjes bedekt en dan weer losgelaten. Hierdoor verandert de frequentie waarmee het instrument resoneert, en zo ontstaan er telkens andere noten.

| Namen | Do | re | mi | fa | sol | la | si | do |
|-----------------------------------|-------|--------|---------|--------|--------|--------|---------|------|
| Interval tussen een toon en de do | f | $9/8f$ | $5/4f$ | $4/3f$ | $3/2f$ | $5/3f$ | $15/8f$ | $2f$ |
| Interval tussen twee tonen | $9/8$ | $10/9$ | $16/15$ | $9/8$ | $10/9$ | $9/8$ | $16/15$ | |

Resonantie kan voorkomen wanneer de frequentie van de trillingen een fractie of een veelvoud van de eigen frequentie bedraagt. De noten die deze resonantie voortbrengen, zijn hoger dan de gespeelde noot en worden bovendien boventonen genoemd (zie later Fourier-analyse)



1.4.3.1 Snaarinstrumenten



Bij snaarinstrumenten wordt klank geproduceerd door op de snaren te tokkelen, te slaan of door ze te buigen. Zo maakt een gitaarspeler geluid door te tokkelen op de snaren, ofwel met zijn vingers, ofwel met een plectrum, een klein stukje plastic dat de nagels spaart. De lengte van de snaar wordt bepaald door de vingers van de speler, die de snaar tegen verhoogde metalen kammetjes op de hals van de gitaar drukken. Deze kammetjes worden 'frets' genoemd. De toonhoogte van de noten hangt af van de spanning die de speler aan de snaar geeft en van de lengte en de dikte ervan. Wordt de spanning op de snaar verhoogd (wordt er harder aan getrokken), dan zal de snaar sneller trillen en zal de toonhoogte hoger zijn. Een dikke snaar brengt lage noten voort. De snaren van een gitaar zijn zo ontwikkeld dat ze bij de juiste spanning zes noten kunnen spelen: mi, la, re, sol, si en de volgende mi.

Elke muzieknoot bestaat uit een grondtoon en een aantal hogere tonen, de zogenaamde boventonen. Op een gitaar heeft elke snaar die niet wordt opgespannen, haar eigen grondtoon en een eigen reeks boventonen. Het oor hoort een combinatie van die tonen, maar de grondtoon is de sterkste. Door op verschillende plaatsen op de snaar te drukken, kan de speler verschillende grondtonen vormen, die elk hun eigen reeks boventonen hebben. De combinatie van grondtonen en boventonen zorgt voor de typische klank (geluidskwaliteit) van elk instrument.

1.4.3.2 Blaasinstrumenten



Blaasinstrumenten werken volgens een ander principe dan snaarinstrumenten. Er wordt lucht in geblazen en de luchtkolom die in het instrument is ingesloten, begint te trillen. De frequentie van het geproduceerde geluid wordt bepaald door de lengte van de buis. De langste blaasinstrumenten produceren de laagste noten. Een korte luchtkolom, zoals die van een dwarsfluit, trilt sneller en brengt dus hogere tonen voort dan instrumenten met een grotere luchtkolom, zoals een hobo. In een trombone gebeurt dat door een U-vormige buis in en uit te schuiven, waardoor de luchtruimte inderdaad wordt verkleind of vergroot. De meeste blaasinstrumenten daarentegen hebben een buis met vaste lengte, zo bijvoorbeeld de blokfluit of de klarinet. De lengte van de buis wordt gewijzigd door de gaatjes in de wand te sluiten of te openen, waardoor de correcte muzieknoot wordt geproduceerd.

1.4.3.3 Slaginstrumenten

Slaginstrumenten behoren tot de oudste muziekinstrumenten. Ze komen overal ter wereld voor, in diverse vormen en maten. De bouw is bijzonder eenvoudig: een holle doos waar een vel over is gespannen. De trommelslag ontstaat door met de hand of met een trommelstok op het vel te slaan, waardoor het vel en de lucht in de doos aan het trillen gaan. De toon van de geproduceerde noot hangt van een aantal factoren af, onder meer van de grootte van het slagvel. Hoe groter het vel, hoe lager de noot. Ook de spanning van het vel beïnvloedt de noot. Een strakker gespannen vel zorgt ervoor dat de duur van de klank wordt beperkt en dat de klank iets hoger is. Ook de hoeveelheid lucht binnenin beïnvloedt de toon van de noot. Hoe meer lucht de trommel bevat, hoe lager de noot.

1.5 Geluid opnemen en afspelen

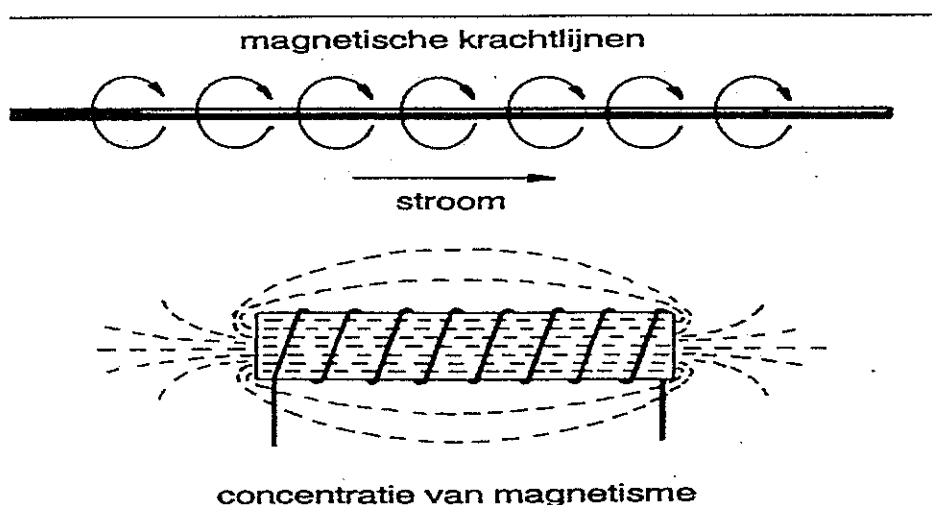
We weten dat geluiden eenvoudigweg een opeenvolging van drukgolven zijn die luchtdrukverschillen veroorzaken (zie uitleg geluid). Als men de wijzigende luchtdruk nu nauwkeurig kan opmeten, kan het geluid later opnieuw worden afgespeeld. Tot voor kort was dit een louter analogo gebeuren (zie uitleg cd). We vonden het daarom zeer belangrijk en nuttig om opname op band nog eens kort uit te leggen (we bespreken alleen de cassette omdat andere analoge opname technieken hierop gebaseerd zijn (zoals de videoband, de floppydisk, de diskette...)). Maar om de kwaliteit van de geluidswaergave te verbeteren, wordt er bij het opnemen en afspelen steeds vaker gebruik gemaakt van digitale technieken (zie uitleg cd).

1.5.1 Geluid op band

Nu nog even kort wat uileg over elektriciteit en magnetisme, dit om het opnemen op band beter te begrijpen.

1.5.1.1 Elektriciteit en magnetisme

De zojuist besproken microfoon (zie 1.3.2 microfoons) is gebaseerd op het even eenvoudige als belangrijke elektromagnetisch principe: waar elektriciteit is, is magnetisme. Zodra er een elektrische stroom door een draad vloeit, ontstaat rondom die draad een magneetveld. En andersom: zodra een elektrische geleider zoals een elektrische draad. Zich in een variërend magneetveld bevindt, wordt in die draad een elektrische spanning opgewekt.

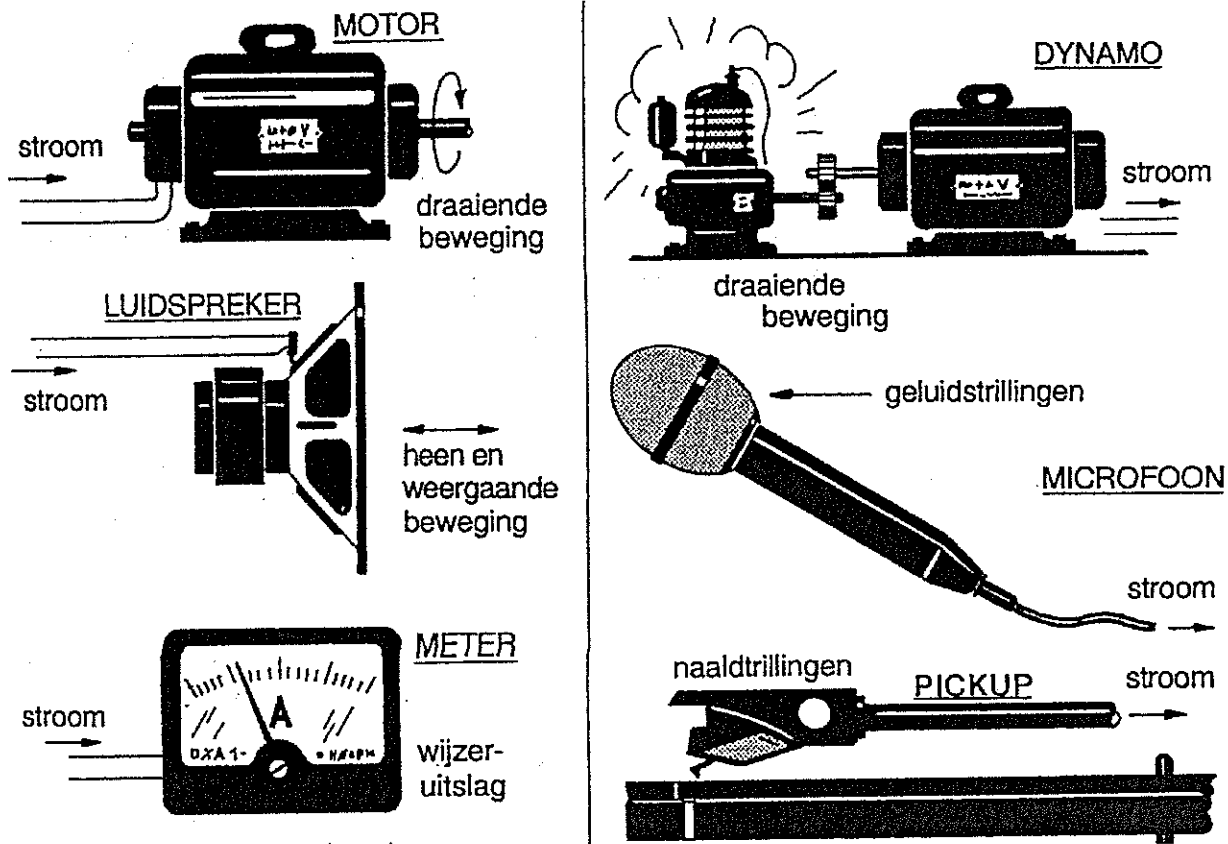


Waar stroom is, is magnetisme.

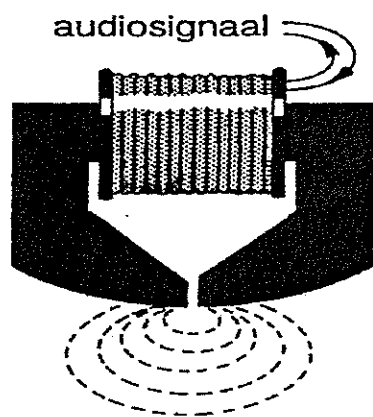
Let wel: het moet een variërend magneetveld zijn, want indien dit ook gold voor een vast magneetveld was er al stroom op te wekken louter door een draadje naast een magneet te leggen. Nee, de magneet moet wel in beweging zijn. Of de geleider, dat maakt niets uit. Het gaat er maar om, dat de een beweegt ten opzichte van de ander. Om een zo groot mogelijke draadlengte binnen de invloedssfeer van het magneetveld te kunnen brengen wordt een lange draad tot spoel opgerold. De zojuist besproken microfoon met zijn spoel in een magneetveld is daarvan een voorbeeld.

Het elektromagnetisch principe wordt zeer veel toegepast. Elektromotoren, dynamo's, microfoons, luidsprekers, de meeste pick-upelementen, magneetkoppen, relais, transformatoren, etc., zo zijn al deze componenten op dit principe gebaseerd. Deze componenten vormen ook een brug tussen de elektronenwereld en onze wereld van kracht, beweging, geluid en noem maar op. Dankzij het magnetisme heeft de mens toegang tot de elektronenwereld. De elektromotor levert als gevolg van tegengestelde magnetische krachten tussen vaste en draaibare spoelen kracht en de dynamo levert op basis van diezelfde krachten stroom. Omdat ze allebei op hetzelfde principe zijn gebaseerd, zijn ze ook andersom te gebruiken: de motor als dynamo en de dynamo als motor.

Net zoals de dynamo is de microfoon in feite een generator. Alleen worden nu geen draaiende, maar heen en weer gaande bewegingen in stroom omgezet. En zo is de luidspreker in feite een elektrische motor die heen en weer gaande bewegingen produceert. Ook deze componenten kunnen in principe andersom worden gebruikt: de microfoon als luidspreker en de luidspreker als microfoon. Want in beide gevallen hangt een beweegbaar spoeltje (bij de luidspreker spreekspoel geheten) in een krachtig magneetveld. De hoofdtelefoon werkt ook volgens dit principe. Bij de opname/ weergavekop van het cassettedeck zijn wel magnetische bewegingen in het spel. Bij opname wordt de elektrische trilling, het audiosignaal dus, aan een spoeltje toegevoerd dat is aangebracht om een rondgebogen kern van gemakkelijk magnetiseerbaar materiaal. De kern heeft aan de voorzijde een nauwe spleet, de kopspleet, waardoorheen het opgewekte magneetveld naar buiten treedt en, de erlangs lopende magnetiseerbare geluidsband magnetiseert. Bij weergave geschiedt de procedure precies andersom.



Enkele voorbeelden van omzetting van stroom in mechanische energie (motorprincipe) en van mechanische energie in stroom.



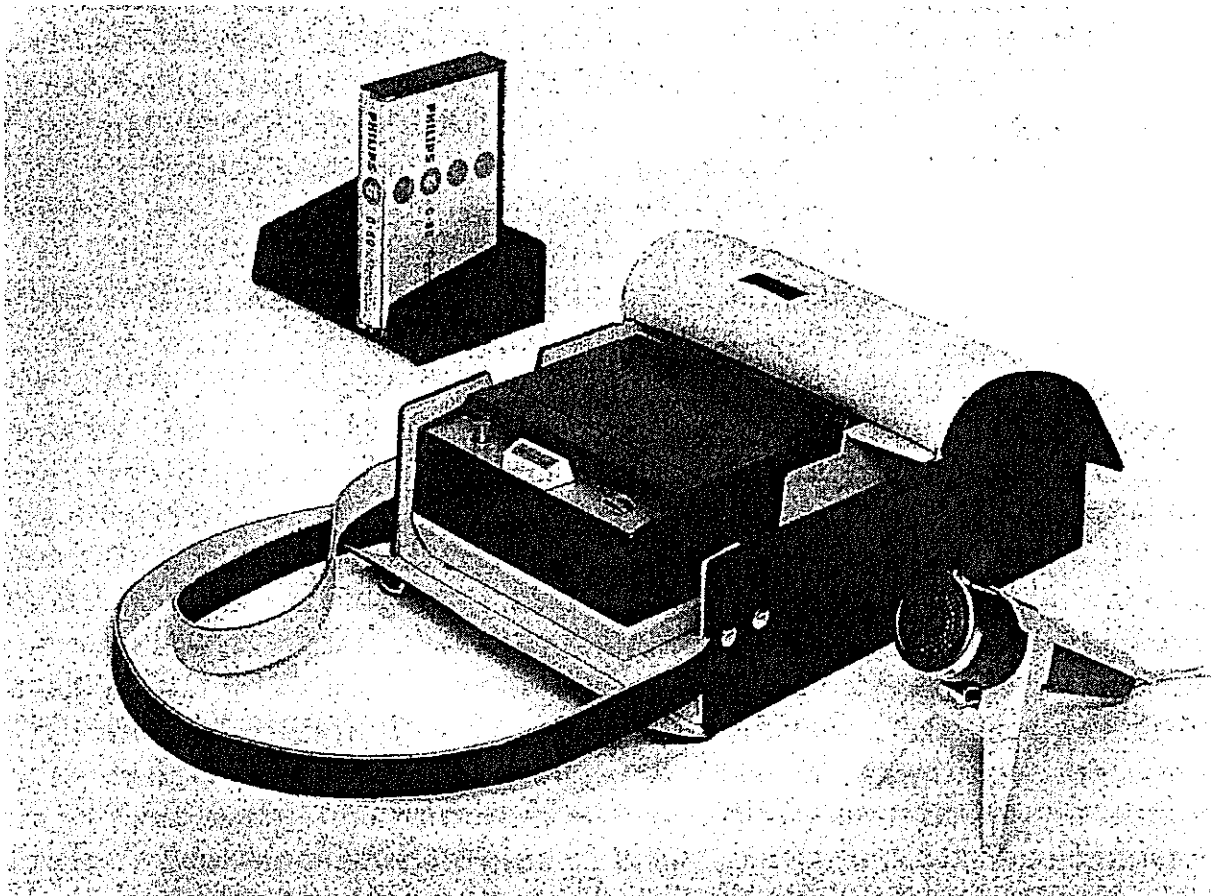
Bij de opname-weergave-kop wordt het in de kopkern opgewekte magnetisch veld bij opname via een smalle spleet aan de band overgedragen. Bij weergave wordt het bandmagnetisme via diezelfde spleet aan de kern overgedragen.

1.5.1.2 Geluid op band

Geluid op band, dat kan van alles zijn tegenwoordig: de band van de band- of spoelenrecorder, het cassetdeck, de videorecorder, de digitale DAT-recorder of de eveneens digitale DCC-recorder. Welk recordertype ook, het heeft altijd tot doel de audiosignalen van microfoon, tuner, Cd-speler, platenspeler en dergelijke op een magnetiseerbare band, kortweg de magneetband, vast te leggen om die op een willekeurig later tijdstip weer te geven.

Bij geluid op band had men eerst de spoelenrecorder die nu al van de markt verdwenen is, en daarna het cassetdeck. Tussen spoelenrecorder en cassetrecorder zijn er wel enkele essentiële verschillen, waardoor de spoelenrecorder toch wel zijn specifieke waarde heeft. Er is bijvoorbeeld bandmontage mogelijk, men kan op verschillende snelheden werken, Het begrip spoelenrecorder staat eigenlijk voor een apparaat met ingebouwde versterker en luidspreker(s). Op zeker moment kwamen er bandapparaten zonder eindversterker, ze moesten op een audio-installatie aangesloten worden. Deze werden tapedeck genoemd. Wanneer er niet met spoelen maar met cassettes werd gewerkt, heette dat apparaat cassetdeck. Dit wordt nu nog steeds gebruikt.

Het voordeel van een cassetdeck is het bedieningsgemak. Het eerste cassetdeck werd door Philips ontwikkeld en gepromoot in 1963 onder de klik-klak-klaar slogan en was meteen een wereldwijd succes.



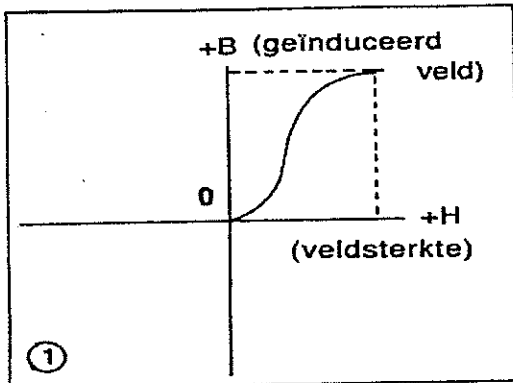
Eerste, door Philips gecommercialiseerde, cassetrecorder (EL3301)

Sinds de eerste cassetterecorder heeft het cassettemedium inmiddels een zeer hoge kwaliteit bereikt. Wanneer men een goed cassettedeck heeft merkt men nauwelijks een verschil tussen de opname en weergave. Ondanks de hoge kwaliteit zijn cassettes niet duur

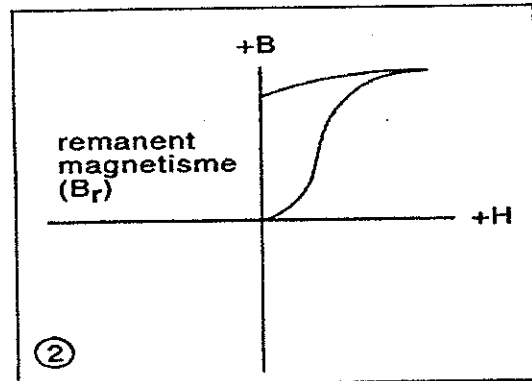
1.5.1.3 Het magnetisatieproces

Zonder voormagnetisatie zou de band niet alleen zeer zwak worden gemagnetiseerd, maar zou er ook ernstige vervorming ontstaan. Want het magnetisatieproces verloopt van nature allerm minst lineair (veranderd niet volgens een vaste constante). Hoe dit precies zit toont de speciale grafiek, de zogenoemde hysteresislus (is een speciale lus waarbij de waarden nooit meer in de oorsprong terug komen; zie onderaan), die vaak bij de technische gegevens van bandmateriaal wordt aangetroffen.

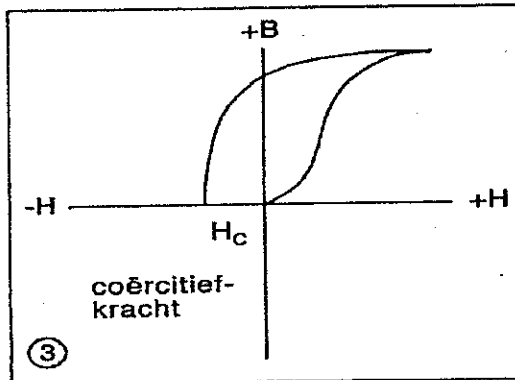
De hysteresislus of -kromme geeft over vier velden het verband aan tussen het opwekkende veld H (magnetische veldsterkte *) en het opgewekte veld B (het geïnduceerde veld **). Hysteresis betekent nawerking en inderdaad laat de hysteresislus zien hoe B altijd achterblijft. Altijd blijft er dus een zeker remanent (= blijvend) magnetisme achter. Het uitgangspunt 0 wordt niet meer bereikt; altijd zal B bij het sturen van een wisselstroom door de spoel die de veldsterkte H opwekt de buiten het nulpunt liggende lus doorlopen. Altijd, gedurende elke periode. Het eerste deel van de kromme, de maagdelijke kromme, wordt pas weer gepasseerd wanneer het magnetische materiaal op een of andere manier in de neutrale toestand is teruggebracht. Want dat gebeurt in de praktijk natuurlijk. We willen immers kunnen opnemen, maar ook kunnen wissen. De hysteresislus laat ook zien hoe de geleidelijk toenemende veldsterkte H geenszins nauwkeurig door het geïnduceerde veld B wordt gevolgd. Er ontstaat een allerm minst rechte lijn. Alleen in het midden is die recht. Alleen in dit kleine gebied is er sprake van een lineaire verhouding tussen H en B . Door bij het wissen een hoge wisselfrequentie toe te passen (de al eerder vermelde 100 kHz) wordt de band in hoog tempo van de ene in de andere verzadigingstoestand (toestand waarbij het maximum aan magnetisme bereikt is) gebracht. Tegelijkertijd wordt de band geleidelijk uit het veld weggetrokken, met als gevolg dat hij volkomen neutraal zijn weg vervolgt. Het nulpunt in de hysteresiskromme is op dat moment bereikt, de band is volledig gewist.



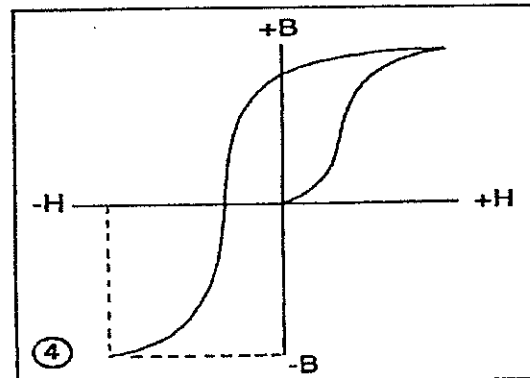
① De maagdelijke kromme. Er is geen lineair verband tussen H en B.



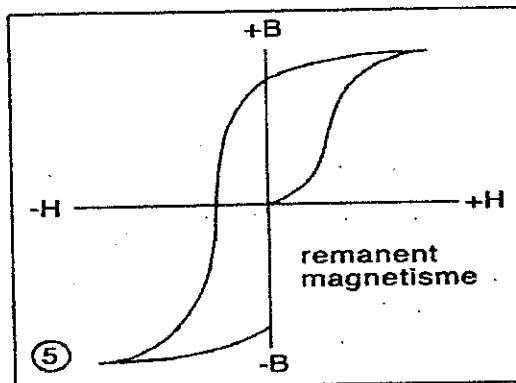
② H is tot 0 teruggebracht, er blijft remanent magnetisme achter.



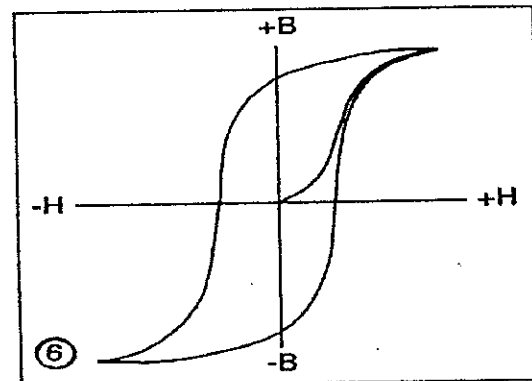
③ H wordt zoveel negatief gemaakt dat B tot 0 daalt. Deze negatieve waarde is de coërcitiefkracht H_c .



④ We laten H in negatieve richting toenemen. B neemt eveneens in negatieve richting toe.



⑤ H wordt weer naar 0 teruggebracht en weer blijft een zeker remanent magnetisme achter.



⑥ Nu laten we de veldsterkte in positieve richting toenemen. De lus sluit zich, er is een hysteresislus ontstaan.

Zo ontstaat de hysteresislus. De basis is een assenkruis dat in feite uit vier grafieken bestaat, waardoor zowel positieve als negatieve waarden kunnen worden uitgezet.

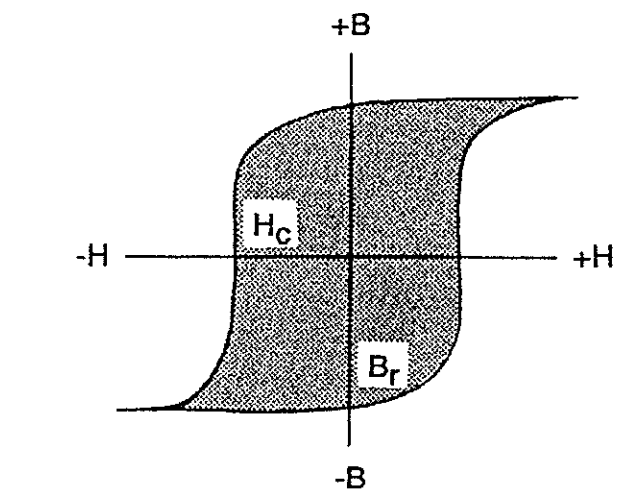
*magnetische veldsterkte (H) = De kracht per eenheid van magnetische lading die deze lading zou ondervinden indien men ze in het magnetisch veld zou brengen.

**geïnduceerd veld (B) = $\mu \cdot H \cdot \cos \alpha$

Met μ = permeabiliteit = een grootte die aangeeft hoe goed een bepaalde stof bepaalde magnetische veldlijnen doorlaat.

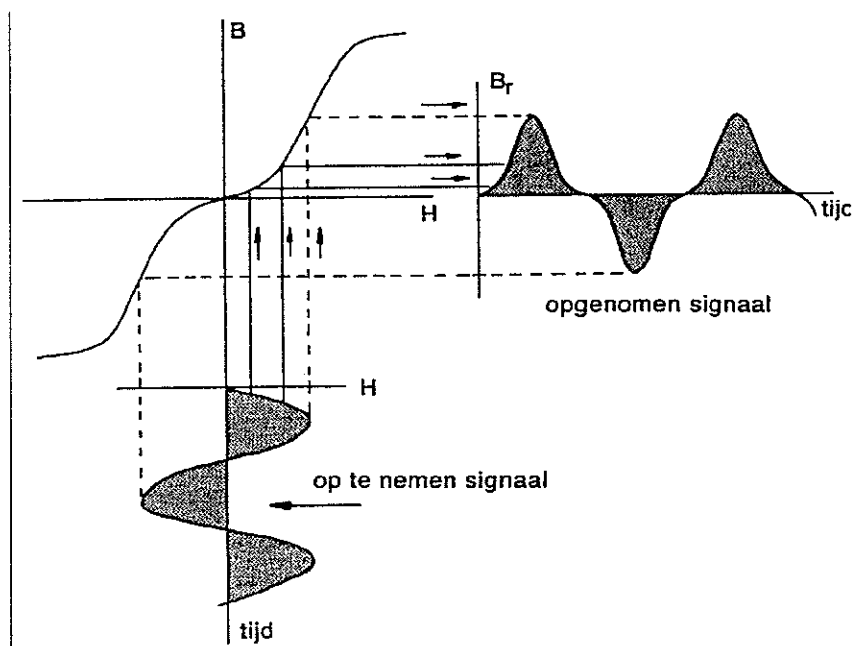
1.5.1.4 Vierkantsverhouding

Er is een zeker verschil tussen de maximaal mogelijke magnetische bandverzadiging en het magnetisme dat na het magnetiseren in het materiaal achterblijft, de remanentie dus. Hoe kleiner dat verschil hoe beter. Dit gegeven wordt als vierkantsverhouding aangegeven, wat te maken heeft met de vorm van de hysteresislus; als er geen verschil zou zijn, zou de vierkantsverhouding (tussen B en H) 1,00 bedragen. Dat is ondenkbaar. Een goede gemiddelde waarde is 0,80. Het praktische traject ligt tussen 0,75 en 0,88. Hoe hoger deze factor, hoe beter de opname-eigenschappen en hoe geringere vervorming. De vierkantsverhouding staat los van de grootte van de kromme. Het kan heel goed voorkomen dat de vierkantsverhouding van een goedkoop bandtype, zoals de ijzeroxidenband beter is dan van een dure metaalband.



1.5.1.5 Vormingsvrije opname.

Zodra we gaan proberen een audiosignaal op de band vast te leggen gaat het niet-lineaire gedrag van het magnetisatieproces ons parten spelen. Niet elk deeltje van het toegevoerde geluidssignaal wordt op exact dezelfde wijze verwerkt en daardoor ontstaat vervorming. Vooral de gebieden rond de nullijn, waar de maagdelijke kromme het minst recht is, worden ernstig aangetast. We moeten dus zorgen het proces via het rechte middengedeelte van het maagdelijke kromme te laten verlopen. Dat is bijvoorbeeld te verwezenlijken door het audiosignaal door middel van een extra, constant magnetisch veld naar rechts te verplaatsen. Er is nu sprake van gelijkstroomvoormagnetisatie, die vroeger wel werd toegepast, maar die vanwege de daarbij optredende ruis geheel van het toneel is verdwenen. In plaats daarvan wordt uitsluitend hoogfrequent- ofwel HF-voormagnetisatie toegepast.

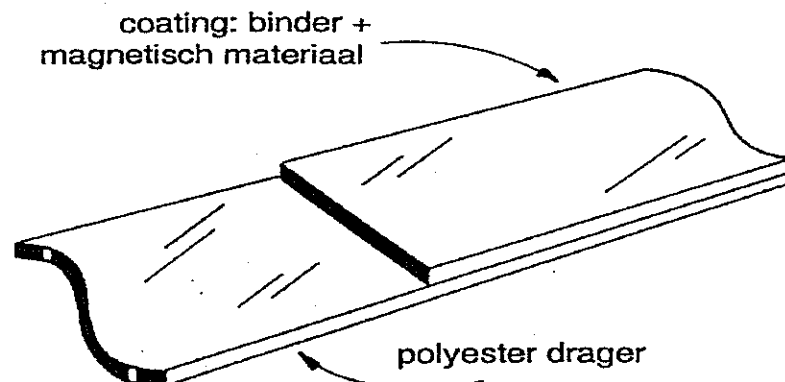


1.5.1.6 Geluidsband voor spoel en cassette

De stof waaruit beide banden zijn vervaardigd is dezelfde, maar bij een cassette is die band veel dunner en wordt op een tragere snelheid (4.75 cm/s)afgespeeld. Bij een spoel is die snelheid dubbel zo hoog (9.5 cm/s).

1.5.1.7 De opbouw van de band

De geluidsband bestaat uit een dunne kunststofdrager, een plastic band, waarop een zeer dun laagje verpulverd ijzeroxide, chroomdioxide of ander magnetiseerbaar materiaal is aangebracht. Het dragermateriaal moet mechanisch zeer sterk en stabiel zijn, een uiterst glad oppervlak hebben, een aantal extreme temperaturen kunnen verdragen zonder ongewenste bijverschijnselen en rekvrij zijn. Dat zijn zeer hoge eisen, waaraan polyester(PE) in hoge mate voldoet. De op de drager aangebrachte magnetiseerbare laag (de emulsie of coating) bestaat uit miljarden kleine naaldvormige kristallen die een lengte hebben van circa $0.5\mu\text{m}$ en een breedte van circa $0.05\mu\text{m}$. Hoe kleiner deze partikeltjes zijn, hoe beter, want des te ruisarmer wordt de band. Maar dit is nog niet voldoende. De deeltjes moeten ook uitermate homogeen over de gehele coating worden verdeeld. De mate waarop dat gebeurt bepaalt voor een belangrijk deel de kwaliteit van de band. De low noise band, die een zeer fijne korrelstructuur had, was het resultaat van de eerste verfijning. Vervolgens kwam de LH-band (low noise - high output), die een zeer lage eigenruis aan een hoge uitstuurbaarheid paarde. Sindsdien is de geluidsband, voor spoelen- en cassetterecorder, steeds verder verfijnd en verbeterd.



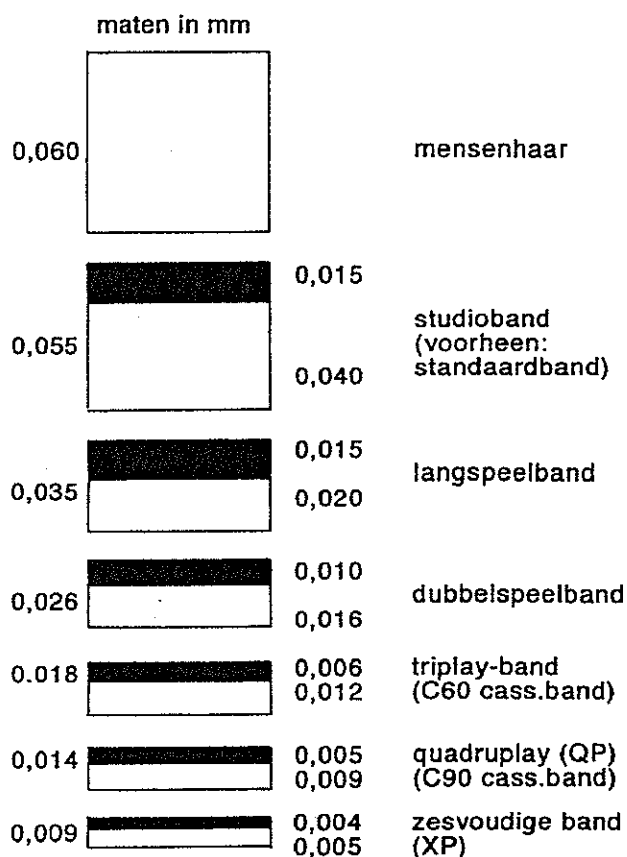
Bij het kwaliteitsaspect speelt het bindmiddel, waarmee de coating hecht verankerd op de drager wordt aangebracht, een zeer belangrijke rol. Verder wordt het hele fabricageproces gekenmerkt door een ongelooflijke precisie en reinheid. Met de meest moderne middelen worden stof en statische ladingen geweerd. En bij elk merk gaat het productieproces gepaard met werkelijk honderden kwaliteitscontroles. Een fout in een band of cassette is dan ook in feite absoluut ondenkbaar.

1.5.1.8 De banddikte

Geluidsband wordt in verschillende dikten gemaakt. De indeling is afgeleid van de spoelenband. De eerste, lang geleden ontwikkelde geluidsband, was de standaard- of normaalband met een dikte van 50 μm ofwel 50 micron. (1 micron 1/1000 millimeter). Deze band wordt nu nog alleen in studio's gebruikt. Door toepassing van sterkere kunststoffen kwam vervolgens de 35 μm dikke langspeelband en daarna de dunnere (26 μm) dubbelspeelband. Deze heeft bij een zelfde spoeldiameter een twee keer zo lange speelduur als de standaardband. Weer wat later kwam de dunne (18 μm) triple- of super-langspeelband, een band met drievoudige speelduur ten opzichte van de standaardband. Deze TP-band werd een tijdlang gebruikt op portable spoelenrecorders. De zeer dunne band werkt een zeer goed kop/bandcontact in de hand en waarborgt ook bij een lage bandsnelheid een goede hoge tonenweergave. Daarom wordt deze band ook in de compactcassette toegepast en wel speciaal in de C60-cassette. In dit verband spreekt men van een 'vrij dikke' band!

Er is een nog dunnere band dan de TP-band, een met een dikte van slechts 12 μm . Deze tape wordt niet op spoel, maar in de C90-cassette toegepast. Er hebben ook C120-cassettes bestaan met een nog weer dunnere band van 9 μm dikte. Omdat er met deze kwetsbare band te gemakkelijk wat fout kon gaan is de C120-cassette uit het marktbeeld verdwenen.

De C60-band is wat verder uit te sturen dan de dunnere C90-band. Daardoor is hij voor de hoogst mogelijke kwaliteitsopnamen aan te bevelen, in het bijzonder op driekoppendecks.



Maar we blijven nog even bij de spoelenrecorder. De daarop toe te passen langspeelband (LP) is te beschouwen als de band voor hifi-doeleinden. Dankzij de betrekkelijk dikke magneetlaag kan deze bandsoort een grote dynamiekomvang verwerken. Een bandsnelheid van 19 cm/s is voor hifi-doeleinden aan te bevelen.

De magnetische eigenschappen van de dubbelspeelband (DP) zijn iets minder gunstig. Zo is de signaal-ruisafstand wat ongunstiger. En de te verwerken dynamiekomvang is wat geringer. Het afgegeven signaal is ook iets minder krachtig. Maar de dunne magneetlaag heeft als voordeel dat de hoge tonenweergave ook bij lage bandsnelheden goed is. Zelfs bij een bandsnelheid van 4,75 cm/s zijn ook met deze band goede resultaten te behalen. Alleen heeft het dan geen zin meer om met een bandrecorder te werken. Dan kan beter een goed cassettedeck worden gebruikt!

De spoelenband kent nog eens twee versies: tape met een glanzende en tape met een matte rugzijde. Deze laatste soort heeft een zgn. back-coating, die voor evenwichtiger snelspoelen zorgdraagt. In het enigszins ruwe oppervlak zijn namelijk als het ware allemaal kleine luchtkussentjes verwerkt die bij zeer snel spoelen, zoals dat bij de duurdere recorders wel gebeurt een zekere demping bewerkstelligen. Daardoor vlijen de wikkelingen zich mooi flexibel tegen elkaar aan. Ook wordt voorkomen dat de spoel te strak wordt opgewikkeld, waardoor magnetische doordrukverschijnselen die voor- en na-echo's veroorzaken achterwege blijven. Maar matte band is alleen te gebruiken op spoelenrecorders die zonder bandaandrukuiltjes werken.

Zoals je ziet zijn er veel nadelen aan het gebruik van een band. Daarom heeft men een nieuwe technologie ontwikkeld nl. de cd. In het volgende deel willen wij daarom eens toelichten hoe men tot de cd gekomen is en welk principe er achter steekt.

1.5.2 Geluid op cd

1.5.2.1 Inleiding

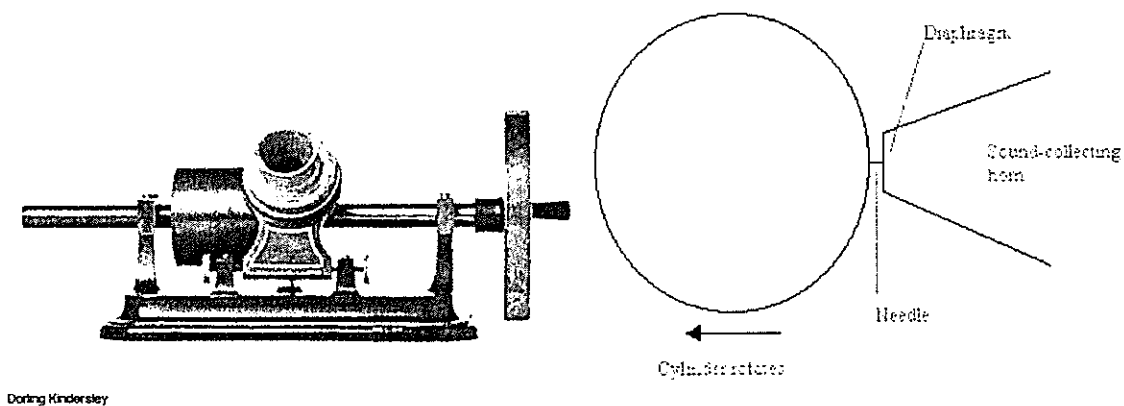
Vandaag zijn de cd's overal te vinden. Als ze nu gebruikt worden om muziek, data of computer software op te slaan, steeds kan men zeggen dat ze een standaard medium geworden zijn voor het verspreiden van grote hoeveelheden informatie die op een betrouwbare manier moeten worden opgeslagen. Daarom zullen we de werking van cd's en cd-drives eens nader belichten.

1.5.2.2 Analoge en digitaal opnemen

Toen de cd's geïntroduceerd werden in de vroege jaren 80, was hun enige doel muziek te bewaren. Dus moet je om te verstaan hoe een cd werkt eerst weten hoe digitaal opnemen van analoge opnemen verschilt.

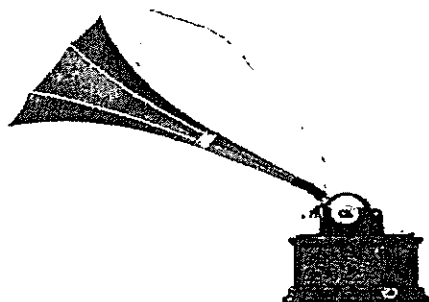
In 1877 vond Thomas Edison het eerste apparaat uit om geluid vast te leggen. Zijn toestel was van erg simpele makelij, het sloeg namelijk een analoge golf op een mechanische wijze op. Zo werd een naald rechtstreeks door een diafragma gecontroleerd waardoor de naald een analoge signaal kraste in een met tin omwonden cilinder. Dit toestel noemde men een fonograaf.

Hieronder staat een foto en het principe.



Dorling Kindersley

Je moest in Edison's toestel praten en terzelfder tijd de cilinder doen ronddraaien, en zo nam de naald het signaal op. Als het diafragma bewoog, bewoog de naald mee en werden de analoge signalen in het tin gekrast en opgenomen. Om het signaal dan weer af te spelen schoof de naald over het analoge opgenomen signaal. Op zijn beurt bewoog dan het diafragma en produceerde het een geluid.

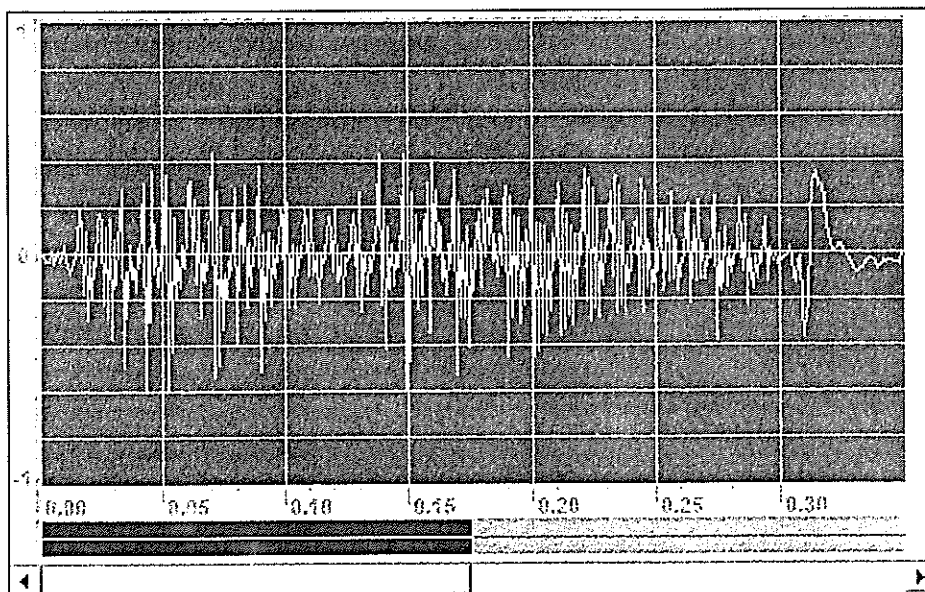


Dorling Kindersley

Dit systeem werd in 1887 verbeterd door Emil Berliner en op die manier ontstond de grammofoonplaat, wat ook een puur mechanisch apparaat was dat gebruik maakt van een naald en een diafragma. De grote verbetering van de grammofoonplaat was dat er geen gebruik meer gemaakt werd van een cilinder maar van een spiraalvormige groef wat massaproductie vergemakkelijkte. De moderne grammofoonplaten werken op een zelfde manier, maar het gelezen signaal wordt elektronisch versterkt in plaats van mechanisch.

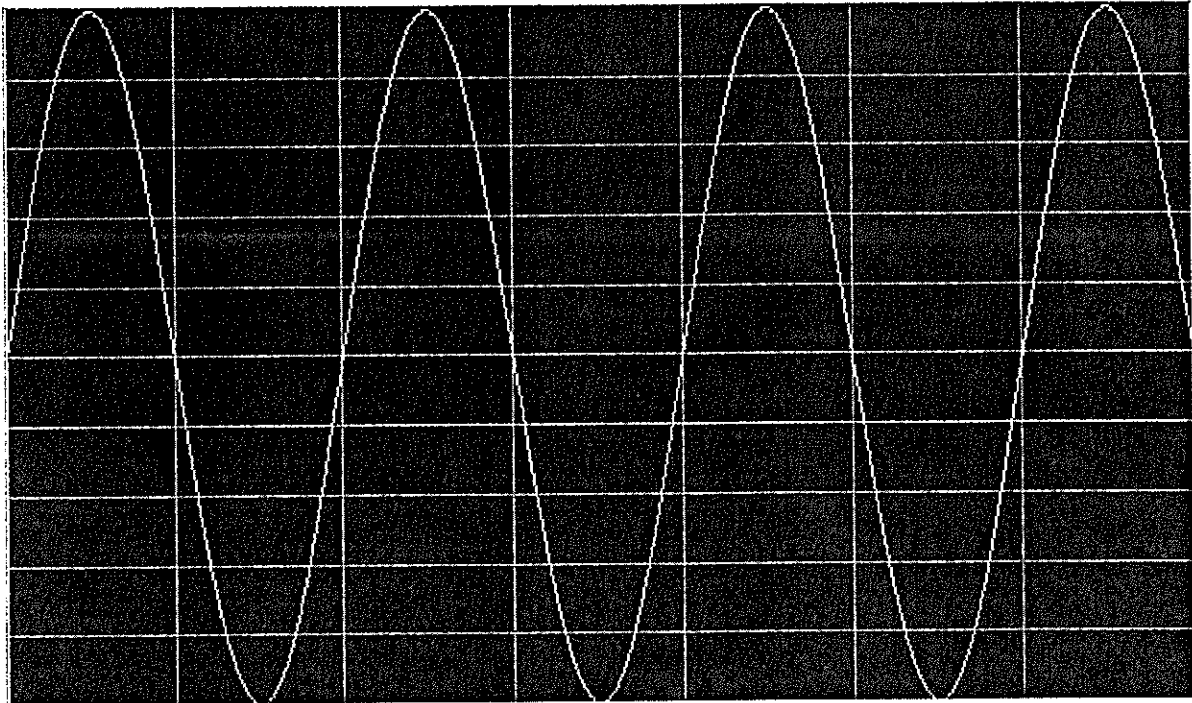
1.5.2.3 Wat is het dat de naald van Edison's fonograaf in de cilinder krast?

Het is een analoge golf die de vibraties weergeeft die de stem veroorzaakt. Bv. hier wordt het woord hallo weergegeven.



Deze golfvorm is elektronisch opgenomen in plaats van op een cilinder maar het principe is hetzelfde. Wat deze grafiek eigenlijk weergeeft is de positie van het diafragma ten opzichte van de ruststand uitgestippelt over een x- en y- as. Het diafragma vibreert tegen een tempo van 1.000 oscillaties per seconde, dit is zeer snel.

Een zuivere toon bestaat uit enkel een sinusvormige golf die vibreert tegen een bepaalde frequentie, zoals deze 500 Hertz golf (500 Hertz = 500 oscillaties per seconde).
Je ziet wel dat het opslaan en afspelen van een analoge golf zeer simpel kan zijn – in een met

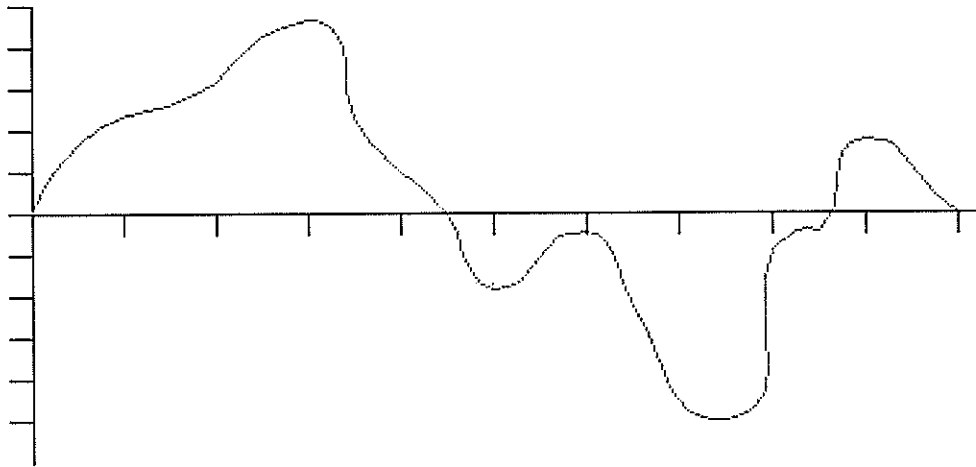


tin omwonden cilinder krassen is een zeer directe en simpele aanpak. Het probleem met deze simpele aanpak is dat de betrouwbaarheid niet erg goed is. Bijvoorbeeld, wanneer je gebruik maakt van Edison's fonograaf wordt er zeer veel ruis opgeslagen samen met het oorspronkelijk bedoelde signaal, en het signaal wordt ook op diverse manieren gestoord. Je zult ook zien dat na herhaald gebruik van de cilinder de informatie zal verdwijnen. Telkens wanneer de naald over de cilinder glijdt verandert de naald de informatie lichtjes en zo is na enige tijd de informatie volledig verdwenen of zodanig verandert dat het niet meer verstaanbaar is.

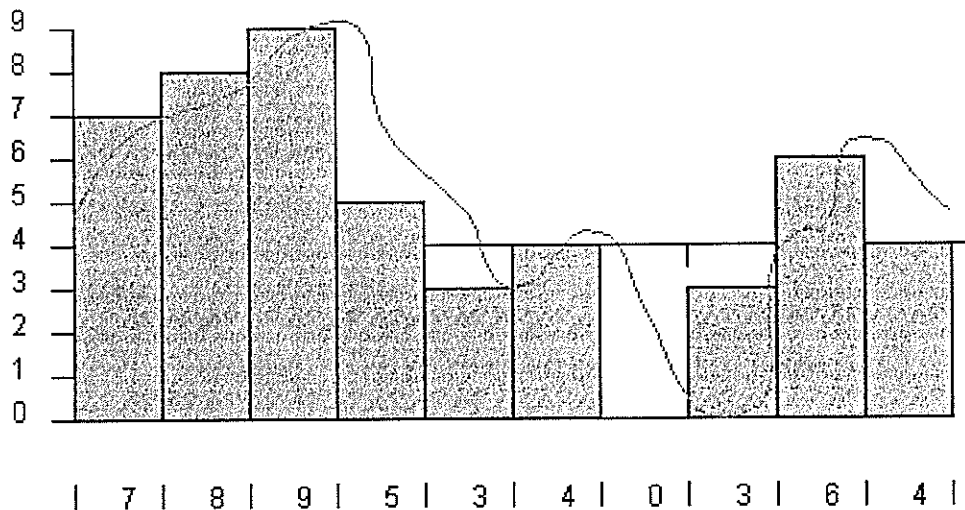
Bij een cd (en elke andere digitale opname technologie) is het doel een zeer betrouwbare opslagplaats te creëren samen met een perfecte afspeling van de data erop. Om deze 2 doelen te bereiken, verandert de digitale opname de analoge golf in een reeks van nummers en slaat deze reeks nummers dan op in plaats van de golf. Deze verandering gebeurt door een apparaat genoemd de analoog-naar-digitaal convertor. Om dan terug de muziek af te spelen, wordt de reeks nummers terug geconverteerd naar een analoge golf wat gebeurt door de digitaal-naar-analoog convertor (DAC). De analoge golf geproduceerd door de DAC wordt versterkt en gevoed naar de luidsprekers gebracht om dan zo het geluid te produceren.

De analoge golf die door de DAC geproduceerd wordt zal iedere keer dezelfde zijn, zolang de nummers niet verstoord worden. De analoge golf zal tevens ook gelijkaardig zijn aan de originele golf als de analoog-naar-digitaal convertor op een hoog niveau samplede en accuraat de nummers schrijft.

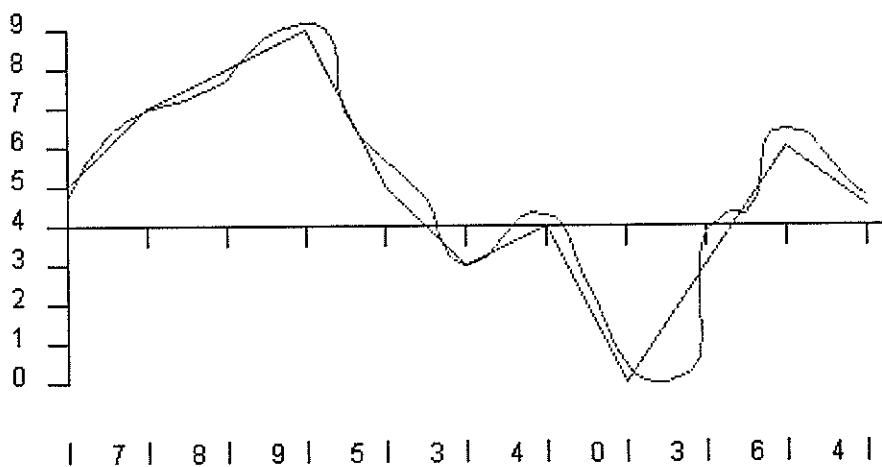
Je verstaat beter waarom cd's een zo hoge betrouwbaarheid hebben als je het analoog-naar-digitaal converteer proces begrijpt. We illustreren dit proces met een voorbeeld en starten met de weergave van een typische analoge golf (op de horizontale as wordt en 1000^{ste} van een seconde voorgesteld)



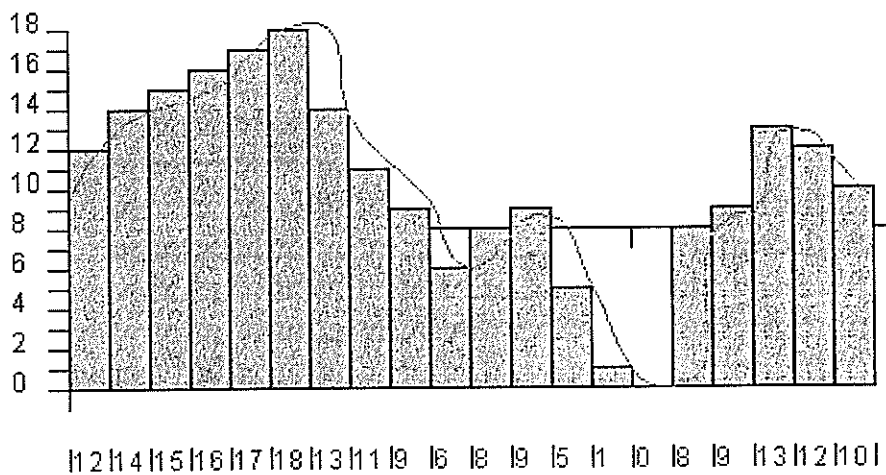
Wanneer je een golf sampelt met een A-naar-D convertor heb je de controle over 2 variabelen. De eerste is de samplingsnelheid. Deze snelheid controleert hoeveel samples er genomen worden per seconde. De 2^{de} is de samplingsprecisie. Deze precisie controleert hoeveel verschillende gradatie niveaus er mogelijk zijn wanneer er gesampeld wordt. In de volgende figuur nemen we aan dat de samplingsnelheid 1000 per seconde is en de precisie 10:



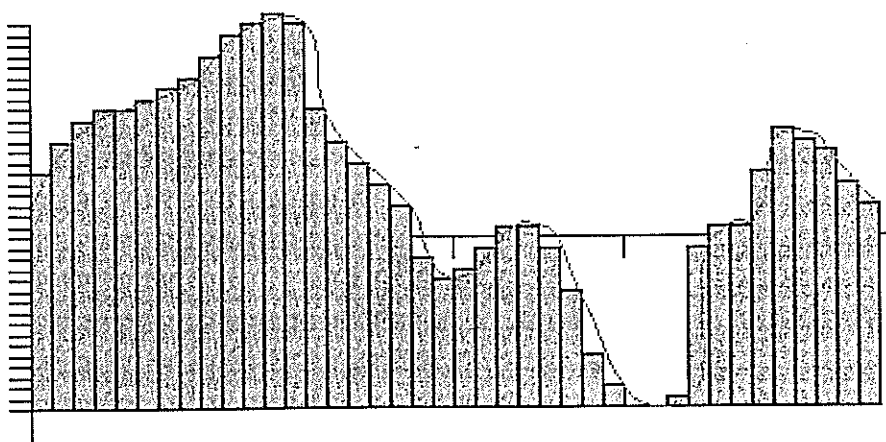
De rechthoeken stellen de samples voor. Ieder 1000^{ste} van een seconde converteert de A-naar-D convertor naar de golf en kiest een nummer tussen 0 en 9. Het nummer dat gekozen wordt staat onder aan de grafiek. Deze nummers zijn digitaal en stellen de originele golf voor. Wanneer de DAC de golf herconstrueert vanuit deze getallen, dan bekom je de hoekige lijn in de volgende figuur:



Je kunt zien dat de hoekige lijn heel wat van de meer gedetailleerde originele golf afwijkt. Dit betekent dat de reproductie van de golf niet zo goed is. Dit noemt men een sampling fout. Je kunt deze fout minimaliseren door de samplingsnelheid en de precisie te verhogen. In de volgende figuur zijn de samplingsnelheid en de precisie met een factor 2 verhoogd.



In de volgende figuur zijn ze nog eens verdubbeld



Hieruit blijkt dat hoe hoger beiden worden hoe hoger de precisie wordt, hoe beter de benadering wordt. Zo verbetert ook de betrouwbaarheid. In het geval van cd-geluid is betrouwbaarheid een belangrijk doel en is de samplingsnelheid 44.100 samples per seconde en het aantal gradaties 65.536. Op dit niveau van de DAC output is de benadering van het origineel zo goed dat het in de oren van de mens origineel klinkt zelfs al is er een afwijking.

Een ding van de cd's samplingsnelheid en de precisie is dat er veel data wordt geproduceerd. Op een cd worden de nummers opgeslagen als bytes en er zijn 2 bytes nodig om 65.536 gradaties op te slaan. Er worden ook terzelfder tijd 2 signalen opgenomen (een voor elke luidspreker van het stereosysteem). Op een cd kun je 74 minuten muziek opnemen, en dus is het totale aantal van digitale data dat op een cd kan opgeslagen worden:

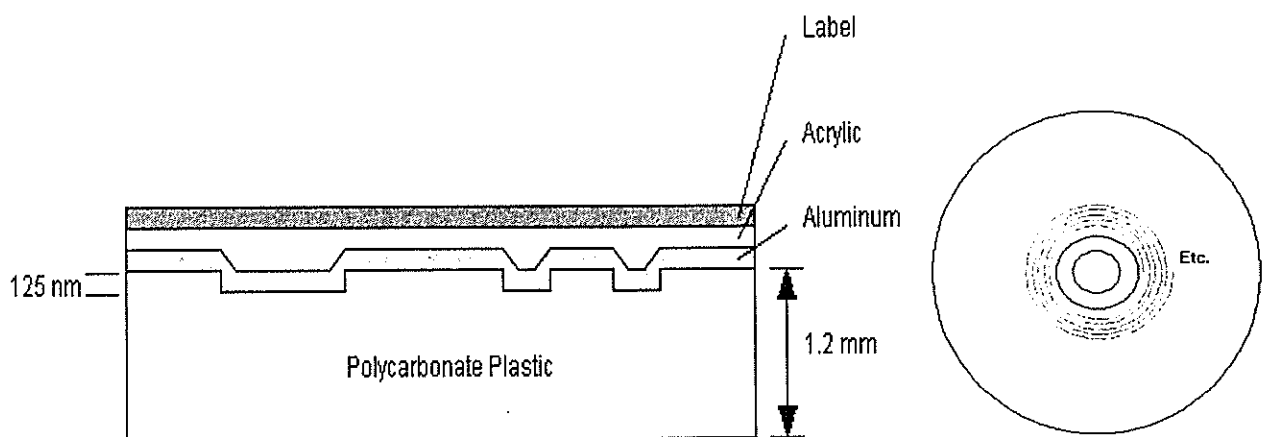
$$44.100 \text{ samples/ kanaal/seconde} * 2 \text{ bytes/sample} * 2 \text{ kanalen} * 74 \text{ minuten} * 60 \text{ seconden/minuut} = 783.216.000 \text{ bytes} \Rightarrow 74 \text{ minuten}$$

1.5.2.4 De cd verstaan

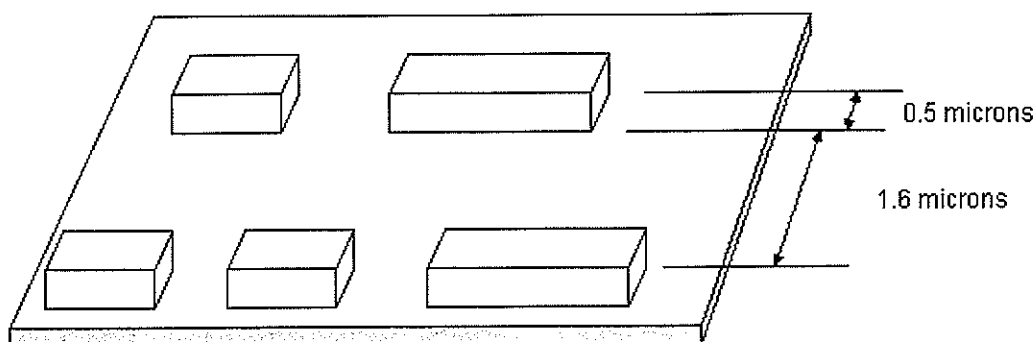
Om deze hoeveelheid data op te slaan op een schijfje met maar een diameter van 12 centimeter moeten de individuele bytes erg smal moeten zijn. Als we kijken naar de opbouw van een cd zullen we zien hoe smal ze eigenlijk wel zijn.

Een cd is eigenlijk maar een simpel stuk plastic van 1.2 millimeter dik. De cd bestaat hoofdzakelijk uit polycarbonaat. Tijdens het maken van een cd wordt op de plasticen schijf een laag aluminium aangebracht waarop dan een laag acryl wordt gelegd om de informatie op het aluminium te beschermen. De bultjes die dan in de aluminiumlaag gebrand worden stellen een binaire code voor. Deze binaire code is aangebracht in een spiraalvorm die vanuit het centrum van de cd vertrekt.

Hier ziet men de opbouw van de cd



Je ziet op de tekening dat de laag aluminium 125 nm dik is maar wat ook belangrijk is dat de spiraal maar 0.5 micron dik is en de spiraal een tussenafstand heeft van 1.6 micron. Het zou er dan moeten uitzien zoals hieronder geïllustreerd.



Deze kleine afstanden zorgen er wel voor dat als je de spiraal zou rechten je een afstand van \pm 7 km zou bekommen.

1.5.2.5 De cd- speler

De job van de cd-speler is de bultjes in de cd lezen en ze vertalen. Daarvoor bestaat de cd speler uit 3 hoofdcomponenten:

1. Een motor die ervoor zorgt dat de cd ronddraait. Deze motor is precies afgesteld om te draaien tussen 200 en 500 tpm (naargelang wat er gelezen wordt)
2. Een laser met lens om de bulten in het aluminium te lezen
3. Een schuifmechanisme om de laser te verplaatsen zodat overal op de cd gelezen kan worden. Dit moet met micron precisie gebeuren.

In de cd-speler is er ook nog een groot deel computertechnologie om de blokken data verstaanbaar te maken voor de DAC of voor de computer

Maar de fundamentele job van de cd-speler is de laser te richten op de bultjes. De laserstraal gaat door de polycarbonaatlaag, kaatst dan af tegen de aluminiumlaag en keert dan terug naar een opto-elektronisch apparaat dat lichtveranderingen detecteert. De bulten reflecteren op een andere manier dan de rest van de aluminium laag en zo kan het apparaat de bits herkennen en bytes vormen. Dit principe wordt tevens niet enkel gebruikt bij de cd-speler maar is ook vertegenwoordigd in de dvd-speler

1.5.2.6 DVD: de nieuwe generatie van dataopslag

Wat is het verschil met de cd-speler? Het grote verschil is dat er op een dvd -rom 7.5 keer meer data kan dan op een normale cd-rom. Hoe komt dit?

Wel, een dvd-rom is fysisch zeer gelijkend aan een cd. Ze hebben dezelfde diameter, dezelfde dikte, gebruiken dezelfde technologie waarin een laser reflecteert tegen een aluminiumlaag. Maar de bultjes die op een dvd-rom staan zijn veel wel kleiner dan op een gewone cd. De afstand tussen de spiraal is maar 0.74 micron (in plaats van de 1.6 micron bij de cd) en ook de dikte van de bulten en de lengte ervan zijn ongeveer met hetzelfde percentage vermindert.

Met al deze opslagcapaciteit is de dvd-rom beter geschikt om films op te slaan dan op een cd.



Hoofdstuk 2. Spraaktechnologie

2.1 Principe

2.1.1 Waarom

De robot C-3PO treedt in de StarWars-films op als vertaler. Het verstaan, vertalen en spreken van 6 miljoen verschillende vreemde talen gaat hem in de film moeiteloos af.



De kleine, teddybeerachtige wezens in de StarWars-film 'Return of the Jedi' zijn volstrekt onverstaanbaar. Luke Skywalker roept hiervoor de hulp in van de robot C-3PO. Die vertaalt feilloos het gebrabbel, een opmerkelijke demonstratie van taal- en spraaktechnologie die in werkelijkheid nog niet is geëvenaard. Allereerst identificeert C-3PO de taak waarin Luke Skywalker hem vraagt het gebrabbel te vertalen. Dan herkent hij de woorden en begrijpt hij de inhoud van het verzoek. Vervolgens doet hij hetzelfde voor de woorden in de onbekende taal: taal identificeren, woorden herkennen en betekenis begrijpen. Dan vertaalt hij het gebrabbel in de taal van de vraagsteller en spreekt hij de vertaalde mededelingen uit.

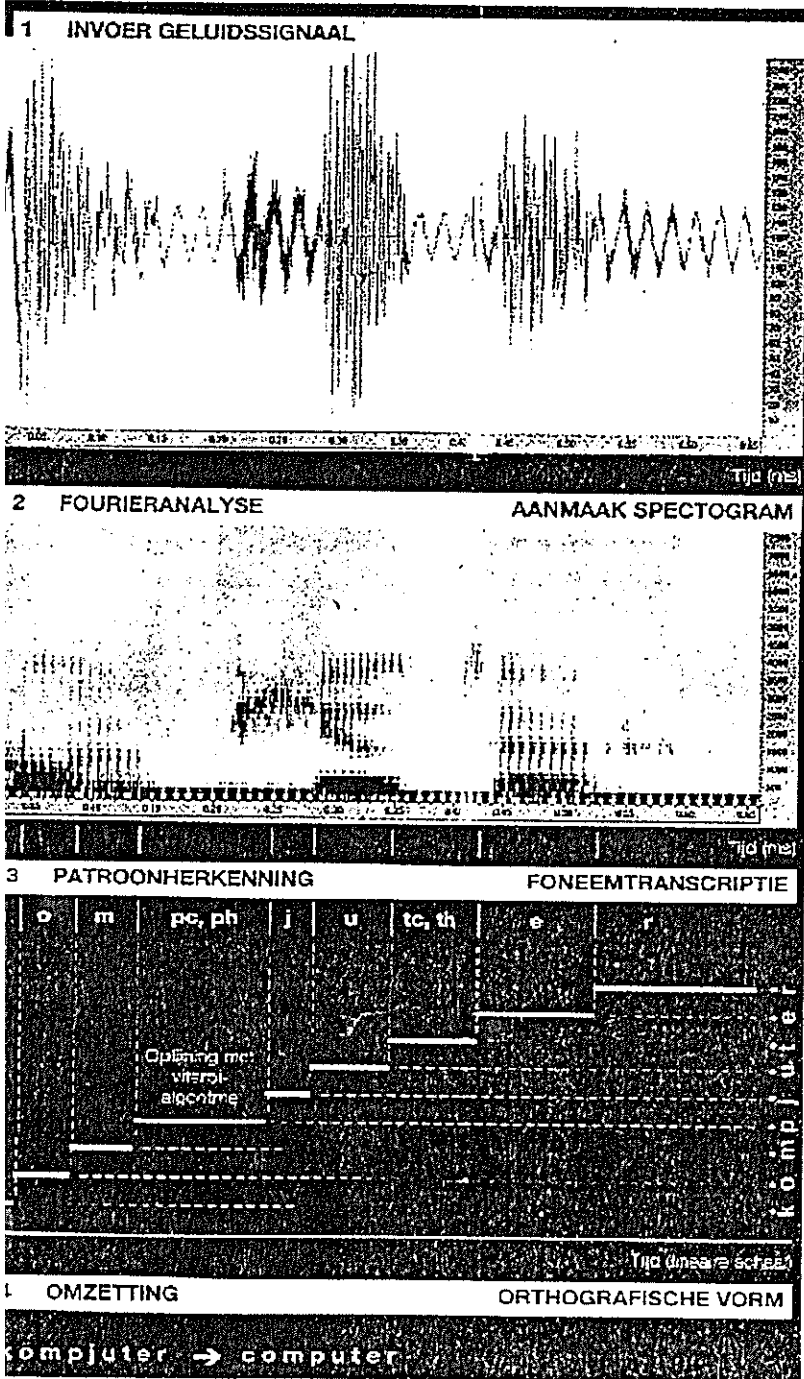
Spraakherkenning en spraaksynthese zijn voorbeelden van taal- en spraaktechnologie. Uiteindelijk moeten deze technieken een vloeiend gesprek tussen mens en machine mogelijk maken. Zo'n systeem noemt men gesprokendialoogstelsel. Denk hierbij bijvoorbeeld aan Captain Kirk uit de film Star Trek, die regelmatig met zijn boordcomputer praat. Inmiddels zijn gesprokendialoogsystemen, die aanvankelijk alleen in sciencefictionfilms voorkwamen, werkelijkheid geworden.

2.1.2 Het nut van spraaktechnologie

Tussen de vele definities die aan automatische spraakherkenning gegeven kunnen worden is diegene die zegt dat het een manier is om data in de computer in te voeren waarschijnlijk de meest juiste.

Er worden reeds vele jaren verschillende invoerapparaten gebruikt. Deze apparaten hebben elk hun eigen toepassingsgebied. Het meest bekende invoerapparaat is ongetwijfeld het toetsenbord, wat wordt gebruikt om commando's en tekst in te voeren. Met de komst van de grafische programma's zoals Robocad, enz... heeft de muis een groot deel van de invoer langs het toetsenbord overgenomen. En bijna dezelfde manier is spraakherkenning nog een ander invoerapparaat dat kan gebruikt worden wanneer het nodig of gepast is.

Invoer langs de stem zal waarschijnlijk nooit het toetsenbord of de muis volledig vervangen, net zoals de muis het toetsenbord niet heeft vervangen. De invoerapparaten zullen veeleer



Klinkers

| | |
|----|-------|
| i | liep |
| e: | leeg |
| a: | laat |
| o: | boom |
| y | buut |
| ɜ: | deuk |
| ɛ | wijs |
| ɔy | huys |
| ɪ | lɪp |
| ɛ | leg |
| ʌ | læt |
| ɒ | bom |
| ʏ | put |
| ɛ | geijk |
| u | boek |
| Au | koud |

Medeklinkers

| | |
|---|-------|
| p | put |
| b | bad |
| t | tak |
| d | dak |
| k | kat |
| N | lang |
| m | mal |
| n | nat |
| l | lat |
| r | rat |
| f | fiets |
| v | vat |
| s | sap |
| z | zat |
| S | sjaal |
| J | jas |
| x | licht |
| h | had |
| w | wat |

Fonemen, of spraakklanken, zijn de bouwstenen van de spraak. OVIS is getraind in bovenstaande 35 fonemen. Links steeds de fonem, rechts een voorbeeld van een woord waarin de klank voorkomt.

samenwerken om zo efficiënter te werken want in sommige toepassingen zal het toetsenbord meer gepast zijn en in andere de muis of de spraakherkenning. Zo zal er een dialoog gevormd worden tussen de computer en de mens.

2.1.3 Oplijning

Om deze dialoog te kunnen voeren moet de computer allereerst in staat zijn de gesproken vraag te herkennen. Uiteraard herkent de computer alleen de woorden die in zijn woordenboek staan. Het nieuwe spraaktechnologie systeem van IBM, ViaVoice, beschikt bijvoorbeeld, afhankelijk van de toepassing, over een woordenschat tot ruim 128.000 woorden. Zo'n woordenboek bevat van ieder woord twee vormen: ten eerste de orthografische vorm, oftewel het woord zoals het geschreven wordt, en ten tweede het woord zoals het uitgesproken wordt. Dit is een reeks van foneemsymbolen, en heet daarom foneemtranscriptie. Ieder foneemsymbool staat voor een bepaalde klank.

Fonetische transcripties zijn een essentieel hulpmiddel in de ontcijfering van de menselijke spraak. Het doel van de fonetische transcriptie is een methode te geven om de uitspraak van een woord in een geschreven vorm op te slaan.

Het meest gebruikte fonetische alfabet is vastgelegd door de Internationale Fonetische Vereniging (IFV). Dit alfabet bestaat vooral uit letters (tekens) afkomstig uit het Romaanse alfabet samen met nog enkele letters van het Griekse alfabet. Het IFV heeft zelf ook enkele letters toegevoegd omdat er in de loop der jaren enkele nieuwe klanken zijn ontstaan die niet konden worden uitgedrukt met het oorspronkelijke alfabet.

Bv. cynisme → fonetische vorm → sienism θ

Het principe van dit alfabet is dat een welbepaald geluid wordt voorgesteld door een enkel symbool, zoals in ons alfabet de letter a samengaat met een bepaalde klank. Het IFV systeem biedt tevens een reeks diakritische tekens aan die werden toegevoegd aan het alfabet

Bv. het accent grave

Het doel van deze diakritische tekens is de omschrijving van de uitspraak nog nader te omschrijven, indien dit nodig is en men dubbelzinnigheden kan vermijden. Het IFV alfabet is niet verkrijgbaar op een standaardtoetsenbord en het wordt niet ondersteund op een gewone computer met een karakterset zoals ASCII, om dit probleem te verhelpen heeft men programma's ontworpen die vanuit de standaardtekens de omzetting maken naar hun fonetische vorm. Maar ieder persoon heeft ook een eigen stem die volgens de plaats waar hij/zij woont een specifieke uitspraak geeft aan sommige woorden. Hiervoor moet er meestal in een spraakprogramma een trainingsfase worden doorlopen.

In de trainingsfase maakt het systeem kennis met een grote verzameling spraakfragmenten, waarbij ieder fragment vergezeld gaat van een bijhorende tekst. Een dergelijke verzameling heet een corpus. Het systeem zoekt bij alle fragmenten uit het corpus de woorden in het woordenboek en vervangt ieder woord door zijn foneemtranscriptie.

Vervolgens bepaald het zogenaamde Viterbi-algoritme de optimale oplijning tussen het spraaksignaal en de foneemtranscriptie. Deze oplijning is in feite een segmentering, die in het spraaksignaal de grenzen van ieder van de elementen uit de foneemtranscriptie bepaalt.

Bovendien berekent het Viterbi-algoritme voor de gevonden oplijning de kans dat het spraaksignaal en de foneemtranscriptie bij elkaar horen. Het algoritme kiest de oplijning met de grootste kans: de optimale oplijning. In de trainingsfase worden het Viterbi-algoritme en een trainingswoordenboek gebruikt om een trainingscorpus geheel te segmenteren. Na segmentatie kan voor iedere basiseenheid (foneem) worden opgezocht welke stukken spraaksignaal uit het trainingscorpus hierbij horen. Voor ieder foneem worden alle bijhorende stukken spraaksignaal statisch verwerkt en wordt een model berekend: het hidden Markov model. Ieder aldus berekend akoestisch model is een statische beschrijving van alles dat bij dat foneem horende stukken spraaksignaal uit het trainingscorpus.

2.1.4 Hypothesen

Voor succesvolle spraakherkenning zijn niet alleen akoestische modellen nodig, maar ook taalmodellen. Het huidige systeem gebruikt een unigram (de kans op ieder woord) en een bigram (de kans op een woordpaar) taalmodel. Deze taalmodellen worden getraind door in het corpus te tellen hoe vaak ieder woord, respectievelijk iedere combinatie van twee woorden, voorkomt. Dit aantal gedeeld door het totaal aantal woorden, respectievelijk woordparen, levert de kans op.

In de herkenningfase wordt tot slot geprobeerd een onbekende uiting te herkennen. De spraakherkenner genereert alle mogelijke opeenvolgingen van woorden. Omdat van tevoren niet bekend is uit hoeveel woorden een zin bestaat, is het aantal hypothesen gigantisch groot, zeker als de herkenner over een grote woordenschat beschikt. Gelukkig scoort de spraakherkenner alle hypothesen van begin af aan, dat wil zeggen dat het bepaalt hoe waarschijnlijk iedere hypothese is, gegeven het binnenkomende (onbekende) signaal.

Hiervoor maakt het programma weer gebruik van het Viterbi-algoritme, dat de optimale oplijning en de bijhorende kans bepaalt. Het merendeel van de hypothesen blijkt dan al na een paar stappen zoveel minder waarschijnlijk te zijn dan de favorieten, dat ze zonder enig gevaar uit de lijst van mogelijke oplossingen kunnen worden geschrapt. Op die manier blijven geheugenbeslag en rekentijd voor het scoren van de hypothesen binnen redelijke grenzen. Dit is belangrijk omdat een spraakherkenner in de praktijk real-time moet werken. Uiteindelijk valt de keuze op de opeenvolging van woorden met de grootste kans – zo heeft de computer de aangeboden zin verstaan.

2.1.5 Opel astra

Zodra de vraag verstaan is – spraak is omgezet in tekst – moet de computer de vraag interpreteren en er een antwoord op verzinnen. Er bestaan vele dialoogstrategieën om een computer een gesprek te laten voeren. Een tijd lang was het bijvoorbeeld mogelijk dat je in Nederland een automaat van de ANWB belde om de waarde van je auto te bepalen. Een mogelijke dialoog verloopt dan als volgt: “voor Opel toets 0, voor Peugeot toets 1,...” enzovoort. Nadat je een automeerk had gekozen, moest je het type en bouwjaar kiezen. Hierbij doorliep je een vast menu, met op iedere plek een beperkt aantal keuzen. In plaats van het intoetsen van nummers zou ook spraak kunnen worden gebruikt: “voor Opel zeg 0, voor Peugeot zeg 1,...”. Dergelijke dialoog structuren zijn in het begin daadwerkelijk toegepast, omdat de beschikbare spraakherkenners alleen de getallen konden herkennen. Iets natuurlijker wordt het als je direct “Opel” of “Peugeot” kunt zeggen, maar nog steeds moet je dan een vast menu doorlopen.

Ofschoon je op deze manier tot de gewenste informatie kunt komen, vinden veel bellers dit een omslachtig en onnatuurlijke manier. Wat je het liefst zou zeggen is: “Wat is de waarde van mijn Opel Astra uit 1994?” In dat geval zul je een dialoogstrategie moeten gebruiken die gemengd-initiatiefdialoog genoemd wordt, omdat zowel het systeem als de beller het initiatief kunnen nemen. Dit is dus niet het geval bij een menugebaseerd systeem waar het systeem altijd het initiatief neemt door vragen te stellen waarop de beller alleen zeer specifieke antwoorden kan geven. Het Openbaar Vervoer Informatie Systeem voert een gemengd-initiatiefdialoog.

2.1.6 Krant

Zodra de computer het antwoord op de gestelde vraag heeft geformuleerd, moet het nog worden uitgesproken. Er is een aantal mogelijkheden om dit te doen. Welke mogelijkheid wordt gekozen hangt af van de hoeveelheid tekst in de applicatie. Als de hoeveelheid mededelingen in een toepassing klein is, volstaat het terugspelen van opgenomen boodschappen. Vijftien jaar geleden was de afdeling Taal en Spraak van de Katholieke Universiteit Nijmegen betrokken bij de ontwikkeling van een sprekende lift, die (met name voor visueel gehandicapten) belangrijke informatie uitspreekt, zoals “de lift komt” en “u nadert de eerste verdieping”. Deze techniek wordt nu ook toegepast in speelgoed, zoals de sprekende pop.

Het werken met opgenomen boodschappen is ondoenlijk als de hoeveelheid tekst heel groot is. In dat geval biedt de zogenaamde spraaksynthese uitkomst. Een mooi voorbeeld is de elektronisch leesbare krant. Zo verstrekt het Centrum voor Gesproken Lectuur in Grave voor visueel gehandicapten elektronische versies van kranten. De gebruiker haalt via de telefoonlijn de digitale versie op de eigen computer binnen. Via speciale software en spraaksynthese kan hij vervolgens de krant beluisteren.

Bij spraaksynthese wordt de binnenkomende tekst, een reeks van letters, allereerst omgezet in een reeks van foneemsymbolen. In het woord 'rekenwerk' bijvoorbeeld zitten drie verschillende klanken 'e'. Die dus elk met een ander foneem moeten worden beschreven.

Vervolgens wordt de reeks fonemen omgezet in een spraaksignaal. Ofschoon deze spraak (zeker na enige gewenning) goed verstaanbaar is, valt het meteen op dat het geen menselijke spraak is maar een computerstem. In de meeste commerciële applicaties geeft men nog de voorkeur aan een meer natuurlijk klinkende stem, zo ook voor het verstrekken van reisinformatie door het Openbaar Vervoer Informatie Systeem (OVIS). Zelfs het beperkte aantal mededelingen dat OVIS moet uitspreken, is al te groot om dat allemaal apart op te nemen. Het terugspelen van een volledige verzameling opgenomen boodschappen is dan onmogelijk. Denk hierbij aan het grote aantal combinaties van stations en reistijden. De lijsten van afzonderlijke onderdelen (stationsnamen, dagen, data en tijdstippen) is echter te overzien en kan wel van tevoren worden opgenomen; dit is bij OVIS gebeurd. Naar gelang de informatie die de gebruiker nodig heeft, plakt het systeem simpelweg spraakfragmenten aan elkaar vast. Deze techniek noemt men spraakconcentratie. Het nadeel van deze techniek is natuurlijk dat als je woorden wil toevoegen je deze eerst moet opnemen.

2.1.7 PIN- code

Een populaire toepassing van spraakherkenning is de dicteersoftware. Wie niet op spreektempo kan typen, kan met zulke programmatuur enorme tijdswinst boeken. Een andere toepassing is het handvrije telefoneren. Onlangs stelde een Nederlandse politieke partij voor om telefoneren in de auto alleen toe te staan als deze techniek wordt gebruikt. Zonder de handen van het stuur te nemen vertel je de telefoon wie je wilt bellen. – De spraakherkenning in de telefoon doet de rest. Voor toepassingen in bijvoorbeeld beveiliging is alleen het herkennen van gesproken tekst niet voldoende. De identiteit van de spreker moet ook op basis van zijn spraak worden vastgesteld. Een toepassing die binnenkort zeker gebruikt gaat worden, is telefonisch thuisbankieren. De klant noemt zijn rekeningnummer en een pincode, de spraakherkenner herkent dit, weet dan wiens rekening het betreft en controleert of de spraak die van de rekeninghouder is. Een andere toepassing van spraakherkenning en sprekerverificatie is stemgestuurde toegangscontrole: de deur van een ruimte gaat alleen open als de juiste persoon de juiste boodschap uitspreekt.

Bij Openbaar Vervoer Reisinformatie is het OVIS-gesproken-dialogosysteem al operationeel en talrijke toepassingen zullen spoedig volgen. In de toekomst zal het steeds vaker gebeuren dat je niet alleen tussen neus en lippen door iets tegen de machine zegt, maar dat hij je ook nog eens begrijpt.

2.1.8 Wat is spraaksynthese?

Spraakconversie is de omzetting/conversie van geschreven input naar gesproken output door het automatisch genereren van synthetische spraak. Als de synthese het hele traject omvat van geschreven tekst naar uitgesproken tekst, wordt deze vaak "Tekst-naar-spraak"-conversie (TTS, Text-to-Speech conversion) genoemd. De afbeelding hieronder toont het hele traject. De fonetische input kan uiteindelijk op veel manieren gesynthetiseerd worden. Bij sommige systemen is het zelfs mogelijk de verschillende synthesesmogelijkheden te mengen. De Fluent Dutch API bijvoorbeeld biedt ontwikkelaars de mogelijkheid om pure tekst-naar-spraak te mengen met foneem-naar-spraak en resynthese van natuurlijke prosodie. De factor die het meeste verschil uitmaakt is de eenheid die wordt gekozen als het basiselement dat de spraak opbouwt. We beschrijven hieronder drie verschillende methoden: allofoonspraak, difoonspraak en aaneengeketende zinnen.

2.1.8.1 Allofoonspraak

In sommige systemen wordt de allofoon gebruikt als basiseenheid. Dit is kleinste eenheid in de spraak; in de synthesesystemen wordt de gewoonlijk automatisch gegenereerd. De spraak die hiervan het resultaat is, is daarom niet gebaseerd op de karakteristieken/kenmerken van een menselijke spreker. Deze spraak zal gewoonlijk mechanischer klinken dan andere vormen van spraaksynthese. Aan de andere kant kunnen de eigenschappen van de stem eenvoudig worden verandert en aangepast.

2.1.8.2 Difoonspraak

Voor dit type synthese worden alle mogelijke overgangen tussen 2 fonemen van één spreker opgenomen, bijvoorbeeld de overgang tussen de 'b' en de 'a' in het woord 'bad'. Deze overgangen, die difonen worden genoemd, kunnen vervolgens in iedere gewenste volgorde worden geplakt, en op deze manier spraak produceren. De spraak die hieruit voorkomt, klinkt vaak natuurlijker dan allofoonspraak, maar behoudt ook de karakteristieke kenmerken van de stem van de spreker. Het is tot op zekere hoogte mogelijk om de parameters van de stem (zoals toonhoogte en snelheid) te manipuleren.

2.1.8.3 Aaneengeschakelde zinnen

In sommige gevallen worden volledige zinnen opgenomen en gedigitaliseerd. Dit kan gebeuren in de vorm van sjablonen, waarin stukken worden opengelaten om andere zinnen of woorden tussen te voegen. Van de 3 soorten synthese klinkt deze het meest natuurlijk, aangezien de meeste kenmerken van de stem van de spreker behouden blijven. Dit systeem is echter beperkt tot de zinnen die zijn opgenomen. Iedere nieuwe zin of ieder nieuw deel van een zin moet door dezelfde spreker worden opgenomen.

2.1.9 Verschillende types van spraakherkenning

Volgens de criteria die nodig zijn kan men verschillende types onderscheiden. Wanneer we naar de spraaksnelheid kijken kunnen we een onderscheid maken tussen discrete en opeenvolgende spraakherkenning.

Discrete spraakherkenning herkent woord per woord. Het nadeel bij deze soort van spraakherkenning is dat ze zeer tijdsrovend is door het feit dat je tussen ieder woord een korte pauze moet inlassen. Het heeft dan wel een voorlopig groot voordeel dat het zeer juist werkt.

Opeenvolgende spraakherkenning herkent zinnen. Deze manier van dicteren gaat zeer vlug maar is nogal onnauwkeurig.

Het type van de toepassing speelt natuurlijk ook nog een rol. Hier kan men de keuze maken tussen Command-and-control programma's en dictatie-programma's. Verder kan er nog een onderscheid gemaakt worden tussen de programma's die een training nodig hebben (speaker-dependant) en programma's die geen training nodig hebben (speaker-independant).

Tussen al deze types kan men afgaande op de engines eigenlijk 2 grote types onderscheiden namelijk de programma's met als basis woordmodellen en die met als basis foneemmodellen. In de volgende paragraaf zullen de werking van de foneemmodellen eens verduidelijken (de woordmodellen laten we achterwege want deze zijn verouderd).

2.2 Werking

2.2.1 Fonetische engines

Wanneer de engine een bepaalde uitspraak heeft herkend en een herkenningsresultaat voor handen is verwittigd het de toepassing. Terzelfder tijd geeft de engine een betrouwbaarheidsniveau aan de herkende zin. Dit betrouwbaarheidsniveau is een indicatie van hoe zeker de engine is dat het de gesproken zin juist heeft herkend.

Daarna wordt het betrouwbaarheidsniveau vergeleken met de aanvaardingparameter. Deze aanvaardingsparameter is een soort grensgebied waarin het betrouwbaarheidsniveau moet liggen zodat de toepassing wordt verwittigd.

Als men deze aanvaardingsparameter gelijkstelt aan 0 dan aanvaardt de engine slechts één specifieke uitspraak maar enkel wanneer het 100% zeker is van het herkenningsresultaat. Het betrouwbaarheidsniveau van het herkenningsresultaat wordt berekend en vergeleken met de aanvaardingsparameter die voordien ingesteld is op basis van de training. Als deze waarde hoger is dan de waarde van de aanvaardingsparameter dan wordt het resultaat aanvaardt, zoniet wordt het afgewezen en wacht de engine op nieuwe spraak.

Maar niet tegenstaande zijn er ook nog een aantal niet te vergeten eisen.

2.2.2 Functionele eisen

De functionele eisen bepalen de fundamentele taken van het systeem. Ze leggen vast met welke uitvoer het systeem reageert op bepaalde invoer. De functionaliteit van het systeem wordt beschreven vanuit de volgende invalshoeken.

1. Dialoogstructuur
2. Softwarestructuur
3. Afhandeling van fouten

2.2.2.1 Dialoogstructuur

De dialoog wordt gemaakt om er voor te zorgen dat de communicatie tussen mens en computer zo goed mogelijk verloopt. In de dialoog wordt geprobeerd om de antwoorden van de persoon zo veel mogelijk te sturen, zodat een computer begrijpt wat er geantwoord wordt.

Bij het opstellen van de dialoog zal er worden uitgegaan van de volgende punten:

- Om vervelende, onnatuurlijk aandoende vragen te voorkomen wordt in principe impliciete verificatie gebruikt om vast te stellen of de antwoorden van de gebruiker goed herkend zijn.
- Op ieder moment moet de gebruiker om inlichtingen kunnen vragen
- Er wordt van uitgegaan dat de gebruiker totaal geen ervaring heeft met computers en spraakprogramma's

2.2.2.2 Softwarestructuur

De software die gemaakt wordt is een brug tussen dialoog (en dus de gebruiker) en de spraakherkennings en synthese software. Een andere taak van de software is het bijhouden van de stemmen die uitgebracht zijn. Dit gebeurt door het bijhouden van een logbestand waar alle uitgebrachte stemmen in staan.

Het programma is grofweg ingedeeld in drie delen:

1. De stemmodule is het hoofdgedeelte van het programma, hier worden de andere delen onder geïnitieerd en worden de uitgebrachte stemmen bijgehouden.
2. De dialoogmodule leest een dialoog bestand in en zorgt ervoor dat de dialoog in de juiste volgorde wordt afgehandeld.
3. De communicatiemodule zorgt voor de communicatie met de spraakherkennings- en -synthese-software, zodat het programma kan "luisteren" en "spreken".

2.2.2.3 Afhandeling van fouten

Het optreden van een fout door toedoen van de gebruiker mag nooit leiden tot stoppen van het systeem. De spraakherkenningssoftware die gebruikt wordt, geeft met een percentage aan hoeveel zekerheid er aan het herkende woord gehecht kan worden. Door middel van tests zal geprobeerd worden een waarde te vinden voor dit percentage die een ideale balans is tussen betrouwbaarheid en tolerantie t.o.v. afwijkend stemgeluid.

2.2.3 De structuur van de stemcomputerprogrammatuur

2.2.3.1 Stemmodule

De basis van de programmatuur wordt gevormd door de stemmodule. Deze module is verantwoordelijk voor de initialisatie van de overige modules en voor het beginnen van de stemprocedure. De instellingen voor de initialisatie van de overige modules worden in de stemmodule gedaan. Ook zorgt dit deel van de software voor de registratie van de uitgebrachte stemmen. Initialisatie moet vanzelfsprekend plaatsvinden en ook moet er de mogelijkheid zijn om het stemproces in gang te zetten. Verder is het voor testdoeleinden erg praktisch dat er al een elementaire registratie van de uitgebrachte stemmen aanwezig is. Maar een belangrijke reden om dit niet te doen is het feit dat bestaande computers ook onderdelen hebben met vergelijkbare functionaliteit (dit geldt met name voor de registratie van stemmen). De stemmodule kan zo mogelijk (gedeeltelijk) vervangen worden door stukken software die al aanwezig zijn in de computer.

2.2.3.2 Dialoogmodule

De dialoogmodule wordt door de stemmodule aangeroepen wanneer een stemprocedure gestart wordt. De module begint dan een dialoog met de gebruiker. De uiteindelijke communicatie vindt plaats met behulp van een communicatie module, maar de dialoogmodule bepaalt de structuur van de dialoog. Om de dialoog zo flexibel mogelijk te houden, wordt er meestal gekozen om de structuur van een stemdialoog vast te leggen in een afzonderlijk bestand in plaats van een programmacode in de dialoogmodule zelf. In feite is de dialoogmodule een soort van dialoog-interpreter die op basis van aanwijzingen afkomstig uit een dialoogbestand een bepaalde dialoog voert met de gebruiker.

Er zijn meerdere redenen waarom voor een dergelijke flexibele opstelling wordt gekozen.

- Een (eenvoudige) wijziging in de dialoogstructuur vereist geen wijziging van het programma zelf.
- Ook een onervaren gebruiker kan gemakkelijk een nieuwe gewijzigde dialoog laden.
- Het programma is geschikt voor de meest uiteenlopende doeleinden waarbij een dialoog tussen mens en computer gewenst is.
- Eventueel kan het programma ook voor niet-Nederlandstaligen bruikbaar worden gemaakt, door andere taalversies van de spraakherkenning- en spraaksynthese software te installeren.

2.2.3.3 Communicatiemodule

De communicatiemodule voorziet grofweg in twee functies: een in- en een uitvoerfunctie. De uitvoerfunctie maakt het mogelijk om de gebruiker iets mee te delen, de invoerfunctie stelt de gebruiker in staat om de computer iets mee te delen.

Deze manier van gebruikersinvoer zal uiteindelijk plaatsvinden met behulp van spraaksynthese en spraakherkenning. De toepassing van spraaktechnologie is echter software technisch niet een vereiste voor de communicatie module. In principe mag de module op iedere mogelijke manier communiceren met de gebruiker. Het is daardoor mogelijk om voor testdoeleinden een communicatie module te schrijven die mededelingen gewoon op het scherm zet en invoer eenvoudigweg langs het toetsenbord gebeurt.

Indien er geen stemprocedure actief is, dat wil zeggen, wanneer er niemand bezig is, kan er een nieuwe stemdialoog gestart worden met behulp van een knop bv. "start stemprocedure". De dialoogmodule wordt samengeroepen voor het uitvoeren van zo'n stemdialoog en het resultaat wordt samen met twee tijdsmarkeringen opgeslagen in het logbestand. De tijdsmarkeringen geven respectievelijk het begin en het einde van de stemprocedure aan. Na het eindigen van een stemdialoog kan een volgende stemprocedure begonnen worden.

In principe moet er speciaal iemand beschikbaar zijn om door een druk op de "start stemprocedure"-knop een stemdialoog te starten. De implementatie van de stemmodule kan vrij direct zijn. De overige modules zijn meestal geïmplementeerd pascal of C++ objecten en communicatie met die modules wordt gerealiseerd door het samenvoegen van een dialoog- en een communicatie-object en het aanroepen van methoden inclusief specificaties van hun eigenschappen en methoden.

2.2.3.4 Dialoogconcept

Om een dialoog te beschrijven zijn nu twee dingen noodzakelijk. In de eerste plaats moeten alle mogelijke dialoogdocumenten worden vastgelegd. In de tweede plaats moet worden aangegeven in welke volgorde de verschillende dialoogelementen elkaar opvolgen. Die volgorde is deels van tevoren vast gelegd, maar kan voor een ander deel ook afhankelijk zijn van de gebruikersinvoer.

Soms is de reactie van de gebruiker alleen van belang om te bepalen hoe de dialoog verder zal gaan (dat wil zeggen, wat het volgende dialoog element zal zijn), in andere gevallen is dat niet zo. Een antwoord moet dus ergens bewaard worden voor het gebruik in latere dialoog elementen. Daarom kunnen in een dialoogbestand variabelen worden gedeclareerd. Een dialoogbestand is een tekstbestand. Globaal gezien heeft het bestand drie delen: een variabelen-, een dialoogelement en een defaultelement. Deze delen zijn gescheiden door een lege regel.

De variabelensectie declareert een aantal variabelen voor opslag en gegevens invoer. Deze variabelen kunnen in de rest van het dialoogbestand gebruikt worden. De naam van elke variabele moet op een aparte regel komen, zonder verder gegevens.

Na de variabelen komt de dialoogelementensectie. Daarin worden alle mogelijke dialoogelementen gedefinieerd. Na ieder element zijn de volgende gegevens terug te vinden in het dialoogbestand.

- Naam: ieder dialoogelement moet een unieke naam hebben
- Type: is het dialoogelement een mededeling of een vraag?
- Meldingstekst die het systeem aan de gebruiker doet bij de uitvoering van dit dialoogelement. De meldingstekst kan zowel gewone tekst als fonetische transcriptie zijn.
- Meldingstype: geeft aan of de invoertekst als tekst of als fonetische transcriptie is opgeslagen. Het meldingstype kan eventueel vooraf gegaan door een aanduiding ter indicatie weergegeven worden
- Onderbreekbaarheid: kan de gebruiker de uitvoer van het systeem onderbreken of moet hij altijd de hele melding anhoren?
- Standaard verwijzing: de naam van het volgende element in de dialoog. Bij een vraag wordt eerst gekeken of voor de binnenkomende gebruiksinvoer een speciale verwijzing aanwezig is, zoniet dan wordt de standaardverwijzing gebruikt.

In het geval van een vraag element moeten de volgende bijkomende gegevens in het dialoogbestand staan.

- Lexicon: geeft het lexicon met daarin de mogelijke antwoorden op de gestelde vraag. De naam van het Lexicon wordt letterlijk doorgegeven aan de communicatiemodule. De lexiconnaam moet voorkomen in het lexiconmappingbestand. Zie voor meer informatie de paragraaf over de communicatiemodule.
- Opslagvariabele: de naam van de variabele waarin de gebruikersinvoer wordt opgeslagen.
- Verwijzingen: geven aan wat bij een specifiek antwoord van de gebruiker het volgende element in de dialoog moet worden. Iedere verwijzing koppelt één antwoord van de gebruiker aan een dialoog element.
- Foutverwijzing: de naam van het volgende element in de dialoog, wanneer de invoer van de gebruiker niet herkend kan worden

2.2.4 Lexicon-, uitspraak- en mappingbestanden

De communicatiemodule maakt gebruik van een vrij groot aantal bestanden om zo goed mogelijk spraakherkenning en –synthese op verwerken. In de eerste plaats zijn er de lexiconbestanden. Deze bestanden worden oa. door het programma SpeechPearl gebruikt bij het herkennen van de spraak van de gebruiker. De lexiconbestanden hebben twee functies.

Ten eerste bepalen ze welke woorden herkend kunnen worden. Alleen woorden die in het actieve lexicon zijn opgenomen, worden door SpeechPearl herkend. Wanneer bijvoorbeeld gevraagd wordt naar een politieke partij dan zal het lexiconbestand actief worden gemaakt met daarin de namen van alle politieke partijen waaruit gekozen kan worden. Kiest de gebruiker vervolgens een bepaalde partij en vraagt de computer daarna naar een kandidaat van die partij, dan wordt een lexiconbestand actief met daarin de namen van alle kandidaten van de betreffende partij.

De reden om niet alle mogelijke woorden in een keer te herkennen (dat wil zeggen, alle woorden in één en hetzelfde lexiconbestand te plaatsen), is de volgende. De herkenninggraad van spraakherkenningssoftware wordt hoger, naarmate er minder verschillende woorden mogelijk zijn op een bepaald moment. Wanneer een door de gebruiker uitgesproken kandidaatnaam voor moet komen op de (relatief korte) lijst van een bepaalde partij, is de kans dat die naam correct wordt herkend veel groter dan wanneer het een naam uit de lijst van 'alle' kandidaten van 'alle' partijen kan zijn.

De tweede functie van de lexiconbestanden is om aan te geven hoe alle woorden in zo'n bestand uitgesproken worden. Dit gebeurt door niet alleen de woorden zelf, maar ook hun fonetische weergave op te nemen.

Een tweede groep bestanden zijn de uitspraakbestanden. In feite lijken deze bestanden sterk op de lexiconbestanden, hun doel is vrijwel gelijk aan de tweede functie van de lexiconbestanden: aangeven hoe de uitspraak van een bepaald woord is. Het enige verschil is, dat de uitspraakbestanden gebruikt worden voor spraaksynthese, terwijl de lexiconbestanden voor spraakherkenning bedoeld zijn. De uitspraakbestanden bevatten daarom ook aanduidingen voor klemtoon en intonatie; SpeechPearl-lexiconbestanden bieden niet de mogelijkheid om dergelijke informatie op te slaan.

Een derde soort bestanden die gebruikt worden, zijn gevormd door de mappingbestanden. Soms verwijzen verschillende woorden (of woordgroepen) naar hetzelfde voorwerp of dezelfde persoon. Voor sommige woorden zijn er allerlei synoniemen. De functie van de mappingbestanden is om al zulke woorden te vervangen door één unieke aanduiding. De woorden 'Kok.', 'W. Kok' en 'Wim Kok' bijvoorbeeld verwijzen allemaal naar dezelfde persoon. Om ervoor te zorgen dat het geen verschil maakt welk van deze aanduidingen de kiezer gebruikt, vindt meteen na de spraakherkenning een mappingproces plaats 'Kok' en 'W. Kok' worden afgebeeld ('gemapt') op 'Wim Kok'. Het gevolg is dat het na het mappingproces niet meer uitmaakt wat de gebruiker precies gezegd heeft. De communicatiemodule geeft als resultaat altijd 'Wim Kok' ook als er eigenlijk 'Kok-' of 'W. Kok' herkend werd door .SpeechPearl.

2.2.4.1 Het Lexiconmappingbestand

De invoerfunctie heeft als een van zijn parameters het te gebruiken lexicon. 'SpeechPearl' werkt echter niet rechtevoren met lexiconnamen, maar met talen. Aan iedere taal is vervolgens (onder andere) een lexiconbestand gekoppeld. De SpeechPearl-implementatie van de communicatiemodule gebruikt een lexiconmappingbestand om een lexiconnaam die wordt meegegeven aan de invoerfunctie, te vertalen in een SpeechPearl-taal. Het lexiconmappingbestand is een tekstbestand dat een regel bevat voor iedere mogelijke lexiconnaam. Waarbij lexicon de naam is van een lexicon zoals dat aan de invoerfunctie wordt meegegeven en taalnaam de naam van een resource (taal) die gedefinieerd is voor SpeechPearl.

Door toepassing van een lexiconmappingbestand kan een dialoogbestand worden geschreven (zie vorige paragraaf) zonder dat men weet hoeft te hebben van de werking van 'SpeechPearl'. Ieder lexicon kan een zinvolle naam krijgen, zonder dat rekening gehouden hoeft te worden met de naam van een resource in 'SpeechPearl'. De enige voorwaarde is, dat het lexiconmappingbestand voor ieder lexicon de bijbehorende taalnaam bevat. Deze constructie is niet alleen handig, maar in feite zelfs noodzakelijk: de communicatiemodule is immers niet verplicht om 'SpeechPearl' te gebruiken en dus kan moeilijk één van zijn invoerparameters wel in SpeechPearl-termen gedefinieerd worden.

2.2.5 Programma verloop

2.2.5.1 Programmafases

Gedurende de uitvoering van het programma worden globaal gezien vier fasen doorlopen:

1. initialisatie
2. wachtfase
3. dialoogafhandeling
4. beëindiging

Initialisatie vindt plaats bij het opstarten van het programma; bestanden worden geopend, modules geladen, etc. Vervolgens begint de wachtfase, waarin niets gedaan wordt tot de gebruiker aangeeft wat er moet gebeuren. Wanneer die er voor kiest om een stemprocedure in werking te zetten, begint de derde fase, de dialoogafhandeling. In die fase wordt de stemdialoog met de gebruiker gevoerd. Wanneer die is afgelopen, begint opnieuw een wachtfase. In plaats van een stemprocedure te beginnen kan de gebruiker er ook voor kiezen het programma te beëindigen. In dat geval komt het programma in de vierde fase, waarin de tijdens de initialisatie geopende bestanden weer worden gesloten.

2.2.5.2 Programmaverloop

Programmaverloop is ingedeeld in vier fasen. Wat er gebeurt in drie van die fasen is het onderwerp van het volgende onderdeel. De resterende fase, de wachtfase, blijft buiten beschouwing om de simpele reden dat er eigenlijk nauwelijks iets over te vertellen valt. In die fase wordt namelijk slechts gewacht op een handeling van de gebruiker. Verder gebeurt er in de wachtfase niets, want zodra er actie moet worden ondernomen naar aanleiding van de gebruikersinvoer, komt het programma in een andere fase.

Er wordt heel optimistisch aangenomen, dat er zich tijdens de programma-uitvoering geen problemen voordoen. Helaas zijn fouten in werkelijkheid natuurlijk onvermijdelijk. In ieder geval moet voorkomen worden, dat het programma ten gevolge van een fout vastloopt of ongecontroleerd wordt afgebroken.

2.2.5.3 Communicatie tussen modules

De communicatie tussen de drie modules is betrekkelijk eenvoudig. Bij het opstarten van het programma is in feite alleen de stemmodule nog maar beschikbaar. Door het creëren van een communicatie- en een dialoogobject, worden ook de andere modules echt geladen.

Communicatie met een bepaalde module vindt plaats door het aanroepen van methoden of het aanspreken van eigenschappen van het object dat die module vertegenwoordigt. In enkele gevallen wordt gecommuniceerd via berichten en gebeurtenissen, om asynchroon optredende gebeurtenissen af te handelen. Dit is met name noodzakelijk vanwege de werking van 'SpeechPearl' en 'Fluent Dutch', die beide asynchrone functies kennen. De dialoogmodule laat de stemmodule bijvoorbeeld via een bericht weten dat een dialoog volledig afgehandeld is.

De programma's kunnen gewone Nederlandse tekst omzetten in SAMPA (dit wordt ook wel transcriberen genoemd). SAMPA is een fonetisch schrift dat beschrijft hoe een woord uitgesproken wordt met behulp van tekens die elk een bepaalde klank voorstellen (fonemen). 'SpeechParser' zet Nederlandse tekst alleen om in klanktekens die door 'SpeechPearl' gebruikt worden om een woord te kunnen herkennen. 'Speech Editor' zet de tekst om in zowel klanktekens als accenttekens die door 'Fluent Dutch TTS' gebruikt worden om een woord uit te spreken. Overigens kunnen ook transcripties die door 'SpeechParser' worden gegenereerd met 'Fluent Dutch TTS' uitgesproken worden, zodat enigszins gecontroleerd kan worden of de transcriptie goed ingevoerd is, maar zulke transcripties klinken niet optimaal. In het vervolg wordt gesproken over herkenningstranscripties, die alleen klanktekens bevatten, en uitspraaktranscripties, die zowel klanktekens als accenttekens bevatten.

Herkenningstranscripties worden opgeslagen in lexiconbestanden en uitspraaktranscripties in uitspraakbestanden.

2.2.5.4 Lexiconbestanden

Lexiconbestanden worden door SpeechPearl gebruikt om namen te kunnen herkennen. Wanneer een naam herkend moet worden tijdens het stemproces, dan wordt een bepaalde taal (met één lexiconbestand, één mappingbestand en één uitspraakbestand) geladen. In het lexiconbestand van die taal staan alle woorden die herkend moeten kunnen worden. Een woord wordt opgeslagen door middel van een label en een herkenningstranscriptie. SpeechPearl kan één van de transcripties in het lexiconbestand herkennen en geeft dan het bijbehorende label terug als resultaat.

Hiervoor worden ook variabelen gebruikt, meerbepaald globale variabelen.

2.2.6 Globale variabelen

Er zijn grofweg drie verschillende redenen die hebben geleid tot het besluit om een bepaalde variabele globaal te declareren:

1. standaard-Delphi-gebruik;
2. efficiëntie;
3. call-back-functies van 'SpeechPearl'.

Deze drie redenen zullen hier kort worden toegelicht. Om te beginnen het standaardgebruik in Delphi-programma's. Deze reden is van toepassing op de variabelen die verwijzen naar een vensterobject. Hoewel niet verplicht, is het in Delphi gebruikelijk om voor vensters een globale variabele te reserveren. Het venster is dan overal in het programma "aanspreekbaar". Wanneer in de Delphi-omgeving een venster wordt ontworpen, wordt automatisch in de corresponderende code een globale variabele gedeclareerd.

In de meeste gevallen bestaat er geen zwaarwegende reden om dit gebruik niet toe te passen. Daarom zijn de meeste vensters te bereiken via een globale variabele. Een uitzondering is het stemdialoogvenster, het venster dat getoond wordt tijdens de uitvoering van een dialoog. Dit venster is onlosmakelijk gekoppeld aan een dialoogobject en wordt daarom gerepresenteerd door een (private) eigenschap van een dialoogobject.

In sommige gevallen was een hogere efficiëntie de reden om een bepaalde variabele globaal te maken. Dat geldt bijvoorbeeld voor de variabele 'Logbestand', die zoals de naam al doet vermoeden, het bestand aanwijst waarnaar alle logmeldingen worden weggeschreven. Het logbestand wordt op vele plaatsen in het programma gebruikt. Zonder de globale variabele 'Logbestand' zou iedere procedure/functie die naar het logbestand schrijft (of die een procedure/functie aanroept die dat doet) er een parameter voor dat bestand bijkrijgen. Zomin de efficiëntie van het programma, noch de overzichtelijkheid van de code zouden daarmee gediend zijn.

De derde reden om globale variabelen te gebruiken wordt gevormd door het gebruik van call-back-functies door 'SpeechPearl'. De spraakherkenningsfunctie van 'SpeechPearl' is asynchroon van opzet. Dat wil zeggen dat een programma dat gebruik wil maken van die functie, een 'SpeechPearl' aanroep doet die onmiddellijk terugkeert, vóórdat de eigenlijke herkenning voltooid is. 'SpeechPearl' start een procedure van het aanroepende programma, wanneer het resultaat beschikbaar is. Deze aanroep is asynchroon en valt (daardoor) enigszins buiten de objectgeoriënteerde opzet van een Delphi-programma. De aangeropen procedure kan dan ook (in eerste instantie) alleen gebruik maken van globale variabelen en van lokale procedurevariabelen. Er kan geen sprake zijn van het zonder meer lezen of schrijven van de eigenschappen van een object. Daarvoor moet eerst een verwijzing naar een object zijn verkregen middels een globale variabele.

Aangezien de communicatiemodule gebruik maakt van 'SpeechPearl' voor spraakherkenning en daarom een call-back-functie bevat, is het gebruik van globale variabelen onontkoombaar.

2.2.6.1 Normale condities

Achtereenvolgens zullen we hier drie van de programmafases bekijken: de initialisatie, de dialoogafhandeling en beëindiging van het programma. We zetten de belangrijkste acties die gedurende een bepaalde fase worden ondernomen, op een rijtje. Ook wordt aangegeven welke belangrijke procedures en functies bij die acties betrokken zijn.

2.2.6.2 Initialisatie

De eerste fase, de initialisatie programma, wordt begonnen met het inlezen van de instellingen uit het register. Vervolgens vinden achtereenvolgens de volgende zaken plaats.

1. Het logbestand wordt geopend. Vanaf dit moment worden alle belangrijke gebeurtenissen in het logbestand geregistreerd.
2. Het dialoogbestand wordt geopend.
3. Alle stemstatistieken worden op beginwaarden (meestal nul) ingesteld
4. Een TTS-event wordt gecreëerd. (Met behulp van deze gebeurtenis kan de communicatiemodule vaststellen wanneer Fluent Dutch is "uitgesproken. ")
5. De communicatiemodule wordt geladen.
6. De dialoogmodule wordt geladen.
7. Het dialoogbestand wordt weer afgesloten.
8. Het hoofdvenster wordt op het scherm gezet.

Gedurende al deze gebeurtenissen staat een speciaal opstartvenster op het scherm, waarin de voortgang van de initialisatie zichtbaar wordt gemaakt. Het opstartvenster verdwijnt zodra de initialisatiefase is afgerond.

2.2.6.3 Dialoogafhandeling

De dialoogafhandeling is in feite de fase waarom het allemaal te doen is. In deze fase wordt met de gebruiker gecommuniceerd aan de hand van de aanwijzingen en regels in het dialoogbestand. Het voeren van een dialoog wordt helemaal geregeld door de methode Execute van een dialoogobject. Hoe deze methode werkt en welke andere routines zij aanroept,

In het programma is voor de volgende oplossing gekozen. De methode Execute (in feite dus de hele fase van dialoogafhandeling) wordt uitgevoerd in een eigen, nieuwe stijl. De originele stijl van het programma wordt nu niet meer geblokkeerd door Execute, waardoor voortdurend de message-handler actief kan zijn. Het scherm wordt daardoor keurig bijgewerkt zodra dat nodig is.

Met het asynchroon laten verlopen van de dialoogafhandeling en “de rest van het programma,” is niet alleen een probleem opgelost, maar ook een nieuw probleem geïntroduceerd: hoe weet het hoofdprogramma wanneer een dialoog voltooid is? Dat is van belang omdat het resultaat van de dialoog geregistreerd moet worden en omdat pas een nieuwe dialoog begonnen mag worden zodra de oude is afgelopen. De oplossing is het gebruik van een message, de al genoemde WM-DIAL-FINISHED-boodschap. De laatste handeling binnen de dialoogstijl, de laatste handeling van Execute dus, is het sturen van WM-DIAL-FINISHED naar de hoofdstijl. Zodra de stemmodule die boodschap signaleert, weet hij dat de dialoog beëindigd is, dat het resultaat ervan klaarstaat en dat eventueel een nieuwe dialoog gestart mag worden.

Bij het afsluiten van het programma gebeurt in feite precies het tegenovergestelde van wat er tijdens de initialisatie gebeurt. In de routine `TstemmoduleVenster.Finalise` worden de volgende handelingen verricht.

1. Enige laatste gegevens worden naar het logboek geschreven.
2. Het logboek wordt gesloten.
3. De dialoogmodule wordt beëindigd
4. De communicatiemodule wordt beëindigd.
5. Het TTS-event dat tijdens de initialisatie is gecreëerd, wordt vrijgegeven.

2.2.7 Praktische kant

2.2.7.1 Instelling van de apparatuur

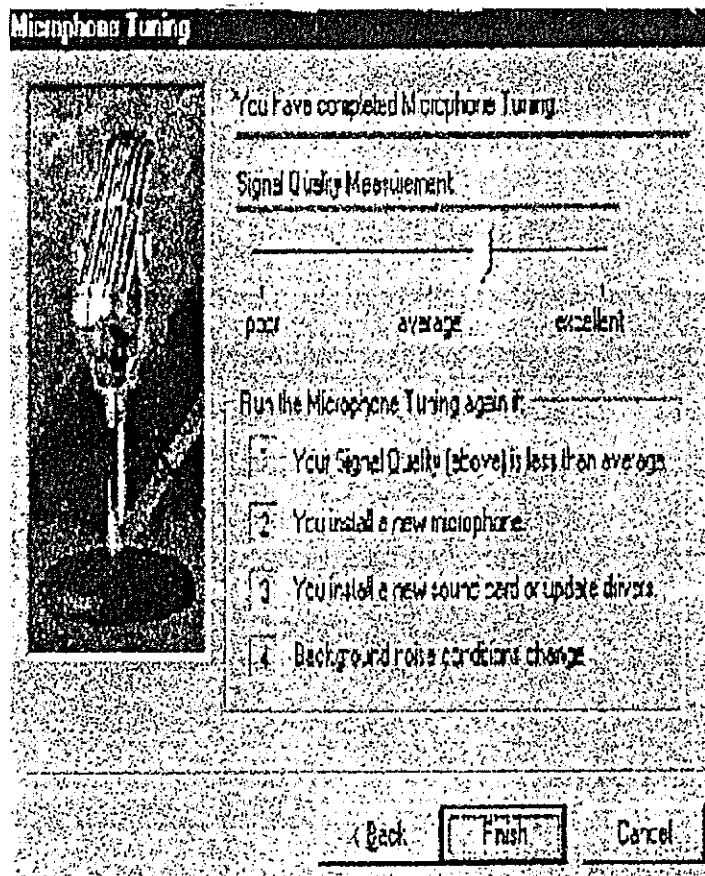
Opdat het programma de stem van de gebruiker juist verstaat moet er vooraf een microfoon testfase zijn. In deze fase wordt het achtergrond geluid van de kamer waarin de computer staat gemeten. Hierdoor weet het programma welk geluid het als ruis moet beschouwen en niet moet beschouwen als een geluid. Hierdoor moet het programma geen onnodige tijd steken in het opzoeken van woorden die afkomstig zijn door achtergrond geluid.

Deze fase moet steeds heruitgevoerd worden als:

- de signaalkwaliteit zeer zwak is
- als er een nieuwe microfoon wordt geïnstalleerd
- als een nieuwe geluidskaart of drivers worden geïnstalleerd
- als de sterkte van het achtergrondgeluid

In deze testfase moet men 2 stappen doorlopen namelijk:

1. men moet de computer een 10tal seconden laten opnemen terwijl er niet gesproken wordt om zo het achtergrond geluid te laten meten
2. Er moet een tekstje worden in gesproken om de sterkte van de stem te laten registreren.
3. Er wordt een kort tekstje uitgesproken door de computer om te horen of de luidsprekers op de goede geluidssterkte staan.



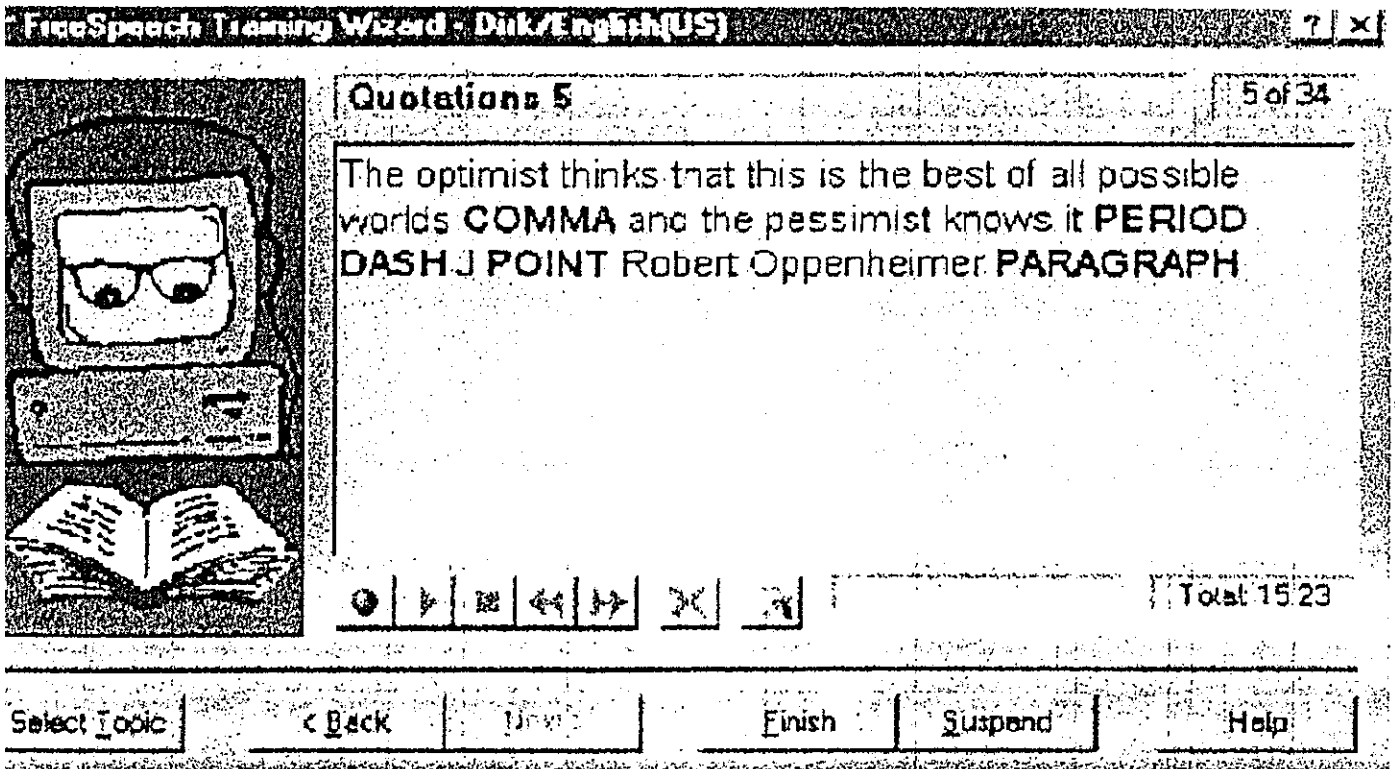
2.2.7.2 Trainingsfase

na het instellen van de apparatuur volgt er nog een trainingsfase. Deze fase zorgt ervoor dat het programma niet enkel je stem verstaat maar ook begrijpt.

Om je stem te begrijpen wordt je verplicht om een bepaalde tijd deze training te volgen. Deze training gaat meestal als volgt:

Je moet een tekst lezen die wordt gegeven door het programma, daarna vergelijkt het programma de tekst met het geluidsfragment dat het heeft opgenomen en dit voor elke tekst.

Na deze training heeft het programma een bepaalde tijd nodig om te leren. In deze tijd vergelijkt het programma alle woorden die uitgesproken zijn met de tekst en stelt zo een eerder vermelde corpus samen.



In freespeech heeft men er een manier op gevonden, zelfs nadat je de trainingsfase hebt volbracht, om het herkenningsniveau te verbeteren:

De adaptatie, dit onderdeel doet net hetzelfde als de training met dat verschil dat het een tekst is die je zelf hebt gedictreed en waarop dan de training wordt op toegepast.

Hoofdstuk 3. Toepassingen en toekomst

3.1 Voxtron Flanders

Voxtron Flanders is een nieuwe divisie van de internationale Voxtron groep. Die bezit kantoren in India, Singapore, Manila, en België en stelt wereldwijd een 80-tal mensen te werk. Het bedrijf is gespecialiseerd in de ontwikkeling van “computer telephony” toepassingen en maakte al naam als pionier op het vlak van de phone banking.

In 1995 besliste Voxtron om één van zijn nieuwe toepassingen (vb.: de earmail) te ontwikkelen op de FLV (Flanders Language Valley), in de nabijheid van de spraak- en taaltechnologie van Lernout & Hauspie. Dit zorgde voor een versnelde ontwikkeling en lancering. Voxtron waardeert ook de wisselwerking tussen de verschillende bedrijven die op de FLV-campus heerst want het is ook nodig in deze snel evoluerende wereld om permanent bij te benen en niet achterop te geraken.

Het principe waarop Voxtron zich baseert is zeer simpel: u vraagt informatie op en u voert routine taken uit via een computersysteem, en dit aan de hand van uw eigen stem en een telefoon.

De programma's die Voxtron geproduceerd heeft zijn: Telebutler, Digilog, Earmail en Call'em

Telebutler:

Een software pakket, waardoor binnenkomende telefonische oproepen snel en automatisch behandeld worden. Het is Voxtron's hoofd programma en is eigenlijk een “automatische secretaresse”. Telebutler is baanbrekend op het vlak van CTI (Computer Telephony Integration). De software is toegankelijk voor niet-programmeurs en het pakket bevat spraakherkenning. Zo wordt de beller automatisch doorverbonden met de persoon waarmee hij wenst te spreken, na het uitspreken van de naam. Telebutler werd voorgesteld op de CT expo '98 in Los Angeles en mag zich nu “best of CT expo98” noemen.

Digilog:

Een softwarepakket waarmee radiofragmenten en telefoonconversaties digitaal worden opgeslagen. De gesprekken zijn nadien eenvoudig op te zoeken en te herbeluisteren.

Earmail:

Earmail is een eenvoudig en gebruikersvriendelijk softwarepakket om e-mail boodschappen via de telefoon te kunnen beluisteren én beantwoorden. E-mail wint razendsnel aan belang als communicatiemiddel, ook de vraag om op de hoogte te blijven van de nieuwe berichten – buiten de kantooromgeving – stijgt. Earmail stuurt naar de GSM een signaal, dankzij integratie van text – to – Speechtechnologie van Lernout & Hauspie Speech Products (LHSP) worden de berichten dan voorgelezen. Momenteel zijn al 8 talen geïntegreerd.

Call'Em:

Nog een ander product van Voxtron heet Call'Em. Dit product gebruikt de L&H Text-to-Speech technologie om e-mails voor te lezen. Nu dat vandaag de dag meer en meer mensen geregeld buiten het kantoor werken, is er duidelijk een groeiende vraag tot de mogelijkheid om e-mails buiten het kantoor te ontvangen. Call'Em is een server die de mensen toelaat een bel verbinding vast te leggen, hun e-mails te downloaden en hen te beluisteren. Je kan zelfs een reactie opnemen en opsturen als wave-bestand.

3.2. SyVox

SyVox Europe (voorheen TCE) ontwikkelt reeds 10 jaar oplossingen voor een perfecte integratie van automatische identificatiesystemen. Vandaag zijn dat spraakherkenning, barcode systemen, draagbare datacommunicatie, handterminals en radio-identificatie.

SyVox biedt hoogtechnologische apparatuur en het nodige advies aan op het vlak van goederen- en personenidentificatie, kwaliteits en productiecontrole. SyVox bespreekt vooraf, tijdens en na het proces nog steeds met het bedrijf als alles in orde is, om te kijken hoe het rendement van het bedrijf gestegen is, indien er nog behoeften zijn,... Er zijn door dat nauw overleg effectief verbeteringen, dat is merkbaar en meetbaar aan het rendement van de arbeid, aan de verlaging van de kosten,... Zo krijgen de mensen de handen vrij om beter en effectiever te presteren.

SyVox presenteert een productenmix die stuk voor stuk van hoogtechnologische aard is. Zo was het de eerste die toestellen met stemherkenning installeerde.

Tevens is Syvox de deskundig in de juiste en veilige integratie van automatische identificatiesystemen. Ze helpen de in- en uitvoer te beheersen op het vlak van de identificatie van goederen en personen in toepassingen zoals: kwaliteits- en productiecontrole, receptie van goederen, stokbeheer en het verzamelen van orders bij parkeerretributie en het lezen van metergegevens.

SyVox is actief in de volgende sectoren:

- automobielandustrie
- chemie en petrochemie
- nutsbedrijven (elektriciteit, gas, water)
- havengebonden activiteiten
- opslag- en overslag
- farmaceutica
- voeding en drank
- meubelindustrie
- politie en overheidsdiensten
- retail

met de volgende toepassingen:

- logistiek
- meterlezing en -beheer
- onderhoud en inspectie
- productie- en kwaliteitscontrole
- order entry
- distributie
- veiligheid

3.3 Softamed nv.

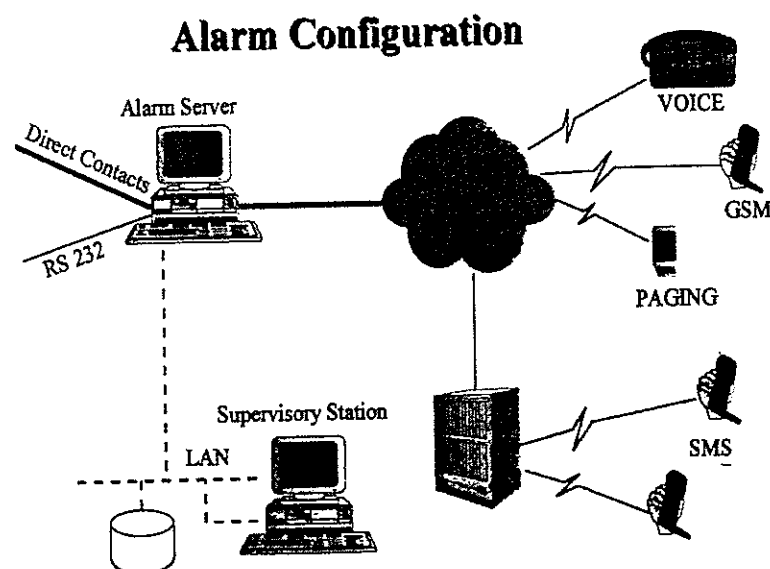
3.3.1 Inleiding

Brand, inbraak, technische problemen, serrebewaking,...zo kan men honderden redenen vinden waarom mensen gewaarschuwd moeten worden. In zulke gevallen is het zeer belangrijk zeker te zijn dat de juiste mensen gewaarschuwd worden op het juiste tijdstip en met het juiste bericht. Vroeger moest een persoon, opgezadeld met die bepaalde taak, al de nodige acties doen. Maar, een persoon kan maar één telefoon plegen naar één plaats per keer. In geval van hoge nood kan dit een zeer zenuwslopende bezigheid zijn, zeker wanneer er meerdere alarmsystemen geactiveerd zijn.

De meeste rampindicators werken op verschillende manieren. In sommige gevallen wordt het binnenkomend signaal gegenereerd door een fysische verandering (binnen of buiten), wat wel meestal het geval is vb. bij brand en/of inbraak systemen. In grotere bedrijven wordt het alarm systeem meestal geëvalueerd door een complex PLC-systeem. De uitvoer van deze systemen is beschikbaar in databases die je kan vinden in een LAN-netwerk.

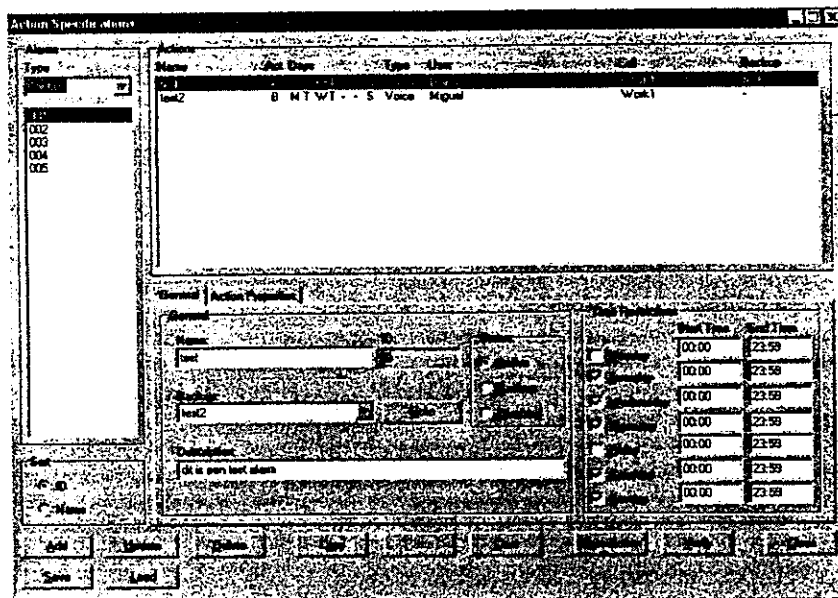
Afhankelijk van de aard van het alarm, tijd van de dag, de toegangsmogelijkheid van de persoon,... zijn er verschillende scenario's mogelijk i.v.m. het soort alarm. Deze scenario's moeten niet alleen rekening houden met het feit dat meerdere personen gewaarschuwd moeten worden, maar ook indien er niet opgenomen wordt, hoe iemand gewaarschuwd moet worden zoals vb. op zijn bureau, thuis, via GSM, via fax,...

Voor al deze doelen, biedt het alarm systeem gebaseerd, op VoiceLink, de ideale configuratie.

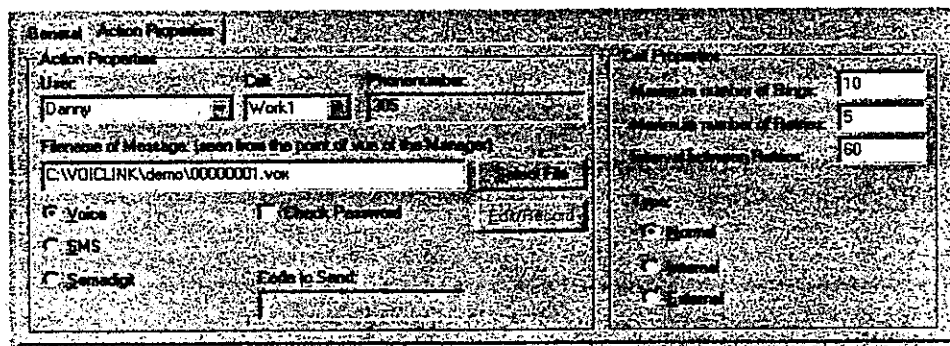


3.3.2 een korte voorstelling

3.3.2.1 Het actiedefinitie veld



In dit scherm is er een overzicht gegeven van de actie definitie voor een bepaald soort alarm. Er kunnen meerdere acties gedefinieerd worden. Iedere actie kan gedefinieerd worden als "active", "backup" of "inactive".



3.3.2.2 de gebruikers defenitie

Om de flexibiliteit te optimaliseren i.v.m. de telefoon nummers, kan er een lijst van nummers gedefinieerd worden voor iedere mogelijke gebruiker. Via een gemakkelijk setup scherm kunnen deze getallen ingegeven en behouden worden.

The screenshot shows a 'User Management' window with the following fields and values:

- General:**
 - Name: Danny
 - Extension: 305
 - Password: 8647
 - Language: English
- PhoneNumbers:**
 - Work 1: 050230300
 - Work 2: 050230305
 - Private 1: 050123456
 - Private 2: (empty)
 - Home 1: 0018665332
 - Home 2: (empty)
 - GSM 1: 075955555
 - GSM 2: 095592332

Buttons on the right side include: Add, Update, Delete, and Close.

3.3.2.3 Alarm prioriteit

Indien er twee alarmsystemen gelijktijdig afgaan, is het ene in het meeste gevallen meer belangrijker dan het andere. Omdat het systeem bedoeld is om op zichzelf te opereren, moeten er prioriteitsregels ingegeven worden in de toepassing. Zo kan er dus voor iedere alarmbron een prioriteitsniveau gedefinieerd worden. Een minder belangrijk alarm zal dus altijd voorrang geven aan een of meer belangrijker alarm.

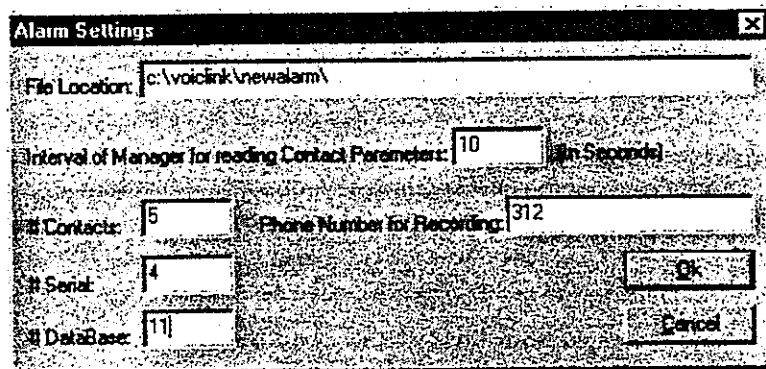
The screenshot shows an 'Alarm Properties' window with the following configuration:

- Type:** Contact
- Sort:** ID (selected), Name
- Alarms List:** 001, 002, 003, 004, 005
- Properties:**
 - Name: Fire
 - Priority: 5 (slider between Low and High)

Buttons at the bottom include: Update and Close.

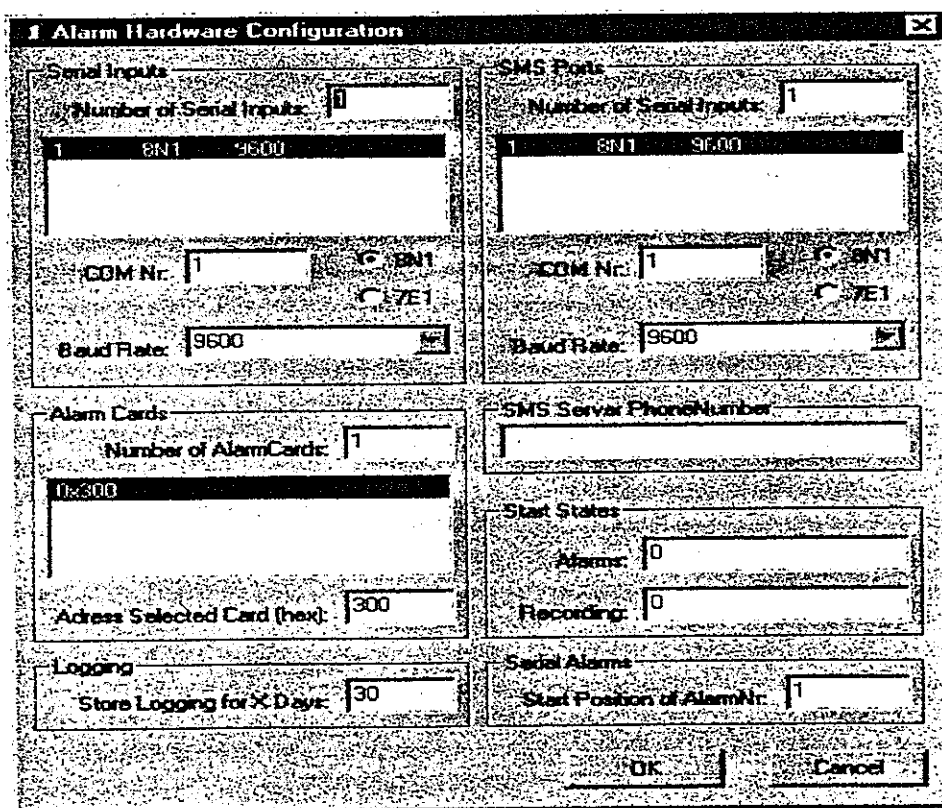
3.3.2.4 Flexibel systeem setup

Doordat alarm bronnen zeer afhankelijk zijn van plaats tot plaats, is er een flexibele setup vereist. Dankzij een installatie hulpmiddel kan dit zeer gemakkelijk geconfigureerd worden.



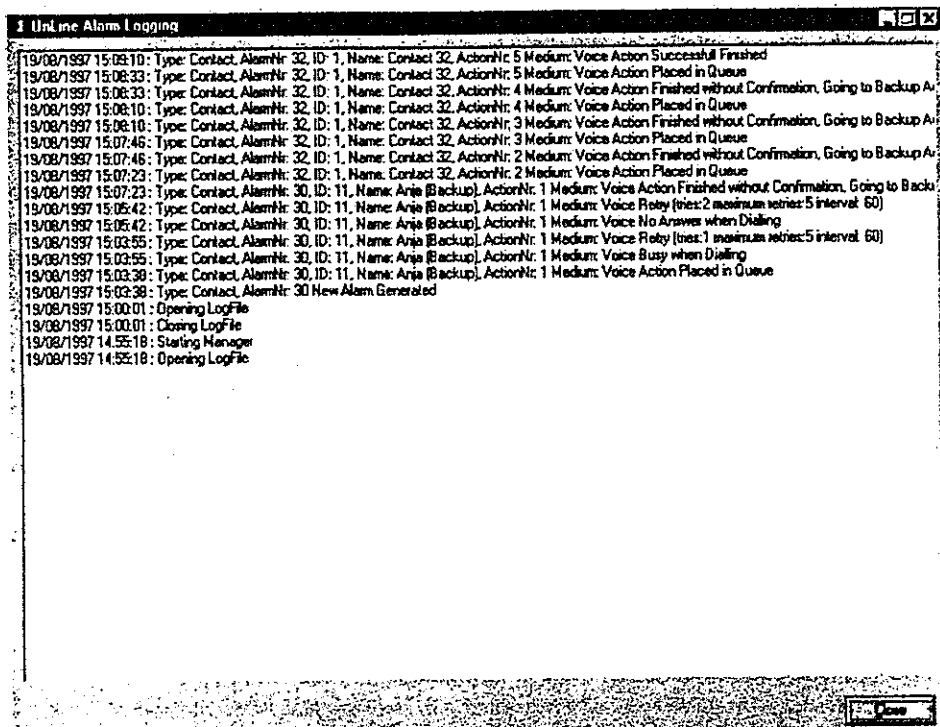
3.3.2.5 systeem configuratie

Omdat meerdere functies ondersteund zijn, is een open interface nodig om te bepalen welke COM-poort gebruikt moet worden voor de inkomende signalen van de alarmen, op welke adressen de alarm-contact-borden zich bevinden,...Hieronder vindt je dit setup hulpmiddel weergegeven.



3.3.3 Logboek

Wanneer de alarm toepassing operationeel wordt, is het zeer belangrijk om een gedetailleerd overzicht bij te houden van wanneer er alarm was in het verleden. De onderliggende figuur geeft dit logboek weer.



```
19/08/1997 15:09:10: Type: Contact, AlarmNr: 32, ID: 1, Name: Contact 32, ActionNr: 5 Medium: Voice Action Successful Finished
19/08/1997 15:06:33: Type: Contact, AlarmNr: 32, ID: 1, Name: Contact 32, ActionNr: 5 Medium: Voice Action Placed in Queue
19/08/1997 15:06:33: Type: Contact, AlarmNr: 32, ID: 1, Name: Contact 32, ActionNr: 4 Medium: Voice Action Finished without Confirmation, Going to Backup A
19/08/1997 15:06:10: Type: Contact, AlarmNr: 32, ID: 1, Name: Contact 32, ActionNr: 4 Medium: Voice Action Placed in Queue
19/08/1997 15:06:10: Type: Contact, AlarmNr: 32, ID: 1, Name: Contact 32, ActionNr: 3 Medium: Voice Action Finished without Confirmation, Going to Backup A
19/08/1997 15:07:46: Type: Contact, AlarmNr: 32, ID: 1, Name: Contact 32, ActionNr: 3 Medium: Voice Action Placed in Queue
19/08/1997 15:07:46: Type: Contact, AlarmNr: 32, ID: 1, Name: Contact 32, ActionNr: 2 Medium: Voice Action Finished without Confirmation, Going to Backup A
19/08/1997 15:07:23: Type: Contact, AlarmNr: 32, ID: 1, Name: Contact 32, ActionNr: 2 Medium: Voice Action Placed in Queue
19/08/1997 15:07:23: Type: Contact, AlarmNr: 30, ID: 11, Name: Anja (Backup), ActionNr: 1 Medium: Voice Action Finished without Confirmation, Going to Backu
19/08/1997 15:05:42: Type: Contact, AlarmNr: 30, ID: 11, Name: Anja (Backup), ActionNr: 1 Medium: Voice Retry (tries:2 maximum retries:5 interval: 60)
19/08/1997 15:05:42: Type: Contact, AlarmNr: 30, ID: 11, Name: Anja (Backup), ActionNr: 1 Medium: Voice No Answer when Dialling
19/08/1997 15:03:55: Type: Contact, AlarmNr: 30, ID: 11, Name: Anja (Backup), ActionNr: 1 Medium: Voice Retry (tries:1 maximum retries:5 interval: 60)
19/08/1997 15:03:55: Type: Contact, AlarmNr: 30, ID: 11, Name: Anja (Backup), ActionNr: 1 Medium: Voice Busy when Dialling
19/08/1997 15:03:39: Type: Contact, AlarmNr: 30, ID: 11, Name: Anja (Backup), ActionNr: 1 Medium: Voice Action Placed in Queue
19/08/1997 15:03:38: Type: Contact, AlarmNr: 30 New Alarm Generated
19/08/1997 15:00:01: Opening LogFile
19/08/1997 15:00:01: Closing LogFile
19/08/1997 14:55:18: Starting Manager
19/08/1997 14:55:18: Opening LogFile
```

3.3.4 technische vereisten

- Windows NT 4.0 PC-stuurprogramma : Pentium 166 MHz with 48 MB RAM
- Genoeg COM-poorten om alarm signalen te ontvangen
- LAN-netwerk voor database toegang
- Meer dan 6 "alarm contact" borden (16 contacten per bord)
- Genoeg telefoon lijnen (analoog of digitaal)

3.4 Creative Labs, Incorporated

De soundblaster 32 PnP is de volgende generatie soundblaster. Deze betaalbare, professionele synthesekaart maakt spelgeluid nog beter en maakt muziek geluid nog realistischer door het gebruik van instrumentale geluid stukken.

In de soundblaster is zelfs een Creative tekstassistentie tekst-to-speechsoftware module geïntegreerd die gebruikmaakt van L&H technologie om 9 voorgeprogrammeerde realistische synthese te maken.

De technologie

- Automatische spraakherkenning

- Tekst-to-speech

- Talen: Amerikaans, Engels, Frans, Duits, Spaans, Italiaans, Koreaans en Nederlands

Het L&H voordeel

- Industieel leidende fonetisch gebaseerde producten

- Combinatie van ASR en TTS op eenzelfde soundblaster

- Meerdere talen

- Een compleet product met ontwikkelingshulpmiddelen

- Onbepaalde, gebruiksvriendelijke woordenlijsten

- Volledige controle over volume, snelheid en sampelsnelheid

Resultaten

De gebruiker kan de instellingen van de tekstlezer en de tekstassistent van binnenin specificeren. Hij kan een nieuwe stem creëren en het bewaren voor verdere toepassingen. De gebruiker kan extreem realistische spraak genereren

3.5 Creator Ltd.

Creator Ltd ontwikkelt en voorziet in levendige objecttechnologieën, een geavanceerde technologie die objecten toelaat om de processor kracht van de pc draadloos te gebruiken op basis van stemgebaseerde, interactieve communicatie met haar gebruikers. De objecten zijn interactief en intelligent: ze spreken met een menselijke stem, verstaan gesproken woorden en reageren met spraak en beweging.

Het levendige object kan een speeltje zijn, een elektronisch toestel, of gelijk welk ander object. Een basis die verbonden is met de computer reguleert de datastroom tussen het object en de pc.

Primaire doelen voor de leef objecttechnologie zijn speelgoed, opvoeding, huisautomatisatie, themagebaseerde bezigheid,...

Creator's productielijn bestaat erin:

Ontwikkelingshulpmiddelen te ontwikkelen

Besturingssoftware te schrijven

Het tekenen, ontwikkelen en integreren van de objecttechnologie

Creator heeft een patent op de levende objecttechnologie.

3.6 probleemstellingen

Wat bij een verkoudheid ?

Als u last heeft van een verkoudheid dan zal het spraakprogramma nog altijd uw opdrachten begrijpen zonder verlies van grote nauwkeurigheid. De snelheid van het programma zal wel een klein beetje dalen maar het is onbeduidend.

Wat bij een accent ?

Een spraakprogramma heeft geen probleem met accenten. Gedurende de training van een spraakprogramma stelt het een lexicon samen gebaseerd op de gebruiker zijn eigen stem. Niet tegenstaande word aangeraden als het niet de gebruiker zijn moedertaal is om meer training te doorlopen dan gevraagd is bij het programma.

Wat bij veel achtergrondgeluid ?

Bij veel achtergrondgeluid krijgt een spraakprogramma het moeilijk om de opdracht te herkennen. Het programma kan een deel van het geluid eruit filteren. Tijdens de microfoon setup word dit achtergrondgeluid gemeten en vergeleken met de spreeksterkte.

Kan ik mijn programma door meerdere personen laten gebruiken ?

Naargelang welk spraakprogramma wordt gebruikt kan men een of meerdere personen hiervan gebruik laten maken. Bv. bij het programma freespeech kan men meerdere personen de trainingsfase laten door lopen , bij het opstarten van freespeech moet men dan zijn eigen lexicon inladen. Er zijn ook programma's die geen training nodig hebben en die dus door praktisch iedereen kunnen gebruikt worden.

Kan ik spraaktechnologie in mijn huidige programma's integreren?

De meeste spraakprogramma's kunnen gebruikt worden in de programma's die reeds in de computer aanwezig zijn. Dit komt doordat het spraakprogramma meestal zelf de gesproken woorden omzet naar tekst en hiervoor is het tekstprogramma niet verantwoordelijk.

Hoeveel geld moet ik uitgeven om een goed spraakprogramma te verkrijgen?

Op dit moment staat spraaktechnologie nog in zijn kinderschoenen, het duurt een hele tijd totdat het programma je goed begint te begrijpen. Er zou dus best nog een tijdje gewacht worden met het aanschaffen van een dergelijk programma, er is niettegenstaande een sterk merkbare evolutie in de spraaktechnologie die zeker nog naar duizelingwekkende hoogte punten zal gaan maar voorlopig kunnen we hier alleen van dromen

3.7 Voordelen van spraaktechnologie

Vaak ontstaan er problemen door het feit dat waarbij communicatie door tekst gewoonweg niet mogelijk is. Een aantal oorzaken kunnen zijn:

- De communicatie is beperkt door de telefoon.
- De persoon waarmee gecommuniceerd wordt is bv. visueel gehandicapt, heel jonge kinderen...
- De gebruiker heeft zijn ogen nodig voor iets anders bv. rijden met de auto.

Spraaksynthese is een techniek die een hele reeks van producten mogelijk maakt. Bovendien zijn er situaties waarin uw product aanzienlijk verbeterd wordt door de toevoeging van spraak-uitvoer . bijvoorbeeld:

- voor e-mail of faxvoorlezer
- als spraakcomponent in uw cd-rompublicatie
- op het gebied van spelletjes (zie bv. Team Fortress 2 waarin je de karakters kunt laten praten)

een ander voordeel van spraaktechnologie is dat personen die geen typervaring hebben kunnen toch tegen een aanzienlijk tempo een tekst op hun computer krijgen. Dit vergde vroeger voor hen veel tijd doordat ze telkens de letters afzonderlijk moesten zoeken.

Besluit

De computer heeft in de voorbije jaren een sterke revolutie meegemaakt. Dit kan geïllustreerd worden door de stap van Edison's fonograaf naar de huidige mp3- technologie.

Samen met deze vooruitgang in de opname is de computersoftware geëvolueerd van de basiswiskunde (+,-,*,/) naar de spraaktechnologie. Over dit laatste zal er in de nabije toekomst zeker nog veel tumult ontstaan. Het is wel duidelijk dat de spraaktechnologie nog in zijn kinderschoenen staat maar de evolutie hierin is zo groot dat we nog veel mogen verwachten.

To be continued...

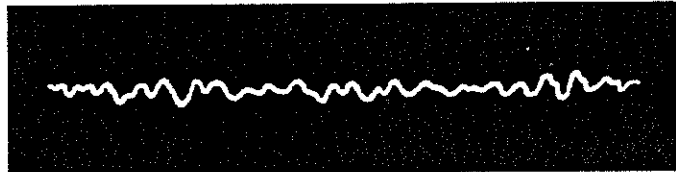
appendix

Appendix: Fourier reeksen

1. Inleiding

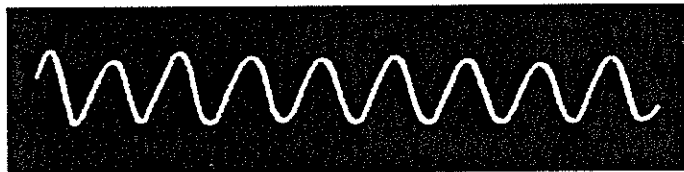
Bij sommige geluiden ontbreekt alle periodiciteit: men spreekt over ruis.

Vb.: Het scheuren van een blad papier voor een microfoon; de weergave van dit geluid op een oscilloscoop ziet eruit als volgt.

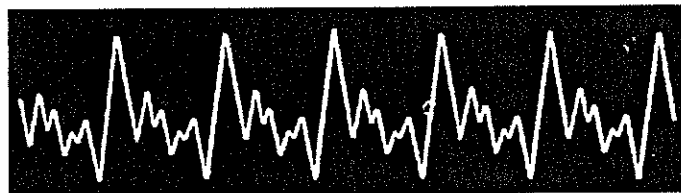


Bij andere geluiden kan men dan weer een periodiek terugkerend patroon waarnemen: men spreekt over tonen(zie uitleg geluid).

Vb.: Het aanslaan van een stemvork geeft een zuiver sinusoïdale voorstelling op een oscilloscoop. Een dergelijke toon noemt men een enkelvoudige toon.

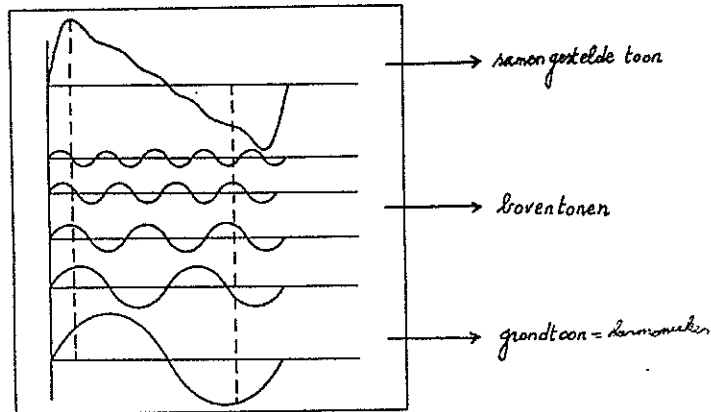


Vb.: Een toon voortgebracht door een muziekinstrument heeft eveneens een periodiek verloop dat echter niet meer sinusoïdaal is. Een dergelijke toon noemt men een samengestelde toon.



Welnu, zo'n samengestelde toon blijkt uit één welbepaalde grondtoon en een reeks van boventonen te bestaan waarvan de frequenties gehele veelvouden van de frequentie van de grondtoon zijn.

De grondtoon heeft over het algemeen de grootste amplitude en bepaalt daardoor de algemene indruk van de samengestelde toon.



Het theorema van Fourier leert ons hoe we de amplitudo's (en frequenties) van de grondtoon en de verschillende boventonen kunnen vinden.

2. Theorema van Fourier

Onderstel dat f een periodieke functie is met periode T , d.w.z. $\forall x \in \text{dom } f: f(x + T) = f(x)$ (1)
De functie f kan dan ontwikkeld worden in een som van cosinussen en sinussen:

$$F(x) = a_0/2 + \sum(a_k \cdot \cos k \cdot 2\pi/T \cdot x + b_k \cdot \sin k \cdot 2\pi/T \cdot x) \quad (2)$$

$$a_k = 2/T \int f(x) \cdot \cos k \cdot 2\pi/T \cdot x \cdot dx \quad (3)$$

$$b_k = 2/T \int f(x) \cdot \sin k \cdot 2\pi/T \cdot x \cdot dx \quad (4)$$

Indien we stellen dat $b_k/a_k = \text{tg } \varphi_k$, dan kan (2) herschreven worden als

$$F(x) = C_0 + \sum C_k \cdot \cos(k \cdot 2\pi/T \cdot x - \varphi_k) \quad (5)$$

Met $C_0 = a_0/2$ en $C_k = a_k/\cos\varphi_k$

De φ_k stellen faseverschillen voor.

3. Spectrum

Wat is nu de periode T_k van de k -de component uit de Fourierreeks (5) ?

$$\cos [k \cdot 2\pi/T \cdot (x + T_k) - \varphi_k] = \cos(2\pi/T \cdot x - \varphi_k)$$

$$\Rightarrow k \cdot 2\pi/T \cdot T_k = 2\pi$$

$$\Rightarrow T_k = T/k$$

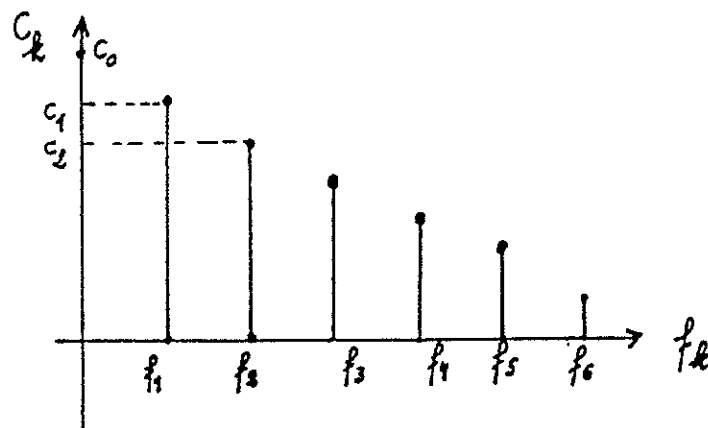
Stellen we de frequentie van de opgegeven functie f voor door : $F_0 = 1/T$

Dan is dus de frequentie F_k van de k -de component uit (5) gegeven door:

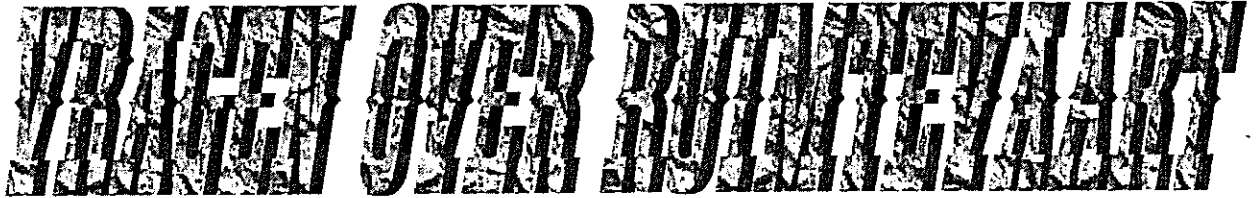
$$F_k = 1/T_k = k/T = k \cdot F_0$$

De eerste component ($k = 1$) uit de Fourierreeks heeft dus steeds dezelfde frequentie als de opgegeven functie f : vandaar de naam grondtoon of fundamentele toon. Men spreekt ook over de eerste harmoniek. De k -de component uit de Fourierreeks noemt men dan de k -de boventoon of k -de harmoniek.

Het zogenoemde spectrum van de periodieke functie f wordt bekomen door in abscis de frequenties f_k uit te zetten en in ordinaat de amplitudo's C_k van de corresponderende harmonieken.



De amplitudo's C_k uit formule (5) kunnen steeds positief gemaakt worden door een extra faseverschil π op te nemen in φ_k . Hierbij nemen de waarden van de amplitudo's C_k af voor toenemende waarden van de frequentie F_k zodat het gerechtvaardigd is de Fourier reeks af te kappen na een voldoende aantal harmonieken.



4. Enkele eenvoudige oefeningen

4.1 Bereken de omlooptijd, de cirkelsnelheid en de ontsnappingsnelheid van een satelliet op 222 km boven het aardoppervlak.

a. Omloopsnelheid

$$\begin{aligned} a &= 222 + 6370 \\ &= 6592 \text{ km} \end{aligned}$$

$$\begin{aligned} T &= 2\pi \sqrt{a^3 / \mu_s} \\ &= 2\pi \sqrt{6592^3 / 398603} \\ &= 5326.41971 \text{ sec} \\ &= 1.47956103 \text{ uren} \\ &= 1\text{h } 28\text{min } 46.42 \text{ sec} \end{aligned}$$

b. Cirkelsnelheid

$$\begin{aligned} V_{\text{cirkel}} &= \sqrt{\mu_s / r} \\ &= \sqrt{398603 / 6592} \\ &= 7.776097228 \text{ km/s} \\ &= 27993.95002 \text{ km/h} \end{aligned}$$

c. Ontsnappingsnelheid

$$\begin{aligned} V_{\text{ont}} &= \sqrt{2\mu / r} \\ &= \sqrt{2 \cdot 398603 / 6592} \\ &= 10.99706216 \text{ km/s} \\ &= 39589.423776 \text{ km/h} \end{aligned}$$

4.2 Hoe ver staat een satelliet in het evenaarsvlak van de aarde die dezelfde omlooptijd heeft als die aarde en bovendien een cirkelbaan heeft (de zgn. Geostationaire baan)?
+ de cirkelsnelheid op de geostationaire baan.

a. Afstand op de evenaar

$$\begin{aligned} 23\text{h } 56\text{min } 4.090\text{ sec} \\ 23\text{h} &= 828009\text{ sec} \\ 56\text{min} &= 33609\text{ sec} \\ 4.090\text{sec} &= \underline{4.090\text{ sec}} \\ \text{som} &= 86164.09\text{ sec} \end{aligned}$$

$$\begin{aligned} T &= 2\pi \sqrt{(a^3 / \mu_s)} \\ \rightarrow a^3 &= (T / 2\pi)^2 \cdot 398603 \\ &= 7.43963872 \cdot 10^{13} \text{ km}^3 \end{aligned}$$

$$\rightarrow a = 42059.19356 \text{ km}$$

Baan op de evenaar is $a - 6370 = 35688.19356 \text{ km}$ vanop het aardoppervlak.

b. Cirkelsnelheid op geostationaire baan

$$\begin{aligned} V_{\text{cirkel}} &= \sqrt{(\mu_s / r)} \\ &= \sqrt{(398603 / 42058.19356)} \\ &= 3.078541226 \text{ km/s} \\ &= 11082.7484136 \text{ km/h} \end{aligned}$$

5. Hohmannbanen

Om van een lage baan (LEO = Low Earth Orbit) naar een hogere orbit (GEO = Geostationary Orbit) te vliegen, maakt men gebruik van een (halve) ellipsbaan aan de LEO en de GEO.

Deze baan werd in 1925 gepubliceerd door de Duitse stadsarchitect en ruimtevaarttheoreticus Walter Hohmann (1880 – 1945) in zijn *Erreichbarkeit der Himmelskörper*. Hij toonde aan dat de orbiswijziging technisch eenvoudig uit te voeren is en bovendien een minimum aan energie vergt.

Nemen we als voorbeeld een satelliet die in het evenaarsvlak op 222 km hoogte boven het aardoppervlak beweegt die we geostationair 35786km boven de evenaar) wensen te maken.

a. Bereken grote as ellips

$$\begin{aligned} r_1 + r_2 &= 6592 + (6370 + 35786) \\ &= 48748 \text{ km} \end{aligned}$$

$$\begin{aligned} r_{\text{per}} &= 6592 \text{ km} \\ r_{\text{ap}} &= 42156 \text{ km} \end{aligned}$$

b. Bereken de cirkelsnelheid in het perigeum

$$\begin{aligned}V_{\text{per cirkel}} &= \sqrt{(\mu_s / r)} \\ &= \sqrt{(398603 / 6592)} \\ &= 7.7786097228 \text{ km/s} \\ &= 27993.95002 \text{ km/h}\end{aligned}$$

c. Bereken de perigeumsnelheid van de Hohmannellips

$$\begin{aligned}a &= (r_1 + r_2)/2 \\ &= 24374\end{aligned}$$

$$\begin{aligned}V_{\text{per}} &= \sqrt{(\mu \cdot r_{\text{ap}} / (a \cdot r_{\text{per}}))} \\ &= \sqrt{(398603 \cdot 42156 / (24374 \cdot 6370))} \\ &= \sqrt{108.2265264} \\ &= 10.40319789 \text{ km/s} \\ &= 37451.51241 \text{ km/h}\end{aligned}$$

d. Bereken het verschil van snelheden in perigeum

$$\begin{aligned}\text{verschil}_{\text{per}} &= V_{\text{per}} - V_{\text{per cirkel}} \\ &= 10.40319789 - 7.7786097288 \\ &= 2.324588162 \text{ km/s} \\ &= 8368.517383 \text{ km/h}\end{aligned}$$

e. Bereken cirkelsnelheid in apogeum

$$\begin{aligned}V_{\text{ap cirkel}} &= \sqrt{(\mu_s / r_{\text{ap}})} \\ &= \sqrt{(398603 / 42156)} \\ &= 3.074967879 \text{ km/s} \\ &= 11069.88437 \text{ km/h}\end{aligned}$$

f. Bereken apogeumsnelheid van de Hohmannbaan

$$\begin{aligned}V_{\text{ap}} &= \sqrt{((\mu_s \cdot r_{\text{per}}) / (a \cdot r_{\text{ap}}))} \\ &= \sqrt{((398603 \cdot 6592) / (24374 \cdot 42156))} \\ &= 1.599125224 \text{ km/s} \\ &= 5756.850806 \text{ km/h}\end{aligned}$$

g. Bereken verschil van snelheden in apogeum

$$\begin{aligned}\text{verschil}_{\text{ap}} &= V_{\text{op cirkel}} - V_{\text{ap}} \\ &= 3.074967879 - 1.599125224 \\ &= 1.475752655 \text{ km/s} \\ &= 5312.709558 \text{ km/h}\end{aligned}$$

h. Bereken totale snelheidswijziging

$$\begin{aligned}\text{Totaal} &= \text{verschil}_{\text{per}} + \text{verschil}_{\text{ap}} \\ &= 2.324588162 + 1.475752655 \\ &= 3.800340817 \text{ km/s} \\ &= 13681.22694 \text{ km/h}\end{aligned}$$

Wiskunde en Ruimtevaart

4.1 Bereken de omlooptijd, cirkelsnelheid en ontsnappingssnelheid van een satelliet op 222 km boven het aardoppervlak.

*Omlooptijd: $T = 2\pi\sqrt{a^3/\mu_s}$
 $a = 222 \text{ km} + 6370 \text{ km} = 6592 \text{ km}$
 $\mu_s = 398603$

$$\begin{aligned}\Rightarrow T &= 2\pi\sqrt{(6592)^3/398603} \\ &= 5326.41371 \text{ sec} \\ &= 1.47956103 \text{ uren} \\ &= 1 \text{ uur } 28 \text{ min } 46.42 \text{ sec}\end{aligned}$$

*omlooptijd: $V_{\text{cirkel}} = \sqrt{\mu_s/R}$
 $= \sqrt{398603/6592}$
 $= 7.776097228 \text{ km/s}$

*ontsnappingssnelheid: $V_{\text{ont}} = \sqrt{2\mu_s/R}$
 $= \sqrt{2(398603)/6592} = 10.99706216 \text{ km}$
 $= \sqrt{2} \cdot 7.776097228 = 10.99706225 \text{ km}$

4.2 Hoe ver staat een satelliet in het evenaarsvlak van de aarde die dezelfde omlooptijd heeft als de aarde en bovendien een cirkelbaan heeft (de zgn. geostationaire baan) + bereken de cirkelsnelheid op geostationaire baan ?

*omlooptijd van de aarde = 23 uren 56 min 4.09 sec

$$23 \text{ uren} = 1380 \text{ min} = 82800 \text{ sec}$$

$$56 \text{ min} = 3360 \text{ sec}$$

$$\Rightarrow \text{omlooptijd van de aarde} = 82800 + 3360 + 4.09 = 86164.09 \text{ sec}$$

$$\begin{aligned}T &= 2\pi\sqrt{a^3/\mu_s} \\ \Rightarrow a^3 &= (T/2\pi)^2 \cdot \mu_s \\ &= (86164.09/2\pi)^2 \cdot 398603 \\ &= 7.43963072 \times 10^{13} \\ a &= 42058.19356 \text{ km}\end{aligned}$$

$$\text{baan van op de evenaar} = a - 6370 \text{ km}$$

$$= 42058.19356 \text{ km} - 6370 \text{ km}$$

$$= 35688.19356 \text{ km van op het aardoppervlak gezien.}$$

*cirkelsnelheid op geostationaire baan =

$$\begin{aligned}V &= \sqrt{\mu_s/R} \\ &= \sqrt{398603/42058.19356} \\ &= 3.078541226 \text{ km/s}\end{aligned}$$

5.a. Bereken de grote as van de ellips

$$\begin{aligned}R &= R_1 + R_2 \\R_1 &= 6370 \text{ km} + 222 \text{ km} = 6592 \text{ km} \\R_2 &= 6370 \text{ km} + 35786 \text{ km} = 42156 \text{ km} \\ \Rightarrow R &= 6592 \text{ km} + 42156 \text{ km} = 48748 \text{ km}\end{aligned}$$

5.b. Bereken de cirkelsnelheid in perigeum

$$\begin{aligned}V_{\text{cirkel}} &= \sqrt{\mu_s/R} \\ &= \sqrt{398603/6592} \\ &= 7.776097228 \text{ km/s}\end{aligned}$$

5.c. Bereken de perigeumsnelheid van de Hohmannellips

$$\begin{aligned}V_{\text{per}} &= \sqrt{\mu \cdot (1 + e)/a \cdot (1 - e)} \\ &= \sqrt{\mu \cdot R_{\text{ap}}/a \cdot R_{\text{per}}} \\ &= \sqrt{398603 \cdot 42164/24382 \cdot 6600} \\ &= 10.21961541 \text{ km/s}\end{aligned}$$

5.d. Bereken het verschil van snelheden in perigeum

$$\begin{aligned}\text{verschil perigeum} &= V_{\text{per}} - V_{\text{per cirkel}} \\ &= 10.21961541 - 7.776097228 \\ &= 2.443518182 \text{ km/s}\end{aligned}$$

5.e. Bereken de cirkelsnelheid in het apogeum

$$\begin{aligned}V_{\text{ap}} &= \sqrt{\mu/R_{\text{ap}}} \\ &= \sqrt{398603/42164} \\ &= 3.074676151 \text{ km/s}\end{aligned}$$

5.f. Bereken de apogeumsnelheid van de Hohmannbaan

$$\begin{aligned}V_{\text{ap}} &= \sqrt{\mu_s \cdot R_{\text{par}}/a \cdot R_{\text{ap}}} \\ &= \sqrt{398603 \cdot 6600/24382 \cdot 42164} \\ &= 1.599693144 \text{ km/s}\end{aligned}$$

5.g. Bereken het verschil van de snelheden in het apogeum

$$\begin{aligned}\text{Verschil ap} &= V_{\text{ap cirkel}} - V_{\text{ap}} \\ &= 3.074676151 - 1.599693144 \\ &= 1.474983007 \text{ km/s}\end{aligned}$$

5.h. Bereken de totale snelheidswijziging

$$\begin{aligned}\text{verschil ap} - \text{verschil per} &= 2.443518182 + 1.474983007 \\ &= 3.918501189 \text{ km/s}\end{aligned}$$

Vertaling Engels

CorrectEnglish

Een schrijfsysteem met Engels als tweede taal.
Chinees, Engels, Frans, Duits, Japans en Spaans kunnen we vertalen naar het Engels

In de hedendaagse internationale computer- en softwareomgeving, is Engels de dominante taal. Om wereldwijd te slagen in de zakenwereld moeten mensen hun geschreven ideeën duidelijk, effectief en in het Engels uit kunnen drukken. Aangezien dit een moeilijke taak is voor de niet-Engels-sprekenden, is er een groeiende vraag naar softwaremiddelen die helpen met het maken van nauwkeurige Engelse taaldocumenten.

CorrectEnglish biedt softwareontwikkelaars een complete Engelse schrijvende hulp die gemakkelijk geïntegreerd kan worden in bijna alle softwareproducten. Het enige waterdichte systeem speciaal ontwikkeld voor personen die als tweede taal Engels hebben. CorrectEnglish helpt meteen de gevorderde niet-Engelssprekenden om hun eigen documenten te maken die duidelijk en effectief communiceren met hun ideeën.

Ga met de leider

Lernout & Hauspie is de marktleider in de productie van de juiste benodigdheden. L&H correctiemiddelen zijn geïnstalleerd op meer dan 70 miljoen computers in de wereld en de spellingscontroletechnologie van L&H is inbegrepen in de standaardproducten van het bedrijf.

Het CorrectEnglish voordeel

CorrectEnglish is uniek. Het specialiseert zich op de speciale noden van niet-Engelstaligen. Iedere taal editie (chinees, Frans, Duits, Japans en Spaans) identificeert de typische grammatica en woordkeuze fouten die gemaakt worden door de personen met Engels als moedertaal wanneer ze Engels schrijven. Een ESL editie is ook verkrijgbaar voor personen met een andere taal.

Plaatselijke voordelen

Alleen CorrectEnglish biedt hulp in de moedertaal van de gebruiker. Dit unieke aspect laat de gebruiker toe om vlug en gemakkelijk van zijn fouten te leren. CorrectEnglish identificeert duidelijk de fouten en legt de regels uit in de gekozen taal en dit in bekende termen. Dus kunnen gebruikers al lezend het probleem en de oplossing begrijpen. Voorbeelden, in het Engels geschreven, bieden meer duidelijkheid en helpen de gebruikers om hun Engels schrijven te verbeteren.

Fout / correctie type voorbeeld

De typische fouten die typisch niet-Engelstaligen maken worden onderverdeeld in 5 categorieën. CorrectEnglish zoekt naar deze fouten en stelt juiste, gemakkelijk verstaanbare, alternatieven voor in de moedertaal van de gebruiker.

1. Werkwoordfout: *"I finished the project if you had asked."*
Correctie: *"I would have finished the project if you had asked."*
2. Voorvoegselfout: *"We will make a decision within the end of the week."*
Correctie: *"We will make a decision by the end of the week."*
3. Woordvormfout: *"I look forward to become associates."*
Correctie: *"I look forward to becoming associates."*
4. Woordvolgordefout: *"What languages speaks she"*
Correctie: *"What languages does she speak"*
5. Woordkeuzefout: *"She is actually applying for a new position."*
Correctie: *"She is currently applying for a new position."*

De mogelijkheid om aan te passen.

Gebruikers kunnen de CorrectEnglish veranderen door de toegang tot documenten gemakkelijker te maken en zo hun productiviteit te verhogen. Om documenten te maken die overeenkomstig zijn met een bepaalde manier van schrijven kunnen de gebruikers de systeemregels aan- en afzetten als dit nodig is. CorrectEnglish het de gebruikers mogelijk hun eigen correctieregels te selecteren, wat dan automatisch op de toekomstige documenten toepast wordt.

Proofing stijlen

Spellingsoptie

De spellingsoptie overloopt vlug een document om fouten te ontdekken in spelling, contextgevoeligheid, kapitalisatie, afkortingen, overeenkomsten en dubbele woorden.

Quikproofoptie

De Quikproof bezit al de functies van de spellingsoptie én corrigeert grammaticale fouten, stiptheidsproblemen, fouten in data, tijd, nummers, formaten,...

Fullproofoptie

De Fullproofoptie biedt correcties aan voor alle schrijfonderwerpen die geïdentificeerd zijn in de Quickproof- en Spellingsoptie én geeft ook hulp bij de meer subjectieve aspecten van het schrijven. Deze bevatten vragen van stijl zoals overbodige clichés en woordexpressies.

Formality Prefences

Gebruikers kunnen kiezen uit drie niveaus naar keuze voor de gewenste toon en stijl van het document.

- Formele contracten, doctorale verhandelingen, voorstellen
- Standaard zakencorrespondentie, persberichten
- Informele persoonlijke brieven, memo's en e-mail

OEM Voordelen

- Een goed geprogrammeerde API laat gemakkelijk de integratie in ieder softwareproduct toe.
- Iedere taaleditie is gelokaliseerd voor de taal die het ondersteunt, dus kunnen OEM's de kostprijs verminderen. De fabricage- en levertijd van een gemakkelijk te gebruiken product heeft geen standaardtaal nodig
- Beide speel- en compileertijdopties zijn ingeschakeld bij de OEM zodat het product kan aangepast worden aan de specifieke noden van de markt
- OEM's kunnen dit product gebruiken om, in de markt met Engels als tweede taal uit te breiden.

Gebruiksvoordelen

- Versnellen van het schrijfproces door het identificeren van fouten en het geven van oplossingen voor verbetering van de taal.
- Verhogen van de productiviteit door het verminderen van de nood om documenten te laten controleren door personen met Engels met moedertaal.
- Verbeteren van de kwaliteit van de communicaties door het voor de gebruikers mogelijk te maken om hun teksten juist te verbeteren
- Helpt gebruikers te leren van hun fouten en om hun Engels schrijven te verbeteren
- Biedt correctie voor Britse én Amerikaanse schrijfovereenkomsten

*©1998 Lernout & Hauspie praat producten N.V. Alle rechten voorbehouden
kopieer- en gebruiksrechten

Woordverklaring:

| Woord zoals gevonden | Grondvorm | Betekenis |
|----------------------|----------------|---------------------|
| Environment | Environment | Omgeving |
| Predominant | Predominate | Overheersen |
| Succeed | To succeed | Slagen |
| Effectively | Effective | Effectief |
| Demand | To demand | Vragen |
| Aid | To aid | Helpen |
| Accurate | Accurate | Nauwkeurig |
| Integrate | To integrate | Integreren |
| Virtually | Virtual | Virtueel |
| Proofing | Proof | Bewijs |
| Intermediate | Intermediate | Meteen |
| Developer | To develop | Ontwikkelen |
| Focuses | To focus | Concentreren |
| Native | Native | Moeder(taal) |
| Easily | Easy | Gemakkelijk |
| Familiar | Familiar | Gekende |
| Readily | ready | bereid |
| Provide | To provide | Voorzien |
| Clarification | Clarification | Verklaring |
| Alternatives | Alternative | Alternatief |
| Explanations | To explain | Uitleggen |
| Tailor | Tailor | Kleermaker |
| Increase | To increase | Verbeteren |
| Acceptable | To accept | Aanvaarden |
| Conforming | To conform | Bevestigen |
| Reflect | To reflect | Reflecteren |
| Capitalization | To capitalize | Kapitalisatie |
| Abbreviation | Abbreviation | Afkorting |
| Punctuation | Punctuation | Punctueel |
| Subjective | Subjective | Subjectief |
| Redundancies | Redundancy | Overtollig |
| Desired | To desire | Verlangen |
| Dissertations | Dissertation | Verhandeling |
| Implementation | Implementation | Integratie |
| Localized | Localize | Locale |
| Reviewed | To review | Herbekijken |
| Enabling | To enable | In staat stellen |
| Specific | Specific | Specifieke |
| Refine | To refine | Zuiveren, verfijnen |
| Conventions | Convention | Overeenkomst |

Vertaling Duits

De revolutionaire technologie van IFABO: Mobilkom presenteert de eerste gesproken e-mail over A1.

Mobilkom Australië AG brengt een technische innovatie van buitengewone kwaliteit op de Oostenrijkse markt. De "Voice Butler" in A1-pocketnet zal in de toekomst de mogelijkheid bieden om e-mails met de "A1-Handy" te horen in gesproken vorm en langs dezelfde weg terug te sturen. Met de internationaal beproefde tekst-naar-spraak-technologie zet dit nieuwe prototype schriftelijke meldingen om in woorden, die door een vrouwelijke computerstem omgezet worden.

Binnenkort wordt deze technische vernieuwing op de IFABO van dit jaar voorgesteld. Met een druk op een knop kan men de gewenste e-mails, die men in het huis of op het bureau ontvangt, door "Handy" laten voorlezen. Voorwaarde is: de eenmalige instelling van het bevel "alle e-mails automatisch doorsturen" aan het persoonlijke e-mail adres, dat elke "A1-Handy"- bezitter automatisch heeft. Het bericht wordt dan gehoord op het nummer 0664/68000. Wanneer je dan wilt antwoorden, spreek dan de willekeurige tekst in met "Handy" en die stuurt hem dan als e-mail terug. Heeft de ontvanger een multimedia-pc, dan kan hij het antwoord ook horen.

In de toekomst zal men daarvoor noch communicator noch laptop nodig hebben.- Dat alleen al is door "Handy" mogelijk. Door A1-pocketnet is "A1-Handy" automatisch met het internet verbonden. De toegang is kosteloos, dit betekent dat, de klant zonder een bijkomend contract met een service-provider, het volle internet Account over het EU-net heeft. De kosten zijn alleen de verbingskosten aan deze A1-tarieven.

Op dit moment wordt deze nieuwe technologie nog getest. De productie-aanvang is voor de zomer van 1998 gepland.

| Woord zoals gevonden | Grondwoord | Betekenis zoals ik denk | Controle |
|----------------------|----------------|-------------------------|-----------------------|
| Revolutionäre | Revolutionär | revolutionair | Revolutionair |
| präsentiert | präsentieren | presenteren | Presenteren |
| sprechende | sprechen | spreken | Spreken |
| Zukunft | Zukunft | toekomst | Toekomst |
| Möglichkeit | Möglichkeit | mogelijkheid | mogelijkheid |
| abzuhören | abhören | verhoren | beluisteren |
| zurückzuschicken | zurückschicken | terugsturen | terugsturen |
| erproben | erproben | proberen | beproeven |
| übersetzt | übersetzen | overzetten | toezenden |
| weiblichen | weiblich | vrouwelijk | vrouwelijk |
| gewünschten | wünschten | gewoonten | * |
| Voraussetzung | Voraussetzung | veronderstellen | onderstelling |
| weiterleiten | weiterleiten | verder doen | verder leiden |
| Nachricht | Nachricht | * | tijding, bericht |
| sofort | sofort | zo verder | dadelijk, terstond |
| spricht | sprechen | spreken | spreken |
| schickt | schicken | * | zenden, sturen |
| Empfänger | Empfänger | ontvanger | ontvanger |
| anhören | anhören | aanhoren | beluisteren, aanhoren |
| braucht | brauchen | * | nodig hebben |
| einfache | einfach | eenvoud | eenvoud |
| möglich | möglich | mogelijk | mogelijk |
| Zugang | Zugang | toegang | toegang |
| zusätzlichen | zusätzlich | aanvullend | aanvullend |
| entstehen | entstehen | ontstaan | ontstaan |
| lediglich | lediglich | enkel | enkel |
| Derzeit | Derzeit | nu | nu, destijds |
| getestet | testen | getest | testen |
| erstmals | erstmals | Binnenkort | Voor de eerste maal |

* = stond niet in woordenboek

Logboek Seys Michiel

| | | |
|-------------------|---------|---|
| 5 september 1998 | 5 uur | Informatie verzamelen bij jobevent L&H te Ieper |
| 6 september 1998 | 4 uur | Informatie verzamelen en voorlopige inhoudstabel maken |
| 8 september 1998 | 3 uur | Informatie lezen |
| 12 september 1998 | 3 uur | Informatie verzamelen over geluid in bib. Torhout |
| 12 september 1998 | 4 uur | Informatie lezen |
| 13 september 1998 | 6 uur | samenvatten boeken over band |
| 20 september 1998 | 7 uur | samenvatten boeken over cd |
| 24 september 1998 | 4 uur | Vragen ruimtevaart oplossen |
| 28 september 1998 | 3 uur | Vragen ruimtevaart oplossen |
| 3 oktober 1998 | 2.5 uur | Infomatie zoeken op internet in bib. Torhout |
| 14 oktober 1998 | 2 uur | lezen informatie |
| 24 oktober 1998 | 2 uur | Informatie verzamelen over geluid in bib. Torhout |
| 24 oktober 1998 | 4 uur | Lezen informatie band |
| 29 oktober 1998 | 4 uur | Informatie verzamelen over geluid en spraaktechnologie in bib. Brugge |
| 3 november 1998 | 5 uur | Intikken inleiding 'Geluid op band' |
| 4 november 1998 | 3 uur | Informatie verzamelen bij Keyware technologies te Brugge |
| 4 november 1998 | 4 uur | Vertalen informatie verkregen bij Keyware |
| 11 november 1998 | 4 uur | Lezen infomatie toepassingen |
| 25 november 1998 | 4 uur | Infomatie verzamelen bij Softamend te Torhout |
| 5 december 1998 | 3.5 uur | Informatie zoeken op het internet over toepassingen |
| 6 december 1998 | 2 uur | Voorlopige informatie over toepassigen intikken |
| 12 december 1998 | 3 uur | Lezen informatie Softamend |
| 21 december 1998 | 4 uur | Inscannen foto's 'Geluid op band' |
| 20 december 1998 | 2.5 | Lezen boeken band + intikken |

| | | |
|------------------|---------|--|
| 21 december 1998 | 2.5 | Intikken boeken band |
| 22 december 1998 | 2 uur | Intikken boeken band |
| 23 december 1998 | 2.5 uur | Intikken boeken band |
| 24 december 1998 | 2 uur | Intikken boeken band |
| 27 december 1998 | 2.5 uur | Intikken boeken band |
| 29 december 1998 | 2 uur | Intikken boeken band |
| 2 januari 1999 | 2.5 uur | Intikken boeken band |
| 5 februari 1999 | 2 uur | Controleren |
| 27 februari 1999 | 6 uur | Geluid + band uitwerken |
| 28 februari 1999 | 5 uur | Band + cd uitwerken |
| 6 maart 1999 | 3 uur | Controleren van voorlopige inleiding |
| 10 maart 1999 | 3 uur | Opzoeken internet i.v.m. L&H |
| 13 maart 1999 | 8 uur | Naar L&H te ieper om info te verzamelen over fonetische engines + toebehoren |
| 21 maart 1999 | 3 uur | Opstellen voorlopige inhoudstabel |
| 2 april 1999 | 5 uur | Intikken 1 ^e deel over 'digitaal en analoog openemen' |
| 24 april 1999 | 10 uur | Vertalen + samenvatten info verkregen van Dieter Boulard |
| 4 mei 1999 | 11 uur | Spraaktechnologie uitbreiden op basis van verkregen info |
| 5 mei 1999 | 4 uur | Spraaktechnologie uitbreiden op basis van verkregen info |
| 9 mei 1999 | 4 uur | Toepassingen verbeteren + toevoegen |
| 15 mei 1999 | 3 uur | Studeren Fourierreeksen |
| 16 mei 1999 | 5 uur | Intikken toepassingen |
| 17 mei 1999 | 3 uur | Studeren Fourierreeksen |
| 22 mei 1999 | 5 uur | Fourierreeksen + uitdiepen geluid |
| 30 mei 1999 | 13 uur | Alles aanvullen + volledige opmaak + woordvooraf |

Logboek Simpelaere Dennis

| | | |
|-------------------|---------|---|
| 5 september 1998 | 5 uur | Informatie verzamelen bij jobevent L&H te Ieper |
| 6 september 1998 | 4 uur | Informatie verzamelen en voorlopige inhoudstabel maken |
| 9 september 1998 | 3 uur | Informatie lezen |
| 12 september 1998 | 3 uur | Informatie verzamelen over geluid in bib. Torhout |
| 12 september 1998 | 4 uur | Informatie lezen |
| 13 september 1998 | 6 uur | samenvatten boeken over band |
| 20 september 1998 | 7 uur | samenvatten boeken over cd |
| 24 september 1998 | 4 uur | Vragen ruimtevaart oplossen |
| 28 september 1998 | 3 uur | Vragen ruimtevaart oplossen |
| 3 oktober 1998 | 2.5 uur | Infomatie zoeken op internet in bib. Torhout |
| 14 oktober 1998 | 2.5 uur | Vertalen informatie |
| 24 oktober 1998 | 2 uur | Informatie verzamelen over geluid in bib. Torhout |
| 24 oktober 1998 | 4 uur | Lezen informatie cd-speler en dvd-speler |
| 29 oktober 1998 | 4 uur | Informatie verzamelen over geluid en spraaktechnologie in bib. Brugge |
| 3 november 1998 | 5 uur | Intikken inleiding 'Geluid op band' |
| 4 november 1998 | 3 uur | Informatie verzamelen bij Keyware technologies te Brugge |
| 4 november 1998 | 4 uur | Vertalen informatie verkregen bij Keyware |
| 11 november 1998 | 4 uur | Lezen infomatie toepassingen |
| 25 november 1998 | 4 uur | Infomatie verzamelen bij Softamend te Torhout |
| 5 december 1998 | 3.5 uur | Informatie zoeken op het internet over toepassingen |
| 6 december 1998 | 2 uur | Voorlopige informatie over toepassigen intikken |
| 12 december 1998 | 3 uur | Lezen informatie Softamend |
| 21 december 1998 | 4 uur | Inscannen foto's 'Geluid op band' |

| | | |
|------------------|---------|--|
| 24 december 1998 | 4 uur | Contact maken met Patrick Tomas bij L&H education |
| 5 februari 1999 | 2 uur | Controleren |
| 27 februari 1999 | 6 uur | Geluid + band uitwerken |
| 27 februari 1999 | 1 uur | Corresponderen met Dieter Boullard bij L&H education |
| 28 februari 1999 | 5 uur | Band + cd uitwerken |
| 6 maart 1999 | 3 uur | Controleren van voorlopige inleiding |
| 10 maart 1999 | 3 uur | Opzoeken internet i.v.m. L&H |
| 13 maart 1999 | 8 uur | Naar L&H te ieper om info te verzamelen over fonetische engines + toebehoren |
| 21 maart 1999 | 3 uur | Opstellen voorlopige inhoudstabel |
| 23 maart 1999 | 2.5 uur | Lezen informatie L&H |
| 24 maart 1999 | 3 uur | Lezen informatie L&H |
| 23 maart | 1 uur | Downloaden internetprogramma : GetRight |
| 28 maart 1999 | 1.5 uur | Afhalen bestand Cool Edit (analyse programma)+ |
| 2 april 1999 | 5 uur | Intikken 1 ^e deel over 'digitaal en analog openemen' |
| 8 april 1999 | 2 uur | Begin intikken uit natuur en techniek |
| 12 april 1999 | 4 uur | Afwerken natuur en techniek |
| 14 april 1999 | 2.5 uur | Uitbreiden met eerder verkregen info van natuur en techniek |
| 17 april 1999 | 3 uur | Programma Cool Edit uitpluizen |
| 24 april 1999 | 10 uur | Vertalen + samenvatten info verkregen van Dieter Boulard |
| 4 mei 1999 | 11 uur | Spraaktechnologie uitbreiden op basis van verkregen info |
| 9 mei 1999 | 4 uur | Toepassingen verbeteren + toevoegen |
| 15 mei 1999 | 3 uur | Studeren Fourierreeksen |
| 17 mei 1999 | 3 uur | Studeren Fourierreeksen |
| 22 mei 1999 | 5 uur | Fourierreeksen + uitdiepen geluid |
| 30 mei 1999 | 13 uur | Alles aanvullen + volledige opmaak + woordvooraf |

Bronnen

- * Encarta '98
- * Het menselijk lichaam
Tony smith
Davidsfonds/ Leuven
- * Internet
- * Cursus fysica 6^e jaar
- * Cursus wiskunde 6^e jaar
- * Geluid op band
Plantyn
- * Studie cursus L&H education