

**Näkövammaisten ihmisten ottaminen huomioon
puhekäyttöliittymän suunnittelussa – Tapaus ”Aistiboxi”**

Liisa Huotari

Tampereen yliopisto
Tietojenkäsittelytieteiden laitos
Tietojenkäsittelyoppi
Pro gradu –tutkielma
Ohjaaja: Markku Turunen
Lokakuu 2006

Tampereen yliopisto

Tietojenkäsittelytieteiden laitos

Tietojenkäsittelyoppi

Liisa Huotari: Näkövammaisten ihmisten ottaminen huomioon puhekäyttöliittymän suunnittelussa – Tapaus ”Aistiboxi”

Pro gradu -tutkielma, 59 sivua, 1 liitesivu

Lokakuu 2006

Tässä tutkielmassa käsitellään sitä, miten näkövammaiset käyttäjät tulee ottaa huomioon puhekäyttöliittymän suunnittelussa. Tutkielmassa luodaan katsaus nykyiseen puheteknologiaan, puhekäyttöliittymien suunnitteluperiaatteisiin ja arviointimenetelmiin, sekä näkövammaisten erityistarpeisiin puhekäyttöliittymien käyttäjinä. Erityistä huomiota kiinnitetään sellaisiin puhekäyttöliittymiin, jotka sisältävät sekä puhetta että visuaalisia elementtejä. Esimerkkinä puhekäyttöliittymän suunnittelusta näkövammaisille esitellään Aistiboxi-hanke sekä hankkeen puitteissa toteutetun näkövammaisten haastattelun tulokset. Hankkeen aikana toteutetun digiboksin käyttöliittymän prototyypin pohjalta esitellään suunnitelma näkövammaiset huomioon ottavasta käyttöliittymästä.

Avainsanat ja -sanonnat: puhekäyttöliittymät, suunnittelu, toteutus, saavutettavat käyttöliittymät, näkövammaisten käyttöliittymät.

Sisällys

| | | |
|----------|---|-----------|
| 1 | Johdanto | 1 |
| 2 | Nykyinen puheteknologia | 3 |
| 2.1 | Puheentunnistus | 4 |
| 2.2 | Puhesynteesi | 7 |
| 3 | Puhekäyttöliittymien suunnitteluperiaatteet | 10 |
| 3.1 | Dialoginhallinta | 11 |
| 3.1.1 | Keskustelun kontrolli | 11 |
| 3.1.2 | Keskustelun aloite | 13 |
| 3.2 | Käyttäjän ohjaaminen puhetulosteiden avulla..... | 13 |
| 3.2.1 | Kehotusten muotoilu | 14 |
| 3.2.2 | Varmennusmenetelmät..... | 15 |
| 3.3 | Virheiden käsittely..... | 17 |
| 3.4 | Puhuvat valikot..... | 20 |
| 3.5 | Puhekäyttöliittymät TV-ympäristössä | 22 |
| 4 | Puhekäyttöliittymien arviointimenetelmät | 24 |
| 4.1 | Teknologian arviointi..... | 24 |
| 4.2 | Puhesovellusten arviointi | 25 |
| 5 | Näkövammaisten ihmisten ottaminen huomioon puhekäyttöliittymän toteutuksessa | 27 |
| 5.1 | Näkövammaisuuden määrittely | 27 |
| 5.2 | Puhekäyttöliittymien suunnitteleminen näkövammaisille | 28 |
| 6 | Tapaus ”Aistiboxi” | 31 |
| 6.1 | Kohderyhmän haastattelu..... | 31 |
| 6.1.1 | Television katselu | 32 |
| 6.1.2 | Television ja digiboksin ohjaaminen kaukosäätimen tai etupaneelin avulla | 32 |
| 6.1.3 | Digiboksin ominaisuudet ja erityistoiminnot | 34 |
| 6.1.4 | Digiboksin valikot..... | 35 |
| 6.1.5 | Yhteenveto haastattelun tuloksista..... | 36 |
| 6.2 | Käyttöliittymän suunnitellut ominaisuudet | 37 |
| 6.3 | Käyttöliittymän toteutus..... | 37 |
| 6.3.1 | Laitteen käynnistäminen ja sammuttaminen | 38 |
| 6.3.2 | Valikot ja niissä navigointi yleisesti | 39 |
| 6.3.3 | Päähakemisto | 43 |
| 6.3.4 | Alavalikot | 44 |
| 6.3.5 | Kaukosäädin ja etupaneeli | 48 |
| 6.3.6 | Käyttöliittymän ohjaaminen puheella | 50 |
| 6.3.7 | Yhteenveto suunnitelluista toiminnoista ja valikoista..... | 50 |
| 6.4 | Tulevaisuuden näkymiä | 53 |

| | |
|---|-----------|
| 7 Yhteenveto | 54 |
| Viiteluettelo..... | 56 |
| Liite 1. Aistiboxin suunnitellut ominaisuudet..... | 60 |

1 Johdanto

Teknologian kehittymisen myötä yhteiskunta sähköistyy kovaa vauhtia. Yritysten ja julkisen sektorin toiminta nojaa lisääntyvässä määrin tietojärjestelmiin ja yleensäkin tietotekniikan käyttöön. Internetin välityksellä ihmiset pääsevät palvelujen äärelle, joiden käyttäminen ilman Internetiä olisi joko mahdotonta tai todella vaikeaa. Kehitysaskelaita on tapahtunut myös televisiotoiminnassa digitelevision tuleminen myötä: digitelevision käyttöliittymän toiminnot, muun muassa ohjelmaopas, tuovat lisäarvoa television katseleluun, ja nyt myös television kautta on mahdollista saavuttaa uusia palveluita. Jotta mahdollisimman monet erityisryhmät, kuten näkövammaiset, olisivat tasavertaisessa asemassa yhteiskunnassa, täytyy olla saatavilla sellaisia laitteita ja ohjelmistoja, joiden avulla myös erityisryhmät pääsevät tietotekniikan mahdollistamien palveluiden äärelle. Palveluiden tulee siis olla saavutettavia.

Bergmanin ja Johnsonin [1995] mukaan saavutettavuus on sellaisten esteiden poistamista, jotka estävät eri tavoin vammaisia ihmisiä osallistumasta olennaisiin elämän aktiviteetteihin, kuten palveluiden, tuotteiden ja tiedon käyttämiseen. Käyttöliittymissä saavutettavuuden täytyy tarkoittaa myös sellaisten käyttöliittymien suunnittelua, jotka ovat helppoja käyttää sekä vammaisille käyttäjille, että käyttäjille, joilla ei ole erityisiä vammoja. Tällainen suunnittelu onnistuu ottamalla huomioon eri ryhmien tarpeet siinä vaiheessa, kun ohjelmistoa suunnitellaan.

Näkövammaisten on mahdollista käyttää erilaisia apuvälineitä, kuten pistenäytöt ja ruudunlukuohjelmat, joiden avulla he pystyvät käyttämään monia niin sanottuja tavallisia käyttöliittymiä. Tutkielmassani keskityn kuitenkin tutkimaan sellaisia käyttöliittymiä, joita myös näkövammaiset voivat käyttää ilman edellä mainitun kaltaisia välikäsiä. Puhekkäyttöliittymät on käyttöliittymätyyppi, jota myös sokeat ja vaikeasti heikkonäköiset voivat käyttää ilman apuvälineitä ja -ohjelmistoja.

Puhekkäyttöliittymillä tarkoitetaan sellaisia käyttöliittymiä, joissa vuorovaikutuksen välineenä käytetään joko puhetulostusta, puhesyötteitä tai näitä molempia. Puhekkäyttöliittymiä voidaan siis ohjata joko puheella tai jonkin tyyppisillä näppäimillä. Jotta myös näkövammaiset pystyisivät käyttämään näppäimillä ohjattavia puhekkäyttöliittymiä sujuvasti, täytyy niiden suunnittelijoiden ottaa huomioon näkövammaisten erityistarpeet tietotekniikan käyttäjinä. Tällainen panostus on hyödyllinen muillekin kuin näkövammaisille, sillä erityisryhmät huomioon ottavan ohjelmiston suunnitteleminen tekee ohjelmistosta usein käytettävämmän myös tavallisille käyttäjille, joilla ei ole merkittäviä tietokoneen käyttöä haittaavia vajavaisuuksia [Bergman and Johnson, 1995].

Tutkielmassani selvitän puhekäyttöliittymien ominaisuuksia ja suunnitteluperiaatteita sekä sitä, miten niiden toteutuksessa tulee ottaa näkövammaiset ihmiset huomioon. Esimerkkinä myös näkövammaisille suunnatun käyttöliittymän suunnittelusta ja toteutuksesta esittelen ”Digital Television for All” -nimisen hankkeen, jota tutkielmassani kutsun lyhyemmin nimellä ”Aistiboxi”. Arvioin hankkeen pilottivaiheessa toteutetun käyttöliittymän käytettävyyttä tutkielmani teoriaosuuden sekä hankkeessa tehtyjen käyttäjäröhmän haastattelujen perusteella. Hanke toteutetaan yhteistyössä Turun yliopiston, Taideteollisen korkeakoulun, Tampereen yliopiston ja Teknillisen korkeakoulun kanssa, ja pilottivaiheessa hankkeen tarkoituksena on suunnitella erityisesti näkövammaisten ja vanhusten käyttöön soveltuva digisovitin, eli digiboksi.

Tutkielmassani luon ensiksi katsauksen nykyiseen puheteknologiaan ja kartoitan kirjallisuuden perusteella puhekäyttöliittymien suunnitteluperiaatteita ja arviointimenetelmiä. Tämän jälkeen selvennän tutkimukseni kannalta tärkeää käsitettä näkövammaisuus. Niin ikään kirjallisuustutkimuksen perusteella selvitän, miten näkövammaiset ihmiset tulee ottaa huomioon puhekäyttöliittymien suunnittelussa ja toteutuksessa, eli mitä erityistarpeita näkövammaisilla on puhekäyttöliittymien käyttäjinä. Erityistä huomiota kiinnitän sellaisiin puhekäyttöliittymiin, jotka sisältävät sekä puhetta että visuaalisia elementtejä. Lopuksi esittelen ”Aistiboxi”-hankkeen, jonka puitteissa toteutetun alustavan esittelyprototyypin pohjalta pyrin antamaan ehdotuksen näkövammaiset ihmiset huomioon ottavasta käyttöliittymästä. Suunnittelun apuna käytän tutkielmani teoriaosuutta sekä hankkeen aikana toteutetussa kohderyhmän haastattelussa ilmenneitä seikkoja.

2 Nykyinen puheteknologia

Puheteknologian kehittyessä ohjelmistoala saa yhä enemmän paineita ja samalla myös mahdollisuuksia suunnitella uusia puhekäyttöliittymiä. Tämä suuntaus on todennäköisesti mieluinen erityisesti näkövammaisille ja mahdollisesti sellaisille erityisryhmille, joille tarkka fyysinen toiminta on hankalaa.

Puheteknologiasta voivat hyötyä sekä teknologian ja palveluiden tuottajat, että loppukäyttäjät. Teknologian ja palveluiden tuottajille pääasiallinen hyöty on se, että puheteknologia antaa heille mahdollisuuden tarjota laajemman valikoiman palveluita matalammilla kustannuksilla. Esimerkiksi puheteknologian käyttö puhelinvaihteessa työntekijöiden sijaan voi pienentää vaihteeseen soitettujen puheluiden hintaa jopa 90 prosenttia. Puhe voi myös mahdollistaa sellaisten palveluiden tarjoamisen, jotka erottavat palvelun tuottajan kilpailijoistaan parantaen myös asiakastyytyvää. [McTear, 2004]

Loppukäyttäjille pääasiallinen hyöty on käytön mukavuus: puheen avulla tietoon ja palveluihin voi päästä käsiksi ajasta ja paikasta riippumatta. Puhe on myös varteen otettava vaihtoehto navigointitavaksi pieniin kannettaviin laitteisiin, kuten Internet-yhteydellä varustettuihin matkapuhelimiin ja PDA-laitteisiin, sillä kynien ja tekstinäppäinten käyttö voi olla hankalaa tällaisilla pienillä laitteilla. Puheteknologiasta voi olla hyötyä loppukäyttäjälle myös siinä, että esimerkiksi puhelinvaihteeseen soittaessaan soittaja joutuu epätodennäköisemmin jonottamaan, jos rutiininomaiset kyselyt hoitaa automaatti. Samalla myös puhelun hinta laskee. [McTear, 2004]

Puhe on luonnollinen tapa kommunikoida, ja se on nopea tapa tiedon syöttämiseen. Toisaalta tiedon syöttäminen puheella voi olla nopeaa vain silloin, kun sovellus pystyy käsittelemään nopeankin puheen ilman virheitä. Puheteknologian käyttöön liittyykin lukuisia haasteita, joihin McTearin [2004] mukaan kuuluvat muun muassa puheteknologian tämänhetkiset puutteet ja käyttäjien epärealistiset odotukset. Puheentunnistuksen virheet tulevat todennäköisesti aina olemaan haaste puhekäyttöliittymien suunnittelijoille. Tulostuksessa puolestaan voi olla ongelmia puhesynteesin virheiden takia, esimerkiksi ennalta tuntemattomien nimien ääntämisessä. Myös käyttäjien epärealistiset odotukset ovat suuri haaste, koska käyttäjät saattavat odottaa puhekäyttöliittymien toimivan yhtä edistyneesti kuin elokuvista ja tieteissarjoista tutut puhekäyttöliittymät.

Puheen yksi huono puoli käyttöliittymissä on sen hitaus tulostuksessa: ihmiset osaavat yleensä lukea nopeammin kuin kuunnella. Toisaalta esimerkiksi näkövammaisten kannalta puhe on oletusarvoisesti hyvä asia käyttöliittymässä, jos he eivät joka tapauksessa näe lukea.

2.1 Puheentunnistus

Puheentunnistuksen päätehtävä on käyttäjän puheen vastaanottaminen ja sen muuntaminen sanoiksi. Puheentunnistusjärjestelmän antamat vastaukset perustuvat aina sanastoon ja kielioppiin, jotka järjestelmään on määritelty. Tästä johtuen puheentunnistusjärjestelmän sanaston ja kieliopin suunnitteleminen on tärkeä tehtävä puhekäyttöliittymien suunnittelijoille. Puhekäyttöliittymiä suunnitellessa on myös hyvä tuntee puheentunnistusprosessi pääpiirteissään, jotta syntyy ymmärrys siitä, miksi puheentunnistus on vaikeaa, miksi tunnistusvirheet ovat väistämättömiä, ja miten sovelluksia voi suunnitella niin, että niitä käytettäessä virheitä esiintyy mahdollisimman vähän ja että tapahtuneet virheet saadaan käsiteltyä onnistuneesti [McTear, 2004]. Puheentunnistusprosessin tarkka tekninen kuvaus ei ole tutkielmani kannalta tarpeellista, joten selvitan asian vain keskeisiltä osin.

Jotta puheen käsittely tietokoneella olisi mahdollista, on puhe ensin muunnettava digitaalseksi. Tämä tapahtuu äänittämällä akustinen signaali eli käyttäjän puhe mikrofoniin tai puhelimen avulla. Tunnistettavasta puhesignaalista poistetaan aluksi epäoleelliset seikat, kuten taustamelu ja puhujan painotukset. Näin puhesignaali saadaan kuvattua keskeisten piirteidensä avulla käyrällä tai aikasarjalla: digitoitu puhesignaali muunnetaan sarjaksi tietokoneen ymmärtämiä piirteitä sopivan pituisissa jaksoissa, jotka ovat yleensä 10-20 millisekuntia pitkiä. [Zue et al., 1997]

Seuraava prosessin vaihe on akustisesta signaalista poimittujen piirteiden luokittelu ään-teiksi ja tämän jälkeen ään-teiden yhdistäminen sanoiksi. Tähän tehtävään tarvitaan kaksi erilaista mallia: akustinen malli esittää sitä, miten sanat koostuvat ään-teiden sarjoista, ja kielimalli puolestaan määrittää sallitut sanojen sarjat. Signaalista poimittuja piirteitä verrataan edellä mainittuihin malleihin, jotta löydettäisiin piirteitä todennäköisimmin vastaava sana tai lause. Laajimmin nykyisissä puheentunnistusjärjestelmissä käytetty akustinen malli perustuu piilotettuihin Markovin malleihin (Hidden Markov Models, HMM), jotka perustuvat tilastollisiin todennäköisyyksiin puheen yksiköiden peräkkäisyydestä. [McTear, 2004]

Kielimalleja on yleisessä käytössä kahdenlaisia. Yksinkertaisin kielimalli voidaan määritellä graafiksi, jossa on määritelty tarkasti sanat, jotka saavat seurata toisiaan - toisin sanoen kieliopiksi, jossa sallitut fraasit on ennalta määrätty [Zue et al., 1997; McTear, 2004]. Tämä tapa sopii lähinnä järjestelmiin, joissa on käytössä pieni sanasto, tai joissa käsitellään tarkasti määriteltyjä sanojen jaksoja, kuten päivämääriä tai osoitteita. Kielimallit voivat olla myös monimutkaisempia "N-gram"-malleja, jotka sisältävät tilastollista tietoa sanojen sarjoista. Ne kertovat, kuinka todennäköisiä eri sanat ovat annetuissa yhteyksissä.

Tällaisia malleja käytetään enimmäkseen suurten sanastojen sovelluksissa, joissa ei voida etukäteen määrittellä kaikkia sallittuja lauseita ja sanojen yhdistelmiä. [McTear, 2004]

Kun puheentunnistus on tehty, voidaan tunnistuksen tulos vielä analysoida kieltä ymmärtävän komponentin avulla. Kieltä ymmärtävän komponentin tarkoitus on antaa tunnistuksen tulokselle merkitys, jota dialogimanageri eli keskustelua ohjaava komponentti voi käyttää. Jos kyseessä on yksinkertainen sovellus, jossa on pieni sanasto ja yksinkertainen kieli-malli, ei kieltä ymmärtävää komponenttia tarvita. [McTear, 2004]

Kielen ymmärtäminen puheeseen perustuvissa järjestelmissä on McTearin [2004] mukaan ongelmallista kahdesta syystä: virheellisten syötteiden ja luonnollisen kielen moniselitteisyyden takia. Virheellisiä syötteitä tulee järjestelmälle puhuttaessa helposti, sillä puhe-kieli eroaa usein kirjoitetusta kielestä, ja puheentunnistusjärjestelmät on yleensä suunniteltu ymmärtämään kieliopillisesti virheetöntä kieltä. Luonnollinen kieli voi myös olla moniselitteistä monella tavalla. Jotkut sanat voivat toimia eri lauseenosina, esimerkiksi suomen kielen sana ”kuusi” voi olla sekä substantiivi että adjektiivi. Tällainen moniselitteisyys selviää tosin yleensä lauseen muiden sanojen muodostaman kontekstin avulla. Sanalla voi olla myös toisella tavalla eri merkityksiä, esimerkiksi englannin kielen sana ”bank” voi tarkoittaa sekä pankkia että joen reunustaa. Edellä mainittujen moniselitteisyyden muotojen lisäksi myös lauseen fraasiin suhde voi olla moniselitteinen. Suomen kielessä tämä ei ole taiputusäänteiden takia yhtä yleistä kuin englannin kielessä. Suomessakin kuitenkin esimerkiksi lause ”ostaisin kolme pizzaa ja jäätelöä”, voidaan ymmärtää joko niin, että halutaan ostaa 3 pizzaa ja 3 jäätelöä, tai niin, että ostoslistalla on 3 pizzaa ja jokin epämääräinen määrä jäätelöä.

Jotkut puheentunnistusjärjestelmät vaativat puhujan ”sisäänkirjoittautumista”. Tällaisten puhujasta riippuvaisten järjestelmien käyttäjien täytyy antaa näytteitä puheestaan ennen järjestelmän käyttämistä, ja näytteiden perusteella järjestelmä ”koulutetaan” eri käyttäjien puheeseen [Schmandt, 1994]. Puhujasta riippumattomat järjestelmät puolestaan eivät tarvitse tällaisia etukäteisnäytteitä, ja ne on yleensä suunnattu suurelle joukolle erilaisia käyttäjiä [Weinschenk, 2000; McTear, 2004]. Puhujasta riippumattomia järjestelmiä voidaan kouluttaa etukäteen useilta erilaisilta puhujilta kerättyjen puhenäytteiden avulla. Puhenäytteiden antajien tulisi vastata puhetaivaltaan järjestelmän mahdollisia käyttäjiä [McTear, 2004]. Weinschenkin [2000] mukaan puhujasta riippumattomat järjestelmät eivät ole luotettavia tehtäviin, jotka vaativat suuren sanaston, vaan niitä käytetään yleensä erikoistuneissa sovelluksissa, joilla on yksi, tarkkaan määritelty tehtävä.

Puheentunnistusjärjestelmiä on erilaisia myös sallitun puhemuodon sekä puhetyylin suhteen. Puheentunnistusteknologian ollessa nykyistä kehittymättömämpää monet puheen-

tunnistusjärjestelmät vaativat, että puhuja pitää pienen tauon sanojen välillä. Useimmat nykyaikaiset järjestelmät pystyvät käsittelemään sujuvaa, jatkuvaa puhetta, joten niitä käyttäessään käyttäjä voi puhua luonnollisemmin. Suurin ongelma jatkuvaa puhetta ymmärtävissä järjestelmissä on sanavälien virheetön erottaminen puheesta [McTear, 2004]. Sanavälien erottaminen on hankalaa etenkin sen vuoksi, että tauot yksittäisen sanan tavujen välillä voivat usein olla pidempiä kuin tauot sanojen välissä [Schmandt, 1994].

Sallitun puhetyylin perusteella puheentunnistusjärjestelmät voidaan puolestaan jakaa sellaisiin järjestelmiin, jotka on suunnattu käsittelemään luettua puhetta ja järjestelmiin, jotka on suunnattu käsittelemään spontaania puhetta. Luettu puhe on luonnollisesti sujuvampaa ja virheettömämpää kuin spontaani puhe. Puhuttuun dialogiin perustuvan sovelluksen käyttöön parempi vaihtoehto on spontaania puhetta tukeva puheentunnistusjärjestelmä. [McTear, 2004]

Eräs puheentunnistusjärjestelmien mahdollinen ominaisuus on päällepuhuminen (Barge-in). Tämä ominaisuus tarkoittaa sitä, että käyttäjä voi keskeyttää järjestelmän puheen, ja on hyödyllinen etenkin kokeneille käyttäjille, jotka eivät halua kuunnella kaikkia tuttuja tulosteita ja saavat päällepuhumisella nopeutettua dialogin kulkua. Päällepuhumisominaisuudessa on kuitenkin joitakin huonoja puolia, jotka tulee ottaa huomioon suunnitteluvaiheessa. Yksi huomioon otettava seikka on se, että ominaisuuden ollessa käytössä voi prosessori joutua kovalle koetukselle, sillä puheentunnistus joutuu olemaan koko ajan aktiivisena, eikä voi levätä missään välissä. Tämä puolestaan voi hidastaa sovelluksen toimintaa. Toinen mahdollinen ongelma on se, että jos käyttäjä esimerkiksi sattuu puhumaan toiselle henkilölle, tai kuuluu joku muu ylimääräinen ääni, keskeytyy sovelluksen tulostus ja käyttäjä voi hämmentyä. Lisäksi käyttäjä saattaa päällepuhuessaan käyttää kovempaa ääntä kuin normaalisti, tai toistaa päällepuhnutun asian sitten, kun sovellus on lopettanut puheenvuoronsa. Nämä seikat saattavat vaikeuttaa puheentunnistusta. [McTear, 2004]

Päällepuhuminen voidaan toteuttaa ainakin kahdella eri tavalla. Yksi mahdollisuus on se, että päällepuhutusta puheesta tehdään puheentunnistus ja tulostus lopetetaan vasta, kun käyttäjä on sanonut kokonaisen sanan tai fraasin. Keskeytys voi tapahtua myös niin, että käyttäjän alkaessa päällepuhumisen lopetetaan tulostus heti, ja käyttäjälle annetaan mahdollisuus toistaa sanomansa. [McTear, 2004]

Puheentunnistustehtävän vaikeusastetta voidaan mitata kompleksisuusluvulla (perplexity). Tarkemmin tämä tarkoittaa sen mittaamista, miten monia mahdollisia sanojen yhdistelmiä käyttäjä voi sanoa. Jos kompleksisuus on korkea, voi käyttäjä yhdistellä sanoja melko vapaasti, jolloin puheentunnistuksen tehtävä vaikeutuu. [Turunen, 2004]

Yksi puheentunnistusjärjestelmien tärkeimmistä ominaisuuksista on sanavaraston laajuus, joka vaihtelee järjestelmien välillä tarkkaan kontrolloitujen järjestelmien muutamasta sanasta sanelusovellusten yli 200 000 sanaan [McTear, 2004]. Puheentunnistuksessa tapahtuneiden virheiden määrä kasvaa luonnollisesti sanavaraston kasvaessa: sitä mukaa kun sanaston koko kasvaa, kasvaa myös samankaltaisten sanojen määrä. Tämä puolestaan luonnollisesti lisää virheellisten tunnistusten määrää. Toisaalta sanaston ollessa pieni kasvaa helposti myös sanaston ulkopuolisista sanoista johtuvien virheiden määrä.

Virhetunnistuksia voi olla kahdenlaisia: voi tapahtua niin, että jos käyttäjän sanoma sana ei sisälly sanastoon, se tunnistetaan joksikin sanastossa olevaksi samankuuloiseksi sanaksi. Toisentyypinen virhe on, jos puheentunnistus tekee virheen ja palauttaa vastaukseksi väärän sanan, vaikka oikea sana olisi ollut sanastossa. Jälkimmäisiä virheitä tapahtuu pääosin puheen vaihtelevuudesta johtuen. Puheentunnistuksen ongelmista suuri osa aiheutuukin puhesignaaliin liittyvistä vaihtelevuustekijöistä. Ensiksikin se, miltä äänteet kuulostavat, riippuu paljon siitä, missä yhteydessä ne esiintyvät. Esimerkiksi englanninkielessä äänne ”t” kuulostaa erilaiselta sanoissa ”toy”, ”stir”, ”strong” ja ”bit”. Tämän lisäksi esiintyy käyttäjäkohtaista vaihtelua johtuen puhujan fyysisen tai emotionaalisen tilan muutoksista, puheen nopeuden vaihtelusta, tai puhujan äänen laadun vaihtelusta. Edellä mainittujen tekijöiden lisäksi myös käyttäjien välillä on vaihtelua esimerkiksi eri murteiden vaikutuksesta. Myös äänen siirtokanavan – mikrofonin tai puhelimen – vaihtelut sekä ympäristön muutokset, esimerkiksi taustahäly, ovat puheentunnistusta vaikeuttavia tekijöitä. [Zue et al., 1997; McTear, 2004]

2.2 Puhesynteesi

Puhesynteesissä on kyse tekstin muuttamisesta synteettiseksi puhesignaaliksi. Weinschenkin [2000] mukaan puhesynteesi voidaan toteuttaa kahdella eri tavalla: yhdistelevällä synteesillä (concatenated synthesis), tai formanttisynteesillä (formant synthesis). D'Alessandro ja Liénard [1997] esittävät näiden lisäksi kolmannen mahdollisen puhesynteesitekniikan, artikulatorisen synteesin (articulatory synthesis), jossa puhetta muodostetaan mallintamalla ihmisen puheentuottamisen fysiologiaa. Weinschenkin mukaan [2000] nykyään yleisimmin käytetty tapa on yhdistelevä synteesi, joka käyttää puheen tuottamiseen äänitetystä ihmisen puheesta erotettuja palasia. Palaset voivat olla joko kahden äänten yhdistelmiä tai puolikkaita tavuja. Periaatteessa voitaisiin käyttää kokonaisia sanoja, tavuja, tai ääniteitä. Ne ovat kuitenkin epäkäytännöllisiä luonnollisen kuuloksen puheen rakennuspalikoiksi, sillä niiden ääntäminen voi muuttua useampia yksiköitä yhdistettäessä. Yhdistelevällä synteesillä tuotettu puhe kuulostaakin luonnollisemmalta

kuin formanttisynteesillä tuotettu puhe, jossa puheääni on kokonaan keinotekoisesti tuotettu.

Puhesynteesiin liittyy kuusi vaihetta: tekstin syöttö, normalisointi, tekstin muuttaminen foneemeiksi, poikkeuksien käsittely, prosodinen analyysi, ja puheen tulostus. Tekstisyöte voidaan saada käsittelyyn esimerkiksi käyttäjän syöttämänä, tietokannasta, tai tiedostosta. Aluksi saatu tekstisyöte normalisoidaan, eli lyhenteet muutetaan kokonaisiksi sanoiksi. Tämän jälkeen sanojen joukot muutetaan foneemeiksi, eli ne muutetaan sanakirjan ja muun muassa asiayhteyden analysoinnin avulla siihen muotoon, miten ne lausutaan. Poikkeuksien käsittely joudutaan käymään läpi esimerkiksi silloin, jos lausumismuotojen etsimiseen käytetty sanakirja ei anna minkäänlaista tulosta. Tällöin päätökset ääntämisestä tehdään käyttämällä avuksi kirjainten ääntämiseen liittyviä sääntöjä. Lopuksi prosodinen tieto, kuten intonaatio ja kesto, lasketaan ja sisällytetään lopulliseen ääntämisestä koskevaan määrittelyyn, jonka jälkeen puhe on valmiina tulostettavaksi. [Weinschenk, 2000]

Tekijät, jotka määräävät syntetisoidun puheen laadun ja samalla menestyksen, ovat ymmärrettävyys ja miellyttävyys. Yksi tärkeä ymmärrettävyyteen ja miellyttävyysvaikutusta vaikuttava tekijä on prosodian käyttö. Tärkeimmät puheen prosodiset piirteet ovat puheäänien korkeus, voimakkuus, nopeus, sekä tauot puheessa [Hakulinen et al., 1999]. Näiden piirteiden vaihtelu tuo puheeseen elävyyttä, ja välittää myös tietoa, jota ei voi saada millään muulla tavalla. Esimerkiksi eri sanojen painottaminen lauseessa voi muuttaa lauseen merkitystä olennaisesti. Myös monitahoisten elementtien, kuten listojen tai osoitteiden, esittämisessä prosodian käyttö on olennaista, sillä tavallisessakin ihmisten välisessä kommunikaatiossa niitä voi olla vaikeaa ymmärtää. Hyvin toteutettu taukojen käyttö puheessa voi Hakulisen et al. [1999] mukaan parantaa puheen ymmärrettävyyttä huomattavasti. Heidän tutkimuksensa mukaan äänen korkeus ja voimakkuus puolestaan vaikuttavat enemmän puheen luonnollisuuteen, eivät niinkään ymmärrettävyyteen.

Prosodinen vaihtelu syntetisoituun puheeseen voidaan toteuttaa SSML-ohjauskielen avulla. SSML tarjoaa suunnittelijoille standardoidun tavan kontrolloida puhetulosteiden ominaisuuksia, ja käytännössä se tarkoittaa sitä, että syntetisoitavan tekstin joukkoon lisätään SSML-kielen elementtejä kuvaamaan haluttua prosodista rakennetta. [W3C SSML, 2006]. Hakulisen et al. [1999] mukaan koodin käyttäminen ei kuitenkaan kaikissa tapauksissa ole mahdollista. Lisäksi ollakseen luonnollista ja tehokasta puhe ei saisi olla staattista, vaan sen pitäisi olla asiayhteyden ja dialogihistorian perusteella dynaamisesti rakentunutta. He esittävätkin toisen mahdollisen tavan prosodisen tiedon lisäämiseen puhetulosteisiin: sovellukseen toteutetaan useita erilaisia ”henkilöiksi” kutsuttuja moduuleita, jotka tuottavat kaikki tulosteet. Jokaisella ”henkilöllä” on omat tyypilliset luonteenpiirteensä ja käyttäytymismallinsa, ja käyttäjä voi valita sellaisen henkilön, jonka

ominaisuuksista hän pitää. Koska nämä ”henkilöt” käyttävät hyväkseen tietoa keskustelun historiasta ja tietävät asiayhteyden, ne pystyvät tuottamaan prosodisesti rikkaita, ja samalla siis laadukkaita, tulosteita. [Hakulinen et al., 1999]

3 Puhekäyttöliittymien suunnitteluperiaatteet

Käyttöliittymien suunnittelussa yksi keskeinen päämäärä on hyvän käytettävyyden saavuttaminen. Käytettävyyteen vaikuttavat monet asiat. Constantinen ja Lockwoodin [2000] mukaan korkean käytettävyyden sovelluksia on helppo oppia käyttämään, ja niiden käyttäminen myös tuntuu käyttäjistä hyvältä. Käytettävyyteen sisältyy myös helppous muistaa käyttökerrasta toiseen, miten sovellusta käytetään. Lisäksi käytettävä sovellus auttaa käyttäjää tehokkaassa käytössä, jonka aikana syntyy vain vähän virheitä.

Puhekäyttöliittymät voidaan jakaa erilaisiin ryhmiin sen perusteella, kuinka merkittävä rooli puheella on käyttöliittymässä. Weinschenk [2000] jakaa puhekäyttöliittymät puheeseen perustuviin käyttöliittymiin ja käyttöliittymiin, joissa puhe on osa käyttöliittymää muiden, usein visuaalisten, elementtien rinnalla. Visuaaliset elementit voivat tarkoittaa esimerkiksi graafista käyttöliittymäikkunaa tai WWW-sivua.

Aistiboxi-hankkeen puitteissa suunnitellussa käyttöliittymässä puhetta käytetään lähinnä tulosteisiin. Luon kuitenkin perusteellisen katsauksen myös puheeseen perustuviin käyttöliittymiin, jotta tutkielmani perusteella muodostuisi kokonaiskuva puhekäyttöliittymistä. Lisäksi myös Aistiboxi-hankkeen kannalta on hyödyllistä kartoittaa, mitä kaikkea tarvitsee ottaa huomioon, jos hankkeessa aletaan jossain tulevassa vaiheessa harkita myös suuremmassa määrin puheeseen perustuvan käyttöliittymän tai käyttöliittymän osan toteutusta.

Puheentunnistuksen käyttäminen käyttöliittymässä on monella tavalla monimutkaisempi asia kuin puheen käyttäminen pelkkään tulostukseen. Kun käyttäjän syöteenä käytetään puhetta, täytyy ottaa huomioon huomattava tunnistusvirheiden mahdollisuus ja niihin reagoiminen, mikä puolestaan monimutkaistaa käyttöliittymän suunnittelua. Jos taas syöte tapahtuu esimerkiksi näppäimistön tai kaukosäätimen kautta, eivät virhemahdollisuudet ole yhtä moninaisia, ja käyttöliittymän suunnittelu on tämänkin vuoksi yksinkertaisempaa.

Tässä luvussa käsittelen puhekäyttöliittymien suunnitteluperiaatteita liittyen dialogin hallintaan, käyttäjän ohjaamiseen puhetulosteiden avulla, sekä virheiden käsittelyyn. Nämä aihepiirit ovat puhekäyttöliittymien suunnittelussa keskeisiä. Niiden lisäksi käsittelen myös puhuvien valikoiden suunnittelua sekä puhekäyttöliittymien suunnittelua TV-ympäristöön, sillä nämä aihepiirit ovat olennaisia erityisesti Aistiboxi-hankkeen kannalta.

3.1 Dialoginhallinta

Sovelluksissa, joissa puhetta käytetään sekä tulostukseen että tiedon syöttämiseen, tarvitaan koko keskustelun ajan dialoginhallintaa: sovelluksen tulee olla koko ajan perillä siitä, mitä keskustelussa tapahtuu ja mitä siinä voi seuraavaksi tapahtua. Puhesovellukset voidaan jakaa erilaisiin ryhmiin muun muassa keskustelun kontrollointityylin sekä keskustelun aloitteen tekijän mukaan.

3.1.1 Keskustelun kontrolli

Puhesovellukset voidaan jakaa erilaisiin ryhmiin sen perusteella, millä keinoin keskustelua kontrolloidaan. McTear [2004] jakaa puhesovellukset dialogin kontrollin suhteen tilapohjaisiin (finite-state-based), lomakepohjaisiin (frame-based), ja agenttipohjaisiin (agent-based) järjestelmiin. Nykyään käytetyin tapa dialogin kontrolloimiseen on tilapohjainen lähestymistapa. Rosenfeld et al. [2001] puolestaan jakavat puhesovellukset luonnollista kieltä käyttäviin järjestelmiin, keskustelupuujärjestelmiin (dialog-tree systems), ja komentopohjaisiin (command-and-control) järjestelmiin. Näistä luonnollista kieltä käyttävät järjestelmät pitävät sisällään edellä mainitut McTearin [2004] määrittelemät agenttipohjaiset järjestelmät sekä lomakepohjaiset järjestelmät. Keskustelupuujärjestelmät puolestaan ovat monilta osin tilapohjaisten järjestelmien kaltaisia, mutta nämä järjestelmätyypit eivät kuitenkaan täysin vastaa toisiaan.

Komentopohjaisissa järjestelmissä on tiukka kielioppi, joka määrittelee kaikki mahdolliset syötteet. Tällainen tapa sopii hyvin rajattujen alojen sovelluksiin. Koska käytetty kieli on ennustettavissa, vähentyvät puheentunnistusongelmat. Toisaalta, jotta tällaiseen tilanteeseen päästään, täytyy käyttäjän käyttää aikaa järjestelmän kielen opettelemiseen, mikä onkin komentopohjaisten järjestelmien huono puoli. Toinen heikko kohta on se, että komentopohjaisten järjestelmien rakenne ei itsessään ohjaa käyttäjää oikeanlaisiin syötteisiin, vaan käyttäjän täytyy itse opetella miten sovellusta käytetään. [Rosenfeld et al., 2001]

Tilapohjaisten sovellusten dialogin kulku voidaan määritellä etukäteen tilakaavion avulla, jossa solmut esittävät järjestelmän puheenvuoroja, ja solmujen väliset siirtymät kuvaavat kaikkia mahdollisia polkuja, joita keskustelu voi edetä. McTearin [2004] mukaan tilapohjaisen mallin tärkeä etu on sen yksinkertaisuus: tilakaaviot ovat luonnollinen tapa mallintaa järjestelmän kontrolloimia, ennalta määritellyn kaavan mukaan eteneviä keskusteluja. Lisäksi esimerkiksi puheentunnistukselta ei vaadita paljon, kun käyttäjän syötteitä voidaan rajoittaa ja ennustaa. On myös olemassa tilapohjaista mallia hyödyntäviä työkaluja, kuten esimerkiksi CSLU toolkit kehitysympäristö. Tällaisilla työkaluilla

käyttöliittymän suunnittelu on miellyttävää, sillä dialogin kulku voidaan esittää suunnittelijalle intuitiivisella tavalla graafisesti.

Myös keskustelupuujärjestelmät helpottavat puheentunnistusta jakamalla dialogin pisteisiin, joissa jokaisessa käyttäjä vastaa yhteen järjestelmän esittämään kysymykseen. Rosenfeld et al. [2001] mukaan tällaisten sovellusten huono puoli on käyttäjän näkökulmasta se, että käyttäjällä ei ole suoraa pääsyä haluamiinsa osiin sovellusta, vaan koko keskustelupuu täytyy käydä läpi siinä järjestyksessä kuin suunnittelija on määritellyt. Suunnittelijan näkökulmasta keskustelupuujärjestelmät voivat puolestaan olla vaikeita toteuttaa, jos keskustelussa on paljon erilaisia etenemismahdollisuuksia. Sama ongelma voi esiintyä ylläpidossa, sillä uusia toimintoja lisätessä joudutaan mahdollisesti muokkaamaan kokonaista keskustelupuuta. Toisaalta keskustelupuu yksinkertaistaa vuorovaikutusta ja vähentää käyttäjien kouluttamisen tarvetta.

Lomakepohjaiset sovellukset ovat tilapohjaisten sovellusten tavoin sopivia tehtäviin, joissa järjestelmä kerää käyttäjältä tietoja kysymällä kysymyksiä. Lomakepohjaiset sovellukset eroavat tilapohjaisista sovelluksista siinä, että niissä kysymysten järjestys ei ole ennalta määrätty ja käyttäjä voi vastata useampaan kysymykseen kerralla. Kun käyttäjä on antanut syötteensä, sovellus tarkistaa, mitkä lomakkeen kohdat ovat täyttämättä, ja päättää seuraavan kysymyksen tämän perusteella. Tällaisissa sovelluksissa keskustelun aloite vaihtelee käyttäjän ja sovelluksen välillä. [McTear, 2004]. Keskustelun aloitteeseen liittyvää asiaa on käsitelty enemmän kohdassa 3.1.2.

Käyttäjien kannalta lomakepohjaisten sovellusten etu tilapohjaisiin verrattuna on niiden joustavuus ja mahdollisuus käyttää luonnollista kieltä vuorovaikutuksessa. Sovelluksen kehittäjän näkökulmasta lomakepohjaiset sovellukset ovat monimutkaisia toteuttaa. [McTear, 2004]

Agenttipohjaisissa järjestelmissä käyttäjät ovat vuorovaikutuksessa ohjelmistoagenttien kanssa. Tällaiset järjestelmät perustuvat tekoälyn käyttämiseen ja ovat sopivia monimutkaisiin tehtäviin, kuten neuvotteluun ja ongelmanratkaisuun. Dialogi etenee dynaamisesti niin, että puheenvuorot rakentuvat edellisten puheenvuorojen pohjalta. Yleensä agenttipohjaisissa järjestelmissä keskustelun aloite vaihtelee käyttäjän ja sovelluksen välillä. Tämän takia käyttäjän syötettä ei voida etukäteen ennustaa, ja monimutkaisimmissa sovelluksissa tarvitaan luonnollista kieltä ymmärtävää komponenttia. [McTear, 2004]

Rosenfeld et al. [2001] pitävät luonnollisen kielen käyttämistä sovelluksissa äärimmäisen haastavana tehtävänä. Käyttäjille tällaiset sovellukset ovat siinä mielessä helppoja käyttää, että niiden käyttäminen ei vaadi erityistä opettelua. Keskustelupuujärjestelmissä sekä

komentopohjaisissa järjestelmissä voidaan parantaa järjestelmän tehokkuutta rajoittamalla sitä, mitä käyttäjä voi sanoa.

3.1.2 Keskustelun aloite

Puhesovelluksissa keskustelun aloite ei yleensä vaihtelee samalla tavalla kuin ihmisten välisessä keskustelussa. Puhesovellukset voidaankin jakaa aloitteen tekijän mukaan: aloite voi olla joko järjestelmällä, käyttäjällä, tai molemmilla. [McTear, 2004]

Kun aloite on järjestelmällä, on käyttäjän syöte yleensä rajoitettu yksittäisiin sanoihin tai fraaseihin, ja järjestelmän tulosteet on tarkkaan suunniteltu. Tämä tapa sopii hyvin yksinkertaisiin keskusteluihin, sillä jokaiselle käyttäjän vastaukselle voidaan määrittää sanasto ja kielioppi etukäteen. Toisaalta väärin tulkintojen korjaus on vaikeampaa, kun aloite on järjestelmällä. [McTear, 2004]

Kun aloite on käyttäjällä, käyttäjä voi kysyä yhden tai useampia kysymyksiä, jonka jälkeen järjestelmä tulkitsee kysymykset ja vastaa niihin. Käyttäjä voi myös antaa komentoja, joiden mukaan järjestelmä sitten toimii. Käyttäjä ikään kuin tekee luonnollisen kielen kyselyitä tietokantaan. Joissain sovelluksissa järjestelmä voi esittää selventäviä kysymyksiä, jos käyttäjän syöte on ollut jollain tapaa epäselvä. Yksi tällaisten käyttäjävetoisten keskustelujärjestelmien huono puoli on se, että järjestelmän täytyy pystyä kattavasti käsittelemään monenlaisia syötteitä. Myös käyttäjän täytyy olla tietoinen sanoista ja fraaseista, jotka järjestelmä ymmärtää. [McTear, 2004]

Jos aloite vaihtelee käyttäjän ja järjestelmän välillä, kumpikin osapuoli voi esimerkiksi tehdä kysymyksiä. Tämän tyyppisen aloitteen vaihtelun käyttö nykyisissä järjestelmissä on kuitenkin harvinaista. Aloitteen vaihtelua on käytetty lähinnä rajoitetummalla tavalla, jossa järjestelmä on päävastuussa keskustelusta, mutta käyttäjällä on mahdollisuus antaa järjestelmälle enemmän tietoa, kuin mitä on kysytty. Eli käyttäjä voi vastata useampaan kysymykseen kerralla, tai vastata kysymyksiin itse haluamassaan järjestyksessä. Tämän tyyppinen keskustelu toimii hyvin, jos järjestelmä pysyy selvillä siitä, mihin kysymyksiin on milläkin hetkellä vastattu. [McTear, 2004]

3.2 Käyttäjän ohjaaminen puhetulosteiden avulla

Olenainen asia käyttäjän ja sovelluksen vuorovaikutuksessa on palaute: käyttäjän tulee saada toimistaan palautetta välittömästi. Vaikka sovelluksella kestäisikin tehtävän suorittaminen jonkin aikaa, tulee palautteen antamisen olla ensimmäisenä tehtävällä [Schmandt, 1994].

Käyttäjän ja sovelluksen välisessä dialogissa on monissa tapauksissa tärkeää se, että sovellus ohjaa omilla tulosteillaan käyttäjää oikeanlaisiin syötteisiin, etenkin, jos sovelluksella on ongelmia tulkita käyttäjän syötteitä. Puheentunnistuksen monimutkaisuuden huomioon ottaen on myös tärkeää varmistua siitä, että käyttäjän syötteet tulevat tulkituiksi oikein. Kohdissa 3.2.1 ja 3.2.2 käsittelen erilaisia tapoja sovelluksen kehotusten muotoiluun sekä sitä, millaisin menetelmin sovellus voi varmistaa ymmärtäneensä käyttäjän syötteen oikein.

3.2.1 Kehotusten muotoilu

Puhekäyttöliittymän tulosteiden muotoiluun kehotuksien tarkkuuden suhteen on monenlaisia vaihtoehtoja. Voidaan käyttää esimerkiksi täsmällisiä kehotuksia (explicit prompts), epäsuoria kehotuksia (implicit prompts), lisääntyviä kehotuksia (incremental prompts), laajennettuja kehotuksia (expanded prompts), kaventamista (tapering), tai vihjeitä (hints) [Yankelovich, 1996].

Pienen sanaston sovelluksissa käytetään useimmiten täsmällisiä kehotuksia, jotka kertovat käyttäjälle tarkalleen sanat, jotka tämän kuuluu sanoa, tai vaihtoehdot, joista tämän kuuluu valita. Epäsuorat kehotukset puolestaan ovat vastakohta täsmällisille kehotuksille. Niitä käyttävät sovellukset eivät kerro käyttäjälle tarkkaan, mitä tämän kuuluu sanoa, vaan antavat käyttäjän vapaasti muotoilla vastauksensa. Tehokkaimmat puhetulosteiden muotoilun menetelmät yhdistelevät näitä kahta edellä mainittua menetelmää. [Yankelovich, 1996]

Lisääntyvien kehotusten menetelmässä järjestelmä antaa käyttäjälle ensin epäsuoran kehotuksen ja jää odottamaan käyttäjän syötettä. Jos käyttäjä ei anna minkäänlaista syötettä tai jos syöte on virheellinen, annetaan uusi, täsmällisempi kehotus [Yankelovich, 1996]. Weinschenkin [2000] mukaan tällainen kehittyvä tulostustapa on yksi keino vähentää virheitä ja lisätä käyttäjien tyytyväisyyttä.

Järjestelmä: Welcome to ABC Bank. What would you like to do?

Käyttäjä: (hiljaisuus)

Järjestelmä: You can check an account balance, transfer funds or pay a bill. What would you like to do?

Käyttäjä: (hiljaisuus)

Järjestelmä: Say one of the following choices: check balance, transfer funds or pay bills.

Esimerkki 1. Lisääntyvien kehotusten menetelmä. [Yankelovich, 1996]

Laajennetut kehotukset ovat variaatio lisääntyvistä kehotuksista. Niitä voidaan käyttää, kun käyttäjä vastaa epäsuoraan kehotukseen jotain, mitä järjestelmä ei osaa tulkita. Tällöin seuraava versio järjestelmän tulosteesta on edellistä tarkempi. [Yankelovich, 1996]

Järjestelmä: You have heard all messages, please give an instruction.

Käyttäjä: Mmmh, I don't know.

Järjestelmä: The accepted speech commands are "replay", "delete", "new announcement", and "turn off".

Esimerkki 2. Laajennettujen kehotusten menetelmä. [Yankelovich, 1996]

Lisääntyvät ja laajennetut kehotukset auttavat käyttäjiä niin, ettei vuorovaikutus hidastu käyttäjän osatessa antaa oikeanlaisen syötteen heti vuorovaikutuksen alussa. [Yankelovich, 1996]

Kaventamisen avulla pyritään lyhentämään vuorovaikutusta käyttäjien kanssa sitä mukaa, kun heidän kokemuksensa järjestelmän kanssa lisääntyy. Käyttäjälle annetaan aluksi täsmällinen kehotus, ja vuorovaikutuksen edetessä, esimerkiksi jos käyttäjä tekee samalla käyttökerralla toisen samanlaisen kyselyn, kehotusta kavennetaan epätarkemmaksi kehotukseksi. [Yankelovich, 1996; Weinschenk, 2000]

Järjestelmä: Please say the first and last name of the person you want.

Toisella kertaa sama kehotus olisi kavennettuna:

Järjestelmä: Say the name of the person you want.

Esimerkki 3. Kaventaminen. [Yankelovich, 1996]

Kaventamisen yksi muunnelma on vihjeiden käyttö: uusille käyttäjille annetaan ohjeita, jotka sisältävät epäsuoran kysymyksen lisäksi tarkan vihjeen. Täsmällisten kehotusten kattavan valintamahdollisuuslistan sijaan ne sisältävät vain joitakin hyödyllisimpiä tai tavallisimpia vaihtoehtoja. Ne ovat myös suunniteltuja poistettaviksi sitten, kun käyttäjä on useita kertoja reagoinut kehotuksiin antamalla oikeanlaisen syötteen. [Yankelovich, 1996]

3.2.2 Varmennusmenetelmät

Sovelluksen ja käyttäjän välisessä vuorovaikutuksessa on tärkeää, että sovellus varmistaa ymmärtäneensä käyttäjän syötteen oikein. Erityisen tarpeellista tällainen varmennus on silloin, jos käsitellään tärkeää ja yksityiskohtaista tietoa, kuten luottokorttien numeroita tai

osoitteita. Varmennukseen on olemassa kaksi yleisesti käytettyä menetelmää: täsmällinen varmennus ja epäsuora varmennus. [McTear, 2004]

Täsmällinen varmennus tarkoittaa sitä, että käyttäjältä kysytään varmennusta syötteelle täsmällisellä kysymyksellä. Täsmällinen varmennus voi olla käyttäjälle pitkäveteistä, kun jokainen asia varmistetaan erillisellä kysymyksellä, ja samalla dialogissa käytettävien puheenvuorojen määrä kasvaa. Varmennus voidaan tosin tehdä myös yhdellä kertaa siinä vaiheessa, kun kaikki tiedot on kerätty. Myös tällainen tapa saattaa pidentää dialogia, jos käyttäjä haluaa muuttaa jonkin varmennuksessa esitetyistä arvoista, kuten esimerkissä 4. [McTear, 2004]

Järjestelmä: Where are you travelling to?

Käyttäjä: London

Järjestelmä: What day?

Käyttäjä: Friday.

Järjestelmä: So you want to travel to London on Sunday?

Käyttäjä: No.

Järjestelmä: Which value do you wish to correct: the destination or the day?

Käyttäjä: Day

Järjestelmä: What day?

Esimerkki 4. Täsmällinen varmennus. [McTear, 2004]

Epäsuoran varmennuksen menetelmässä sovellus sisällyttää varmennuksen seuraavaan kysymykseensä, kuten esimerkissä 5, eikä varmentamiseen käytetä erillisiä puheenvuoroja. Käyttäjä voi korjata haluamansa arvon, mutta jos hän vastaa seuraavaan kysymykseen korjaamatta mitään, arvo on epäsuorasti varmennettu.

Käyttäjä: I want to fly from Belfast to London.

Järjestelmä: At what time do you want to fly from Belfast to London?

Käyttäjä: Seven in the evening.

Esimerkki 5. Epäsuora varmennus. [McTear, 2004]

Epäsuoran varmennuksen käyttäminen voi lyhentää keskustelun kestoa verrattuna täsmälliseen varmennukseen, sillä kysyttäviä kysymyksiä on vähemmän. Keskustelu voi sujua myös luonnollisemman tuntuisesti. Keskustelun sujuvuus riippuu kuitenkin myös siitä, kuinka hyvin puheentunnistus toimii: epäsuora varmennus vaatii paljon puheentunnistukselta, sillä käyttäjän erilaisten vastausvaihtoehtojen määrä on suuri. Korjauksenkin käyttäjä voi tehdä monella eri tavalla, erityisesti jos korjattavia arvoja on useampia. [McTear, 2004]

Epäsuoraa varmennusta käytettäessä luotetaan siihen, että käyttäjä korjaa virheet, jos niitä sovelluksen epäsuorassa kysymyksessä ilmenee. Tämä voi kuitenkin johtaa myös ongelmiin, sillä käyttäjä ei välttämättä ymmärrä tai huomaa, että korjaamista tarvitaan, ja että korjaus kuuluu tehdä heti eikä vasta myöhemmässä vaiheessa. [McTear, 2004]

Yksi mahdollinen varmennustapa on sellainen, että yhdistellään täsmällistä ja epäsuoraa varmennusta. Esimerkiksi aluksi voidaan käyttää epäsuoraa varmennusta, ja tämän jälkeen siirtyä täsmälliseen varmennukseen, ja mahdollisesti vielä rajoitetumpiin kysymyksiin, kuten esimerkissä 6. [McTear, 2004]

Käyttäjä: I want to fly from Belfast to Luton.

Järjestelmä: At what time do you want to fly from Belfast to London?

Käyttäjä: No, not London, Luton.

Järjestelmä: Do you want to fly to London? Please answer yes or no.

Käyttäjä: No.

Järjestelmä: Please spell your destination.

Käyttäjä: L-U-T-O-N.

Järjestelmä: So you want to fly to Luton?

Käyttäjä: Yes.

Järjestelmä: At what time do you want to fly from Belfast to Luton?

Esimerkki 6. Varmennusmenetelmien yhdisteleminen. [McTear, 2004]

3.3 Virheiden käsittely

Weinschenkin [2000] ohjeet virheiden käsittelyyn ja ennaltaehkäisyyn ovat tiivistetysti seuraavat:

- Käytä tarkkoja virheilmoituksia.
- Rajoita taustahälyä.
- Anna käyttäjälle mahdollisuus sammuttaa syötelaite (esim. mikrofoni).
- Tarjoa mahdollisuus toimintojen peruuttamiseen.
- Käytä äänimerkkejä virheen sattuessa.
- Käytä multimodaalisia vihjeitä virheiden sattuessa, jos mahdollista.
- Harkitse toistotoiminnon tarjoamista.
- Älä oletta, että käyttäjät kuulevat kaiken. Sijoita tärkeä tieto ensimmäiseksi tai viimeiseksi, jolloin se todennäköisimmin kuullaan ja muistetaan.

McTearin [2004] mukaan puhejärjestelmän dialogimanageri joutuu kohtaamaan sellaisia ongelmia, joita ei normaalisti tule vastaan graafisten käyttöliittymien ollessa kyseessä. Dialogimanageri saa puheentunnistukselta ja kielen ymmärtämiskomponentilta tulkinnan siitä, mitä käyttäjä on sanonut, ja päättää tämän perusteella, mitä järjestelmä tekee seuraavaksi. Järjestelmä ei kuitenkaan voi taata, että puheentunnistuksen ja kielen ymmärtämiskomponentin käsittelyn jälkeen saatu käyttäjän syöte on virheetön. Saattaa myös tapahtua niin, että järjestelmässä ei ole saatavilla tietoa, jota käyttäjä kyselyssään pyytää.

McTearin [2004] mukaan käyttäjältä saatu syöte voi olla epäselvä tai vajavainen monesta syystä. Puheentunnistus saattaa jättää kokonaan palauttamatta käyttäjän antaman syötteen, jos puhe on ollut liian hiljaista ja jäänyt sen takia rekisteröimättä. Voi käydä myös niin, että vain osa käyttäjän syötteestä tunnistetaan ja palautetaan sovellukselle. Näin voi käydä esimerkiksi sellaisessa tapauksessa, että käyttäjä alkaa puhua ennen kuin puheentunnistus on käynnistynyt, eli ennen äänimerkkiä, tai jos järjestelmä tulkitsee käyttäjän pitämän tauon puheenvuoron lopetuksiksi ja sen takia lopettaa puheentunnistuksen. Yksi mahdollinen virhetilanne on sellainen, että koko syöte tunnistetaan, mutta joko osa syötteestä tai koko syöte tunnistetaan virheellisesti. Voi käydä myös niin, että vaikka kaikki sanat tunnistetaan oikein, kielenymmärryskomponentti ei saa määriteltyä oikeaa merkitystä syötelle.

Turunen ja Hakulinen [2001] esittävät virheiden käsittelyn prosessina, johon sisältyy seuraavat seitsemän vaihetta: (1) virheen havaitseminen, (2) syyn selvittäminen, (3) virheenkorjausmenetelmän valitseminen, (4) korjauksen suorittaminen, (5) virheestä tiedottaminen käyttäjälle, (6) virheen korjaamisen lopettaminen ja palaaminen varsinaiseen dialogiin, ja (7) virheiden ennaltaehkäisy.

Virheen voi havaita joko käyttäjä tai järjestelmä. Yksinkertaisimmissa tapauksissa järjestelmä havaitsee virheen sen perusteella, että puheentunnistuksen tunnistustulokselle annama luottamusluku (confidence score) alittaa asetetun raja-arvon. Tällöin puheentunnistuskomponentti on kovin epävarma siitä, onko saatu tunnistustulos oikea. [Turunen and Hakulinen, 2001]

Virheen syyn selvittämisessä olennaista on se, että syyn ollessa selvillä järjestelmä voi paremmin korjata virheen ja myös ennaltaehkäistä virheitä jatkossa. Syyn selvittämisprosessissa voidaan käyttää hyväksi eri lähteistä, kuten käyttäjämallista tai dialogihistoriasta, saatua tietoa. [Turunen and Hakulinen, 2001]

Virheen korjaaminen tarkoittaa pohjimmiltaan dialogin suunnan muuttamista. Kun virheen syy on selvillä, kartoitetaan virheen tyyppi, dialogin tila, sekä tiedot edellisistä virheistä.

Näiden tietojen perusteella voidaan valita sopiva virheenkorjausmenetelmä. Virheiden korjaamisessa vaarana on se, että järjestelmä joutuu virheenkorjauskierteeseen, jossa yhden virheen korjaaminen johtaa toiseen virheeseen. Tämän ongelman välttämiseksi virheenkorjausmenetelmää valitessa onkin tärkeää tarkistaa, miten virheitä on aikaisemmin korjattu kyseisessä dialogissa, ja minkälaisin tuloksin. [Turunen and Hakulinen, 2001]

Korjauksen suorittaminen on yleensä itsessään pieni dialogi, mikä voi tarkoittaa esimerkiksi yksinkertaista varmistusta käyttäjältä [Turunen and Hakulinen, 2001]. Käyttäjältä voidaan esimerkiksi kysyä kysymys, johon tämä voi vastata kyllä tai ei. Käyttäjälle voidaan myös esimerkiksi esittää lyhyt lista, josta tämän tulee valita sopiva vaihtoehto. McTearin [2004] mukaan yksinkertaisin tapa käsitellä käyttäjän antamaa vääränlaista tai vajaata syötettä on asian ilmoittaminen käyttäjälle ja käyttäjän kehottaminen muotoilemaan syöte uudelleen. Tällaisen lähestymistavan ongelma on siinä, että se ei millään tavalla erottele eri tapoja, joilla syöte voi olla virheellinen. Eri virhetyyppien erottelu olisi tärkeää, jotta käyttäjä tietäisi paremmin, millä tavalla syöte tulee muotoilla uudelleen. Esimerkiksi, jos syötteenä on saatu hiljaisuutta, järjestelmä voi antaa esimerkin 7 mukaisen kehotuksen. Jos taas käyttäjän syötettä ei saada tulkittua, voi järjestelmä antaa esimerkin 8 mukaisen kehotuksen.

Järjestelmä: Sorry I did not hear anything. Would you please speak a little louder.

Esimerkki 7. Järjestelmän mahdollinen tuloste, kun käyttäjä on puhunut liian hiljaa. [McTear, 2004]

Järjestelmä: Sorry I didn't quite understand that. Would you please rephrase?

Esimerkki 8. Järjestelmän mahdollinen tuloste, kun käyttäjän tulostetta ei osata tulkita. [McTear, 2004]

Kun käyttäjälle annetaan tieto virheestä, voidaan myös kertoa, mitkä seikat johtivat virheeseen. Käyttäjän on myös tärkeää saada tietää, mitä tapahtuu seuraavaksi, kun virheen korjaaminen lopetetaan ja palataan takaisin dialogiin.

Virheiden ennaltaehkäisyssä voidaan käyttää hyödyksi tietoa aikaisemmista virhetilanteista. Joissain tapauksissa järjestelmä voi mukauttaa toimintaansa paremmin sopivaksi käyttäjän käyttötapoihin niin, että jatkossa mahdollisten virheiden tapahtuminen pystytään estämään. Esimerkiksi jos tietyn käyttäjän kohdalla on havaittu paljon virhetilanteita, voidaan kyseisen käyttäjän ollessa kyseessä muotoilla dialogia yksinkertaisemmaksi. Tämä ei välttämättä ole kovin tehokasta, mutta monimutkaisessa tilanteessa se voi olla ainoa keino dialogin onnistuneeseen läpivientiin. [Turunen and Hakulinen, 2001]

3.4 Puhuvat valikot

Puhuvissa valikoissa voidaan käyttää joko äänitettyä ihmisääntä tai syntetisoitua ääntä, joissa molemmissa on omat hyvät ja huonot puolensa. Nykyään yleisemmin käytetty puhuvien valikoiden toteutustapa on äänitetyn ihmisäänen käyttäminen [Schmidt and Wlodkowski, 2003]. Eri tavoilla syntetisoitua puhetta on käsitelty tarkemmin kohdassa 2.2. On kuitenkin huomioitava, että kun kyse on puhuvista valikoista, puhetta tarvitsee tuottaa vain ennalta määrättyjen valikkokomentojen ja mahdollisten ohjeiden verran. Tämä on monelta osin eri asia kuin kehotusten muodostaminen käyttäjän kanssa käytävään dynaamiseen dialogiin. Tämän takia myös ihmisäänen ja syntetisoidun puheen välinen vertailu tulee tehdä erilaiselta pohjalta, erityisesti puhuvien valikoiden ominaisuudet huomioidaan ottaen.

Schmidtin ja Wlodkowskin [2003] mukaan äänitetyn ihmisäänen käyttäminen puhuvissa valikoissa tapahtuu yksinkertaisesti niin, että tietyn ihmisen puhumana äänitetään kaikki valikoiden sisältämät sanat tai fraasit, jotka sitten soitetaan käyttäjälle vasteena tämän käyttöliittymässä tekemiin valintoihin. Puhesynteesiä ei siis kirjoittajien esittämässä toteutustavassa tarvita. Schmidt ja Wlodkowski [2003] jättävät kuitenkin mainitsematta, että myös yhdistelevän, ihmispuheen palasia käyttävän puhesynteesin käyttäminen olisi mahdollista myös puhuvien valikoiden tapauksessa. Yhdistelevää synteesiä käyttämällä päästäisiin lähelle puheen luonnollisuutta, ja joustavuus lisääntyisi, kun valikoita voitaisiin muokata ohjelmallisesti. Toisaalta Schmidt ja Wlodkowski [2003] esittävät äänitetyn ihmisäänen käyttämisen hyvänä puolena sen, että se on yhteensopiva nykyisen teknologian kanssa. Tämä tarkoittaa sitä, että äänitetty puhe voidaan tallentaa esimerkiksi digiboksin muistiin ilman erityisiä muutoksia laitteeseen. Yhdistelevän synteesin käyttäminen ei olisi todennäköisesti yhtä yksinkertaista.

Kokonaan syntetisoidun puheen käyttäminen puhuvissa valikoissa tuo mukanaan joustavuutta: kun halutaan vaihtaa puhuttua valikon osaa, voidaan syntetisaattorille yksinkertaisesti antaa uusi teksti luettavaksi. Lisäksi puheen ominaisuuksia, kuten nopeutta, voidaan helposti muuttaa, ja valikoiden luomisen kustannukset ovat pieniä, kun ei tarvitse palkata ihmistä niitä puhumaan [Schmidt and Wlodkowski, 2003]. Syntetisoidun puheen käytön kustannukset eivät ole suuret, sillä esimerkiksi suomenkielisiä syntetisaattoreita, kuten Mikropuhe (<http://www.mikropuhe.com/mikropuhe.asp>) on saatavilla edullisesti. Englanninkielisiä syntetisaattoreita sekä suomenkielisiä tutkimusprototyyppejä on saatavilla myös ilmaiseksi. Toisaalta uskottavinta puhetta tuottavat syntetisaattorit ovat kalliita, eikä niitä ole vielä valmiiksi upotettuna esimerkiksi digibokseihin [Schmidt and Wlodkowski, 2003].

Puhuvia valikoita käytävällä sovelluksella saattaa olla useita erilaisia käyttäjiä, joista toiset tarvitsevat puhenavigointia, ja toisia tämä ominaisuus pahimmassa tapauksessa häiritsee. Tämän takia puhenavigointi tulee aina voida kytkeä myös pois päältä. Schmidtin ja Wlodkowskin [2003] mukaan käyttäjälle tulee antaa heti käynnistämisen yhteydessä puhutut ohjeet siitä, miten puhenavigointi saadaan pois päältä. Käyttäjälle tulee samalla mainita myös siitä, miten puhenavigointi saadaan pois päältä missä tahansa myöhemmässä vaiheessa. Jos mahdollista, kaukosäätimeen kannattaa myös asettaa näppäin, josta puhenavigointi saadaan päälle tai pois päältä.

Käyttäjälle annettava palaute on kaikenlaisissa käyttöliittymissä ensiarvoisen tärkeä asia. Puhuvien valikoiden tapauksessa palaute tulee luonnollisesti aina antaa äänellisessä muodossa, joko puheena tai ei-kielellisinä ääнинä. Tämän lisäksi voidaan käyttää graafisesti esitettävää palautetta, jos käyttöliittymässä on myös graafisia elementtejä. Schmidt ja Wlodkowski [2003] luettelevat sääntöjä, joita puhuvan valikon tulee noudattaa palautteen antamisessa. Ensiksikin uudelle ruudulle siirryttäessä käyttäjälle tulee aina ilmoittaa ruudun otsikko. Käyttäjän valintaan tulee aina myös vastata osoittamalla jollain tavalla, että valinta on vastaanotettu. Esimerkiksi, kun siirrytään valikon kohdasta toiseen, tulee uuden kohdan nimi aina ilmoittaa käyttäjälle. Toisaalta käyttäjällä on hyvä olla myös mahdollisuus keskeyttää sovelluksen puhetulosteet esimerkiksi kaukosäätimen tai näppäimistön näppäintä painamalla.

Käyttäjälle on hyvä antaa mahdollisuus pyytää sovellusta toistamaan edellinen tuloste. Tämä on parempi vaihtoehto kuin tulosteiden toistaminen automaattisesti kerta toisensa jälkeen, jos käyttäjä ei reagoi mitenkään. Edellisen tulosteen toistamisen yhteydessä on hyvä toistaa myös sen hetkisen ruudun nimi, jotta käyttäjä pysyisi paremmin tietoisena siitä, missä kohtaa valikkoa on navigoimassa. [Schmidt and Wlodkowski, 2003]

Puhuvia valikoita käytettäessä kaukosäätimen tai näppäimistön käytön tulee olla mahdollisimman yksinkertaista. Schmidtin ja Wlodkowskin [2003] mukaan kaikkien välttämättömien toimintojen olisi hyvä olla käytettävissä pelkästään nuoli- ja valintanäppäimiä käyttämällä. Heidän mukaansa kannattaa myös harkita sellaista toiminnallisuutta, että käyttäjän painaessa esimerkiksi nuolta alaspäin valikon viimeisessä kohdassa, siirryttäisiin valikon ensimmäiseen kohtaan. Jos tällaiset ”kiertävät” valikot eivät ole mahdollisia, toinen mahdollinen vaihtoehto on jonkinlaisen äänimerkin tai muunlaisen palautteen antaminen valikon viimeisen kohdan jälkeen. On myös sellainen mahdollisuus, että valikon kohdat numeroidaan ja kohtien määrä mainitaan valikon nimen yhteydessä, jolloin käyttäjä tietää, kuinka monta kohtaa on tiedossa ja osaa odottaa valikon loppumista. Tällaisessakaan tilanteessa ei ole kuitenkaan hyvä jättää käyttäjää kokonaan vaille palautetta valikon loppuessa.

Puhuvien valikoiden suunnittelussa on pitkälti kyse visuaalisen informaation muuttamisesta puhutuksi. Ensimmäinen askel graafisen valikon muuttamisessa puhuvaksi valikoksi on sen rakenteen päättäminen. Valikon rakenteellinen malli voi olla esimerkiksi lineaarinen lista tai hierarkkinen rakenne. Se voi olla myös ruudukko, eli ryhmä listoja, joiden kohdilla on tiettyjä yhteisiä ominaisuuksia.

Kun suunnitellaan puhuvia valikoita, on erittäin tärkeää ottaa huomioon, että käyttäjää ei kuormiteta turhalla informaatiolla [Schmidt and Wlodkowski, 2003]. Kun informaatio on visuaalisessa muodossa, käyttäjän on helpompi keskittää huomionsa haluamiinsa asioihin. Informaation ollessa pelkästään äänenä jakaantuu käyttäjän huomio tasaisemmin koko kuultuun asiaan, eikä siitä niin helposti pystytä suodattamaan epäolennaisia asioita pois. Käyttöliittymän tulee siis esittää vain sen verran tietoa, kuin käyttäjä pystyy yhdellä kertaa käsittelemään. Edwardsin [2002] mukaan suunnittelijan täytyy päättää, mikä on käyttäjälle välittömästi esitettävää olennaisinta tietoa, ja mikä on vähemmän olennaista. Hänen mukaansa toissijaisen informaation ei tarvitse olla yhtä helposti saavutettavissa, vaan se voi esimerkiksi olla jonkin tietyn käyttäjän komennon takana, mutta kuitenkin saatavilla.

3.5 Puhekäyttöliittymät TV-ympäristössä

Aikaisemmin tässä luvussa olen käsitellyt puhekäyttöliittymien yleisiä suunnitteluperiaatteita, ja samat suunnitteluperiaatteet koskevat eri ympäristöihin suunniteltavia puhekäyttöliittymiä. Näen kuitenkin tarpeelliseksi tuoda esiin vielä joitakin näkökulmia, jotka koskevat erityisesti TV-ympäristöön suunniteltavia puhekäyttöliittymiä.

TV-ympäristön ja tietokoneympäristön erot tulee ottaa huomioon suunniteltaessa käyttöliittymiä TV-ympäristöön. Ibrahim ja Johanssonin [2002] mukaan TV-ympäristö eroaa perinteisestä tietokoneympäristöstä neljällä tavalla: (1) Välimatka TV-ruudun ja käyttäjän välillä rajoittaa ruudulla esitettävän informaation määrää. (2) TV-ruudulla on matalampi resoluutio kuin tietokoneen ruudulla, mikä vaikuttaa graafiseen ulkoasuun. (3) TV:tä käytetään yleensä viihdetarkoituksessa rennossa ilmapiirissä, usein sohvalla tai nojatuolissa istuen, kun taas tietokonetta käytetään usein myös työtarkoituksessa tai asioiden hoitamisessa. (4) TV pitää ääntä, mikä voi tehdä puhetulostuksen käytön sopimattomaksi. Kohdissa 1 ja 2 mainitut asiat eivät luonnollisesti vaikuta mitenkään pelkkään puheeseen perustuviin käyttöliittymiin, mutta koskevat sellaisia puhekäyttöliittymiä, joissa on myös graafisia elementtejä.

Ibrahim ja Johansson [2002] suosittelivat väitteen 4 (TV pitää ääntä, mikä voi tehdä puhetulostuksen käytön sopimattomaksi) pohjalta käyttämään TV-ympäristössä multimodaalisia käyttöliittymiä, joissa käyttäjän syöte tapahtuu puheella tai kaukosäätimellä, ja tulos-

tus on visuaalista. Kohdan 4 väite ei kuitenkaan ole oleellinen, kun suunnitellaan puhekäyttöliittymiä näkövammaisille, sillä etenkin vaikeasti näkövammaisille puhetulostus on joka tapauksessa hyödyllistä. Väitettä 4 tulkitessa tulee ottaa huomioon myös se mahdollisuus, että itse TV:n ääni saadaan mahdollisesti hiljennettyä siksi aikaa, kun selataan esimerkiksi digi-TV:n käyttöliittymään toteutettuja puhuvia valikoita. Joka tapauksessa on hyödyllistä, jos puhetulostuksen lisäksi käyttöliittymä on toteutettu myös visuaalisesti, sillä kaikki saman laitteen käyttäjät eivät välttämättä tarvitse tai halua käyttää puheohjausta, ja jotkut käyttäjät voivat myös arvostaa mahdollisuutta käyttää molempia ominaisuuksia.

Berglund ja Johansson [2004] suosittelevat, että TV-ympäristöön toteutetun käyttöliittymän navigoinnissa ei kaukosäädintä kokonaan korvattaisi puheohjauksella. Parempi toteutustapa olisi puheohjauksen mahdollistaminen ja kaukosäätimen pitäminen vaihtoehtoisena vuorovaikutusvälineenä. Puheohjaus on hyödyllinen monesta syystä ihmisille, joille kaukosäätimen käyttö on vaikeaa esimerkiksi näkövammaisen tai motoristen ongelmien takia. Puheohjaus tuo myös joustavuutta käyttöliittymän käyttämiseen aivan tavallisillekin käyttäjille. Puheohjausta voi kuitenkin olla joissain tilanteissa hankala käyttää, erityisesti TV-ympäristön luonteen huomioon ottaen: televisiota katsoo usein enemmän kuin yksi ihminen kerrallaan, ja he saattavat myös puhua keskenään. Tämä seikka vaikeuttaa puheohjauksen käyttämistä, samoin kuin myös se seikka, että joskus televisiota saatetaan haluta katsoa hiljaa sellaiseen aikaan, jolloin osa saman asunnon väestä nukkuu.

Rinnetmäen et al. [2004] mukaan suunniteltaessa sovelluksia televisioympäristöön, on tärkeää huomioida käyttäjän vireystila. Sovelluksilla tulee olla todella helppo ja selkeä käyttöliittymä, sillä käyttäjien vireystila saattaa vaihdella hyvinkin väsyneestä kiihtyneeseen esimerkiksi urheilua seurattaessa. Lisäksi, jotta sovellusten käyttö olisi mahdollista television seuraamisen ohessa, ei käyttäjältä saisi edellyttää muistamista. Tämän takia kaikkien toimintojen tulisi olla joko heti käytettävissä tai ainakin helposti löydettävissä näkyvien viittausten avulla. [Rinnetmäki et al., 2004]

Ihminen pystyy hahmottamaan kerralla viidestä yhdeksään näytöllä olevaa valintaa, riippuen yksilöstä ja tilanteesta. Koska television käyttäjäkunta on niin moninainen, ja käyttötilanteet vaihtelevat, on hyvä käyttää vain viidestä seitsemään valintaa kerralla. Selkeä ryhmittely auttaa myös hahmottamista, joten suuremmat ryhmät kannattaa jakaa kahteen tai useampaan selkeästi erotettavaan valintojen ryhmään, jos mahdollista [Rinnetmäki et al., 2004]. Näitä suosituksia voidaan pitää pätevinä myös puhetulostusta käyttävien käyttöliittymien ollessa kyseessä.

4 Puhekäyttöliittymien arviointimenetelmät

Kaikenlaisten sovellusten kehittämisessä yksi tärkeä vaihe on sovellusten arviointi. Puhe-sovelluksia voidaan arvioida kahdesta eri näkökulmasta. Ensiksikin voidaan arvioida teknologisten komponenttien selviytymistä niille osoitetuista tehtävistä. Toinen lähestymistapa arviointiin on käytettävyyden arviointi, joka ottaa huomioon käyttöliittymään liittyvät asiat ja inhimillisen näkökulman.

4.1 Teknologian arviointi

Puheentunnistuksen arviointiin yleisimmin käytetty menetelmä on sanojen virheasteen (word error rate, WER) tai sanojen tarkkuusasteen (word accuracy rate) mittaaminen. Virheaste tarkoittaa väärin tunnistettujen sanojen osuutta puhutuista sanoista. Tarkkuusaste puolestaan tarkoittaa vastaavasti oikein tunnistettujen sanojen osuutta puhutuista sanoista. Vastaavanlaiset prosenttiosuudet voidaan laskea myös lauseille tai käsitteille, tai muille puheen osatekijöille. [Turunen, 2004]

Ymmärrettävyys ja miellyttävyys ovat tekijät, joiden perusteella puhetulosteiden laatua voidaan arvioida. Ymmärrettävyys voidaan jakaa segmenttiseen ymmärrettävyyteen (segmental intelligibility) ja ymmärrykseen (comprehension). Segmenttinen ymmärrettävyys kertoo, kuinka tarkasti käyttäjä on vastaanottanut tulostetut lauseet, ja sen mittaaminen voidaan toteuttaa samaan tapaan kuin virheasteen mittaaminen [Turunen, 2004]. Esimerkiksi sovelluksen tuottamien sanojen määrää voidaan verrata käyttäjien oikein tulkittujen sanojen määrään. Sama vertailu voidaan toteuttaa myös foneemien, tavujen tai fraasien perusteella [D'Alessandro and Liénard, 1997]. Ymmärrys puolestaan kertoo, kuinka hyvin käyttäjä ymmärtää tulostetut lauseet. Sitä voidaan mitata kysymyksillä tai tehtävillä, joita tehdessä käyttäjän täytyy ymmärtää viestien merkitys. [Turunen, 2004].

Miellyttävyys on subjektiivinen määre, joka voi olla käyttäjän kannalta yhtä tärkeässä asemassa puhekäyttöliittymän laatua arvioitaessa kuin ymmärrettävyys. Miellyttävyyteen liittyy monia tekijöitä, joista yksi on luonnollisuus. Luonnollisuus viittaa siihen, kuinka paljon käyttäjän tulkinnan mukaan syntetisoitu puhe vastaa ihmisen puhetta. Toisaalta se, kuinka luonnolliselta käyttäjät haluavat puheen kuulostavan, riippuu käyttäjästä, sovelluksen tehtävästä ja asiayhteydestä [D'Alessandro and Liénard, 1997].

Tulostetun puheen miellyttävyyttä voidaan mitata käyttäjien mielipiteitä keräämällä. Kun kerätään mielipide suurelta määrältä ihmisiä, saadaan laskettua mielipiteiden keskiarvo (mean opinion score), jonka perusteella voidaan mitata puheen laatua yleisesti [Turunen,

2004; Pols, 1997]. Miellyttävyys ja ymmärrettävyys ovat yhteydessä toisiinsa, mutta eivät kuitenkaan täysin suoraviivaisesti vaikuta toisiinsa. Näillä tekijöillä on joka tapauksessa suuri merkitys siihen, miten käyttäjät ottavat sovelluksen vastaan, sillä epämiellyttävä puhetulos voi johtaa käyttäjien tyytymättömyyteen muuten pitkälle kehitetyn järjestelmän käytössä [Turunen, 2004].

4.2 Puhesovellusten arviointi

Kun puhesovelluksia kehitetään, teknologian arvioinnin tuloksia tulee tulkita myös käytettävyyden näkökulmasta. Tämä ei ole kovinkaan yksinkertaista, sillä esimerkiksi monissa sovelluksissa 20 prosentin virheaste voi olla hyväksyttävä, mutta toisissa puolestaan viidenkin prosentin virheaste on mahdotonta hyväksyä. [Turunen, 2004]

Puhesovelluksia voidaan arvioida sekä subjektiivisten että objektiivisten kriteerien perusteella. Turunen [2004] mukaan objektiivisia kriteereitä, kuten tehtävien suoritusasteita ja suoritusajoja, pidetään yleensä ensisijaisina kriteereinä sekä puhesovelluksia että muunlaisia sovelluksia arvioitaessa. Objektiivista suorituskyvyn mittaamista voidaan Turunen mukaan käyttää hyväksi puhejärjestelmien hienosäädössä, mutta myös subjektiiviset kriteerit täytyy ottaa arvioinnissa huomioon, sillä ne vaikuttavat suuremmassa määrin käytettävyyteen.

Subjektiiviset kriteerit liittyvät käyttäjien tyytyväisyyteen ja mieltymyksiin. Hartikaisen et al. [2004] mukaan subjektiivisella arvioinnilla on lukuisia etuja objektiiviseen arviointiin verrattuna. Yksi etu on se, että subjektiiviset menetelmät ovat ainoita tapoja saada selville käyttäjien mielipiteitä sovelluksesta. Subjektiiviselle arvioinnille ei myöskään välttämättä ole vaihtoehtoja, jos arvioitava sovellus käsittelee esimerkiksi potilastietojen tai kirjeenvaihdon kaltaisia luottamuksellisia tietoja. Näiden seikkojen lisäksi pelkän subjektiivisen arvioinnin toteuttaminen on kustannuksiltaan edullista, eikä subjektiivisen kyselyn analysoimiseen tarvita mitään erityistaitoja.

Wizard of Oz (WOZ) -tutkimukset ovat tärkeä keino puhekäyttöliittymien käytettävyydestä ja suunnitteluprosessille hyödylliseen tiedonkeruuseen. Niissä simuloidaan sovellusten toimintaa siten, että käyttäjä uskoo käyttävänsä oikeaa järjestelmää, mutta todellisuudessa järjestelmän toimintaa, joko kokonaan tai osittain, ohjaa ihminen [Bernsen et al., 1999]. Suunnitteluprosessin aikana tehtyä WOZ-tutkimusta voidaan käyttää varmistamaan suunniteltujen vuorovaikutustapojen sopivuutta tehtävänsä ennen niiden toteuttamista. WOZ-tutkimusten yksi ongelma on se, että järjestelmän toiminnan simuloiminen on monimutkaista, kun samalla pitää vastata käyttäjälle tarkasti ja riittävän

nopeasti [Turunen, 2004]. Lisäksi virheiden ja muiden teknologisista seikoista riippuvien rajoitusten simuloiminen voi olla vaikeaa [Turunen, 2004; Yankelovich, 1997].

Käytettävyyystutkimusten toteuttaminen toimivalla tai simuloidulla järjestelmällä on olennainen osa sovellusten kehittämistä, mutta myös ihmisten välisen vuorovaikutuksen tutkiminen voi oikein toteutettuna antaa hyödyllistä tietoa ihmisen ja tietokoneen välisen vuorovaikutuksen suunnitteluun. Ihmisten välistä vuorovaikutusta voidaan tutkia tarkkailemalla ihmisten tapaa suorittaa erilaisia tehtäviä, vaikkapa tallentamalla osapuolten välinen keskustelu esimerkiksi jonkinlaisen aikataulupalvelun käyttämisen yhteydessä. Tällaisten tutkimusten ongelma on kuitenkin se, että ihmisten välinen kommunikointi eroaa monella tavoin ihmisen ja tietokoneen välisestä kommunikoinnista. Ihmisten välisen vuorovaikutustutkimuksen tulosten sopivuus tuleekin varmistaa perusteellisesti ennen kuin tuloksia käytetään perustana ihmisen ja tietokoneen välisen vuorovaikutuksen suunnittelussa. [Turunen, 2004]

Puhesovelluksen kehittäminen on usein iteratiivista, ja toteutetun sovelluksen oikea käyttäminen voi tuoda esiin monia sellaisia vuorovaikutuksen ongelmia, joita ei ilmenisi simuloiduissa käyttötilanteissa tai ihmisten välisessä vuorovaikutuksessa. Ihmisen ja tietokoneen välisiä tutkimuksia käytetään tiedonhankinnan ja järjestelmätestauksen lisäksi sopivien ja epäsovivien sovellusalueiden löytämiseen puhesovellusten kannalta [Turunen, 2004]. Yankelovich [1997] kutsuu tällaisia tutkimuksia esitutkimuksiksi (pre-design studies). Hänen mukaansa esitutkimukset ovat hyödyllisiä seuraavissa tehtävissä:

1. Sovelluksen vaatimusten ja toiminnallisuuden jalostaminen.
2. Sopivan sanaston kerääminen.
3. Yleisesti käytettyjen kieliopillisten rakenteiden määrittäminen.
4. Tehokkaiden vuorovaikutusmallien löytäminen.
5. Tulosteiden ja palautteen suunnittelun helpottaminen.
6. Keskustelun sävyn löytäminen.

Ihmisen ja tietokoneen välisen vuorovaikutuksen tutkimisessa voi myös olla ongelmia, kuten se, että monet tutkimukset olettavat teknologian olevan täydellistä, eivätkä ota teknologisia rajoituksia huomioon. Tämänkin takia tulee tutkimusten tuloksia analysoida erityisen huolellisesti. [Turunen, 2004]

5 Näkövammaisten ihmisten ottaminen huomioon puhe- käyttöliittymän toteutuksessa

Näkövammaisuus on melko laaja käsite, sillä näkövammoja on eriasteisia ja erityyppisiä. Lisäksi näkövammaiset ovat erittäin heterogeeninen ryhmä, sillä heitä löytyy muun muassa kaikista ikäluokista. Tässä luvussa määrittelen lyhyesti, mitä näkövammaisuus tarkoittaa ja selvitän, miten näkövammaiset tulisi huomioida puhekäyttöliittymän toteutuksessa.

5.1 Näkövammaisuuden määrittely

Näkövammaisuus voi merkitä joko sokeutta tai heikkonäköisyyttä. Ihmistä, jonka näkökyky saadaan laseilla korjattua normaaliksi, tai jolla on toisessa silmässä normaali näkö, ei luokitella näkövammaiseksi. [NKL, 2006]

Näkövammaisuus voi ilmetä muun muassa näöntarkkuuden tai näkökentän muutoksina. Myös kontrastien erotuskyvyssä, värinäössä, silmien sopeutumisessa muuttuviin valaistusolosuhteisiin, silmälihasten toiminnassa sekä silmän mukautumisessa eri etäisyyksille voi olla muutoksia [Huvio et al., 2004]. Voi olla, että heikkonäköinen ihminen ei näe lukea, mutta pystyy liikkumaan ilman valkoista keppiä. Toisaalta ihminen voi kyetä lukemaan, mutta ei näe ympäristöään [NKL, 2006].

Monet sokeat voivat nähdä valon tai jopa hahmoja, joten täydellinen sokeus on harvinaista. Toiminnallisesti sokeiksi määritellyt henkilöt ovat menettäneet muun muassa suuntausnäön vieraassa ympäristössä sekä lukunäön tavanomaisessa mielessä, mutta saattavat kuitenkin pystyä lukemaan erityisapuvälineiden avulla. [NKL, 2006]

On arvioitu, että Suomessa on noin 80 000 näkövammaista ihmistä, eli 1,55 prosenttia väestöstä. Noin 10 000 heistä on sokeita ja loput ovat eri tavoin heikkonäköisiä. Noin 80 prosenttia Suomen näkövammaisista on yli 65-vuotiaita, ja nuorten ja lasten osuus on vain noin 5 prosenttia. Yleisin näkövammaisuuden aiheuttaja Suomessa onkin ikääntyneiden ihmisten silmänpohjan rappeuma. Muita aikuisten näkövammojen aiheuttajia ovat verkkokalvon perinnölliset rappeumat, diabeteksesta johtuvat silmänpohjamuutokset, sekä glaukooma eli silmänpainetauti. Työsuojelun ja lääketieteen kehittymisen ansiosta työtapaturmien ja tulehdusten aiheuttamien näkövammojen määrä on pienentynyt. Lasten näkövammojen yleisimpiä aiheuttajia ovat synnynnäiset kehityshäiriöt ja näköhermoston viat. Keskosuus on lapsilla neljänneksi yleisin näkövammaisuuden aiheuttaja, vaikka keskosuuteen liittyvä näön-

menetyks on vähentynyt. Noin 75 prosentilla vuosittain syntyvistä näkövammaisista lapsista on myös muita vammoja. [NKL, 2006]

5.2 Puhekäyttöliittymien suunnittelemisen näkövammaisille

Näkövammaiset ovat monessa mielessä laaja kohderyhmä, joten käyttöliittymien suunnittelemisen heille on haasteellinen tehtävä. Näkövammoja on monen tasoisia ja monen laatuista ja näkövammaisia on kaikissa ikäluokissa: näin ollen käyttötottumuksia ja -tarpeita on todella monenlaisia. Esimerkiksi nuoret ja keski-ikäiset näkövammaiset ovat monesti tottuneita ja halukkaita käyttämään monenlaisia laitteita, ja vanhemmat näkövammaiset mahdollisesti aivan päinvastoin. Koska näkövammaisista noin 80 prosenttia on ikääntyneitä, täytyy vanhuksien kuitenkin ottaa erityisesti huomioon kohderyhmänä.

Oman tutkimukseni piiristä rajaan pois sellaiset näkövammaiset henkilöt, joilla on näkövamman lisäksi myös muita vammoja, jotka vaikeuttavat normaalia vuorovaikutusta ja tietokoneen käyttöä. On kuitenkin huomattava, että esimerkiksi vanhuuden aiheuttama motorinen kankeus on jossain määrin verrattavissa heikkonäköisyyteen: molemmille ryhmille ovat hyödyksi esimerkiksi laitteiden suuret näppäimet.

Kuten aiemminkin tässä tutkielmassa on tullut ilmi, pelkkään puheeseen perustuvien puhe-käyttöliittymien lisäksi puhe-käyttöliittymiä voi olla myös sellaisia, joissa tulostukseen käytetään puhetta ja käyttäjä ohjaa käyttöliittymää esimerkiksi näppäimistön tai kaukosäätimen avulla. Kolmas mahdollinen puhe-käyttöliittymätyyppi on sellainen, jota ohjataan puheella, mutta tulostus tapahtuu visuaalisesti. Näitä tyyppisiä voidaan myös yhdistää niin, että käyttöliittymän tulostus voi tapahtua sekä puheella että visuaalisesti, tai niin, että käyttöliittymän ohjaaminen voi tapahtua sekä puheella että näppäimien kautta. Täysin sokeille ihmisille ovat luonnollisesti käyttöliittymän visuaaliset elementit turhia, ja puhe-tulostus on tarpeen. Sokeatkin pystyvät kuitenkin käyttämään esimerkiksi kaukosäädintä, jos siinä on suurikokoiset näppäimet, joissa on hyvä tuntopalaute. Suurikokoisten näppäimien ohella myös selkeä ryhmittely auttaa näkövammaisia muistamaan paremmin näppäimien paikat, vaikka näkökykyä ei olisi ollenkaan jäljellä.

Sellaisia näkövammaisia ihmisiä varten, joilla on heikentymiä näöntarkkuudessa, mutta jotka näkevät kuitenkin jonkin verran, voi puhe-käyttöliittymässä olla myös graafisia elementtejä. Parhaiten heidät voidaan ottaa huomioon siten, että graafisia elementtejä, esimerkiksi fonttia, suurennetaan. Ne voivat mahdollisesti olla jo oletuksena normaalia suurempia, mutta erityisen tärkeä on mahdollisuus elementtien koon muuttamiseen tai yleiseen suurentamiseen. Myös esimerkiksi fontin tyyli on hyvä olla muutettavissa, sillä näkökyvyn mukaan voi vaihdella se, mikä fontti koetaan selkeimmäksi ja miellyttävimmäksi.

Näkövammaisuus voi ilmetä muun muassa alentuneena kontrastien erotuskykynä. Kontrastien erotuskyvyllä tarkoitetaan kykyä erottaa vierekkäisten pintojen vaaleusaste-eroja [Huvio et al., 2004]. Käyttöliittymässä käytetyillä väreillä tulee siis olla selkeä kontrastiero, ja värejä tulee myös pystyä muuttamaan. Myös värinäkö, eli silmän kyky erottaa värisävyjä, voi olla heikko. Tämänkin takia käyttöliittymän värisävyjä tulee miettiä tarkkaan. Kontrastien erotuskyvyn ja värinäön lisäksi silmien sopeutuminen muuttuviin valaistusolosuhteisiin on heikkonäköisillä usein hidastunut ja heikentynyt [Huvio et al., 2004]. Tämän takia käyttöliittymän värien kirkkaudessa ei saa tapahtua suuria muutoksia.

Silmälihasten hallintavaikeudet huonontavat etäisyyksien arviointikykyä [Huvio et al., 2004]. Tämän takia käyttöliittymän visuaalisten elementtien selkeä erottelu on tärkeää. Silmän mukautuminen eri etäisyyksille heikkenee yleensä iän myötä, mutta näkövammaisella saattaa olla heikko silmien mukautumiskyky varhaislapsuudesta lähtien [Huvio et al., 2004].

Vaikka käyttöliittymässä olisi myös visuaalisia elementtejä, monet näkövammaiset eivät niitä näe. Tämän vuoksi kaikki käyttöliittymässä saatu palaute tulisi olla mahdollista saada myös äänenä. Äänellinen palaute on kuitenkin hyvä saada myös kytkettyä pois päältä, sillä se voi olla rasittavaa niille käyttäjille, jotka eivät sitä tarvitse. Joka tapauksessa käyttäjän tulisi saada jonkinlainen palaute kaikista toimistaan käyttöliittymässä, jotta hän ei joutuisi miettimään, menikö hänen lähettämänsä viesti perille vai ei. Esimerkiksi jo virtanäppäintä painaessa tulisi kuulua jonkinlainen ilmoitus joko puheella tai muunlaisella äänimerkillä siitä, että laite on käynnistymässä, sillä näkövammaisen ei todennäköisesti laitteeseen syttyvää valoa tai muuta visuaalista palautetta näe.

Jos puheikäyttöliittymässä ei ole visuaalisia elementtejä, ja se perustuu pelkkään puheeseen, eivät näkövammaiset suuresti eroa käyttäjinä ihmisistä, jotka näkevät normaalisti. Siinä näkövammaiset tosin saattavat erota niin sanotuista normaaleista käyttäjistä, että he ovat tottuneempia käsittelemään kuultua informaatiota ilman visuaalista lisäinformaatiota. Tämän vuoksi näkövammaiset saattavat toivoa puheikäyttöliittymältä esimerkiksi nopeampaa puhetta, jotta sovelluksen käyttäminen olisi tehokkaampaa. Tämän seikan takia myös puheen nopeuden muuttamisen mahdollisuutta on hyvä harkita puheikäyttöliittymää suunnitellessa.

Puheikäyttöliittymien suunnittelijoiden on hyvä suorittaa käyttäjätestejä oikealla kohderyhmällä. Suunnittelijat, joilla ei useinkaan itsellä ole näkövammoja, eivät kovin helposti pysty kuvittelemaan sitä, millaista on olla näkövammaisen. Pelkästään silmien sulkeminen ei todennäköisesti kerro koko totuutta. Esimerkiksi Ahon et al. [2005] mukaan enemmistö näkövammaisista kannattaa televisio-ohjelmien tekstityksen puhetulkkauksessa mieluum-

min hieman monotonisempaa puhetta, kun kyse on puheen elävyydestä. Monotoninen puhe häiritsee keskittymistä eläväistä puhetta vähemmän. Ihmisestä, jolla on normaali näkö, voi puolestaan monotoninen puhe tuntua kovin robottimaiselta ja oudolta, kun esimerkiksi juuri tekstityksen puhetulkkaukseen ei jouduta omassa arkielämässä turvautumaan.

6 Tapaus ”Aistiboxi”

Huvion et al. [2004] mukaan digitaalisen television käyttöliittymien tulee olla yksinkertaisia käyttää, sillä katsojien taso erilaisten käyttöliittymien käyttämisessä vaihtelee suuresti. Käyttöliittymiä kehitettäessä on muistettava ottaa huomioon myös erityisryhmien tarpeet. Kustannuksia ei heidän mukaansa tule olennaisesti lisää, jos vaatimukset otetaan huomioon jo suunnitteluvaiheessa. Näihin vaatimuksiin pyrkii vastaamaan ”Digital Television for All” -hanke, jota tutkielmassani kutsun lyhyemmin nimellä ”Aistiboxi”.

Hanke toteutetaan yhteistyössä Turun yliopiston, Taideteollisen korkeakoulun, Tampereen yliopiston ja Teknillisen korkeakoulun kesken. Hankkeen tavoitteena on tuoda digitaalinen televisio ja sen kautta lähitulevaisuuden tietoyhteiskuntapalvelut kaikkien ulottuville. Pilottivaiheessa hankkeen tarkoituksena on suunnitella erityisesti näkövammaisten ja vanhusten käyttöön soveltuva digiboksi. Tuotteen teknisestä toteutuksesta vastaa turkulainen Axel Technologies, joka on myös toteuttanut tuotteesta alustavan esittelyprototyypin.

Tässä luvussa esittelen kohderyhmän haastattelun tulokset, hankkeen päämäärät ja tutkielmani tekovaiheessa voimassa olevat suunnitelmat. Tutkielmani aikaisempien lukujen sekä kohderyhmän haastattelussa ilmenneiden seikkojen perusteella pyrin alustavasta esittelyprototyypistä liikkeelle lähtien antamaan ehdotuksen sellaisesta käyttöliittymästä, joka on käytettävyydeltään hyvä myös näkövammaisille ihmisille.

6.1 Kohderyhmän haastattelu

”Aistiboxi”-hankkeen puitteissa kaksi henkilöä toteutti näkövammaisten haastattelun, jossa pyrittiin selvittämään näkövammaisten toiveita helppokäyttöiseen digiboksin käyttöliittymään ja kaukosäätimeen liittyen. Molemmat haastattelijat ovat korkeakouluopiskelijoita, joista toinen teki Taideteolliseen korkeakouluun opinnäytetyötä saavutettavasta kaukosäätimestä ja toinen oli projektiluontoisesti toteuttamassa haastattelua Näkövammaisten Keskusliiton toimeksiannosta. Haastattelu toteutettiin ryhmähaastatteluna, jossa kaikkia kuutta haastateltavaa haastateltiin yhtä aikaa samassa tilassa puheenvuoroja jakaen. Analysoin kysymykset nauhoitetun materiaalin perusteella, jonka sain käyttöni tutkielmaani varten.

Kaikki haastatellut ovat työikäisiä, noin 30-50-vuotiaita. Viisi haastatelluista on miehiä ja yksi nainen. Neljä haastatelluista on täysin sokeita, ja yhdellä miehellä ja yhdellä naisella on lievempi näkövamma, mutta he molemmat ovat kuitenkin erittäin heikkonäköisiä.

Haastateltavilta kysyttiin kokemuksia ja mielipiteitä liittyen seuraaviin osa-alueisiin: television katselu, television ja digiboksin ohjaaminen kaukosäätimen tai etupaneelin avulla, digiboksin ominaisuudet ja erityistoiminnot, ja digiboksin valikot.

6.1.1 Television katselu

Haastateltavilta kysyttiin, mitä televisio-ohjelmia he yleensä seuraavat ja milloin. Heiltä kysyttiin myös, että käyttävätkö he saatavilla olevaa äänitekstityspalvelua ja mistä syistä he sitä käyttävät tai eivät käytä.

Kaikkien haastateltavien televisionkatselu keskittyy pääsääntöisesti ilta-aikaan. Neljä haastatelluista miehistä seuraa enimmäkseen urheilua, uutisia ja ajankohtaisohjelmia. Yksi miehistä seuraa lähinnä kotimaista asiaviihdettä ja dokumentteja, eikä seuraa uutisia, koska kokee niiden olevan niin visuaalisia, että hän sokeana menettää suuren osan informaatiosta. Haastateltu nainen puolestaan katsoo enimmäkseen elokuvia, sarjoja ja dokumentteja.

Haastatelluista viidellä on digiboksi, ja yhdellä ei ole. Digiboksittoman henkilön televisionkatselutottumukset eivät haastattelun perusteella erityisemmin eroa muiden henkilöiden televisionkatselutottumuksista. Yleisradion tarjoaman äänitekstityspalvelun käyttöön digiboksin olemassaolo luonnollisesti vaikuttaa, sillä palvelu on käytettävissä vain digiboksin kautta. Äänitekstityspalvelussa TV:n tekstitys muunnetaan synteettiseksi puheeksi, joka on valinnaisesti kuultavissa normaaliäänän lisäksi [Aho et al., 2005]. Neljä haastatelluista henkilöistä ei käytä äänitekstityspalvelua. Tärkein syy siihen, miksi äänitekstitystä ei käytetä, on se, että monet kokevat puhesynteesin äänenvoimakkuuden liian kovaksi: ohjelmaan joutuu keskittymään todella tarkasti, jos haluaa saada selvän myös ohjelman äänistä puhesynteesin taustalla. Kaksi haastatelluista henkilöistä kuitenkin pitää äänitekstitystä käyttökelpoisena erityisesti dokumenttien katselemisen yhteydessä.

6.1.2 Television ja digiboksin ohjaaminen kaukosäätimen tai etupaneelin avulla

Kolme haastateltavista käyttää television ja digiboksin ohjaamiseen ainoastaan kaukosäädintä. Kaksi haastateltavista käyttää pääsääntöisesti kaukosäädintä, ja silloin tällöin etupaneelia. Yksi sokea haastateltava käyttää laitteiden ohjaamiseen etupaneelia, ja ohjaa kaukosäätimellä ainoastaan digiboksia, jossa ei ole etupaneelia.

Toiminnot, joita haastateltavat pääsääntöisesti käyttävät kaukosäätimessä, ovat virran kytkeminen päälle tai pois, äänenvoimakkuuden säätö sekä kanavien vaihto. Kanavien vaihtamiseen he käyttävät vaihtelevasti sekä numeroita että nuolinäppäimiä. Toinen

haastatelluista heikkonäköisistä käyttää myös ohjelmaopasta sekä digiboksin tallennusominaisuuden liittyviä näppäimiä. Kaksi haastatelluista käyttää myös kielen valintaa.

Ainoastaan yhdellä haastatelluista ei ole ollut ongelmia kaukosäätimien kanssa. Kyseinen henkilö näkee jonkin verran, ja hänellä on vain yksi kaukosäädin käytössä. Muilla haastatelluilla on ollut monenlaisia ongelmia kaukosäätimien käytössä. Haastattelun perusteella runsaasti ongelmia näkövammaisille käyttäjille aiheuttaa käytössä olevien kaukosäädinten suuri määrä ja sen tunnistaminen, mitä laitetta milläkin kaukosäätimellä ohjataan. Etenkin se on ongelma, että useimpien kaukosäätimien näppäimet on ryhmitelty eri tavoilla. Esimerkiksi jossain kaukosäätimessä numeronäppäimistö on ylhäällä ja toisessa alhaalla, ja virtanäppäimen paikka vaihtelee kaukosäädinten yläreunassa. Jos kaukosäätimiä olisi käytössä vain yksi, sen näppäinten paikat oppisi helpommin muistamaan. Myös laitteiden etupaneelien käyttö auttaa ongelmiin jonkin verran, sillä etupaneelia käyttäessä tietää helpommin, mitä laitetta käyttää, vaikka sitä ei näkisikään: itse laitteen sijainnin näkövammaisenkin tietää helposti, kun taas kaukosäätimien paikat saattavat vaihdella.

Myös näppäinten monitilaisuus ja palautteen vähyyt vaikeuttavat näkövammaisten haastateltujen kaukosäätimen käyttöä. Vaikeasti näkövammaisen ei saa kaukosäätimellä annettujen komentojen perille menosta muuta palautetta kuin sen, että televisiosta tuleva ääni mahdollisesti muuttuu. Tämän takia monien kaukosäätimellä ohjattavien toimintojen käyttäminen on näkövammaisille hankalaa, tai jopa mahdotonta. Jonkinlainen äänellinen palaute auttaisi näkövammaisia monitilaisten näppäimien käytössä, jos he sen avulla saisivat tiedon siitä, mitä ominaisuutta tai toimintoa tällaisilla näppäimillä kulloinkin ohjataan. Monitilaiset näppäimet ovat joka tapauksessa lähes välttämättömiä, sillä jos jokaisella toiminnolla olisi oma näppäimensä, tulisi kaukosäätimistä suuresta näppäinten määrästä johtuen helposti joko täyteen ahdettuja tai liian suuria yhdellä kädellä käytettäviksi. Kaikki haastatellut pitivätkin tärkeänä sitä, että kaukosäädintä pystyy käyttämään yhdellä kädellä.

Useimpien haastateltujen mielestä kaukosäädintä ei tarvitse pystyä käyttämään heti ensi kosketuksella ilman minkäänlaista opastusta. Yksi haastateltava mainitsikin, että oman kaukosäätimensä oppii tuntemaan, kunhan se on selkeä ja sen käyttöön saa aluksi opastusta. Haastattelussa tuli myös moneen kertaan esille, että huonostakin kaukosäätimestä myös vaikeasti näkövammaisen ihminen oppii käyttämään perustoimintoja, mutta valikoiden käyttämiseen näkövammaiset tarvitsisivat jonkinlaista erityistä palautetta, sillä nykyisissä käyttöliittymissä ei valikoiden käyttäminen yleensä onnistu.

Haastatellut olivat yhtä mieltä siitä, että kaukosäätimen näppäinten ryhmittely on ensisijaisen tärkeä asia niiden käytettävyydessä. Hyvää kaukosäädintä käyttäessä näppäimet pystyy

helposti löytämään tunnustelemalla. Esimerkiksi numeronäppäimet löytyvät helposti, jos ne on selkeästi ryhmitelty omaksi kokonaisuudekseen ja erotettu muista näppäimistä jollain keinolla, kuten tilan käytöllä tai kaukosäätimen pinnan muodoilla. Numeronäppäimissä on myös hyvä olla jonkinlainen pieni merkki numeron viisi näppäimessä, sillä osaksi sen avulla ovat ainakin monet haastatelluista näkövammaisista tottuneet etsimään haluamansa numeronäppäimet. Myös näppäinten kokoeroja pidettiin yhtenä kaukosäätimen hahmottamista helpottavana tekijänä.

Heikkonäköiset haastatellut toivoivat myös selkeää kontrastia näppäinten ja taustan välille. Muuten he pitivät värien käyttöä ongelmallisena asiana, sillä heikkonäköiset näkevät värejä niin monilla eri tavoilla, että on vaikeaa löytää kaikille sopivia väriyhdistelmiä.

Suhtautuminen digiboksin etupaneeliin vaihteli kovasti haastateltavien keskuudessa. Kaksi haastatelluista koki etupaneelin erittäin tärkeäksi, kaksi merkityksettömäksi, yksi melko merkityksettömäksi ja yksi neutraaliksi. Etupaneelin näppäinten ryhmittelyyn osa haastateltavista ei ottanut ollenkaan kantaa, mutta he painottivat, että näppäinten tulee selkeästi erottua toisistaan. He, jotka ryhmittelyyn ottivat kantaa, kannattivat melko yksimielisesti perinteistä ryhmittelyä: virtakytkin vasemmassa reunassa ja valikkonäppäin, nuolinäppäimet sekä OK-näppäin oikeassa reunassa.

6.1.3 Digiboksin ominaisuudet ja erityistoiminnot

Haastateltavilta kysyttiin myös, miten tärkeinä he pitävät digiboksi-laitteen erilaisia mahdollisia ominaisuuksia, joita ovat digi-TV, DVD-soitin, webradio, webkamera, WWW-selain, ja tallennusmahdollisuus. Tärkeimmäksi ominaisuudeksi koettiin luonnollisesti digi-TV. Tallennusmahdollisuuden sekä DVD-soittimen haastateltavat kokivat keskimäärin myös melko tärkeiksi ominaisuuksiksi. Webkamera ja webradio eivät olleet haastateltavien mielestä tärkeitä ominaisuuksia, mutta eivät aivan merkityksettömiäkään. Vaikka kaikki haastatellut henkilöt käyttävät Internetiä, pitivät he WWW-selainta melko merkityksettömänä ominaisuutena digiboksissa.

Digiboksin mahdollisten ominaisuuksien lisäksi haastateltavilta kysyttiin, miten tärkeinä he pitävät digiboksin mahdollisia erityistoimintoja, kuten suurennettavaa tekstitystä, puheesynteesiä palautteena valikkonavigoinnista tai mistä tahansa elektronisen ohjelmaoppaan navigoinnista, ja puheohjausta puheentunnistuksen perusteella.

Suurennettavaa tekstitystä suurin osa haastateltavista piti melko tärkeänä tai tärkeänä. Sokeat ihmiset eivät luonnollisesti tarvitse suurennettavaa tekstitystä, mutta ainakin yksi so-

kea haastateltava mainitsi, että pitää silti asiaa tärkeänä vierailmassa käyvien ”heikkareiden” eli heikkonäköisten ystäviensä takia.

Puhesynteesiin palautteena navigoinnista kaikki haastateltavat suhtautuivat erittäin myönteisesti: kaikki antoivat sille numeron 5, eli erittäin tärkeä. Toinen heikkonäköisistä haastateltavista myös mainitsi, että jos tämän saisi, niin muuta heiltä ei tarvitsisikaan enää kysyä.

Puheentunnistuksen perusteella tapahtuvaan digiboksin puheohjaukseen haastateltavat puolestaan suhtautuivat varauksellisesti, ja tämä toiminto sai kaikilta arvosanan 1 tai 2, eli merkityksetön tai melko merkityksetön. Yksi sokea haastateltava mainitsi, että näppäimien löytäminen ohjauspaneelista tai kaukosäätimestä voi olla jopa helpompaa kuin muistaa puheohjaukseen tarvittavat komennot. Hänen mielestään ratkaiseva asia on se, että näppäimien järjestys olisi selkeä ja siihen olisi saatavilla palaute esimerkiksi puhesynteesillä. Mielipide sai myös muiden haastateltavien hyväksynnän. Osa haastateltavista epäroï myös puheentunnistusteknologian toimivuutta, kuten esimerkiksi puheentunnistuksen reagoïntia taustahälyyn tai television kaiuttimista tulevaan puheeseen.

Haastateltavilta kysyttiin myös, tekevätkö he digiboksiin suosikkilistoja kanavista. Kuukaan haastatelluista henkilöistä ei ole tehnyt digiboksiinsa tällaisia suosikkilistoja. He olivat yksimielisesti sitä mieltä, että suosikkilistojen tekeminen vaatii näkökykyä.

6.1.4 Digiboksin valikot

Haastateltavilta kysyttiin, miten he järjestäisivät eri toiminnot sekä asetukset digiboksin päävalikossa. Viisi haastatelluista järjesti TV:n ensimmäiseksi valinnaksi, ja yksi radion. Toisena valintana useimmilla oli radio. Yksi sokea haastateltava järjesti WWW-selaimen toiseksi valinnaksi, mutta mainitsi kuitenkin, että hänen mielestään selaimen ”sun muiden” yhdistäminen samaan laitteeseen on epärealistista, sillä tuloksena on niin monimutkainen kokonaisuus. Kolmantena valintana haastateltavilla vaihtelivat tallennus, DVD-soitin ja WWW-selain. Asetusvalikon paikka jäi siis päävalikon loppupäähän.

Haastateltavilta kysyttiin myös, mihin järjestykseen he haluaisivat asetusvalikon kohdat, joita olisivat käyttöasetukset (ääni ja tekstitys), asennus (automaattihaku ja manuaalihaku, kanavien muokkaus, kanavien poisto listasta), sekä järjestelmätiedot eli ohjelmistopäivitykset. Eli mitkä näistä he kokevat niin tärkeiksi, että niiden pitäisi olla valikon alkupäässä. Haastateltavat olivat enimmäkseen sitä mieltä, että käyttöasetusten olisi hyvä olla ensimmäisenä, mutta että järjestyksellä ei ole kuitenkaan suurta merkitystä. Jos valikossa on hyvä puhepalaute, tarvittavat kohdat löytyvät valikosta järjestyksestä riippumatta, mutta

jos puhepalautetta ei ole, eivät sokeat ja vaikeasti heikkonäköiset joka tapauksessa pysty käyttämään valikoita. Yksi haastateltavista esitti ehdotuksena, että jos digiboksi on myös näkövammaisille suunnattu, niin ensimmäisenä asetusvalikon kohtana voisi olla esimerkiksi ”kytke päälle puhetuki ja jyrkkäkontrastiset värit”.

Haastatellut heikkonäköiset olivat sitä mieltä, että fontilla ja värityksellä voidaan vaikuttaa valikoiden käytettävyyteen. Etenkin suuret kontrastit voivat vaikuttaa ratkaisevasti siihen, saako heikkonäköinen käyttäjä selvää käyttöliittymän teksteistä.

6.1.5 Yhteenveto haastattelun tuloksista

Haastattelun perusteella myös näkövammaiset seuraavat televisiota päivittäin. Tämän vuoksi näkövammaiset ovat selkeä käyttäjäryhmä digitelevision palveluille, joten heitäkään ei sovi unohtaa laitteita ja palveluita kehitettäessä. Äänitekstityspalvelu helpottaa jos-sain määrin ainakin joidenkin näkövammaisten television katselua.

Kaukosäätimien kanssa näkövammaisilla tulee monesti ongelmia, mutta lähinnä sen takia, että kaukosäätimiä on käytössä useita. Hyvän opastuksen avulla oman kaukosäätimen tärkeimpiä toimintoja oppii todennäköisesti käyttämään sujuvasti näkövammasta huolimatta, ja erityisesti näppäinten selkeä ryhmittely helpottaa tätä tehtävää. Suuri osa kaukosäätimien näppäimistä vaikuttaa kuitenkin olevan turhia näkövammaisille, kuten myös oman kokemukseni mukaan monille näkevillekin käyttäjille. Laitteiden etupaneelit ovat myös tärkeitä ohjausvälineitä joillekin käyttäjille, joten etupaneelia ei kannata ainakaan kevyin perustein jättää pois laitteesta.

Digiboksin ominaisuuksista tärkeimpiä haastatelluille näkövammaisille ovat digi-TV, tallennusmahdollisuus ja DVD-soitin. Muutkaan ominaisuudet, lukuun ottamatta WWW-selainta, eivät ole heille täysin merkityksellisiä. Suurennettava tekstitys ja puhepalaute ovat haastattelun perusteella näkövammaisille erittäin tervetulleita ja tärkeitä ominaisuuksia, mutta puheohjausta ainakin haastatellut näkövammaiset epäilivät epäkäytännölliseksi ominaisuudeksi.

Haastattelun perusteella hyvä vaihtoehto digiboksin päävalikon kohtien järjestykseksi olisi seuraavanlainen: ensimmäisenä kohtana olisi tv, toisena radio. Näiden jälkeen tulisivat jonkinlaisessa järjestyksessä tallennus, DVD-soitin ja WWW-selain. Viimeisenä kohtana olisi asetukset. Haastattelun kysymyksissä ei ollut kuitenkaan mukana kaikkia mahdollisia digiboksin ominaisuuksia, joten käytettävyydeltään hyvä valikon järjestys voi vaihdella paljonkin sen mukaan, mitä ominaisuuksia digiboksiin on toteutettu, ja mitä mahdollisesti jätetty pois.

Asetusvalikon kohtien järjestykseen haastateltavat eivät erityisemmin ottaneet kantaa, muuten kuin, että käyttöasetukset olisi hyvä olla valikon alkupäässä. Haastateltujen mukaan valikoita eivät kuitenkaan ainakaan sokeat henkilöt pysty käyttämään, ellei niistä saa puhepalautetta, eli puhepalaute oli erityisen toivottu valikoiden ominaisuus. Haastattelun perusteella myös valikoiden värien suuri kontrastiero parantaa käytettävyyttä heikkonäköisten kannalta.

6.2 Käyttöliittymän suunnitellut ominaisuudet

Aistiboxin suunnitellulla käyttöliittymällä voi tehdä liitteessä 1 lueteltuja asioita. Prototyyppi sekä tämän tutkielman suunnitelmat ovat kuitenkin pelkistettyjä versioita alkuperäisistä suunnitelmista. Seuraavat listan kohdat ovat mukana myös prototyypissä sekä tässä tutkielmassa esitetyissä suunnitelmissa:

- Katsoa digitaalisia televisiokanavia sekä kuunnella digitaalisia radiokanavia
- Selata ohjelmatietoja ohjelmaoppaasta
- Nauhoittaa ja tallentaa TV-ohjelmia sekä ajatuksen avulla että suosikkiohjelman valinnan avulla
- Toistaa tallennettuja ohjelmia
- Katsoa kuvia
- Kuunnella äänitiedostoja
- Selata muita verkkoon kytkettyjä laitteita
- Katsoa Internet-TV-ohjelmia ja kuunnella Internet-radiokanavia.

Aistiboxiin on integroitu puhesyntetisaattori, joka päälle kytkettynä puhuu valikoiden tekstit, ohjelmatiedot ja tekstitiedostot. Tekstin kokoa voi myös vaihtaa ja alaotsikot ovat skaalautuvia. Myös fontin ja taustan värit ovat vaihdettavissa. Valikoiden kuvakkeet on pyritty suunnittelemaan niin, että ne on helppo tunnistaa. Suunniteltu on myös toimintoa, jossa käyttäjä voi antaa komentoja puheella mikrofonin välityksellä, mutta tätä toiminnallisuutta ei ole prototyyppiin vielä toteutettu. Myös viittomakielistä tulkkauksruutua on mahdollista käyttää ohjelmien katselemisen yhteydessä [Axel, 2006]. Tulkkauksruutua ei ole kuitenkaan käsitelty tässä tutkielmassa esitetyissä suunnitelmissa.

6.3 Käyttöliittymän toteutus

Aistiboxissa on dynaaminen HTML-pohjainen käyttöliittymä, jossa valikkorakenteen muuttaminen on sovelluksen kehittäjälle helppoa. Toistaiseksi Aistiboxi on ollut

esittelykäytössä, ja siinä on paljon ominaisuuksia, jotka eivät ole välttämättömiä lopullisessa versiossa. Esittelyprototyypin valikkorakenne ei siis myöskään ole lopullinen.

Tutkielmaani varten minulla oli mahdollista saada käyttööni kuvaruutukaappauksia prototyypin käyttöliittymästä, sekä luettelo ominaisuuksista. Esittelen oman suunnitelmani käyttöliittymästä siten, että alan prototyypin kuvaruutukaappauksien pohjalta pohtimaan, millainen käyttöliittymän tulisi olla näkövammaiset ihmiset huomioon ottaen. Kohdassa 6.3.5 esittelen käyttöliittymän suunnittelun ohella syntyneen ehdotukseni Aistiboxin kaukosäätimestä. Prototyyppiin ei ole toteutettu puheohjausta, eikä puheohjaus ole ainkaan alkuvaiheen suunnitelmissa ollut mukana muuten kuin mainintana tästä mahdollisuudesta. Itsekin lähdän suunnitelmistani siitä, että käytössä on puhetulostus, mutta ei puheentunnistusta. Perustelen valintaani ja käsittelen Aistiboxin mahdollista puheohjausta kohdassa 6.3.6.

6.3.1 Laitteen käynnistäminen ja sammuttaminen

Aistiboxi-laitteen käynnistäminen voi tapahtua joko laitteen etupaneelin tai kaukosäätimen virtanäppäimestä. Koska televisio-ohjelmien katselu on yleisin toiminto, jota digiboksilla käytetään (haastattelun perusteella myös näkövammaisten keskuudessa), on jonkin televisiokanavan hyvä näkyä heti laitteen käynnistyttyä. Yleisesti käynnistymisen jälkeen näkyvä kanava on se, joka on ollut ennen sulkemista päällä, ja tämä onkin hyvä ratkaisu myös Aistiboxin tapauksessa. Itse televisiolaitteen täytyy luonnollisesti olla päällä, jotta televisiokanavia pystyy katsomaan tai Aistiboxin valikoita selaamaan. Valikoihin tulee päästä kaukosäätimessä olevasta näppäimestä, ja samanlainen näppäin olisi hyvä olla myös laitteen etupaneelissa.

Kun Aistiboxi käynnistetään ja sammutetaan, tulisi kuulua jonkinlainen palauteääni, jotta myös näkövammaiset tietäisivät käskyn menneen perille. Useimmissa digibokseissa palautteena käynnistymisestä on ainoastaan valon syttyminen laitteeseen sekä televisiosta mahdollisesti kuuluva ääni, eikä näistä ole apua vaikeasti näkövammaiselle, jos television äänet on jostain syystä kytketty pois päältä. Palauteääni voisi joko olla lyhyt piippaus, tai jonkinlainen hieman pidempään kestävä käynnistysääni. Yksi vaihtoehto palauteääneksi voisi olla myös syntetisaattorin sanoma lause, esimerkiksi ”Odota hetki, laite käynnistyy” ja ”Laite sammuu”. Jos syntetisaattori puhuisi tässä vaiheessa, voitaisiin Schmidtin ja Wlodkowskin [2003] suosituksen mukaisesti samalla myös mainita, miten puhetuki saadaan pois päältä, esimerkiksi ”Odota hetki, laite käynnistyy. Puhetuen saat kytkettyä pois päältä kaukosäätimen keskellä olevasta pitkästä näppäimestä.”

Vaikeasti näkövammaiset joutuvat haastattelun perusteella yleensä turvautumaan näkevien läheistensä apuun esimerkiksi uuden kaukosäätimen käytön opettelussa. Itsenäistä käyttöä alusta alkaen tukisi kaukosäätimessä oleva ohjenäppäin, jota painamalla puhesyntetisaattori lukisi näytölle ilmestyvät ohjeet kaukosäätimen näppäimien paikoista ja toiminnallisuudesta.

Ohjenäppäimen olemassaolo olisi hyvä mainita laitteen käynnistyksen yhteydessä. Maininta voisi olla esimerkiksi seuraavanlainen: ”Ohjeita kaukosäätimen käyttöön saat kaukosäätimen alarivin keskimmäisestä näppäimestä”. Tällöin laitteen käynnistykseen liitettävä puhetuloste voisi olla kokonaisuudessaan esimerkin 9 mukainen. Toisessa lauseessa ”pois päältä” ja ”takaisin päälle” voisivat vaihdella sen mukaan, onko puhetuki päällä vai ei. Käyttäjälle on tärkeää jollain tavalla ilmoittaa se, onko puhetuki päällä vai ei, sillä käynnistyksen yhteydessä kuultava puhetuloste voisi muuten erehdyttää käyttäjää luulemaan, että puhetuki on päällä, vaikka se ei olisikaan. Lisäksi puhetukinäppäintä painaessa tulisi kuulua jonkinlainen ilmoitus puhetuen tilasta, esimerkiksi ”Puhetuki kytketty päälle” tai ”Puhetuki kytketty pois päältä”.

”Odota hetki, laite käynnistyy. Puhetuen saat kytkettyä pois päältä/takaisin päälle kaukosäätimen keskellä olevasta pitkästä näppäimestä. Ohjeita kaukosäätimen käyttöön saat kaukosäätimen alarivin keskimmäisestä näppäimestä”.

Esimerkki 9. Käynnistyksen yhteydessä annettavat ohjeet.

Asetus-valikossa tulisi myös olla kohta, jossa käyttäjä voisi valita esimerkin 9 kaltaisen tulosteen tilalle myös yksinkertaisen käynnistysäänen, sillä joka käynnistyksen yhteydessä kuultavat ohjeet voivat käydä käyttäjälle rasittaviksi.

6.3.2 Valikot ja niissä navigointi yleisesti

Aistiboxin valikkoon pääsee kaukosäätimen Menu-näppäimestä. Television äänen olisi hyvä mykistyä automaattisesti valikkoon mennessä puheominaisuuden ollessa päällä ja mennä takaisin päälle valikosta poistuttua. Tämä todennäköisesti helpottaisi valikoiden selaamista, sillä jos kaksi puheääntä on äänessä samaan aikaan, voi olla erittäin vaikeaa keskittyä kumpaankaan kunnolla.

Aistiboxin prototyypissä valikkonäytön yläreunassa lukee valikon nimi vaalealla pohjalla. Muu näyttö on tummemman värinen, ja valikon nimen alapuolella ovat valikon valintavaihtoehdot allekkain. Valintavaihtoehtojen vasemmalla puolella ovat niitä kuvaavat kuvakkeet, joiden käytön saa myös kytkettyä pois päältä näytön asetuksista. Näytön

vasemmassa alareunassa on teksti Exit, OK, ja oikeassa alareunassa kellonaika. Teksti ”Exit” on ilmeisesti unohtunut suomentaa prototyypin suomenkielisessä käyttöliittymässä. Suomenkieliseksi vastineeksi sopisi teksti ”Poistu”, ja käytänkin jatkotarkasteluissani tätä suomennettua versiota.



Kuva 1. Aistiboxin prototyypin päähakemisto

Fokus siirtyy kohdasta toiseen kaukosäätimen ja mahdollisesti myös etupaneelin pystysuuntaisilla nuolinäppäimillä. Fokuksen alla olevan kohdan valinta onnistuu ”OK”-näppäimellä eli valintanäppäimellä, jolloin siirrytään kyseiseen alavalikkoon. Jos valikon kohta sisältää valintalistan, voidaan valittua vaihtoehtoa vaihtaa vaakasuuntaisilla nuolinäppäimillä.

Kun laitteen puheominaisuus on kytketty päälle, lukee puhesyntetisaattori jokaisen valikon kohdan niihin näppäimillä siirryttäessä. Jos käyttäjä ei halua kuunnella kaikkia valikon kohtia kokonaan, voi tämä yksinkertaisesti liikkua valikossa sen verran nopeammin, ettei puhesyntetisaattori ehdi niitä sanoa. Ennen ensimmäistä valikon kohtaa tulisi puhesyntetisaattorin sanoa myös valikon nimi, jotta näkövammaisenkin käyttäjä saisi tiedon siitä, missä valikon kohdassa on navigoimassa. Jos käyttäjä pysähtyy johonkin valikon kohtaan, tulee puhesyntetisaattorin toistaa kyseinen kohta pienen viiveen (esimerkiksi 3 sekuntia) jälkeen. Useampia kertoja ei mitään kohtaa kannata kuitenkaan automaattisesti toistaa, sillä silloin tällöin käyttäjä saattaa tarkoituksella pitää navigoinnissa taukoa, ja tällaisissa tilanteissa useat toistot aiheuttaisivat melko todennäköisesti ärsyyntymistä käyttäjässä.

Puheominaisuuden olisi hyvä olla tehdasasetuksissa kytkettynä päälle, jotta näkövammaisilta onnistuisi laitteen käyttöön ottaminen ilman apuvoimia. Näkeviltä käyttäjiltä puheominaisuuden kytkeminen pois päältä onnistuu todennäköisesti paljon helpommin kuin näkövammaisilta käyttäjiltä valikoiden selaaminen ilman puhetukea. Puheominaisuuden kytkemisen päälle ja pois on tärkeää onnistua myös kaukosäätimen näppäimestä: jos näke-

vän käyttäjän jäljiltä puhetuki jää pois kytketyksi, ja seuraavaksi laitetta käyttää sokea käyttäjä, on tämän vaikea alkaa etsimään valintaa, vaikka se sijaitisi valikon ylimmällä tasolla.

Prototyypissä fokuksen alla olevan kohdan tausta on aina tummemman sininen kuin näytön muu tausta, ja sen oikealla puolella näkyy teksti ”OK”. Nuolilla alaspäin mentäessä valikon viimeisen kohdan jälkeen fokus siirtyy kohtaan Poistu, OK, jonka valitseminen vie edelliselle valikon tasolle ja päähakemistossa oltaessa kokonaan pois valikosta. Tämän kohdan jälkeen nuolinäppäimellä alaspäin mentäessä fokuksen olisi hyvä siirtyä takaisin valikon alkuun, missä mainitaan myös valikon nimi. Tällöin näkövammaisinkin käyttäjä tietäisi koko ajan, missä kohtaa valikossa mennään, eikä tarvittaisi ylimääräistä tulostetta tai äänimerkkiä siitä, että kyseistä näppäintä painamalla ei tässä valikon kohdassa pääse liikkumaan.

Valikon kohtien oikealla puolella oleva ”OK” -teksti vaikuttaa turhalta, kun kuitenkin fokuksen alla oleva valikon kohta osoitetaan tummemmalla taustavärillä. Teksti voi aiheuttaa jopa hämmennystä käyttäjässä, sillä voi olla epäselvää, pitääkö valinnan tehdä vielä siirtyä ”OK” -tekstin kohdalle nuolinäppäimien avulla. Samoin kuulostaisi tökeröltä, jos puhesyntetisaattori sanoisi valikon tekstin luettuaan vielä, että ”OK”, eikä sekään ole hyvä vaihtoehto, että puhesyntetisaattori jättäisi lukematta jonkin näytössä näkyvän asian. ”OK” -tekstin jättäminen pois käyttöliittymästä on siis mielestäni perusteltua. Sama koskee ”OK” -tekstiä ”Poistu” -tekstin vieressä. Kuvassa 2 on esitettyä valikkonäytön sisältö tämän muutoksen jälkeen.

| Valikon nimi | |
|------------------------|--------------|
| Valikon kohta 1 | |
| Valikon kohta 2 | |
| Valikon kohta 3 | |
| Valikon kohta 4 | |
| Valikon kohta 5 | |
| Poistu | tt:mm |

Kuva 2. Valikkonäytön sisältö muutosten jälkeen

Fokuksen siirtymistavaksi valikoissa on myös erilaisia vaihtoehtoja. Esimerkiksi puheominaisuuden ollessa päällä voisi syntetisaattori lukea automaattisesti kaikki valikon kohdat vuorotellen, kunnes jokin kohdista valittaisiin kaukosäätimen valintanäppäimellä. Tässä vaihtoehdossa on kuitenkin se huono puoli, että valintanäppäintä pitäisi painaa juuri

oikealla hetkellä, mikä voi olla hankalaa etenkin näkövammaisille vanhuksille. Puhesyntetisaattorin pitämä tauko valikon kohtien välillä ei taas voi olla kovin pitkä, sillä taukojen ollessa pitkiä valikon kuunteleminen taukoineen päivineen voisi käydä tuskastuttavaksi. Valikoissa siirtyminen nuolinäppäinten avulla vaikuttaa käytettävyydeltään paremmalta vaihtoehdolta myös näkövammaiset huomioon ottaen.

Joidenkin käyttöliittymän toimintojen yhteydessä on hyvä ilmoittaa käyttäjälle, että toimenpide on onnistunut tai kysyä käyttäjältä vahvistusta toimenpiteelle. Nämä ilmoitukset ja varmennukset voisivat tapahtua näytön keskelle aukeavassa ikkunassa. Puhetuen ollessa päällä teksti luettaisiin luonnollisesti puhesyntetisaattorilla. Näkövammaisten kannalta erilliset graafiset näppäimet käyttöliittymässä olisivat erittäin huonoja vaihtoehtoja valintojen tekemiseen, sillä ne poikkeaisivat käyttöliittymän muusta linjasta. Käyttäjälle voitaisiinkin tarvittaessa ilmoittaa, mitä kaukosäätimen näppäintä tämän tulee painaa minkäkin vaihtoehdon valitessaan.

Kuvassa 1 näkyvät myös päähakemiston kohtien edessä olevat kuvakkeet. Kuvakkeet eivät näin toteutettuna tuo helppokäyttöisyyden kannalta erityistä lisäarvoa käyttöliittymään. Etenkin näkövammaisille ainakin prototyyppiin toteutetut kuvakkeet ovat turhia, sillä tekstin korkeuteen rajoitettuna ne ovat niin pienen kokoisia, että niiden yksityiskohtia ei huonolla näkökyvyllä voi erottaa. Toki tilanne paranee hieman, jos tekstikokoa suurennettaessa kuvakkeetkin suurenevät. Kuitenkin vaikuttaa siltä, että vaikka tekstikokoa suurennetaan, nojaavat kuvakkeet niin paljon yksityiskohtiin, että heikkonäköiselle kuvien merkitys jää helposti epäselväksi suuremmassakin koossa. Näiden seikkojen takia on hyvä, että näytön asetuksista kuvakkeiden käytön saa kytkettyä pois päältä. Heikkonäköisten kannalta mahdollisimman pelkistetty käyttöliittymä on yleensäkin paras vaihtoehto, sillä silloin rajoitettu näkökyky voidaan keskittää kaikkein olennaisimpiin asioihin. Normaalisti näkeville käyttäjille hyvin suunnitelluista ja toteutetuista kuvakkeista saattaa olla jossakin määrin apua, vaikkakin tärkeää olisi pyrkiä siihen, että valikon tekstit itsessään olisivat niin selkeitä, ettei kuvista tarvittaisi lisäapua niiden ymmärtämiseen. Eri asia olisi sellainen käyttöliittymä, jossa navigointi perustuisi kuviin, ja tekstit olisivat toissijaisena apuna navigoinnissa.

Aistiboxin prototyypin käyttöliittymässä käytetyt värit vaikuttavat miellyttäviltä tavallisesti näkevien käyttäjien näkökulmasta. Heikkonäköisiä käyttäjiä ajatellen valikon tekstien ja sinisen taustan välinen kontrasti saattaa kuitenkin olla liian pieni, joten hieman tummempi taustaväri olisi mahdollisesti parempi vaihtoehto. Myös fontin olisi hyvä olla hieman normaalia suurempi jo oletuksena. Tämän lisäksi käyttäjällä olisi hyvä olla myös mahdollisuus valita käyttöliittymään mahdollisimman jyrkkäkontrastiset värit. Tässä vaihtoehdossa voitaisiin poistaa heikkonäköisiä hämäämästä kaikki ylimääräiset taustan

sävyerot, jolloin tausta olisi yksivärinen ja taustan ja tekstin välillä olisi suuri kontrastiero. Haastattelussa esille tullut ehdotus valikon kohdasta ”kytke päälle puhetuki ja jyrkkäkontrastiset värit” on periaatteessa varteen otettava vaihtoehto. Nämä kaksi toimintoa on kuitenkin hyvä toteuttaa omina valintoinaan, sillä kaikki eivät välttämättä halua käyttää niitä molempia. Jos puhetuki on jo laitetta käyttöön otettaessa kytkettynä päälle, ja sille on myös kaukosäätimessä oma näppäimensä, riittää, jos puhetukeen ja väreihin liittyvät valinnat ovat muutaman valinnan päässä päähakemistosta. Puhetuen avulla värivalinnatkin löytyvät todennäköisesti helposti loogiselta paikaltaan alavalikoista, eikä niitä tarvitse sijoittaa päähakemistoon.

Puhesynteesin sijaan ihmisäänen lukemien valikoiden käyttäminen olisi myös periaatteessa mahdollista. Puhesynteesin käyttäminen on kuitenkin melko selkeä valinta tässä tapauksessa, sillä myös ohjelmaopas luetaan puhesyntetisaattorilla puhetuen ollessa päällä. Ohjelmaopas muuttuu koko ajan, joten sen lukemiseen puhesynteesi on ainoa vaihtoehto. Jos valikot olisivat ihmisäänen lukemat ja käyttäjä liikkuisi valikoiden ja ohjelmaoppaan välillä, vaihtelisi puheääni puhesynteesin ja ihmisäänen lukemien valikoiden välillä. Tämä on sellainen vaihtoehto, joka saattaisi olla käyttäjälle sekavan tuntuinen. Gongin ja Lain [2001] mukaan käyttäjät voivat myös paremmin keskittyä tehtäväänsä, kun käyttöliittymä sisältää vain yhden tyyppistä puhetta. Heidän toteuttamassaan tutkimuksessa käyttäjät suoriutuivat huonommin tehtävästään käyttöliittymällä, joka sisälsi sekä äänitettyjä tulosteita että puhesynteesillä tuotettuja tulosteita, verrattuna pelkkää puhesynteesiä sisältävään käyttöliittymään. Toisaalta molempia puhetyyppejä sisältävän käyttöliittymän käyttäjät pitivät käyttöä helpompana kuin pelkkää puhesynteesiä sisältävän käyttöliittymän käyttäjät.

Miellyttävä äänitetty ihmisääni saa siis Gongin ja Lain [2001] mukaan käyttäjän tuntemaan olonsa mukavammaksi kuin puhesynteesillä tuotettu puhe. Tämän takia käyttöliittymää saattaa olla mukavampi käyttää, jos edes osa puheesta on äänitettyä, vaikkakin tehtävän suoritus saattaa kärsiä puhetyylin vaihtelun vuoksi. Myös Hakulinen et al. [2003] pitävät eri puhetyyppien yhdistämistä hyvänä vaihtoehtona, jos eri puhetyyppejä käytetään selkeästi erillisissä käyttöliittymän osissa ja jos äänitettävien tulosteiden määrä on kohtuullinen. Tällä perusteella Aistiboxin käyttöliittymässä voisi harkita myös äänitettyjen tulosteiden käyttämistä ainakin ohjeissa, sillä ne ovat selkeästi erillinen käyttöliittymän osa eikä niitä varten tarvita useita erilaisia tulosteita.

6.3.3 Pähakemisto

Haastattelun perusteella valikon kahdeksi ensimmäiseksi kohdaksi parhaat vaihtoehdot olisivat Televisio ja Radio, ja kolmeksi seuraavaksi kohdaksi Tallennus, DVD-soitin ja

WWW-selain. Valikon viimeiseksi kohdaksi sopisi Asetukset. Haastattelussa ei kuitenkaan mainittu DVD:n lisäksi muita mahdollisia Aistiboxilla käsiteltäviä medioita, joita ovat kuvat, videot, äänitiedostot ja CD-ROM-levyt. Jos nämä kaikki olisivat mukana laitteessa, ne voitaisiin loogisesti yhdistää DVD-soittimen kanssa yhdeksi ”Media”-kokonaisuudeksi. Myöskään Internet TV vaihtoehtoa ei haastattelussa mainittu. Jos sellainen toteutettaisiin Aistiboxiin, myös tämän vaihtoehdon olisi hyvä olla heti päävalikossa näkyvässä, kuten se prototyypissä onkin.

Haastattelussa käsiteltiin Tallennus-vaihtoehtoa yhtenä päävalikon kohtana. Televisio-ohjelmien tallennus tapahtuu kuitenkin helpoiten ohjelmaoppaan yhteydessä eli niin, että ohjelmaopasta selatessa valitaan ohjelmat, jotka halutaan tallentaa. Tämän lisäksi tarvitaan erillinen valikon kohta, josta päästään tiivistetysti katsomaan tallennettavaksi valittuja ohjelmia. Ohjelmaopas ja Tallennettavat ohjelmat sopisivat loogisimmin Televisio-valikkoon, johon puolestaan tulee päästä päävalikosta. Prototyypin käyttöliittymään onkin toteutettu juuri tällainen järjestys. Käsitellen Televisio-valikkoa sekä muita alavalikoita tarkemmin luvussa 6.3.4.

Prototyypin valikoihin ei ole sisällytetty WWW-selainta, ja haastattelun perusteella ei olisi suuri puute, jos WWW-selain jäisi laitteesta kokonaan pois. WWW-selaimen voi siis jättää ainakin tässä vaiheessa valikosta pois, ja toisaalta sen lisääminen ei suuremmin vaikuttaisi tässä esitettyyn valikkorakenteeseen, sillä se toisi ainoastaan yhden valintavaihtoehdon lisää päähakemistoon. Myös DVD-soitin sekä CD-ROM levyjen käyttäminen ovat tällaisia vaihtoehtoisia ominaisuuksia, jotka olen jättänyt myös oman käsittelyni ulkopuolelle. Yleisesti uusien toimintojen lisäämisessä laitteeseen tulee kuitenkin ottaa huomioon Rinnetmäen et al. [2004] suositus siitä, että televisioympäristössä yhdellä valikkoruudulla on hyvä käyttää korkeintaan 5-7 valintaa kerralla. Jos laitteeseen lisätään useita uusia toimintoja, tulee valikoiden ryhmittely järjestää niin, että tähän tavoitteeseen päästään.

Prototyypin toteutetussa päähakemistossa (kuva 1) on viisi valintavaihtoehtoa: Televisio, Radio, Internet TV, Media, ja Asetukset. Kuten edellä käy ilmi, tämä on useat seikat huomioon ottaen hyvä kokoonpano ja järjestys.

6.3.4 Alavalikot

Käyn hierarkiassa päähakemiston alla olevat valikot läpi yleisluontoisesti, sillä niiden tarkka läpikäyminen ja suunnittelu vaatisi syvempää asiantuntemusta digibokseista ja niiden toiminnasta. En myöskään saanut tutkielmaani varten tarkkaa kuvausta Aistiboxiin suunnitellusta toiminnallisuudesta, enkä kuvia kaikista valikoista, joten valikoiden ja

esimerkiksi ohjelmaoppaan täydellinen suunnittelu ei onnistu tämän laajuisen tutkielman puitteissa.

Televisio-valikko sisältää prototyypissä kohdat Ohjelmaopas, Kanavat ja Äänitettävät ohjelmat. Äänitettävät ohjelmat tarkoittaa käsittääkseni listaa tallennettavista ohjelmista, eikä äänitys-termi tarkoita samaa asiaa, joten Tallennettavat ohjelmat on parempi nimi tälle valikon kohdalle. Kanavat on todennäköisesti harvemmin käytetty valikon kohta kuin Tallennettavat ohjelmat, joten sille parempi paikka on valikon viimeisenä.

| | |
|--|--------------|
| Televisio | |
| Ohjelmaopas Tallennettavat ohjelmat Kanavat | |
| Poistu | 12:00 |

Kuva 3. Televisio-valikon sisältö

Ohjelmaoppaassa voi selata ohjelmatietoja ja valita ohjelmia tallennettaviksi. Puhesyntetisaattori lukee myös ohjelmatiedot ja niitä selataan samalla tavalla kaukosäätimen tai etupaneelin nuolinäppäimillä kuin muitakin valikon osia. Tietyn ohjelman valinta tallennettavaksi olisi hyvä tapahtua yksinkertaisesti painamalla valintanäppäintä kyseisen ohjelman nimen kohdalla. Ohjelman tarkempiin tietoihin voisi siinä tapauksessa päästä erillisellä ”Info”-näppäimellä, joka on yleinen digiboksien kaukosäätimissä, ja käytössä myös televisio-ohjelmia katsottaessa. Käyttäjän valittua ohjelman tallennettavaksi tulee häneltä vielä pyytää varmistus, jotta tallennus ei tapahdu vahingossa, esimerkiksi ”Jos haluat, että Ohjelma xx tallennetaan pp.kk.vvvv kello xx:xx, paina uudelleen OK-näppäintä. Peruuttaaksesi toiminnon paina vasemmanpuoleista nuolinäppäintä”.

Kanavat-valikko sisältää listan televisiokanavista ja listasta näkyy, mitkä kanavat näkyvät kyseisellä digiboksilla. Tallennettavat ohjelmat –valikossa on lista tallennettavaksi valituista ohjelmista. Listan alussa voisi mahdollisesti olla myös vaihtoehto ”Valitse uusi ohjelma tallennettavaksi”, joka veisi ohjelmaoppaaseen. Ohjelmat olisi hyvä olla lueteltu tallennusjärjestyksessä uusimmasta alkaen, ja ohjelman nimen oikealla puolella tulisi näkyä tieto siitä, onko ohjelma tallennettu, onko tallennus kesken, vai odottaako se tallennusta. Tämä voitaisiin ilmoittaa esimerkiksi ”Valmis”, ”Tallentuu” ja ”Odottaa” -maininnoilla.

Tallennus tulee olla mahdollista myös peruuttaa tai poistaa, ja ohjelmia tulee voida valita suosikeiksi eli jatkossakin tallennettaviksi. Käyttäjän tulee myös voida valita tallennettu ohjelma katsottavaksi suoraan tästä listasta, vaikka ohjelmien toistamisen voisi tehdä myös Media-valikon kautta. Nämä valintavaihtoehdot eivät voi olla muiden valintalistojen tapaan selattavissa näytön oikeassa reunassa jokaisen ohjelman kohdalla, sillä ne eivät ole tyypillisiä valintalistan vaihtoehtoja, vaan omia toimintojaan. Tämän vuoksi paras vaihtoehto on, että näille valinnoille tehdään vielä oma valikkonsa. Tällöin käyttäjän valitessa tietyn ohjelman aukeaa valikko, jonka otsikkona on ohjelman nimi ja valikon kohtina esimerkiksi ”Katsele”, ”Poista” ja ”Valitse suosikiksi”.

Ainakin joissain digibokseissa tallennettuja ohjelmia voi alkaa katselemaan jo ennen kuin tallennus on edennyt loppuun asti. Tällainen toiminta on varmasti hyödyllistä käyttäjille, sillä silloin ohjelmaa voi alkaa katsoa alusta vaikkapa 10 minuuttia oikean alkamisajan jälkeen. Tällöin käyttöliittymästä ei myöskään tarvitsisi poistaa Katsele-vaihtoehtoa missään vaiheessa. Ohjelmien katselun olisi hyvä onnistua myös suoraan kaukosäätimen oikeanpuoleista nuolta painamalla, joka toimisi tällöin myös ”Play”- ja ”Pause”-näppäiminä. Samalla ”OK”-näppäin toimisi ”Stop”-näppäimenä.

Jotta käyttäjä ei vahingossa poistaisi tallennettavaa tai jo tallennettua ohjelmaa listasta, olisi käyttäjältä hyvä pyytää jonkinlainen varmistus Poista-valinnan tekemisen jälkeen, esimerkiksi: ”Jos haluat peruuttaa ohjelman xx tallennuksen, paina uudelleen OK-näppäintä. Muussa tapauksessa paina vasemmanpuoleista nuolinäppäintä”. Kokonaan tallennetun ohjelman ollessa kyseessä varmistus voisi sisältää tekstin: ”Jos haluat poistaa ohjelman xx vain tallennettavat ohjelmat –listasta, paina oikeanpuoleista nuolinäppäintä. Jos haluat poistaa ohjelman kokonaan, paina uudestaan OK-näppäintä. Muussa tapauksessa paina vasemmanpuoleista nuolinäppäintä”. Poistamisen jälkeen tulee käyttäjälle antaa myös vahvistus siitä, että toimenpide on onnistunut, esimerkiksi ilmoituksella ”Ohjelma xx poistettu”.

Radio-valikossa voi valita jonkin digitaalisen radiokanavan kuunneltavaksi ja Internet TV –valikossa voi valita jonkin Internet TV –kanavan katsottavaksi. Media-valikko (kuva 4) puolestaan sisältää kohdat Tallennetut ohjelmat, Videot, Kuvat, Musiikki, ja Kotiverkko.

| | |
|--|--------------|
| Media | |
| Tallennetut ohjelmat Videot Kuvat Musiikki Kotiverkko | |
| Poistu | 12:00 |

Kuva 4. Media-valikon sisältö

Tallennetut ohjelmat –valikossa luetellaan tallennetut ohjelmat. Järjestys olisi hyvä olla aakkosellinen, sillä tästä valikosta käyttäjät mahdollisesti etsivät vanhempiakin tallennuksia: käyttäjän voi olla vaikea muistaa, missä järjestyksessä vanhempia tallennuksia on tehty, mutta tallennusten nimet sen sijaan on todennäköisesti helpompi muistaa. Käyttäjän valitessa jonkin ohjelman voisi Tallennettavat ohjelmat –valikon tapaan aueta uusi valikkonäyttö, jossa olisi vaihtoehdot ”Katsele” ja ”Poista”. Poistamisen varmistuksen tulee myös tapahtua samaan tapaan kuin Tallennettavat ohjelmat –valikossa. Sama pätee Videot-valikkoon, jossa luetellaan laitteeseen tallennetut videot ja Musiikki-valikkoon, jossa luetellaan laitteeseen tallennetut kappaleet. Tallennettavat ohjelmat-, Videot- ja Musiikki-valikoissa tulee myös voida käyttää kaukosäätimen nuolinäppäimiä ja OK-näppäintä samaan tapaan kuin Tallennettavat ohjelmat –valikossa. Kotiverkko-valikossa voi selata muita verkkoon kytkettyjä laitteita ja siirtää niistä tiedostoja Aistiboxin kiintolevylle sekä kiintolevyltä toiseen laitteeseen.

Asetukset-valikko sisältää prototyypissä kohdat Asennus, Näytön asetukset, Äänen asetukset, Valitse valikkojen kieli ja Palauta tehdasasetukset. Valikkojen kielen valitsemiseen on kuitenkin turha tehdä omaa valikkoansa, sillä valinnan voi tehdä myös valintalistassa suoraan Asetukset-valikossa, kuten kuvassa 5 näkyy. Tehdasasetusten palautus voi tapahtua prototyypin mukaisesti myös suoraan Asetukset-valikossa. Toimenpiteen tapahtuminen suoraan Asetukset-valikon kohdan valitsemalla voi kuitenkin olla käyttäjälle yllättävää, joten käyttäjän valittua tämän kohdan tulee tältä pyytää vielä varmistus asiasta, esimerkiksi ”Jos haluat palauttaa tehdasasetukset, paina uudelleen OK-näppäintä. Peruuttaaksesi toiminnon paina vasemmanpuoleista nuoli-näppäintä”.

| Asetukset | |
|--------------------------------|--------------|
| Asennus | |
| Näytön asetukset | |
| Äänen asetukset | |
| Valikkojen kieli | suomi |
| Palauta tehdasasetukset | |
| Poistu | 12:00 |

Kuva 5. Asetukset-valikon sisältö

Asennus-valikko sisältää laitteen asennukseen liittyviä valintavaihtoehtoja. Näytön asetukset –valikko (kuva 6) sisältää tekstikoon, kirjasinlajin, kuvakkeiden ja tekstityksen koon valinnat. Äänen asetuksissa puolestaan voidaan muuttaa erilaisia ääneen liittyviä asetuksia, ja tähän valikkoon sopisi myös käynnistys- ja sammumisäänen valinta.

| Näytön asetukset | |
|--------------------------|---------------|
| Tekstikoko | Koko1 |
| Kirjasinlaji | Arial |
| Kuvakkeet | Päällä |
| Tekstityksen koko | x1 |
| Poistu | 12:00 |

Kuva 6. Näytön asetukset –valikon sisältö

6.3.5 Kaukosäädin ja etupaneeli

Näkövammaisille suunnatussa kaukosäätimessä sekä etupaneelissa tärkeintä on, että näppäimet on selkeästi eroteltu toisistaan ja että ne ovat käytön kannalta loogisessa järjestyksessä. Näppäinten määrä on hyvä rajoittaa kohtuulliseksi niin, että vain tärkeimmille toiminnoille on omat näppäimensä. Mahdolliset lisänäppäimet voidaan kaukosäätimessä piilottaa kaukosäätimen kannen alle, mistä näkevät käyttäjät pystyvät niitä käyttämään. Myös näppäimien ja taustan väriä välinen kontrastiero tulee olla mahdollisimman suuri. Lisäksi kaikkiin näppäimiin olisi näkövammaisten käyttäjien kannalta tärkeää yhdistää äänimerkki, joka kuuluu puhutuen päällä ollessa, kun näppäintä painaessa ei tapahdu mitään.

Näppäimet, jotka edellisissä kohdissa suunnitellun käyttöliittymän perusteella olisi hyvä sisällyttää kaukosäätimeen, ovat virtanäppäin, äänen mykistys, numeronäppäimet, puhetuki päälle/pois, nuolinäppäimet ylös ja alas, valintanäppäin, ohje, info ja valikko (lyhyesti Menu). Ohje-näppäimestä aukeavia ohjeita ei ole suunniteltu tässä tutkielmassa, sillä ne ovat selkeästi oma kokonaisuutensa eikä niiden suunnitteleminen ollut mahdollista tutkielman tekoaikataulun puitteissa. Kuvassa 7 on esitettyä luonnos kaukosäätimestä. Luonnoksessa ei oteta kantaa muotoilullisiin seikkoihin, joilla voidaan vaikuttaa kaukosäätimen ergonomiaan sekä ulkonäköön.



Kuva 7. Kaukosäädin

Haastattelun perusteella etupaneelissa olisi hyvä olla seuraavat näppäimet: virtakytin vasemmassa reunassa ja valikkonäppäin, nuolinäppäimet sekä valintanäppäin oikeassa reu-

nassa. Nuolinäppäimet ja valintanäppäin voisivat olla samoin ryhmitellyt kuin kaukosäätimessä, jolloin ne olisivat helposti tunnistettavissa.

6.3.6 Käyttöliittymän ohjaaminen puheella

Haastatellut näkövammaiset suhtautuivat varauksellisesti digiboksin puheohjaukseen. He olivat sitä mieltä, että puheohjauksessa käytettävien komentojen muistaminen olisi vaikeampaa kuin kaukosäätimen näppäinten paikkojen muistaminen. Myös sitä pohdittiin, että televisiosta tuleva ääni saattaisi häiritä puheentunnistusta. Yleisesti ottaen haastateltavat tuntuivat pitävän juuri puhetulostuksen ja kaukosäätimen käytön yhdistelmää optimaalisena vaihtoehtona digiboksin ohjaamiseen.

Ibrahim ja Johansson [2002] puolestaan suhtautuvat myönteisesti puheohjaukseen TV-ympäristössä. Myös Berglund ja Johansson [2004] suosittelevat puheohjauksen mahdollistamista ja kaukosäätimen pitämistä vaihtoehtoisena vuorovaikutusvälineenä. He kuitenkin pitävät puheohjausta joissain tilanteissa ongelmallisena vaihtoehtona TV-ympäristön luonteen takia: käyttäjät saattavat puhua keskenään, ja toisaalta illalla televisiota katsova käyttäjä saattaa haluta muut asukkaat huomioidakseen olla kokonaan puhumatta.

Kun otetaan huomioon puheentunnistuksen tunnetut ongelmat sekä TV-ympäristön puheentunnistukseen tuomat haasteet, vaikuttaa puheohjauksen lisääminen Aistiboxiin epärealistiselta. Tulee myös ottaa huomioon, että erityisesti näkövammaisille suunnitelluista laitteista ei markkinoilla ole paljoa vaihtoehtoja. Puheohjaus monimutkaistaisi laitteen ohjelmistoa paljon, ja tämän myötä myös kustannukset ja samalla laitteen hinta todennäköisesti kasvaisivat. Tasavertaisuutta ei edistäisi se, että ainoa näkövammaisille suunnattu laite olisi selvästi kalliimpi kuin niin sanotut tavalliset digiboksit, ja korkea hinta luonnollisesti nostaisi kynnyksen laitteen hankkimiseen. Jos näkövammaiset pystyvät käyttämään digiboksia sujuvasti puhetulostuksen ja kaukosäätimen avulla, on turhaa lisätä mukaan puheohjausta vain sen itsensä vuoksi. Kuitenkin, jos digiboksin käyttämisessä ilman puheohjausta ilmenee ongelmia, kannattaa puheohjauksen toteutusmahdollisuuksia vähintäänkin tutkia.

6.3.7 Yhteenveto suunnitelluista toiminnoista ja valikoista

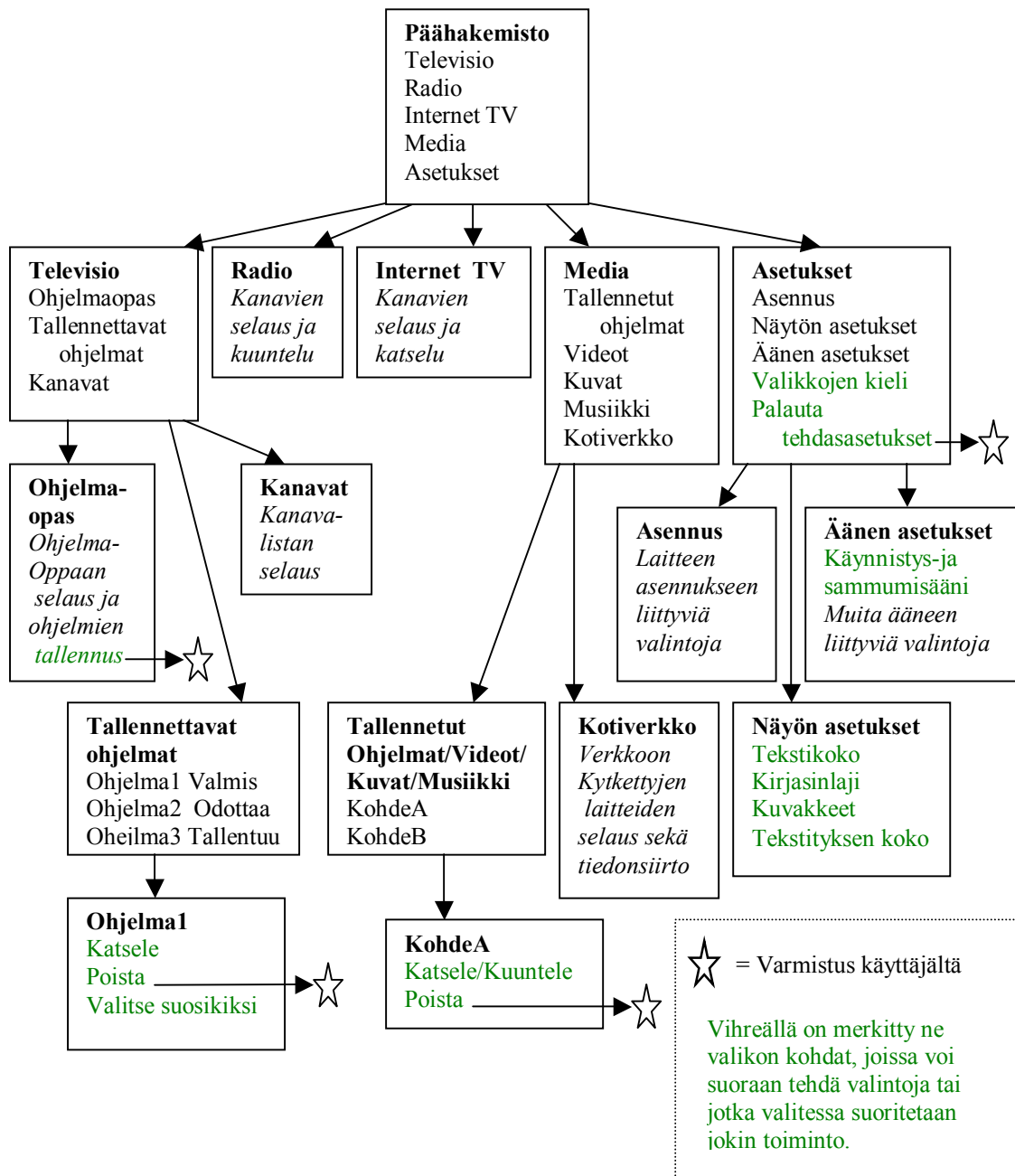
Tämän tutkielman kohdissa 6.3.1-6.3.5 on suunniteltu Aistiboxin prototyypin pohjalta näkövammaisille soveltuva käyttöliittymä ja kaukosäädin pääpiirteissään. Taulukossa 1 on vielä esitetty tiivistetysti kaukosäätimen näppäimiä painettaessa tapahtuvat toiminnot ja

kaaviossa 1 on kuvattuna suunniteltu valikkorakenne. Kaikissa taulukossa ja kaaviossa esitetyissä tilanteissa on oletuksena, että televisio on auki: muussa tapauksessa ei kuulu eikä näy mitään. Käyttäjän voidaan myös olettaa tietävän tämän asian.

| Näppäin | Missä tilanteessa käyttäjä painaa näppäintä | Mitä tapahtuu |
|---|--|---|
| Virta- kytkin | Laite pois päältä. | Laite käynnistyy. Kuuluu käynnistysääni, joka tehdasasetuksissa on puhetuloste (ks. esimerkki 9). Viimeksi valittuna ollut televisiokanava tulee näkyviin. |
| | Laite päällä | Kuuluu sammumisääni, joka tehdasasetuksissa on puhetuloste ”Laite sammuu”, ja sen jälkeen laite sammuu. |
| Äänen mykistys | Televisiokanava näkyvissä tai valikko auki ilman puhetukea | Television ääni mykistyy, toista kertaa painaessa ääni palaa takaisin. |
| | Valikko auki puhetuen ollessa päällä | Ei tapahdu mitään, sillä television ääni mykistyy automaattisesti valikkoon mennessä puhetuen ollessa päällä. |
| Numero- näppäimet | Televisiokanava näkyvissä | Televisiokanava vaihtuu. |
| | Valikko auki | Ei tapahdu mitään. |
| Puhetuki- näppäin | Puhetuki päällä | Puhetuki menee pois päältä. Kuuluu ilmoitus ”Puhetuki kytketty pois päältä” |
| | Puhetuki pois päältä | Puhetuki menee päälle. Kuuluu ilmoitus ”Puhetuki kytketty päälle”. |
| Vaaka- suuntaiset nuoli- näppäimet | Televisiokanava näkyvissä | Television ääni hiljenee tai voimistuu. Jos infoikkuna näytön alareunassa on auki, oikeanpuoleisella nuolella tulevat näkyviin seuraavan ohjelman tiedot (tästä tarkemmin taulukon kohdassa ”Infonäppäin”). |
| | Valikko auki | Jos fokuksen alla olevan valikon kohdan oikeassa reunassa on valintalista, valintalista ”rullaa” oikealle tai vasemmalle. Muussa tapauksessa ei tapahdu mitään. |
| | Valikko auki, video / tallennettu ohjelma/ musiikkikappale valittuna | Oikeanpuoleisesta nuolesta alkaa toistaa kyseistä mediaa. Toista kertaa painaessa pysäyttää toiston, eli toimii kuten ”Play/Pause”-näppäin. |
| Pysty- suuntaiset nuoli- näppäimet | Televisiokanava näkyvissä | Televisiokanava vaihtuu. |
| | Valikko auki | Fokus siirtyy valikossa ylös- tai alaspäin. |

| | | |
|-----------------|--|---|
| OK/stop-näppäin | Televisiokanava näkyvissä | Ei tapahdu mitään |
| | Valikko auki | Fokuksen alla oleva valikon kohta tulee valituksi, eli siirrytään seuraavalle valikon tasolle tai valikon kohdan mukaiseen toimintoon (katsele/poista tms.). Jos taso on viimeinen, ei tapahdu mitään. |
| | Valikko auki, videon / tallennetun ohjelman / musiikkikappaleen toisto käynnissä | Toisto pysähtyy ja fokus siirtyy valikkoon kyseisen videon / tallennetun ohjelman / musiikkikappaleen kohdalle. |
| Ohje-näppäin | Puhetuki päällä | Näyttöön ilmestyvät ohjeet omaan ikkunaan ja puhesyntetisaattori alkaa lukea niitä. Television ääni mykistyy. Toista kertaa painaessa ohjeikkuna sulkeutuu ja puhe loppuu, television ääni palaa kuuluviin. |
| | Puhetuki pois päältä | Näyttöön ilmestyvät ohjeet omaan ikkunaan. Toista kertaa painaessa ohjeikkuna sulkeutuu. |
| Info-näppäin | Televisiokanava näkyvissä | Kanavalla näkyvän ohjelman ohjelmatiedot tulevat näkyviin näytön alareunaan. Jos puhetuki on päällä, puhesyntetisaattori lukee tiedot ja television ääni mykistyy. Toista kertaa infonäppäintä painaessa infoikkunaan ilmestyvät ohjelman tarkemmat tiedot. Kolmatta kertaa painaessa infoikkuna sulkeutuu (ja puhe loppuu, television ääni palaa kuuluviin). |
| | Valikossa auki Ohjelmaopas, Tallennetut ohjelmat tai Tallennettavat ohjelmat | Valitun ohjelman tarkemmat tiedot tulevat näkyviin (ja puhutuen päällä ollessa puhesyntetisaattori lukee tiedot). Toista kertaa painaessa fokus siirtyy takaisin ohjelmalistaan. |
| Valikko-näppäin | Televisiokanava näkyvissä | Valikko aukeaa. Jos puhesyntetisaattori on päällä, se alkaa lukea valikon kohtia ja television ääni mykistyy. |
| | Valikko auki | Fokus siirtyy valikon edelliselle tasolle. Jos fokus on päähakemistossa, valikko sulkeutuu. (Samaan tapaan toimii myös ”Poistu”-valinta valikon alareunassa) |

Taulukko 1. Kaukosäätimen näppäimistä tapahtuvat toiminnot



Kaavio 1. Suunniteltu valikkorakenne pääpiirteissään.

6.4 Tulevaisuuden näkymiä

Tämän tutkielman tekoaikana selvisi, ettei Aistiboxi-hanke saa rahallista tukea Teknologian kehittämiskeskukselta (Tekes). Tämän vuoksi hankkeen tulevaisuudesta ei ole tarkkaa tietoa. Aistiboxille olisi kuitenkin olemassa tilausta etenkin näkövammaisten keskuudessa, ja hankkeen jatkaminen olisi tärkeää. Laitteiden suunnitteleminen ja toteuttaminen erityisryhmille ei todennäköisesti ole taloudellisesti kannattavaa, ja siksi valtion tulisikin tukea tällaisia hankkeita.

7 Yhteenveto

Tässä tutkielmassa selvitettiin, miten näkövammaiset käyttäjät tulee ottaa huomioon puhe-käyttöliittymän suunnittelussa. Tutkielman alussa luotiin katsaus nykyiseen puheteknologi-
aan ja kartoitettiin kirjallisuuden perusteella puhekäyttöliittymien suunnitteluperiaatteita
sekä sitä, millainen on käytettävyydeltään hyvä puhekäyttöliittymä. Tämän jälkeen
selvennettiin näkövammaisuuden käsitettä ja kirjallisuustutkimuksen perusteella selvitet-
tiin, mitä erityistarpeita näkövammaisilla on puhekäyttöliittymien käyttäjinä. Erityistä huo-
miota kiinnitettiin sellaisiin puhekäyttöliittymiin, jotka sisältävät sekä puhetta että visuaali-
sia elementtejä. Esimerkkinä puhekäyttöliittymän suunnittelusta näkövammaisille esiteltiin
Aistiboxi-hanke sekä hankkeen puitteissa toteutetun näkövammaisten haastattelun tulok-
set. Hankkeen aikana toteutetun digiboksin käyttöliittymän prototyypin pohjalta esitettiin
suunnitelma näkövammaiset huomioon ottavasta käyttöliittymästä.

Tutkielman teon alkuvaiheessa tarkoituksena oli arvioida Aistiboxin prototyypin käytettä-
vyyttä. Kuitenkin siinä vaiheessa, kun arviointi olisi ollut tämän tutkielman aikataulun
puolesta ajankohtaista, olikin prototyypin toteutus vielä kovin varhaisessa vaiheessa, eikä
alkutekijöissään olevan käyttöliittymän arvioiminen tuntunut mielekkäältä. Prototyyppiä ei
myöskään ollut kokonaisuudessaan saatavilla arviointia varten, vaan saatavilla oli ainoas-
taan suunniteltujen ominaisuuksien lista sekä joitakin kuvaruutukaappauksia. Näiden
seikkojen vuoksi tutkielmassa päädyttiin antamaan perusteellisesti suunniteltu ehdotus
siitä, millainen Aistiboxin käyttöliittymän tulisi olla, jotta se olisi hyvä käytettävyydeltään
myös näkövammaisille käyttäjille. Suunnittelun apuna käytettiin tutkielman teoriaosuutta
sekä hankkeen aikana toteutetussa kohderyhmän haastattelussa ilmenneitä seikkoja.

Puheteknologian käytössä on paljon hyviä puolia etenkin näkövammaisten ihmisten kan-
nalta, mutta siinä on myös lukuisia haasteita ja ongelmia. Lisäksi näkövammoja on monen
tasoisia ja monen laatuksia ja näkövammaisia on kaikissa ikäluokissa: näin ollen
käyttötottumuksia ja –tarpeita on todella monenlaisia. Käyttöliittymien suunnittelemine
näkövammaiset huomioon ottaen onkin monimutkainen tehtävä, jota täytettäessä on tär-
keää muistaa kohderyhmän kirjavuus ja suunnitelmien mahdollisimman laaja testaaminen
kohderyhmällä.

Haastattelun perusteella myös useimmat näkövammaiset seuraavat päivittäin televisiota,
joten heitäkään ei sovi unohtaa digitelevisioon liittyviä laitteita ja palveluita kehitettäessä.
Suurennettava tekstitys, suuret kontrastierot ja puhepalaute ovat haastattelun perusteella
näkövammaisille tärkeitä digiboksin käyttöliittymän ominaisuuksia, mutta puheohjausta
ainakin haastatellut näkövammaiset epäilivät epäkäytännölliseksi ominaisuudeksi. He oli-

vat sitä mieltä, että selkeästi ryhmiteltyä kaukosäädintä on näkövammaisenkin helpompi oppia käyttämään kuin puhekomentoja. Tätä väitettä olisikin mielenkiintoista testata käytännössä.

Haastattelun perusteella useimmat näkövammaisten ongelmat kaukosäätimen käytössä johtuvat siitä, että kaukosäätimiä on useita ja ne ovat erilaisia. Digiboksin tai television valikoita eivät ainakaan sokeat henkilöt pysty käyttämään kaukosäätimelläkään, ellei niistä saa puhepalautetta, eli puhepalaute oli haastateltavien keskuudessa erityisen toivottu valikoiden ominaisuus.

Aistiboxin kaltaiselle laitteelle olisi tilausta erityisryhmien, kuten näkövammaisten, keskuudessa. Tässä tutkielmassa on esitetty suunnitelma näkövammaiset huomioon ottavasta käyttöliittymästä, ja sitä voitaisiinkin hyödyntää mahdollisissa jatkohankkeissa. Myös puheohjauksen mahdollisuutta laitteessa kannattaa tutkia tarkemmin, vaikka se tässä tutkielmassa on jäänyt sivuosaan. Laitteiden suunnitteleminen ja toteuttaminen erityisryhmille ei todennäköisesti ole taloudellisesti kannattavaa. Valtion tulisikin tukea tällaisia hankkeita, jotta myös erityisryhmät pystyisivät käyttämään palveluja tasavertaisesti valtaväestön kanssa.

Viiteluettelo

[Aho et al., 2005] Lari Aho, Mikko Lehmuskoski, Helena Pekkarinen, Matti Sarmela, Leila Oravisto, Jussi Keskinarkaus, Petteri Huvio, Petri Ihalainen, Päivi Aarreniemi-Jokipelto, Tomi Nummi, Karin Mariola, Petri Seljavaara, Simon Staffans, Tommy Strandvall, ArviD-palvelunkehityshankkeet - Yhteenvetoraportti. *ArviD-julkaisu* 08/2005. Liikenne- ja viestintäministeriö, 2005.

[Axel, 2006] Axel Technologies, eA Box – easy access digital TV receiver – esite. Axel Technologies, 2006.

[Berglund and Johansson, 2004] Aseel Berglund and Pontus Johansson, Using speech and dialogue for interactive TV navigation. In *Universal Access in the Information Society*, Volume 3, Numbers 3-4, 224 - 238. Springer-Verlag Heidelberg.

[Bergman and Johnson, 1995] Eric Bergman and Earl Johnson, Towards accessible human-computer interaction. In: Jakob Nielsen (Ed.), *Advances in human-computer interaction*, Volume 5. Ablex Publishing, 1995. Available as <http://www.sun.com/access/developers/updt.HCI.advance.pdf> (1.10.2006)

[Bernsen et al., 1999] Niels Ole Bernsen, Hans Dybkjaer and Laila Dybkjaer, *Designing Interactive Speech Systems*, Springer-Verlag, London, 1999, 127-160.

[Constantine and Lockwood, 2000] L. Constantine and L. Lockwood, *Software for use – A Practical guide to the models and methods of usage-centered design*. Addison-Wesley, 2000.

[d'Alessandro and Liénard, 1997] Christophe d'Alessandro and Jean-Sylvain Liénard, Synthetic Speech Generation. In: (eds. Ron Cole (Editor in Chief), Joseph Mariani, Hans Uszkoreit, Giovanni Batista Varile (Managing Editor), Annie Zaenen Antonio Zampolli (Managing Editor), Victor Zue, *Survey of the State of the Art in Human Language Technology*, Cambridge University Press and Giardini 1997. Available as <http://cslu.cse.ogi.edu/HLTsurvey/> (1.10.2006)

[Edwards, 2002] A. D. N Edwards, Multimodal interaction and people with disabilities. In: Björn Granström, David House and Inger Karlsson (eds.), *Multimodality in Language and Speech Systems*. Kluwer Academic Publishers, 2002, 73-92.

[Gong and Lai, 2001] Li Gong and Jennifer Lai, Shall We Mix Synthetic Speech and Human Speech? Impact on User's Performance, Perception and Attitude. In *Proceedings of CHI 2001*: 158-166, 2001. Available as

http://www.geocities.com/lai_photography/283Gong_Lai_CameraReady.pdf (1.10.2006)

[Hakulinen et al., 1999] Jaakko Hakulinen, Markku Turunen, and Kari-Jouko Rähkä. The use of prosodic features to help users extract information from structured elements in spoken dialogue systems. In *Proceedings of ESCA Tutorial and Research Workshop on Dialogue and Prosody*: 65-70, 1999. Available as

http://www.cs.uta.fi/hci/spi/reports/the_use_of_prosodic_features.pdf (1.10.2006)

[Hakulinen et al., 1999] Jaakko Hakulinen, Markku Turunen, and Esa-Pekka Salonen, Agents for Integrated Tutoring in Spoken Dialogue Systems. In *Proceedings of Eurospeech 2003*: 757-760, 2003. Available as

<http://www.cs.uta.fi/hci/spi/reports/Tutoring-Eurospeech2003.pdf> (1.10.2006)

[Hartikainen et al., 2004] Mikko Hartikainen, Esa-Pekka Salonen and Markku Turunen, Subjective Evaluation of Spoken Dialogue Systems Using SERVQUAL Method. In *Proceedings of ICSLP 2004*: 2273-2276. Available as

<http://www.cs.uta.fi/hci/spi/reports/SESUSM.pdf> (1.10.2006)

[Huvio et al., 2004] Petteri Huvio, Marjukka Saarijärvi, Pekka Nykänen ja Arto Saikanmäki, Minor-hanke, Vähemmistöryhmien ja vähemmistöjen yhteisölliseen viestintään liittyvä esiselvitys. *ArviD-julkaisu* 04/2004. Liikenne- ja viestintäministeriö, 2004.

[Ibrahim and Johansson, 2002] Aseel Ibrahim and Pontus Johansson, Multimodal Dialogue Systems for Interactive TV Applications. *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces 2002 (ICMI'02)*, Pittsburgh, USA, 117-122. Available as

<http://www.ida.liu.se/~asebe/publications/icmi02.pdf> (1.10.2006)

[McTear, 2004] Michael F. McTear, Spoken Dialogue Technology: Toward the Conversational User Interface. Springer, 2004.

[NKL, 2006] Näkövammaisten keskusliitto ry:n WWW-sivut, 2006. <http://www.nkl.fi/index.htm> (1.10.2006)

[Pols, 1997] Louis C. W. Pols, Speech Synthesis Evaluation. In: (eds. Ron Cole (Editor in Chief), Joseph Mariani, Hans Uszkoreit, Giovanni Batista Varile (Managing Editor), Annie Zaenen Antonio Zampolli (Managing Editor), Victor Zue), *Survey of the State of the*

Art in Human Language Technology, Cambridge University Press and Giardini 1997. Available as <http://cslu.cse.ogi.edu/HLTsurvey/> (1.10.2006)

[Rinnetmäki et al., 2004] Mikael Rinnetmäki, Marko Heikkinen, Ismo Kosonen, Marjukka Saarijärvi, Ari Pöyhtäri, Pekka Nykänen ja Arto Saikanmäki, Digi-tv:n palveluntekijän opas. *ArviD-julkaisuja* 01/2004. Liikenne- ja viestintäministeriö, 2004.

[Rosenfeld et al., 2001] Ronald Rosenfeld, Dan Olsen and Alex Rudnicky, Universal Speech Interfaces. In: *ACM Interactions* 8, 6 (Nov./Dec. 2001), 34-44.

[Schmandt, 1994] Christopher Schmandt, *Voice Communication with Computers*. Van Nostrand Reinhold, New York, 1994.

[Schmidt and Wlodkowski, 2003] Chris Schmidt and Tom Wlodkowski, *A Developer's Guide to Creating Talking Menus for Set-top Boxes and DVDs*. WGBH Educational Foundation, 2003. Available as http://ncam.wgbh.org/resources/talkingmenus/all_print.html (1.10.2006)

[Turunen, 2004] Markku Turunen, *Jaspis – A spoken Dialogue Architecture and its Applications*. University of Tampere, 2004. Available as <http://acta.uta.fi/pdf/951-44-5909-1.pdf> (1.10.2006)

[Turunen and Hakulinen, 2001] Markku Turunen and Jaakko Hakulinen. Agent-based Error Handling in Spoken Dialogue Systems. *Proceedings of the 7th European Conference On Speech Communication and Technology (Eurospeech 2001)*, Aalborg, Denmark, September 3-7, 2001, 2189-2192. Available as http://www.cs.uta.fi/hci/spi/reports/Agent-based_Error_Handling.pdf (1.10.2006)

[W3C SSML, 2006] Speech Synthesis Markup Language (SSML) Version 1.0 W3C Recommendation 7 September 2004. Available as <http://www.w3.org/TR/speech-synthesis/> (1.10.2006)

[Weinshenck, 2000] Susan Weinschenk, *Designing Effective Speech Interfaces*. Wiley, 2000.

[Yankelovich, 1997] Nicole Yankelovich, *Using Natural Dialogs as the Basis for Speech Interface Design*, Submitted to MIT Press as a chapter for the upcoming book "Automated Spoken Dialogue Systems", edited by Susann Luperfoy. Available as

<http://research.sun.com/speech/publications/mit-1998/MITPressChapter.v3.html>
(1.10.2006)

[Yankelovich, 1996] Nicole Yankelovich, How Do Users Know What to Say? In: *ACM Interactions* 3, 6 (Nov./Dec. 1996), 32-43.

[YLE, 2006] Yleisradion WWW-sivut, 2006. <http://www.yle.fi> (1.10.2006)

[Zue et al., 1997] Victor Zue, Ron Cole and Wayne Ward, Speech Recognition. In: (eds. Ron Cole (Editor in Chief), Joseph Mariani, Hans Uszkoreit, Giovanni Batista Varile (Managing Editor), Annie Zaenen Antonio Zampolli (Managing Editor), Victor Zue), *Survey of the State of the Art in Human Language Technology*, Cambridge University Press and Giardini 1997. Available as <http://cslu.cse.ogi.edu/HLTsurvey/> (1.10.2006)

Liite 1. Aistiboxin suunnitellut ominaisuudet

Aistiboxin suunnitellulla käyttöliittymällä voi tehdä seuraavia asioita [Axel, 2006]:

- Katsoa digitaalisia televisiokanavia
- Kuunnella digitaalisia radiokanavia
- Selata ohjelmatietoja ohjelmaoppaasta
- Nauhoittaa ja tallentaa TV-ohjelmia sekä ajastuksen avulla että suosikkiohjelman valinnan avulla
- Toistaa tallennettuja ohjelmia.
- Katsoa kuvia (GIF, PNG, JPEG)
- Kuunnella äänitiedostoja (MPEG 1&2, AVI, DivX, Xvid jne)
- Käyttää DVD-soitinta.
- Käyttää CD-ROM-soitinta
- Käyttää muita oheislaitteita ilman erillistä ajureiden lataamista
- Ladata kuvia ja videoita matkapuhelimesta USB- ja Bluetooth-yhteyksien välityksellä
- Selata muita verkkoon kytkettyjä laitteita
- Katsoa Internet-TV-ohjelmia
- Kuunnella Internet-radiokanavia
- Käyttää Podcast-palveluja ja RSS-syöttöä. Podcasting tarkoittaa multimediatiedostojen, yleisimmin mp3-tiedostojen, julkaisemista verkossa RSS-syötteen liitteenä. RSS on puolestaan XML-kielen laji, jonka avulla tiedon tuottajat voivat välittää materiaalia tiedon kuluttajille koodaamalla verkkosivuillaan tietyt asiat RSS-koodilla [YLE, 2006].
- Lukea äänikirjoja
- Katsoa saavutettavia uutisia