

Anna-Stiina Brusila

**KÄYTTÄJÄN OMINAISUUKSIEN VAIKUTUS VIRTU-
AALIAVUSTAJIEN PUHEENTUNNISTUKSEEN**



JYVÄSKYLÄN YLIOPISTO
INFORMAATIOTEKNOLOGIAN TIEDEKUNTA
2020

TIIVISTELMÄ

Brusila, Anna-Stiina

Käyttäjän ominaisuuksien vaikutus virtuaaliavustajien puheentunnistukseen

Jyväskylä: Jyväskylän yliopisto, 2020, 30 s.

Tietojärjestelmätiede, kandidaatin tutkielma

Ohjaaja(t): Marttiin, Pentti

Ääniohjattavat virtuaaliavustajat kuten Applen Siri, Amazon Alexa ja Google Assistant yleistyvät nopeaan tahtiin. Tämän tutkielman tarkoitus oli kerätä yhteen, miten käyttäjän eri ominaisuudet vaikuttavat virtuaaliavustajan kykyyn ymmärtää puhetta. Tutkielmassa on otettu huomioon käyttäjän kielen, sukupuolen, iän ja kyvykkyyden vaikutus. Usein ongelmat johtuvat siitä, että virtuaaliavustaja on suunniteltu keskiverto aikuiselle, jolloin muut on jätetty valmistuksessa vähemmälle huomiolle. Usein erot johtuvatkin siitä, että virtuaaliavustajan koulutukseen käytetty data on epätasaisesti jakautunut esimerkiksi, silloin kun kyseessä on käyttäjän sukupuoli tai ikä. Virtuaaliavustajille tuottivat vaikeuksia myös puheessa olevat ylimääräiset tauot ja täytesanat, mitkä olivat yleisiä, jos käyttäjä puhui itselleen vierasta kieltä, oli iäkäs tai hänellä oli puhumiseen liittyviä rajoitteita. Samoin vaikeuksia tuottivat puheviat tai väärin lausutut sanat, mitkä olivat yleisiä esimerkiksi lapsilla, vierasta kieltä puhuvilla ja rajoitteista kärsivillä. Tutkielma on kirjallisuuskatsaus.

Asiasanat: Ääniohjattava virtuaaliavustaja, puheentunnistus, kyvykkyys

ABSTRACT

Brusila, Anna-Stiina

The effect of user characteristics on voice assistant's speech recognition

Jyväskylä: University of Jyväskylä, 2020, 30 pp.

Information Systems Science, Bachelor's Thesis

Supervisor(s): Marttiin, Pentti

Voice assistants such as Apple's Siri, Amazon Alexa, and Google Assistant are becoming more common. The purpose of my thesis was to gather how different characteristics of a user affect the voice assistant's ability to understand speech. The thesis focuses on the impact the user's language, gender, age and disability in speech recognition. The voice assistants are usually designed for the average adult leaving everyone else with less attention in developing the product. Usually the differences are caused by the unevenly distributed data used in the training of the voice assistant as for example in the case of the user's gender or age. Voice assistants also had difficulties with the pauses and fill-in words in speech, which were common if the user spoke a language foreign to them, was elderly or had speech disabilities. Voice assistant also couldn't understand users with speech impairments or mispronounced words which were common if the users were children, foreign language speakers or had disabilities. The thesis is a literature review.

Keywords: voice assistant, speech recognition, ability

TAULUKOT

TAULUKKO 1 Ääniohjattavien virtuaaliavustajien osaamat kielet vuonna 2019	13
--	----

SISÄLLYS

TIIVISTELMÄ	2
ABSTRACT	3
TAULUKOT	4
SISÄLLYS.....	5
1 JOHDANTO.....	6
2 ÄÄNIOHJATTAVAT VIRTUAALIIVUSTAJAT	8
2.1 Virtuaaliavustajan toiminta.....	8
2.2 Käyttötarkoitukset	9
3 KÄYTTÄJÄN KIELEN VAIKUTUS	12
3.1 Kielen vaikutus	12
3.2 Aksenttien vaikutus.....	14
4 FYYSISTEN OMINAISUUKSIEN VAIKUTUS PUHEENTUNNISTUKSEEN.....	17
4.1 Sukupuolen vaikutus	17
4.2 Iän vaikutus	18
4.3 Änkyttäminen ja puheviat.....	19
4.4 Puheeseen vaikuttava kyvykkyys	20
4.4.1 Kehitysvammaisuus.....	20
4.4.2 Fyysinen kyvykkyys	22
5 YHTEENVETO	24
LÄHTEET	26

1 Johdanto

Ääniohjattavat virtuaaliavustajat kuten Applen Siri, Google Assistant ja Amazon Alexa yleistyvät nopeaa tahtia ja vuonna 2018 niiden määrän arvioitiin kymmenkertaistuvan vuoteen 2023 mennessä (Juniper Research, 2018). Ääniohjattavalla virtuaaliavustajalla tarkoitetaan puhekomennoilla toimivaa ohjelmistoagenttia, joka voi sijaita esimerkiksi älypuhelimessa tai kaiuttimessa (Hoy, 2018). Ääniohjattavia virtuaaliavustajia voidaan käyttää viihdekäytön lisäksi moneen muuhunkin, minkä takia niiden kykyä ymmärtää käyttäjän puhetta on tärkeää tutkia. Niitä voitaisiin esimerkiksi hyödyntää kouluissa, turistikohteissa oppaina ja dementiapotilaiden kanssa. Lisäksi monet erilaisista rajoitteista kärsivät voisivat hyötyä puheohjattavasta teknologiasta, sillä se auttaisi parantamaan heidän itsenäisyyttään ja turvallisuuttaan, eikä niiden käyttö vaadi näkökykyä, kirjoittamista tai motorisia taitoja. (Pradhan, Mehta, & Findlater, 2018.) Tästä huolimatta eri rajoitteista kärsivät ja iäkkäät ihmiset on jätetty virtuaaliavustajien kehityksessä vähemmälle huomiolle ja heidän kohdallaan esimerkiksi puheentunnistuksessa tapahtuvat virheet ovat yleisiä. Kyseessä on kuitenkin vielä melko uusi teknologia, joten sillä on vielä paljon haasteita liittyen esimerkiksi turvallisuuteen, yksityisyyteen ja puheentunnistukseen. Tässä tutkielmassa keskitytään virtuaaliavustajan puheentunnistukseen liittyviin haasteisiin.

Tutkielman tutkimuskysymyksenä on, miten käyttäjän eri ominaisuudet vaikuttavat ääniohjattavan virtuaaliavustajan kykyyn ymmärtää puhetta. Tarkoituksena on löytää ne käyttäjän ominaisuudet, jotka vaikeuttavat puheentunnistusta ja koota ne yhteen. Puheentunnistuksella tarkoitetaan koneen tai ohjelman kykyä muuttaa puhe koneen ymmärtämään muotoon (Rouse, 2016). Lisäksi tutkielma pyrkii selittämään miten ja miksi nämä ominaisuudet vaikuttavat virtuaaliavustajan puheentunnistukseen. Tutkielmassa keskitytään käyttäjän pysyviin ominaisuuksiin, jolloin jätetään pois tilapäiset tekijät, kuten päihitteiden tai tunteiden vaikutus. Jos virtuaaliavustaja ei ymmärrä käyttäjän antamaa kommentia, voi toiminto jäädä kokonaan tekemättä tai saatetaan suorittaa väärä toiminto. Tämän takia on tärkeää ymmärtää, mitkä tekijät saattavat vaikeuttaa virtuaaliavustajan puheentunnistusta. Vaikka tutkielma keskittyy vir-

tuaaliavustajiin, voidaan sen tuloksia kuitenkin myös soveltaa muunkin ääniteknologian kehittämisessä paremmaksi.

Kaikista parhaiten virtuaaliavustajat toimivat tällä hetkellä englannin kielellä. Lisäksi suurin osa tutkimuksesta on tehty englanniksi, mikä tulee näkymään tutkielmassani. Lisäksi eri kielissä on erilainen tilanne kehityksen kannalta ja osa tutkielmassa esitellyistä ominaisuuksista voivat näkyä eri tavalla eri kielissä. Myös Kiinassa kehitellään virtuaaliavustajia, mutta ne jäävät tutkielman ulkopuolelle. Tutkielmassa keskitytään vain kaikista tunnetuimpiin virtuaaliavustajiin. Tutkielma on kirjallisuuskatsaus ja käytetyt lähteet ovat pääasiassa tieteellisiä artikkeleita joko Google Scholarista tai Jykdokista. Suurin osa tutkielman lähteistä on löydetty käyttämällä hakusanoja "voice assistant", "speech recognition" ja "ability". Tavoitteena oli löytää vertaisarvioituja ja mahdollisimman uusia tieteellisiä artikkeleita. Kyseessä on nopeasti kehittyvä teknologia, joten lähteiden tuoreus oli tutkielman kannalta erityisen tärkeää. Kaikesta tutkielman kannalta oleellisesta tiedosta ei kuitenkaan löytynyt tieteellistä tutkimusta, joten esimerkiksi virtuaaliavustajien kielitaidossa listauksessa on turvauduttu muihinkin lähteisiin. Näissä tilanteissa tieteellisen tutkimuksen löytäminen ei kuitenkaan ollut kovin oleellista.

Tutkielma etenee siten, että aluksi selitetään, mikä on virtuaaliavustaja, miten se toimii, mihin sitä käytetään ja mihin sitä voitaisiin käyttää. Tämän jälkeen kerrotaan, miten käyttäjän kieli vaikuttaa virtuaaliavustajan puheentunnistukseen. Tässä kappaleessa käydään läpi, miten eri virtuaaliavustajat osaavat eri kieliä ja miten ne ovat ottaneet huomioon saman kielen alueelliset erot. Kappaleen lopuksi kerrotaan, miten virtuaaliavustaja ymmärtää käyttäjää, kun hän ei puhu omalla äidinkielellään. Käyttäjän kielen jälkeen siirrytään käyttäjän fyysisiin ominaisuuksiin. Käyttäjällä on erilaisia fyysisiä ominaisuuksia, jotka vaikuttavat virtuaaliavustajan puheentunnistukseen. Näitä ominaisuuksia ovat käyttäjän ikä, sukupuoli ja kyvykkyys. Edellä mainittujen ominaisuuksien vaikutuksia käydään läpi tutkielman neljännessä kappaleessa. Lopuksi tutkielman tärkeimmät asiat, rajoitteet ja tulokset kootaan yhteen yhteenvedossa.

2 Ääniohjattavat virtuaaliavustajat

Ääniohjattavat virtuaaliavustajat yleistyvät ja kehittyvät nopeaan tahtiin. Vaikka ääniohjattava teknologia on saanut alkunsa jo 1960-luvulla, ne alkoivat saavuttaa suurempaa huomiota, kun Apple julkaisi Sirin vuonna 2011 (Mutchler, 2017). Ääniohjattavalla virtuaaliavustajalla tarkoitetaan ohjelmistoagenttia, joka toimii käyttäjältä saaduilla puhekomennoilla. Ääniohjattava virtuaaliavustaja voi sijaita esimerkiksi älypuhelimessa tai kotikaiuttimessa. Näitä ovat mm. Siri, Amazon Alexa, Microsoft Cortana ja Google Assistant. Keskeistä virtuaaliavustajan toiminnan kannalta on se, että virtuaaliavustaja ymmärtää puhetta ja pystyy vastamaan puhesynteesin avulla sekä suorittamaan erilaisia toimintoja kuten vastamaan kysymyksiin tai lähettämään viestejä. (Hoy, 2018.) Virtuaaliavustajien käyttö lisääntyy koko ajan. Tutkimuksen mukaan vuonna 2017 Yhdysvalloissa 27.5 % älypuhelimien omistajista käytti ääniohjattavaa virtuaaliavustajaa kuukausittain (Wu, 2018). Vuonna 2018 puolestaan arveltiin, että ääniohjattavia virtuaaliavustajia olisi käytössä noin 25 miljoonaa ja niiden määrän arvellaan nousevan 275 miljoonaan laitteeseen vuoteen 2023 mennessä (Juniper Research, 2018). Koska kyseessä on vielä melko uusi teknologia, liittyy siihen myös paljon haasteita turvallisuuden, yksityisyyden ja toimivuuden kannalta. Lisäksi pitää myös selvittää, kuka on vastuussa virtuaaliavustajan mahdollisesti aiheuttamasta vahingosta ja miten esimerkiksi sananvapaus voisi vaikuttaa siihen. (Massaro & Norton, 2016.)

2.1 Virtuaaliavustajan toiminta

Tässä kappaleessa selitetään lyhyesti, miten ääniohjattava virtuaaliavustaja toimii. Ääniohjattavat virtuaaliavustajat perustuvat koneoppimiseen. Tekoälylle ominaista on sen kyky oppia. Jotta lopputulos olisi mahdollisimman hyvä, ääniohjattavien virtuaaliavustajien opettamiseen käytetään paljon dataa, jota on saatu eri lähteistä. Virtuaaliavustajien kouluttamiseen käytetyssä datassa pitää olla huomioitu muun muassa sanojen synonyymit ja erilaiset ääntämistavat.

Koska useat eri yritykset valmistavat ääniohjattavia virtuaaliavustajia, käytetään eri yrityksissä erilaisia datan lähteitä, metodeja ja algoritmeja, minkä seurauksena niiden välillä on eroja. Virtuaaliavustajien erot johtuvat myös siitä, että yritykset, jotka valmistavat ääniohjattavia virtuaaliavustajia painottavat eri virtuaaliavustajan ominaisuuksia ja pyrkivät vastaamaan käyttäjien erilaisiin tarpeisiin. (Polyakov, 2018.) Virtuaaliavustajien välinen ero tulee näkymään myös tutkielman edetessä, sillä usein eri virtuaaliavustajilla saatiin erilaisia tuloksia ja niillä oli erilaisia heikkouksia ja vahvuuksia.

Puheentunnistus on yksi tutkielman keskeisiä käsitteitä ja hyvin tärkeä virtuaaliavustajan toiminnan kannalta. Puheentunnistuksella tarkoitetaan koneen tai ohjelman kykyä löytää puheesta sanoja ja ilmauksia ja muuttaa ne koneen ymmärtämään muotoon. Ääniohjattavien virtuaaliavustajien lisäksi puheentunnistusta hyödynnetään muissakin puheella toimivassa teknologiassa. Puheentunnistuksen toiminta perustuu algoritmeihin, joiden avulla se pystyy yhdistämään sanan äänen kanssa ja erottaa samalta kuulostavat sanat toisistaan. (Rouse, 2016.) Puheentunnistukseen kuulu kaksi vaihetta. Ensimmäisessä vaiheessa määritellään, koska käyttäjän puhe alkaa ja loppuu sekä poistetaan saadusta äänisyöttestä turhat taustäänät. Toisessa vaiheessa pyritään ymmärtämään käyttäjän antaman käskyn tarkoitus ja löytää käyttäjälle paras mahdollinen vastaus sekä tietoalueen kyselyhaaraprosessien kartoitus. (Liu & Zhang, 2014.)

Ääniohjattavat virtuaaliavustajat vaativat useaa eri teknologiaa toimiakseen. Näitä teknologioita ovat ääniohjaus, automaattinen puheentunnistus, puheesynteesi, äänibiometriikka, dialogin johtaminen, luonnollisen kielen ymmärtäminen ja nimentunnistus (Polyakov, 2018). Ääniohjattava virtuaaliavustaja aktivoituu kuullessaan avainsanansa kuten "ok Google" tai "hei Alexa". Joissain tilanteissa esimerkiksi älypuhelimessa käyttäjä voi myös aktivoida virtuaaliavustajan painamalla siihen tarkoitettua ikonia. (Hoy, 2018.) Virtuaaliavustajan toimintaprosessiin kuulu äänisyöttö, puheentunnistus, puheesynteesi ja toiminnon suorittaminen. Äänisyötön saamiseen vaaditaan mikrofoni. Äänisyötön jälkeen on puheentunnistuksen kaksi vaihetta. Puheentunnistusta seuraa puheesynteesi, joka mahdollistaa virtuaaliavustajan puhumisen, jolloin virtuaaliavustaja pystyy antamaan käyttäjälle puhutun vastauksen. Lopuksi virtuaaliavustajan pitää suorittaa käyttäjän haluama toiminto. (Liu & Zhang, 2014.) Jos ääniohjattava virtuaaliavustaja ei ymmärrä käyttäjältä saatua äänisyötettä, ei ääniohjattava virtuaaliavustaja pysty suorittamaan käyttäjän haluamaa toimintoa tai se suorittaa väärän toiminnon.

2.2 Käyttötarkoitukset

Ääniohjattavat virtuaaliavustajat voivat tehdä paljon eri asioita ja niistä voidaan myös hyötyä monella eri tavalla. Virtuaaliavustajan käyttö on usein kirjoittamista nopeampaa. Lisäksi puhekomentojen käyttö vapauttaa käyttäjän kädet, jolloin hän voi tehdä samalla muita asioita. Ääniohjattavien virtuaaliavustajien

ominaisuudet vaihtelevat tuotteen mukaan. Yleisesti virtuaaliavustajat kuitenkin pystyvät esimerkiksi soittamaan ja vastaanottamaan puheluita, lähettämään viestejä, lukemaan ja lähettämään sähköposteja, asettamaan hälytyksiä ja tekemään kalenteriin merkintöjä. Ne pystyvät myös tekemään listoja, hallitsemaan mediantoistoa ja esineiden internettiin yhdistettyjä laitteita. Tämän lisäksi ne pystyvät kertomaan tarinoita ja vitsejä sekä vastaamaan käyttäjän kysymyksiin. (Hoy, 2018.) Tällä hetkellä virtuaaliavustajia käytetään paljon viihdekäyttöön ja esimerkiksi Amazon Alexa antaa käyttäjilleen mahdollisuuden ladata erilaisia ominaisuuksia, joilla pyritään antamaan käyttäjälle yksilöllinen kokemus. Näiden ominaisuuksien avulla esimerkiksi perheet voivat leikkiä yhdessä, sillä ne on suunniteltu tukemaan useita käyttäjiä samanaikaisesti. (Beirl, Yuill, & Rogers, 2019.)

Ääniohjattavat virtuaaliavustajat kehittyvät nopeasti ja samalla niille kehitetään uudenlaisia ominaisuuksia. Tulevaisuudessa niitä tullaan hyödyntämään entistä monipuolisemmin. Virtuaaliavustajan ääni on vielä melko robottimainen, mutta tulevaisuudessa se voisi lukea esimerkiksi äänikirjoja. Niitä voitaisiin myös hyödyntää dementia-potilaiden kanssa, sillä ne jaksavat vastata samoihin kysymyksiin turhautumatta. Lisäksi ääniohjattavat virtuaaliavustajat voivat parantaa tiedon saatavuutta eri rajoitteista kärsiville sillä niiden käyttö ei vaadi käsiä tai näkökykyä ja ne eivät väsy tai turhaudu. Yksi mahdollisuus on myös ääniohjattavien virtuaaliavustajien käyttö reaaliaikaisessa kieleltä toiselle kääntämisessä. (Hoy, 2018.)

Puhekomennoilla toimivaa virtuaaliavustajaa voidaan hyödyntää myös opetuksessa. Virtuaaliavustajat pystyvät luomaan tehtäviä, opettamaan pelien avulla sekä vastaamaan oppilaiden kysymyksiin ja toimia luokkahuoneissa opettajan apuna. Ääniohjattavien virtuaaliavustajilla on erityisesti potentiaalia auttaa oppilaita vieraan kielen oppimisessa. Tämä perustuu siihen, että virtuaaliavustajat osaavat useita eri kieliä ja niiden kanssa voidaan keskustella samalla tavalla kuin ihmisen kanssa. Etuna virtuaaliavustajan käytössä opetuksessa on se, että virtuaaliavustajalle voi puhua ja siltä voi kysyä kysymyksiä ilman paineita tai nolostumista. (Franceschin, 2017.)

Ääniohjattavat virtuaaliavustajat voivat olla hyödyksi myös erialojen yrityksille. Tulevaisuudessa virtuaaliavustajat ovat tehokkaampia ja niitä voidaan hyödyntää varastotason tietokoneissa, mikä tehostaa toimintaa ja vähentää kustannuksia (Hauswald ym., 2016). Erilaiset turistikohteet, kuten museot voivat hyödyntää ääniohjattavia virtuaaliavustajia oppaina. Ne tarjoavat asiakkaalle mahdollisuuden kysyä kysymyksiä häntä kiinnostavasta aiheesta omalla kielellään. (Hoy, 2018.) Ääniohjattavia virtuaaliavustajia hyödynnetään myös autoissa. Autoissa ne voivat esimerkiksi hallinnoida esineiden internettiin yhdistettyjä laitteita, olla avuksi navigoinnissa tai toimia ihan vain matkaseurana. (Lugano, 2017.)

Ääniohjattavista virtuaaliavustajista voisi olla myös paljon hyötyä erilaisista rajoitteista kärsiville. Ääniteknologian käyttö ei vaadi motorisia taitoja tai näkökykyä. Lisäksi monelle on kirjoittamiseen liittyviä haasteita, jotka voidaan välttää virtuaaliavustajan avulla. Ääniohjattavaa virtuaaliavustajaa voitaisiin

esimerkiksi käyttää puheterapeuttina. Lisäksi moni kokee niiden parantavan käyttäjän itsenäisyyttä, sillä moni pystyy virtuaaliavustajan avulla suorittamaan tehtäviä, mitkä eivät ole aikaisemmin onnistuneet. Virtuaaliavustaja pystyy parantamaan myös turvallisuutta esimerkiksi, jos käyttäjän pitää hälyttää paikalle apua hätätilanteessa. (Pradhan, Mehta, & Findlater, 2018.) Virtuaaliavustajista voisi olla myös hyötyä esimerkiksi autistisille lapsille. Tutkimusten mukaan lapset, joilla on autismi kokevat robotin kanssa vuorovaikuttamisen helpomaksi kuin ihmisen kanssa. Tätä voidaan hyödyntää myös virtuaaliavustajan kanssa ja käyttää, sitä esimerkiksi autistisen lapsen tukena. Keskustelu virtuaaliavustajan kanssa voisi myös parantaa autistisen lapsen vuorovaikutustaitoja. (Krakovsky, 2018.) Lisäksi sokeat käyttäjät hyötyvät virtuaaliavustajista, sillä sen käyttö ei vaadi näkökykyä. Puhekomennoilla toimiva virtuaaliavustaja on myös muita vaihtoehtoja nopeampi ja vaatii vähemmän työtä käyttäjältä. (Abdolrahmani, Kuber, & Branham, 2018.)

3 Käyttäjän kielen vaikutus

Ääniohjattavat virtuaaliavustajat eroavat toisistaan paljon, kun kyse on niiden kielitaidosta. Kielet, joita virtuaaliavustaja osaa voivat vaihdella muutaman ja usean kymmenen välillä. Virtuaaliavustajat ymmärtävät myös aksentteja erilaisella menestyksellä. Yksi syy, miksi virtuaaliavustajien kannattaa panostaa eri kielten osaamiseen on kilpailuedun saaminen. Tämä näkyy esimerkiksi siten, että Yhdysvalloissa Amazon Alexalla on noin 70 %:n markkinaosuus (Sterling, 2019). Maailman laajuisesti tilanne tasoittuu, sillä Alexalla ja Google Assistantilla on molemmilla noin 30 %:n markkinaosuus ja Sirillä noin 6 %:n markkinaosuus (Statista.com, 2020). Eri kielten kehittäminen tarjoaa kilpailuedun ja mahdollisuuden pärjätä markkinoilla. Seuraavaksi käydään tarkemmin läpi, mitä kieliä virtuaaliavustajat osaavat ja miten ne ymmärtävät eri aksentteja.

3.1 Kielen vaikutus

Ääniohjattavat virtuaaliavustajat eroavat toisistaan niiden osaamien kielten määrässä. Englanti on kuitenkin ääniteknologiassa huomattavasti muita kieliä kehittyneempi kaikkien suurten virtuaaliavustajien kohdalla. Virtuaaliavustajien verkkosivuilta löytyi todella vähän tietoa niiden kielitaidosta, joten niiden kielitaito täytyi löytää muista lähteistä. Niiden kielitaidosta esiintyi myös ristiriitaista tietoa eri lähteiden välillä esim. siinä osaako Google Assistant suomea. Google Assistant osasi vuoden 2019 alkupuolella noin 30 eri kieltä, mikä on huomattavasti enemmän verrattuna siihen, mitä se osasi vuonna 2017. Applen Siri puhuu kahtakymmentäyhtä eri kieltä ja monia eri murteita. Microsoft Cortana osaa kahdeksaa eri kieltä ja Amazon Alexa vain viittä. (Wiggers, 2019.) Yksi syy siihen, että Google Assistant osaa eniten kieliä on se, että Googlen valmistamilla virtuaaliavustajilla on käytössään suuret tietokannat, joiden avulla virtuaaliavustaja oppii uuden kielen ja sen eri murteet muita virtuaaliavustajia nopeammin (Globalme, 2020). Eri kielten ja aksenttien välillä on eroja puheentunnistuksessa kestävässä ajassa. Parhaiten ymmärrettiin englantia eri aksenttil-

la puhuvia. Muita kieliä kuin englantia puhuvilla esimerkiksi ranskaa, thaita tai japania puhuvien puheentunnistuksessa kesti kauemmin. (Pal & Bhowmick, 2017.)

Parhaiten virtuaaliavustajat osaavat puhujamäärältään suuria länsimaisia kieliä ja englantia kaikista parhaiten. Käyttäjä pystyy valitsemaan haluamansa kielen laitteen asetuksista. Taulukkoon 1 on listattu virtuaaliavustajien osaamat kielet vuonna 2019 Wiggersin artikkelin perusteella. Yhdysvalloissa suosittu Amazon Alexa osaa viittä eri kieltä: espanjaa, japania, englantia, saksaa ja ranskaa. Kieli taidollaan Alexa ei erotu muista, sillä Cortana, Siri ja Google Assistant osaavat kaikki nämä viisi kieltä. Microsoft Cortana osaa muutaman kielen enemmän kuin Alexa. Applen Siri ja Google Assistant osaavat molemmat myös sellaisia kieliä, joilla ei ole paljon puhujia ja erottuvat näin myös markkinoilla esimerkiksi Siri osaa hepreaa ja suomea ja Google Assistant norjaa. Ne osaavat myös puhujamäärältään isoja kieliä, joita Alexa ja Cortana eivät ole kehittäneet erityisesti Aasian maissa. (Wiggers, 2019.)

TAULUKKO 1 Ääniohjattavien virtuaaliavustajien osaamat kielet vuonna 2019

Virtuaaliavustaja	Alexa	Cortana	Google Assistant	Siri
Osatut kielet	englanti espanja japani ranska saksa	englanti espanja italia japani kiina (yksinkertais- tettu) portugali ranska saksa	arabia bengali englanti espanja gujarati hindi hollanti indonesia italia japani kannada kiina (perinteinen) korea malajalam marathi norja portugali puola ranska ruotsi saksa tamil tanska telugu thai turkki urdu venäjä	arabia englanti espanja heprea hollanti italia japani kiina (man- dariini, shanghai, kanton) korea malaji norja portugali ranska ruotsi saksa suomi tanska thai turkki venäjä

Kun ääniohjattavalle virtuaaliavustajalle opetetaan uusia kieliä esimerkiksi suomea, pitää ottaa huomioon monia eri asioita. Ääniohjattavan virtuaaliavustajan pitää sekä ymmärtää kieltä että kyetä puhumaan sitä. Uuden kielen opettamisessa ääniohjattavalle virtuaaliavustajalle on otettava huomioon uuden kielen kielioppi, erilaiset termit, puhetyylit ja virtuaaliavustajan opetuksessa käytettävä tietokanta. Kielien välillä saattaa olla isojakin eroja esimerkiksi sanajärjestyksessä ja tarvittavien sanojen määrässä, mikä vaikeuttaa uuden kielen oppimista. Ei kuitenkaan riitä, että ääniohjattava virtuaaliavustaja ymmärtää kieltä, sen pitää myös ymmärtää eri kulttuureita. Virtuaaliavustajan täytyy oppia toimimaan siten, että se ei riko kulttuurin normeja tai tapoja. Jokin virtuaaliavustajan ominaisuus tai sanavalinta saattaa olla todella toimiva joissain kulttuureissa, mutta koetaan loukkaavaksi toisessa. (Wiggers, 2019.) Suomen kieli eroaa paljon englannin kielestä, mikä on usean ääniohjattavan virtuaaliavustajan vahvin kieli. Lisäksi Suomessa keskustelukulttuuri eroaa esimerkiksi Yhdysvaltojen keskustelukulttuurista. Nämä on kaikki huomioitava virtuaaliavustajan suomen kielen kehittämisessä.

Ääniohjattavat virtuaaliavustajat osaavat suomea heikosti. Tällä hetkellä ainoastaan Applen Siri toimii suomeksi (Wiggers, 2019). Google Assistantin suomen kielen taidoista löytyy ristiriitaista tietoa. Wiggersin (2019) mukaan Google Assistant ei osaa suomea, kun taas Karkion (2019) mukaan se osaa suomea, mutta huonosti. Normaalin keskustelun käyminen suomeksi virtuaaliavustajan kanssa on hankalaa ja toimii periaatteessa siten, että käyttäjä sanoo muutaman sanan ja odottaa toiminnon suoriutumista (Karkio, 2019). Amazon Alexaa ei voida käyttää suomeksi. Amazon ei myöskään ole listannut suomea kehitettävien kielten listalle, joten Alexaa ei todennäköisesti hetkeen ole tulossa saatavaksi suomen kielellä. (Wiggers, 2019; Amazon.com & Barr, 2016). Koska virtuaaliavustajilla esimerkiksi Alexalla on tällä hetkellä heikko suomen kielen taito, täytyy niitä käyttäessä turvautua englannin tai jonkun muun kielen puhumiseen, johon liittyy omat ongelmansa.

3.2 Aksenttien vaikutus

Tässä kappaleessa käydään läpi, miten saman kielen alueelliset erot ja vieraalla kielellä puhuminen vaikuttavat ääniohjattavan virtuaaliavustajan puheentunnistukseen. Ääniohjattavien virtuaaliavustajien on otettava tuotteen kehityksessä huomioon, että samaa kieltä voidaan puhua eri tavalla eri alueilla. Esimerkiksi Yhdysvalloissa käytetty englanti eroaa jonkin verran Australiassa käytetystä. Englannin lisäksi myös esimerkiksi ranskaa, espanjaa, saksaa ja portugalia puhutaan omilla aksenteillaan useassa eri maassa. Osa virtuaaliavustajista on ottanut nämä erot huomioon. Lisäksi kaikille kielille ei ole saatavissa virtuaaliavustajaa, jolloin osa käyttäjistä joutuu käyttämään virtuaaliavustajaa, jollain muulla kuin omalla äidinkielellään. Esimerkiksi suomalaisilla on tämä on-

gelma joidenkin ääniohjattavien virtuaaliavustajien kuten Amazon Alexan kanssa.

Erilaiset aksentit ovat haastavia puhekomennoilla toimivien virtuaaliavustajien puheentunnistukselle. Palanican, Thommandramin, Leen, Lin, ja Fossatin vuonna 2019 tekemässä tutkimuksessa verrattiin, miten eri ääniohjattavat virtuaaliavustajat ymmärtävät lääkkeiden nimiä, kun käyttäjällä on aksentti. Tutkimukseen osallistui sekä englantia äidinkielenään puhuvia että niitä, joiden äidinkieli ei ole englanti. Ääniohjattavien virtuaaliavustajien välillä oli isoja eroja siinä, miten ne toimivat aksenttien kanssa. Google Assistant kykenee editoimaan pois tarpeettomia välisanoja ja änkytyksen, kun taas Applen Siri ja Amazon Alexa eivät muokanneet käyttäjän puhetta. Google Assistant pystyi myös ymmärtämään väärin lausuttuja sanoja. Alexalla oli lyhyin vastausaika, mutta samalla huonoin kokonaisuus. Alexa ei aina antanut käyttäjän lopettaa käskyään ennen vastaamista ja käytti myös tiedonhakuun vähän aikaa. Alexalla oli myös vaikeuksia ymmärtää käyttäjän puhetta, jos sanan aikana tuli tauko. Aina tauon tullessa se oletti uuden sanan alkavan ja tällöin jakoi yhden sanan useammaksi sanaksi, mikä johti väärinymmärryksiin. Lisäksi Alexa vaati Siriä ja Google Assistantia pidemmän käskyn toimiakseen. (Palanica, Thommandram, Lee, Li, & Fossat, 2019.)

Ääniohjattavilla virtuaaliavustajilla on eroja siinä, miten ne ymmärtävät englannin kielen eri aksentteja. Amazon Alexa on panostanut monen eri kielen osaamisen sijaan englannin kielen eri aksenttien ja murteiden ymmärtämiseen eri puolella maailmaa esimerkiksi Intiassa (Wiggers, 2019). Tämä kuitenkin on ristiriidassa edellisen kappaleen tutkimuksen kanssa. Erona voi kuitenkin olla se, että Alexa on kehittänyt palveluitaan niille, jotka puhuvat englantia äidinkielenään maasta riippumatta. Google Assistant ymmärtää englannin aksenteista huonointen skotlantilaista korostusta ja parhaiten Yhdysvaltojen eri aksentteja. Parhaiten se kuitenkin ymmärtää Kalifornian aksenttia, sillä Googlella on pääkonttori siellä. Lisäksi Google Assistant ymmärtää hyvin Uuden-Seelannin aksenttia. (Tatman, 2016.) Aksenttien ymmärtäminen riippuu myös paljon siitä, minkälaisia puheentunnistus tekniikoita käytetään ja minkälaista dataa virtuaaliavustajan koulutukseen on käytetty. Puheentunnistuksessa voi olla isotkin erot aksenttien välillä esimerkiksi Yhdysvaltojen eri alueilla, jos aksentteja ei oteta virtuaaliavustajan kehitysvaiheessa huomioon. (Rizwan & Anderson, 2018.)

Englantia äidinkielenään puhuvilla on yleisesti parempi käyttäjä kokemus kuin niillä, jotka käyttävät ääniohjattavaa virtuaaliavustajaa englanniksi, vaikka se ei olisi heidän äidinkieltään. Yksinkertaiset käskyt ääniohjattava virtuaaliavustaja ymmärsi yhtä hyvin huolimatta, siitä puhuiko käyttäjä omalla äidinkielellään vai käyttikö hän jotain muuta kieltä. Monimutkaisten käskyjen ymmärtäminen oli virtuaaliavustajille hankalampaa, kun käyttäjä ei puhunut omalla äidinkielellään. Vierasta kieltä puhuville ovat yleisiä esimerkiksi väärin lausutut sanat, jotka tuottivat vaikeuksia virtuaaliavustajien puheentunnistukselle. Lisäksi käyttäjät, jotka eivät puhuneet omalla äidinkielellään käyttivät normaalista poikkeavia sanavalintoja, joihin ääniohjattavat virtuaaliavustajat

eivät olleet tottuneet. Lisäksi vierasta kieltä puhuvien käskyjen rakenne erosi englantia äidinkielenään puhuvien käyttämistä lauseenrakenteista. Edellä mainitut ongelmat johtivat siihen, ettei ääniohjattava virtuaaliavustaja kyennyt aina ymmärtämään sellaisen käyttäjän puhetta, jonka äidinkieli ei ole englanti. (Pyae & Scifleet, 2018.)

4 Fyysisten ominaisuuksien vaikutus puheentunnistukseen

Käyttäjällä on myös fyysisiä melko pysyviä ominaisuuksia, jotka vaikuttavat heidän puheeseensa ja tätä kautta ääniohjattavan virtuaaliavustajan kykyyn ymmärtää käyttäjän puhetta. Näitä ominaisuuksia ovat esimerkiksi ikä, sukupuoli ja käyttäjän kyvykkyys. Osa näistä ominaisuuksista on myös sidoksissa toisiinsa, esimerkiksi iäkkäillä ihmisillä on todennäköisemmin puheeseen vaikuttavia rajoitteita, kuten dementiaa tai lihasten toimintaan liittyviä haasteita. Ääniohjattava virtuaaliavustaja voisi olla suuri apu iäkkäille ihmisille ja eri rajoitteista kärsiville ja auttaa heitä elämään itsenäisemmin. Tästä huolimatta heidät on usein jätetty virtuaaliavustajien kehityksessä vähemmälle huomiolle. Tässä kappaleessa käydään läpi, miten käyttäjän sukupuoli, ikä, änkyttäminen, puheviat ja kyvykkyys vaikuttavat ääniohjattavan virtuaaliavustajan kykyyn ymmärtää puhetta.

4.1 Sukupuolen vaikutus

Sukupuolen vaikutuksesta virtuaaliavustajien kykyyn ymmärtää käyttäjän puhetta on vaikea löytää tutkimusta. Ääniohjattavat virtuaaliavustajat hyödyntävät kuitenkin puheentunnistusteknologiaa, josta löytyy sukupuoleen liittyvää tutkimusta. Tämän kappaleen tutkimuksissa keskitytään käyttäjän sukupuolen vaikutukseen puheentunnistuksessa. Sukupuoli on ollut puheentunnistuksen ongelma alusta alkaen. Eroja voi olla molempiin suuntiin. Erot sukupuolten välillä ei ole tarkoituksella luotuja vaan ovat syntyneet vahingossa teknologiaa kehittäessä. (Howard & Borenstein, 2018.)

Usein puheentunnistus ymmärtää miehiä naisia paremmin. Ero johtuu osittain siitä, että suurin osa alan työntekijöistä on miehiä. Lisäksi monessa tekoälyn kouluttamiseen käytetyssä tietokannassa on epätasapaino sukupuolten välillä niin, että suurin osa tietokannasta koostuu miesten puheesta. Tällöin on täysin odotettavissa, että tekoäly oppii ymmärtämään miehiä naisia paremmin.

(Tatman, 2016.) Tutkimuksissa on havaittu, että ääniteknologia ymmärtää miesten puhetta paremmin ja tätä kautta teknologian toimivuus on parempi miehillä ja siksi heidän on myös mieluisampi sitä käyttää. (Howard & Borenstein, 2018.) Erot johtuvat myös siitä, että naisten puhe eroaa yleensä miesten puheesta. Naiset puhuvat usein hieman miehiä hitaammin. Naiset myös puhuvat usein miehiä muodollisemmin ja käytettyjen täytesanojen määrät eroavat. Naisten vokaalit ovat miesten vokaaleita pidempiä. Yksi keskeinen tekijä, mikä vaikuttaa myös puheen ymmärrettävyyteen on äänenvoimakkuus. Naisilla on tapana puhua miehiä hiljaisemmalla äänellä, mikä myös hankaloittaa virtuaaliavustajan kykyä ymmärtää heidän puhettaan. (Tatman, 2016.)

Asia ei kuitenkaan ole näin yksinkertainen vaan on myös olemassa tutkimuksia, joiden mukaan tekoäly ymmärtää naisten puhetta paremmin kuin miesten puhetta. Goldwaterin ym. (2008) tutkimuksessa puheentunnistus systeemit tekivät enemmän virheitä miesten puheen ymmärtämisessä kuin naisten. Tähän vaikuttaa se, että naisten ja miesten puhe eroaa toisistaan, kuten edellisessä kappaleessa kävi ilmi. Miehet käyttivät naisiin verrattuna huolimattomaa kieltä ja puhuivat nopeammin, jolloin puheentunnistuksella oli vaikeuksia heidän puheensa ymmärtämisessä. Puheentunnistuksen kykyä ymmärtää puhetta paransivat myös naisille yleisemmät pitkät vokaalit. Miesten lyhyet vokaalit olivat puheentunnistukselle hankalia, mikä vaikutti puheentunnistuksessa tapahtuvien virheiden määrää. (Goldwater, Jurafsky, Manning & Christopher D., 2008.) Tämä tutkimus ja osa edellisen kappaleen tutkimuksissa käytetyistä lähteistä ovat kuitenkin teknologian kehityksen takia hieman vanhoja ja nykytilanne saattaa olla hyvinkin erilainen.

Toisaalta voidaan myös pohtia, paljonko vaikutusta sukupuolella enää tänä päivänä on virtuaaliavustajan kykyyn ymmärtää puhetta ja voisivatko sukupuolien väliset erot johtua muista syistä. Kwoonin, Kimin ja Choehin (2016) tekemässä tutkimuksessa puheentunnistus ymmärsi nuoria naisia paremmin kuin nuoria miehiä, mutta vanhempia naisia vanhempia miehiä huonommin. Tutkijat eivät tässä tapauksessa pitäneet sukupuolta merkitseväenä tekijänä puheentunnistuksessa ja selittivät erojen johtuneen muista tekijöistä. (Kwoon, Kim & Choeh, 2016.) Lisäksi sukupuolten erot puheessa ovat yleistyksiä. Tutkimuksista voidaan kuitenkin päätellä, että sukupuoli vaikuttaa virtuaaliavustajan puheentunnistukseen, jos siihen ei kiinnitetä huomiota kehitysvaiheessa. Puheentunnistuksessa virhettä voi tapahtua molempiin suuntiin, vaikka sitä ei olisi tarkoitettu. Ääniohjattavien virtuaaliavustajien täytyy pystyä huomioimaan sukupuolesta johtuvat erot, jotta tasa-arvo toteutuisi niiden kohdalla mahdollisimman hyvin

4.2 Iän vaikutus

Ihmisen puhe muuttuu iän myötä ja siihen vaikuttavat sekä biologiset tekijät että oppiminen. On siis täysin odotettavissa, että käyttäjän ikä vaikuttaa ääniohjattavan virtuaaliavustajan kykyyn ymmärtää käyttäjän puhetta. Ääniohjattavat

virtuaaliavustajat on luotu ymmärtämään ensisijaisesti keskiverto aikuisen puheetta. Aikuisen puhe eroaa kuitenkin huomattavasti lasten ja iäkkäiden ihmisten puheesta, jolloin myös virtuaaliavustajat ymmärtävät lapsia ja vanhuksia heikommin.

Ero puheentunnistuksen kyvyssä ymmärtää ikäihmisiä verrattuna nuoriin aikuisiin on huomattava. Nuorten aikuisten puheesta puheentunnistus ymmärsi 94 % ja iäkkäiden ihmisten puheesta vain 74 %. Tämä ero koostuu useasta eri syystä. Iäkkäiden ihmisten puhe on hitaampaa ja heillä on erilainen puhetyyli ja artikulaatio kuin nuorilla aikuisilla. Erityisesti vanhemmilla naisilla on usein pitkiä tavujen sisäisiä taukoja, jotka vaikeuttivat puheentunnistusta. Iän myötä biologiset muutokset vaikuttavat äänisignaaliin ja tätä kautta myös ääntämiseen. Tutkimuksen mukaan nuorilla aikuisilla puhe oli yleisesti myös helpommin ymmärrettävää kuin ikäihmisillä. (Kwon, Kim, & Choeh, 2016.) Lisäksi iäkkäät ihmiset eivät välttämättä ole kovin tottuneita käyttämään ääniteknologiaa tai teknologiaa ylipäättänsä, mikä voi vaikuttaa käyttöä rajoittavana tekijänä. Osa iäkkäistä ihmisistä etenkin he, joilla on muistivaikeuksia, puhuivat virtuaaliavustajalle niin kuin toiselle ihmiselle, mikä osaltaan vaikeutti virtuaaliavustajan puheentunnistusta. (Wolters, Kelly, & Kilgour, 2016.)

Ääniohjattavalla virtuaaliavustajalla oli myös vaikeuksia ymmärtää lasten puhetta. Virtuaaliavustaja ymmärsi lapsia paremmin, kun vanhemmat olivat ensin opastaneet lapselle, miten virtuaaliavustajalle pitää puhua. Esimerkiksi vanhemmat opastivat lapsia esittämään yksinkertaisempia kysymyksiä ja puhumaan mahdollisimman selkeästi. (Beirl ym., 2019.) Ääniohjattavien virtuaaliavustajien ongelma on myös se, ettei se erota lapsen puhetta aikuisen puheesta. Tällöin se ei esimerkiksi pysty rajoittamaan sitä, mitä lapsi saa tehdä. Lapsi voi esimerkiksi käyttää virtuaaliavustajaa tilaamaan haluamansa lelun ilman vanhempien suostumusta. (Karppi & Granata, 2019.) Lapsilla on myös tapana käyttää non-verbaalia viestintää, mitä ääniohjattava virtuaaliavustaja ei voi ymmärtää. Esimerkiksi lapsi voi osoittaa jotakin ja haluta tietää, mikä se on. Lapsilla on myös aikuisia enemmän erilaisia puhevikoja, joiden takia puhetta on vaikea ymmärtää, jolloin virtuaaliavustaja ei pysty suorittamaan haluttua toimintaa. (Gordon & Breazeal, 2015.) Virtuaaliavustaja ei usein ymmärtänyt lapsen käskyjä tai kysymyksiä. Se ei myöskään ymmärtänyt lapsien käyttämiä strategioita, mitkä saattavat auttaa aikuisia ymmärtämään lapsia. Osa lapsista esimerkiksi korotti ääntään, piti taukoja sanojen välillä ja osa uhkaili virtuaaliavustajaa. Myös se, että lapsi saa puhua virtuaaliavustajalle, miten hän haluaa ja virtuaaliavustaja suorittaa halutun toiminnon on ongelmallista. Tällöin lapsi voi oppia esimerkiksi puhumaan ei-toivotulla tavalla. (Lovato & Piper, 2015.)

4.3 Änkyttäminen ja puheviat

Ääniohjattavilla virtuaaliavustajilla on myös vaikeuksia ymmärtää änkyttämistä. Änkyttämisestä kärsivät pystyvät käyttämään virtuaaliavustajaa, mutta usein heidän täytyy toistaa antamansa käsky useaan kertaan. (Corcoran, 2018.)

Änkyttäminen on puuuhäiriö, jolle on ominaista äänien, tavujen tai sanojen toisto sekä äänneiden pidentyminen ja puheen keskeytykset. Yksi tekijä, mikä vaikeuttaa virtuaaliavustajien puheentunnistusta on puheeseen tulevat tauot, mitkä ovat änkytyksessä yleisiä. Lisäksi saman äänähdyksen tai tavun toistaminen ovat hämmentäviä virtuaaliavustajan puheentunnistukselle. Ääniohjattavien virtuaaliavustajien nimet, jonka sanominen vaaditaan sen aktivoimiseen, sisältävät usein kirjaimia, jotka ovat vaikeita sanoa. Virtuaaliavustajien koulutuksessa harvoin käytetään änkyttämistä sisältävää puhetta. Asiaa vaikeuttaa se, että änkyttäminen voi olla hyvinkin erilaista käyttäjien välillä tai samalla käyttäjällä päivästä riippuen. (Wheeler, 2018.) Lisäksi virtuaaliavustajien välillä on eroja, esimerkiksi Google Assistant pyrkii editoimaan puhetta niin, että siitä poistetaan kaikki ylimääräinen. Alexa ja Siri eivät puolestaan muokkaa käyttäjän puhetta mitenkään. (Palanica ym., 2019.)

Puheviat ovat melko yleisiä erityisesti lapsilla ja erilaisista rajoitteista kärsivillä. Niitä saattaa olla kuitenkin myös aivan tavallisilla aikuisilla. Useamman tutkimuksen mukaan puheviat vaikeuttavat ääniohjattavan virtuaaliavustajan puheentunnistusta ja tätä kautta sen toimintaa. (Gordon & Breazeal, 2015; Pradhan ym., 2018) Tutkimusta, joka keskittyisi pelkästään virtuaaliavustajan kykyyn ymmärtää puhevikoja on vaikea löytää. On kuitenkin olemassa useita eri tapoja, joilla voidaan parantaa virtuaaliavustajan kykyä ymmärtää käyttäjää. Esimerkiksi jos käyttäjällä on puhevikoja tai muita puheeseen liittyviä vaikeuksia, on mahdollista käyttää eri aktivointisanaa tai käyttää näppäimiä apuna. (McFedries, 2020.)

4.4 Puheeseen vaikuttava kyvykkyys

Tässä kappaleessa käydään läpi käyttäjän kyvykkyuden vaikutusta ääniohjattavan virtuaaliavustajan kykyyn ymmärtää käyttäjän puhetta. Ääniteknologiasta voisi olla paljon hyötyä erilaisista rajoitteista kärsiville, sillä sen käyttämiseen ei tarvitse näkökykyä, motorisia taitoja tai kirjoittamista. Tästä huolimatta kyvykkyys on jätetty vähemmälle huomiolle virtuaaliavustajien kehityksessä. Puheeseen vaikuttava kyvykkyys voi koskea kognitiivisia toimintoja, leuan ja kielen lihaksia, hermoston toimintaa ja aivoalueiden toimintaa. On kuitenkin tärkeä muistaa, että eri rajoitteet voivat ilmetä eri tavoin ja eri voimakkuudella, jolloin tutkimuksista saadut tulokset eivät välttämättä sovellu kaikille. Lisäksi monet tekijät kuten puheviat ja äänenvoimakkuus voivat olla haasteita myös täysin terveillekin henkilöille.

4.4.1 Kehitysvammaisuus

Tässä kappaleessa keskitytään ääniohjattavien virtuaaliavustajien kykyyn ymmärtää kehitysvammaisten puhetta. Kehitysvammaisilla on usein vaikeuksia mm. muistin, ongelmanratkaisun, suunnittelun, monimutkaisten konseptien ymmärtämisen ja ajatteluprosessien kanssa. Kehitysvammoja ovat mm. downin

syndrooma ja autismi. Kehitysvammoihin voi usein liittyä myös muita rajoitteita, kuten puutteelliset motoriset taidot. (Nour, 2015.) Äänellä toimiva teknologia auttaa kehitysvammaisia ylittämään kirjoituksesta koituvat esteet ja suorittamaan erilaisia tehtäviä itsenäisesti (Pradhan ym., 2018).

Kehitysvammaisten puhe on ääniohjattaville virtuaaliavustajille vaikeaa ymmärtää. Teknologian kehityksessä harvoin huomioidaan käyttäjien kyvykkyys. Balasuriyan, Sitbonin, Bayorin, Hoogstraten, & Breretonin (2018) tekemässä tutkimuksessa kognitiivisesti rajoittuneiden käyttäjien piti suorittaa neljä melko yksinkertaista tehtävää esimerkiksi virtuaaliavustajan aktivointi ja verkkoahan suorittaminen. Vain puolet osallistuneista pystyivät suorittamaan ne kaikki ja noin 22 % osallistuneista eivät onnistuneet yhdenkään tehtävän suorittamisessa, koska virtuaaliavustaja ei pystynyt ymmärtämään heidän puhettaan. Yleisesti tutkimukseen osallistuneet pitivät virtuaaliavustajasta, vaikka se ei antanutkaan haluttuja tuloksia. (Balasuriya ym., 2018.) Toisessa tutkimuksessa vain 25 % onnistui suorittamaan tehtävän Googlen puheentunnistuksella, vaikka se olikin kirjoittamista nopeampi. Artikkelissa tuodaan esille, että ongelmat johtuivat siitä, että puheentunnistus keskittyi sanojen ääntämisen oikeellisuuteen, mistä vähäinen onnistuminen johtui. Toisaalta tämä ei välttämättä johdu pelkästään käyttäjän kyvykkyydestä vaan kaikki, joilla on ongelmia ääntämisen kanssa, olisivat todennäköisesti epäonnistuneet tehtävässä. (Rocha, Carvalho, Bessa, Reis, & Magalhães, 2017.)

Virtuaaliavustajalle pitää myös puhua tarpeeksi kovalla äänellä, jotta se ymmärtää, mitä tarkoitetaan. Tämä oli kuitenkin hankalaa joillekin kehitysvammaisille. Monella kehitysvammaisella on myös erilaisia puhevikoja ja vaikeuksia lausua kaikki sanat oikein, jolloin virtuaaliavustaja ei ymmärtänyt heitä. (Pradhan ym., 2018.) Yleensä käyttäjät, joilla on kehitysvamma käyttävät puhuessaan pitkiä kohteliaita lauseita ja odottivat myös ääniohjattavan virtuaaliavustajan puhuvan samalla tavalla. Pitkät lauseet osoittautuivat kuitenkin hankaliksi virtuaaliavustajille, sillä se ei ymmärtänyt, milloin käyttäjä on valmis vaan lopetti kuuntelemisen heti tauon tullessa. Tässä tilanteessa osa käskystä jäi pois, jolloin virtuaaliavustaja ei pystynyt suorittamaan käyttäjän haluamaa toimintoa. Monella meni myös aikaa muotoilla käsky, jolloin virtuaaliavustaja oli lopettanut kuuntelemisen ennen kuin käyttäjä oli aloittanut puhumisen. Google Assistant ei myöskään ilmoittanut selvästi ymmärsikö se annetun käskyn, joka aiheutti hämmennystä. (Balasuriya ym., 2018.)

Toisaalta se, että virtuaaliavustajalle pitää puhua tietyllä tavalla esimerkiksi kuuluvalla äänellä ja selkeästi saattaa myös kehittää käyttäjän puhetaitoja. Virtuaaliavustajaa voidaan myös käyttää eräänlaisena puheterapeutina. Osa käyttäjistä myös koki, että virtuaaliavustaja ymmärtää heitä paremmin kuin muut ihmiset. (Pradhan ym., 2018.) Vaikka käyttäjän puhetaidot kehittyivät, tämä ei kuitenkaan poista sitä, että virtuaaliavustajan viat rajoittavat sitä, miten niitä voidaan käyttää. Esimerkiksi jos käyttäjän pitää kutsua paikalle apua täytyy virtuaaliavustajan pystyä toimimaan. Pitää kuitenkin muistaa, että kehitysvammaisten kognitiiviset kyvyt vaihtelevat paljon ja vaikuttavat eri henkilöihin eri tavalla (Nour, 2015). Tällöin virtuaaliavustaja ymmärtää hyvin joitain käyt-

täjiään ja osaa se ei ymmärrä ollenkaan. Osa käyttäjistä voi hyötyä siitä, että virtuaaliavustajalle pitää puhua tietyllä tavalla ja osalta se estää virtuaaliavustajan käytön.

4.4.2 Fyysinen kyvykkyys

Kaikki puheeseen vaikuttava kyvykkyys ei koske kognitiivisia toimintoja. Rajoitukset voivat liittyä esimerkiksi lihasten, hermoston, aistien ja aivoalueiden toimintaan. Näissäkin tapauksissa pitää muistaa, että rajoitteet vaikuttavat eri tavalla eri henkilöön ja mihin henkilö pystyy, vaihtelee vamman vakavuuden mukaan. Puhevaikeudet voivat esimerkiksi johtua Parkinsonin taudista tai aivohalvauksesta. Usein tällaisissa tapauksissa myös henkilön motoriset taidot ja liikuntakyky ovat rajoittuneita. Henkilö, jolla ei ole kognitiivisia rajoitteita käyttää normaalia kieltä ja kielioppia, joihin virtuaaliavustajat ovat tottuneet. Ongelmana voi olla esimerkiksi se, että leuan ja kielen lihakset eivät toimi kunnolla, jolloin puhetta on vaikeaa ymmärtää. Tällöin puhujan konsonantit ovat epäselviä ja vokaalit sulautuvat yhteen, jolloin virtuaaliavustaja ei ymmärrä käyttäjän puhetta. (Corcoran, 2018.)

Myös käyttäjän kuulo vaikuttaa siihen, miten ääniohjattava virtuaaliavustaja ymmärtää käyttäjän puhetta. Käyttäjät, joilla on heikko kuulo käyttävät usein tavallisesta poikkeavia puhetyylejä. Henkilöille, joilla on kuulovaikeuksia tai on kuuro, ovat yleistä fonologiset ja artikulointi virheet. He myös puhuvat normaalia hitaammin ja venyttävät vokaaleita. Nämä tekijät yhdistettynä puheen poikkeaviin taukoihin tekevät puheesta vaikeamman virtuaaliavustajan ymmärtää. Toisaalta myös ihmisillä on usein vaikeuksia kuurojen puheen ymmärtämisessä. (Fok, Kaur, Palani, Mott, & Lasecki, 2018.)

Lihasten toiminnan ja kuulon lisäksi myös hermostolla on vaikutusta käyttäjän puheeseen. Dysartria tarkoittaa ryhmää neurogeenisia puhehäiriöitä, jotka voivat olla synnynnäisiä tai myöhemmällä iällä saatu joko onnettomuuden tai sairauden seurauksena. Ominaista henkilölle, jolla on dysartria on poikkeavuudet hengityksen, artikulaation ja puheentuottamisen nopeudessa, voimassa, tahdissa ja tasaisuudessa. Taustalla voi olla yksi tai useampi sensorimotorinen ongelma, jolloin äänentuottamiseen lähetetty käsky vääristyy. Lisäksi puhe voi olla hiljaista tai kova äänistä, epäselvää, nasaalia, epävarmaa, epärytmikästä tai monotonista. Nämä vaikuttavat käyttäjän puheeseen ja tätä kautta ääniohjattavan virtuaaliavustajan puheentunnistukseen. Virtuaaliavustajien kyvyt ymmärtää käyttäjiä, joilla on dysartria vaihtelevat paljon tuotteiden välillä. Google Assistant ymmärsi noin 62 % käyttäjien, jolla on dysartia kysymyksistä. Cortana ymmärsi 39.7 % ja Siri vain noin 16 %. (Ballati, Corno, & De Russis, 2018.)

Afasia vaikeuttaa ääniohjattavan virtuaaliavustajan puheentunnistusta. Afasia-oireet vaihtelevat puhujien välillä, jollain voi olla esimerkiksi vaikeuksia löytää oikea sana. Yleistä on myös foneettiset ja artikulaatio virheet sekä epä johdonmukainen puhe. Afasia johtuu aivovammasta ja sen on voinut saada aikaan esimerkiksi aivohalvaus tai erilaiset aivovammat. Ääniohjattavaa virtuaali-

liavustajaa voidaan käyttää hyödyksi afasiasta kärsivien puheterapiassa. Vaikka puheentunnistuksessa on pyritty kehittämään keinoja parantaa henkilöiden, joilla on afasia käyttäjäkokemusta, tapahtuu puheentunnistuksessa silti paljon virheitä. Tutkimuksissa on eroja, mitkä johtuvat myös käytetyistä puheentunnistusmalleista ja puheentunnistuksessa tapahtuneiden virheiden määrä on vaihdellut aiheesta tehdyissä tutkimuksista 21 %:n ja 58.2 %:n välillä. (Jamal, Shanta, Mahmud, & Sha'abani, 2017.) Edellä mainittujen lisäksi on muitakin kyvykkyyteen liittyviä rajoitteita, jotka vaikuttavat henkilön puheeseen. Näiden vaikutuksesta virtuaaliavustajan puheentunnistukseen ei kuitenkaan löytynyt tutkimusta.

5 Yhteenveto

Monet käyttäjän ominaisuudet vaikuttavat ääniohjattavan virtuaaliavustajan kykyyn ymmärtää puhetta. Tutkielman tutkimuskysymys oli, miten käyttäjän eri ominaisuudet vaikuttavat ääniohjattavan virtuaaliavustajan kykyyn ymmärtää puhetta. Tämän tutkielman perusteella tähän voidaan vastata, että käyttäjän kieli, ikä ja kyvykkyys sekä joissain tapauksissa sukupuoli vaikuttavat virtuaaliavustajan puheentunnistukseen.

Kieli on iso tekijä siinä, miten ääniohjattava virtuaaliavustaja ymmärtää käyttäjänsä. Virtuaaliavustajien kielitaito vaihtelee paljon tuotteiden välillä. Kielitaitonsa avulla virtuaaliavustaja pystyy erottumaan toisista virtuaaliavustajista. Niiden kehityksessä täytyy myös ottaa huomioon saman kielen erot eri alueilla sekä tilanteet, joissa käyttäjä ei puhu omalla äidinkielellään. Saman kielen eri aksenttien ymmärtämiseen vaikutti paljon se, millaista dataa virtuaaliavustajan koulutukseen oli käytetty. Kun käyttäjä ei puhunut omalla äidinkielellään puheentunnistusta vaikeuttivat käyttäjän sanavalinnat, virheellinen sanajärjestys, puheeseen tulleet tauot ja väärin lausutut sanat.

Virtuaaliavustajat on ensisijaisesti luotu ymmärtämään tavallisen aikuisen puhetta, jolloin lasten ja iäkkäiden ihmisten puhe tuottaa niille vaikeuksia. Iäkkäiden ihmisten puhe on hitaampaa, heillä on erilainen puhetyyli ja artikulaatio sekä pitkiä tavujen sisäisiä taukoja. Myös vähäinen kokemus teknologiasta vaikeuttaa sen käyttöä, ja osa iäkkäistä käyttäjistä puhuivat virtuaaliavustajalle kuin normaalille keskustelukumppanille. Lapsilla puolestaan ovat yleisiä puheviat. Lapset myös käyttivät non-verbaalia viestintää sekä kysyivät monimutkaisia kysymyksiä. Sukupuoli vaikutti puheentunnistukseen. Virhettä saattoi tapahtua molempiin suuntiin. Tämä johtui osittain siitä, että virtuaaliavustajan koulutukseen oli käytetty epätasapainoisesti jakautuvaa dataa. Erot myös johtuivat siitä, että naiset puhuivat miehiä muodollisemmin, hitaammin, hiljaisemmalla äänellä ja heidän vokaalinsa olivat pidempiä, mikä vaikuttaa puheentunnistukseen.

Käyttäjällä voi olla rajoituksia, jotka vaikuttavat puhumiseen. Rajoitukset voivat koskea kognitiivisia kykyjä, aisteja, aivoalueiden toimintaa, lihaksia ja hermostoa. Kyvykkyys voi aiheuttaa vaikeuksia sanojen oikein lausumisessa,

äänenvoimakkuudessa sekä aiheuttaa erilaisia puhevikoja tai -vaikeuksia. Yleisiä olivat myös pitkät kohteliaat lauseet, epäselvä puhe ja poikkeava puherytmi. Nämä kaikki vaikuttavat ääniohjattavan virtuaaliavustajan puheentunnistukseen. Myös änkyttäminen oli haasteellinen puheentunnistukselle.

Monessa kappaleessa tuli esille, että virheet puheentunnistuksessa johtuvat siitä, ettei esim. sukupuolta, ikää ja kyvykkyyttä ole otettu huomioon virtuaaliavustajan kouluttamisessa. Virtuaaliavustajien välillä oli eroja, mutta usein niille tuottivat vaikeuksia täytesanat, tauot, puheen hiljaisuus ja väärin sanotut sanat. Tässä tutkielmassa koottiin yhteen, mitkä asiat vaikeuttavat virtuaaliavustajan puheentunnistusta ja mitä pitäisi ottaa huomioon, kun tuotetta kehitetään paremmaksi. Moni kehittää ääniteknologiaa virtuaaliavustajien lisäksi, joten tutkielmaan kootuista asioista voisi olla hyötyä myös heille.

Ääniohjattavat virtuaaliavustajat ovat kuitenkin nopeasti kehittyvä teknologia, joten osa tuloksista saattaa vanheta nopeasti. Tästä johtuen jo pari vuotta vanha tutkimus saattaa sisältää tänä päivänä virheellistä tietoa. Osa tutkielmassa käytetyistä artikkeleista tuovat esiin tutkimuksen vähäisyyden. Kaikesta ei löytynyt tieteellistä tutkimusta, jolloin jouduttiin turvautumaan myös ei-tieteellisiin lähteisiin. Lisäksi kaikki tutkimukset eivät suoraan koskeneet virtuaaliavustajia, vaan osa keskittyi siinä hyödynnettävään puheentunnistusteknologiaan. Virtuaaliavustajat eroavat toisistaan ja ne suoriutuvat eri tavalla samassa tilanteessa, mikä hankaloittaa tulosten yleistämistä. Tulevaisuudessa pitäisi keskittyä siihen, miten virtuaaliavustajia voidaan parantaa niin, että se toimisi kaikille mahdollisimman hyvin. Lisäksi kyvykkyys on laaja alue ja kaikista puheetta rajoittavista tekijöistä ei löytynyt tutkimusta tai se oli vähäistä tai vanhaa.

LÄHTEET

- Abdolrahmani, A., Kuber, R., & Branham, S. M. (2018). " Siri talks at you" an empirical investigation of voice-activated personal assistant (VAPA) usage by individuals who are blind. *Paper presented at the Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, 249-258.
- Amazon & Barr, J. (2016). AWS news blog. Amazon Polly – Text to Speech in 47 Voices and 24 Languages. Haettu osoitteesta <https://aws.amazon.com/blogs/aws/polly-text-to-speech-in-47-voices-and-24-languages/>
- Balasuriya, S. S., Sitbon, L., Bayor, A. A., Hoogstrate, M., & Brereton, M. (2018). Use of voice activated interfaces by people with intellectual disability. *Paper presented at the Proceedings of the 30th Australian Conference on Computer-Human Interaction*, 102-112.
- Ballati, F., Corno, F., & De Russis, L. (2018). Assessing virtual assistant capabilities with italian dysarthric speech. Paper presented at the *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, 93-101.
- Beirl, D., Yuill, N., & Rogers, Y. (2019). Using voice assistant skills in family life.
- Corcoran, M.(2018). SLATE. When Alexa can't understand you. Haettu osoitteesta <https://slate.com/technology/2018/10/voice-assistants-alexasiri-speech-disabilities-recognition.html>
- Franceschin, T. (2017). Edy4me. Voice assistants can become powerful allies in the classroom. Haettu osoitteesta <http://edu4.me/voice-assistants-can-become-powerful-allies-in-the-classroom/>
- Fok, R., Kaur, H., Palani, S., Mott, M. E., & Lasecki, W. S. (2018). Towards more robust speech interactions for deaf and hard of hearing users. *Paper presented at the Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*, 57-67.
- Global. (2020). Language Support in Voice Assistants Compared (2019 Update). Haettu osoitteesta https://www.globalme.net/blog/language-support-voice-assistants-compared#Why_is_teaching_voice_assistants_a_new_language_so_hard

- Goldwater, Sharon Jurafsky, Dan Manning, Christopher D. (2008). Which words are hard to recognize? prosodic, lexical, and disfluency factors that increase {ASR} error rates. *Speech Communication* 52 2010, 181-200.
- Gordon, M., & Breazeal, C. (2015). Designing a virtual assistant for in-car child entertainment. *Paper presented at the Proceedings of the 14th International Conference on Interaction Design and Children*, 359-362.
- Hauswald, J., Laurenzano, M. A., Zhang, Y., Yang, H., Kang, Y., Li, C., & Mudge, T. (2016). Designing future warehouse-scale computers for sirius, an end-to-end voice and vision personal assistant. *ACM Transactions on Computer Systems (TOCS)*, 34(1), 1-32.
- Howard, A., & Borenstein, J. (2018). The ugly truth about ourselves and our robot creations: The problem of bias and social inequity. *Science and Engineering Ethics*, 24(5), 1521-1536.
- Hoy, M. (2018). Alexa, siri, cortana, and more: An introduction to voice assistants. *Medical Reference Services Quarterly*, 37 (1) (2018), 81-88.
Retrieved from
- Jamal, N., Shanta, S., Mahmud, F., & Sha'abani, M. (2017). Automatic speech recognition (ASR) based approach for speech therapy of aphasic patients: A review. *Paper presented at the AIP Conference Proceedings*, 1883(1) 020028.
- Juniper Research. (2018). Press release. Voice assistants used in smart homes to grow 1,000%, reaching 275 million by 2023, as Alexa leads the way. Haettu osoitteesta <https://www.juniperresearch.com/press/press-releases/voice-assistants-in-smart-homes-reach-275m-2023>
- Karkio, A. (2019). Googlen apuriohjelma oppi ymmärtämään puhuttua suomea: ymmärrys toimii tosin hitaasti. Mikrobitti. Haettu osoitteesta <https://www.mikrobitti.fi/uutiset/googlen-apuriohjelma-oppi-ymmartamaan-puhuttua-suomea-ymmarrys-toimii-tosin-hitaasti/c67d88b9-65cf-48dd-b315-a13f95f66e30>
- Karppi, T., & Granata, Y. (2019). Non-artificial non-intelligence: Amazon's alexa and the frictions of AI. *AI & Society*, 34(4), 867-876
- Krakovsky, M. (2018). Artificial (emotional) intelligence. *Communications of the ACM*, 61(4), 18-19.
- Kwon, S., Kim, S., & Choeh, J. Y. (2016). Preprocessing for elderly speech recognition of smart devices. *Computer Speech & Language*, 36, 110-121.
- Liu, G., & Zhang, H. B. (2014). Voice assistant – application of speech recognition technology in the android system. *Applied Mechanics and Materials*, 596(Mechatronics and Industrial Informatics II), 384-387.

- Lovato, S., & Piper, A. M. (2015). "Siri, is this you?" understanding young children's interactions with voice input systems. *Paper presented at the Proceedings of the 14th International Conference on Interaction Design and Children*, 335-338.
- Lugano, G. (2017). Virtual assistants and self-driving cars. *Paper presented at the 2017 15th International Conference on ITS Telecommunications (ITST)*, 1-5.
- McFedries, P. (28.4.2020). Dummies. How to Make Amazon Alexa More Speech Accessible. Haettu osoitteesta <https://www.dummies.com/consumer-electronics/smart-devices/how-to-make-amazon-alexa-more-speech-accessible/>
- Martinelli, E., Mencattini, A., Daprati, E., & Di Natale, C. (2016). Strength is in numbers: Can concordant artificial listeners improve prediction of emotion from speech?(research article). *PLoS ONE*, 11(8), e0161752.
- Massaro, T. M., & Norton, H. (2016). Sir-iously? free speech rights and artificial intelligence.(free speech foundations symposium). *Northwestern University Law Review*, 110(5).
- Mutchler, A. (2017). Voicebot.ai. Voice Assistant Timeline: A Short History of the Voice Revolution. Haettu osoitteesta <https://voicebot.ai/2017/07/14/timeline-voice-assistants-short-history-voice-revolution/>
- Nour, R. (2015). Web searching by individuals with cognitive disabilities. *ACM SIGACCESS Accessibility and Computing*, (111), 19-25.
- Pal, S., & Bhowmick, S. (2017). Evaluation framework for disabled students based on speech recognition technology. *International Journal of Modern Education and Computer Science*, 09(10), 10.
- Palanica, A., Thommandram, A., Lee, A., Li, M., & Fossat, Y. (2019). Do you understand the words that are comin outta my mouth? voice assistant comprehension of medication names. *NPJ Digital Medicine*, 2(1), 55.
- Polyakov, E. V., Mazhanov, M. S., Rolich, A. Y., Voskov, L. S., Kachalova, M. V., & Polyakov, S. V. (2018). Investigation and development of the intelligent voice assistant for the internet of things using machine learning. *Paper presented at the 2018 Moscow Workshop on Electronic and Networking Technologies (MWENT)*, 1-5.
- Pradhan, A., Mehta, K., & Findlater, L. (2018). "Accessibility came by accident" use of voice-controlled intelligent personal assistants by people with disabilities. *Paper presented at the Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1-13.

- Pyae, A., & Scifleet, P. (2018). Investigating differences between native english and non-native english speakers in interacting with a voice user interface: A case of google home. *Paper presented at the Proceedings of the 30th Australian Conference on Computer-Human Interaction*, 548-553.
- Rizwan, M., & Anderson, D. V. (2018). A weighted accent classification using multiple words. *Neurocomputing*, 277, 120-128.
- Rocha, T., Carvalho, D., Bessa, M., Reis, S., & Magalhães, L. (2017). Usability evaluation of navigation tasks by people with intellectual disabilities: A google and SAPO comparative study regarding different interaction modalities. *Universal Access in the Information Society*, 16(3), 581-592. doi:10.1007/s10209-016-0489-5
- Rouse, M. (2016). TechTarget. Speech recognition. Haettu osoitteesta <https://searchcustomerexperience.techtarget.com/definition/speech-recognition>
- Statista (9.4.2020). Consumer electronics. Smart speaker with intelligent personal assistant market share in 2018 and 2019, by platform. Haettu osoitteesta <https://www.statista.com/statistics/1005558/worldwide-smart-speaker-market-share/>
- Sterling, G. (2019). Marketing Land. Alexa devices maintain 70% market share in U.S. according to survey. Haettu osoitteesta <https://marketingland.com/alexa-devices-maintain-70-market-share-in-u-s-according-to-survey-265180>
- Tatman, R. (12.7.2016). Making noise & hearing things. Google's speech recognition has a gender bias. Haettu osoitteesta <https://makingnoiseandhearingthings.com/2016/07/12/googles-speech-recognition-has-a-gender-bias/>
- Wheeler, K. (2018). Curbed. Why voice assistants don't understand people who stutter. Haettu osoitteesta <https://www.curbed.com/2018/12/12/18135195/alexa-google-home-tech-stuttering>
- Wiggers, K. (2019). Venturebeat. Which voice assistant speaks the most languages, and why?. Haettu osoitteesta <https://venturebeat.com/2019/02/02/which-voice-assistant-speaks-the-most-languages-and-why/>
- Wolters, M. K., Kelly, F., & Kilgour, J. (2016). Designing a spoken dialogue interface to an intelligent cognitive assistant for people with dementia. *Health Informatics Journal*, 22(4), 854-866.

Wu, K. (2018) Voice assistant. 11-21.