

Kielellisen ja visuaalisen aivokäyttöliittymän potentiaali

TURUN YLIOPISTO
Tietotekniikan laitos
LuK-tutkielma
Tietojenkäsittelytiede
Toukokuu 2024
Sini Salmi

TURUN YLIOPISTO
Tietotekniikan laitos

SINI SALMI: Kielellisen ja visuaalisen aivokäyttöliittymän potentiaali

LuK-tutkielma, 29 s.
Tietojenkäsittelytiede
Toukokuu 2024

Aivokäyttöliittymiä tutkitaan laajasti kommunikaatio-ongelmien ratkaisemiseksi, erityisesti halvaantuneiden avustamiseksi. Lisäksi niillä on potentiaalia helpottaa ihmisen ja koneen välisen vuorovaikutuksen eri ongelmakohtia myös yleisemmällä tasolla. Aivokäyttöliittymät hyödyntävät suoraan aivoista saatuja signaaleja, jolloin perinteistä fyysistä käyttöliittymää ei tarvita. Tämä kirjallisuuskatsaus keskittyy aivokäyttöliittymän mahdollisuuksiin käyttäjän kommunikoinnissa laitteiden kanssa, erityisesti kuviteltujen kuvien ja sanojen avulla. Tutkielmassa tarkastellaan kuutta eri tutkimusta, joissa analysoidaan vokaalien, sanojen, lauseiden ja kuvien kuvittelun tunnistamista eri aivokuvantamismenetelmien avulla. Tulokset viittaavat siihen, että aivokäyttöliittymät voisivat tarjota merkittävää apua halvaantuneille potilaille tulevaisuudessa, mutta niiden yleisessä käytössä ne eivät vielä pärjää perinteisille käyttöliittymille.

Asiasanat: aivokäyttöliittymä, kuviteltu puhe, vuorovaikutusongelmat

Sisällys

1	Johdanto	1
2	Ihmisen ja tietokoneen välisen vuorovaikutuksen haasteita	5
3	Aivokäyttöliittymät	9
3.1	Signaalien kerääminen	9
3.2	Ohjaustavat	11
4	Tutkimukset	13
4.1	Sanojen ja vokaalien tunnistus	13
4.2	Lauseiden tunnistus	16
4.3	Kuvien tunnistus	19
5	Pohdinta	23
6	Yhteenveto	28
	Lähdeluettelo	30

1 Johdanto

Teknologian merkitys elämässämme ja infrastruktuurissamme kasvaa jatkuvasti. Koneet ja tekniset laitteet ovat vähitellen sulautuneet osaksi ihmisten arkea. Jotta ihminen ja kone voivat toimia yhdessä tarvitaan käyttöliittymä, jonka kautta ne voivat kommunikoida keskenään. Tämä käyttöliittymä voi sisältää monenlaisia komponentteja, kuten fyysisiä ja digitaalisia painikkeita, vipuja ja ohjaimia. Jatkossa käytettäessä termiä 'painike' viitataan näihin kaikenlaisiin eri käyttöliittymän komponentteihin.

Ihmisen ja tietokoneen välinen vuorovaikutus (engl. *human computer interactions*, HCI) on tieteenala, joka tutkii ja kehittää tapoja, joilla ihmiset ja tietokonejärjestelmät toimivat yhdessä, erityisesti käyttäjien tarpeiden ymmärtämisen, käytettävyyden ja käyttökokemuksen näkökulmasta (Carroll, 2014). HCI on tärkeä tieteenala, sillä huonosti suunniteltu käyttöliittymä vaikeuttaa käyttöä ja altistaa virheille, jotka voivat pahimmillaan johtaa onnettomuuksiin.

Vaikka HCI tieteenalana syntyi vasta 1980-luvulla, on huonosti suunniteltujen käyttöliittymien seuraukset huomattu jo paljon aikaisemmin. Esimerkkinä tästä voidaan mainita B-17 Flying Fortress lentokoneet. Toisen maailmansodan aikana näiden lentokoneiden laskeutuminen usein epäonnistui ja johti onnettomuuksiin. Nämä onnettomuudet laitettiin siihen aikaan lentäjän syyksi. Psykologi, Alphonse Chapanis kutsuttiin tutkimaan tarkemmin näiden lentäjien tekemiä virheitä. Hän tuli siihen tulokseen, että monet näistä tapaturmista johtuivat huonosta suunnittelus-

ta, eivätkä lentäjien virheistä. Chapanis huomasi, että lentokoneen ohjaamossa kytkimet laskutelineen sisäänvetämiseen ja siipiläppien ohjaamiseen olivat vierekkäin ja täysin samanlaiset. Kahden vuoden aikana tämä suunnitteluvirhe oli aiheuttanut noin 457 lento-onnettomuutta. Chapanisin ehdotuksesta vipujen muodot muutettiin niin, että lentäjät pystyivät tunnistamaan vivut nopeasti toisistaan pelkän vilkaisun tai kosketuksen avulla. Nykyään nämä vipukytkimet on lain mukaan suunniteltava selvästi erilaisiksi toisistaan. Tämä perjaate näkyy myös jokapäiväisessä elämässämme kuten videopelihjaimien ja ajoneuvojen painikkeissa. Erilaiset muodot ja pinnan tekstuurit auttavat käyttäjiä erottamaan kytkimet toisistaan ja vähentävät virheiden mahdollisuutta. (Kuang & Fabricant, 2020).

Ihmisen ja koneen välistä vuorovaikutusta voidaan havainnoida erilaisten mallien avulla. Tämä lähestymistapa mahdollistaa toimintojen hajottamisen pieniin osiin, mikä auttaa analysoimaan ja syventämään ymmärrystä tapahtuvasta vuorovaikutuksesta. Tämän havainnointiin kehitettyjä malleja ovat esimerkiksi TOTE (Miller ym., 1986), GOMS (Card ym., 1983), Normanin seitsemän toiminnan vaihetta (Norman, 1990) ja Plaumannin esittämä vuorovaikutus-silmukka (Plaumann, 2022). Mallit auttavat hahmottamaan vuorovaikutuksen eri vaiheet sekä selventämään, miten ihmisen ja koneen vastuut jakautuvat eri vaiheissa. Ihmisen vastuulla on esimerkiksi vuorovaikutuksen aloittaminen ja oikean syötteen antaminen, kun taas koneen vastuulla on syötteen käsittely ja palautteen antaminen. Tässä tutkielmassa tarkastellaan ihmisen vastuualueita vuorovaikutuksessa sekä niissä ilmeneviä ongelmia ja haasteita, ja voisiko uudenlainen käyttöliittymä, aivokäyttöliittymä (engl. *brain-computer interface*, BCI), olla ratkaisu näihin ongelmiin. Aivokäyttöliittymä mahdollistaa suoran kommunikoinnin käyttäjän ja tietokoneen tai muiden laitteiden välillä hyödyntämällä aivojen signaaleja. Monet kontrolli signaaleina käytetyt toiminnot, kuten välkkyvät valot tai musiikin kuvittelu, eivät kuitenkaan välttämättä ole intuitiivisempia kuin nykyiset käyttöliittymät. Todellinen intuitiivisuus voisi syn-

tyä esimerkiksi sanojen ja kuvien kuvittelun avulla, jolloin käyttäjä voisi pelkästään ajatella haluamaansa toimintoa saadakseen koneen suorittamaan sen. Tässä tutkielmassa perehdytään aivokäyttöliittymiin ja miten ne voisivat helpottaa ihmisten ja koneiden välistä kommunikaatiota.

Tämän tutkielman tutkimuskysymykset ovat:

1. Kuinka aivokäyttöliittymät toimivat?
2. Missä vaiheessa on aivokäyttöliittymä, joka hyödyntää kuviteltua puhetta ja kuvia?
3. Olisiko aivokäyttöliittymä ratkaisu ihmisen ja tietokoneen vuorovaikutuksen haasteisiin?

Tutkielma on kirjallisuuskatsaus aivokäyttöliittymiin. Katsauksen aineiston kerääminen suoritettiin Google Scholar -hakukoneella käyttäen hakutermeinä ”BCI”, ”brain-computer interface”, ”Imagined speech” ja ”Imagined pictures”. Haku rajattiin vuoteen 2020 ja sen jälkeisiin artikkeleihin, mikä tuotti yhteensä 788 tulosta. Näistä tuloksista karsittiin ne, jotka eivät olleet julkisesti saatavilla. Painotehtiin myös aivosignaalien monipuolisuutta, jotta katsaukseen sisältyy muitakin kuin EEG:llä suoritettuja tutkimuksia. Lisäksi vaadin tutkimuksilta moniluokkaista luokitutelta (engl. *multiclass classification*). Näiden kriteerien täyttäneistä tutkimuksista tarkempaan tarkasteluun valittiin 6 tutkimusta; kaksi sanojen ja vokaalien, kaksi lauseiden ja kaksi kuvien kuvittelusta.

Tutkielman luvussa 2 käsitellään ihmisen ja koneen välistä vuorovaikutusta, sen ongelmia sekä miten nämä ongelmat ilmenevät. Luvussa 3 tarkastellaan aivokäyttöliittymien perusteita, kuten signaalien keräystapoja ja kontrollisignaaleja. Luvussa 4 tarkastellaan tutkimuksia sisäisen puheen ja kuvien tunnistamisesta. Luvussa 5 pohditaan aivokäyttöliittymien toimintaa ja niiden potentiaalia muuttaa ihmisten

ja koneiden välistä vuorovaikutusta. Lopuksi kuudennessa 6 on yhteenveto käsitellyistä aiheista.

Tutkielman kirjoittamisessa on käytetty generatiivista tekoälyä apuna kielenhuollossa ja ilmaisutapojen kehittämisessä.

2 Ihmisen ja tietokoneen välisen vuorovaikutuksen haasteita

HCI tutkii ihmisten ja koneiden välistä vuorovaikutusta ja pyrkii tekemään siitä vaivattomampaa (Carroll, 2014). Ihmisen ja koneen välistä vuorovaikutusta on mallinnettu monin erin tavoin. Katrin Plaumann (2022) esittää väitöskirjassaan aikaisempien mallien (Card ym., 1983; Miller ym., 1986; Norman, 1990) pohjalta tehdyn kahdeksan vaiheisen vuorovaikutussilmukan (Plaumann, 2020). Tämä malli esitellään taulukossa 2.1. Tässä tutkielmassa perehdytään vuorovaikutussilmukassa ilmeneviin ongelmiin. Laitteen puolella ohjelmistossa voi olla virheitä, jotka ilmenevät vaiheissa 4-6. Tässä tutkielmassa keskitytään kuitenkin vuorovaikutussilmukan kohtiin 1-4, jossa tärkeässä osassa on ihmisen vuorovaikutus käyttöliittymän kanssa.

Vuorovaikutussilmukka suoritetaan jokaisen käyttäjän toiminnan yhteydessä. Esimerkiksi taulukon 2.1 tilanteen jälkeen käyttäjä voi asettaa tavoitteeksi siirtyä 'Yleiset' kohtaan ja sieltä edelleen 'Ohjelmistopäivitykset'. Tällöin koko vuorovaikutussilmukka käydään uudelleen läpi kahdesti.

Käyttöliittymien suunnittelussa käytettävyyden harkitseminen on tärkeää. Monet erilaiset käyttötilanteet ja käyttäjät on otettava huomioon. Usein tietyt käyttäjäryhmät kuitenkin jäävät harkinnasta helposti pois. Esimerkiksi useimmat käyttöliittymät on suunniteltu niin, että tarvitsemme niiden käyttöön käsiä. Kädet eivät kuitenkaan aina ole käytettävissä, joko sen takia, että ne ovat jo käytössä, tai lii-

Taulukko 2.1: Ihmisen ja tietokoneen vuorovaikutus-silmukka (Plaumann, 2022)

1. Tavoitteen muodostaminen	Käyttäjä muodostaa haluamansa tavoitteen, kuten "Haluan mennä asetuksiin"
2. Vuorovaikutusvälineen valitseminen	Käyttäjän täytyy muistaa minkälaisen toimintojen kautta tavoitteeseen päästään ja valita lähestymistapa kuten hiiri tai kosketusnäyttö.
3. Toiminnan suorittaminen	Käyttäjä liikuttaa käden ja hiiren avulla kursorin 'Asetukset' painikkeen päälle ja painaa tästä.
4. Syötteen vastaanottaminen	Kone ottaa tämän syötteen vastaan.
5. Syötteen käsittely	Kone suorittaa toiminnan syötteen ohjeiden mukaisesti.
6. Palautteen antaminen	Kone antaa ilmoituksen, että se on suorittanut tehtävän.
7. Palautteen käsittely	Käyttäjä käsittelee saamaansa ilmoitusta.
8. Järjestelmän tilan arviointi	Käyttäjä arvioi tekikö kone niin kuin hän halusi.

kuntarajoitteiden takia. Tällöin voi vaihtoehtoisesti käyttää äänikontrollia, joka on kehittynyt paljon vuosien saatossa ja näkyy tekoälyllä toimivien ääni assistenttien, kuten Sirin ja Alexan, käytössä. Myös katseenseurausteknologiaa sovitellaan käyttöliittymiin. (Woods ym., 2017). On kuitenkin tilanteita, joissa henkilö voi olla halvaantunut niin, ettei hän pysty puhumaan tai liikuttamaan silmiään. Tällöin myös ääni- ja katseohjatut käyttöliittymät ovat hyödyttömiä. Laitteen käyttämiseen voi vaikuttaa myös ihmisestä riippumattomat ulkoiset olosuhteet. Esimerkiksi sateella kosketusnäyttö voi kastua, eikä vastaa kosketuskomentoihin tai turbulenssin takia lentäjällä on vaikeuksia ylettää oikeaan painikkeeseen. Näitä kutsutaan nimellä tilannekohtaiset rajoitukset (engl. *situational disabilities*). (Gjøsæter ym., 2019).

Erilaisten käyttäjien ja tilanteiden lisäksi käyttöliittymiä suunnitellessa on tärkeää kiinnittää huomiota moniin yksityiskohtiin. Painikkeiden tulisi olla sijoiteltu niin, että ne ovat helposti ja nopeasti saavutettavissa. Mitä tärkeämpi painike on sitä suurempi ja keskeisemmällä paikalla sen tulisi olla. (Liu ym., 2020). Olennaista on myös välttää käyttäjän muistin kuormittamista liikaa, sillä se lisää virheiden riskiä. Tämä tarkoittaa sitä, että mitä vähemmän käyttäjän tarvitsee muistaa, sitä parempi. Nykyään koneilta odotetaan kuitenkin entistä monipuolisempia toimintoja, mikä lisää käyttöliittymän monimutkaisuutta, ja näin kuormittaa käyttäjän muistia. (Sridevi, 2014).

Edellä mainittuihin ongelmiin voisi ratkaisuna olla aivokäyttöliittymä. Näiden avulla käyttäjä voisi antaa käskyjä suoraan aivoistaan. Aivokäyttöliittymän avulla käyttäjä voisi vain ajatella halauamaansa asiaa ja laite toteuttaisi sen. Tällöin käyttäjän muisti ei kuormittuisi eikä oikeaa painiketta tarvitsisi etsiä. Aivokäyttöliittymät tarjoaisivat myös uuden kommunikointitavan niille, joilla ei esimerkiksi ole käsiä tai ovat halvaantuneet. Lisäksi ne voivat auttaa estämään onnettomuuksia tilanteissa, joissa nopea toiminta on kriittistä.

Esimerkiksi ajoneuvojen jarrutustilanteissa nopea toiminta voi pelastaa henkiä. Tutkimusten mukaan aivoissa syntyy aikomus liikkeestä noin 0,4-1 sekuntia ennen itse liikkeen suorittamisen aloittamista (Kalat, 2018). Sen jälkeen kestää esimerkiksi noin 0,2-0,4 sekuntia siirtää jalka kaasulta jarrulle (Drożdziel ym., 2020). On myös mahdollista, että ajaja vahingossa painaa kaasua jarrun sijasta tai ei osu jarruun ollenkaan. Aivokäyttöliittymän avulla voitaisiin tunnistaa aivoissa liikkeen aikomus, minkä ansiosta auto voisi aloittaa jarrutuksen välittömästi ja näin lyhentää jarrutusmatkaa. Esimerkiksi Nissan ja Mercedes-Benz ovat jo tutkineet aivokäyttöliittymien soveltuvuutta autoihin (Bitbrain, 2018; Mercedes-Benz, 2021).

3 Aivokäyttöliittymät

Aivokäyttöliittymä on käyttöliittymä, joka mahdollistaa suoran kommunikaation aivojen ja ulkoisen koneen välillä. Yhdistämällä viimeisimmät teknologiat ja aivoista kerätty data, pystytään tulkitsemaan aivojen eri signaaleja. Tiedetään esimerkiksi, että takaraivolohko vastaa visuaalisen informaation käsittelystä ja otsalohkossa oleva motorinen aivokuori nimensä mukaisesti vastaa motorisista liikkeistä. Tämän ansioista tiedetään mitä aivojen osaa tarkkailla, kun halutaan seurata tietynlaista signaalia. (Kalat, 2018). Aivojen toimintaa voidaan seurata monin eri tavoin, kuten käyttäen elektrodeja tai magneetteja. Neuroverkkojen ja koneoppimisen avulla voidaan kouluttaa järjestelmä tunnistamaan tiettyjä piirteitä kerätyistä signaaleista. Ne voidaan opettaa luokittelemaan signaalien piirteet eri ryhmiin, mikä mahdollistaa henkilön mahdollisten ajatusten tunnistamisen. Tavoitteena on, että kone toimii ihmisen ajatuksien kautta antamien käskyjen mukaisesti. Yksi tärkeimmistä käyttötarkoituksista tällaiselle teknologialle on esimerkiksi sen hyödyllisyys halvaantuneille henkilöille, jotka ovat menettäneet kaikki muut tavat kommunikoida.

3.1 Signaalien kerääminen

Aivoista voidaan kerätä signaaleja invasiivisin tai ei-invasiivisin tavoin. Invasiivisissa aivokäyttöliittymissä elektrodit asetetaan aivoihin kallon alle. Invasiivisen aivokäyttöliittymän etuna on se, että sen läheisyys tarkasteltaviin neuroneihin mahdollistaa tarkan spatiaalisen ja ajallisen resoluution. Ongelmana on, että siihen tarvitaan ai-

na leikkaus, joka on kallista, aikaa vievää ja leikkauksessa on aina komplikaatioiden riski. Elektrodiä ei myöskään saa siirrettyä paikasta toiseen ilman toista leikkausta, mikä vaikeuttaa sen siirtämistä, jos halutaankin kerätä signaaleja toisesta kohdasta. Elektrodien ympärille voi myös ajan myötä kertyä arpikudosta, joka voi haitata kerätyn signaalin laatua. (Ramadan & Vasilakos, 2016).

Ei-invasiivisissa menetelmissä aivosignaaleja tarkkaillaan pään ulkopuolelta eri tavoin, kuten asettamalla elektroneja päänahalle tai käyttämällä laitetta, joka ei vaadi fyysistä kosketusta käyttäjään. Aivojen ja kuvauslaitteen välinen etäisyys kuitenkin haittaa signaalien laatua. Tästä huolimatta tämä signaalien keräysmuoto on varsin suosittu, sillä sen käyttämiseen ei tarvita leikkausta. (Ramadan & Vasilakos, 2016).

Sähkökortikografia (ECoG) on invasiivinen tapa mitata aivoissa tapahtuvia sähköisiä signaaleja elektrodien avulla. Signaalit ovat hyvälaatuisia, spatiaalinen resoluutio on yksi millimetri ja ajallinen resoluutio on 3ms, mutta invasiivinen metodi rajoittaa käyttöä. (Ramadan & Vasilakos, 2016).

Elektroenkefalografia (EEG) käyttää elektrodeja mittaamaan neuronien toimintaa seuraamalla näiden aiheuttamia sähköisiä signaaleja päänahan päältä. EEG:n ajallinen resoluutio on yksi millisekunti ja sen spatiaalinen resoluutio on 10 mm. Jotta EEG:llä saataisiin parhaat mittaustulokset, tulee elektrodeissa käyttää geeliä ja niiden olla 1-2 cm etäisyydellä toisistaan. EEG:n avulla pystytään havaitsemaan muutoksia aivoissa millisekunnin tarkkuudella. EEG on suosittu, koska sillä on korkea ajallinen resoluutio, se on helppo, turvallinen ja halpa käyttää. (Ramadan & Vasilakos, 2016).

Magnetoenkefalografia (MEG) on ei-invasiivinen tapa havainnoida aivojen magneettisia kenttiä käyttäen magneettintureita. Anturien etäisyys päänahasta on vähintään 2 cm (Brookes ym., 2022). Sen ajallinen resoluutio on 50 ms ja spatiaalinen resoluutio on 5 mm. (Ramadan & Vasilakos, 2016).

Funktionaalinen magneettiresonanssikuvaus (fMRI) seuraa ei-invasiivisin tavoin veren happipitoisuutta magneetteja käyttäen. Näin pystytään paikantamaan aivoissa aktiivisia alueita, kun happipitoista verta virtaa aivojen aktiivisille alueille. fMRI on helppokäyttöinen, ja sillä pystytään paikantamaan aktiivinen alue jopa 1 millimetrin tarkkuudella. Verellä menee kuitenkin useita sekunteja siirtää hapetta aktiiviselle alueelle, jolloin sen ajallinen resoluutio on huono. (Glover, 2011).

Nämä ovat yleisimmät aivokäyttöliittymissä käytetyt aivokuvantamismenetelät. Nämä neljä menetelmää ovat myös käytössä luvussa 4 mainittavissa tutkimuksissa. Muitakin tapoja on, kuten yksifotoniemissiotomografia (SPECT), joka perustuu gammasäteisiin. Henkilölle annetaan gammasäteilyä lähettävää radioisotooppia, jonka kiertämistä aivoissa seurataan. Tämän metodin spatiaalinen resoluutio on 1 cm ja ajallinen resoluutio on useita sekunteja. Lähi-infrapunaspektroskopia (NIRS) puolestaan on aivokuvannustapa, jossa verenhapenpitoisuutta seurataan lähettämällä aivoihin infrapunavaloa ja mittaamalla kuinka paljon sitä heijastuu takaisin. Tällä on kuitenkin suuri ajallinen resoluutio. (Ramadan & Vasilakos, 2016).

3.2 Ohjaustavat

Aivokäyttöliittymiä ohjataan erilaisien ohjaussignaalien avulla. Nämä jaetaan evokoituihin signaaleihin, joita kutsutaan myös herätetyiksi signaaleiksi, ja spontaaneihin signaaleihin. Yleisimpiä hyödynnettyjä evokoituja signaaleja ovat SSEP ja P300, jotka ovat käytössä esimerkiksi kirjainten valinnassa ruudulta. Menetelmä perustuu siihen, että ruudulla vilkkuvat kirjaimet aiheuttavat aivoissa erilaisia reaktioita, jotka voidaan havaita, kun käyttäjä keskittyy haluamaansa kirjaimeseen. Vilkkuvien valojen lisäksi voidaan käyttää myös kuulo tai tuntoaistimuksia. Evokoidut signaalit ovat siis signaaleja, jotka syntyvät tiedostamatta käyttäjän aivoissa, kun tapahtuu jonkinlainen ulkoinen ärsyke. Käytännössä evokoidut signaalit tavat voivat kuitenkin

kin olla epämukavia tai rasittavia käyttäjille. (Min ym., 2017; Ramadan & Vasilakos, 2016)

Spontaanit signaalit ovat signaaleja, joita henkilö tuottaa vapaaehtoisesti ilman ulkoisia ärsykeitä. Näissä käyttäjä käyttää mielikuvitustaan aivokäyttöliittymän ohjaamiseen. Spontaaneja signaaleja ovat esimerkiksi motoriset ja sensorimotoriset rytmit sekä ei-motoriset kognitiiviset tehtävät. Motoriset ja sensorimotoriset rytmit liittyvät motorisiin toimintoihin, kuten käsien, sormien tai jalkojen liikuttamiseen. Tällaiset signaalit voidaan havaita primaarisella sensorimotorisella aivokuorella, kun henkilö kuvittelee näitä toimintoja (Min ym., 2017). Näitä voidaan käyttää kontrollisignaaleina esimerkiksi niin, että käyttäjän kuvitellessa oikean käden liikuttamista, ruudulla kursori tai pelihahmo voi vastaavasti liikkua oikealle.

Ei-motorisilla tehtävillä toimivaa aivokäyttöliittymää ohjataan kuvittelemalla esimerkiksi musiikkia, hajuja, kuution pyöritystä tai matemaattisten laskujen laske-
mista, jotka toimivat erilaisina käskyinä laitteelle. Näillä tavoilla toimivan aivokäyttöliittymän ohjaaminen ei kuitenkaan ole kovin intuitiivista, sillä käyttäjän täytyy opetella eri komentokäskyt ulkoa, mikä kuormittaa käyttäjän muistia. Tämän takia tutkitaan myös ajateltujen sanojen ja kuvien avulla toimivaa aivokäyttöliittymää. Tämä eroaa edellä mainituista tavoista niin, että esimerkiksi luonnollisella kielellä kuviteltu ajatus, kuten ”haluan avata sähköpostini”, voi toimia käskynä laitteelle. Tämä mahdollistaisi hyvin intuitiivisen tavan kommunikoida laitteiden kanssa. Tähän ohjaustapaan juuri perehdytään tarkemmin tässä tutkielmassa.

4 Tutkimukset

Tässä luvussa perehdytään tutkimuksiin, joissa tavoitteena on tunnistaa kuviteltuja vokaaleja, sanoja, lauseita ja kuvia aivosignaaleista. Näiden tutkimuksien kautta nähdään minkälaisia erilaisia menetelmiä käytetään ja missä tilanteessa tämänhetkinen ajatusten tunnistus teknologia on.

4.1 Sanojen ja vokaalien tunnistus

Tutkimuksissaan Agarwal ja Kumar (2023) sekä Mahapatra ja Bhuyan (2022) tarkastelivat sanojen ja vokaalien tunnistamista EEG-signaaleista. Mahapatra ja Bhuyan käyttivät tutkimuksessa vokaaleja 'a', 'e', 'i', 'o' ja 'u', sekä espanjankielisiä sanoja 'ylös', 'alas', 'oikea', 'vasen', 'eteenpäin' ja 'taaksepäin'. Kun taas Agarwalin ja Kumarin tutkimuksessa käytettiin viittä englanninkielistä sanaa: 'sos', 'lopetä', 'lääke', 'tule tänne' ja 'pesuhuone'. Molemmissa tutkimuksissa koehenkilöille esitettiin ruudulla yksi vokaali tai sana kerrallaan, jonka jälkeen heidän tuli toistaa se hiljaa mielessään. Mahapatran ja Bhuyanin tutkimuksessa vertailtiin muutaman eri laskennallisen menetelmän tarkkuutta. Agarwal ja Kumar mittasivat lisäksi eri menetelmien suoritusnopeutta.

Mahapatra ja Bhuyan (2022) vertailivat TCN:tä sekä CNN–TCN-yhdistelmää eri esikäsittelymenetelmien kanssa. TCN–CNN yhdistelmän tarkkuus ylitti pelkän TCN:n tarkkuuden 12 prosentilla. Paras suoritus saavutettiin hyödyntämällä TCN- ja CNN-yhdistelmää yhdessä pelkän DWT-esikäsittelyn kanssa, mikä johti 96,49

prosentin kokonaistarkkuuteen. Sanojen ja vokaalien tunnistustarkkuuksissa ei ollut merkittäviä eroja.

Tutkimuksessa käytetyt syväoppimismallit olivat ajallinen konvoluutioverkko (engl. temporal convolutional networks, TCN) ja konvoluutiohermoverkko (engl. convolutional neural network, CNN). TCN on hyvä ajallisten ominaisuuksientunnistamisessa, kun taas CNN on parempi tilaominaisuuksien tunnistamisessa. Nämä yhdistämällä, malli pystyi oppimaan signaalien sekä ajalliset että spatiaaliset ominaisuudet. Esikäsittelymenetelminä olivat diskreetti aallokemuunnos (engl. Discrete Wavelet Transformation, DWT) ja itsenäinen komponenttianalyysi (engl. Independent Component Analysis, ICA). Esikäsittelymenetelmiä käytetään piirteiden erottamiseen ja artefaktien poistamiseen. Artefaktit ovat signaaleja, jotka ovat peräisin muualta kuin aivoista, kuten silmien liikkeistä tai muusta lihastoiminnasta.

Agarwal ja Kumar (2023) vertailivat tutkimuksessaan kolmen syvä neuroverkon (CNN, LSTM, GRU) ja kuuden muun koneoppimisluokittelijan tarkkuutta ja nopeutta. Syvä verkkojen suoriutumiset ovat taulukossa 4.1. Tutkimuksessa osoitettiin, että syvä neuroverkot suoriutuivat koneoppimisluokittelijoita huomattavasti paremmin. Suurimmaksi osaksi koneoppimisluokittajien tarkkuus oli alle 30 eli vain hiukan sattuman yläpuolella. Sattuman raja viiden luokan luokituksessa on 20

Taulukko 4.1: Agarwalin ja Kumarin (2023) vertailemien syvä neuroverkkojen tarkkuus ja nopeus

Neuroverkot	Tarkkuus	Nopeus
LSTM	63,88 %	0,15s
GRU	61,66 %	0,14
CNN	71,75 %	0,57s

Tutkimusten vertailu on haastavaa, koska niissä käytetään erilaisia mittareita tulosten esittämiseen. Esimerkiksi Agarwalin ja Kumarin (2023) tutkimuksessa raportoidaan keskimääräinen tarkkuus, joka lasketaan jakamalla ennustettujen kohteiden tarkkuuden summa luokkien lukumäärällä. Mahapatran ja Bhuyanin (2022) tutki-

muksessa tulokset esitetään puolestaan kokonaistarkkuutena, mikä ilmaisee oikein ennustettujen kohteiden lukumäärän suhteessa ennustettavan kohteen kokonaismäärään.

Agarwalin ja Kumarin (2023) tutkimuksen mukaan parhaan tarkkuustuloksen saavutti CCN. Mahapatran ja Bhuyanin (2022) tutkimuksessa sama syväverkko yhdistettiin TCN:n kanssa, mikä johti 96,49 prosentin kokonaistarkkuuteen. Kuitenkin Agarwalin ja Kumarin (2023) tutkimuksen perusteella CCN oli selvästi hitain kolmesta vertailusta syvä verkosta. TCN:n lisääminen voi vielä hidastaa prosessia entisestään.

Tutkimuksissa elektrodien määrät ja näytteenottotaajuudet vaihtelivat valtavasti. Agarwal ja Kumar (2023) käyttivät kokeessaan 30 elektrodia, jotka oli sijoitettu kansainvälisen 10-20 elektrodien sijoittelumallin mukaisesti. Näytteenottotaajuus pidettiin 250 hertsissä. Mahapatran ja Bhuyanin (2022) kokeessa elektrodit sijoitettiin myös kansainvälisen mallin mukaan, mutta käytettiin vain kuutta elektrodia ja näytteenottotaajuus oli 1024 hertsiä. Tutkimuksissa on havaittu, että elektrodien määrän vähentyessä, laskee myös saatujen aivokuvien laatu. Montoya-Martínez ym. (2021) tutkivat miten elektrodien määrä vaikuttaa kuullun puheen tunnistukseen. He havaitsivat, että käytettäessä vähemmän kuin 20-22 elektrodia suorituskyky alkoi selvästi heikentyä. Toisessa tutkimuksessa Panachakel ja Ramakrishnan (2021) kokoavat 18 tutkimusta, jotka käsittelevät kuvitellun puheen tunnistamista EEG:n avulla. He tarkastelevat, kuinka monta elektrodia ja mitä näytteenottotaajuutta oli käytetty kussakin tutkimuksessa. Tutkimusten perusteella 64 elektrodia oli ylivertaisesti suosituin elektrodien määrä, ja alin havaittu elektrodien määrä oli 14. Suosituin näytteenottotaajuus oli 1KHz. Näiden perusteella vaikuttaa siltä, että 250 hertsiä ja kuusi elektrodia eivät ole riittäviä. Toisaalta se voi myös osoittaa, että aiempaa pienemmällä elektrodin määrällä ja taajuudella voidaan saada myös lupaavia tuloksia. Tämä tekisi aivokäyttöliittymästä helpommin käytettävän.

Agarwalin ja Kumarin (2022) tutkimuksessa käytettiin henkilöstä riippumatonta-asetelmaa, jossa mallia ei koulutettu henkilökohtaisesti yksittäisen koehenkilön datalla, vaan kaikkien tutkittavien yhteisellä datalla. Tässä mallissa käytettiin siis samaa koulutusdataa kaikkien koehenkilöiden kohdalla. Yleisesti on huomattu, että henkilökohtaisesti koulutetut mallit, joissa koulutusdatana on käytetty ainoastaan yksilökohtaista dataa, ovat suoriutuneet paremmin. Yksilöiden välillä voi olla merkittäviä eroja aivojen toiminnassa, jotka eivät tule huomioiduiksi, kun mallia koulutetaan kaikille yhteisellä datalla. Mahapatran ja Bhuyanin (2022) tutkimuksessa ei ole mainittu, onko heidän tutkimuksensa tehty henkilökohtaisesti vai ei.

4.2 Lauseiden tunnistus

Tutkimuksissaan Dash ym. (2021) ja Tang ym. (2023) tarkastelivat lauseiden de-koodaamista aivojen signaaleista. Dashin ym. tutkimuksessa oli viisi valmiiksi valittua lausetta, jotka pyrittiin tunnistamaan MEG:iä käyttäen. Tang ym. puolestaan pyrkivät tunnistamaan vapaata sisäistä puhetta käyttäen fMRI:tä. Dash ym. tutkivat, onko uudenlainen OPM-MEG yhtä tehokas sisäisen puheen tunnistamisessa, kuin perinteinen MEG. OPM-MEG teknologia mahdollistaa kevyemmän ja liikuteltavamman MEG-kuvantamismenetelmän, joka on tärkeä läpimurto aivokäyttöliittymien kannalta. OPM-MEG käyttää magnetometrejä, jotka havaitsevat syvältä aivoista tulevat magneettikentät. Magnetometrit ovat kuitenkin herkkiä ympäristömelusta aiheutuvalla kohinalle. Tämän takia MEG-tutkimuksissa gradiometrit ovat suositumpia, sillä ne tarjoavat parempaa melunvaimennusta. (Dash ym., 2021).

Magnetometrejä käyttämällä saatiin kuvitellun puheen tunnistuksen tarkkuudeksi 42.67 % ja Gradiometreillä 43.59 %. Nämä tulokset ovat merkittävästi korkeampia kuin sattumantaso, joka 5-luokan luokittelussa on 20 %. Tutkimuksen tulokset osoittavat, että sisäisen puheen tunnistuksessa magnetometrit ovat hyvin lähellä gradiometriä tarkkuuksia.

Tang ym. (2023) tutkivat kuvitellun puheen tunnistamista fMRI-tekniikan avulla, joka perustuu veren hapenpitoisuuden seurantaan. Hermoimpulssi saa veren tuomaan alueelle happea noin 10 sekunnin ajan. Puhutun englannin nopeus on kuitenkin yli 2 sanaa sekunnissa. Tämä tarkoittaa, että puheen muodostuksen seuraaminen tällä tavalla on varsin hidasta verrattuna puheen muodostuksen nopeuteen. Yhteen fMRI-tekniikalla otettuun aivokuvaan voi siis vaikuttaa noin 20 sanaa.

Tutkimuksessa käytetty koodausmalli koulutettiin datalla, joka kerättiin koehenkilöiden kuunnellessa puhuttuja kertomuksia. Mallin generoimien lauseiden sujuvuuden parantamiseksi hyödynnettiin generatiivista kielimallia, joka oli opetettu englanninkielisillä sanajonoilla. Se pyrkii ennustamaan edeltävien sanojen pohjalta, mikä sana saattaisi tulla seuraavaksi.

Kuvitellun puheen tunnistamistestiä varten, koehenkilöt kertoivat mielessään viisi yhden minuutin pituista tarinaa fMRI-kuvauskoneessa, sekä ääneen sen ulkopuolella. Näin saatiin fMRI-kuvat kuvitellusta tarinan kerronnasta ja ääneen puhutut mallitranskriptit, joita vertailla keskenään. Vertailemalla dekooderin fMRI-kuvista generoimia ennustuksia mallitranskripteihin havaittiin, että ne voitiin yhdistää oikein kaikissa tapauksissa. Tarinoiden dekooderin ennustukset olivat merkittävästi samankaltaisempia referenssitranskriptien kanssa kuin olisi voitu odottaa sattumanvaraisesti. Laadullisen analyysin perusteella dekooderi kykenee tunnistamaan kuvitellun puheen merkityksen. Se kuitenkin kykenee vain harvoin tunnistamaan tarkat sanat. WER, eli word error rate, olikin korkea, noin 93 %. Tästä syystä otettiin käyttöön uudempi mittari, BERTScore, joka käyttää koneoppimista mitatakseen, tarkoittavatko kaksi lausetta samaa asiaa, vaikka käytetyt sanat olisivatkin erilaiset. BERTScore antoi merkittävästi korkeampia tuloksia suurimmassa osassa ajasta, kuin sattumanvaraisesti odotettaisiin. Taulukossa 4.2 on muutama esimerkki tämän toiminnasta.

Taulukko 4.2: Tangin ym. (2023) mallin toiminta

Referenssitranskripti	Mallin käännös
look for a message from my wife saying the she had changed her mind and that she was coming back	to see her for some reason i thought maybe she would come to me and say she misses me
went on a dirt road through a field of wheat and over a stream and by some log buildings	he had to walk across a bridge to the other side and a very large building in the distance

Tutkimuksessa koodausmalli koulutettiin yksilöllisesti jokaiselle koehenkilölle, käyttäen heidän omaa dataa. Tutkijat kokeilivat mallia, joka oli koulutettu toisen koehenkilön datalla. Tulokset osoittivat, että toisten henkilöiden datalla koulutettu dekooderi toimi vain hieman paremmin kuin sattumanvarainen arvaus.

Tutkimuksissa Dash ym. (2021) ja Tang ym. (2023) esitetään kaksi erilaista lähestymistapaa lauseiden tulkintaan, joista toisessa käytettiin MEG- ja toisessa fMRI-menetelmää. Dash ym. (2021) pyrkivät tunnistamaan viiden ennalta valitun lauseen välillä, mikä on hyvin samanlainen asetelma kuin sanojen tunnistamiskokeissa. Lauseet olivat myös hyvin lyhyitä, vain 2-4 sanaa. Näillä lauseilla pystyy ilmaisemaan vain hyvin yksinkertaiset perustarpeet, mikä on hyväksi niille potilaille, jotka pikaisesti tarvitsevat tavan kommunikoida hoitajien kanssa. Tämä ei kuitenkaan ole vielä lähelläkään sellaista aivokäyttöliittymää, jonka olisi tarkoitus ymmärtää kaikki käyttäjän ajatukset. Tang ym. (2023) pääsivät tutkimuksessaan lähemmäksi tällaista aivokäyttöliittymää. Koehenkilöt saivat vapaasti sanoa mielessään mitä halusivat ja koodausmalli pyrki tunnistamaan tämän. Koodausmalli usein onnistuikin dekoodaamaan ajatuksen pääpiirteet.

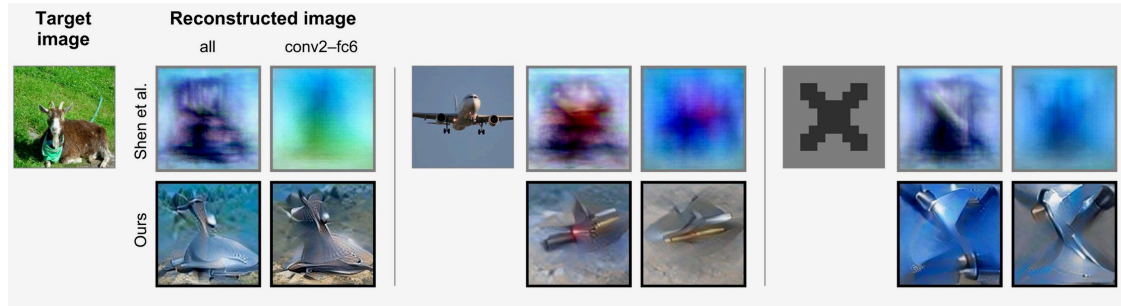
4.3 Kuvien tunnistus

Koide-Majima ym. (2024) tutkivat nähtyjen ja kuviteltujen kuvien rekonstruktioista aivosignaaleista fMRI-kuvauksen avulla. Heidän tutkimuksensa perustuu aiempaan Shen ym. (2019) tekemään tutkimukseen, jonka menetelmää Koide-Majima ym. pyrkivät kehittämään eteenpäin käyttämällä uudenlaisia koneoppimismalleja. Shenin ym. käyttämä malli perustui matalan tason visuaalisen tiedon tulkitsemiseen. Matalana tason visuaalisen tiedon uskotan sisältävän tietoa väreistä ja kontrasteista. Koide-Majima ym. päättivät puolestaan korostaa mallissaan korkean tason visuaaliseen informaatioon käyttöä, joka puolestaan abstraktoi visuaalisen syötteen eri kategorioihin. Koide-Majiman ym. mukaan aiemmat aivokuvantamistutkimukset ovat osoittaneet, että aivojen korkean tason visuaaliselta alueelta saadut tiedot soveltuvat paremmin kuviteltujen kohteiden tunnistamiseen.

Koide-Majima ym. (2024) käyttivät mallinsa kouluttamiseen ja testaamiseen samaa fMRI-dataa kuin Shen ym. Mallin koulutukseen käytetty fMRI-data oli kerätty esittämällä koehenkilöille 1200 luonnollista kuvaa, eli reaailmailman muokkaamattomia kuvia, kuten esimerkiksi kuvia eläimistä tai esineistä. Jokaiselle koehenkilölle luotiin henkilökohtaiset mallit, jotka koulutettiin ja testattiin heidän omalla datallaan.

Sekä Shenin ym. että Koide-Majiman ym. mallia testattiin käyttäen pelkää korkean tason visuaalista informaatiota nähtyjen ja kuviteltujen kuvien rekonstruktioihin. Rekonstruoitujen kuvien laatua arvioitiin käyttäen kuvaparin analyysiä (engl. *pairwise image identification analysis*.), jossa vertaillaan onko rekonstruoitu kuva samankaltaisempi kohdekuvan kanssa, kuin toisen satunnaisesti valitun kuvan kanssa. Koide-Majima ym. käyttämä lähestymistapa pystyi tunnistamaan nähdyt ja kuvitellut kuvat. (nähdyt: 90,7 %, kuvitellut: 75,6 %). Sitä vastoin edellinen menetelmä (Shen ym., 2019) pystyi tunnistamaan vain nähdyt kuvat (nähty: 64,3 %, kuvitellut: 50,4 %). Tämä osoittaa, että Shenin ym. malli tarvitsee matalan tason

visuaalista tietoa toimiakseen, kun taas Koide-Majima ym. tekemä malli hyödyntää paremmin korkean tason visuaalista tietoa. Kuvassa 4.1 on esimerkki rekonstruoituista kuvista.



Kuva 4.1: Kuviteltujen kuvien rekonstruktointi. Kuvassa esitetään miten Shenin ym. ja Koide-Majiman ym. mallit suoriutuivat käyttäessä tietoa kaikkilta aivojen osilta (all) ja vain korkean tason visuaaliselta alueelta (conv2-fc6). Kuva: Koide-Majima ym. (2024), CC BY 4.0 Deed.

Vaikka aivodekooderit koulutettiin käyttämällä vain aivojen reaktioita luonnollisiin kuviin, onnistuneita rekonstruktioita saatiin jopa keinotekoisille muodoille. Koide-majiman ym. mukaan nämä tulokset osoittavat, että dekooderilla on vahva yleistyskyky, eikä se luo kuvia poimimalla niitä vain rajallisista esimerkeistä.

Useissa ajatuksien tunnistamista käsittelevissä tutkimuksissa, koehenkilöiden pyydetään pitävän silmänsä suljettuina tai tarkkailevan mustaa/valkoista ruutua kuvittelutehtävän aikaa. Näin pyritään minimoimaan ristiriitaiset häiriösignaalit. Jos aivokäyttöliittymä tulisi käyttöön jokapäiväisessä elämässä, täytyy sen pystyä tunnistamaan nähdyn kuvan ja kuvitellun kuvan välillä. Fukuma ym. tutkivatkin voiko aivoista tunnistaa ECoG:in avulla kuviteltu kuva, vaikka koehenkilö katsoo samanaikaisesti toisenlaista kuvaa.

Kokeeseen osallistuneilla henkilöillä oli implantoitu elektrodeja ohimo- tai takaraivolohkolle, joiden läpi ventraalinen näkörata kulkee. Tutkimuksessa suoritettiin useita erilaisia kokeita. Yhdessä kokeessa tehtävänä oli kuvitella erilaisia asioita, kuten maisemia, sanoja tai kasvoja, ja näin saada aivokäyttöliittymän kautta heidän edessään oleva näyttö näyttämään vastaavan kategorian kuvia. Erityisesti koe-

henkilöitä pyydettiin pitämään näytöllä tietyn kategorian kuvia mahdollisimman pitkään. Koehenkilöitä kehoitettiin kuvittelemaan vapaasti kuvia, jotka heidän mielestään edustivat hyvin haluttua kategoriaa. Näytölle lähetettävät kuvat valittiin 250 ms välein harjoitusvideossa näytetyistä kuvista ECoG-tulosten samanlaisuusk-sien perusteella. Koehenkilöiden keskimääräinen tarkkuus oli 44,79 %, mikä ylittää sattumanvaraisuuden rajan, 33,3 %.

Viimeisimmässä kokeessa koehenkilöt kuvittelivat tietyn kategorian kuvaa samalla, kun he katsoivat toisen kategorian kuvaa. Koe aloitettiin näyttämällä koehenkilölle sanan tai maiseman kuva, jonka jälkeen heille esitettiin toisen kategorian kuva. Tarkoituksena oli kuvittela ensimmäistä kuvaa samalla, kun näytöllä näkyi toisen kategorian kuva. Dekooderin arvio siirtyi merkittävästi lähemmäksi kuvittelua luokkaa, vaikka he samanaikaisesti katselivatkin toisenlaista kuvaa. Tarkkuus kuitenkin vaihteli eri kategorioiden välillä. Dekooderi onnistui tilastollisesti merkitsevästi tunnistamaan, kun koehenkilö kuvitteli sanaa samalla, kun hän katsoi maiseman kuvaa ($p=0,0006$). Sen sijaan, maiseman kuvittelun tunnistaminen samalla, kun katsottiin sanan kuvaa, ei saavuttanut tilastollisesti merkitsevän rajaa ($p=0,57$, merkitsevyyden raja= $0,05$).

Tutkimuksien asetelmat ja tavoitteet olivat erilaisia. Koide-Majima ym. (2024) tutkimuksessa pyrittiin rekonstruoimaan nähtyjää ja kuviteltuja kuvia. Fukuma ym. (2022) tutkimuksessa tutkittiin mahdollisuutta tunnistaa kuviteltu kuva, vaikka koehenkilö samanaikaisesti katsoisi toisenlaista kuvaa.

Koide-Majima ym. tutkimuksessa kokeiltiin käyttää korkean tason visuaalista tietoa, joka tuotti lupaavia tuloksia. Fukuma ym. tutkimuksessa parhaat tulokset saavutettiin tarkastelemalla korkean tason visuaalista tietoa. Kummatkin tutkimukset siis osoittavat, että korkean tason visuaalinen tieto sopii hyvin kuviteltujen kuvien tunnistamiseen. Fukuma ym. tekemässä tutkimuksessa koehenkilöihin sijoitettujen elektrodien määrä vaihteli. Subduraalisia elektrodeja oli 47-81 ja syväelekt-

rodeja oli 0-20. Elektrodien lukumääristä aiheutuvia eroavaisuuksia ei kuitenkaan mainita tutkimuksessa.

5 Pohdinta

Tässä tutkielmassa tarkastellaan, voiko aivokäyttöliittymä olla ratkaisu HCI:n ongelmiin. Tutkielman alussa luvussa 2 käsiteltiin HCI:hin liittyviä ongelmia kuten käytettävyyden helppous eri ihmisryhmille, tilannekohtaiset haitat, käyttämiseen kuluva aika ja virhepainallukset. Aivokäyttöliittymät voisivat mahdollisesti tarjota ratkaisun näihin haasteisiin muodostamalla suoran linjan koneen ja käyttäjän aivojen välillä. Kaaviosta 5.1 näkyy kuinka tulevaisuuden aivokäyttöliittymät voisivat lyhentää vuorovaikutussilmukkaa poistamalla tarpeen 'vuorovaikutusvälineen valinta' sekä 'toiminnan suorittaminen' vaiheille. Näin käyttäjän ei tarvitse tuhlaa aikaa oikean toiminnon tai painikkeen etsimiseen. Nykyisiin aivokäyttöliittymiin liittyy vielä vuorovaikutussilmukan kolmas vaihe, joka näkyy kaaviossa 5.1 lihavoituna. Aivokäyttöliittymä voisi mahdollistaa kaikkien kannalta helpommin käytettävän, nopeamman ja turvallisemman tavan työskennellä koneiden kanssa.

Virheet laitteita käyttäessä voivat pahimmassa tapauksessa johtaa vakaviin onnettomuuksiin. Tapaturmien estämisessä reaktionopeudella on merkittävä rooli. Teoriassa aivokäyttöliittymien pitäisi nopeuttaa toimintaa, kun käsky tulee suoraan aivoista. Teknologia kuitenkin asettaa rajoitteita sen nopeudelle.

Aivosignaalien kuvausmenetelmät havaitsevat aivosignaaleja eri nopeuksilla. Esimerkiksi fMRI:n ajallinen resoluutio on useita sekunteja. Harkittaessa aivokäyttöliittymiä tapaturmien ehkäisyn kannalta, on fMRI siis liian hidas tapa saada tietoa

Taulukko 5.1: Nykyisten käyttöliittymien ja tulevaisuuden aivokäyttöliittymien vuorovaikutus-silmukka

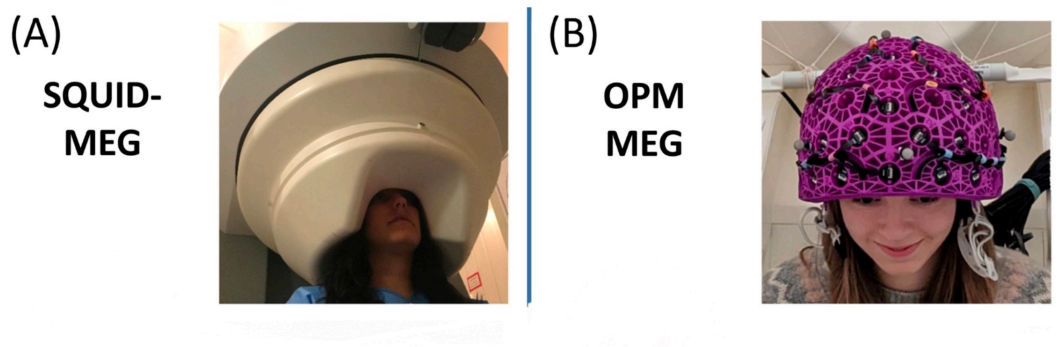
Vuorovaikutus-silmukka	Nykyiset käyttöliittymät	Tulevaisuuden aivokäyttöliittymät
1. Tavoitteen muodostaminen	Käyttäjä muodostaa tavoitteen, minkä haluaa saavuttaa	Käyttäjä muodostaa tavoitteen, minkä haluaa saavuttaa
2. Vuorovaikutusvälineen valitseminen	Hiiren, kosketusnäytön tai ääni komennon valinta	
3. Toiminnan suorittaminen	Käyttäjä siirtää hiirtä, sormea tai sanoo komennon ääneen. Käyttäjän täytyy muistaa oikeat komennot.	käyttäjä kuvittelee hajuja, ääniä tai raajojen liikuttelua. Käyttäjän täytyy muistaa oikeat komennot.
4. Syötteen vastaanottaminen	Kone ottaa tämän syötteen vastaan.	Kone ottaa komennon vastaan

aivoista. EEG, ECoG ja MEG ovat puolestaan selvästi nopeampia ja näin parempia vaihtoehtoja.

Mittausmenetelmän tarvitseman ajan lisäksi neuroverkoilla kuluu aikaa saatujen signaalien tulkitsemiseen. Agarwalin ja Kumarin (2023) tutkimuksessa mitattiin juuri verkon ennustusaikaa (NPT). Parhaimmillaan NPT oli tutkimuksessa noin 150 ms. Näiden nopeimpien verkkojen keskimääräiset tarkkuudet olivat noin 10-8 % pienemmät verrattuna parhaimman tarkkuuden saavuttaneeseen verkkoon. Parhaimman tarkkuuden saaneella verkolla NPT oli kuitenkin selkeästi suurempi, noin 570ms. Fukuma ym. (2022) suorittamassa tutkimuksessa mainitaan heidän käyttämän järjestelmän aika viiveen olleen noin 200 ms. Nämä viiveet ovat vielä varsin suuria, kun verrataan viivettä esimerkiksi näppäimistön ja näytön välillä, joka on noin 40 ms (Floridi, 2021). Ihminen alkaa havaitsemaan viiveen 100 millisekunnin jälkeen, joten käyttömukavuuden kannalta olisi hyvä pitää viive tätä pienempänä (O'Hara & Fleger, 2020). Agarwalin ja Kumarin (2023) sekä Mahapatran ja Bhuyan

(2022) tutkimuksien pohjalta näyttää siltä, että nykyisin joudutaan vielä tasapainoilemaan tarkkuuden ja laskenta-ajan välillä. Tulevaisuuden kannalta on tärkeää pyrkiä vielä parantamaan sekä tarkkuutta että nopeutta entisestään.

Aivokuvausmenetelmien laitteistot ovat hyvin erilaisia ja niiden käytettävyys vaihtelee suuresti. EEG on suosituin aivojenkuvaus menetelmä, sillä se on helppokäyttöinen, turvallinen, edullinen ja sen ajallinen resoluutio on hyvä. Se kuitenkin tarvitsee päänahalle sijoitettavaa geeliä tai suolaliuosta toimiakseen parhaiten. Geeli voi olla monelle epämukava arjessa liikkeessä. Geeli myös kuivuu ajanmyötä. Tähän ongelmaan on kehitelty viime vuosina kuivaelektrodeja, jotka voisivat ratkaista tämän ongelman. (Ramadan & Vasilakos, 2016). MEG on myös suosittu tapa. Perinteinen MEG-kuvauskone kuitenkin painaa noin 450 kg. Kevyempiä MEG-koneita (OPM-MEG) on kehitelty ja se voi olla lupaava vaihtoehto. Kuvassa 5.1 näiden kahden koko ero. Dashin ym. tutkimuksessa OPM-MEG osoitti lähes saman tarkkuuden kuin SQUID-MEG, jääden vain muutaman prosentin päähän. fMRI-kuvantaminen tarvitsee myös ison laitteen, joten sen käytettävyys ei ole todennäköistä. Tosin uudenlaisia kevyempiä MRI laitteita kehitellään, mikä voisi viitoittaa tietä myös fMRI:lle (Deoni ym., 2022).



Kuva 5.1: Perinteisen SQUID-MEG:in ja uudenlaisen OPM-MEG:in koon vertailu. Kuva: Brookes ym. (2022), CC BY 4.0 Deed.

ECoG:in avulla saadaan tarkkoja kuvia aivojen toiminnasta sillä kallo ja iho eivät ole tiellä vaimentamassa signaalia. Tähän kuitenkin tarvitaan leikkaus ja elektrodien ympärille voi ajan kuluessa syntyä arpikudosta, joka haittaa signaalia. Tutkimuksissa käytetään usein ei-invasiivisia tapoja niiden helppouden takia. Neuralink yhtiö kuitenkin aloitti hiljattain, 2023, ihmiskokeet heidän invasiivisesti toimivalle aivokäyttöliittymälle. Aivojen motoriselle aivokuorelle sijoitettu N1-implantti seuraa 1024 elektrodin avulla motoristen liikkeiden kuvittelua, joka muunnetaan kursorin liikkeiksi. (Neuralink, 2023). Tämä voi näyttää suuntaa tulevaisuuden aivokäyttöliittymille. Neuralink kuitenkin keskittyy vasta motoriseen aivokuoreen. Tulevaisuudessa saattaa olla tarpeen kuvata koko aivojen toimintaa, mikä edellyttäisi useiden implanttien sijoittamista ympäri aivoja ja siten useita leikkauksia. Tulevaisuus näyttää mikä tai mitkä tavat lopulta vakiintuvat aivokäyttöliittymiin.

Aivokäyttöliittymät ovat lisäksi varsin herkkiä häiriösignaaleille. Useimmissa tutkimuksissa koehenkilöitä opastetaan olemaan mahdollisimman paikallaan ja katsoa tyhjää ruutua. Jokapäiväisessä elämässä kuitenkin ihmisten kaikista aisteista tulee häiriösignaaleja. Aivokäyttöliittymän tulisi toimia näistä asioista huolimatta. Fukuma ym. (2022) tutkivat sitä miten hyvin aivokäyttöliittymä pystyy tunnistamaan ajatellun asian ristiriitaisesta näköaistimuksesta riippumatta. Tutkimuksessa saatiin vaihtelevia tuloksia. Aivokäyttöliittymien toimimista aistien aiheuttamista häiriöistä huolimatta täytyy vielä tutkia enemmän.

Vaikka aikaisemmat ongelmat saataisiin ratkaistua, tulee aivokäyttöliittymissä toisenlaisia ongelmia vastaan. Käyttöliittymien suunnittelussa ja käytössä on oleellista määritellä, miten laite tietää, milloin sen odotetaan reagoivan. Ongelmat tämän kanssa ilmenevät nykyisissä käyttöliittymissä esimerkiksi taskupuheluna ja Alexan vastatessa kysymyksiin tai komentoihin, jotka ei ole suunnattu sille. Kaikki mahdolliset päivittäiset ajatukset lähtevät aivoista, johon aivokäyttöliittymällä on suora yhteys. Yksi ongelma olisikin se miten aivokäyttöliittymä tunnistaa sille tarkoite-

tun käskyn kaikkien muiden ajatusten keskeltä. Esimerkiksi ongelmia voi aiheuttaa tunkeilevat ajatukset (engl. *intrusive thoughts*). Nämä ovat usein tilanteseen sopimattomia ajatuksia, joiden mukaan ihminen ei kuitenkaan oikeasti toimisi. Esimerkiksi ajatukset kuten, ”entä jos ajaisin tuota puuta päin” tai ”entäpä jos lähettäisin pomolle viestin jossa lukee ’haista paska’”. Aivokäyttöliittymä saattaisi tulkita nämä halutuiksi komentokäskyiksi ja suorittaa kyseiset toiminnot. Aivokäyttöliittymää on tutkittu myös ajamisen avuksi, esimerkiksi jarrutus aikomuksen tunnistamiseen. Tässä ongelmia voi kuitenkin ilmetä tilanteessa, jossa ajaja näkee jotain liikettä silmäkulmassa, säikähtää ja hänen ensimmäinen ajatuksensa on jarruttaa. Nykyään siirtäessä jalkaa kaasulta jarrulle ajajalla on aikaa huomata, että tarvetta jarruttaa ei olekaan. Aivokäyttöliittymä kuitenkin aloittaisi jarruttamisen heti ajatuksen tunnistaessaan. Tästä mahdollisesti syntyvät tarpeettomat äkkipysähdykset voivat myös johtaa onnettomuuksiin. Tämän ongelman ratkaisemiseksi voisi harkita aivokäyttöliittymään varmistuskäskyn lisäämistä. Tämä lisäisi kuitenkin reagointiaikaa ja se todennäköisesti olisi tehoton tunkeilevia ajatuksia vastaan. Kuten ”älä ajattele vaaleanpunaista norsua” lausahduksen kanssa, väkisin jonkin ajatuksen esteleminen päinvastoin tuo ne esiin. Tämä periaate voisi ilmetä myös varmistuskäskyn käytössä.

Aivokäyttöliittymien jatkuva ajatusten seuranta herättää myös kysymyksiä yksityisyydestä. Esimerkiksi Amazonin Alexan tapauksessa on noussut esiin kysymys siitä, ovatko yksityiset keskustelut todella yksityisiä. Kun käyttäjä tekee pyynnön Alexalle, Amazon tallentaa tästä äänitteen sekä kirjallisen kopion. Käyttäjillä on myös ollut vaikeuksia poistaa heistä tallennetut tiedot. (DOJ, 2023). Samat huolenaiheet korostuvat entisestään, kun pohditaan aivokäyttöliittymää, jolla voi olla pääsy kaikkiin käyttäjän ajatuksiin. On mahdollista, että yritykset, valtiot tai muut ulkopuoliset voisivat saada pääsyn näihin tietoihin.

6 Yhteenveto

Tutkielmassa perehdyttiin käyttöliittymien haasteisiin ja mitä ratkaisuja aivokäyttöliittymät voisivat tulevaisuudessa tarjota näihin ongelmiin. Erityisesti erilaisilla käyttäjäryhmillä, kuten ikääntyneillä ja vammaisilla, voi olla vaikeuksia perinteisten käyttöliittymien kanssa. Lisäksi tilannekohtaiset tekijät, kuten sade tai turbulenssi, vaikuttavat käyttökokemukseen ja voivat johtaa onnettomuuksiin.

Aivokäyttöliittymä mahdollistaa suoran kommunikoinnin käyttäjän aivojen ja ulkoisen laitteen välillä. Aivokäyttöliittymän ohjaamiseen voidaan käyttää motoristen toimintojen, hajujen, musiikin, sanojen ja kuvien kuvittelusta tai välkkyvistä valoista syntyviä aivosignaaleja. Näissä aivokäyttöliittymissä on monia samoja ongelmia kuin perinteisissä käyttöliittymissä. Esimerkiksi kaikkien kontrolli signaalien ulkoa opettelu kuormittaa käyttäjän muistia. Näiden sijasta aivokäyttöliittymän ohjaamiseen voisi käyttää vapaata sisäisen puheen ja kuvien kuvittelua. Tämä avaisi mahdollisuuden erittäin intuitiiviseen kommunikaatioon ihmisen ja koneen välillä, mikä voisi eliminoida nykyisten käyttöliittymien käytettävyysongelmat.

Ajatuksienlukuteknologia on vielä kokeellisessa kehitysvaiheessa ja tarvitaan paljon tutkimuksia, jotta se saataisiin toimivaksi. Aivokäyttöliittymiä kokeillaan pääasiassa tilanteisiin, joissa muita kommunikaatiomenetelmiä ei ole saatavilla. Koneoppimisen ja neuroverkkojen kehittyessä aivojen signaalien tunnistaminen paranee. Vaikka aivokäyttöliittymästä voi lähitulevaisuudessa tulla hyödyllinen apuväline esimerkiksi halvaantuneille, sen yleistymiseen arkikäytössä menee vielä aikaa. Aivo-

käyttöliittymien täytyy vielä tarkemmin ja monipuolisemmin pystyä tunnistamaan ajatuksia. Laitteiston täytyy olla helpommin käytettävissä ja tehokkaampi, jotta ne voivat tuoda ratkaisun HCI:n ongelmiin.

Lähdeluettelo

- Agarwal, P., & Kumar, S. (2023). EEG-based imagined words classification using Hilbert transform and deep networks. *Multimedia Tools and Applications*, *83*, 2725–2748. <https://doi.org/10.1007/s11042-023-15664-8>
- Bitbrain. (2018). The Nissan Brain-to-vehicle technology. Haettu maaliskuuta 17, 2024, osoitteesta <https://www.bitbrain.com/blog/nissan-brain-to-vehicle-technology>
- Brookes, M. J., Leggett, J., Rea, M., Hill, R. M., Holmes, N., Boto, E., & Bowtell, R. (2022). Magnetoencephalography with optically pumped magnetometers (OPM-Meg): The next generation of functional neuroimaging. *Trends in Neurosciences*, *45*(8), 621–634. <https://doi.org/10.1016/j.tins.2022.05.008>
- Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Lawrence Erlbaum Associates.
- Carroll, J. (2014). Human computer interaction - brief intro. Haettu maaliskuuta 2, 2024, osoitteesta <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/human-computer-interaction-brief-intro>
- Dash, D., Ferrari, P., Babajani-Feremi, A., Borna, A., Schwindt, P. D. D., & Wang, J. (2021). Magnetometers vs Gradiometers for Neural Speech Decoding. *2021 43rd Annual International Conference of the IEEE Engineering in Medicine*

- Biology Society (EMBC)*, 6543–6546. <https://doi.org/10.1109/EMBC46164.2021.9630489>
- Deoni, S. C., Medeiros, P., Deoni, A. T., Burton, P., Beauchemin, J., D'Sa, V., Boskamp, E., By, S., McNulty, C., Mileski, W., & et al. (2022). Development of a mobile low-field MRI scanner. *Scientific Reports*, *12*(1). <https://doi.org/10.1038/s41598-022-09760-2>
- DOJ. (2023). Amazon agrees to injunctive relief and \$ 25 million civil penalty for alleged violations of children's privacy law relating to Alexa. *U.S. Department of Justice DOJ*. Haettu huhtikuuta 2, 2024, osoitteesta <https://www.justice.gov/opa/pr/amazon-agrees-injunctive-relief-and-25-million-civil-penalty-alleged-violations-childrens>
- Drożdziel, P., Tarkowski, S., Rybicka, I., & Wrona, R. (2020). Drivers' reaction time research in the conditions in the real traffic. *Open Engineering*, *10*(1), 35–47. <https://doi.org/10.1515/eng-2020-0004>
- Floridi, L. (2021). Digital Time: Latency, real-time, and the Onlife experience of Everyday Time. *Philosophy amp; Technology*, *34*(3), 407–412. <https://doi.org/10.1007/s13347-021-00472-5>
- Fukuma, R., Yanagisawa, T., Nishimoto, S., Sugano, H., Tamura, K., Yamamoto, S., Imura, Y., Fujita, Y., Oshino, S., Tani, N., Koide–Majima, N., Kamitani, Y., & Haruhiko, K. (2022). Voluntary control of semantic neural representations by imagery with conflicting visual stimulation. *Communications Biology*, *5*. <https://doi.org/10.1038/s42003-022-03137-x>
- Gjøsaeter, T., Radianti, J., & Chen, W. (2019). Understanding situational disabilities and situational awareness in disasters. https://www.researchgate.net/publication/333448990_Understanding_Situational_Disabilities_and_Situational_Awareness_in_Disasters

- Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurgery Clinics of North America*, 22(2), 133–139. <https://doi.org/10.1016/j.nec.2010.11.001>
- Kalat, J. W. (2018). Biological psychology [sivut 245-246]. Cengage.
- Koide-Majima, N., Nishimoto, S., & Majima, K. (2024). Mental image reconstruction from human brain activity: Neural decoding of mental imagery via deep neural network-based Bayesian estimation. *Neural Networks*, 170, 349–363. <https://doi.org/https://doi.org/10.1016/j.neunet.2023.11.024>
- Kuang, C., & Fabricant, R. (2020). User friendly how the hidden rules of design are changing the way we live, work, and play [luku 3.Error]. WH Allen.
- Liu, W., Gori, J., Rioul, O., Beaudouin-Lafon, M., & Guiard, Y. (2020). How Relevant is Hick’s Law for HCI?, 1–11. <https://doi.org/10.1145/3313831.3376878>
- Mahapatra, N., & Bhuyan, P. (2022). Multiclass Classification of Imagined Speech Vowels and Words of Electroencephalography Signals Using Deep Learning. *Advances in Human-Computer Interaction*, 2022, 1–10. <https://doi.org/10.1155/2022/1374880>
- Mercedes-Benz. (2021). Benz Vision AVTR: Mercedes-Benz Group. Haettu maaliskuuta 18, 2024, osoitteesta <https://group.mercedes-benz.com/innovation/product-innovation/design/vision-avtr-bci.html>
- Miller, G. A., Galanter, E., & Pribram, K. H. (1986). *Plans and the structure of behavior*. Adams-Bannister-Cox.
- Min, B.-K., Chavarriaga, R., & Millán, J. d. (2017). Harnessing prefrontal cognitive signals for brain–machine interfaces. *Trends in Biotechnology*, 35(7), 585–597. <https://doi.org/10.1016/j.tibtech.2017.03.008>
- Montoya-Martínez, J., Vanthornhout, J., Bertrand, A., & Francart, T. (2021). Effect of number and placement of EEG electrodes on measurement of neural

- tracking of speech. *PloS one*, *16*(2), e0246769. <https://doi.org/10.1371/journal.pone.0246769>
- Neuralink. (2023). Prime-Study-Brochure.pdf. Haettu maaliskuuta 24, 2024, osoitteesta <https://neuralink.com/pdfs/PRIME-Study-Brochure.pdf>
- Norman, D. A. (1990). *The design of everyday things*. Currency/Doubleday.
- O'Hara, J. M., & Fleger, S. (2020). Human-System Interface Design Review Guidelines. *U.S. Department of Energy*. <https://doi.org/10.2172/1644018>
- Panachakel, J. T., & Ramakrishnan, A. G. (2021). Decoding Covert Speech From EEG-A Comprehensive Review. *Frontiers in Neuroscience*, *15*. <https://doi.org/10.3389/fnins.2021.642251>
- Plaumann, K. (2022). *Challenging selections in mobile and pervasive computing* [Väitöskirja, University of Ulm, Germany]. <https://d-nb.info/1255703989>
- Ramadan, R., & Vasilakos, A. (2016). Brain Computer Interface: Control Signals Review. *Neurocomputing*, *223*. <https://doi.org/10.1016/j.neucom.2016.10.024>
- Shen, G., Dwivedi, K., Majima, K., Horikawa, T., & Kamitani, Y. (2019). End-to-End Deep Image Reconstruction From Human Brain Activity. *Frontiers in Computational Neuroscience*, *13*. <https://doi.org/10.3389/fncom.2019.00021>
- Sridevi, S. (2014). USER INTERFACE DESIGN. *ISSN 2348-1196 (print) International Journal of Computer Science and Information Technology Research*, *2*(2), 415–426. <https://doi.org/ISSN2348-1196>
- Tang, J., LeBel, A., Jain, S., & Huth, A. (2023). Semantic reconstruction of continuous language from non-invasive brain recordings. *Nature Neuroscience*, *26*, 1–9. <https://doi.org/10.1038/s41593-023-01304-9>
- Woods, D. D., Dekker, S., Cook, R., Johannesen, L., & Sarter, N. (2017). *Behind human error*. <https://doi.org/10.1201/9781315568935>