



Vaasan yliopisto  
UNIVERSITY OF VAASA

Teemu Hietikko

# **Puhesynteesi huoltovideolla arki- ja ammattikäyttäjien arvioimana**

Markkinoinnin ja viestinnän yksikkö  
Digitaalinen media  
Viestinnän monialainen maisteriohjelma

Vaasa 2019

---

**VAASAN YLIOPISTO****Markkinoinnin ja viestinnän yksikkö**

<b>Tekijä:</b>	Teemu Hietikko	
<b>Tutkielman nimi:</b>	Puhesynteesi huoltovideolla arki- ja ammattikäyttäjien arvioimana	
<b>Tutkinto:</b>	Filosofian maisteri	
<b>Oppiaine:</b>	Viestinnän monialainen maisteriohjelma	
<b>Työn ohjaaja:</b>	Merja Koskela	
<b>Valmistumisvuosi:</b>	2019	<b>Sivumäärä:</b> 114

---

**TIIVISTELMÄ:**

Tämän tutkielman tavoite on selvittää, miten ammattikäyttäjät ja arkikäyttäjät kokevat puhesynteessin käytön Oras Groupin termostaattihanan huoltoa ohjeistavalla videolla sekä arvioida tulosten perusteella puhesynteessin mahdollista käyttöönottoa ohjevideoissa tulevaisuudessa. Puhesynteessin keskeisinä arviointikriteereinä tutkielmassa toimivat luonnollisuus ja ymmärrettävyys. Nämä jaetaan alakäsitteisiin, joiden avulla puhesynteessin kokemista voidaan tarkastella yksityiskohtaisemmin. Vastauksia puhesynteessin käytöstä ohjevideolla saadaan kyselytutkimuksella, johon vastaavat ammatti- ja arkikäyttäjät.

Puhesynteessin tutkimuksessa ohjevideoiden kautta esiin nousevat kuuntelukokemuksen rooli, informatiivisuus, audiovisuaalisuus sekä onnistuneen viestinnän kriteerit informaation välityksessä. Tutkielmassa tarkastellaan puhesynteessin roolia informaation onnistuneena välittäjänä. Määrittäjinä onnistuneelle viestinnälle toimivat kyselytutkimukseen vastanneiden henkilöiden kokemukset puhesynteesistä.

Tutkielman päämenetelmänä käytetään kvantitatiivista kyselytutkimusta, jonka vastauksista lasketaan Mean Opinion Score (MOS). Huomioon otetaan myös vastaajien taustamuuttujat: ikä, sukupuoli ja englannin kielen taito, sekä itse puhesynteessin muuttujat: mies- ja naisääni. Näiden keskiarvoista suoritetaan lopuksi varianssianalyysi, jonka avulla tarkastellaan yhtäläisyyksiä ja eroavaisuuksia.

Kyselytutkimukseen vastasi yhteensä 146 henkilöä, 77 ammattikäyttäjien kyselyyn ja 69 arkikäyttäjien kyselyyn. Ammattikäyttäjät antoivat puhesynteesistä positiivisemmän arvion kuin arkikäyttäjät. Puhesynteessin ymmärrettävyys oli hyvällä ja luonnollisuus hyväksyttävällä tasolla. Vanhimmat ikäluokat arvioivat puhesynteessin paremmaksi kuin nuoremmat, ja naiset antoivat puhesynteesistä paremman arvion kuin miehet. Englannin kielen taito oli tärkeä tekijä puhesynteessin arvioinnissa. Huonon englannin taidon omaavat vastaajat arvioivat puhesynteessin heikomaksi kuin hyvän englannin kielen omanneet vastaajat. Myös puhesynteessin äänellä oli merkittävä vaikutus puhesynteessin kokemisessa. Miesääni miellettiin vastaajien mukaan naisääntä parempana, niin kuuntelukokemuksena kuin informaation välittäjänäkin.

Puhesynteesi sai hyvän arvion ammattikäyttäjien ja hyväksyttävän arvion arkikäyttäjien tuloksissa. Arkikäyttäjien arvion perusteella puhesynteessissä on vielä kehittämisen varaa ihmispuheen kaltaisuudessa sekä oikeaoppisessa puheessa ja puheen rytmittämisessä. Ammattikäyttäjät kokivat erityisesti puhesynteessin informatiivisuuden, kuuntelukokemuksen ja kuuntelun helppouden hyvänä. Puhesynteessin käyttöönotto ohjevideoissa on kyselyn tulosten perusteella mahdollista.

---

**AVAINSANAT:** puhesynteesi, kyselytutkimus, informaation välitys, ohjevideot, audiovisuaalisuus, tekninen viestintä

## Sisällys

1	Johdanto	7
1.1	Tutkimuksen tavoite	8
1.2	Tutkimusaineisto	10
1.3	Tutkimusmenetelmät	12
1.3.1	Kyselytutkimus	13
1.3.2	Mean opinion score (MOS)	17
2	Informaation välitystä onnistuneella viestinnällä	20
2.1	Viestintämallit	21
2.1.1	Shannonin ja Weaverin tiedonsiirron malli ja sen kritiikki	22
2.1.2	Semioottinen viestintämalli	24
2.2	Onnistuneen viestinnän mittarit	25
3	Tekniset käyttöohjeet ja ohjevideot	28
3.1	Teknisestä ohjeesta tekniseksi ohjevideoksi	28
3.1.1	Audiovisuaalisuus	29
3.1.2	Kuuntelukokemuksen rooli informaation välityksessä	30
3.2	Huolto- ja käyttöohjeet Oras Group Oy:ssä	31
3.3	Tekniset ohjevideot ja niiden kohderyhmät Oras Group Oy:ssä	33
4	Puhesynteesin käyttö ja arviointi	35
4.1	Puhesynteesijärjestelmät	35
4.2	Puhesynteesin käyttötavat	38
4.3	Puhesynteesi teknisissä ohjevideoissa Oras Group Oy:ssä	40
4.4	Puhesynteesin arvioinnin kriteerit	42
4.4.1	Luonnollisuus puhesynteesin kriteerinä	43
4.4.2	Ymmärrettävyys puhesynteesin kriteerinä	44
5	Kyselytutkimuksen tulokset	45
5.1	Ammattikäyttäjien MOS-tulokset	48
5.1.1	Miesääni ammattikäyttäjien arvioimana	60
5.1.2	Naisääni ammattikäyttäjien arvioimana	63

5.2 Arkikäyttäjien MOS-tulokset	66
5.2.1 Miesääni arkikäyttäjien arvioimana	78
5.2.2 Naisääni arkikäyttäjien arvioimana	82
5.3 Varianssianalyysin tulokset	85
5.3.1 Luonnollisuuden ja ymmärrettävyyden alakäsitteet	87
5.3.2 Mies- ja naisäänen vaikutus tuloksiin	90
6 Pohdinta	99
Lähteet	102
Liitteet	108
Liite 1. Kyselylomake : Arkikäyttäjät	108
Liite 2. Kyselylomake: Ammattikäyttäjät	113

## **Kuvat**

Kuva 1. Esimerkki hanan huolto-ohjeesta (Oras Ltd. 2019).	32
Kuva 2. Kuvankaappaus ohjevideosta (Oras Ltd. 2019).	33

## **Kuviot**

Kuvio 1. Kyselytutkimuksen prosessit	14
Kuvio 2. Muuttujat kyselytutkimuksessa	16
Kuvio 3. Kyselytutkimuksen käsitteet	18
Kuvio 4. Shannonin ja Weaverin viestinnän malli	23
Kuvio 5. Esimerkki MOS-asteikosta käytännössä	37
Kuvio 6. Kyselytutkimuksen käsitteet	42
Kuvio 7. Arkikäyttäjien englannin kielen taito	46
Kuvio 8. Ammattikäyttäjien englannin kielen taito	47
Kuvio 9. Ammattikäyttäjien englannin kielen taito ikäluokkien mukaan	49

Kuvio 10. Ammattikäyttäjien MOS-tulokset	49
Kuvio 11. Arkikäyttäjien englannin kielen taito ikäluokkien mukaan	67
Kuvio 12. Arkikäyttäjien MOS-tulokset	68

## **Taulukot**

Taulukko 1. Kyselyyn vastanneet	45
Taulukko 2. MOS-tulokset	47
Taulukko 3. Ammattikäyttäjien muuttujien MOS-tulokset	50
Taulukko 4. Ammattikäyttäjät: Muuttujat miesvastaajien osalta	51
Taulukko 5. Ammattikäyttäjät: Muuttujat naisvastaajien osalta	52
Taulukko 6. Ammattikäyttäjien luonnollisuuden alakäsitteet	53
Taulukko 7. Ammattikäyttäjien luonnollisuuden alakäsitteet: Miesvastaajat	54
Taulukko 8. Ammattikäyttäjien luonnollisuuden alakäsitteet: Naisvastaajat	55
Taulukko 9. Ammattikäyttäjien ymmärrettävyyden alakäsitteet	56
Taulukko 10. Ammattikäyttäjien ymmärrettävyyden alakäsitteet: Miesvastaajat	58
Taulukko 11. Ammattikäyttäjien ymmärrettävyyden alakäsitteet: Naisvastaajat	59
Taulukko 12. Puhesynteesi miesäänellä: ammattikäyttäjien kokonaisarvot	60
Taulukko 13. Miesääni: ammattikäyttäjien luonnollisuuden alakäsitteet	61
Taulukko 14. Miesääni: ammattikäyttäjien ymmärrettävyyden alakäsitteet	62
Taulukko 15. Puhesynteesi naisäänellä: ammattikäyttäjien kokonaisarviot	63
Taulukko 16. Naisääni: ammattikäyttäjien luonnollisuuden alakäsitteet	64
Taulukko 17. Naisääni: ammattikäyttäjien ymmärrettävyyden alakäsitteet	65
Taulukko 18. Arkikäyttäjien muuttujien MOS-tulokset	69

Taulukko 19. Arkikäyttäjät: Muuttajat miesvastaajien osalta	70
Taulukko 20. Arkikäyttäjät: Muuttajat naisvastaajien osalta	71
Taulukko 21. Arkikäyttäjät: luonnollisuuden alakäsitteet	72
Taulukko 22. Arkikäyttäjien luonnollisuuden alakäsitteet: Miesvastaajat	73
Taulukko 23. Arkikäyttäjien luonnollisuuden alakäsitteet: Naisvastaajat	74
Taulukko 24. Arkikäyttäjät: ymmärrettävyyden alakäsitteet	75
Taulukko 25. Arkikäyttäjien ymmärrettävyyden alakäsitteet: Miesvastaajat	76
Taulukko 26. Arkikäyttäjien ymmärrettävyyden alakäsitteet: Naisvastaajat	77
Taulukko 27. Puhesynteesi miesäänellä: arkikäyttäjien kokonaisarviot	78
Taulukko 28. Miesääni: arkikäyttäjien luonnollisuuden alakäsitteet	79
Taulukko 29. Miesääni: arkikäyttäjien ymmärrettävyyden alakäsitteet	81
Taulukko 30. Puhesynteesi naisäänellä: arkikäyttäjien kokonaisarviot	82
Taulukko 31. Naisääni: arkikäyttäjien luonnollisuuden alakäsitteet	83
Taulukko 32. Naisääni: arkikäyttäjien ymmärrettävyyden alakäsitteet	84
Taulukko 33. Varianssianalyysi: kokonaistulokset	85
Taulukko 34. Varianssianalyysi: luonnollisuuden alakäsitteet	87
Taulukko 35. Varianssianalyysi: ymmärrettävyyden alakäsitteet	89
Taulukko 36. Varianssianalyysi: miesäänen kokonaistulokset	91
Taulukko 37. Varianssianalyysi: naisäänen kokonaistulokset	92
Taulukko 38. Varianssianalyysi: miesäänen luonnollisuuden alakäsitteet	93
Taulukko 39. Varianssianalyysi: naisäänen luonnollisuuden alakäsitteet	94
Taulukko 40. Varianssianalyysi: miesäänen ymmärrettävyyden alakäsitteet	95
Taulukko 41. Varianssianalyysi: naisäänen ymmärrettävyyden alakäsitteet	96

# 1 Johdanto

Keinotekoisesta puheesta, eli tarkemmin puhesynteesistä, on tullut viime vuosien aikana vakiintunut ilmiö maailmalla. Sitä kuullaan yleisökuulutuksissa julkisissa kulkuneuvoissa, auton GPS-navigointi äänenä, vieraan kielen kääntämisen puheavustajana sekä monessa muussa harrastustoiminnassa sekä kasvavassa määrin ammattimaisessa toiminnassa. (Mehler & Romary 2012: 189) Esimerkiksi monen yrityksen asiakaspalvelun on syrjäyttämässä puhesynteesiä avuksi käytettävä järjestelmä perinteisen puhelinvaihteen ja vastaanoton sijaan (Tung 2018).

Puhesynteesin käyttömahdollisuudet ovat myös laajentuneet, ja esimerkiksi puhesynteesin käyttö puhelinvaihteessa herättää kysymyksiä puhesynteesin ja kuulijan välisestä yhteydestä. Puhesynteesin vastaanotto on tärkeää onnistuneen viestinnän kannalta, mutta se nostaa esiin myös mielenkiintoisen kysymyksen siitä, miten kuulija tarkalleen kokee puhesynteesin käytön ja kärsiikö informaation välitys puhesynteesin vuoksi.

Työskentelen hanoja valmistavassa Oras Group Oy:ssä teknisen viestinnän asiantuntijana, ja työni kautta valikoitui myös tutkielman aihe. Tuotan ohjevideoita osana työtäni ja puhesynteesin käyttö nousi esiin keskustellessamme keinoista optimoida ohjevideoiden tuotantoa. Tästä heräsi ajatus siitä, että puhesynteesin käyttöä ohjevideolla testattaisiin kohderyhmillä ja heidän kokemuksiaan arvioitaisiin tilastollisesti.

Puhesynteesin tuottaminen on tällä hetkellä kustannustehokasta, ja se on myös helpompaa kuin ammattipuhujan palkkaaminen. Äänen optimointi ja muokkaaminen onnistuu huomattavasti helpommin puhesynteesillä, ja mahdollisia virheitä korjaavat hintavat uudelleen nauhoitukset ammattipuhujan tekemänä jäävät myös kokonaan pois puhesynteesillä. Näiden huomioiden kautta heräsi kiinnostus lähteä tutkimaan puhesynteesin käyttöä ohjevideoilla tarkemmin. Puhesynteesi on myös monipuolisen käyttönsä vuoksi hyödyllinen väline moderneille yhtiöille, joissa tuotetaan monenlaista mediasisältöä sekä videoita ja äänitteitä sisältäviä e-koulutuslustoja. Puhesynteesi on myös hyvin välitöntä tuottaa ja näin ollen mahdolliset korjaukset videon puheeseen saadaan tehtyä

välittömästi ja asiakkaille saadaan päivitettyä tietoa tuotteen asentamisesta tai huoltamisesta nopeammin.

Puhesynteesi, joka tunnetaan myös nimellä tekstistä-puheeksi (engl. *text-to-speech*, TTS) on ihmispuheen kaltaisen puheen tuottamista keinotekoisin tavoin. Tekstistä-puheeksi voidaan tarkentaa tarkoittavan kirjoitetusta tekstistä puhutuksi puheeksi tuotettua ääntä. (Taylor 2007: 25) Yleisin mielikuva, joka ihmiselle tulee mieleen puhesynteesistä, on koomisen olinen robottiääni, joka on kaikkea muuta kuin uskottavan tai ihmismäisen kuuloinen. Nykytilanne puhesynteesissä on kuitenkin toinen. Puhesynteesi on kehittynyt mekaanisesti luoduista tai alkeellisesti tietokoneella luoduista, muun muassa Stephen Hawkingin kuuluisaksi tekemistä, äänikoneiden päivistä. Näillä äänikoneilla luotu puhe oli hyvin karua ja vaikkakin se oli ymmärrettävää, puheen luonnollisuus oli kaukana ihmispuheesta. Äänikoneiden käyttäminen oli myös hidasta ja hankalaa. (Medeiros 2015)

Puhesynteesin perinteiset ja vakiintuneet arvioinnin määritteet ovat olleet luonnollisuus ja ymmärrettävyys. Toisin sanoen sen samankaltaisuus ihmispuheen kanssa sekä sen äänenlaadun selkeys ja kieliopillisuus. (Le Maguer 2018: 10-12, 25-26) Puhesynteesin aiempi tutkimus on keskittynyt näihin kahteen, mutta enemmän keinotekoisien puheen tuotannon ja siihen liittyvien teknisten ominaisuuksien ja aidolta kuulostavan puheen luonnin näkökulmasta, kun taas tutkimus siitä, miten ihmiset vastaanottavat puhesynteesiä ja mitä tunteita tai ajatuksia heille siitä syntyy, on jäänyt vähemmälle huomiolle. (Chung, Wang, Hsu, Zhang & Skerry-Ryan 2018: 1-5) Tässä tutkielmassa keskitytään käyttäjän näkökulmaan ja nimenomaan siihen, miten arkikäyttäjä tai alan ammattilainen vastaanottaa puhesynteesin.

## **1.1 Tutkimuksen tavoite**

Tämän tutkielman tavoitteena on selvittää, miten ammattikäyttäjät ja arkikäyttäjät kokevat puhesynteesin käytön Oras Groupin termostaattihanan huoltoa ohjeistavalla videolla,



sekä arvioida tulosten perusteella puhesynteesin mahdollista käyttöönottoa ohjevideossa tulevaisuudessa. Tutkielmalla haen vastausta kysymyksiin:

1. Miten puhesynteesin luonnollisuus ja ymmärrettävyys koetaan teknisessä ohjevideossa?
2. Onko ammattilaisen ja arkikäyttäjän kokemuksissa puhesynteesistä eroja?
3. Miten muuttujat kuten puhesynteesin mies- ja naisääni tai vastaajien ikä, sukupuoli tai englannin kielen osaaminen vaikuttavat puhesynteesin kokemiseen?

Oras Group Oy:n tekniset ohjevideot on suunnattu niin ammattikäyttäjälle kuin arkikäyttäjällekin, ja tästä syystä on tärkeää tutkia, miten nämä kaksi ryhmää vastaanottavat saman teknisen materiaalin ja onko puhesynteesin vastaanotossa eroa näillä ryhmillä. Ammattikäyttäjien ryhmä on saanut vähintään yleisen koulutuksen hanan asentamisesta ja huollosta työn aloittamisen yhteydessä ja on näin ollen lähtökohtaisesti valmiimpi vastaanottamaan ohjevideon tietoa. Arkikäyttäjien ryhmä tuntee hanat pelkästään loppukäyttäjän näkökulmasta, ilman kokemusta hanojen asentamisesta tai huollosta.

Tekninen dokumentti, esimerkiksi käyttöohje tai ohjevideo, on väline, jolla välitetään lukijalle tietoa mahdollisimman yksiselitteisesti ja jonka avulla lukijan on tarkoitus suorittaa konkreettisia toimintoja (Mehler & Romary 2012:40). Näitä toimintoja voivat olla erinäisten tuotteiden asennus, huolto tai käyttöönotto. Tutkielmassa keskitytään esittämään hanan huoltotoiminto teknisen dokumentaation avulla ja tarkemmin käyttöohjeen ja siitä tuotetun ohjevideon keinoin.

Puhesynteesin arviointikriteereinä toimivat aiemmin mainitut luonnollisuus ja ymmärrettävyys. Luonnollisuuden ja ymmärrettävyyden mittarit haarautuvat alakäsitteisiin, joilla haetaan yksityiskohtaisempaa tarkennusta mittareille. Käsitteet ja alakäsitteet avataan tarkemmin tutkielman luvuissa 1.3.1 ja 1.3.2.

Käsitteiden avulla arvioin arkikäyttäjien sekä ammattikäyttäjien kokemusta puhesynteesistä ja siitä, miten se toimii heidän mielestään huoltovideolla. Kyselytutkimuksella saadaan arkikäyttäjän ja ammattikäyttäjän puhesynteesin vastaanoton tulos. Näitä vertailaan keskenään ja niistä selvitetään puhesynteesin toimivuus teknisessä huoltovideossa kummassakin ryhmässä. Vertailulla saadaan selville myös, vaikuttaako alan tuntemus puhesynteesin kokemiseen ja siihen, mitä tuntemuksia se herättää verrattuna toiseen ryhmään, jolla ei ole kokemusta alasta muuten kuin kuluttajana.

## 1.2 Tutkimusaineisto

Tutkimusaineisto on kerätty kyselyllä, johon ovat vastanneet sekä ammatti- että arkikäyttäjät. Kysely suoritettiin kahdessa samaan aikaan avoimena olevassa kyselyssä. Toisessa kyselyyn vastasivat Oras Group-yhtiön työntekijät, eli alan ammattilaiset, ja toiseen vastasivat ns. arkikäyttäjät eli tavalliset kuluttajat, joilla ei ole kokemusta hanojen huollosta tai hanojen asennuksesta. Arkikäyttäjärühmän kyselyyn lisättiin kyselyn alkuun kyllä/ei-kysymys, jolla vastaajat selventävät, omaavatko he kokemusta hanojen asentamisesta tai huollosta. Vain vastaajat, jotka vastasivat ”ei kokemusta” sisällytettiin tutkielmaan. Kyselyyn vastataan viisiportaisen asteikon mukaan huonoimmasta (1) parhaimpaan (5).

Arkikäyttäjien ryhmää ja heidän vastauksiaan nimitetään tilastollisen tutkimuksen teorioita mukaillen *näytteeksi*. Koska en ole ollut yhteydessä vastaajiin ennakolta, enkä voi myöskään tietää yksittäisten vastaajien henkilöllisyyttä, ei tässä tapauksessa voida puhua otoksesta vaan nimenomaan näytteestä. Tilastollisten teorioiden mukaan tämä on oleellista mainita kyselyn ryhmistä puhuttaessa. (Valli 2015: 49)

Alan ammattilaisten ryhmästä ja heidän tuloksistaan puhuttaessa käytetään nimitystä *harkinnanvarainen näyte*. Tämä siitä syystä, että ryhmä on valikoitu etukäteen ja kyselyssä on ryhmä, jonka etuja kyselyllä haetaan ja ryhmä, jonka tarpeita täyttävä kysely on. Kyselyssä esillä oleva tuote on myös entuudestaan tuttu Oras Groupin henkilöstölle, joten ryhmä on tästäkin syystä ennakolta harkinnanvarainen. (Vehkalahti 2014: 46)

Kysely toteutettiin kyselynetti.fi-sivustolla arkikäyttäjille, kun taas ammattikäyttäjien kysely toteutettiin yhtiön sisäisen palvelimen sekä kyselypohjan kautta. Arkikäyttäjien ryhmän kysely jaettiin omien Facebook- sekä Instagram-tilien kautta vastattavaksi. Ammattikäyttäjien kysely jaettiin Oras Groupin intranetin sekä sisäisen sähköpostin kautta koko organisaation vastattavaksi.

Kyselynetti.fi palvelu oli käytössä kyselyn ensimmäisen version esitestausta suorittaessani keväällä 2019, ja se osoittautui päteväksi alustaksi niin kyselyn luonnin kannalta kuin kyselyn toteutuksen ja siitä saatujen tulosten tilastoinnissakin. Palvelu täyttää siis verkkokyselyn tärkeimpiä vaatimuksia (Valli 2015: 47). Yrityksen sisäistä kyselypohjaa ja alustaa oli myös helppo käyttää kyselyn luomiseen ja toteutukseen. Alusta tuottaa saman datan kuin kyselynetti.fi-alusta, joten eroavaisuuksia ei ole datan tilastoinnissa. Molemmat palvelut tuottavat raakadatan kyselyn vastauksista, mutta keskiarvon laskemisen suoritan kuitenkin itse.

Tietosuoja on myös tärkeä tekijä kyselylomaketta tehtäessä ja kerätessä tietoja ihmisiltä, vaikkakin tässä kyselyssä ei kysytä henkilötietoja muuten kuin iän ja sukupuolen osalta. Kyselyn saatekirjeessä on kuitenkin maininta, että tiedot kerätään tutkielmaa varten ja osallistumalla kyselyyn osallistuja hyväksyy tiedon käyttämisen tutkielman aineistona. Tämä on hyvän kyselylomakkeen käytäntö ja vastaajien yksityisyys, eli anonymiteetti saadaan näin turvattua. Saatekirje on muutenkin hyvin tärkeä osa kyselytutkimusta, koska sillä saadaan parhaimmassa tapauksessa motivoitua osallistujia vastaamaan kyselyyn, tai vaihtoehtoisesti lannistettua hänet olemaan vastaamatta kyselyyn. Saatekirjeessä tulee myös olla selvästi esillä kyselyn tavoite, tutkimuksen tausta ja ohjeet kyselyn tekemiseen. (Tähtinen, Laakkonen & Broberg 2011: 24-25)

Kysely koostuu 12 kysymyksestä, joita edeltää englanninkielinen videonäyte teknisestä ohjevideosta, jossa kuvataan Oras Esteta Wellfit-termostaattihanan vaihtimen kahvan asennon tunnistimen vaihto. Kuten ylläolevasta virkkeestä käy ilmi kyseessä on tekninen

ja erikoiskieltä sisältävä video. Tämä on tietoinen valinta, koska tämä kuvastaa tyypillistä teknistä ohjevideota. Jos olisin valinnut helpommin ymmärrettävän videonäytteen ja huolto-operaation, olisi se antanut mielestäni väärän kuvan teknisistä ohjevideoista, ja se olisi ollut liian yleistajuinen selvittääkseen arkikäyttäjän ja ammattikäyttäjän eroja aineistosta.

Videonäytteitä on yhteensä kaksi, joissa ainoana erona on, että toisessa videossa on miesääni ja toisessa naisääni. Videonäyte on 1 minuutti 30 sekuntia pitkä ja siinä käydään sanallisesti ja kuvallisesti läpi yllä mainittu huolto-operaatio. Videonäyte vaihtui aina loppuun suoritettun vastauksen jälkeen, joten molemmille videonäytteille saatiin näin sama määrä vastaajia. Vastaukset identifioitiin myös sen mukaan, minkä version videosta kyselyn tekijä katsoi. Tämän lisäksi vastaajilta kysytään kyselyn aluksi ikä, sukupuoli sekä heitä pyydetään arvioimaan englannin kielen osaamistaan. Näitä muuttujia ja niiden eroja tutkitaan vastausten eroavaisuuksien kautta aineistossa. Näitä käytetään myös apuna aineiston analyysissä ja tulkinnassa siitä, onko eroavaisuuksia eri sukupuolten, ikäluokkien tai englannin kielen taidon välillä puhesynteesin kokemisessa.

### 1.3 Tutkimusmenetelmät

Tutkielman menetelmänä toimii kvantitatiivinen kyselytutkimus, jonka vastauksista lasketaan *Mean Opinion Score (MOS)*, eli keskiarvoinen tulos luonnollisuuden ja ymmärrettävyyden käsitteille ja kaikille muuttujille: vastaajien ikä, sukupuoli ja englannin kielen taito sekä lisäksi puhesynteesin muuttujat: miesääni ja naisääni. Näistä keskiarvoista suoritetaan varianssianalyysi eri muuttujien välillä ja muuttujien tuloksista pyritään löytämään yhtäläisyyksiä ja eroavaisuuksia (ks. Heikkilä 2014: 210).

Kyselyn pohjana toimii siis mittausasteikon ja strukturoitujen vastausvaihtoehtojen puolesta *Mean Opinion Score (MOS)* -metodi. MOS-metodi on otettu alun perin käyttöön puheluissa käytetyn äänenlaadun arvioimiseen (International Telecommunication Union 1994: 85). Alun perin puheluiden puheen äänenlaadun ja selkeyden arvioimiseen luotu

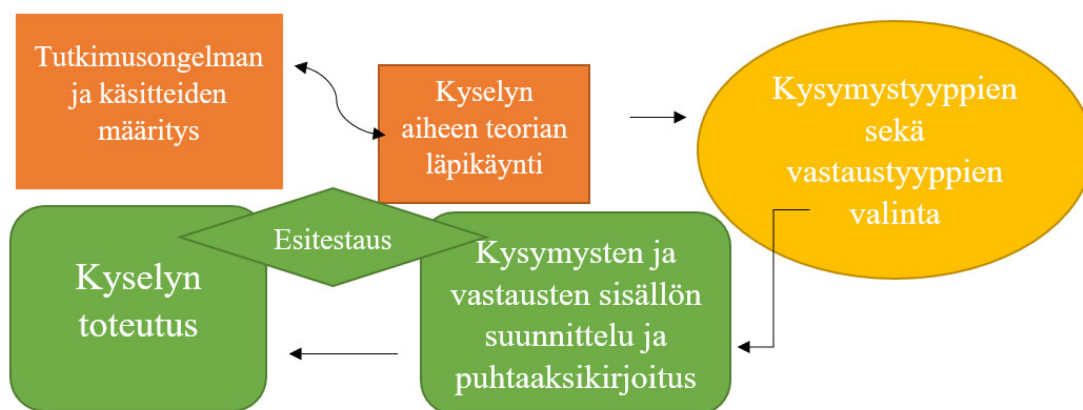
kysely ja arvoasteikko on myöhemmin otettu käyttöön käyttäjän kokemuksia arvioivaan tutkimukseen ja tätä tutkielmaa silmällä pitäen myös puhesynteetin ja sen puheen laadun arvostelun metodiksi. (Streijl, Winkler, & Hands 2016: 213)

Kyseessä on peruseriaatteiltaan kyselylomake, jossa jokaisessa kysymyksessä on viisi eri vastausvaihtoehtoa. Kysymysten vastausten arvot lasketaan yhteen, nämä jaetaan osallistujamäärällä, ja näin ollen saadaan keskiarvo, josta tulee MOS-asteikko 1-5, jossa 1 on huonoin ja 5 paras. (Viswanathan & Viswanathan 2005: 56) MOS käydään kokonaisuutena läpi tarkemmin alaluvussa 1.3.2.

Teoreettiseen viitekehykseen liittyvät vahvasti vertailevan tutkimuksen teoriat tilastollisesta tutkimuksesta. Käytän tutkielmassa varianssianalyysiä, jolla analysoidaan yhden tai useamman muuttujan vaikutusta vasteeseen, eli tässä tutkielmassa MOS-metodilla saatuun keskiarvoon. (Ks. Tähtinen ym. 2011: 102) Muuttujina on ylimmällä tasolla kaksi käyttäjäryhmää: ammattikäyttäjä ja arkikäyttäjä. Seuraavan tason muuttujia ovat käyttäjäryhmien sisäiset muuttujat sukupuoli ja ikä sekä englannin kielen taitotaso, jotka tilastoidaan myös. Teoreettisena pohjana toimivat luonnollisuuden ja ymmärrettävyyden käsitteiden tarkentamisessa onnistuneen viestinnän periaatteet, informaation välitys, audiovisuaalinen viestintä, kuuntelukokemus sekä tekninen viestintä.

### **1.3.1 Kyselytutkimus**

Kyselytutkimuksen toteuttamisen prosessit on tiivistetty Kuviossa 1.



**Kuvio 1.** Kyselytutkimuksen prosessit

Prosessit kulkevat käytännössä tutkimusongelman kohdentamisesta teorian avulla, kyselyn luontiin, sen esitestaukseen ja toteutukseen (Vilkkä 2007: 79). Testaaminen on jo suoritettu keväällä 2019, jolloin loin alustavan version kyselystä. Testin pohjalta muokattulle kyselylle suoritettiin lisäksi esitestaus, johon osallistui kaksi kollegaani Oras Group Oy:ssä. Tämän avulla pyrittiin löytämään mahdollisia virheitä teknisellä tai ulkoasullisella puolella sekä myös löytämään korjausehdotuksia kysymyksiin tai vastauksiin.

Kysymysten sekä kysymystyyppien laadintaa pohjustivat tutkimuskysymykset ja arvioitiin käytetyt käsitteet. Tärkein seikka kysymyksiä laatiessa olikin, että jokainen kysymys palvelee jotain haetuista käsitteistä, joita tutkitaan. Tärkeää kyselyn kysymysten luonnissa on myös luonnollisesti se, että kysymykset ovat toimivia ja että niihin on helppo vastata. Tähän vaikuttavat paljon vastausvaihtoehdot ja niiden selkeys. (Heikkilä 2014: 54-55)

Kysymysten luonti pohjautui tilastollisessa tutkimuksessa käytettävään hyvän kysymyksen käsitteeseen. Tähän kuuluu ymmärrettävyys (osallistuja ymmärtää kysymyksen), tieto asiasta (osallistuja omaa tiedon vastaukseen), vastaushalukkuus (on halukas vastaamaan) ja yksiselitteisyys (kysymykset ovat yksiselitteisiä). (Kananen 2014: 142)

Hyvän kysymyksen on siis oltava kieliasultaan moitteeton ja siinä on kysyttävä selkeästi ja tarkasti vain yhtä asiaa. Hyvän kysymyksen on myös mahdollistettava tuloksen saanti

halutulla tavalla, eli toisin sanoen kyselyyn osallistujan on pystyttävä vastaamaan kysymykseen. Kysymys ei saa myöskään olla liian pitkä tai liian vaikeasti ymmärrettävä, vaan sen on oltava mahdollisimman yksiselitteinen. (Heikkilä 2014: 54-55)

Hyvän kysymyksen piirteitä on myös selkeä kysymystyyppi, suoraviivaiset vastausvaihtoehdot sekä toimiva mittausasteikko. Onnistuneeseen kyselyyn kuuluu kuitenkin monia muitakin tekijöitä kuin vain nämä yllämainitut tekijät. Näitä ovat muun muassa vastaajan mielenkiinto kyselyä kohtaan, fyysinen mahdollisuus vastata kyselyyn, saatekirjeen/kyselyn ohjeistuksen riittävyys sekä muut inhimilliset tekijät, jotka voivat vaikuttaa kyselyn tekoon. (Vilkkä 2007: 63-70)

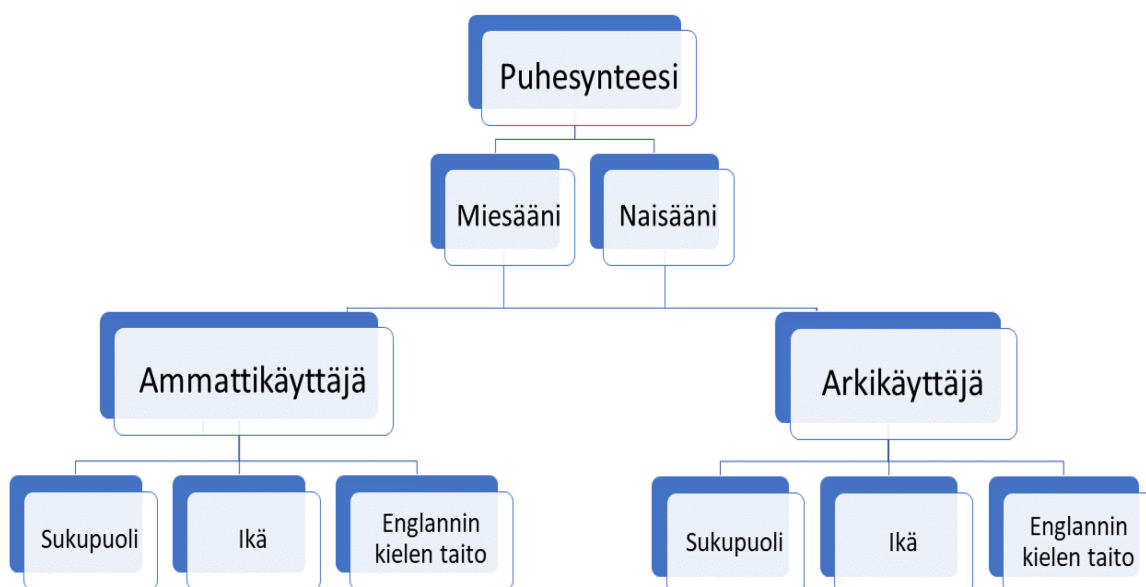
Kysymystyyppit voidaan jakaa karkeasti kolmeen isoon ryhmään: monivalintakysymykset, avoimet kysymykset sekä sekamuotoiset kysymykset. Tutkielmassani keskityn ensimmäiseen. Monivalintakysymykset ja niiden vastausvaihtoehdot ovat strukturoidut sekä ennalta määräytyt, niiden kysymysmuoto on myös vakioitu. (Vilkkä 2007: 62) Tämä noudattaa kyselytutkimuksen laadinnan perusteorioita ja on hyvin suoraviivainen ja tehokas tapa kerätä dataa.

Hyvien kysymysten luonti on ehkä tärkein seikka kyselylomaketta laadittaessa, mutta hyvät kysymykset voidaan tehdä tarpeettomiksi huonoilla tai riittämättömillä vastausvaihtoehdoilla. Tutkielmassani käytän jo aiemmin mainittua MOS-asteikkoa, joka on verrattavissa tilastollisessa tutkimuksessa yleisesti käytettävään Likertin skaalaan. Likert-asteikoiset kysymykset täyttävät vuorostaan välimatka-asteikon kriteerit, joka on myös vakiintunut mitta-asteikko tilastollisessa tutkimuksessa. Välimatka-asteikossa voidaan määrittellä muuttujien eroavaisuus toisistaan. Muuttujien arvojen järjestys on myös ennalta määrätty ja yksi sen suositeltavista testausmenetelmistä on myös MOS-metodissa käytettävä keskiarvon laskeminen. (Tähtinen ym. 2011: 26)

Vastausvaihtoehdot ja mittausasteikko noudattivat tutkielmassani siis MOS-asteikkoa ja menetelmää. MOS-asteikko on viiden strukturoidun vastausvaihtoehdon kysely, jossa

asteikko on useimmiten muotoiltu ääripäästä ääripäähän esim.: *bad, poor, fair, good, excellent* (Viswanathan & Viswanathan 2005: 63). Käytän tässä tutkielmassa kyseisestä asteikosta suomennosta erittäin huono, huono, kelvollinen, hyvä, erinomainen. Tulokset lasketaan laskemalla keskiarvo osallistujien vastauksien arvoista.

Kun MOS-keskiarvot on laskettu kaikista kyselyn pääkäsitteistä ja muuttujista, suoritetaan näille sovellettu varianssianalyysi, jolla vertaillaan eri keskiarvoja toisiinsa. Varianssianalyysillä tutkitaan nimenomaan keskiarvojen välisiä eroja eikä varianssien välisiä eroja kuten nimestä voisi päätellä (Heikkilä 2014: 210). Käytän tutkielmassani yksisuuntaista varianssianalyysiä, eli jaan muuttujat yksittäisiin ryhmiin ja vertaan näitä keskenään. Käyttäjärühmien lisäksi muuttujia ovat osallistujien ikä, sukupuoli sekä englannin kielen taitotaso. Puhesynteesin muuttujia ovat miesääni ja naisääni. Muuttujat kuvattu kuviossa 2.



**Kuvio 2.** Muuttujat kyselytutkimuksessa

Jokaiselle muuttujalle lasketaan MOS-tulos kaikista vastauksista sekä erikseen oma arvo ymmärrettävyydelle, luonnollisuudelle sekä näiden alakäsitteille. Suoritan varianssianalyysin yllämainittujen muuttujien keskiarvojen (MOS-tulos) välillä.



Tärkeäksi osaksi varianssianalyysissä nousee tilastojen tarkka listaaminen ja tulosten tarkka raportoiminen. Tuloksia raportoitaessa varianssianalyysissä on tarpeellista esittää vain oleelliset tiedot. Käytän omassa tutkielmassani ja raportoinnissani keskiarvojen vertailua, ryhmäkokojen ilmoitusta sekä tietoa siitä onko eroavaisuus ollut merkittävä. (ks. Heikkilä 2014: 212)

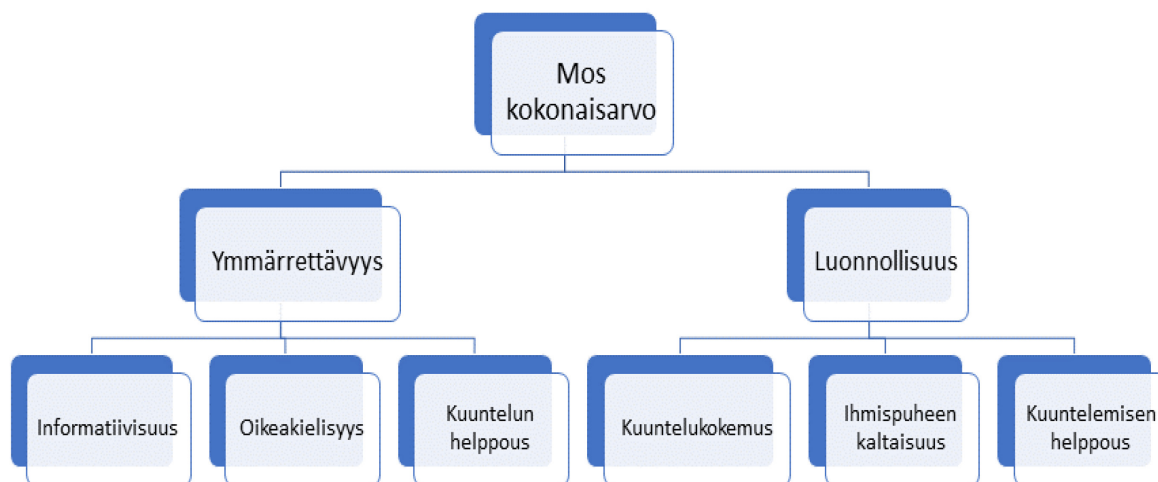
### **1.3.2 Mean opinion score (MOS)**

Puhesynteesin alalla vakiintuneeksi tutkimusmenetelmäksi on muodostunut MOS-kyselytutkimus, eli Mean opinion score -tutkimus, jolla pyritään löytämään keskiarvo halutuista laatuksista. MOS-metodi on otettu alun perin käyttöön puheluissa käytetyn äänenlaadun arvioimiseen (International Telecommunication Union 1994: 85). MOS on sittemmin otettu puhesynteesin arviointimetodiksi ja laadun määrittämisen konseptiksi.

MOS-kyselyn voi suorittaa subjektiivisen tai objektiivisen metodin avulla. Tässä tutkielmassa käytetään käyttäjälähtöisempää subjektiivisen testin menetelmää, koska tällä menetelmällä tarkastellaan enemmän puhesynteesin vastaanottajan kokemusta. Objektiivisessä testissä huomio kohdistuu puhesynteesin teknisiin ominaisuuksiin ja laatuun. (Strejil, Winkler, & Hands 2016: 213)

Kyselyni ei ole täysin perinteinen MOS-kysely vaan metodi toimii pohjana, jota sovellan omaan tutkimukseeni. Kriteerit, joilla tutkin kyselyn tuloksia ovat perinteiset

luonnollisuus ja ymmärrettävyys, jotka haarautuvat alakäsitteisiin Kuvion 3 osoittamalla tavalla.



### Kuvio 3. Kyselytutkimuksen käsitteet

Jokainen ymmärrettävyyden ja luonnollisuuden alakäsite sisältää kaksi kysymystä kyselylomakkeessa. Näin syntyy 12 kysymyksen kyselylomake. Jokaiselle käsitteelle saadaan näin myös laskettua oma yksilöity tulos. Kuvion 2 mukaan tulokset yhdistellään kolmesta alakäsitteestä ja näin ollen saadaan ensin ymmärrettävyyden ja luonnollisuuden tulokset ja nämä yhteen laskettuna MOS-kokonaisarvo. Luonnollisuuden ja ymmärrettävyyden käsitteitä ja niiden alakäsitteitä tarkennetaan luvussa 4.4.

Omassa tutkielmassani käytän mitta-asteikkona: erittäin huono, huono, kelvollinen, hyvä, erinomainen-skaalaa. Mittausasteikosta löytyy variaatioita mm. 7- ja 9-osaisia asteikkoja, mutta koin osuvammaksi yksinkertaisimman ja selkeimmän viiden asteikon metodin. Isomman asteikon mittauksissa on luonnollisesti tarkemmat ja hienojakoisemmat määritteet, mutta viisiasteikkoinen järjestelmä oli käytössä jo aiemmin suoritettussa esitetauksessa ja se toimi tarpeellisesti, joten en kokenut hyödylliseksi lähteä muuttamaan sitä tähän kyselyyn.

Kaikista esitetyistä käsitteistä lasketaan mittausasteikon mukaisesti keskiarvo jakamalla yhteenlaskettu vastausten MOS-arvo (1-5) vastaajien määrällä, alla olevan kaavan mukaisesti.

(1)

$$A_n = \frac{a_1 + a_2 + \dots + a_n}{n}$$

(De 2016: 1119)

Laskettu keskiarvo kysymyksistä tilastoidaan taulukoilla ja kuvioilla, joista käyvät ilmi jokaisen käsitteen MOS-arvo. Mitä suurempi arvo, sitä laadukkaampana on kyseistä kriteeriä pidetty. Kyselyn tuloksista ja listatuista tilastoista tehdään yksittäiset analyysit sekä vertaileva varianssianalyysi, ja tästä tehdään johtopäätökset ja pohdinta puhesynteesin kokemisesta ohjevideolla.

## 2 Informaation välitystä onnistuneella viestinnällä

Yksi viestinnän tärkeimmistä, ellei tärkein, tehtävä on välittää informaatiota eteenpäin (Akilandeswari 2015: 44). Onnistunut ja tarkka informaation välittäminen on myös teknisen viestinnän ja näin myös oman työtehtäväni ohjevideoitten tekijänä, tärkein osa.

Teknisen viestinnän kehitys ja itsenäistyminen on ollut sen lähihistoriassa jopa tarpeettoman hidasta, ja tähän hitauteen suurimpana syynä oli teknisen viestinnän käsitteen määrittämisen vaikeus itse alan harjoittajien ja ammattilaisten toimesta. Käsitteen määrittämisen vaikeus hankaloitti vuorostaan alan rajaamista selkeästi omaksi tieteenalaksi. (Hundleby & Allen 2010: 7-8)

Teknistä viestintää on alana ja käsitteenä kasvanut tästä epäselvyyden ajasta ja Suomalakin on nykyään oma Teknisen viestinnän yhdistys (STVY) määrittelee teknisen viestinnän näin:

*Sekä laitteiden että ohjelmistojen tehokas ja turvallinen käyttö edellyttää, että pystymme tutustumaan niin niiden uusiin kuin vanhoihinkin toimintoihin vaikkapa käyttöohjeen avulla. Tällaisen käyttäjille tarkoitettun informaation suunnittelu ja tuottaminen on ala nimeltään tekninen viestintä, ja alan asiantuntija on nimeltään tekninen viestijä. Käyttöohjeiden suunnittelu ja toteuttaminen on vain yksi osa teknistä viestintää.  
(STVY 2019)*

Tekninen viestintä ja sen ammattilaisia on ollut vaikea kuvata juuri siksi, että tekninen viestijä voi toimia hyvin monella eri alalla, ja teknistä viestintää käytetään nykypäivänä oikeastaan joka alalla jossain muodossa.

Tekninen viestintä on myös viime vuosina kasvanut tärkeäksi osaksi modernien organisaatioiden viestintää. Kuvaavinta tästä on se, että monella yrityksellä on nykyään täysin oma jaoksensa tekniselle viestinnälle, ja toimintoja ei sisällytetä esimerkiksi markkinointiin tai viestintään, vaan se on kasvavissa määrin täysin oma ja itsenäinen jaoksensa

organisaation sisällä. Oras Group Oy:ssä tekninen viestintä tuli omaksi itsenäiseksi jaoksekseen vuonna 2015.

Oras Groupissa teknisen viestinnän jaoksen ensisijainen tehtävä on välittää teknistä tietoa eteenpäin niin sisäisesti organisaation sisällä, kuin ulkoisestikin asiakkaille, alihankkijoille sekä yhteistyökumppaneille. Teknistä tietoa välitetään pääasiallisesti käyttöohjeiden sekä ohjevideoiden kautta. Näistä tarkemmin luvussa 3.

Alustoja ja välineitä on monia viestinnässä, mutta tässä tutkielmassa keskityn audiovisuaalisen viestinnän elementteihin ohjevideoissa sekä teknisen viestinnän ominaisuuksiin käyttöohjeissa. Audiovisuaalinen ja tekninen viestintä pohjautuvat kuitenkin viestinnän perustaviin viestintämalleihin, jotka käydään läpi alaluvuissa.

## **2.1 Viestintämallit**

Teknisellä viestinnällä välitetään teknistä informaatiota ja tämän vuoksi informaation välitys ja itse viestinnän aspekti voi vaikuttaa hyvin yksinkertaiselta ja suoraviivaiselta prosessilta. Viestintä on kuitenkin harvoin, jos koskaan, vain näin yksinkertainen informaation siirto lähettäjältä vastaanottajalle. Tutkielman kannalta onkin tärkeää käydä läpi eri viestintämallit ja niiden peruseräatteen.

Viestintämallien ajatuksena on kuvata, miten siirretään informaatiota henkilöltä toiselle. Informaatio taas muuttuu tiedoksi, kun vastaanottaja tulkitsee saamansa informaation ja antaa sille merkityksen. Viestinnässä eri mallit tarjoavat tiedonsiirtoon kehyksen, jossa voidaan tarkastella kommunikaatioprosessin ongelmia. Hyvä malli onkin siis sellainen, joka selventää ja yksinkertaistaa kommunikaation rakennetta. Mallit, jotka perustuvat Shannonin ja Weaverin viestintämalliin ovat kaikista yleisimmin käytettyjä teknisessä viestinnässä. (Al-Fedaghi 2012) Nämä Shannonin ja Weaverin viestintämalliin perustuvat mallit ovat melko lineaarisia, kulkiessa lähettäjältä koodaimen kautta vastaanottajaan.

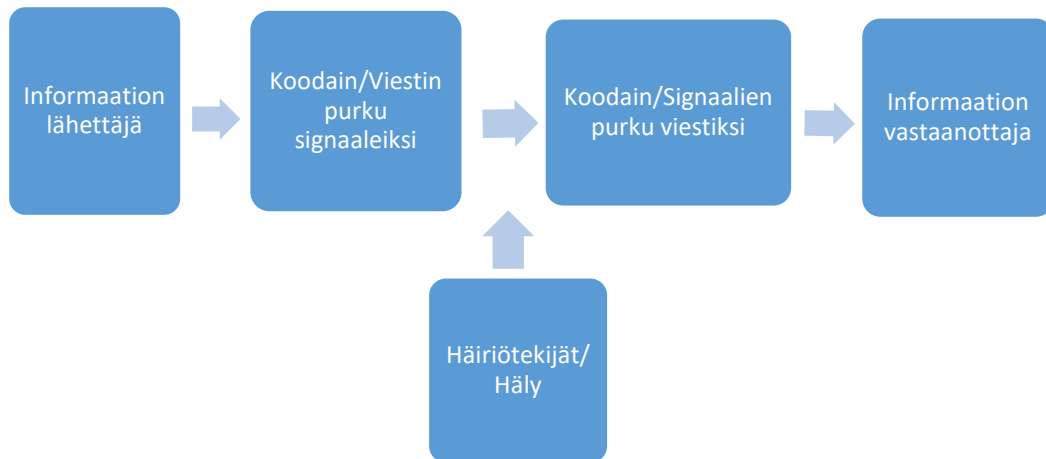
Viestintämalleja voidaan lähestyä myös semiotiikan sekä transaktionaalisuuden kautta. Näistä semiotiikka keskittyy merkkien ja merkitysten rooliin viestinnässä, kun taas transaktionaalisuus keskittyy vuorovaikutukseen. Linearisessa mallissa vuorovaikutus usein jää vain lähettäjän ja vastaanottajan väliseksi toiminnaksi. Vuorovaikutus voidaan kuitenkin nähdä myös joissakin viestintämalleissa useamman henkilön ja eri toimijoiden välisenä toimintana, joka vaikuttaa keskeisellä tasolla viestinnän kulkuun. (Van Ruler 2018: 370)

### **2.1.1 Shannonin ja Weaverin tiedonsiirron malli ja sen kritiikki**

Shannonin ja Weaverin viestintämalli kehitettiin vuonna 1948, jolloin se mainittiin ensimmäisen kerran artikkelissa ”A Mathematical Theory of Communication”, joka julkaistiin Bell System Technical Journal -lehdessä. Mallin laativat Claude Shannon ja Warren Weaver. Heidän laatimansa malli on yksi viestinnän tärkeimpiä malleja, jota on kutsuttu esimerkiksi viestinnän matemaattiseksi teoriaksi ja kaikkien muiden mallien äidiksi. (JMC Study Hub 2019)

Tämä viestinnän malli ei ole suoraan lineaarinen, vaan siinä on monta tekijää, jotka muodostavat yhteisen kokonaisuuden informaation siirrosta. Shannonin ja Weaverin mallin mukaan viestinnän prosessissa on viisi tärkeää tekijää. Näitä ovat:

1. lähettäjä (informaation lähde), 2. koodain (lähetin), 3. häiriötekijät (häly), 4. koodin purkaja (vastaanottaja), sekä 5. vastaanottaja (vastaanotin). Näitä havainnollistetaan tarkemmin Kuviossa 4.



**Kuvio 4.** Shannonin ja Weaverin viestinnän malli

Shannonin ja Weaverin mallissa informaation lähettäjä valitsee, minkä tiedon se päättää viestinnässä lähettää eteenpäin. Tämä viesti kulkee koodaimen, eli lähettimen läpi. Tämä voi olla esimerkiksi jokin kone, joka muuttaa viestin muodon signaaleiksi, kuten puhelimen äänen aalloiksi, jotka kulkevat puhelimen kuulokkeen läpi. Häiriötekijät, joita voidaan kutsua myös hälyksi, ovat kaikki se mikä tuottaa epämiellyttävää ääntä ja häiriötä viestintään. Tätä voivat olla esimerkiksi muut ihmiset ja ympäristö, jotka voivat muuttaa kulkevan viestin todellista tarkoitusta. (JMC Study Hub 2019)

Kun viesti saapuu signaaleina vastaanottajalle, se muuttuu taas koodaimessa, eli vastaanottimessa, ymmärrettävään muotoon. Esimerkiksi puhe kuuluu puheena ja tekstiviesti näkyy selvänä tekstinä eikä koodikirjoituksena. Kun viesti on ymmärrettävänä perillä, saavuttaa se lopulta vastaanottajan. Hyvänä esimerkkinä näistä viidestä tekijästä ovat ihmispuheen eri vaiheet. Aivot toimivat lähettäjänä, suu koodaimena, vastaanottajan huono kuulo häiriötekijänä, vastaanottajan aivot koodaimen vastaanottajana ja toinen henkilö itse vastaanottajana. (JMC Study Hub 2019)

Vaikka edellä mainittu viestintämalli on yleisin malli viestinnän saralla ja sen pohjalta ovat melkein kaikki muut mallit saaneet alkunsa, on Shannonin ja Weaverin viestintämalli saanut myös kritiikkiä. Vaikka Shannonin ja Weaverin malli on inspiroinut tietoliikenteen viestintää, on se alun perin suunniteltu verbaalista kommunikaatiota varten (Al-Fedaghi 2012: 12). Mallin kritiikkinä mainitaan muun muassa palautteen puuttuminen sekä mallin soveltumattomuus massakommunikaatioon. Vastaanottajalla on myös passiivinen rooli tässä mallissa, kun taas lähettäjä toimii tärkeimpänä tekijänä. (JMC Study Hub 2019). Al-Fedaghi (2012) mainitsee mallin olevan enemmän informaation kulun malli kuin itse viestinnän malli. Viestintä ei oikeasti ole niin suoraa kuin mitä viestintämalli antaa ymmärtää, vaan enemmänkin epäsuoraa.

Shannon ja Weaver itse mainitsevat kolme tasoa, joissa voi esiintyä ongelmia viestinnän aikana. Näitä ovat tekniset ongelmat, semanttiset ongelmat ja tehokkuuteen liittyvät ongelmat. Tekniset ongelmat ovat näistä helpoimmin selvitettävissä, ja näitä ongelmia varten malli on alun perin kehitetty. Semanttiset ongelmat viestinnässä on helppo löytää, mutta hankalampi selvittää. Näihin liittyvät esimerkiksi ongelmat sanojen ymmärtämisessä ja kulttuurillisissa eroissa, jotka voivat vaikuttaa viestinnän ymmärrettävyyteen. Tehokkuuteen liittyvät ongelmat taas keskittyvät siihen, miten viestintämallin kautta viestintä voidaan ymmärtää manipulaationa, sillä vastaanottajan rooli on hyvin passiivinen ja viestin sisällöstä vastaa nimenomaan lähettäjä. (Fiske 1990: 7)

Semanttisiin ongelmiin liittyy vahvasti myös se, ettei semanttisia järjestelmiä ole määritelty yhtä tarkkaan kuin teknisiin ongelmiin liittyviä signaaleja. Tämä tekee informaation mittaamisesta semanttisella tasolla hankalampaa, sillä merkitystä ei ole yhtä helppo mitata kuin teknistä, numeraalista tietoa. (Fiske 1990: 9)

### **2.1.2 Semioottinen viestintämalli**

Viestinnän merkitystä voidaan mallintaa ja tutkia monilla eri käsitteillä. Näitä voivat olla esimerkiksi merkki, merkitys, ja symbolit. Mallit, jotka perustuvat näihin, eroavat



erityisesti Shannonin ja Weaverin viestintämallista siinä, että ne eivät ole viestinnässä lineaarisia. Ne keskittyvät enemmänkin analysoimaan suhteita, jotka auttavat viestin merkityksen välittämisessä. (Fiske 1990: 39)

Semioottinen malli on yksi näistä viestinnän malleista. Semioottisessa mallissa on kolme pääaluetta. Näitä ovat 1. Merkki itsessään. Merkit ovat ihmisten kehittämiä, ja niihin liitetyt merkitykset ovat ihmisten laatimia. 2. Järjestelmät joihin merkit on järjestetty. Tällä tarkoitetaan jotakin koodia tai järjestelmää, joka on luotu jotakin laajempaa kokonaisuutta varten. Esimerkiksi liikennemerkki luovat tällaisen järjestelmän. 3. Kulttuuri, jossa nämä järjestelmät toimivat. Kulttuurissa valloilla olevat normit ja käytännöt. (Fiske 1990: 40).

Semiotiikka keskittyy pääasiassa tekstiin, toisin kuin lineaariset viestintämallit tekevät. Toinen suuri ero on vastaanottajan asema, joka semiotiikassa on paljon aktiivisempi kuin Shannonin ja Weaverin viestintämalliin perustuvissa malleissa. Semiotiikassa myös suositaan termiä "lukija", "vastaanottajan" sijaan. Tämä viittaa suurempaan aktiviteetin määrään kuin pelkkä vastaanottaminen. Lukeminen on myös kulttuurillista toimintaa, ja esimerkiksi samaa valokuvaa voidaan lukea usealla eri tavalla. (Fiske 1990: 40)

Semiotiikassa viestintä nähdään merkityksen tuottamisena viesteissä, kun taas Shannonin ja Weaverin viestintämallissa se nähdään mainitun koodaimen tuottamana. Merkitys ei ole koskaan yksiselitteinen käsite viesteissä, vaan se on aktiivista tulkintaa. Viestittely ei ole semiotiikassa vain viestin välittämistä pisteestä A pisteeseen B, vaan se on enemmänkin keskustelua ja neuvottelua, jossa kuljetaan koko ajan henkilöiden ja viestien välillä. (Fiske 1990: 46)

## **2.2 Onnistuneen viestinnän mittarit**

Onnistunut viestintä ja informaation välitys on hyvin moniulotteinen prosessi ja siinä on monta muuttujaa, jotka pitää ottaa huomioon, kuten aiemmista alaluvuista kävi ilmi.

Kyse ei ole pelkästä informaation välittämisestä erinäisten medioitten tai alustojen kautta vastaanottajille. Teknisen viestinnän informaation välitys on ymmärrettävän ja vastaanottajan tarpeita ja käyttäytymistä mukailevan tiedon luomista. Tällaisen viestinnän kannalta on otettava huomioon eri näkökulmia. Esimerkkeinä näistä on tiedon käytettävyys, tiedon ulkomuoto, vastaanottajan kognitiiviset ja psykologiset ominaisuudet, tietokone-ihminen-kommunikaation kysymykset sekä muita muuttujia kontekstin mukaan. (Albers 2012: 13)

Tiedon vastaanottaminen on hyvin subjektiivista, ja jokainen henkilö vastaanottaa tietoa oman filterinsä kautta. Tärkeimmäksi piirteeksi nouseekin aiemmin mainittu välitettävän informaation ulkomuodon ja sisällön mahdollisimman virheetön tila. Onnistuneen viestinnän ja tiedonvälityksen voidaan sanoa sisältävän seuraavat piirteet:

1. Selkeä ajatus siitä mitä tietoa välität ja miten
2. Tiedon virheettömyys ja soveliaisuus
3. Tiedon tiivis esittäminen
4. Tiedon täysi valmius ja aukottomuus
5. Palaute tiedon vastaanottavalta osapuolelta

Viestijällä on oltava selkeä ajatus siitä mikä hänen tavoitteensa on informaation välittämisellä ja myös siitä, miten hän tavoitteeseen haluaa päästä. Mitä parempi ymmärrys ja ajatus välitettävästä tiedosta ja tavasta lähettää sitä eteenpäin, sitä tehokkaammin sen onnistuu viestimään muille. (Akilandeswari 2015: 44)

Tiedon virheettömyys on tärkeää oikeakielisyyden, tiedon oikeellisuuden sekä myös oikean alustan valinnan kautta. Informaatio ja sen välitystapa on oltava soveliasta kulttuurillisesti sekä sosiaalisesti ja huomioon on otettava myös, että ajankohta ja väline, jolla informaatiota välitetään, on sovelias. On varmistettava myös, että välitettävä informaatio on soveliasta sitä vastaanottavalle kohderyhmälle. (Akilandeswari 2015: 44)

Tehokkaassa teknisessä viestinnässä informaation täytyy sisältää tarkka ja oikea sisältö oikeassa kontekstissa, mutta ei muuta. Sen on myös oltava järjestetty loogisesti ja sen on

siirryttävä mahdollisimman nopeasti sen tarkoitteeseen. Informaation on oltava tiiviiden lisäksi myös täysin aukotonta. Sen on sisällettävä kaikki se informaatio, mitä vastaanottaja haluaa tietää. Tiivistämistä ei voida tehdä sillä hinnalla, että informaation aukottomuus ja oikeellisuus kärsii. (Akilandeswari 2015: 45) Tämä on erityisen tärkeää ohjevideoissa ja manuaaleissa. Jokainen prosessi on kuvattava aukottomasti ja oikein, jos näin ei ole ja yksikin kohta on puutteellinen, on koko ohjevideo väärä koska pahimmassa tapauksessa tämä johtaa osan hajoamiseen tai muuhun vahinkoon.

Palaute välitetystä informaatiosta on keskeinen käsite tämän tutkielman kannalta ja tärkeä ominaisuus onnistuneessa viestinnässä. Tutkielman kyselyn perimmäinen tarkoitus on saada palautetta ohjevideon viestinnästä puhesynteesiä käytettäessä. Palaute on myös hieman irrallinen käsite muista listatuista piirteistä, koska se ei ole suoranaisesti onnistuneen viestinnän piirre vaan enemmänkin toimenpide, jolla varmistetaan, että viestintä on ollut onnistunutta. Informaatio voi muuttaa radikaalistikin ulkomuotoa siitä hetkestä, kun se lähetetään, siihen hetkeen, kun se vastaanotetaan. Sosiaalipsykologit ovat laskeneet, että tässä on keskimäärin 40-60% muutos viestin tarkoitteessa lähettäjältä vastaanottajalle. Tästä syystä on kriittisen tärkeää saada palautetta välittämstäsi tiedosta ja nimenomaan siitä, miten se on vastaanotettu. (Akilandeswari 2015: 45)

Esimerkkinä tästä, että jos olisimme sisällyttäneet puhesynteesin ohjevideoihin ilman minkäänlaista palautetta kohderyhmiltä, jotka ohjevideoita käyttää, emme olisi välttämättä koskaan tiedneet, josko puhesynteesi esimerkiksi haittaisi tiedon vastaanottamista häiritsevän paljon. Tekemällä kyselytutkimuksen ja saamalla palautteen kohderyhmiltä saamme vastauksen näihin kysymyksiin.

### **3 Tekniset käyttöohjeet ja ohjevideot**

Tekniset manuaalit ja kasvavissa määrin myös ohjevideot, ovat tuttu käsite käytännössä kaikille ihmisille, jotka ovat ostaneet jonkin uuden tuotteen. Olipa tämä tuote Ikean huonekalu, polkupyörä tai jopa valmisateria, kaikki nämä sisältävät jonkinasteisen asennus, huolto tai käyttöohjeen. Huonekalun kokoamisohje, polkupyörän huolto-ohje tai valmisaterian lämmitys ohjeet ovat kaikki omalla tavallaan teknistä viestintää sekä käyttö-ohjeita, joilla ohjeistetaan tuotteen käyttämisessä. Käyttöohje on näin ollen jopa pakollinen osa tuotetta. (Pöyhönen & Tiusanen 1991: 7)

Yleisin mielikuva, joka ihmisille tulee tuotteen mukana tulevista ohjeista, on valitettavan usein välinpitämätön tai jopa negatiivinen. Tämä johtuu pitkälti siitä, että manuaaleihin ja teknisiin ohjeisiin ei ole panostettu ja niitä ei ole koettu ensisijaisen tärkeänä uuden tuotteen lanseerauksessa ja toimituksessa (Maijala, Nordlund & Virolainen 1987: 5). Teknisen viestinnän vakinaistettua asemansa, tämä on pikkuhiljaa muuttumassa. Yritykset kokevat nykyään ensisijaisen tärkeäksi panostaa laadukkaaseen asennusohjeeseen, koska tämä kulkee käsi kädessä tuotteen ja tätä kautta yhtiön imagon kanssa.

Esimerkkinä käyttöohjeen laadun merkityksestä voi käyttää halvalla tuotettua ja halvalla hinnalla ostettua tuotetta, jossa on hyvin pelkistetty ja usein vajavainen ohje mukana. Jos tämä samantasoinen ohje tulisi vaikka Applen, Samsungin tai muun vastaavan ison yrityksen uuden puhelimen mukana, laskisi tuotteen ja yhtiön imago käyttäjän silmissä, vaikka tuote olisikin laadultaan hyvä. Tähän totuuteen ollaan herätty jo monessa yhtiössä ja tätä kautta myös tekninen viestintä on noussut tärkeämmäksi osaksi monen yhtiön hierarkiassa.

#### **3.1 Teknisestä ohjeesta tekniseksi ohjevideoksi**

Teknisissä ohjevideoissa esitetään tarkalleen se tieto, joka on esitettyä teknisissä ohjeissa. Ohjevideot noudattavat myös samoja yksiselitteisyyden ja aukottoman

ohjeistuksen tekemisen periaatteita kuten käyttöohjeetkin. Videoissa etuna on tekstin ja äänen käyttö havainnollistavina keinoina. Käyttöohjeissa pääpaino on mahdollisimman yksiselitteisissä kuvissa ja symboleissa.

Videoissa pyritään näyttämään nämä asennuskuvat ja symbolit elävänä kuvana säilyttämällä kuitenkin mahdollisimman hyvin yksiselitteisyyden myös videokuvassa. Äänen käyttö ja toimenpiteen avaaminen puheella edistävät kuitenkin videon informaation siinä säilyttämistä sekä lisäävät videolle myös arvoa miellyttävämmän kuuntelukokemuksen kautta. Tähän luonnollisesti vaikuttaa se, miten vastaanottaja kokee kuulemansa äänen.

### 3.1.1 Audiovisuaalisuus

Audiovisuaalisuus liitetään usein elokuvaan, televisioon ja digitaaliseen mediaan, mutta sen tarkka määritelmä on Mikko Lehtosen (2007) mukaan enemmän ruumiillinen ja aistillinen kokemus. Termi audiovisuaalisuus itsessään koostuu sanoista *audio* ja *visuaalisuus*. Audiolla viitataan nimenomaan muotoon ”minä kuulen”, eli toimintaan, jossa kuulija itse on toimijana. Visuaalisuus taas viittaa katseeseen, eli siihen mihin kuulija katsoo, eli visualisoi (Lehtonen 2007: 30).

Audiovisuaalisuudessa yhdistyvät eri aistit, kuuleminen ja näkeminen, ja se onkin tämän vuoksi ruumiillista toimintaa. Näkijän ja nähdyn suhde ei kuitenkaan kaikessa audiovisuaalisuudessa ole välttämättä samanlainen kuin kuulijan ja kuullun suhde. Näkeminen edellyttää aina samaa tilaa, ellei kyseessä ole jokin digitaalisen median väylä. Kuuleminen taas onnistuu esimerkiksi viereisestä huoneesta asti. (Lehtonen 2007: 30)

Kuulemista aistina ei usein pidetä yhtä autenttisena kuin näkemistä. Tästä hyvä esimerkki on tunnettu sanonta ”näin sen omin silmin”. Visuaalisuuteen liittyy siis vahvoja läsnäolon mielikuvia, joissa näkijä on ollut itse paikalla. Kuulemista taas voidaan usein vähätellä huhupuheeksi tai väärin muistetuksi tiedoksi. (Lehtonen 2007: 31)

Yhdessä kuuleminen ja näkeminen, eli audiovisuaalisuus, viittaa Lehtosen mukaan ennen kaikkea välitettyyn kuvaan ja ääneen. Hän tuo ilmi sen, miten audiovisuaalisuutta usein tarkastellaan nimenomaan digitaalisen median valossa. Lehtosen mukaan harvoin taas on esillä se, että audiovisuaalisuus liittyy vahvasti ihmisen ruumiillisuuteen ja aistimellisuuteen. (Lehtonen 2007: 31)

Tästä audiovisuaalisuuden merkityksen ymmärtämisestä huolimatta audiovisuaalisuus liitetään hyvin vahvasti elokuvaan, televisioon ja muuhun videomateriaaliin. Audiovisuaalista mediaa pidetään joustavana välineenä, jonka avulla on helppo levittää tarvittavaa sanomaa kuulemisen ja näkemisen aistien kautta. (Aaltonen 2018: 17) Audiovisuaalisuus digitaalisen median väylänä poikkeaa Lehtosen aikaisemmin kuvaamasta läsnäolosta eniten ehkä siinä, että digitaalisen median audiovisuaalinen tuote, esimerkiksi ohjevideo, on tarkkaan suunniteltu jotakin tarvetta varten, käsikirjoitettu, sekä tuotettu helposti levitettävään muotoon. Läsnäoloon liittyvä audiovisuaalinen media saattaa taas olla spontaani hetki, eikä sitä pystytä uudelleen toistamaan.

### **3.1.2 Kuuntelukokemuksen rooli informaation välityksessä**

Audiota ja näin myös audiovisuaalisuutta voidaan mitata eri tavoin informaation välityksen yhteydessä. Yksi näistä mittareista on kuuntelukokemus *engl. overall listening experience*, jonka perusteella voidaan mitata kuuntelijan kiinnostusta kuunneltavaa audiota kohtaan. Kuuntelukokemuksen käsitteen avulla voidaan yhdistää useita eri kuunteluun liittyviä tekijöitä, jotka saattaisivat vaikuttaa kuuntelijan arviointiin, mielentilaan ja jäljentämisen, eli muistamisen, mahdollisuutta. (Walton & Evans 2018)

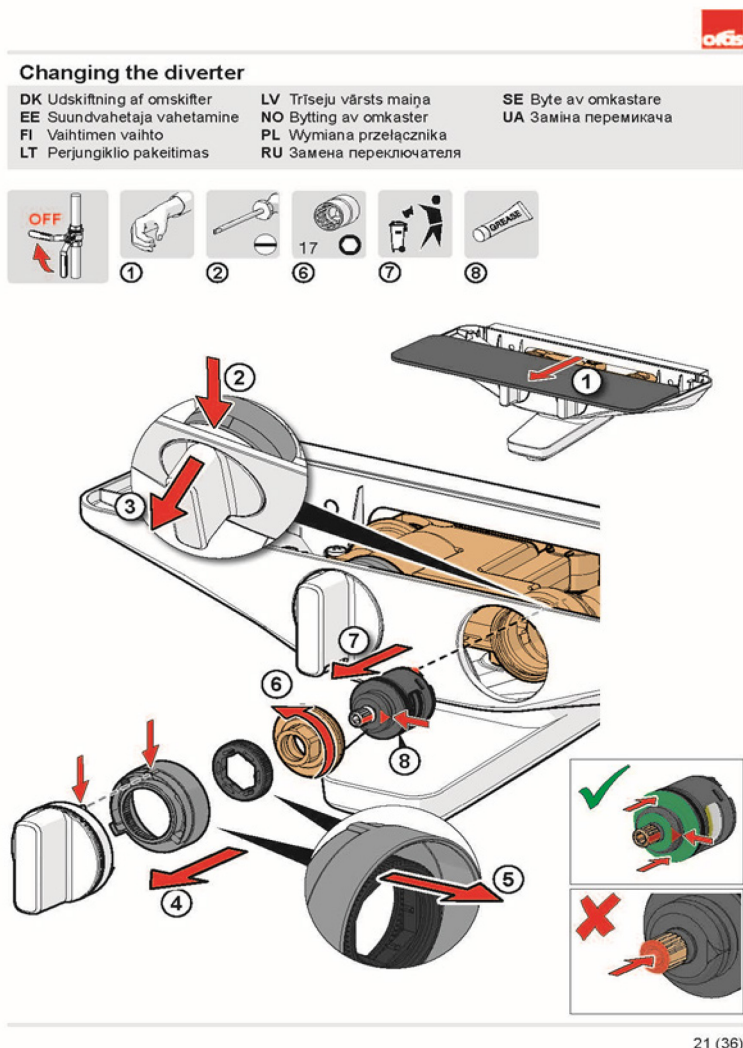
Kuuntelukokemus on audion laadun arvioimisen keino, jonka Schoeffler ja Herre määrittelevät käytettävän kuvaamaan nautittavuutta tiettyä audiota kuunnellessa. Tähän arviointiin otetaan mukaan kaikki mahdolliset tekijät, joilla voidaan parantaa kuuntelijan nauttimista audion kuuntelussa. Näitä ovat esimerkiksi erilaiset aistilliset-, kognitiiviset-, sekä psykologiset seikat. (Schoeffler & Herre 2014)

Kuuntelukokemusta (*Overall listening experience*) verrataan, ja siitä jopa puhutaan usein kokemuksen laatuna (*Quality of experience*) kun on kyse audion arvioinnista. Toisaalta näissä kahdessa kokemuksessa yhdistyvät teoreettisesti orientointuneet mallit, joissa pääpaino on kognitiivisella ja psykologisella näkökulmalla. Jotkut mallit ovat kuitenkin hyvin käytännönläheisiä, ja ne keskittyvät usein empiiriseen dataan audion kuuntelusta. On tärkeää huomioida mitä tahansa audiota arvioitaessa, että kuuntelijan omat mieltymykset ja luonne vaikuttavat kuuntelukokemukseen aina jonkin verran. (Schoeffler & Herre 2014)

Audiossa ja audiovisuaalisuudessa kuuntelukokemuksen mittaaminen on aina riippuvainen siitä, että arviointiin saadaan kuulijoita (Walton & Evans 2018). Audiovisuaalisuudessa mukana taas on myös visuaalisuus, joka tukee mukana olevaa audiota. Kun jotakin palvelua tai järjestelmää suunnitellaan, voidaan kuuntelukokemusta käyttää kohderyhmän tarkentamisessa ja suunnittelussa. (Walton & Evans 2018) Näin ollen myös audiovisuaalisuuden kanavia kuten ohjevideoita suunniteltaessa ja tehdessä, arvioimalla ja mitaamalla kuuntelukokemusta voidaan koko videon laatua parantaa.

### **3.2 Huolto- ja käyttöohjeet Oras Group Oy:ssä**

Huolto- ja käyttöohjeet ovat Oras Group Oy:ssä isossa roolissa. Syy tähän on tuotekategoria sekä LVI-alan luonne, jossa käytännössä aina asennetaan tai huolletaan jotain. Tästä syystä erityisesti on tärkeää, että ohjeistus materiaali on riittävää ja selkeää. Yhtiö on panostanut huolto-ohjeiden päivittämiseen viime vuodet ja tämän päivityksen tuloksia nähdään nykyisissä ohjeissa (ks. Kuva 1).



**Kuva 1.** Esimerkki hanan huolto-ohjeesta (Oras Ltd. 2019).

Kuten kuva 1 osoittaa Oraksen ohjeissa on tekstiä pelkästään otsikoissa ja käytössä on tekstin sijaan universaaleja symboleita ja havainnollistavia kuvia. Tekstittömyys on kustannustehokasta, koska sillä vältetään käännöskustannukset. Lisäksi se on myös suoraviivaisempaa ja välittömämpää kommunikaatiota kuin pelkkä teksti.

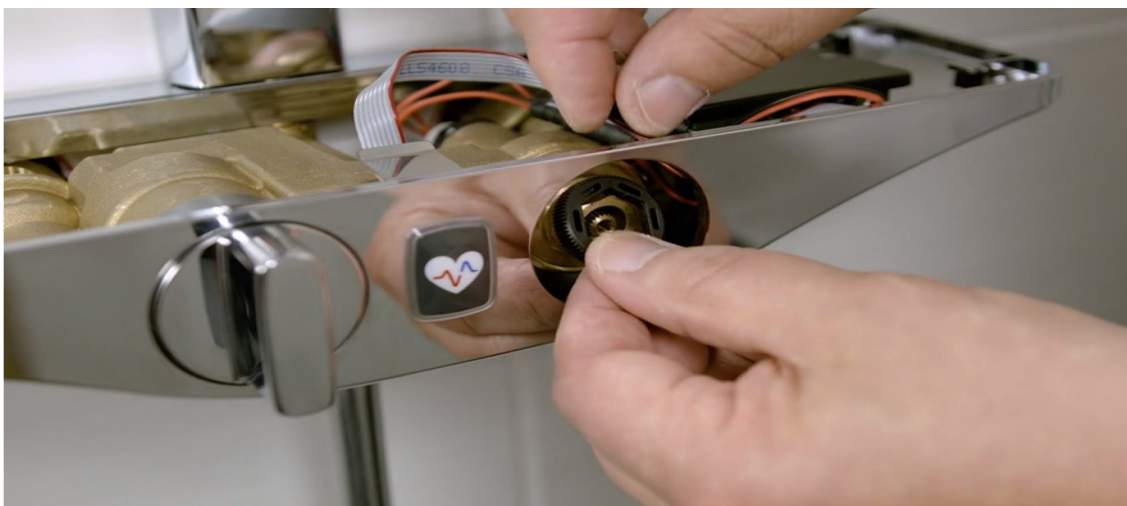
Huolto-ohjeiden ja käyttöohjeiden on aina oltava mahdollisimman yksiselitteisesti ymmärrettäviä, mutta LVI-alalla tämä korostuu erityisesti, koska väärinkäytettynä hanatuotteissa on itsensä tai muiden vahingoittamisen riski. Tästä esimerkkinä sairaalassa käytetyt hygieniahuuhtelun sisältävät hanatuotteet, joissa huuhteluveden lämpötila voi ylittää



75 astetta. Kyseinen tuote on asennettava täysin oikealla tavalla, ja sitä on myös käytettävä oikealla tavalla. Tämä varmistetaan luomalla aukoton ja yksiselitteinen ohjeistus niiden käytön, asennuksen kuin huollonkin kannalta. Luonnollisesti on aina olemassa riski ihmisten tekemiin virheisiin, mutta tätäkin riskiä alennetaan sillä, että ohjeistus on mahdollisimman helposti ymmärrettävä ja että väärinkäsityksille ei ole mahdollisuutta.

### 3.3 Tekniset ohjevideot ja niiden kohderyhmät Oras Group Oy:ssä

Vastaan ohjevideoiden luonnista Oras Groupissa ja työhöni kuuluu tietokoneella tehdyt animoinnit sekä perinteiset ”hands-on”-kuvaukset asennuksista tai muista tuotteisiin liittyvistä toimenpiteistä. Tärkeää videoiden luonnissa on se, että videon sisältö on täysin käyttöohjetta vastaava. Toisin sanoen videolla havainnollistetaan käytännön tasolla se mitä käyttöohjeessa ohjeistetaan. Videoilla näytetään myös asioita, joita käyttöohjeissa ei esitetä tai esitetään käyttöohjeessa hyvin yleisesti. Näitä ovat esimerkiksi elektronisen hanan tai suihkun erilaiset toiminnot, päivitettyjen tai kokonaan uusien varaosien toiminta sekä erinäiset huoltotoimenpiteet.



**Kuva 2.** Kuvankaappaus ohjevideosta (Oras Ltd. 2019).

Ohjevideoiden kohderyhmät voidaan karkeasti jakaa ammattikäyttäjiiin: putkimiehet, huoltomiehet, myynti, koulutus, asiakaspalvelu jne. Tämä käyttäjäryhmä edustaa

ammattikäyttäjiksi luokiteltavaa ryhmää. Toinen kohderyhmä on arkikäyttäjät eli hanan omistajat ja loppukäyttäjät. Tästä ryhmästä karsitaan yllä olevat ammattikäyttäjät; arki-käyttäjien ryhmään kuuluvat siis ne henkilöt, joilla ei ole alan koulutusta tai muuta koke-musta hanojen huoltoon tai asennukseen liittyvistä toimenpiteistä. Ohjevideot ovat kui-tenkin suunnattu molemmille ryhmille.

Videoilla tapahtuva toimenpide voi vaihdella hyvin paljon vaikeustasoltaan, joten tästä syystä on erittäin tärkeää pitää mielessä käyttäjäryhmien välisen kokemustason ero. Vi-deon, kuten ohjeen, on oltava virheetön, yksiselitteinen ja helppo seurata. Karkeasti sa-nottuna kuitenkin huoltoon ja käyttöön liittyvät videot ovat enemmän arkiryhmän käy-tössä, kun taas asennusvideot ovat enemmän ammattiryhmän käytössä. Tähän vaikuttaa esimerkiksi kerrostaloasumisen yleisyys Suomessa, joissa hanojen omatoiminen asenta-minen ei ole suotavaa.

## 4 Puhesynteesin käyttö ja arviointi

Puhesynteesi, joka tunnetaan myös nimellä tekstistä-puheeksi (engl. text-to-speech, TTS) on ihmispuheen kaltaisen puheen tuottamista keinotekoisin tavoin. Tekstistä-puheeksi voidaan tarkentaa tarkoittavan kaikkea tekstistä puhutuksi puheeksi tuotettua ääntä. Taylor (2007: 25) tiivistää puhesynteesin (ja TTS:n) tutkimuksen näin:

*This field of study is known both as speech synthesis, that is the “synthetic” (computer) generation of speech, and text-to-speech or TTS; the process of converting written text into speech.*

Puhesynteesi on alana siinä mielessä mielenkiintoinen, että suurin osa ihmisistä tietää keinotekoisesta robottiäänestä, mutta välttämättä he eivät tiedä tarkalleen, mistä on kyse. Ihmiset ovat varmasti myös tulleet kontaktiin puhesynteesin kanssa jossain vaiheessa elämäänsä, esimerkiksi viihdemedian tai sci-fi-elokuvien yms. kautta. Puhesynteesin luonti ja sen mekaniikka eivät myöskään välttämättä ole vielä yleistietoa ihmisten keskuudessa.

### 4.1 Puhesynteesijärjestelmät

Teknisesti puhesynteesin tuottamiseen on nykypäivänä monia eri tapoja. Puhesynteesin historian voi kuvata alkaneen vuonna 1779 Pietarissa, Venäjällä. Professori Christian Katzenstein halusi kuvata viiden eri ”pitkän” vokaalin eroa, joten hän päätti luoda koneen, joka tuottaa keinotekoisesti nämä viisi vokaalia (a,e,i,o,u). Tämä kone käytti akustisia resonaattoreita, jotka aktivoitiin ruokojen värinällä. (Flanagan 1972: 205)

Kehitys on jatkunut näistä mekaanisista ääntä tuottavista laitteista, elektronisiin laitteisiin, josta kehitys siirtyi tietokoneellistumisen myötä järjestelmiin ja ohjelmistoihin. Nykyiset käytetyimmät menetelmät puhesynteesissä ovat artikulatorinen synteesi, Hidden Markov Method (HMM)-pohjainen synteesi, formanttisynteesi ja konkatenaatiosynteesi (Siddhi, Verghese & Bhavik 2017: 26-30).

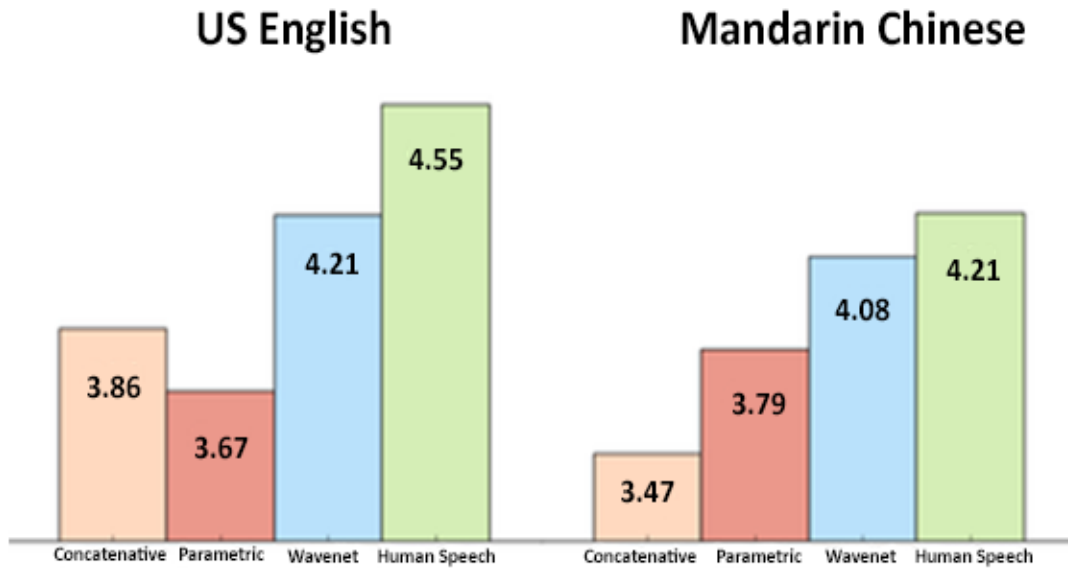
Näistä eniten käytössä on 1990-luvulla esiin nousseeseen Hidden Markov-metodiin perustuvat ns. parametriset puhesynteesijärjestelmät. Tässä metodissa teksti ja siitä tuotettu puhesynteesi luodaan käyttämällä hyväksi stokastisesti luotuja akustisia malleja. Toisin sanoen tekstistä luotu puhesynteesin ääni on sattumanvaraista tiettyjen parametrien sisällä. Puhesynteesi luodaan hierarkisella mallilla, jossa käydään taso tasolta läpi lähdetekstistä otetut parametrit. Näitä parametrejä ovat kielen ominaisuudet: prosodia, fonetiikka sekä kieliopillisuus. Näistä parametreistä valitaan hierarkisessa mallissa todennäköisimmät akustiset ominaisuudet ja nämä lähetetään ääniaallon syntetisaattoriin, josta lopulta tuotetaan puhesynteesin puhe. Tällä mallilla tuotettu puhe on kuitenkin hyvin satunnaista olomuodoltaan ja tasoltaan häilyvää. Suurin ongelma tässä mallissa on tuotetun puhesynteesin äänen luonnollisuuden puute verrattuna ihmispuheeseen. (Deng & Yu 2014: 286-287).

Uusin ja nopeasti suosiota saava metodi puhesynteesin alalla on kuitenkin ns. Deep learning ja sitä käyttävät järjestelmät. Tällä metodilla on kohdennettu vanhan parametrisen mallin ongelmakohtia akustisessa mallinnuksessa ja äänen luonnollisuudessa. Deep learning-puhesynteesin metodissa järjestelmä ”oppii” tarjolla olevasta datasta akustisen kartoituksen ihmispuheelle ja parametreille, jonka avulla se lähtee työstämään annettua tekstipohjaista syöttöä puheeksi. (Deng & Yu 2014: 286-288)

Deep learning juontaa Deep Neural Network (DNN)-tutkimussuunnasta, joka pyrkii mallintamaan biologisia neuraaliverkostoja. Deep learning käsitteenä tarkoittaa koneoppimista monilla eri tasoilla yhtäaikaisesti ja apuna käytetään keinotekoisia neuraaliverkostoja (Deng & Yu 2012: 5).

Aikaisempaa tutkimusta puhesynteesien tuottamiseen käytetyistä malleista on tehty muun muassa puhesynteesin tämän hetkisen kehityksen kärkeä edustavan Googlen Deep Mind-projektin puhesynteesiä luovasta järjestelmästä. Kyselytutkimuksessa vertailtiin Deep learning-konseptilla luotua puhesynteesiä perinteisiin tapoihin (konkatenatio, HMM) käyttäen verrokkina ihmispuhetta. Kuvion 5 mukaisesti Deep learning-

konseptilla tuotettu puhesynteesi todettiin perinteisiä puhesynteesejä laadukkaammaksi MOS-asteikkoa (1-5) hyväksikäyttäen. (Wavenet: A Generative Model for Raw Audio 2016)



**Kuvio 5.** Esimerkki MOS-asteikosta käytännössä

Kuvio osoittaa miten Deep Learning-konseptien avulla luotu WaveNet sai huomattavasti paremman MOS-arvon (4.55 ja 4.21) verrattuna perinteisiin konkatenaatio- ja parametrisiin puhesynteesi ääniin.

Tutkimuksessa käytetään Deep learning-konseptia käyttävää Amazonin Polly järjestelmää. Amazon Polly valittiin kyselyn ohjevideossa käytettäväksi puhesynteesiksi taustatyön tuloksena. Taustatyönä tutkin puhesynteesijärjestelmien markkinatilannetta ja puhesynteesijärjestelmien kyvykkyyttä ja suosiota. Amazon Polly osoittautui yhdeksi suosituimmista ja laadukkaimmista järjestelmistä tällä hetkellä, joten oletuksena oli, että se antaisi parhaan yleiskuvan nykyisestä puhesynteesin tasosta. Amazon Pollyn verkkosivut korostavat myös Pollyn käytön monipuolisuutta niin henkilökohtaisessa, mutta erityisesti

ammattimaisessa käytössä, mikä oli tärkeä tekijä valintaa tehdessä. (Amazon Web Services 2019)

Amazon Pollyn käyttö ei tunnu rajoittuvan myöskään mihinkään tiettyyn alaan, vaan sen käyttö on varsin laajaa. Sitä käytetään muun muassa Duolingo-sivustolla ja-palvelussa vieraan kielen opetukseen, kodin turvajärjestelmien äänenä Y-Cam-turvajärjestelmäytössä sekä Intian ensimmäisessä henkilökohtaisen avustajan sovelluksessa Haptikissa. Ohjevideoita eniten vastaava palvelu on ehkä kuitenkin GoAnimate-palvelu, joka on *do-it-yourself*-sivusto animoitujen videoiden luomiseen. Kyseinen sivusto käyttää Amazon Pollyä hyödykseen äänenluomiseen videoille ja palvelua käytetäänkin paljon juuri ammattimaisten videoiden luomiseen. GoAnimate:n perustaja Alvin Hung korostaakin lainauksessaan Amazon Pollyn sivuilla puhesynteesijärjestelmän äänenluonnin nopeutta ja kustannustehokkuutta eritoten verrattuna aidon ihmisäänen nauhoittamiseen. (Amazon Web Services 2019)

Yllä mainittu kustannustehokkuus ja puhesynteesin puheen laatu olivatkin tärkeitä tekijöitä puhesynteesi-järjestelmää valittaessa. Amazon Polly osoittautui alustavien testien ja tutkielmassa käytettyjen videoiden luonnissa hyvin tehokkaaksi ja laadukkaaksi järjestelmäksi, joten valinta oli lopulta helppo tehdä.

## 4.2 Puhesynteesin käyttötavat

Puhesynteesin käytön voi luokitella nykypäivänä hyvin yleiseksi. Sen yleisimmät käyttökohteet ovat paikkoja tai palveluita, joita ihmiset käyttävät melkein joka päivä, joten sen läsnäolosta on tullut monella tapaa jopa huomaamatonta. Esimerkkejä on kerrosten kullutukset hisseissä, auton navigaattorin ääni ja erinäiset käyttöliittymät (Siri, Alexa, Google Assistant). Puhesynteesi on levinnyt kuitenkin näistä yleisimmistä tavoista laajalle alueelle niin ammatillisessa kuin harrastekäytössäkin.

Ehkä suurin yksittäinen hyöty, joka puhesynteesillä on saatu aikaan, on erinäisistä aistivammoista kärsivien ihmisten elämän helpottaminen. Sokeille ihmisille on olemassa jo vakiona selaimissa, niin tietokoneella kuin matkapuhelimellakin ohjelmia, jotka lukevat ääneen verkkosivuja sekä muita medioita. Tietokoneen käyttöjärjestelmän ohjaaminen ”kommunikoimalla” sen kanssa puhesynteesin ja puheen tunnistuksen yhdistävällä ohjelmalla on myös kätevämpää kuin esimerkiksi sokeille tarkoitetun pistekirjoitusta sisältävän näppäimistön käyttö. Samalla tapaa kuurot ja mykät hyötyvät puhesynteesiä luovista järjestelmistä, joiden avulla he voivat kommunikoida helpommin esimerkiksi sellaisten ihmisten kanssa, jotka eivät osaa viittomakieltä. Molemmissa näissä tapauksissa nousee tärkeäksi puhesynteesin äänen laatu. Tutkielmassa pääroolissa olevien ymmärrettävyyden ja luonnollisuuden on oltava tarpeeksi hyvällä tasolla, jotta edellä mainittuja avustavia ohjelmia ja järjestelmiä voidaan käyttää. (Hande 2014: 8061-8062)

E-oppimisalustat ovat myös hyvin yleinen käyttökohde puhesynteesille. Kielten oppiminen e-alustojen kautta on helpottunut huomattavasti, ja esimerkiksi ääntämisen tarkistaminen onnistuu helpommin puhesynteesin ja puheen tunnistuksen avulla. E-oppimisalustojen peruseriaate on ollut opintojen helpottaminen ja niiden välitön saavuttaminen. Näille alustoille voi mennä koska tahansa ajasta ja paikasta riippumatta, ilman että opetuksen taso laskee. (Ren & Miao 2008: 432-433) Puhesynteesi vaikuttaa tähän, koska paikalla on aina ”opettaja”, joka ohjeistaa ja pohjustaa puheellaan luentoja sekä tehtäviä harjoituksia.

Asiakaspalvelualat ovat ottaneet laajenevissa määrin käyttöönsä puhesynteesin teknologioita. Esimerkiksi monessa suomalaisessa teleoperaattorissa asiakaspalveluun soitettaessa vastaa puhesynteesin ääni. Monet julkisen paikan käyttöliittymät sisältävät myös usein puhesynteesiä. Tästä yksi esimerkki on Taiwanissa kansallisessa teknologisessa yliopistossa toteutettu palvelurobotille tehty ”meet and greet”-käyttöliittymä, joka yhdistää puhesynteesin sekä puhetunnistuksen. Robotti vastaanottaa yliopistolle tulijat puhesynteesillä, ja tietyt avainsanat kuultuaan vastaa tulijalle sopivalla tavalla. Robotin ja yliopistolle tulijoiden välisestä kommunikoinnista tehdyssä tutkimuksessa koettiin

tärkeimmäksi seikaksi robotin ihmismäisyys, mikä oli suoraan yhteydessä sen äänen luonnollisuuteen ja ymmärrettävyyteen. Tutkimuksessa suoritettiin kysely robotin käyttäjillä ja tässä tapauksessa luonnollisuus ja ymmärrettävyys oli hyvällä tasolla, eli puhe-synteesi oli onnistunut tehtävässään. Mielenkiintoista tutkimuksen tuloksissa oli se, että ihmiset, jotka kommunikoivat pidempään robotin kanssa antoivat korkeamman arvosa-nan puhesynteesille. (Huang & Lu 2014: 150-155)

Tämä herättää kysymyksiä puhesynteesin kokemisesta ja siitä, miten suuri vaikutus sillä on miten ja millä välineellä puhesynteesiä vastaanotetaan. Tässä tapauksessa kyseessä oli fyysinen robotti ihmismäisellä kasvolla, jonka kanssa kommunikoida. Tutkielman ta-pauksessa kyse on pelkästään ohjevideon äänestä, jolla ei ole fyysistä representaatiota.

### **4.3 Puhesynteesi teknisissä ohjevideoissa Oras Group Oy:ssä**

Puhesynteesin käyttöä teknisissä ohjevideoissa on kartoitettu Oraksella taloudelliselta ja tekniseltä kannalta. Tämän tutkielman avulla selvitetään sitä, miten käyttäjät vastaanot-tavat puhesynteesin.

Kartoitusta on jo tehty kustannusten sekä teknisten mahdollisuuksien osalta. Puhesyn-teesin kustannustehokkuus on yhtiön kannalta selvästi suurin etu sen mahdollisessa käyttöönnotossa. Puhesynteesin käyttö ammattimaisesti on hyvin edullista, varsinkin jos vertaa voice-overien tilaamisen kustannuksia alihankkijalta, verrattuna puheen luomi-seen Amazon Pollyllä. Amazon Pollyn hintataso vaihtelee ilmaisesta tasosta, johon kuu-luu 5 miljoonaa kirjoitusmerkkiä kuukaudessa ensimmäisen vuoden ajan, maksulli-seen ”Pay-as-you-go”-malliin, jossa 1 miljoona kirjoitusmerkkiä maksaa keskimäärin 4 dollaria. Hinta määräytyy käytön mukaan, joten hinta on hyvin alhainen per kirjoitus-merkki. Alihankkijalta hankittuna kirjoitusmerkkiä kohden hinta nousee Amazon Pollyyn verrattuna, hyvin korkeaksi. (Amazon Web Services 2019)



Esimerkkinä Amazon Pollyn Prices-sivulla annetaan Mark Twainin 224 sivuinen klassikko Huckleberry Finnin seikkailut-teos, jonka konvertoiminen puhesynteesin puheeksi maksaisi 2.40 dollaria. Teos sisältää 600 000 kirjoitusmerkkiä 224 sivussa ja käyttäen tätä pohjana, Oraksen keskiverto ohjevideo sisältää noin 1-2 sivun verran puhetta ja noin 2000-7000 kirjoitusmerkkiä per video. Kustannukset ovat näin per video 0.01-0.05 sentin välissä. Uudet projektit toki sisältävät useita videoita, mutta jopa yliarvioiden lasketulla 10 uuden videon projektijulkaisulla kustannukset nousevat maksimissaan puolen dollarin luokkaan. (ks. Amazon Web Services 2019)

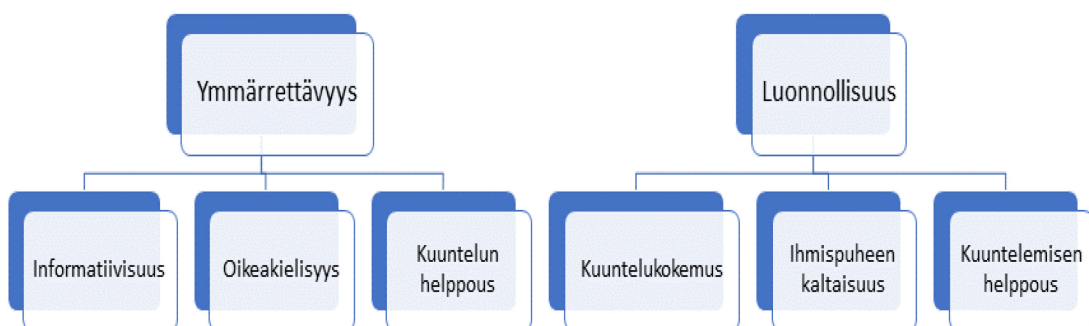
Teknisten mahdollisuuksien kartoitus tarkoitti käytännössä oikeiden ratkaisujen löytämistä puhesynteesin sisällyttämiselle ohjevideoihin. Tämä kattoi äänitiedostojen luomisen vaikeustason, äänityypin sekä äänen muokkaamisen vaihtoehdot, äänitiedostojen sisällyttämisen ohjevideoihin tekniseltä kannalta, sekä muita yleisen tason kysymyksiä kuten tiedostokoko ja tiedostotyypit. Tekniset kysymykset osoittautuvat helposti ratkaistaviksi, ja Amazon Pollyn käyttöjärjestelmä sekä ohjeistus on hyvin selkeää erityisesti äänen optimoinnin ja äänityypin kannalta. Kielivaihtoehtoja löytyy 28, mikä on myös iso etu kansainvälisessä yhtiössä. Keskiverto tiedostokoko on 0,5-1,5 mbit, mikä on hyvin kompakti koko tallentamisen ja säilyttämisen kannalta. Tuettuja tiedostomuotoja Amazon Pollyssä on: MP3, Vorbis sekä raaka PCM audio. Näistä MP3 on yleisin ja universaalisti käytetyin ja tätä käytämme myös Oras Groupissa. (ks. Amazon Web Services 2019)

Tekniseltä ja taloudelliselta kannalta puhesynteesin käyttöönotto on näin ollen yhtiön kannalta hyödyllistä. Puhesynteesin käytettävyys, sen välittömyys sekä käyttämisen helppous ovat myös tärkeitä piirteitä puhesynteesin mahdollisessa käyttöönotossa. Tärkein aspekti on kuitenkin se, miten alalla olevat ihmiset ja tuotteiden kuluttajat ottavat puhesynteesin vastaan.

#### 4.4 Puhesynteesin arvioinnin kriteerit

Puhesynteesin arviointiin on vakiintunut MOS-skaala, kuten aiemmin tutkielmassa mainittiin. Tämän skaalan avulla tutkitaan kahta selkeästi puhesynteesissä vakiintunutta kriteeriä eli luonnollisuutta ja ymmärrettävyyttä. (Le Maguer 2018: 10-12, 25-26)

Puhesynteesin aiempi tutkimus on keskittynyt näihin kahteen, mutta enemmän keinotekoisien puheen tuotannon ja siihen liittyvien teknisten ominaisuuksien ja aidolta kuulostavan puheen luonnin näkökulmasta, kun taas tutkimus siitä, miten ihmiset vastaanottavat puhesynteesiä ja mitä tunteita tai ajatuksia heille siitä syntyy, on jäänyt vähemmälle huomiolle. (Chung, Wang, Hsu, Zhang, Skerry-Ryan 2018: 1-5) Tutkittaessa ymmärrettävyyden ja luonnollisuuden näkökulmasta sitä, miten ihmiset vastaanottavat puhesynteesiä, on tärkeää jakaa nämä kaksi pääkäsitettä pienempiin ja tarkempiin alakäsitteisiin. Käsitteet ja alakäsitteet on kuvattu kuviossa 6.



**Kuvio 6.** Kyselytutkimuksen käsitteet

Kuviossa 6 on esitelty ymmärrettävyyden ja luonnollisuuden kaikki alakäsitteet. Esittelen seuraavissa alaluvuissa tarkemmin kaikki alakäsitteet ja sen, miten näitä arvioidaan kyselyssä.

#### 4.4.1 Luonnollisuus puhesynteesin kriteerinä

Luonnollisuuden käsitteen määrittelyyn käytän tässä tutkielmassa Viswanathanin määrittelyä, jossa luokitellaan luonnollisuutta ihmisperheen kaltaisuuden, kuuntelemisen helppouden sekä kuuntelun mukavuuden ja äänen seljyvyyden avulla. Yhdistän kuuntelemisen mukavuuden sekä äänen seljyvyyden tässä tutkielmassa kuuntelukokemuksen käsitteen alle. Nämä kaikki, kuten koko tutkielmakin, on tehty ihmisen näkökulmasta, koska siinä selvitetään mitä tunteita puhesynteesi herättää ja miten osallistujat vastaanottavat puhesynteesin. Luonnollisuuden käsite on pohja mittaamiselle yhdessä ymmärrettävyyden kanssa. (Viswanathan 2005: 63)

Luonnollisuutta tarkennetaan kuuntelukokemuksen, ihmisperheen samankaltaisuuden sekä kuuntelemisen helppouden kautta. Kuuntelukokemus tarkoittaa puhesynteesin kuuntelemisen kokemuksen miellyttävyyttä. Ihmisperheen samankaltaisuus kattaa alleen miten samanlaisena puhesynteesi mielletään ihmisperheeseen verrattuna. Kuuntelemisen helppoudella haetaan tässä tapauksessa puhesynteesin puheen rytmitystä sekä sointuvuutta ja sitä kautta sen kuuntelemiseen vaadittavaa ponnistelua. (Viswanathan 2005: 63)

Luonnollisuuden alakäsitteitä on mitattu kyselyssä kahdella kysymyksellä. Jokaisella alakäsitteellä on kaksi kysymystä luokiteltuna, ja näiden kahden kysymysten vastauksien tuloksista lasketaan jokaiselle käsitteelle keskiarvo. Tällä keskiarvolla mitataan, miten alakäsite on koettu kyselyn vastaajien toimesta. Liitteessä 1 ja 2 on esitetty kyselylomakkeet kokonaisuutena, sieltä voi tarkistaa alakäsitteille omistetut kysymykset. Kuuntelukokemuksesta on mitattu kysymyksillä 1 ja 6, ihmisperheen kaltaisuutta on arvioitu kysymyksillä 4 sekä 8 ja kuuntelemisen helppoutta on identifioitu kysymyksillä 3 ja 10. Yhdessä näiden kolmen alakäsitteen kysymysten tulokset luovat luonnollisuuden pääkäsitteen tuloksen.

#### 4.4.2 Ymmärrettävyys puhesynteesin kriteerinä

Ymmärrettävyyttä tarkennetaan informatiivisuuden, oikeakielisyyden sekä kuuntelun helppouden avulla. On tärkeää huomata samankaltaisuudet luonnollisuuden alakäsitteiden kanssa. Kuuntelun helppous/kuuntelemisen helppous on yhtäläinen käsite molemmissa. Näissä erona toimii se, että luonnollisuudessa käsitellään kuuntelun helppoutta itse puheen rytmityksen ja miellyttävyyden kautta. Ymmärrettävyydessä kyse on enemmänkin mahdollisista kieliopillisista tai rakenteellisista virheistä, joita puheessa tulee ja jotka saattavat vaikeuttaa kuuntelua. (Viswanathan 2005: 63)

Tarkentaen ymmärrettävyydessä kyse on sanojen ja lauseiden erottumisesta, joita mitataan kuuntelun vaatimalla ponnistelulla (kuuntelun helppous), esitetyn informaation vastaanottamisen vaikeudella (informatiivisuus), sanojen oikealla painotuksella ja ääntämisellä sekä kielen virheettömyydellä (oikeakielisyys). (Viswanathan 2005: 63) Tiivistäen ymmärrettävyyden käsite tarkoittaa sitä, miten selkeästi ääni rekisteröityy kuulijalle.

Oikeakielisyys vaikuttaa luonnollisesti puhesynteesijärjestelmään syötettävän tekstin kieliopillinen virheettömyys ja puheen optimaalinen kalibrointi. Olen käyttänyt Amazon Polly ohjesivuston protokollaa audion säätämiseen, ja kieliopillisuus on tarkistettu jo alkuperäisen Oras Groupin huoltovideon käsikirjoitusta tehdessä, joten tällä tapaa mahdolliset virheet tulivat suoraan puhesynteesijärjestelmästä ja sen luomasta puheesta.

Ymmärrettävyyden alakäsitteitä on myös mitattu kahdella kysymyksellä. Jokaisella alakäsitteellä on kaksi omaa kysymystä kyselylomakkeessa ja näiden kysymysten vastausten perusteella lasketaan alakäsitteen keskiarvollinen tulos. Puhesynteesin kuuntelun helppoutta on arvioitu kysymyksillä 9 ja 11, puhesynteesin informatiivisuutta on mitattu kysymyksillä 5 sekä 7 ja oikeakielisyyttä on arvioitu kysymyksillä 2 ja 12 (ks. Liite 1). Ymmärrettävyyden pääkäsitteen tulos lasketaan näiden kolmen alakäsitteen yhteenlasketuista tuloksista.

## 5 Kyselytutkimuksen tulokset

Tutkielman tavoite on mitata ammatti- ja arkikäyttäjien kokemusta puhesynteessin käytöstä teknisellä ohjevideolla ja näiden kokemusten tulosten perusteella miettiä puhesynteessin käyttöönottoa vakituisesti Oras Group Oy:n ohjevideoissa. Kokemusten mitaamiseen käytettiin kyselytutkimusta, jossa kyselyyn vastaajat arvioivat kokemuksiaan puhesynteessistä luonnollisuuden sekä ymmärrettävyyden pääkäsitteiden sekä näitä tarkentavien alakäsitteiden avulla. Kyselyssä vastaajat katsoivat ensin puhesynteesiä sisältävän ohjevideon, jonka jälkeen he vastasivat kysymyksiin ohjevideolla esiintyneestä puhesynteessistä. Kyselyn tuloksista laskettiin pääkäsitteille sekä jokaiselle alakäsitteelle ja muuttujalle laskettiin keskiarvollinen tulos. Muuttujia kyselyssä ovat ohjevideolla käytetty puhesynteessin mies- tai naisääni sekä vastaajien ikä, sukupuoli ja englannin kielen taito. Käsitteiden ja muuttujien tulokset listataan ylös taulukoihin kategorioiden mukaan ja näistä tuloksista suoritetaan lopussa varianssianalyysi.

Ammatti- ja arkikäyttäjille järjestettiin molemmille omat kyselyt, jotka olivat auki 30.9.-8.10.2019 välisenä ajanjaksona. Kyselyihin vastasi yhteensä 146 ihmistä, joista 69 vastasi arkikäyttäjien kyselyyn ja 77 osallistui ammattikäyttäjien kyselyyn. Taulukossa 1 on esitetty molemmat vastaajaryhmät tarkemmin.

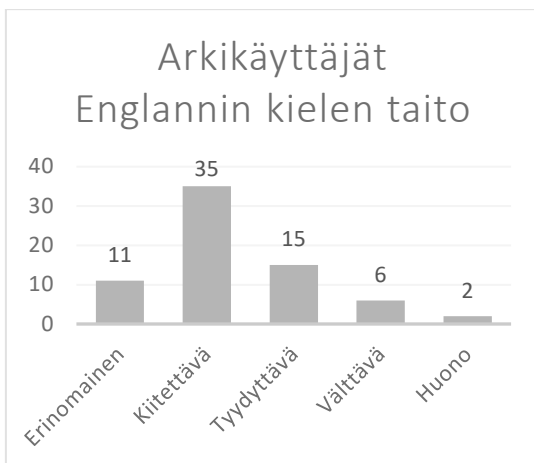
**Taulukko 1.** Kyselyyn vastanneet

Arkikäyttäjä						Ammattikäyttäjä					
Sukupuoli						Sukupuoli					
Miehiä			Naisia			Miehiä			Naisia		
42			27			44			33		
Ikä						Ikä					
Alle 35v		35-55v		Yli 55v		Alle 35v		35-55v		Yli 55v	
M	N	M	N	M	N	M	N	M	N	M	N
23	12	12	10	7	5	10	8	28	22	6	3

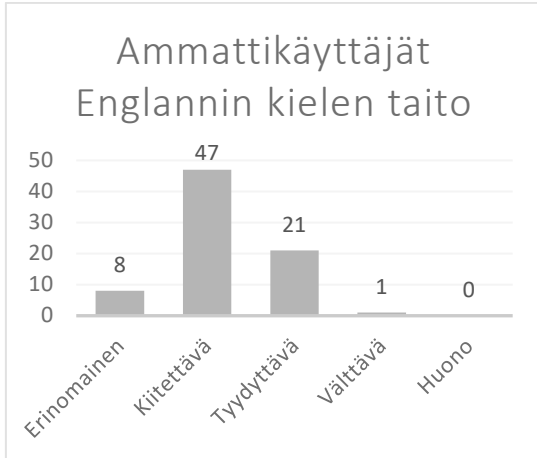
<b>Yhteensä</b>	<b>Yhteensä</b>	<b>Yhteensä</b>	<b>Yhteensä</b>	<b>Yhteensä</b>	<b>Yhteensä</b>
35	22	12	18	50	9
<b>69 osallistujaa</b>			<b>77 osallistujaa</b>		

Kyselyyn saatiin monipuolinen otanta ihmisistä iän puolesta ja se sisälsi myös hyvän tasapainon miesten ja naisten vastauksia. Valtaosa ammattikäyttäjistä sijoittui 35-55-vuotiaisiin, mikä on vallitseva ikäluokka myös Oras Groupin työntekijöissä, joten tästä syystä kyseinen ikäluokka on näin suuri. Arkikäyttäjryhmä oli tasaisemmin jakautunut ja alle 35-vuotiaita oli eniten.

Vastaajia pyydettiin arvioimaan englannin kielen taito ennen kyselyyn vastaamista, ja kuvioissa 7 ja 8 on tarkennettu englannin kielen taidon jakaumaa. Taulukoissa on kuvattu englannin kielen taidot lukumäärien mukaan.



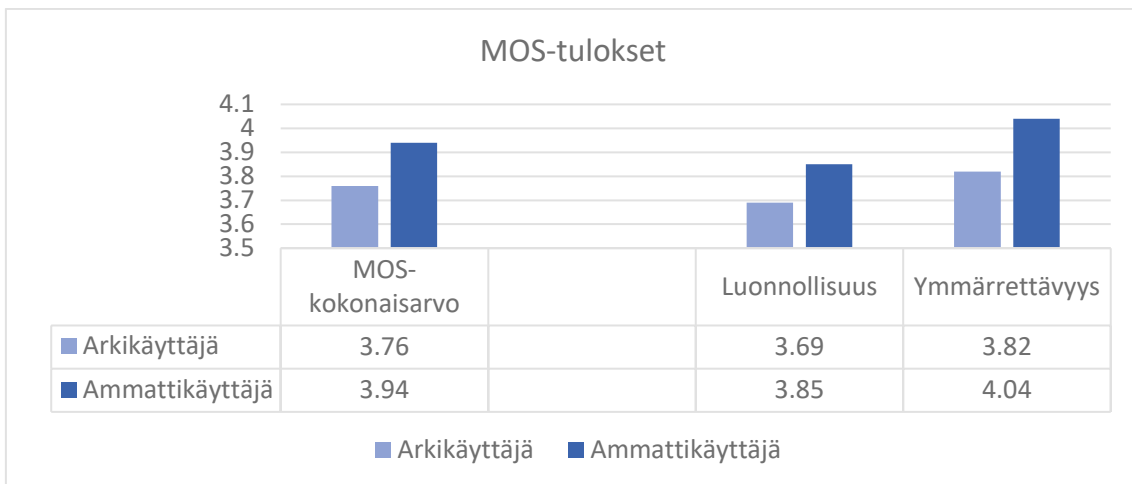
**Kuvio 7.** Arkikäyttäjien englannin kielen taito



**Kuvio 8.** Ammattikäyttäjien englannin kielen taito

Englannin kielen osaaminen on hyvällä tasolla kyselyn vastaajilla, kiitettävän ollessa selvästi valituin kielen taso. Merkittävää on huomata myös, että molempien kyselyjen vastaajista vain 2 luokitteli englannin kielen osaamisensa huonoksi. Jaottelen analyysissäni tästä eteenpäin tyydyttävästä-huonoon valinneet heikompana englannin kielen osaamisen kategoriana ja hyvän englannin kielen osaamisen kategoriaan kuuluvat kaikki kiitettävän ja erinomaisen valinneet vastaajat.

**Taulukko 2.** MOS-tulokset

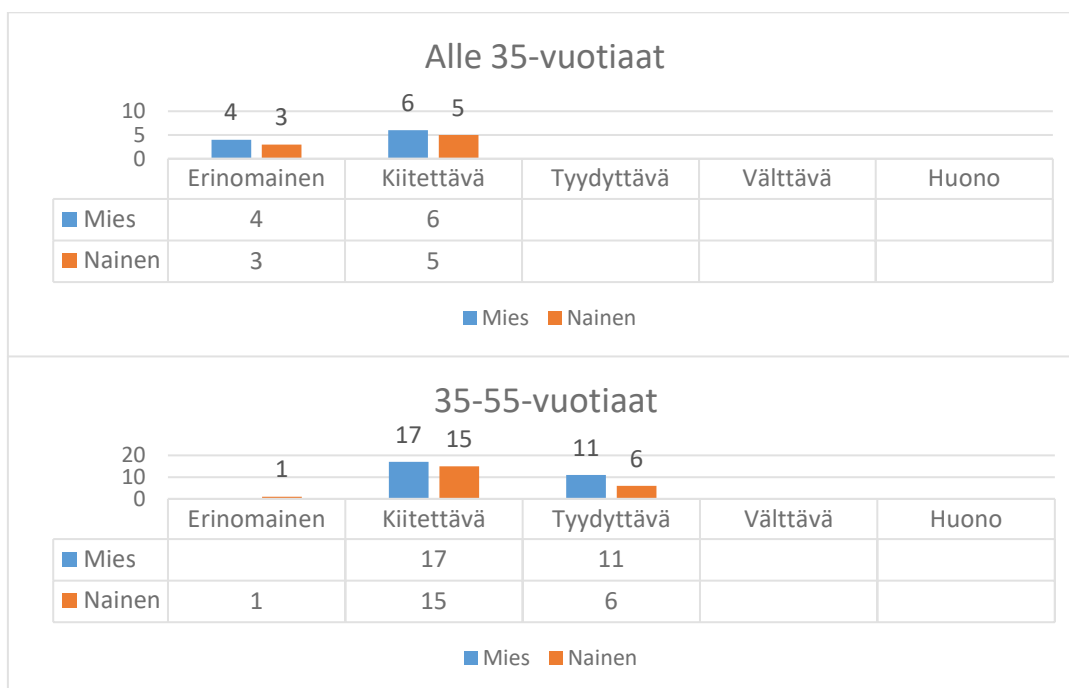


Taulukossa 2 on listattu arkikäyttäjien ja ammattikäyttäjien MOS-kokonaisarvo sekä luonnollisuuden ja ymmärrettävyyden arviointikriteerien tulos. Taulukko antaa yleiskuvan kyselyjen tuloksista, ja siitä käy ilmi alustava ero kahden käyttäjäryhmän välillä. Luvussa 4.1 esitetty Kuvio 5 toimi pohjana MOS-tulosten vertaamiselle. Kuviossa 5 Deep learning-

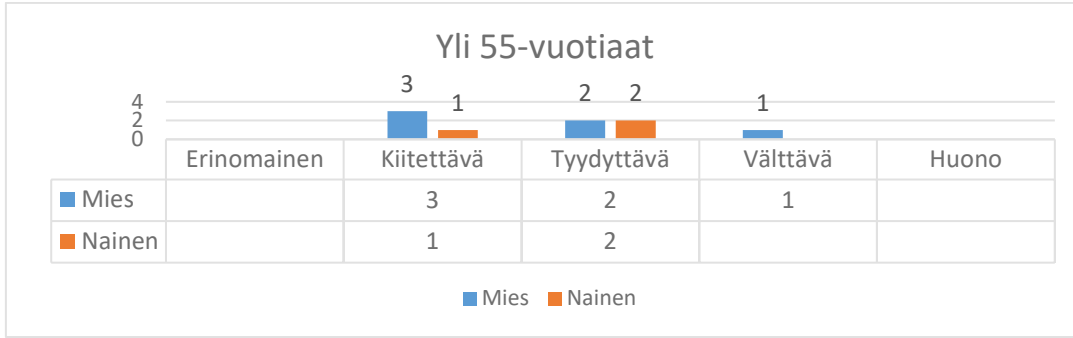
konseptilla toteutettu puhesynteesi sai kyselyssä englannin kielestä MOS-tuloksen 4.21. Perinteisellä parametrisellä järjestelmällä tuotettu puhesynteesi sai kuvion 5. kyselyssä englanninkielestä MOS-tuloksen 3.67. Asetan hyväksyttävän rajan tälle 3.67-4.21 välille MOS-tuloksissa. Tulokset, jotka asettuvat tälle välille luokitellaan hyväksyttävinä tuloksina. Alle 3.67 MOS-tulokset luokitellaan puutteellisiksi tuloksiksi ja yli 4.21 MOS-tulokset ovat erinomaisia tuloksia. Tulokset käydään tarkemmin läpi alaluvuissa käyttäjäryhmien, alakäsitteiden ja muuttujien kautta.

## 5.1 Ammattikäyttäjien MOS-tulokset

Ammattikäyttäjien kyselyyn vastasi yhteensä 77 ihmistä ja vastaajina toimi Oras Groupin henkilöstö, eli alan ammattilaiset. Kuviossa 9 on jaoteltu kaikki ammattikäyttäjryhmän kyselyyn vastanneet henkilöt ikäluokan, sukupuolen ja valitun englannin kielen osaamisen mukaan.

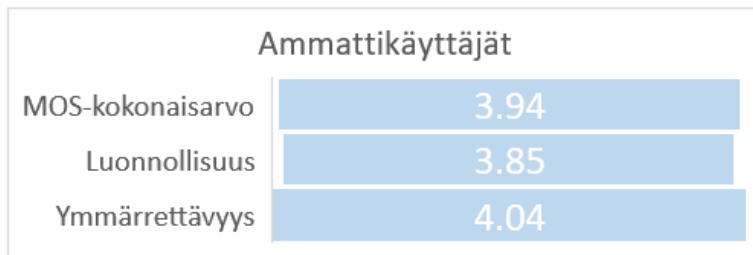






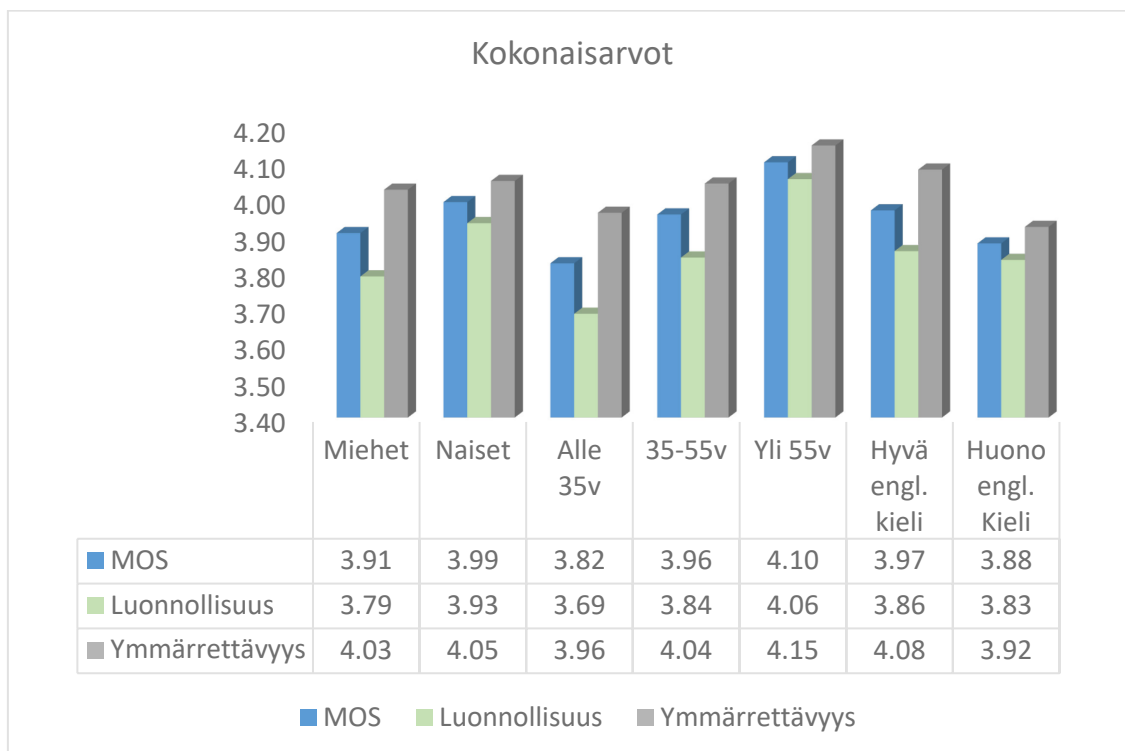
**Kuvio 9.** Ammattikäyttäjien englannin kielen taito ikäluokkien mukaan

Englannin kielen osaaminen jakautui yhteydessä ikäluokan mukaan. Alle 35-vuotiaat arvioivat englannin kielen tasonsa parhaimmaksi, 35-55-vuotiaat arvioivat itsensä toiseksi parhaimpana ja yli 55-vuotiaat heikoimpana. Oras Group on kansainvälinen yritys, joka panostaa työntekijöidensä englannin kielen kehittämiseen, ja tämä vaikutti alhaiseen määrään välttävän tai huonon englannin kielen tason valinnoita vastaajia.



**Kuvio 10.** Ammattikäyttäjien MOS-tulokset

Kuviossa 10 on listattuna ammattikäyttäjien koko kyselyn MOS-tulos sekä luonnollisuuden ja ymmärrettävyyden pääkriteereiden MOS-tulokset. Ammattikäyttäjien tulokset asettuvat aiemmin esitetyn hyväksyttävän MOS-tuloksen (3.67-4.21) välille. Luonnollisuuden ja ymmärrettävyyden kriteereiden tuloksissa on havaittava ero (3.85 ja 4.04) ja seuraavaksi käsitellenkin nämä kaksi pääkriteeriä tarkemmin alakäsitteiden ja muuttujien MOS-tulosten avulla.

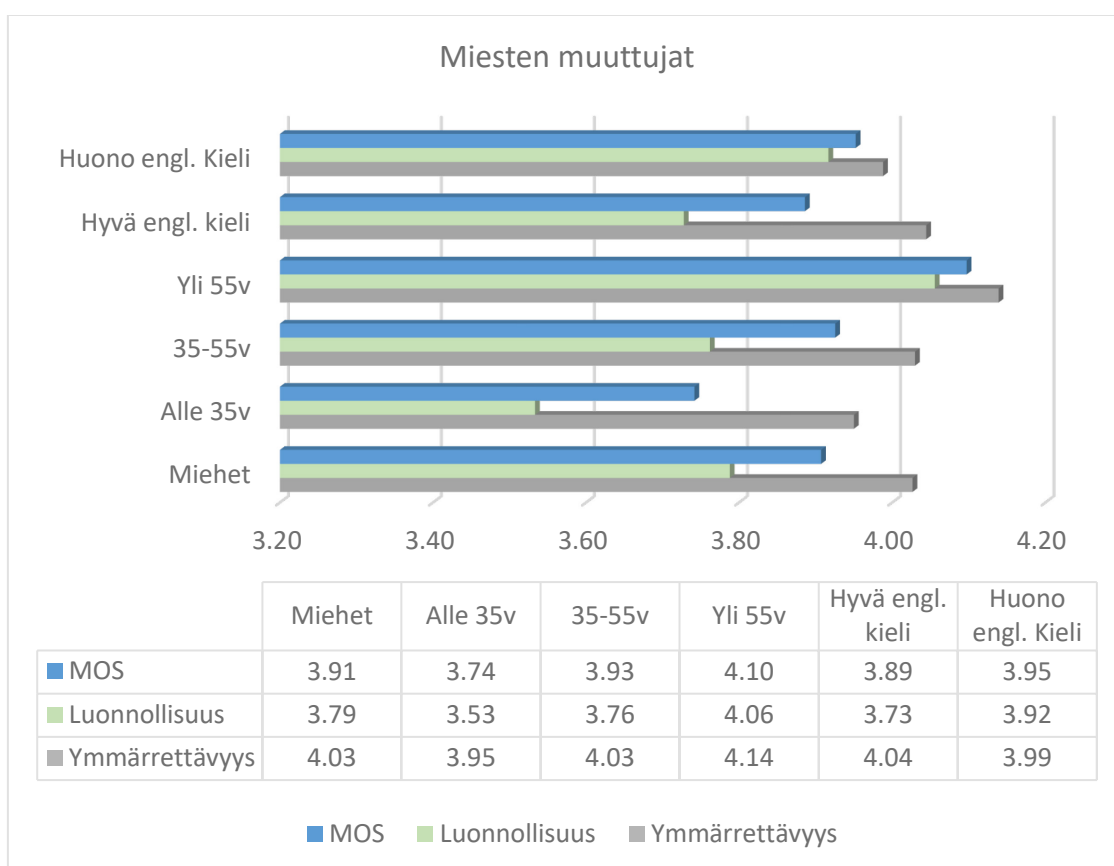
**Taulukko 3.** Ammattikäyttäjien muuttujien MOS-tulokset

Taulukossa 3 on esitetty ammattikäyttäjien kyselyyn osallistuneilta henkilöiltä kysytyt tiedot, eli tutkielmassa käytettävät muuttujat. Kokonaisarvot ovat hyvin tasaisia eri muuttujien välillä, mutta selkein ero on kuitenkin ikäluokissa. Vanhemmat ikäluokat arvostelivat puhesynteesin paremmaksi molemmilla pääkriteereillä sekä myös MOS-kokonaistuloksissa. Huomattavin ero tulee vanhimman ja nuorimman ikäluokan välillä. Vanhin ikäluokka arvosteli puhesynteesin selvästi positiivisemmin kuin nuorin ikäluokka. Nuorimman ikäluokan alin tulos 3.69 luonnollisuudelle ylittää niukasti hyväksytyyn MOS-tuloksen rajan (3.67). Nuorin ikäluokka on myös arvioinut kriittisimmin pääkriteerit verrattuna kahteen muuhun ikäluokkaan. Kaikki käsitteet muuttujissa pysyivät kuitenkin hyväksytyyn MOS-tuloksen rajoissa ja tulokset eivät alita puutteellisen tuloksen rajaa, toisaalta ne eivät myöskään ylitä erinomaisen tuloksen rajaa (4.21).

Englannin kielen taso näyttää vaikuttavan jonkin verran puhesynteesin vastaanottoon. Hyvän englannin kielen tason omaavat henkilöt arvioivat puhesynteesin hieman paremmaksi kuin huonon englannin kielen tason omaavat. Ero ei ole kuitenkaan huomattava, ja näin ollen ammattikäyttäjien osalta englannin kielen tasolla näyttäisi olevan vain pieni

vaikutus puhesynteesin vastaanottoon kokonaistulosten kannalta. Miesten ja naisten väliset tulokset ovat myös samantasoisia, suurin ero myös sukupuolten välillä on luonnollisuudessa. Naiset arvioivat puhesynteesin luonnollisemmaksi kuin miehet ja naiset arvioivat myös kokonaisuutena puhesynteesin hieman paremmaksi kuin miehet. Alla olevissa alaluvuissa käydään läpi tarkemmin miesten ja naisten vastaukset ikäluokkien ja englannin kielen tason mukaan.

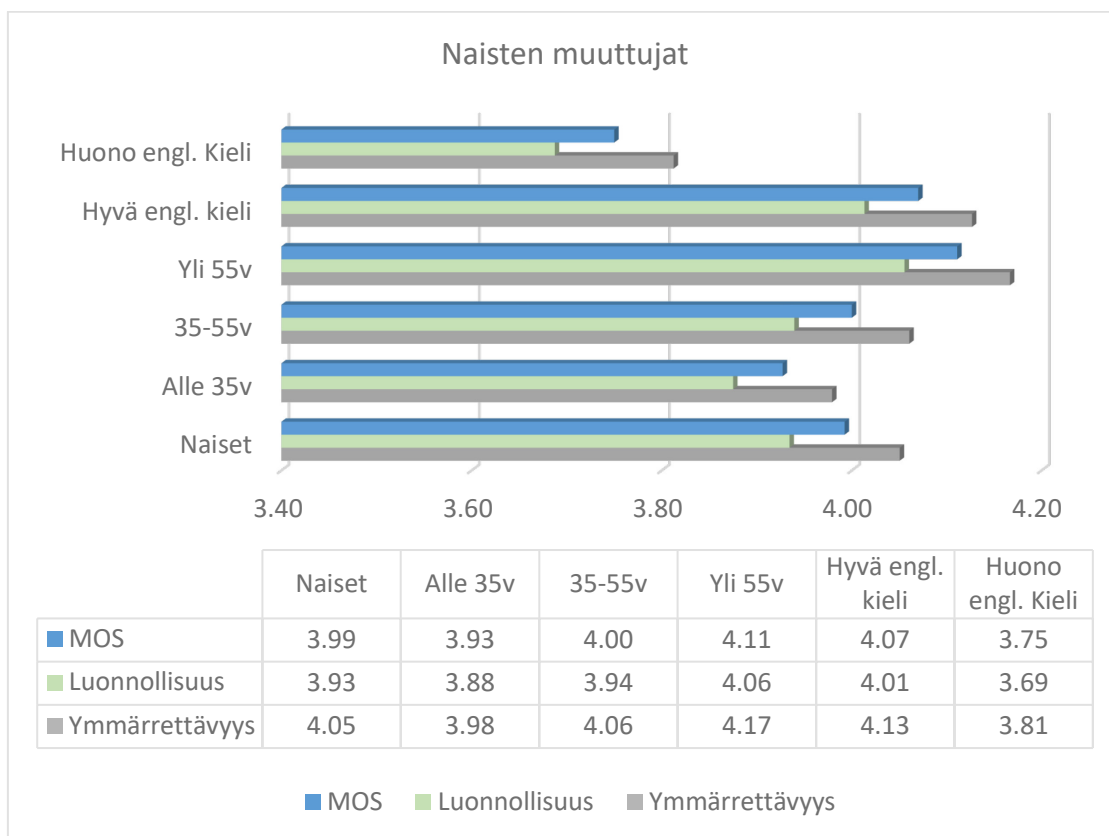
**Taulukko 4.** Ammattikäyttäjät: Muuttujat miesvastaajien osalta



Taulukossa 4 on jaettu miesten vastaukset ikäluokan ja englannin kielen tason mukaan. Miesten muuttujat noudattavat samaa kaavaa kuin aiemmin listatut tulokset. Suurimmat erot tulevat ikäluokkien välillä, mutta miesten tuloksissa nämä erot ovat huomattavasti suuremmat. Alle 35-vuotiaat miehet arvostelivat puhesynteesin puutteelliseksi luonnollisuuden kategoriassa, ja ero yli 55-vuotiaiden tulokseen luonnollisuudessa on huomattava. Englannin kielen tasossa on myös ero hyvän englannin kielen omaavien tuloksessa

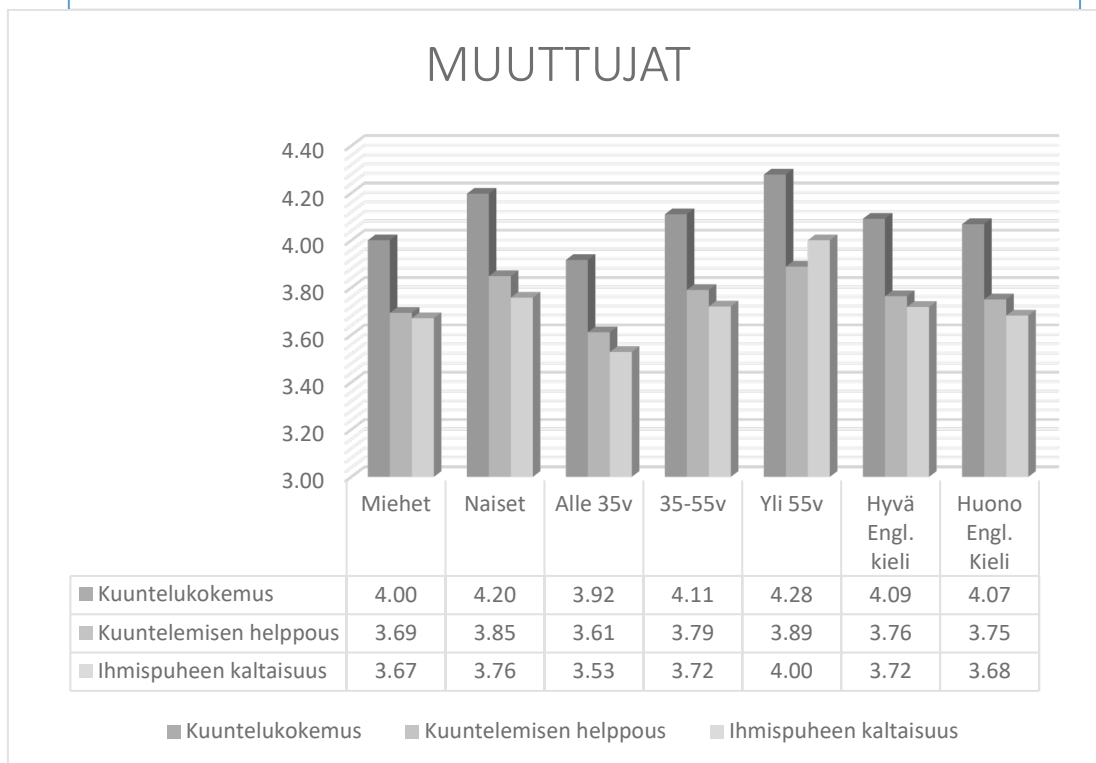
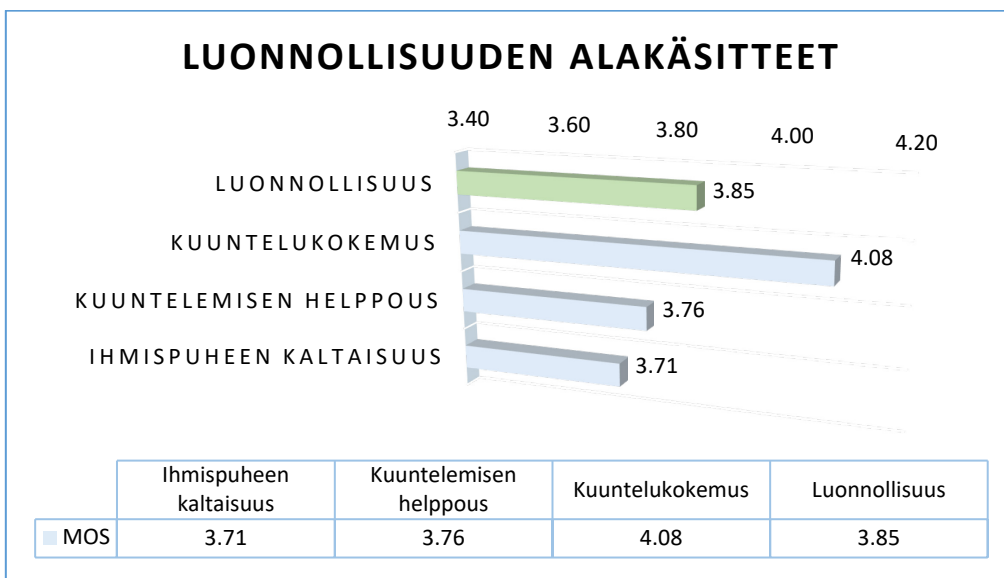
luonnollisuudesta verrattuna huonon englannin kielen taidon omaaviin. Hyvän englannin kielen taidon omaavat arvostelivat luonnollisuuden heikompana kuin huonon kieli- taidon omaavat henkilöt. Kaikki miesten tulokset yhtä lukuun ottamatta sijoittuvat hyväksytyyn MOS-tuloksen välille.

**Taulukko 5.** Ammattikäyttäjät: Muuttajat naisvastaajien osalta



Naisten muuttajien tulokset ovat kuvattu taulukossa 5 ja niissä merkittävin ero on englannin kielen tason suhteen. Paremmalla englannin kielen tason omaavat arvostelivat kaikilla kategorioilla puhesynteesin paremmaksi kuin heikomman tason omaavat naisten tuloksissa. Miesten tuloksissa hyvän englannin kielen omaavat vastaajat arvostelivat pelkästään ymmärrettävyyden paremmin kuin huonon englannin kielen taidon omaavat vastaajat. Ikäluokissa on samaa eroavaisuutta tuloksissa kuin miehissä vanhimmasta nuorimpaan ikäluokkaan, pienemmällä marginaalilla.

Taulukko 6. Ammattikäyttäjien luonnollisuuden alakäsitteet

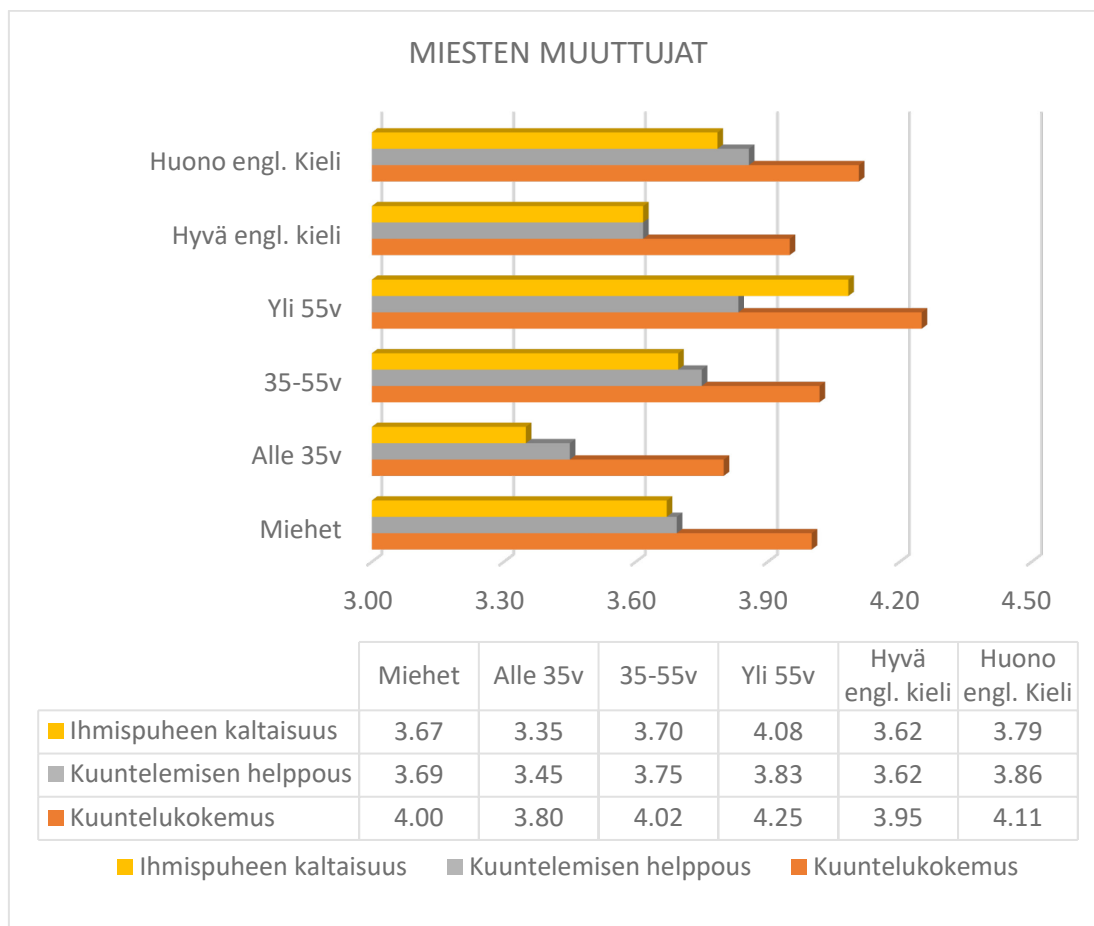


Taulukossa 6 on ensin näytetty luonnollisuuden ja sen alakäsitteiden MOS-kokonaisarvo, jonka jälkeen alakäsitteille on laskettu muuttuja kohtaisesti MOS-tulos. Alakäsitteet sijoittuvat kaikki hyväksytyyn MOS-arvon välille. Ihmispuheen kaltaisuus on kuitenkin lähellä alarajaa ja näin ollen myös heikoimmaksi arvosteltu piirre. Kuuntelemisen helppous on myös saanut hiukan paremman tuloksen 3.76, vaikka sekin sijoittuu lähelle

puutteellisen arvostelun rajaa. Kuuntelukokemus oli selvästi parhaimmaksi arvioitu alakäsite, ja tämän perusteella vaikuttaisi, että ihmispuheen kaltaisuuden heikohko arviointi ei vaikuttanut kuitenkaan kuuntelukokemukseen kokonaisuutena.

Muuttujista kriittisimmin alakäsitteet ovat arvioineet alle 35-vuotiaat, mikä noudattaa aiempien tulosten linjaa. Kuuntelemisen helppous sekä ihmispuheen kaltaisuus on molemmat arvioitu puutteellisiksi (3.61, 3.53) alle 35-vuotiaiden ryhmässä. Vanhemmat ikäluokat ovat arvioineet kauttaaltaan alakäsitteet parempina. Englanninkielen taidon tasolla ei tulosten valossa näytä olleen vaikutusta arviointiin. Naiset ovat arvioineet myös käsitteet hieman miehiä paremmaksi.

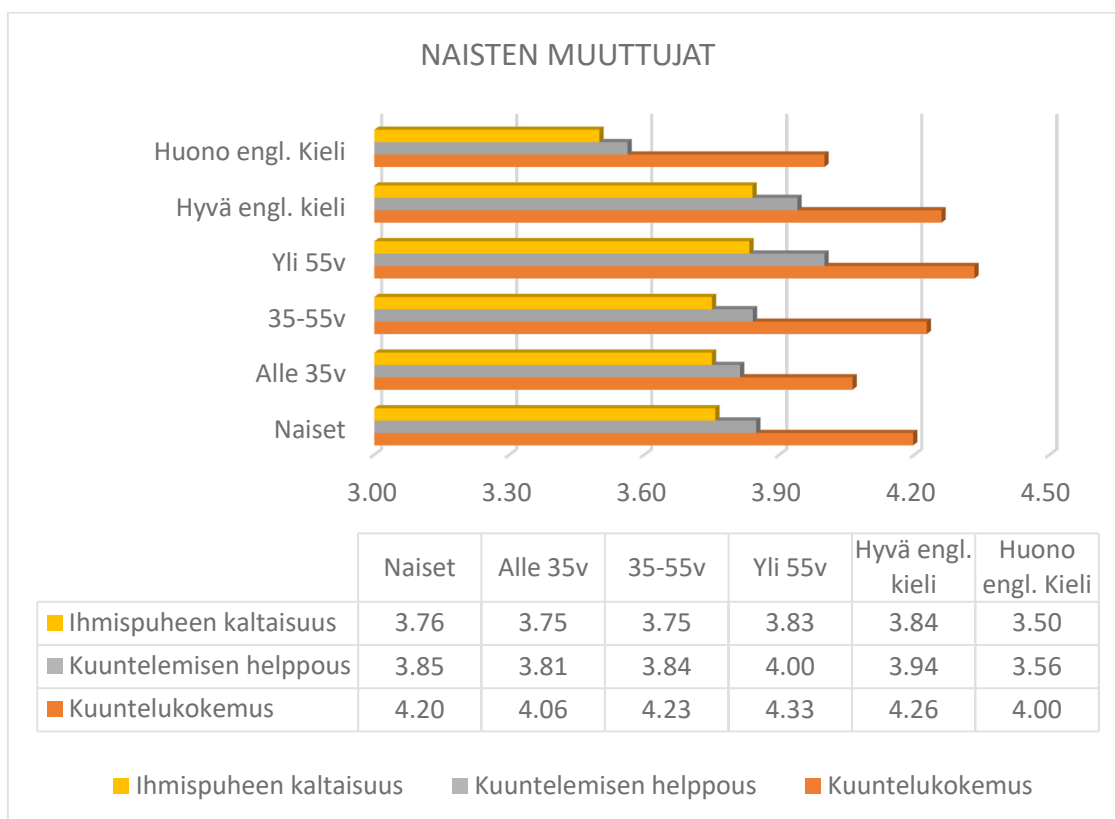
**Taulukko 7.** Ammattikäyttäjien luonnollisuuden alakäsitteet: Miesvastaajat



Taulukossa 7 on tarkennettu miesten muuttujia ja taulukko korostaa aiempia havaintoja siitä, että alle 35-vuotiaat ovat kriittisempiä arvioissaan. Nuorin ikäluokka on arvioinut

ihmispuheen kaltaisuuden sekä kuuntelemisen helppouden selvästi puutteellisen rajan puolelle. Näiden lisäksi myös hyvän englanninkielen taidon omaavat henkilöt ovat arvioineet ihmispuheen kaltaisuuden sekä kuuntelemisen helppouden puutteellisen arvioinnin puolelle. Vanhemman ikäluokan vastaajat ovat kuitenkin arvioineet kyseiset kaksi alakäsitettä selvästi paremmaksi myös pelkästään miesten arvioinneissa. Kokonaisarviona kaikki kolme alakäsitettä ovat kuitenkin hyväksytyyn MOS-tuloksen asteikolla, kuuntelukokemuksen ollessa parhaimmaksi arvosteltu alakäsite myös miesvastaajien toimesta.

**Taulukko 8.** Ammattikäyttäjien luonnollisuuden alakäsitteet: Naisvastaajat

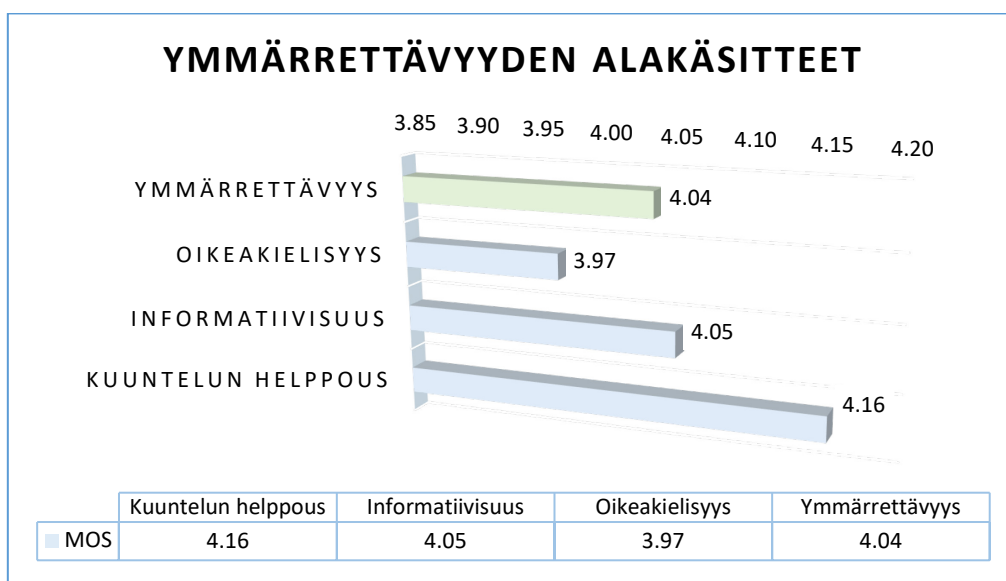


Naispuolisten vastaajien taustamuuttujat on eritelty taulukossa 8. Suurin ero on huonon englannin kielen taidon omaavien vastaajien arvioinnissa ihmispuheen kaltaisuudesta sekä kuuntelemisen helppoudesta. Molemmat käsitteet luokitellaan puutteelliseksi kuuntelemisen helppouden ollessa 3.56 ja ihmispuheen kaltaisuuden 3.50. Englannin kielen taso vaikuttaa kauttaaltaan myös eniten arviointiin. Hyvän englannin kielen omaavat henkilöt ovat arvioineet kaikki kolme alakäsitettä selvästi heikomman englannin

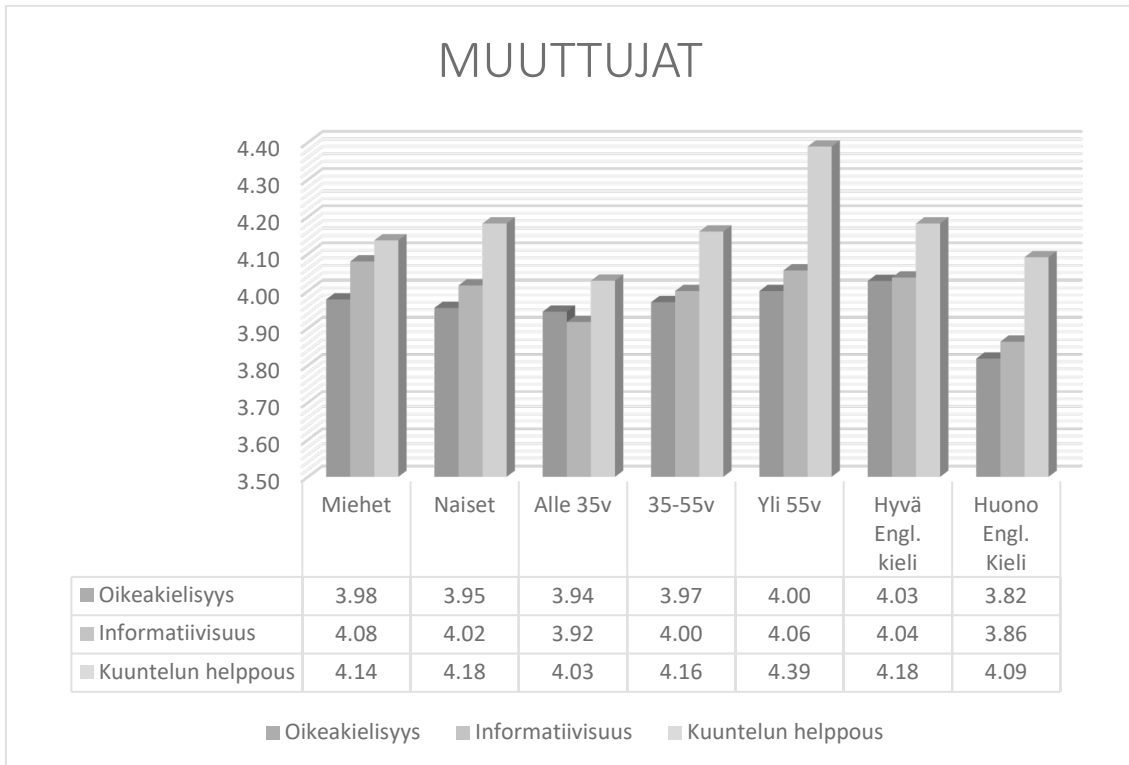
kielen taidon omaavia paremmaksi. Naisten tuloksissa on vanhimman ja nuorimman ikäluokan välillä eroavaisuutta, ja tämä noudattaa aiempaa linjaa siitä, että nuorin ikäluokka arvioi puhesynteesiä kriittisemmin, kun taas vanhin ikäluokka antaa hyvän arvioinnin. 35-55-vuotiaat- ja yli 55-vuotiaat ikäluokan naiset arvioivat myös ensimmäiset erinomaiset tulokset kuuntelukokemukselle 4.23 ja 4.33. Hyvän englanninkielen taidon omaavat naisvastaajat arvioivat myös kuuntelukokemuksen erinomaisena (4.26).

Luonnollisuus ja sen alakäsite olivat pääpiirteiltään arvosteltu hyväksyttävänä tuloksina, mutta tarkemmin arvioituina ne sisälsivät myös puutteellisen, mutta myös erinomaisen tuloksen saaneita arviointeja. Seuraavassa luvussa käyn läpi tarkemmin toisen pääkriteerin ymmärrettävyys ja sen alakäsitteet.

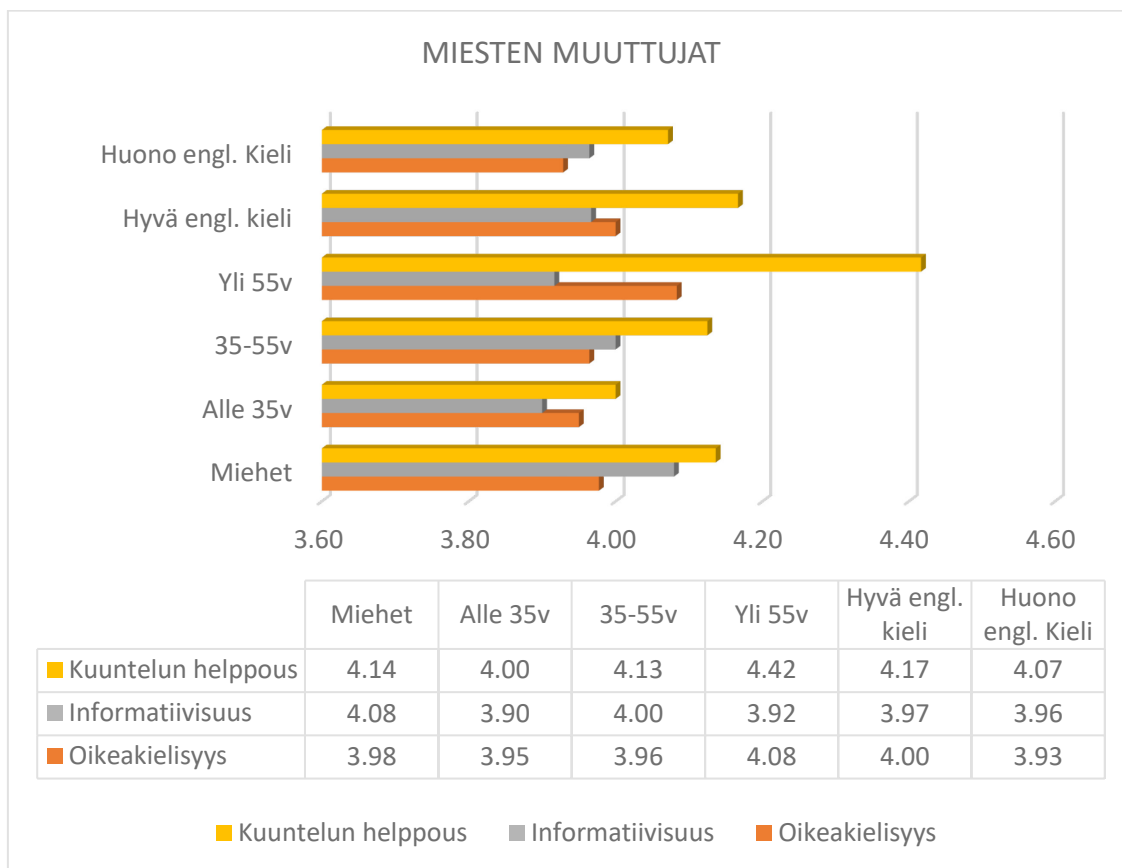
**Taulukko 9.** Ammattikäyttäjien ymmärrettävyyden alakäsitteet



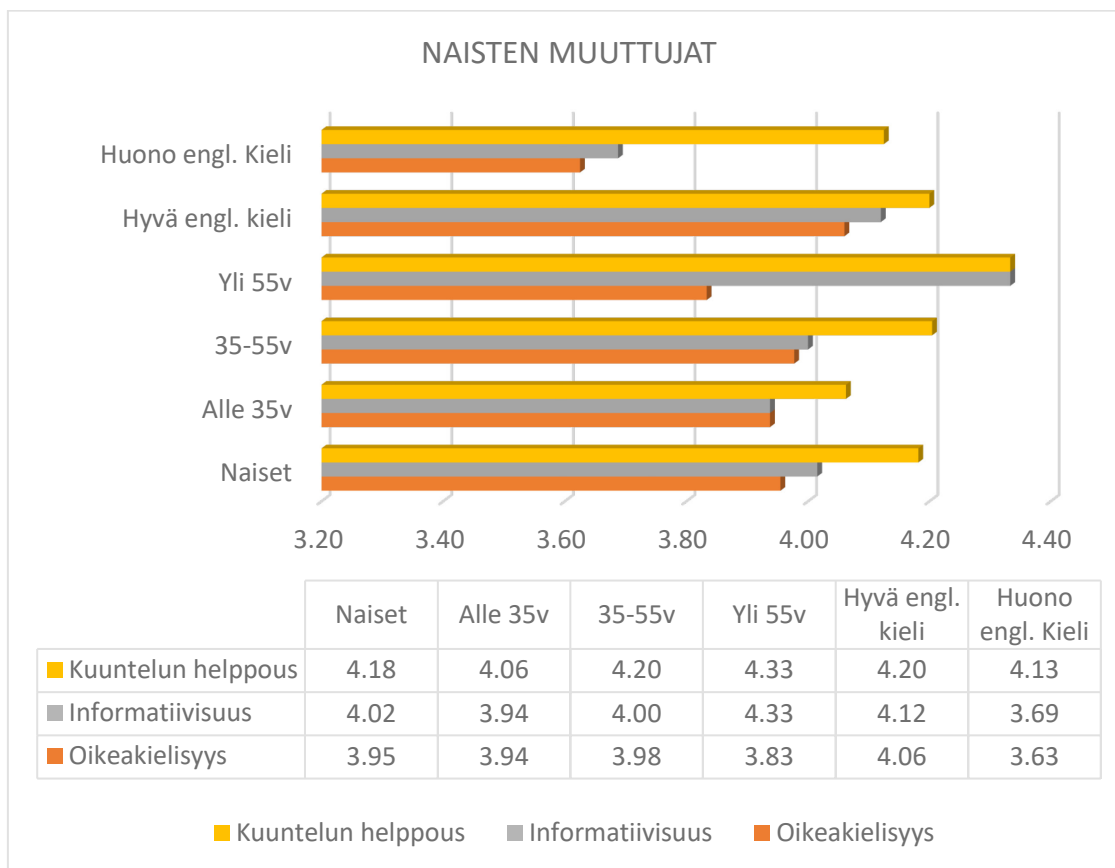




Ymmärrettävyyden alakäsitteet on esitetty taulukossa 9. MOS-kokonaisarviot ovat selvästi hyväksyttävän arvioinnin puolella ja kuuntelun helppous on saanut parhaimman arvioinnin 4.16. Muuttujien arvioinnit ovat myös selvästi hyväksyttävien sekä erinomaisen välisellä tasolla. Erinomaisen arvioinnin on saanut kuuntelun helppous yli 55-vuotiaiden vastaajien arvioinnissa, kaikki muut arvot ovat selvästi hyväksyttävän arvioinnin sisällä. Arvioinnit ovat myös tasapainoiset verrattuna aiempiin taulukoihin, suurin ero on englannin kielen taidon omaavien välillä. Huonon taidon omaavat vastaajat ovat arvoitelleet kaikki kolme alakäsitettä huonommin kuin hyvän taidon omaavat vastaajat. Iso yksittäinen ero on myös erinomaisen 4.39 arvioinnin yli 55-vuotiailta saaneen kuuntelun helppouden käsitteessä, verrattuna nuorimman alle 35-vuotiaiden ikäluokan arviointiin 4.03. Ikäluokkien välillä on näin ollen havaittavissa sama jako kuin luonnollisuuden ja sen alakäsitteiden arvioinnissa.

**Taulukko 10.** Ammattikäyttäjien ymmärrettävyyden alakäsitteet: Miesvastaajat

Miesvastaajien taustamuuttajat ymmärrettävyyden alakäsitteet on kuvattu yllä olevassa taulukossa 10. Tulokset tässä ryhmässä ovat varsin tasapainoiset ja isoja eroja ei juurikaan ole. Yksi selkeä ero ja poikkeama kuitenkin löytyy: yli 55-vuotiaiden arviointi kuuntelun helppoudesta, joka on saanut erinomaisen 4.42 arvioinnin. Tämä poikkeaa itsessään suuren MOS-tuloksensa vuoksi, mutta se erottuu myös muiden ikäluokkien kuuntelun helppouden arvioinnista (4.00 ja 4.13). Tämä kuitenkin noudattaa havaittua puunaista lankaa siitä, että vanhin ikäluokka on arvioinut puhesynteessin varsin positiivisella tavalla. Miesten ymmärrettävyyden sekä sen alakäsitteiden arvioinnit ovat kauttaaltaan myös positiivisia, ja ne eivät sisällä ainoatakaan puutteellisen arvioinnin saanutta tulosta.

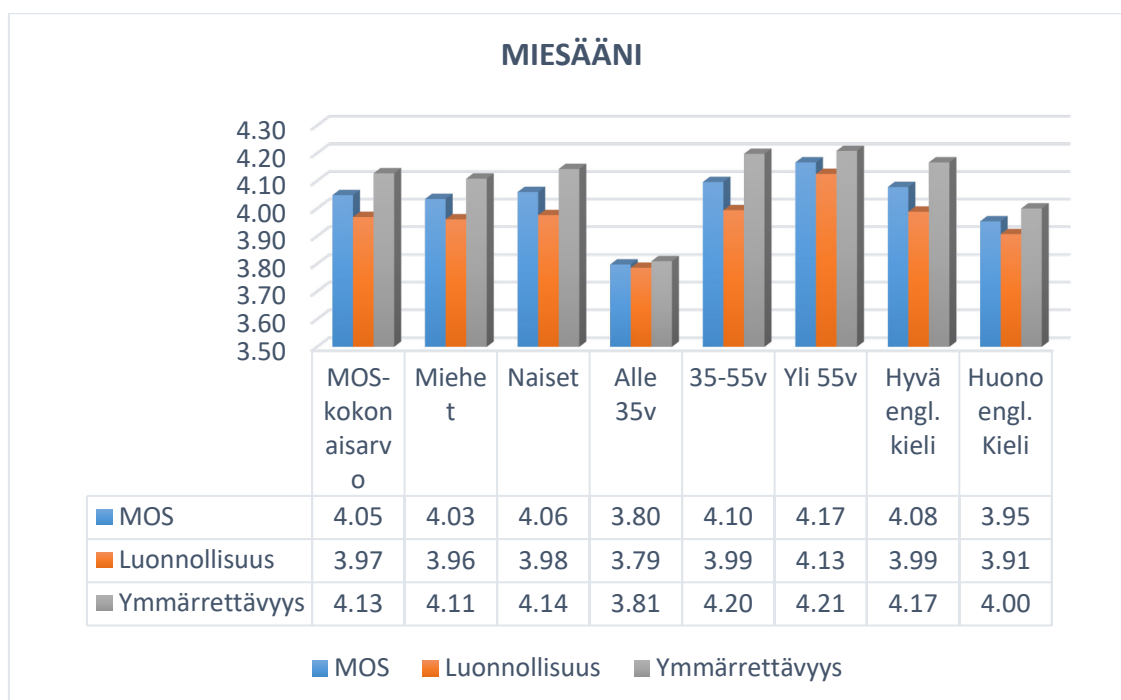
**Taulukko 11.** Ammattikäyttäjien ymmärrettävyyden alakäsitteet: Naisvastaajat

Naisten vastaavat tulokset on listattu taulukossa 11. Tulokset mukailevat miesten tuloksia siinä, että ovat korkeita, mutta muutama poikkeus löytyy myös. Englannin kielen taidon tuloksissa on selvä ero informatiivisuuden ja oikeakielisyyden piirteissä. Huonon englannin kielen taidon omaavat vastaajat ovat arvioineet oikeakielisyyden puutteelliseksi (3.63), ja informatiivisuus on myös saanut alhaisen arvion (3.69). Hyvän englannin kielen taidon omaavat ovat arvioineet informatiivisuuden 4.12 ja oikeakielisyyden 4.06. Ero on näin ollen huomattava. Kuuntelun helppous on kuitenkin saanut molemmilta englannin kielen taidon ryhmiltä arviokseen korkeat 4.20 sekä 4.13 arviot. Tämän lisäksi myös ikäluokkien välillä on sama jatkumo kaikissa alakäsitteissä, arvioiden kulkiessa ikäluokkien mukaan nuorimmasta vanhimpaan, eli alhaisesta korkeampaan arvioon. Yli 55-vuotiaiden ryhmä on arvioinut erinomaisena kuuntelun helppouden (4.33) sekä informatiivisuuden (4.33), mikä kuvastaa hyvin vanhempien ikäluokkien erittäin positiivisia arvioita.

### 5.1.1 Miesääni ammattikäyttäjien arvioimana

Puhesynteesi esitettiin kahdella eri videolla kyselyn vastaajille. Yhdessä videossa oli naisääni ja toisessa miesääni, puhesynteesin äänenä. Käyn läpi tässä luvussa miesäänellä tehdyn videon katsoneiden tulokset. Tutkielmassa tästä eteenpäin tulevista miesäänen ja naisäänen tuloksista on karsittu pois aiemmin esitetyt sukupuolen sisäiset muuttujat, tilastollisesti vastaajamäärä putoaisi nolnaan tai muutama osallistujaan tietyissä muuttujaryhmissä, joten näistä ei näin ollen saataisi tarpeellista tutkittavaa tietoa. Seuraavaksi esittelen puhesynteesin miesäänellä nähneen videon vastaajien tulokset pääkriteereiden kokonaisarvoiltaan sekä pääkriteereiden alakäsitteiden tulosten avulla.

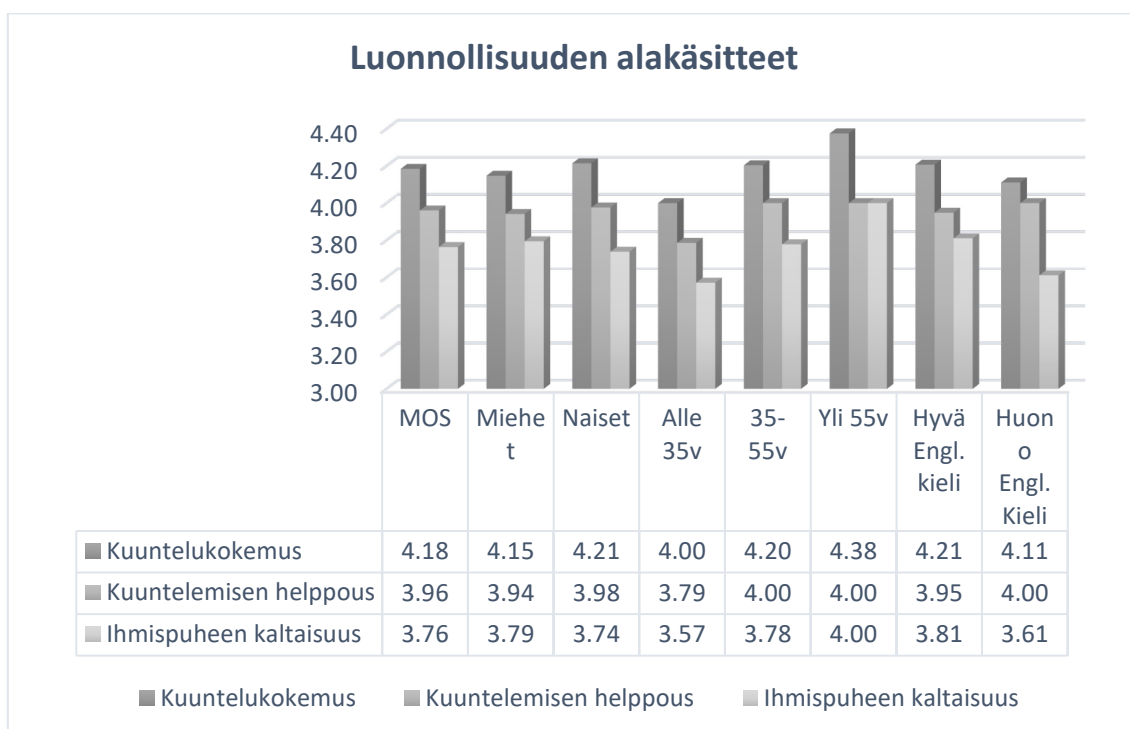
**Taulukko 12.** Puhesynteesi miesäänellä: ammattikäyttäjien kokonaisarvot



Taulukossa 12 on esitetty MOS-kokonaisarvo sekä luonnollisuuden ja ymmärrettävyyden tulokset. Kuten taulukosta käy ilmi kokonaisarvot ovat hyvät kaikista pääkriteereistä. Ymmärrettävyyden käsite on arvioitu parhaiten 4.13 tuloksella ja luonnollisuuden käsite on alhaisin 3.97 tuloksella. Tulokset ovat näin ollen selvästi positiiviset kokonaisarvioiden kannalta puhesynteesin miesäänen arvioinnissa.

Naisten ja miesten arviot ovat hyvin samantasoisia kaikissa kolmessa kriteerissä, selvimmät muutokset ovatkin ikäluokan ja englannin kielen taidon tason kategorioissa. Vanhin ikäluokka on arvioinut puhesynteesin miesäänisen videon parhaiten joka kriteerillä, kun taas nuorin ikäluokka on antanut alhaisimman arvion jokaisessa kolmessa kriteerissä. 35-55-vuotiaiden ryhmä asettuu näiden kahden ryhmän väliin arvioinnissaan. Iällä näyttää näin ollen olevan selvä vaikutus puhesynteesin vastaanottamiseen ja näyttäisi, että vanhemmat ikäluokat mieluummin vastaanottavat puhesynteesin miesäänellä. Englanninkielen taidon tasolla on myös vaikutusta puhesynteesin vastaanotossa. Ero on pienempi kuin ikäluokissa, mutta tulosten valossa hyvän englanninkielen taidon omaavat arvioivat paremmin puhesynteesin.

**Taulukko 13.** Miesääni: ammattikäyttäjien luonnollisuuden alakäsitteet



Luonnollisuuden alakäsitteet on esitelty taulukossa 13. Kuuntelukokemus on arvioitu korkeimmalla tuloksella 4.18, mutta alhaisimpaan ihmispuheen kaltaisuuden tulokseen 3.76 on merkittävä pudotus kuuntelukokemuksen tuloksesta. Ihmispuheen kaltaisuus on myös arvioitu kahteen kertaan puutteellisena. Alle 35-vuotiaat arvioivat ihmispuheen kaltaisuuden 3.57 ja huonon englannin kielen taidon omaavat 3.61 eli puutteelliseksi

merkittävänä tuloksina. Ikäluokkien välillä on sama eroavaisuus kuin taulukossa 16, eli nuoret arvioivat puhesynteesiä kriittisemmin kuin vanhemmat ikäluokat. Yli 55-vuotiaat arvioivat kuuntelukokemuksen selvästi erinomaisen rajan ylittävällä tuloksella 4.38, kun nuorin ikäluokka antoi samalle alakäsitteelle arvion 4.00.

**Taulukko 14.** Miesääni: ammattikäyttäjien ymmärrettävyyden alakäsitteet



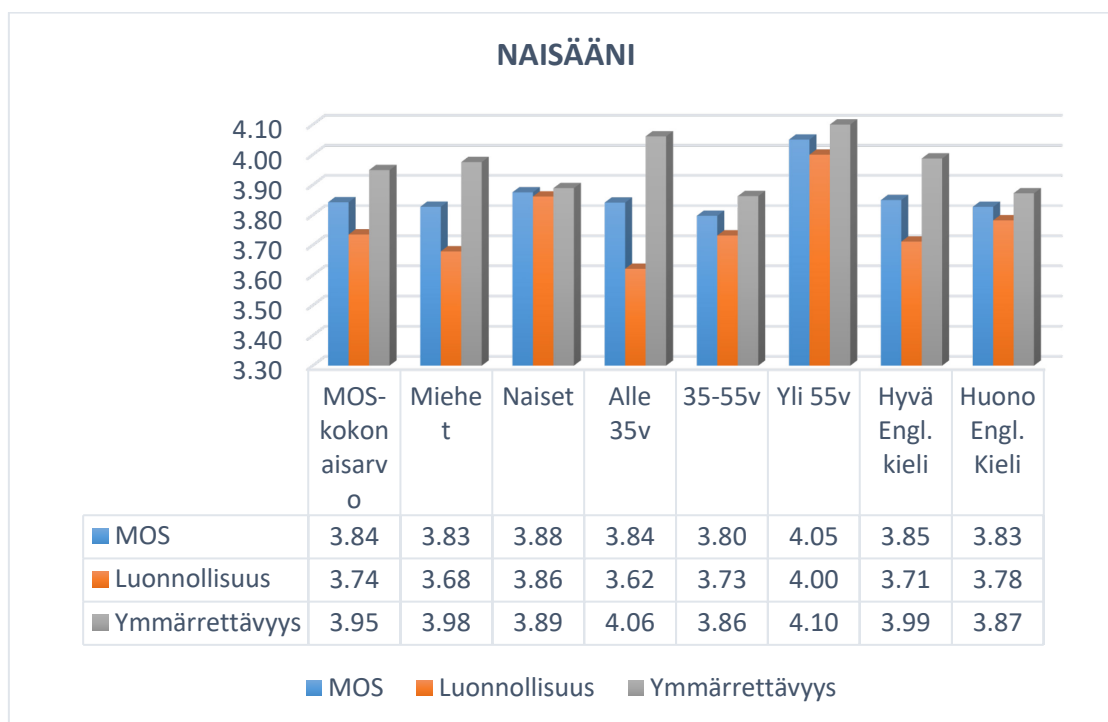
Ymmärrettävyyden alakäsitteet on käyty läpi taulukossa 14. MOS-kokonaisarviot ovat kaikissa kolmessa käsitteessä hyvällä tasolla. Oikeakielisyys on alhaisin 3.99 tuloksella, mutta korkein tulos kuuntelun helppoudessa ylittää erinomaisen tuloksen rajan (4.28). Erinomaisen raja ylitetään myös kuuntelun helppoudessa naisten arvioimana 4.29, 35-55-vuotiaiden arvioimana 4.35, yli 55-vuotiaiden tuloksessa 4.50 sekä hyvän, että huonon englannin kielen omaavien vastaajien 4.29 ja 4.22. Erinomaisen raja ylitetään myös yli 55-vuotiaiden arviossa informatiivisuudesta 4.25 tuloksella. Informatiivisuutta ja kuuntelun helppoutta pidetään hyvänä. Oikeakielisyys jää jonkin verran jälkeen tuloksissa, mutta myös sen tulokset ovat positiiviset. Alhaisimmat tulokset ovat alle 35-vuotiaiden arvioissa kaikista kolmesta alakäsitteestä sekä huonon englannin kielen omaavien vastaajien arviossa oikeakielisydestä. Ymmärrettävyyden alakäsitteiden tulokset olivat

kuitenkin varsin korkealla tasolla kokonaisuutena, niin kuin oli valtaosa miesäänen puhesynteesin katsoneiden vastaajien tuloksista. Varsinkin kokonaisarviot ylsivät korkeisiin lukemiin ja näin ollen puhesynteesin miesäänen vastaanotto oli positiivinen.

### 5.1.2 Naisääni ammattikäyttäjien arvioimana

Tässä luvussa käydään läpi tarkemmin puhesynteesin naisäänellä tehdyn videon arvioinnin tulokset.

**Taulukko 15.** Puhesynteesi naisäänellä: ammattikäyttäjien kokonaisarviot

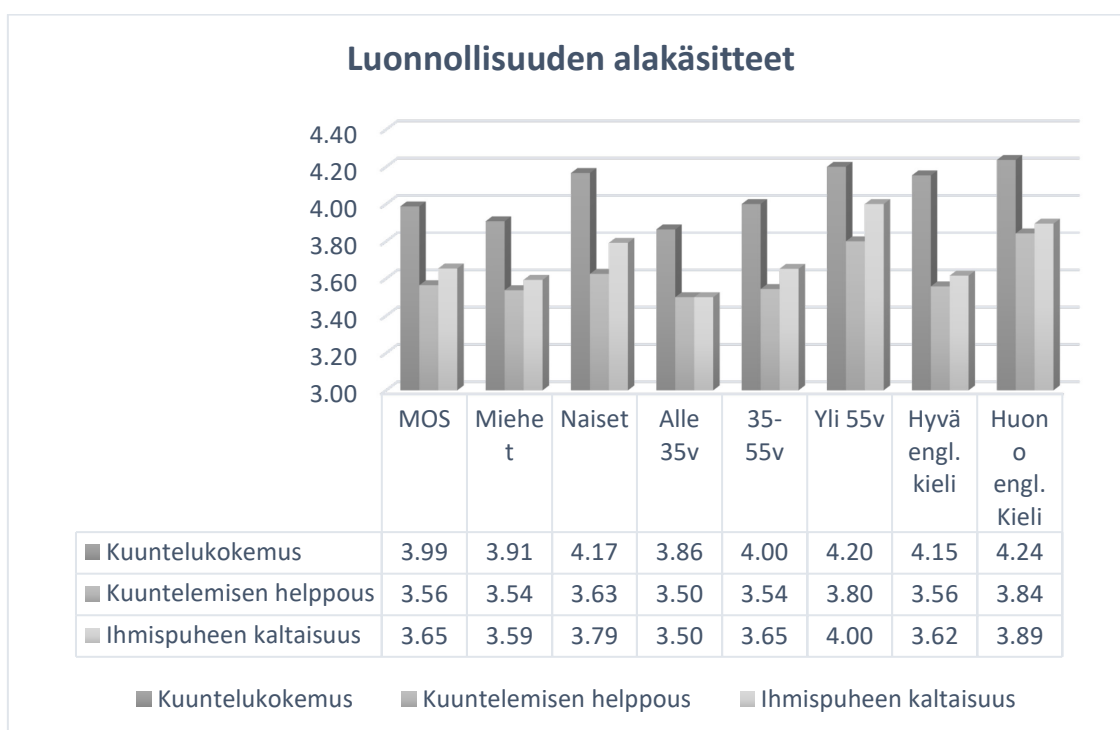


Kokonaisarviot ovat kaikki hyväksyttävän tuloksen rajoissa, ymmärrettävyyden ollessa jälleen parhaimmaksi arvioitu kriteeri. Kokonaisarviot jäivät kuitenkin alhaisemmaksi kuin miesäänen tulokset ja muuttujista löytyy myös enemmän eroavaisuuksia kuin miesäänen muuttujien tuloksissa.

Sukupuolten vertailussa nähdään, että miehet mieltävät naisäänen ymmärrettävämpänä, kun taas naisvastaajat arvioivat naisäänen enemmän luonnollisena kuin miesvastaajat.

Ikäluokissa vanhimmat ikäluokat arvioivat puhesynteesin paremmaksi kuin nuoremmat ikäluokat. Poikkeuksena tähän ymmärrettävyyden arviointi, jonka alle 35-vuotiaat ovat arvioineet lähes yhtä korkeana kuin vanhin yli 55-vuotiaiden ryhmä. Alle 35-vuotiaat ovat arvioineet luonnollisuuden kuitenkin vain tuloksella 3.62, joka on alle puutteellisen tuloksen rajan. Englannin kielen taidon tasossa on ymmärrettävyyden kannalta selvin ero. Paremman taidon omaavat arvioivat puhesynteesin naisäänen ymmärrettävämmäksi. Puhesynteesin naisäänen tulokset ovat kokonaisuutena hyväksyttävällä tasolla, mutta tulokset ovat kuitenkin hyväksyttävän tuloksen alarajalla.

**Taulukko 16.** Naisääni: ammattikäyttäjien luonnollisuuden alakäsitteet



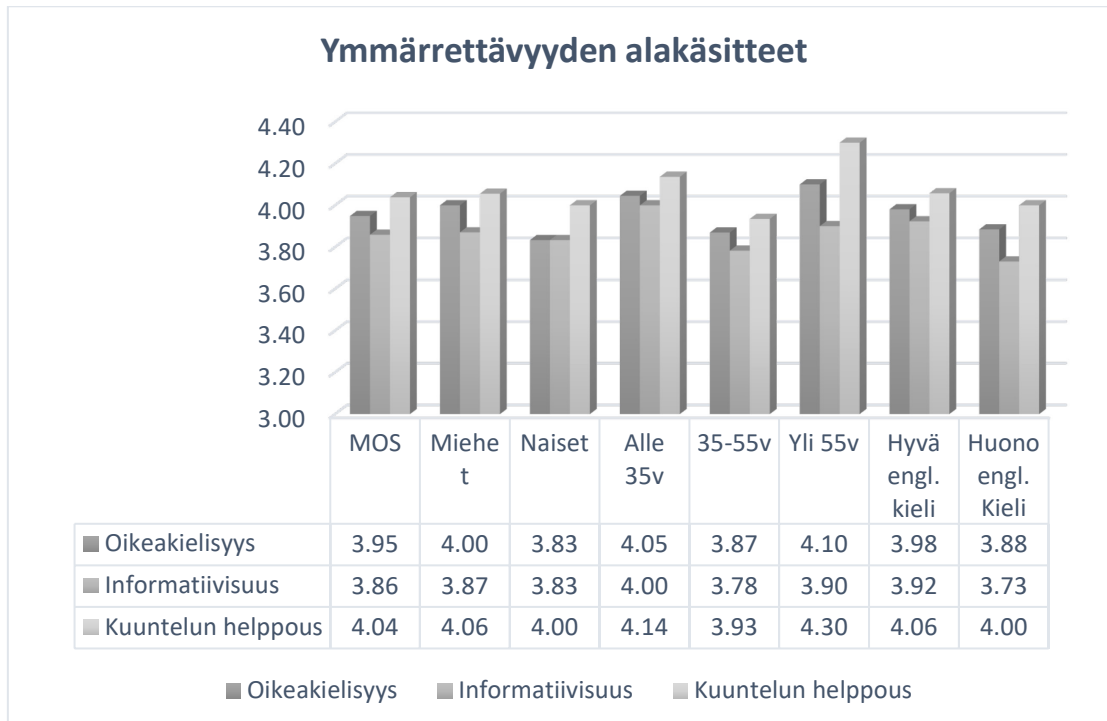
Luonnollisuuden alakäsitteiden tulokset jäivät selvästi alhaisiksi, kuten taulukosta 16 käy ilmi. Puutteellisen tuloksen saaneita arvioita on lukuisa määrä. Kuuntelemisen helppouden alakäsitteessä, kaikki ryhmät ovat arvioineet kuuntelemisen helppouden puutteelliseksi, ainoina poikkeuksina huonon englannin kielen taidon omaavat (3.89) ja yli 55-vuotiaat vastaajat (3.80). Ihmispuheen kaltaisuus on saanut myös hyvin alhaiset tulokset. Kolme muuttujaa ylittää hyväksyttävään tulokseen, ja nämä ovat naisvastaajat (3.79), yli 55-vuotiaat (4.00) ja huonon englannin kielen taidon omaavat (3.89). Loput ryhmät ovat arvioineet ihmispuheen kaltaisuuden puutteellisena. Kuuntelukokemus sai



kokonaisuutena positiiviset arviot ja kaikki sen tulokset ylittävät hyväksytyt tulokset rajan. Huonon englannin kielen taidon omaavat vastaajat arvostelivat puhesynteesin naisäänen myös erinomaisella 4.24 tuloksella. Huomioitava havainto luonnollisuuden alakäsitteiden tuloksissa on myös se, että huonon englannin kielen omaavat vastaajat ovat mieltäneet naisäänen luonnollisempana kaikissa kolmessa alakäsitteessä, kuin hyvän englannin kielen taidon omaavat.

Puhesynteesin naisäänen luonnollisuuden alakäsitteiden tuloksia tarkasteltaessa suurimpana huomiona nousi puutteellisen tuloksen saaneiden arvioiden määrä. Kaksi kolmesta alakäsitteestä sai myös kokonaistuloksestaan puutteellisen tuloksen, joten vaikka luonnollisuuden pääkäsite on hyväksyttävä (3.74), on huomioitava se fakta, että kuuntelemisen helppous ja ihmispuheen kaltaisuus ei ollut tarvittavalla tasolla puhesynteesin naisäänen vastaanotossa.

**Taulukko 17.** Naisääni: ammattikäyttäjien ymmärrettävyyden alakäsitteet

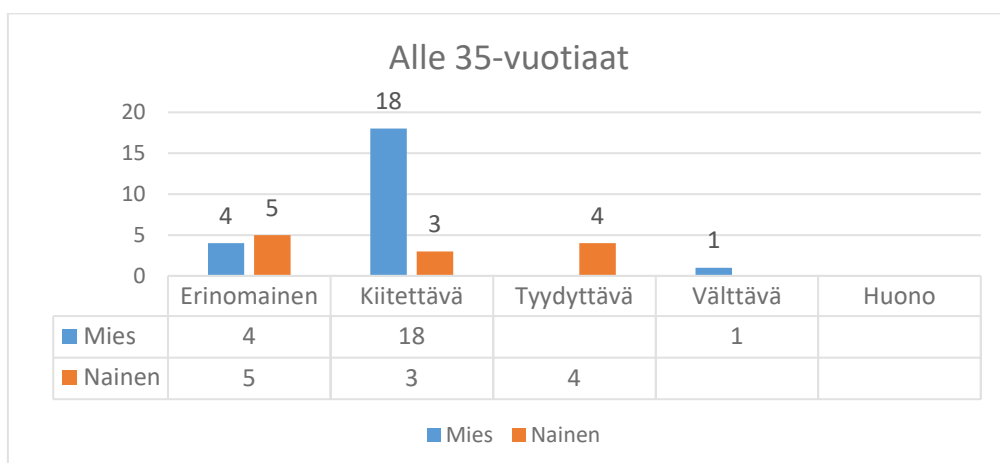


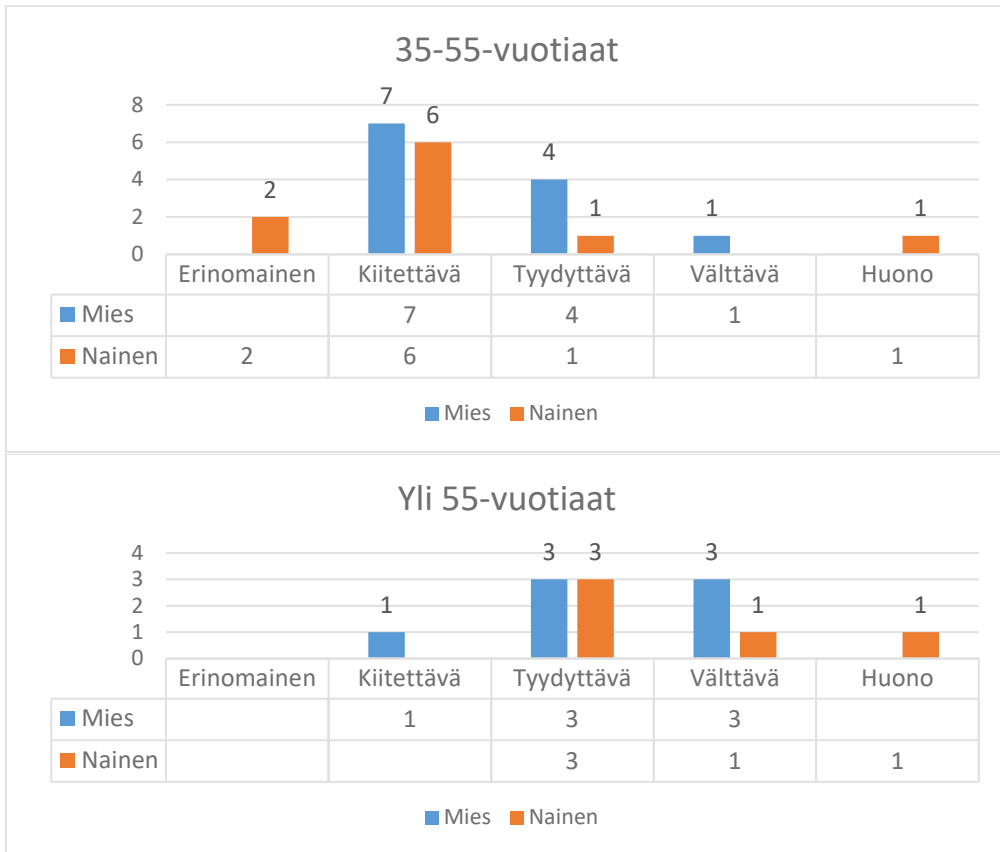
Puhesynteesin naisäänen ymmärrettävyyden arvioinnit on listattu taulukossa 17. Kokonaisarviot ylittävät kaikki kolme hyväksyttävän tuloksen rajoihin, ja erityisesti kuuntelun

helppous on saanut positiivisen vastaanoton, sisältäen vain yhden alle 4 menevän tuloksen 3.93 ja yhden erinomaisen tuloksen 4.30 yli 55-vuotiaiden arvioimana. Ymmärrettävyyden alakäsitteiden tulokset eivät myöskään sisällä ainoatakaan puutteellisen arvioitua tulosta, joten kokonaisuutena ymmärrettävyys on ollut puhesynteesin naisäänessä hyvä. Aiemmin esille nousseesta alle 35-vuotiaiden kriittisestä arvioinnista ei ole havaintoa ymmärrettävyyden alakäsitteissä. Alle 35-vuotiaat arvioivat kaikki kolme alakäsitettä 4 ylittävinä tuloksina. Englannin kielen taidon muuttujien eroavaisuus on käänteinen taulukossa 20 tehtyyn havaintoon siitä, että huonon englannin kielen tason valinneet vastaajat arvioivat puhesynteesin naisäänen luonnollisempaan. Ymmärrettävyyden tuloksissa, hyvän englannin kielen omaavat vastaajat ovat arvioineet naisäänen ymmärrettävämpänä kuin huonon englannin kielen taidon vastaajat.

## 5.2 Arkikäyttäjien MOS-tulokset

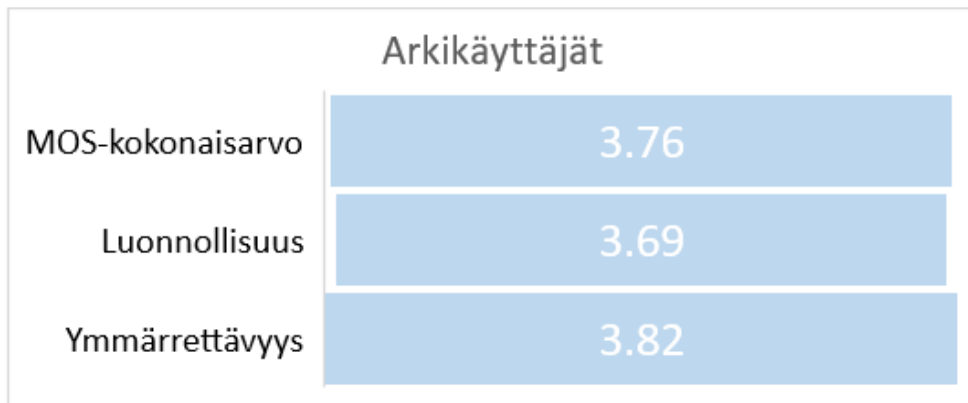
Arkikäyttäjien kyselyyn vastasi yhteensä 69 ihmistä ja vastaajina toimivat hanojen asentamiseen ja huoltamiseen kouluttamattomat henkilöt, eli hanojen arkikäyttäjät. Taulukossa 16 on jaoteltu kaikki arkikäyttäjryhmän kyselyyn vastanneet henkilöt ikäluokan, sukupuolen ja valitun englannin kielen osaamisen mukaan.





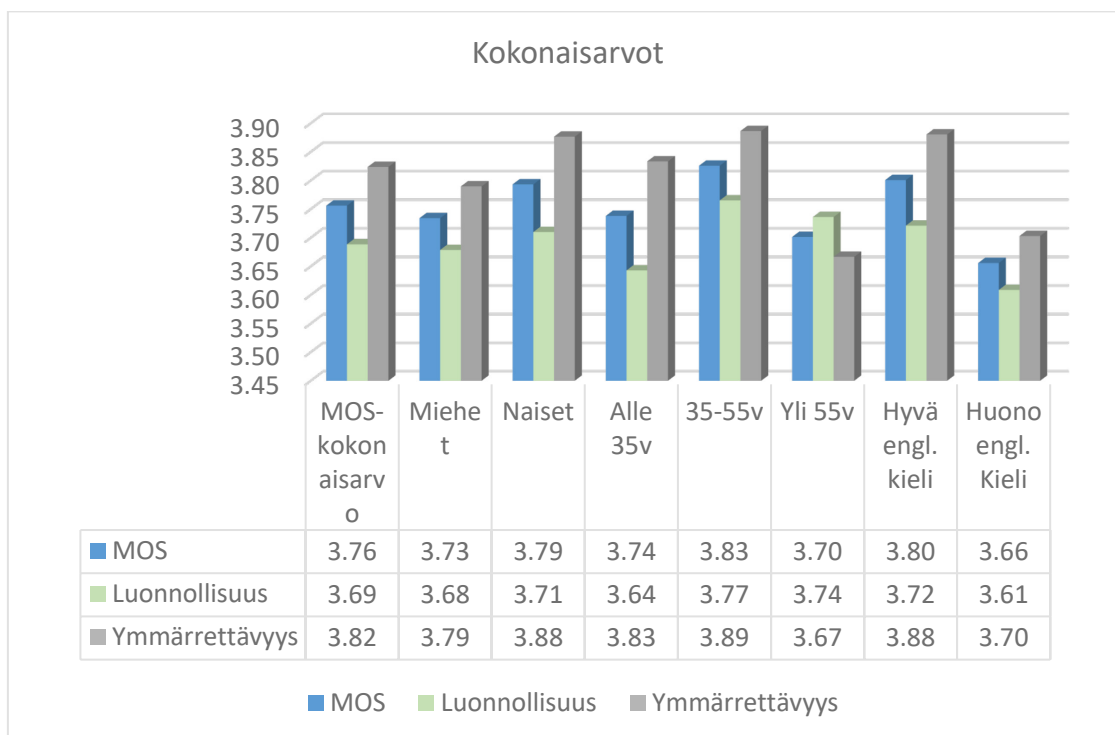
**Kuvio 11.** Arkikäyttäjien englannin kielen taito ikäluokkien mukaan

Kuviossa 11 on esitelty tarkennettu ikäjakauma englannin kielen taidon tason mukaan. Arkikäyttäjien englannin kielen taso jakautui tasaisemmin kuin ammattikäyttäjien vastaavat. Valtaosa vastaajista sijoittuu hyvän englanninkielen taidon tason ryhmään eli erinomaisen tai kiitettävän valinneisiin vastaajiin, mutta huonon tai välttävän valinnoita vastaajia on kuitenkin huomattavasti enemmän kuin ammattikäyttäjissä. Arkikäyttäjissä välttävän valinnoita oli 6 ja huonon valinnoita 2, kun taas ammattikäyttäjissä oli vain yksi välttävän valinnut vastaaja.



**Kuvio 12.** Arkikäyttäjien MOS-tulokset

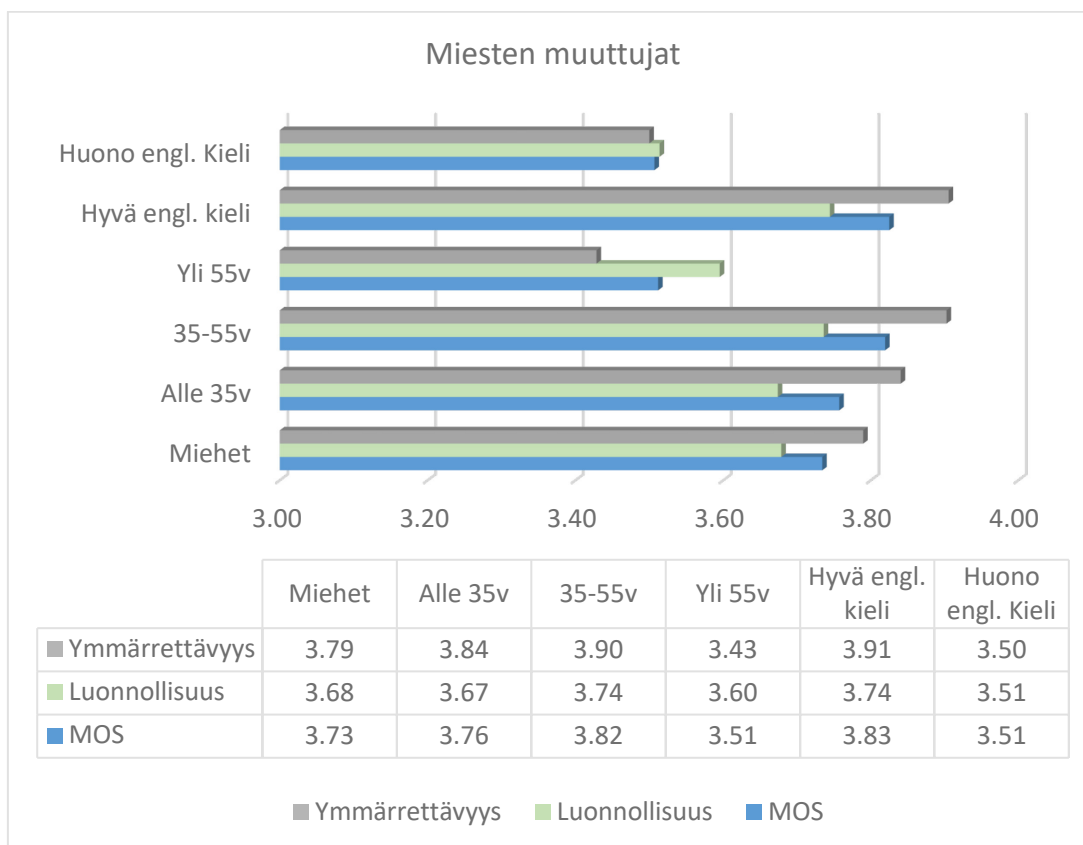
Kuviossa 12 on listattuna arkikäyttäjien koko kyselyn MOS-tulos sekä luonnollisuuden ja ymmärrettävyyden MOS-tulokset. Arkikäyttäjien tulokset saavuttavat hyväksytyt MOS-tulokset, mutta tulokset ovat selvästi lähempänä 3.67 alarajaa kuin ammattikäyttäjien tulokset. Ymmärrettävyyden ja luonnollisuuden kriteereiden välillä on myös eroavaisuus, ymmärrettävyyden ollessa hieman korkeammin arvioitu. Käyn läpi seuraavaksi ymmärrettävyyden ja luonnollisuuden kriteereitä sekä niiden alakäsitteitä, muuttujineen.

**Taulukko 18.** Arkikäyttäjien muuttujien MOS-tulokset

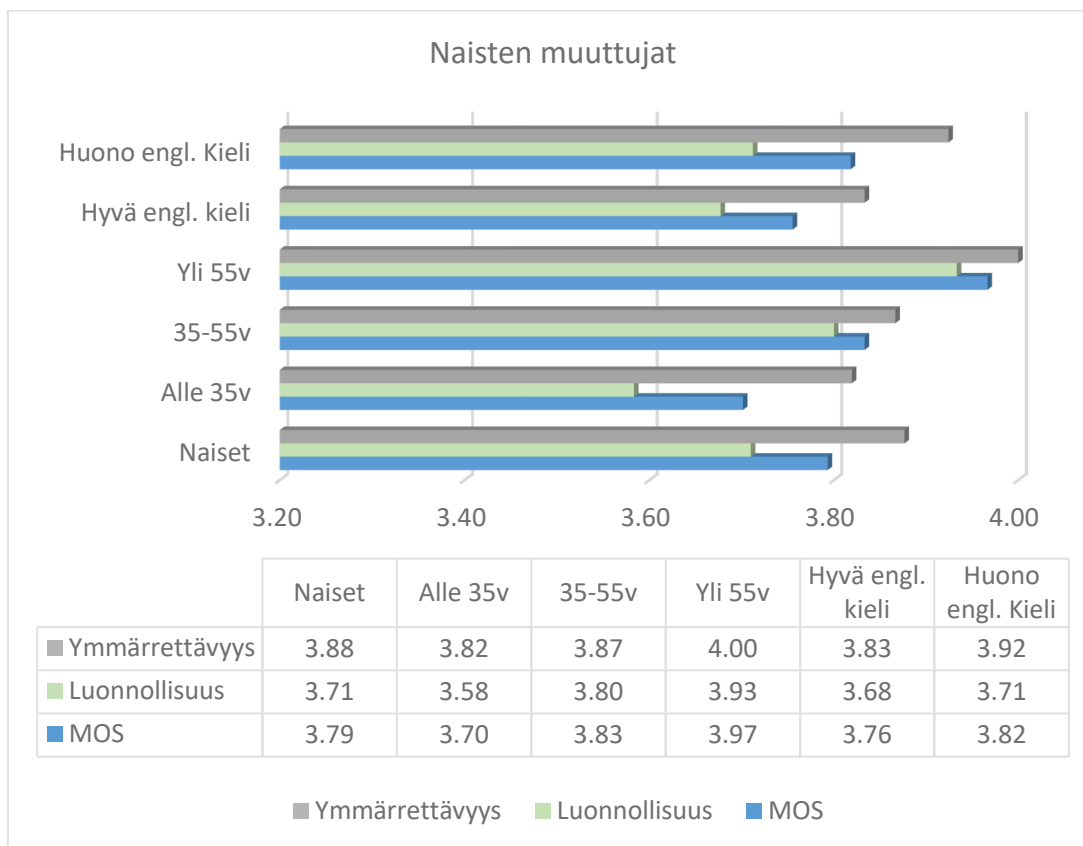
Taulukossa 18 on listattu arkikäyttäjien MOS-tulokset, jokaisen muuttujan mukaan. Arkikäyttäjien tulokset ovat alemmalla tasolla kuin aiemmin esitetyt ammattikäyttäjien tulokset. Tulokset ovat kauttaaltaan alemmalla tasolla, ja tuloksista löytyy kolme puutteellista MOS-tulosta. Arkikäyttäjät, jotka ovat valinneet huonon englanninkielen tason ovat arvioineet puhesynteesin kokonaisuutena puutteelliseksi (3.66). Myös luonnollisuus on arvioitu puutteellisena (3.61). MOS-kokonaistulos on juuri puutteellisen rajan (3.67) alapuolella, mutta luonnollisuus on jo jonkin verran tämän rajan alapuolella. Luonnollisuus on saanut heikoimmat arvioinnit kaikissa arvioinneissa, ja myös alle 35-vuotiaat ovat arvioineet sen puutteelliseksi (3.64). Sukupuolten tulosten välillä ei näyttäisi juurikaan olevan eroa. Ikäluokissa tosin löytyy vaihtelevuutta. MOS-kokonaistuloksen arvioinnissa 35-55-vuotiaiden ikäluokka on arvioinut puhesynteesin parhaimmaksi, nuorin alle 35-vuotiaiden ryhmä toiseksi parhaimmaksi ja heikoimman arvioinnin MOS-tulokselle ovat antaneet vanhin yli 55-vuotiaiden ikäluokka. Luonnollisuudessa parhaimman arvioinnin ovat antaneet 35-55-vuotiaat sekä vanhin ikäluokka, nuorimman ikäluokan arvostellessa luonnollisuuden heikoimmaksi. Ymmärrettävyyden arvioinnissa tämä järjestys on jälleen

vaihtunut, eli vanhin ikäluokka on arvostellut ymmärrettävyyden kriittisimmin ja nuoremmat ikäluokat ovat antaneet parhaimmat arvioinnit. Eri ikäluokkien tuloksissa ei näin ollen näyttäisi olevan samaa yhtenäisyyttä kuin ammattikäyttäjien tuloksissa, vaan tulokset näyttävät olevan hyvin vaihtelevia. Käyn seuraavaksi läpi tarkemmin nämä ikäluokat sekä englannin kielen taidon sukupuolien mukaan.

**Taulukko 19.** Arkikäyttäjät: Muuttujat miesvastaajien osalta

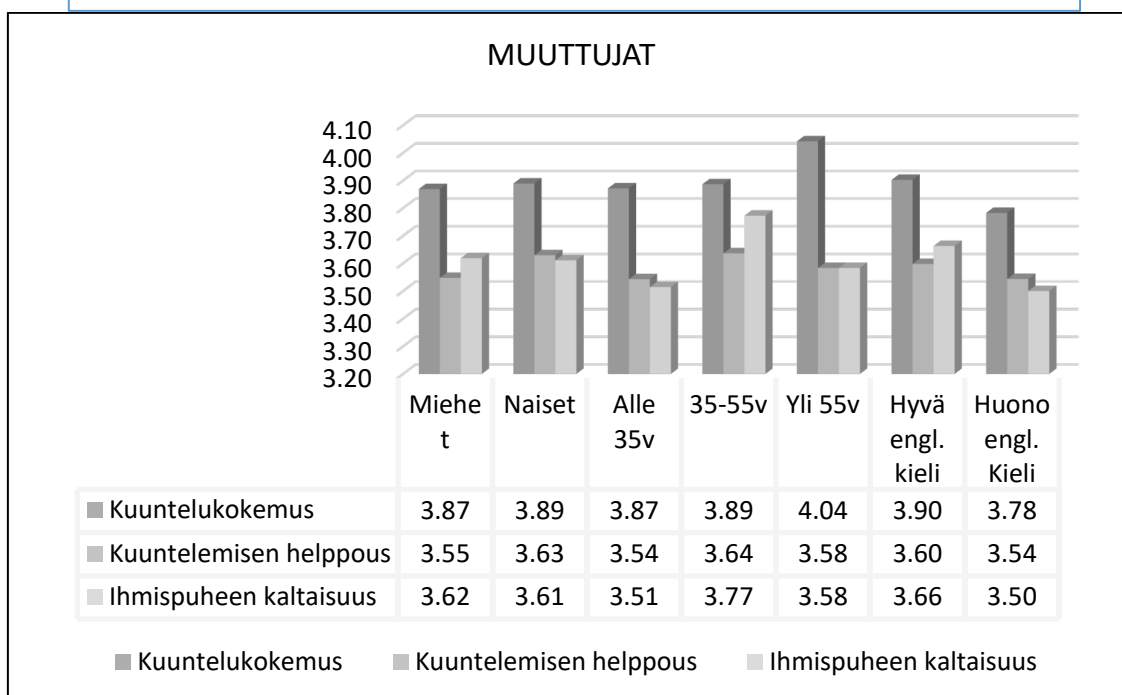
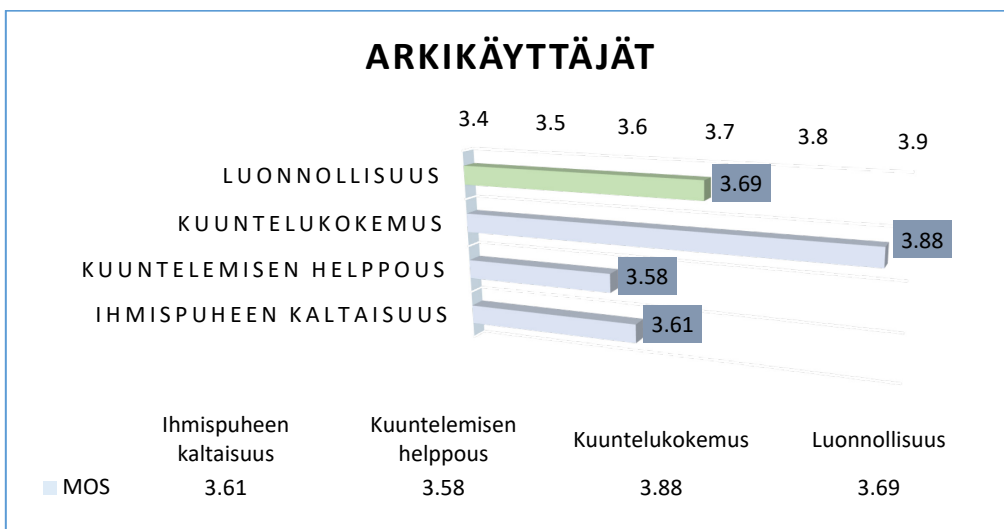


Taulukossa 19 on kuvattu arkikäyttäjien miesten tekemien arviointien ovat muuten tasapainoisia, lukuun ottamatta kahta vastaajaryhmää: yli 55-vuotiaat sekä huonon englannin kielen taidon omaavat vastaajat. Molemmat ryhmät ovat arvioineet puhesynteesin selkeästi puutteelliseksi kaikilla kolmella arvioinnilla. Heikoimmaksi on arvioitu ymmärrettävyys (3.43) yli 55-vuotiaiden tuloksissa. Tulokset ovat kuitenkin suhteessa toisiinsa, kuten kuviossa 11 kävi ilmi, suurin osa huonon englannin kielen tason valinneista vastaajista oli yli 55-vuotiaita, joten tässä mielessä on loogista, että nämä kaksi tulosta ovat samantasoisia.

**Taulukko 20.** Arkikäyttäjät: Muuttujat naisvastaajien osalta

Taulukossa 20 on listattu naisten tekemien arviointien muuttujien tulokset. Tulokset ovat kokonaisvaltaisesti korkeammat kuin miehillä, ja naisten tuloksissa on vain yksi puutteellisen arvioinnin saanut tulos. Alle 35-vuotiaat ovat arvioineet puhesynteesin luonnollisuuden puutteellisenä (3.58) kun taas kaikki muut tulokset yltyvät hyväksytyyn MOS-tuloksen välille. Ikäluokkien välinen arviointi noudattaa enemmän ammattikäyttäjien tuloksissa esiintynyttä mallia, jossa nuorin ikäluokka arvioi puhesynteesin kriittisimmin ja vanhemmat ikäluokat positiivisimmin. Naisten tuloksissa on myös vanhin ikäluokka arvioinut pääkriteerit ja kokonaisarvon parhaimmaksi ja nuorempiin ikäluokkiin mentäessä arvio laskee. Englannin kielessä on myös eroa hyvän englanninkielen ja huonon englanninkielen taitoisten ryhmien välillä. Huonon englanninkielen taidon omaavat arvostelivat puhesynteesin hieman hyvän englanninkielen taidon ryhmää paremmaksi. Järjestys on näin ollen käänteinen miesten ryhmään, jossa huonon englanninkielen taidon ryhmä arvioi puhesynteesin heikompana.

Taulukko 21. Arkikäyttäjät: luonnollisuuden alakäsitteet

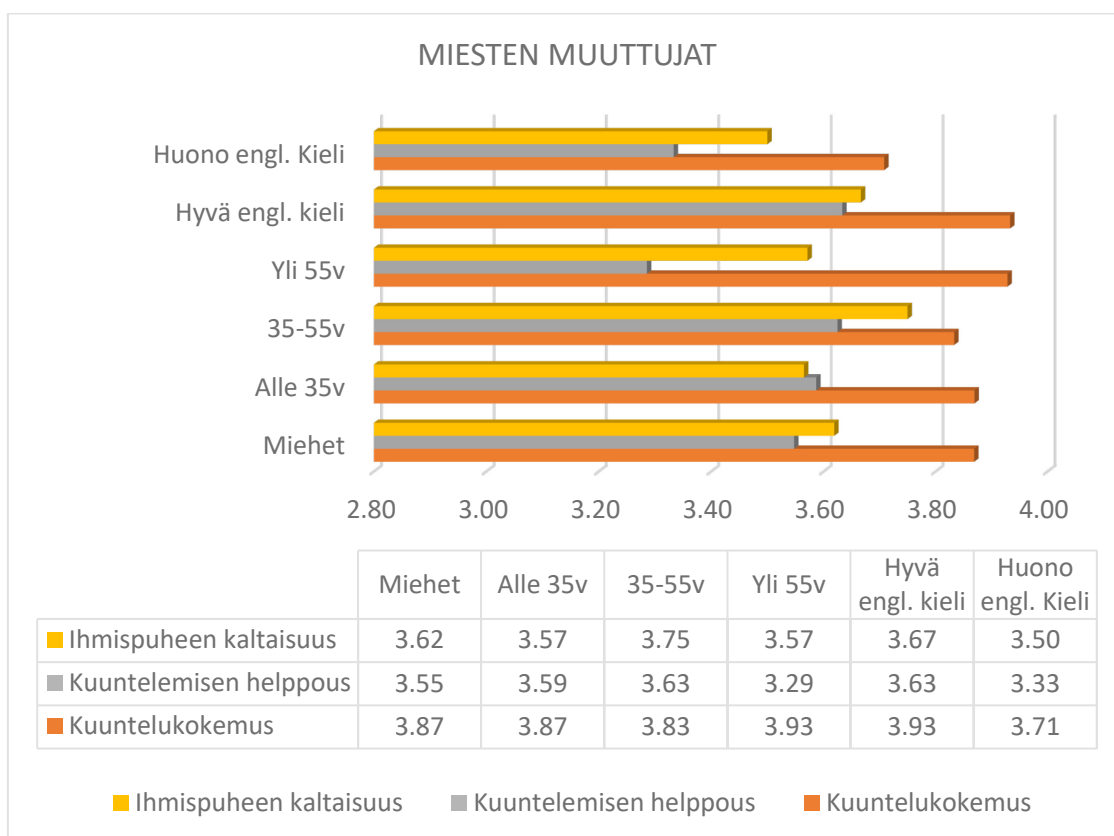


Luonnollisuuden alakäsitteet on kuvattu taulukossa ensin kokonaisarvojen kautta ja tarkemmin jokaisen muuttujan tulosten kautta. Kokonaisarvojen tulokset osoittavat, että kuuntelukokemus oli ainoa hyväksyttävään rajaan päässyt alakäsite. Kuuntelemisen helppous (3.58) ja ihmispuheen kaltaisuus (3.61) ovat molemmat puutteellisen tuloksen asteikolla. Kuuntelukokemuksen positiivinen arviointi osoittaa kuitenkin sen, että esimerkiksi puhesynteesin äänen ihmismäisyys, tai sen puute, ei ole vaikuttanut puhesynteesin kuuntelukokemukseen kokonaisuutena. Alakäsitteiden muuttujat mukailevat



kokonaisarviointien tuloksia. Kuuntelukokemuksen arvioinnit ovat kaikki hyväksyttävien tulosten rajoissa, kun taas kuuntelemisen helppouden ja ihmispuheen kaltaisuuden alakäsitteiden tulokset ovat kaikki, 35-55-vuotiaiden arviota ihmispuheen kaltaisuudesta lukuun ottamatta vain puutteellisia tuloksia. Vaihtelevuus on muutoin hyvin pientä eri muuttujien välillä, englannin kielen taito on näistä selkein eroavaisuus. Hyvän englanninkielen omaavat vastaajat arvioivat luonnollisuuden alakäsitteet paremmaksi kuin huonon englanninkielen taidon omaavat vastaajat. Merkittävin havainto arkikäyttäjien luonnollisuuden ja sen alakäsitteiden arvioinnista on kuitenkin kuuntelemisen helppouden ja ihmispuheen kaltaisuuden alhainen arviointi jokaisessa muuttujaryhmässä. Luonnollisuuden pääkäsite on pysynyt hyväksyttävissä rajoissa, mutta on tärkeää huomioida, että kaksi kolmesta alakäsitteestä on kuitenkin puutteellisen arvioinnin puolella.

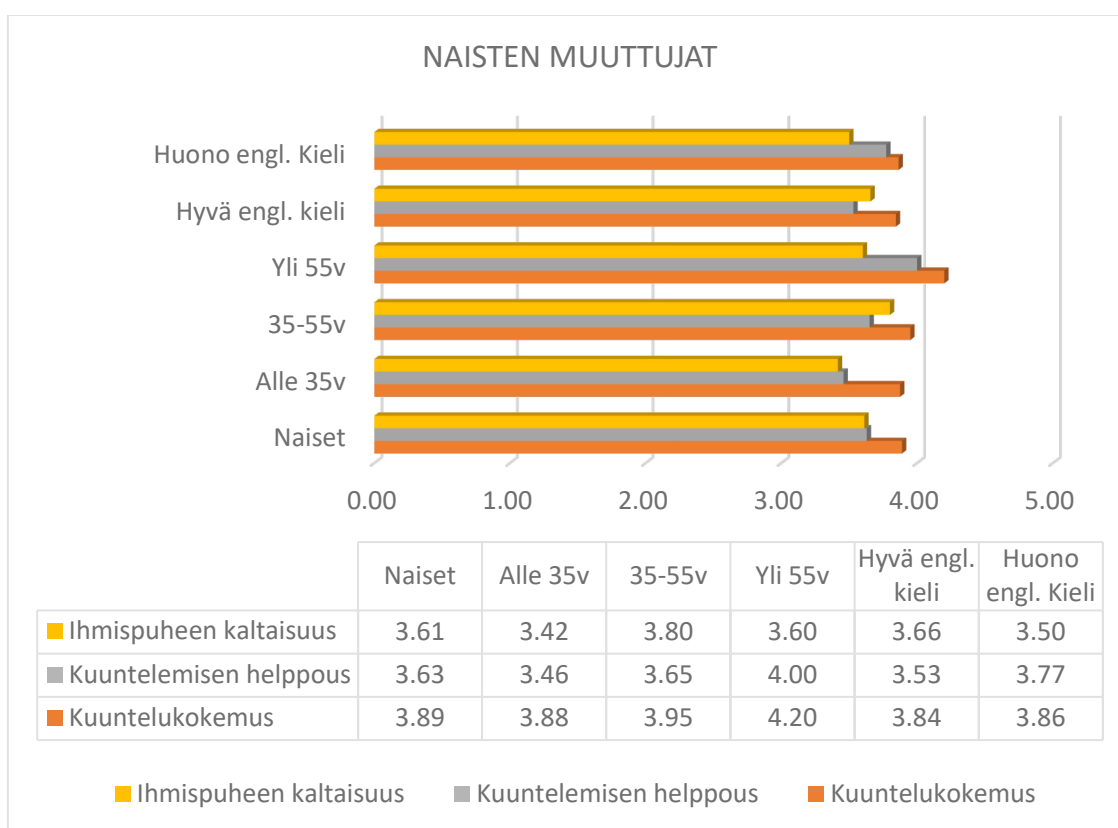
**Taulukko 22.** Miesten muuttujat: arkikäyttäjien luonnollisuuden alakäsitteet



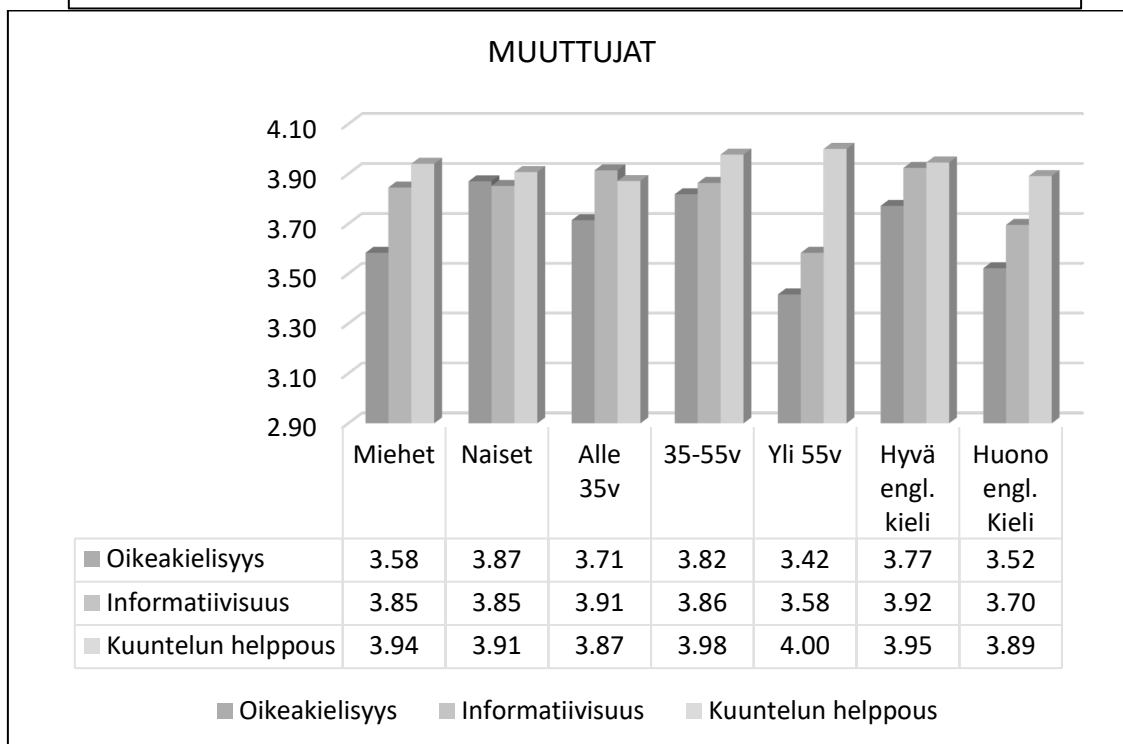
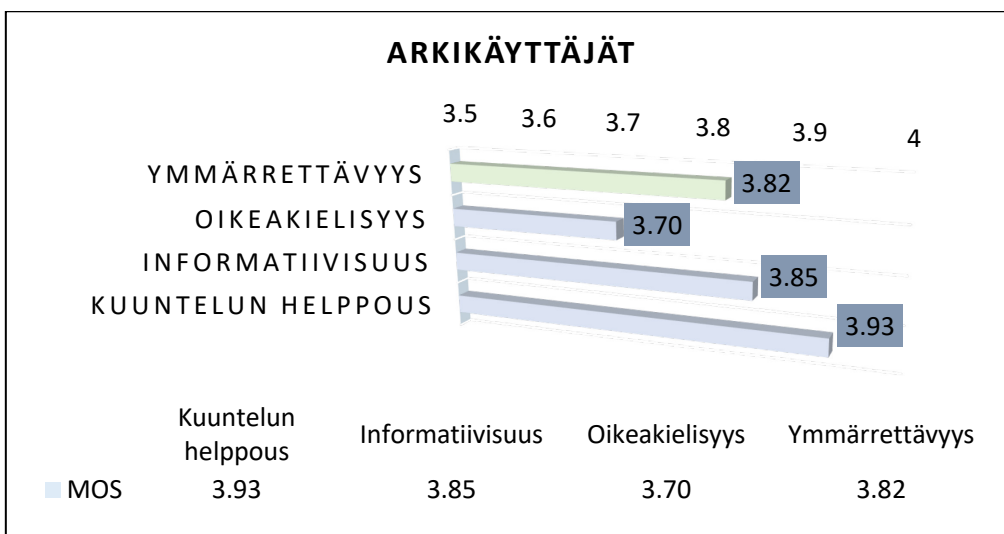
Miesten tekemien arviointien muuttujat luonnollisuuden alakäsitteissä kuvastavat hyvin kokonaisarvoissa listattuja havaintoja taulukossa 22. Kuuntelukokemus on ainoa

kokonaan hyväksyttävien arvioiden puolella oleva alakäsite, ja muut kaksi alakäsitettä on myös miesten arvioimina puutteellisen arvioinnin puolella. Selkein havainto alhaisesta arvioinnista on huonon englannin kielen tason valinneilla vastaajilla ja yli 55-vuotiailla vastaajilla, jotka molemmat arvioivat kuuntelemisen helppouden selvästi puutteellisena (3.33 ja 3.29). Huono englannin kielen taito näyttää vaikuttaneen puheesynteesin vastaanottoon negatiivisesti miesten arvioinneissa, ja erityisesti tämä korostui kuuntelemisen helppoudessa, tai tässä tapauksessa sen vaikeudessa.

**Taulukko 23.** Naisten muuttujat: arkikäyttäjien luonnollisuuden alakäsitteet



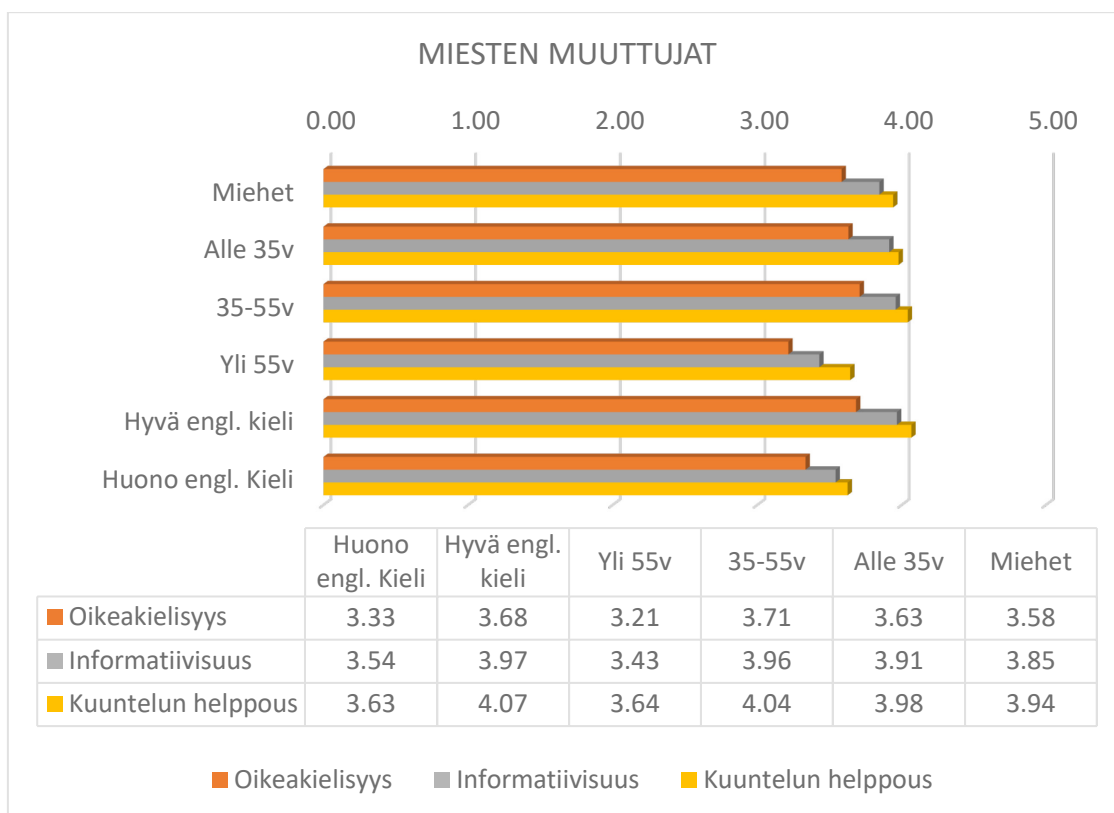
Naisten tekemät arvioinnit luonnollisuuden alakäsitteistä muuttujineen on kuvattu taulukossa 23. Tulokset noudattavat samaa kaavaa kokonaistulosten osalta. Kuuntelukokemus on saanut positiiviset tulokset, kun taas kuuntelemisen helppous ja ihmispuheen kaltaisuus on alhaisella tasolla arvioinneiltaan. Tämä korostui eritoten alle 35-vuotiaiden arvioinneissa, 3.42 ihmispuheen kaltaisuudelle sekä 3.46 arviolle kuuntelemisen helppoudelle. Englannin kielen taidolla ei näytä olleen selkeää ja johdonmukaista vaikutusta vaan tulokset vaihtelevat jokaisessa alakäsitteessä.

**Taulukko 24.** Arkikäyttäjät: ymmärrettävyyden alakäsitteet

Arkikäyttäjien ymmärrettävyyden ja sen alakäsitteiden arviointien tulokset on listattu taulukkoon 24. Ymmärrettävyys ja sen alakäsitteet sijoittuvat kaikki hyväksyttävän tuloksen rajalle ja näin ollen puheesynteesin ymmärrettävyys on koettu positiivisemmin kuin puheesynteesin luonnollisuus. Kuuntelun helppous arvioitu alakäsitteistä parhaimmaksi ja oikeakielisyys on saanut heikoimman arvostelun.

Muuttujien tuloksissa on havaittavissa muutama poikkeus. Yli 55-vuotiaat ovat arvioineet oikeakielisyyden ja informatiivisuuden selvästi muita ikäluokkia alhaisemmaksi ja molemmat näistä alittavat puutteellisen arvion rajan (3.42 ja 3.58). Huono englanninkielen taito on myös vaikuttanut oikeakielisyyden arviointiin (3.52). Miehet ovat arvioineet oikeakielisyyden naisia heikommaksi ja miesten arviointi sijoittuu myös puutteellisen arvion puolelle (3.58). Seuraavaksi käyn läpi miesten ja naisten arvioinnit ja muuttujat erikseen, josta voi nähdä tarkemmin miten iso ero sukupuolten välisessä arvostelussa on ollut.

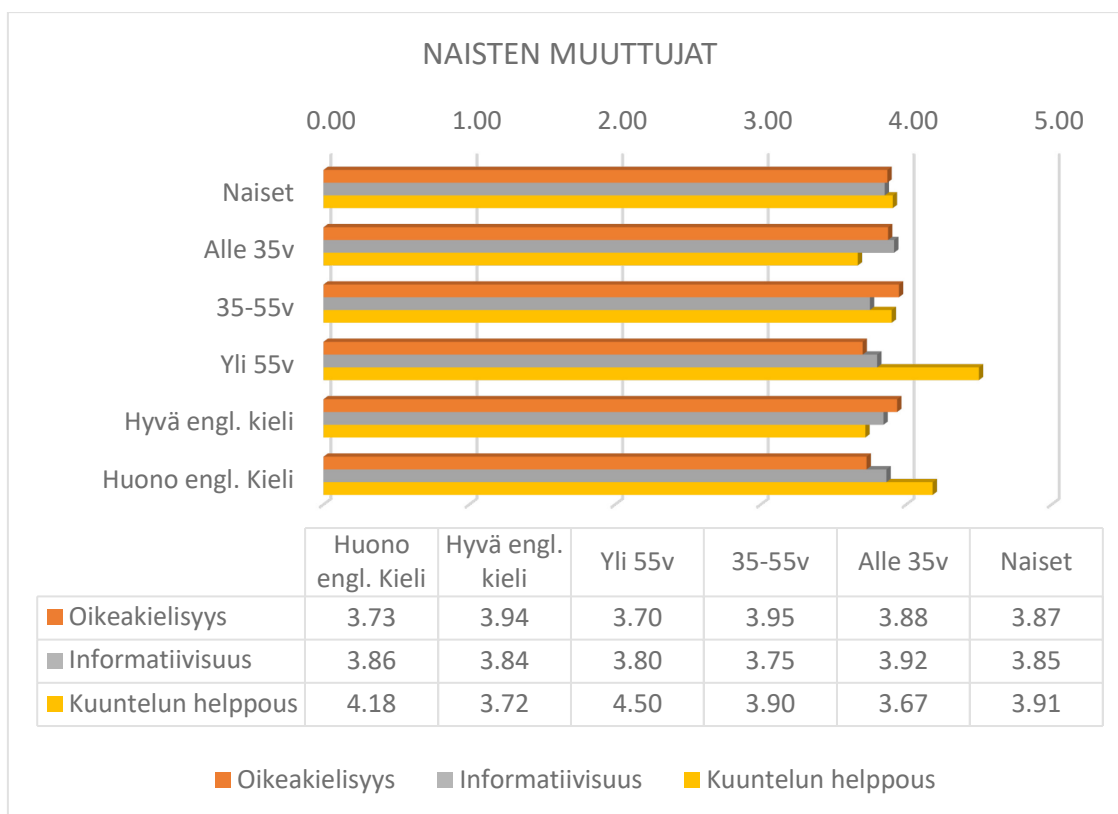
**Taulukko 25.** Arkikäyttäjien ymmärrettävyyden alakäsitteet: Miesvastaajat



Miesten tekemien arviointien tuloksista käy ilmi taulukossa 25, että kriittisimmät ryhmät ovat olleet yli 55-vuotiaat ja huonon englanninkielen tason omaavat vastaajat kaikissa kolmessa alakäsitteessä. Alle 35-vuotiaat ovat myös arvioineet oikeakielisyyden puutteelliseksi (3.63). Muutoin tulokset ovat kaikki hyväksyttävien tulosten puolella. Ikä ja englanninkielen taidon taso ovat näin ollen olleet tärkeimmät vaikuttavat tekijät

oikeakielisyyden ja informatiivisuuden arvioinnissa. Huonolla englanninkielen tasolla ja vanhemmalla ikäluokalla todettiin jo aiemmin olevan yhteys, joten on johdonmukaista, että puhesynteesin ymmärtäminen on kärsinyt, jos englannin kielen ymmärtäminen itsessään on heikommalla tasolla.

**Taulukko 26.** Arkikäyttäjien ymmärrettävyyden alakäsitteet: Naisvastaajat



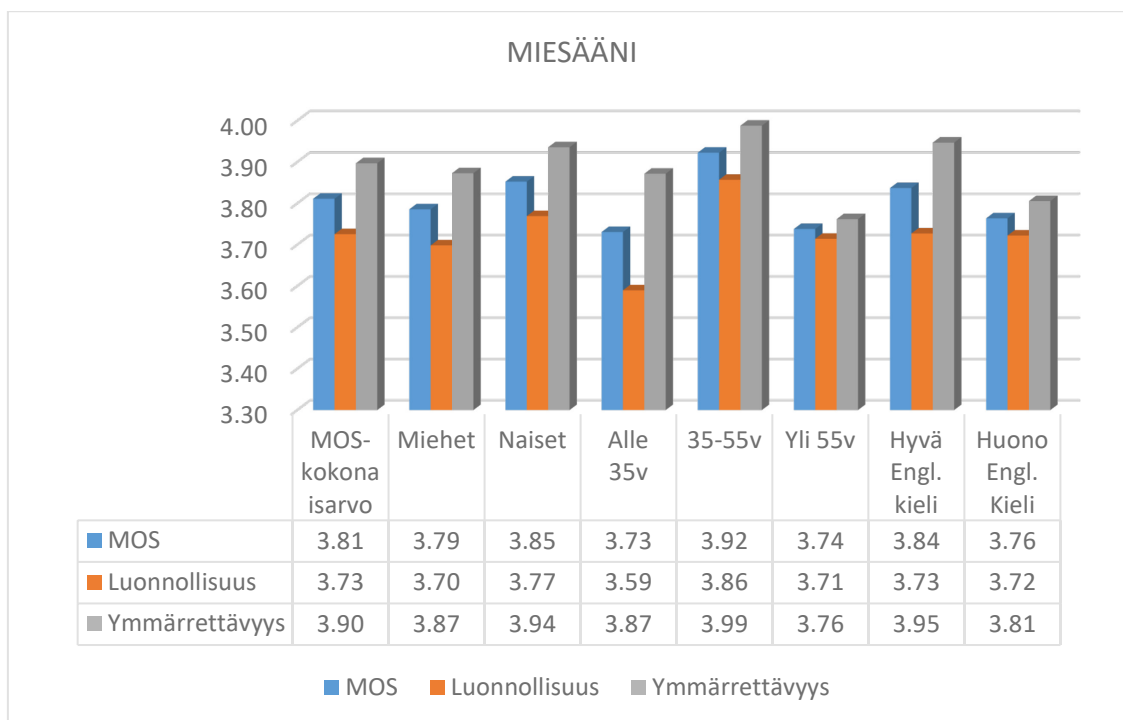
Naisten tekemät arvioinnit puhesynteesin ymmärrettävyydestä ja sen alakäsitteistä ovat kaikki hyväksyttävien tulosten rajoissa, kuten taulukosta 26 voi havaita. Englannin kielen tasolla tai iällä ei näytä olleen täysin samoja vaikutuksia kuin miesten tekemissä arvioinneissa. Molempien ryhmien arviot ovat hyvin vaihtelevat. Esimerkiksi huonon englanninkielen taidon omaavat ovat arvostelleet kuuntelun helppouden huomattavasti hyvän englannin kielen omaavia paremmaksi. Oikeakielisyyden arvioinnissa tämä on taas päinvastoin, ja informatiivisuuden arvioinnit ovat puolestaan melkein identtisiä. Ikä jäljittelee näitä samoja piirteitä. Vanhin ikäluokka arvioi kuuntelun helppouden erinomaisena (4.50) ja oikeakielisyyden heikoimpana (3.70), joten selkeää johdonmukaisuutta ei ikäluokista myöskään löydy. Oikeakielisuus ja yli 55-vuotiaiden ryhmä ovat kuitenkin myös naisten

vastaajissa alhaisimmat arviot antaneet ryhmät, vaikkakin tulokset eivät ole niin alhaiset kuin miesten tulokset.

### 5.2.1 Miesääni arkikäyttäjien arvioimana

Arkikäyttäjät arvioivat ammattikäyttäjien tapaan kahta eri puhesynteesin ääntä: miehen ääntä sekä naisen ääntä. Arkikäyttäjät jaoteltiin kahteen ryhmään sen perusteella minkä sukupuolen äänellä tehdyn videon he katsoivat. Tässä luvussa esitellään arkikäyttäjien puhesynteesin miesäänellä katsoneiden ryhmän tulokset.

**Taulukko 27.** Puhesynteesi miesäänellä: arkikäyttäjien kokonaisarviot

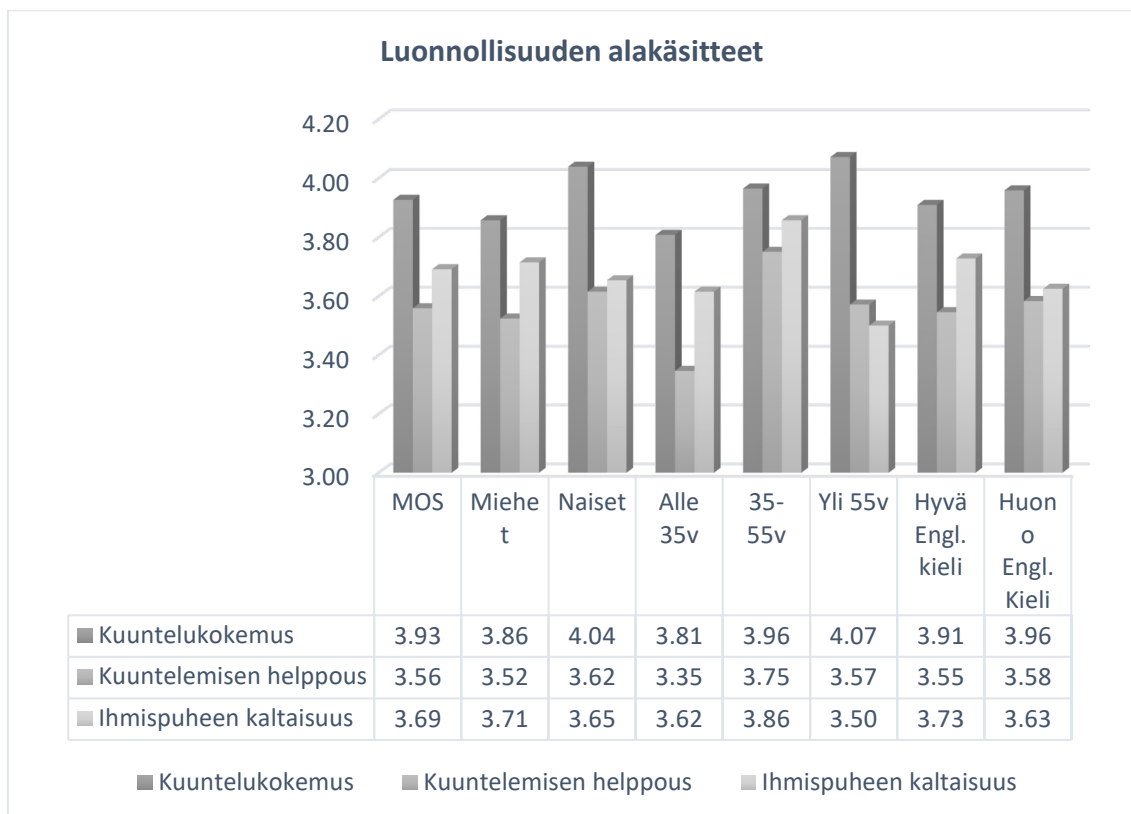


Taulukossa 27 on listattu puhesynteesin miesäänien arvioinnin tulokset. MOS-kokonaisarvot ovat kaikilta kolmelta kriteeriltä hyväksyttävällä tasolla. Ymmärrettävyyden arviointi parhaimpina on johdonmukainen aiempiin tuloksiin verrattuna. Naiset arvioivat myös puhesynteesin hieman paremmaksi kuin miehet, mikä on myös johdonmukaista aiempien arkikäyttäjien tulosten valossa. Alle 35-vuotiaat ovat olleet kriittisimmät arvioinneissaan, ja 35-55-vuotiaiden ikäluokka on arvioinut puhesynteesin miesäänien

parhaimmaksi. Alle 35-vuotiaiden ryhmän arvioinneista löytyy myös ainut puutteellisen arvioinnin tulos (3.59), puhesynteesin miesäänien luonnollisuudesta. Englannin kielen taito vaikuttaa myös hieman puhesynteesin vastaanottoon, hyvän englannin kielen omaavat arvioivat miesäänien kaikilla kriteereillä hieman paremmaksi kuin huonon englannin kielen omaavat vastaajat.

Arkikäyttäjien tulokset puhesynteesin miesäänestä ovat hyvin tasapainoiset, lukuun ottamatta yhtä puutteellisen tuloksen saanutta tulosta. Kaikki tulokset pysyvät myös alle 4, joten tulokset ovat riittävällä tasolla ilman selkeitä kohokohtia tai eroavaisuuksia.

**Taulukko 28.** Miesääni: arkikäyttäjien luonnollisuuden alakäsitteet



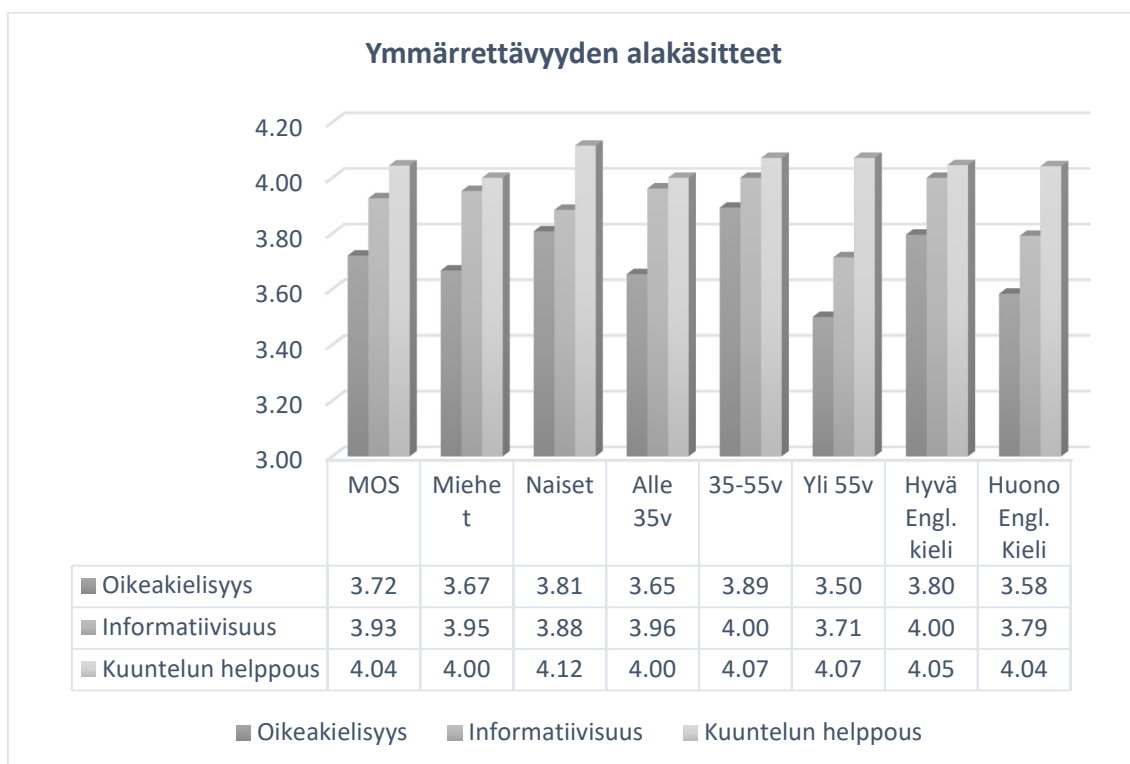
Luonnollisuuden alakäsitteet on arvioitu taulukossa 28. Kuuntelemisen helppous on saanut alhaisimmat tulokset, ja sen MOS-kokonaistulos 3.56 alittaa puutteellisen tuloksen rajan. Ihmispuheen kaltaisuus on saanut myös alhaisen tuloksen 3.69, mutta on vielä hyväksyttävien tulosten rajalla. Kuuntelukokemus on myös tässä ryhmässä arvioitu parhaimpana, ja tämä mukailee analyysin aiempia tuloksia.

Kuuntelemisen helppous on selvästi heikoiten arvioitu alakäsite, ja se sisältääkin vain yhden hyväksyttävän tuloksen 3.75, jonka ovat antaneet 35-55-vuotiaiden vastaajien ryhmä. Kaikki muut tulokset ovat puutteellisen tuloksen puolella ja erityisesti alle 35-vuotiaat ovat arvioineet puhesynteetin miesäänien kuuntelun helppoutta hyvin alhaisena 3.35 tuloksella. Alle 35-vuotiaat ovat tosin myös kriittisimmät arvioijat tässäkin ryhmässä, mutta kuuntelun helppouden 3.35 tulos on samalla myös kyselyn alhaisimpia tuloksia.

Ihmispuheen kaltaisuus on myös saanut neljä puutteellisen tuloksen arviota. Naiset 3.65, alle 35-vuotiaat 3.62, yli 55-vuotiaat 3.50 sekä huonon englannin kielen taidon vastaajat 3.63. Ihmispuheen kaltaisuus ylittää kuitenkin kokonaisarvioltaan juuri ja juuri hyväksyttävän tuloksen rajan, joten ihmispuheen kaltaisuus on kuitenkin riittävällä tasolla.

Kuuntelemisen helppouden alhainen tulos ja kuuntelukokemuksen korkea tulos on havainto, joka on noussut esiin monesti analyysissä. Kuuntelukokemus on mielletty hyvänä, mutta kuuntelemisen helppous on monesti saanut alhaisen tai jopa puutteellisen tuloksen. Ihmispuheen kaltaisuus on myös saanut vaihtelevia tuloksia, mutta on tärkeää havaita, että ihmispuheen kaltaisuuden heikompi taso ja kuuntelemisen vaikeus eivät ole kuitenkaan vaikuttaneet negatiivisesti itse kuuntelukokemukseen ja sen arviointiin.



**Taulukko 29.** Miesääni: arkikäyttäjien ymmärrettävyyden alakäsitteet

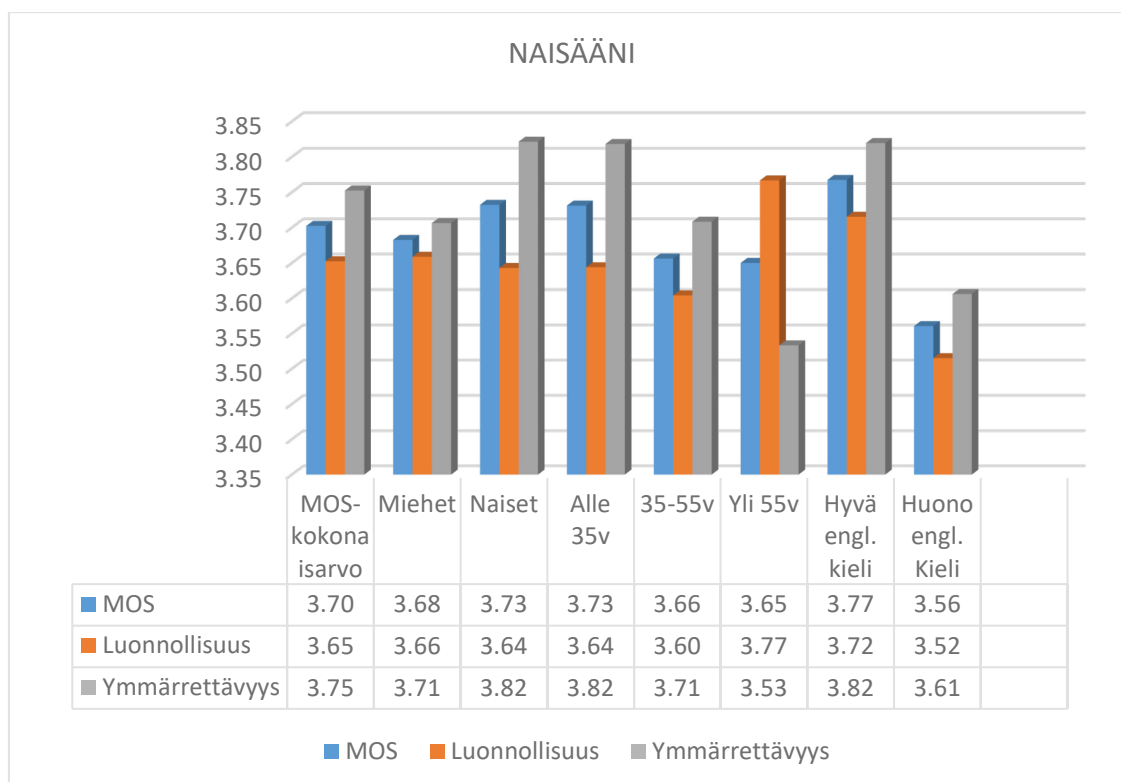
Ymmärrettävyyden alakäsitteet ovat hyväksyttävällä tasolla MOS-tuloksiltaan, kuten taulukosta 29 käy ilmi. Oikeakielisuus on ainut alakäsite, joka on saanut alhaisempia tuloksia arvioinnista. Alle 35-vuotiaat ovat arvioineet puhesynteesin miesäänien oikeakielisyyden puutteelliseksi 3.65 tuloksella, kuten ovat myös yli 55-vuotiaat 3.50 tuloksella ja huonon englannin kielen taidon omaavat vastaajat 3.58 tuloksella. Hyvän englannin kielen taidon omaavat ovat kuitenkin arvioineet oikeakielisyyden hyväksyttävänä 3.80 tuloksella, joten englannin kielen osaamisella näyttäisi tässä tapauksessa olevan merkitystä. Informatiivisuus on hyvällä tasolla. Ainoa poikkeama on yli 55-vuotiaiden muihin informatiivisuuden tuloksiin verrattuna alhainen tulos (3.71). Yli 55-vuotiaat ovat olleet myös kriittisin arvioija-ryhmä puhesynteesin miesäänien ymmärrettävyyden osalta.

Ymmärrettävyyden tulokset ovat kuitenkin hyväksyttävät ja vaikkakin muutama puutteellinen tulos löytyi muuttujista, on MOS-kokonaistulos kaikilta kolmelta alakäsitteeltä hyväksyttävällä tasolla. Arkikäyttäjät ovat arvioineet puhesynteesin miesäänien ymmärrettävyyden hyväksyttävästi.

## 5.2.2 Naisääni arkikäyttäjien arvioimana

Tässä luvussa esitellään puhesynteesin naisäänen arviointi arkikäyttäjien tekemänä MOS-tuloksen sekä luonnollisuuden ja ymmärrettävyyden kriteereiden ja näitä tarkentavien alakäsitteiden kautta.

**Taulukko 30.** Puhesynteesi naisäänellä: arkikäyttäjien kokonaisarviot

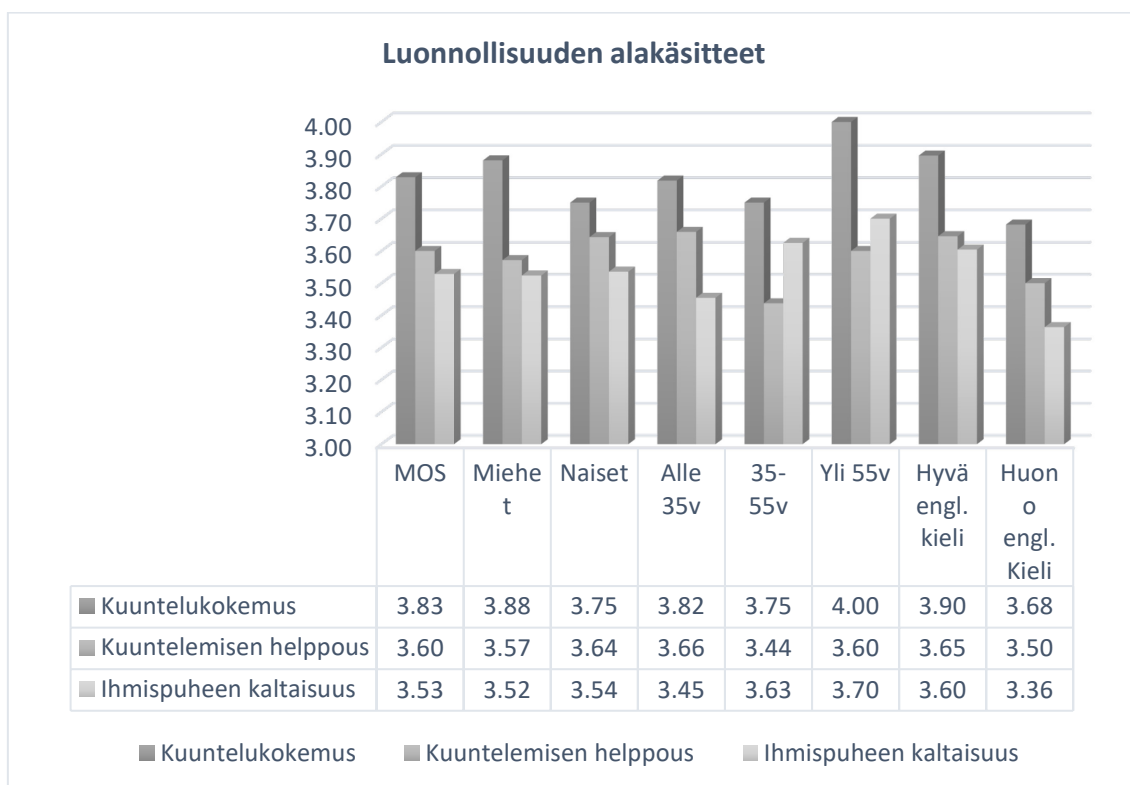


Arkikäyttäjät ovat arvioineet puhesynteesin naisäänen MOS-kokonaistuloksen ja ymmärrettävyyden alhaisella tuloksella, mutta kuitenkin hyväksyttävän tuloksen rajoissa. Luonnollisuus on tuloksella 3.65 puutteellinen. Luonnollisuus on arvioitu hyväksyttäväksi vain yli 55-vuotiaiden ryhmässä tuloksella 3.77 sekä hyvän englannin kielen omaavien arvioissa 3.72. Muut muuttujaryhmät ovat arvioineet luonnollisuuden puutteelliseksi. Ymmärrettävyys on saanut myös kaksi puutteellista tulosta. Yli 55-vuotiaat ovat arvioineet ymmärrettävyyden puutteelliseksi tuloksella 3.53, ja huonon englannin kielen taidon omaavat tuloksella 3.61. MOS-kokonaistulos on saanut kolmelta eri muuttujaryhmältä puutteellisen tuloksen. 35-55-vuotiaiden MOS-kokonaistulos on puutteellinen 3.66, yli

55-vuotiaiden tulos on 3.65, ja myös huonon englannin kielen omaavat vastaajat ovat arvioineet MOS-kokonaistuloksen puutteelliseksi tuloksella 3.56.

Kokonaistulokset ovat näin ollen riittävän ja huonon tuloksen välimaastossa. Puhesynteesin naisääni sai myös ammattikäyttäjien tekemässä arvioissa taulukossa 13 alhaiset arvioinnit, ja arkikäyttäjien tulos näyttäisi noudattavan tätä samaa kaavaa. Arkikäyttäjien arviot ovat kuitenkin kriittisemmät, ja tätä korostaa yhden pääkriteerin arviointi puutteelliseksi.

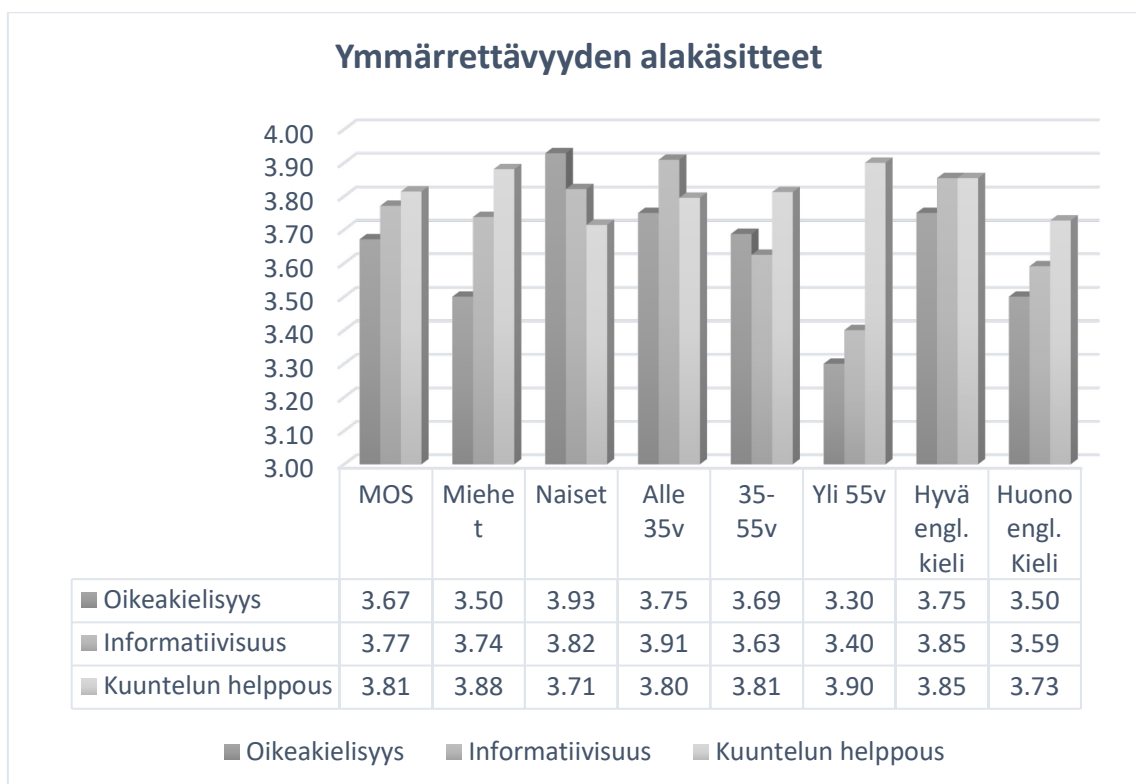
**Taulukko 31.** Naisääni: arkikäyttäjien luonnollisuuden alakäsitteet



Luonnollisuuden alakäsitteet on esitetty taulukossa 31. Kuuntelukokemus on hyväksyttävällä tasolla MOS-tulokseltaan sekä kaikkien muuttujien arvioissa. Aiempien analyysien tapaan kuuntelemisen helppoutta ja ihmispuheen kaltaisuutta on pidetty selvästi kuuntelukokemusta heikompana. Ihmispuheen kaltaisuus sisältää vain yhden hyväksyttävän tuloksen yli 55-vuotiaiden arvioinnissa (3.70). Kuuntelemisen helppous ei ole saanut ainuttakaan hyväksyttävän tason tulosta, mutta se on kuitenkin MOS-kokonaistulokseltaan paremmaksi arvioitu kuin ihmispuheen kaltaisuus. Tämä johtuu huonon

englannin kielen omaavien erittäin alhaisesta arvioinnista puhesynteesin naisäänen ihmispuheen kaltaisuudesta, tuloksella 3.36. Alle 35-vuotiaat ovat myös arvioineet ihmispuheen kaltaisuuden hyvin alhaiseksi tuloksella 3.45. Yksi tärkeä havainto, joka pitää nostaa esiin on myös kuuntelukokemuksen hieman alhaisempi tulos verrattuna esimerkiksi aiempaan taulukkoon 28 puhesynteesin miesäänen luonnollisuuden alakäsitteistä. Kuuntelukokemus on myös lähempänä puutteellisen tuloksen rajaa naisäänen arvioissa kuin muissa aiemmissa tuloksissa. Kuuntelukokemuksen korkeat arviot ovat aiemmissa tuloksissa nostaneet luonnollisuuden hyväksytylle rajalle, mutta naisäänen tuloksissa myös kuuntelukokemuksen alhaiset tulokset laskevat luonnollisuuden puutteelliseksi.

**Taulukko 32.** Naisääni: arkikäyttäjien ymmärrettävyyden alakäsitteet



Ymmärrettävyyden alakäsitteet on kuvattu taulukossa 32. Kaikki kolme alakäsitettä ovat MOS-kokonaistulokseltaan hyväksytyllä tasolla, tosin oikeakielisuus on juuri 3.67 hyväksytyyn rajan tuloksessa. Kuuntelun helppous on parhaimmaksi arvioitu alakäsite ja sen tuloksissa ei ole ainuttakaan puutteellista tulosta. Informatiivisuus ja oikeakielisuus sisältävät molemmat 3 puutteellista tulosta. Yli 55-vuotiaat ovat arvioineet informatiivisuuden 3.40 ja oikeakielisyyden 3.30. Huonon englannin kielen omaavat vastaajat ovat

arvioineet informatiivisuuden 3.59 ja oikeakielisyyden 3.50. Eroavaisuus tulee viimeisissä puutteellisissa tuloksissa, miesvastaajien arvioidessa oikeakielisyyden 3.50 ja 35-55-vuotiaiden arvioidessa informatiivisuuden 3.63 tuloksella puutteelliseksi. Muut tulokset ovat tasaisesti hyväksytyin tuloksen alarajan tuntumassa, ja korkeita yli 4 tuloksen ylittäviä arvioita ei ole lainkaan. Puhesynteesin naisäänen ymmärrettävyys on kuitenkin arvioitu hyväksyttäväksi, koska se ei sisällä kokonaistuloksiltaan puutteellisia tuloksia. Naisäänen tulokset ovat matalat verrattuna luvun 5.2.1 miesäänen tuloksiin ymmärrettävyydessä, ja ne ovat myös koko tutkielman ymmärrettävyyden tuloksiin verrattaessa matalat.

### 5.3 Varianssianalyysin tulokset

Varianssianalyysissä vertailen tutkimuskysymysten mukaan ammattikäyttäjien ja arkikäyttäjien, puhesynteesin mies- ja naisäänen sekä kaikkien näiden sisäisten muuttujien eroavaisuuksia. Aloitan vertailun ammattikäyttäjien ja arkikäyttäjien tuloksista, jonka jälkeen siirryn nais- ja miesäänien vertailuun.

**Taulukko 33.** Varianssianalyysi: kokonaistulokset

AMMATTIKÄYTTÄJÄT	MOS	Luonnollisuus	Ymmärrettävyys
Kokonaistulos	3.94	3.85	4.04
Miehet	3.91	3.79	4.03
Naiset	3.99	3.93	4.05
Alle 35v	3.82	3.69	3.96
35-55v	3.96	3.84	4.04
Yli 55v	4.10	4.06	4.15
Hyvä engl. kieli	3.97	3.86	4.08
Huono engl. kieli	3.88	3.83	3.92
ARKIKÄYTTÄJÄT	MOS	Luonnollisuus	Ymmärrettävyys
Kokonaistulos	3.76	3.69	3.82
Miehet	3.73	3.68	3.79
Naiset	3.79	3.71	3.88
Alle 35v	3.74	3.64	3.83
35-55v	3.83	3.77	3.89
Yli 55v	3.70	3.74	3.67
Hyvä engl. kieli	3.80	3.72	3.88
Huono engl. Kieli	3.66	3.61	3.70

Ammattikäyttäjien ja arkikäyttäjien tulokset muuttujineen on listattu taulukkoon 33. Selkein ja suoraviivaisin ero ammattikäyttäjien ja arkikäyttäjien välillä oli se, että ammattikäyttäjät arvioivat puhesynteesin paremmaksi kuin arkikäyttäjät. Toisin sanoen ammattikäyttäjät kokivat puhesynteesin parempana kuin arkikäyttäjät. Tähän suurimpana syynä on kokonaistulosten kannalta se, että ammattikäyttäjät ovat koulutettuja videolla esitettyihin huoltotoimenpiteisiin ja heillä on näin parempi tietämys videolla esiintyvistä termeistä ja toimenpiteistä. Ero on noin 0.20 pisteen luokkaa kokonaistuloksissa, mitä voidaan pitää jo merkittävänä erona. Kokonaistulosten välillä ei ole myöskään paljon vaihtelua vaan MOS, luonnollisuus ja ymmärrettävyys ovat kaikki tässä 0.20 vaihteluvälillä. Ammattikäyttäjien kaikki tulokset yltyvät myös hyväksytyin tuloksen alarajan 3.67 yläpuolelle, kun taas arkikäyttäjillä on kolme tämän rajan alittavaa puutteellista tulosta. Huono englannin kielen taito on vaikuttanut eniten arkikäyttäjillä ja MOS sekä luonnollisuus ovat molemmat puutteellisen tuloksen puolella (3.66 ja 3.61). Kolmas puutteellinen tulos on alle 35-vuotiaiden arkikäyttäjien arvio puhesynteesin luonnollisuudesta. Kaikki muut tulokset yltyvät kuitenkin hyväksytyin rajaan, joten arkikäyttäjien tuloksia voidaan pitää kokonaisuutena hyväksyttävänä, vaikkakin niissä on selvästi varaa parantaa.

On tärkeää huomata, että myös ammattikäyttäjillä nämä kaksi muuttujaryhmää: alle 35-vuotiaat ja huonon englannin kielen taitoiset ovat antaneet alhaisimmat arviot kaikista ammattikäyttäjien tuloksista. Näin voidaan päätellä, että englannin kielen taidolla oli merkitystä puhesynteesin vastaanotossa ja sen puute on vaikeuttanut puhesynteesin vastaanottoa. Samoin on huomattava, että nuorin ikäluokka on kokenut puhesynteesin heikoiten niin ammattikäyttäjissä kuin arkikäyttäjissäkin. Arkikäyttäjissä vanhin ikäluokka on kokenut puhesynteesin huonommin kuin muut, kun taas ammattikäyttäjissä vanhin yli 55-vuotiaiden ryhmä on arvioinut puhesynteesin parhaimpana. Iässä on näin ollen selvää hajontaa, ja ainoastaan ammattikäyttäjissä näyttäisi olevan selvä johdonmukaisuus puhesynteesin kokemisessa. Nuorin ikäluokka arvioi puhesynteesin huonoimpana ja vanhin ikäluokka parhaimpana.

Miesten ja naisten tuloksissa on myös selvä johdonmukaisuus molemmissa ryhmissä. Naiset arvioivat puhesynteessin paremmaksi kuin miehet. Ero on vaihteleva ja suurin eroavaisuus on ammattikäyttäjien luonnollisuuden tuloksessa, jossa miehet ja naiset eroavat arvioissaan 0.14 pisteen verran. Luonnollisuuden ja ymmärrettävyyden välillä on myös selvästi havaittava ero molemmissa ryhmissä. Ymmärrettävyys on arvioitu paremmaksi jokaisessa muuttujaryhmässä, lukuun ottamatta yli 55-vuotiaita arkikäyttäjiä, jotka arvioivat puhesynteessin luonnollisuuden hiukan puhesynteessin ymmärrettävyyttä paremmaksi. Tarkastelen seuraavaksi tarkemmin luonnollisuuden ja ymmärrettävyyden alakäsitteitä molempien ryhmien tulosten kautta.

### 5.3.1 Luonnollisuuden ja ymmärrettävyyden alakäsitteet

Luonnollisuutta mitattiin kuuntelukokemuksen, kuuntelemisen helppouden ja ihmispuheen kaltaisuuden avulla. Taulukossa 34 on listattu nämä alakäsitteet ja niiden tulokset.

**Taulukko 34.** Varianssianalyysi: luonnollisuuden alakäsitteet

AMMATTIKÄYTTÄJÄT	Kuuntelukokemus	Kuuntelemisen helppous	Ihmispuheen kaltaisuus
Kokonaistulos	4.08	3.76	3.71
Miehet	4.00	3.69	3.67
Naiset	4.20	3.85	3.76
Alle 35v	3.92	3.61	3.53
35-55v	4.11	3.79	3.72
Yli 55v	4.28	3.89	4.00
Hyvä engl. kieli	4.09	3.76	3.72
Huono engl. kieli	4.07	3.75	3.68
ARKIKÄYTTÄJÄT	Kuuntelukokemus	Kuuntelemisen helppous	Ihmispuheen kaltaisuus
Kokonaistulos	3.88	3.58	3.61
Miehet	3.87	3.55	3.62
Naiset	3.89	3.63	3.61
Alle 35v	3.87	3.54	3.51
35-55v	3.89	3.64	3.77
Yli 55v	4.04	3.58	3.58
Hyvä engl. kieli	3.90	3.60	3.66
Huono engl. kieli	3.78	3.54	3.50

Tulokset mukailevat aiemmin esitettyjä kokonaistuloksia taulukossa 33, kaikki tulokset paitsi yksi on arvioitu paremmiksi ammattikäyttäjien ryhmässä. Ainoastaan arkikäyttäjien 35-55-vuotiaiden ryhmä on arvioinut ihmispuheen kaltaisuuden paremmaksi kuin ammattikäyttäjien vastaava ryhmä. Tämä on kuitenkin ”poikkeus, joka vahvistaa säännön”.

Samat johdonmukaisuudet jatkuvat myös tässä taulukossa kuin aiemmassa taulukossa 33. Naiset arvioivat puhesynteetin paremmaksi kuin miehet, ammattikäyttäjien arviointi noudattaa ikäluokkien järjestystä ja arkikäyttäjien vanhin ja nuorin ikäluokka arvioi puhesynteetin heikoimmaksi. Ainoa poikkeus tähän on arkikäyttäjien yli 55-vuotiaiden arvio kuuntelukokemuksesta (4.04), mikä on myös korkein tulos kaikista arkikäyttäjien tuloksista.

Englannin kielen taidon tasolla on myös vaikutusta puhesynteetin vastaanotossa, erityisesti arkikäyttäjien puolella. Ammattikäyttäjien puolella ero on minimaalinen. Selvin eroavaisuus, joka luonnollisuuden alakäsitteissä nousee esiin, on kuitenkin kuuntelemisen helppouden ja ihmispuheen kaltaisuuden matalat arviot. Nämä kaksi alakäsitettä on arvioitu heikommaksi, erityisesti kun sitä vertaa kuuntelukokemuksen tuloksiin. Nämä kaksi heikommaksi arvioitua alakäsitettä sisältävät runsaasti puutteellisen tuloksen rajan alittavia arvioita, jotka on jo 5.2 luvussa nostettu esiin. Ryhmät, joissa nämä alhaiset tulokset korostuvat ovat miehet, alle 35-vuotiaat ja huonon englannin kielen omaavat vastaajat, molemmissa käyttäjäryhmissä. Sanojen ja lauseiden ymmärtäminen on luonnollisesti vaikeampaa tälle ryhmälle, ja näin ollen kuunteleminen vaikeutuu ja myös ihmispuheen kaltaisuus voi olla vaikeasti arvioitavissa, jos englannin kieli on huono. Alle 35-vuotiaiden alhaiset tulokset ovat vaikeammin perusteltavissa, tämä vaatisi tarkempaa tutkimusta vastaajaryhmistä ja mahdollisia jatkohaastatteluja, jonka avulla päästäisiin tarkemmin perille, miksi nuoremman ikäluokan suhtautuminen on kriittisempää.



**Taulukko 35.** Varianssianalyysi: ymmärrettävyyden alakäsitteet

AMMATTIKÄYTTÄJÄT	Oikeakielisuus	Informatiivisuus	Kuuntelun helppous
Kokonaistulos	3.97	3.99	4.16
Miehet	3.98	4.08	4.14
Naiset	3.95	4.02	4.18
Alle 35v	3.94	3.92	4.03
35-55v	3.97	4.00	4.16
Yli 55v	4.00	4.06	4.39
Hyvä engl. kieli	4.03	4.04	4.18
Huono engl. kieli	3.82	3.86	4.09
ARKIKÄYTTÄJÄT	Oikeakielisuus	Informatiivisuus	Kuuntelun helppous
Kokonaistulos	3.70	3.85	3.93
Miehet	3.58	3.85	3.94
Naiset	3.87	3.85	3.91
Alle 35v	3.71	3.91	3.87
35-55v	3.82	3.86	3.98
Yli 55v	3.42	3.58	4.00
Hyvä engl. kieli	3.77	3.92	3.95
Huono engl. kieli	3.52	3.70	3.89

Ymmärrettävyyden alakäsitteet on avattu varianssianalyysin kautta taulukossa 35. Sama punainen lanka kulkee ymmärrettävyydenkin alakäsitteiden tuloksissa, mutta muutamia eroavaisuuksia löytyy myös. Ammattikäyttäjät ovat arvioineet kokonaistulokset paremmin kuin arkikäyttäjät, arkikäyttäjien arviot ovat kuitenkin ymmärrettävyyden alakäsitteissä korkeammat kuin luonnollisuuden alakäsitteissä. Miehet ovat arvioineet oikeakielisyyden ja informatiivisuuden ammattikäyttäjien puolella paremmaksi kuin naiset ja arkikäyttäjien osalta kuuntelun helppous on saanut paremman arvion miehiltä kuin naisilta.

Kuuntelun helppous on saanut korkeimmat arviot molemmissa käyttäjäryhmissä, ja informaation välitys ja oikeakielisuus on myös ollut onnistunutta molemmissa ryhmissä. Ainoana poikkeuksena tähän on yli 55-vuotiaiden arkikäyttäjien ryhmä, joka on antanut selvästi poikkeavan ja alhaisen 3.58 tuloksen informatiivisuudelle ja oikeakielisyydelle. Tämä on yhteyksissä huonoon englannin kielen taitoon. Arkikäyttäjissä huonon englannin kielen valinneet vastaajat ovat arvioineet myös toiseksi heikoimmaksi

informatiivisuuden (3.70) sekä oikeakielisyyden (3.52). Kuviossa 11 kävin läpi arkikäyttäjien englannin kielen taitoja ikäluokan mukaan ja kuten taulukossa kävi ilmi, vanhimmalla ikäluokalla on eniten myös huonon englannin kielen valinneita vastaajia, joten tämä on todennäköisin syy alhaiselle tulokselle arkikäyttäjillä. Huono englannin kielen taso on vaikuttanut myös ammattikäyttäjien tuloksiin oikeakielisyydestä sekä informatiivisuudesta, ja nämä kaksi ovatkin heikoiten arvioidut käsitteet ammattikäyttäjien tuloksissa. Huonolla englannin kielen taidolla on näin ollen selvä vaikutus välitettävän informaation vastaanottamiseen, oli kyseessä ihmispuhe tai puhesynteesi.

Kokonaisuutena puhesynteesin ymmärrettävyys on hyvällä tasolla ja se on saanut paremmat arviot kuin luonnollisuus. Ymmärrettävyyden arviot olivat myös enemmän tasapainossa tuloksien perusteella, se ei sisältänyt isoja pudotuksia tuloksissa muuttujaryhmien sisällä, lukuun ottamatta englannin kielen taidon tason aiheuttamaa pudotusta tuloksissa. Mielenkiintoista olisikin, jos kyselyn olisi pystynyt toteuttamaan yhtä laadukkaalla suomen kieltä luovalla puhesynteesijärjestelmällä, miten esimerkiksi oikeakielisyyss oli arvioitu.

### **5.3.2 Mies- ja naisäänen vaikutus tuloksiin**

Puhesynteesistä esitettiin kaksi versiota ammatti- ja arkikäyttäjille. Nämä versiot olivat puhesynteesin miesäänellä ja naisäänellä luodut puheet ohjevideolle. Käyn tässä luvussa läpi eroavaisuuksia ammatti- ja arkikäyttäjien tuloksissa, mutta koska nämä käytiin jo aiemmassa luvussa läpi, keskityn tarkemmin eroavaisuuksiin miesäänen ja naisäänen versioiden tulosten välillä.

**Taulukko 36.** Varianssianalyysi: miesäänen kokonaistulokset

AMMTTIKÄYTTÄJÄT	MOS	Luonnollisuus	Ymmärrettävyys
Kokonaistulos	4.05	3.97	4.13
Miehet	4.03	3.96	4.11
Naiset	4.06	3.98	4.14
Alle 35v	3.80	3.79	3.81
35-55v	4.10	3.99	4.20
Yli 55v	4.17	4.13	4.21
Hyvä engl. kieli	4.08	3.99	4.17
Huono engl. kieli	3.95	3.91	4.00
ARKIKÄYTTÄJÄT	MOS	Luonnollisuus	Ymmärrettävyys
Kokonaistulos	3.81	3.73	3.90
Miehet	3.79	3.70	3.87
Naiset	3.85	3.77	3.94
Alle 35v	3.73	3.59	3.87
35-55v	3.92	3.86	3.99
Yli 55v	3.74	3.71	3.76
Hyvä engl. kieli	3.84	3.73	3.95
Huono engl. Kieli	3.76	3.72	3.81

Puhesynteesin miesäänen arviointien tulokset on listattu taulukossa 36. Tulokset ovat yhtä lukuun ottamatta kaikki hyväksyttävällä tasolla, ja ainoa puutteellinen tulos on arkikäyttäjien alle 35-vuotiaiden arvio luonnollisuudesta 3.59. Kokonaistulokset mukailevat aiempien lukujen tuloksia, eli arkikäyttäjien ryhmien arviointi on noin 0.20 pistettä alhaisempi kuin ammattikäyttäjien. Samankaltaisuuksia löytyy myös muista tuloksista. Englannin kielen taso vaikuttaa arviointiin, naisten tulokset ovat hieman parempia kuin miesten, ja ikäluokkien arviot ovat vaihtelevia arkikäyttäjissä, ja ammattikäyttäjissä ne ovat johdonmukaisia nuoresta vanhempaan ikäluokkaan mentäessä.

Selkeä havainto miesäänen kokonaistuloksista ovat niiden korkeat arvot. Miesääni on koettu paremmaksi tulosten mukaan molemmissa käyttäjäryhmissä ja tätä korostaa vain yksi puutteellinen tulos koko taulukossa. Käyn seuraavaksi läpi naisäänen vastaavat tulokset ja suoritan samalla vertailun puhesynteesin miesäänen tuloksiin.

**Taulukko 37.** Varianssianalyysi: naisäänen kokonaistulokset

AMMATTIKÄYTTÄJÄT	MOS	Luonnollisuus	Ymmärrettävyys
Kokonaistulos	3.84	3.74	3.95
Miehet	3.83	3.68	3.98
Naiset	3.88	3.86	3.89
Alle 35v	3.84	3.62	4.06
35-55v	3.80	3.73	3.86
Yli 55v	4.05	4.00	4.10
Hyvä engl. kieli	3.85	3.71	3.99
Huono engl. kieli	3.83	3.78	3.87
ARKIKÄYTTÄJÄT	MOS	Luonnollisuus	Ymmärrettävyys
Kokonaistulos	3.70	3.65	3.75
Miehet	3.68	3.66	3.71
Naiset	3.73	3.64	3.82
Alle 35v	3.73	3.64	3.82
35-55v	3.66	3.60	3.71
Yli 55v	3.65	3.77	3.53
Hyvä engl. kieli	3.77	3.72	3.82
Huono engl. Kieli	3.56	3.52	3.61

Taulukossa 37 on listattu puhesynteesin naisäänen arvioinnin tulokset. Puhesynteesin naisäänen tuloksiin pätee hyvin pitkälti samat eroavaisuudet arkikäyttäjien ja ammattikäyttäjien välillä kuin miesäänen tuloksissa. Suurin ero tulee kuitenkin, kun vertailee miesäänen ja naisäänen arvioinnin tuloksia, ja näissä eroavaisuuksia on paljon. Kokonaisuutena naisäänen tulokset ovat selvästi alhaisemmat kuin miesäänen tulokset. Tämä korostuu erityisesti ammattikäyttäjien MOS-kokonaistuloksessa, joka on 0.21 pistettä alhaisempi naisäänen kuin miesäänen arvioinneissa. Puutteellisten tulosten määrä on myös moninkertainen naisäänen tuloksissa verrattuna miesäänen yhteen puutteelliseen tulokseen. Esimerkkinä tästä ovat arkikäyttäjien arviot luonnollisuudesta, jossa on pelkästään kaksi hyväksyttävää tulosta, kun taas kaikki muut ovat puutteellisia tuloksiltaan naisäänen arvioissa. Poikkeus tähän tulee alle 35-vuotiaiden vastaajien tuloksissa. Arkikäyttäjien 35-vuotiaat ovat arvioineet luonnollisuuden paremmaksi naisäänen tuloksissa kuin miesäänen vastaavissa. Alle 35-vuotiaat ammattikäyttäjät ovat arvioineet myös ymmärrettävyyden sekä MOS-kokonaistuloksen paremmaksi naisäänen ryhmässä kuin miesäänen ryhmässä. Mielenkiintoista on ollut huomata, että poikkeavat arviot ovat löytyneet

useimmiten juuri tästä alle 35-vuotiaiden ryhmästä. Alle 35-vuotiaat ovat kuitenkin ainoa ryhmä, jossa tulokset ovat osittain paremmat naisäänen arvioinnin tuloksissa kuin miesäänen tuloksissa. Tärkein huomio ja eroavaisuus onkin miesäänen arvioinnin tulosten paremmuus verrattuna naisäänen tuloksiin. Puhesynteesi miesäänellä on selvästi paremmaksi arvioitu molemmissa käyttäjäryhmissä ja lähes kaikissa kategorioissa.

Käyn seuraavaksi läpi tarkemmin näitä eroavaisuuksia luonnollisuuden ja ymmärrettävyyden alakäsitteiden kautta.

**Taulukko 38.** Varianssianalyysi: miesäänen luonnollisuuden alakäsitteet

AMMATTIKÄYTTÄJÄT	Kuuntelukokemus	Kuuntelemisen helppous	Ihmispuheen kaltaisuus
Kokonaistulos	4.18	3.96	3.76
Miehet	4.15	3.94	3.79
Naiset	4.21	3.98	3.74
Alle 35v	4.00	3.79	3.57
35-55v	4.20	4.00	3.78
Yli 55v	4.38	4.00	4.00
Hyvä engl. kieli	4.21	3.95	3.81
Huono engl. kieli	4.11	4.00	3.61
ARKIKÄYTTÄJÄT	Kuuntelukokemus	Kuuntelemisen helppous	Ihmispuheen kaltaisuus
Kokonaistulos	3.93	3.56	3.69
Miehet	3.86	3.52	3.71
Naiset	4.04	3.62	3.65
Alle 35v	3.81	3.35	3.62
35-55v	3.96	3.75	3.86
Yli 55v	4.07	3.57	3.50
Hyvä engl. kieli	3.91	3.55	3.73
Huono engl. kieli	3.96	3.58	3.63

Luonnollisuuden alakäsitteitä puhesynteesin miesäänen tuloksissa on avattu Taulukossa 38. Ammattikäyttäjien tulokset ovat pääosin hyvällä tasolla, ja mukana on vain kaksi puutteellista tulosta ihmispuheen kaltaisuudesta, jotka tulevat huonon englannin kielen omaavilta vastaajilta sekä alle 35-vuotiaiden vastaajaryhmältä. Arkikäyttäjissä

puutteellisia tuloksia on lukuisa määrä. Nämä keskittyvät kuuntelemisen helppouden sekä ihmispuheen kaltaisuuden alakäsitteisiin. Erityisesti kuuntelemisen helppous on saanut alhaisia arvioita, varsinkin verrattuna ammattikäyttäjien arvioihin vastaavista alakäsitteistä.

Mielenkiintoinen havainto on kuuntelemisen helppouden alakäsitteessä ja englannin kielen tason vastaajissa. Huonon englannin kielen valinneet ovat arvioineet kuuntelemisen helppouden paremmaksi kuin hyvän englannin valinneet vastaajat. Kuuntelemisen helppous keskittyykin enemmän puhesynteesin sävyyn ja miellyttävyyteen, eikä niinkään kieliopillisiin seikkoihin kuten ymmärrettävyyden kuuntelun helppous, joten tässä mielessä eroavaisuus on looginen.

**Taulukko 39.** Varianssianalyysi: naisäänen luonnollisuuden alakäsitteet

AMMATTIKÄYTTÄJÄT	Kuuntelukokemus	Kuuntelemisen helppous	Ihmispuheen kaltaisuus
Kokonaistulos	3.99	3.56	3.65
Miehet	3.91	3.54	3.59
Naiset	4.17	3.63	3.79
Alle 35v	3.86	3.50	3.50
35-55v	4.00	3.54	3.65
Yli 55v	4.20	3.80	4.00
Hyvä engl. kieli	4.15	3.56	3.62
Huono engl. kieli	4.24	3.84	3.89
ARKIKÄYTTÄJÄT	Kuuntelukokemus	Kuuntelemisen helppous	Ihmispuheen kaltaisuus
Kokonaistulos	3.83	3.60	3.53
Miehet	3.88	3.57	3.52
Naiset	3.75	3.64	3.54
Alle 35v	3.82	3.66	3.45
35-55v	3.75	3.44	3.63
Yli 55v	4.00	3.60	3.70
Hyvä engl. kieli	3.90	3.65	3.60
Huono engl. kieli	3.68	3.50	3.36

Puhesynteesin naisäänen arvioinnin tulokset on listattu taulukossa 39. Tärkein huomio naisäänen arvioinnin luonnollisuuden alakäsitteissä on arvioiden matala taso. Ammattikäyttäjienkin arviot on mielletty puutteellisina kuuntelukokemusta lukuun ottamatta, tämä toistuu myös arkikäyttäjissä. Mielenkiintoinen huomio on myös, että kuuntelemisen helppous on korkeampi naisäänen arvioinnin tuloksissa verrattuna miesäänen arvioinnin tuloksiin, muut arvot ovat laskeneet ja jotkin hyvinkin paljon kuten ammattikäyttäjien arvio kuuntelemisen helppoudesta, joka on pudonnut 0.40 pistettä verrattuna miesäänen arvioinnin tuloksiin. Kuuntelemisen helppous on pudonnut myös kaikissa muuttujaryhmissä, mutta pienin pudotus on tullut huonon englannin kielen omaavien vastaajien tuloksissa. Ihmispuheen kaltaisuus on parempi huonon englannin kielen omaavien vastaajien arvioissa naisäänestä verrattuna miesäänen vastaaviin arviointiin. Englannin kielen taito ei ole, kuten aiemmin totesin, yhtä tärkeä puheen luonnollisuutta arvioitaessa, joten tämä on varmasti ainakin osasyynä tähän.

Tärkein huomio kuitenkin on edelleen se, että puhesynteesin miesääni mielletään näinkin paljon parempana kuin puhesynteesin naisääni. Tämä on ollut johdonmukainen teema kaikissa tähänastisissa tuloksissa. Esittelen seuraavaksi vielä ymmärrettävyyden alakäsitteet mies- ja naisäänen tuloksien osalta.

**Taulukko 40.** Varianssianalyysi: miesäänen ymmärrettävyyden alakäsitteet

AMMATTIKÄYTTÄJÄT	Oikeakielisuus	Informatiivisuus	Kuuntelun helppous
Kokonaistulos	3.99	4.12	4.28
Miehet	3.94	4.12	4.26
Naiset	4.02	4.12	4.29
Alle 35v	3.79	3.79	3.86
35-55v	4.06	4.19	4.35
Yli 55v	3.88	4.25	4.50
Hyvä engl. kieli	4.07	4.14	4.29
Huono engl. kieli	3.72	4.06	4.22
ARKIKÄYTTÄJÄT	Oikeakielisuus	Informatiivisuus	Kuuntelun helppous
Kokonaistulos	3.72	3.93	4.04
Miehet	3.67	3.95	4.00
Naiset	3.81	3.88	4.12
Alle 35v	3.65	3.96	4.00
35-55v	3.89	4.00	4.07

Yli 55v	3.50	3.71	4.07
Hyvä engl. kieli	3.80	4.00	4.05
Huono engl. kieli	3.58	3.79	4.04

Miesäänen ymmärrettävyyden arvioinnin tulokset on listattu taulukossa 40. Ammattikäyttäjien osalta tulokset ovat hyvät ja kuuntelun helppous on saanut erinomaiset tulokset kaikilta paitsi alle 35-vuotiaiden muuttujaryhmältä, joka on arvioinut kaikki alakäsitteet alhaisimmin kaikista muista ammattikäyttäjien muuttujaryhmistä. Arkikäyttäjien alle 35-vuotiaat ovat tosin arvioineet informatiivisuuden ja kuuntelun helppouden selvästi ammattikäyttäjien alle 35-vuotiaiden ryhmää paremmiksi. Tässä korostuu myös alle 35-vuotiaiden arvostelun vaihtelevuus ja erikoisuus, joka nousee esiin tässäkin taulukossa.

Arkikäyttäjien arvioinnin tulokset ovat alhaisemmat kuin ammattikäyttäjien, mutta arkikäyttäjien tulokset ovat kuitenkin hyväksyttävällä tasolla. Arkikäyttäjien tuloksista löytyy vain kolme puutteellista tulosta ja nämä ovat alle 35-vuotiaiden, huonon englannin taidon omaavien ja yli 55-vuotiaiden arviot oikeakielisyydestä. Oikeakielisyyden on myös alakäsitteistä heikoimmaksi arvioitu molemmissa käyttäjäryhmissä. Puhesynteesin miesäänen arvioinnin tulokset ovat myös ymmärrettävyyden kannalta varsin positiiviset. Seuraavaksi vertaan naisäänen tuloksia ymmärrettävyydestä miesäänen arvioinnin tuloksiin.

**Taulukko 41.** Varianssianalyysi: naisäänen ymmärrettävyyden alakäsitteet

AMMATTIKÄYTTÄJÄT	Oikeakielisuus	Informatiivisuus	Kuuntelun helppous
Kokonaistulos	3.95	3.86	4.04
Miehet	4.00	3.87	4.06
Naiset	3.83	3.83	4.00
Alle 35v	4.05	4.00	4.14
35-55v	3.87	3.78	3.93
Yli 55v	4.10	3.90	4.30
Hyvä engl. kieli	3.98	3.92	4.06
Huono engl. kieli	3.88	3.73	4.00
ARKIKÄYTTÄJÄT	Oikeakielisuus	Informatiivisuus	Kuuntelun helppous
Kokonaistulos	3.67	3.77	3.81
Miehet	3.50	3.74	3.88
Naiset	3.93	3.82	3.71
Alle 35v	3.75	3.91	3.80



35-55v	3.69	3.63	3.81
Yli 55v	3.30	3.40	3.90
Hyvä engl. kieli	3.75	3.85	3.85
Huono engl. kieli	3.50	3.59	3.73

Puhesynteesin naisäänen arvioinnin tulokset ymmärrettävyyden alakäsitteistä on listattu Taulukkoon 41. Ammattikäyttäjien ja arkikäyttäjien tulokset eroavat siinä, että ammattikäyttäjien tulokset ovat selvästi korkeammat oikeakielisyydessä ja kuuntelun helppouudessa ja jokseenkin korkeammat informatiivisuudessa. Selkein yksittäinen eroavaisuus muuttujaryhmistä löytyy arkikäyttäjien yli 55-vuotiaiden oikeakielisyyden tuloksesta 3.30 verrattuna ammattikäyttäjien vastaavan ryhmän vastaavaan tulokseen 4.10. Päinvastainen ja hieman pienempi eroavaisuus on naisten arvioinneissa oikeakielisyydestä. Arkikäyttäjien naisten tulos on jonkin verran korkeampi kuin ammattikäyttäjien vastaava tulos. Englannin kielen merkitys korostuu molemmissa käyttäjäryhmissä, hyvän englannin kielen omaavat vastaajat ovat arvioineet kaikki kolme ymmärrettävyyden alakäsitettä huonon englannin kielen taidon omaavia korkeammin. Tässä on yksi selkeä ja jatkuva eroavaisuus luonnollisuuden ja ymmärrettävyyden alakäsitteiden tuloksissa.

Miesäänen arvioinnin tuloksiin verrattuna naisäänen tulokset ovat matalammalla tasolla merkittävästi informatiivisuuden ja kuuntelun helppouden alakäsitteissä ja jonkin verran myös oikeakielisyyden alakäsitteen tuloksissa, molemmissa käyttäjäryhmissä. Oikeakielisyyden heikompi arviointi miesäänen tuloksista johtuu alle 35-vuotiaiden ja yli 55-vuotiaiden ammattikäyttäjien naisäänen arvioista, jotka he arvioivat paremmaksi kuin ammattikäyttäjien vastaavat arviot. Ammattikäyttäjien alle 35-vuotiaiden ryhmä arvioi myös naisäänen informatiivisuuden sekä kuuntelun helppouden selvästi paremmin kuin vastaava arkikäyttäjien ryhmä. Nämä ovat myös selvästi poikkeavat tulokset verrattuna muihin muuttujaryhmiin, joissa aiemmin esitettyjä muuttujaryhmiä lukuun ottamatta kaikissa tulos on ollut heikompi verrattaessa miesäänen tuloksia naisäänen tuloksiin.

Ymmärrettävyyden tulokset todistivat sen, että puhesynteesin miesääni sai selvästi naisääntä paremmat arviot ammatti- ja arkikäyttäjien toimesta. Tämä nousi esiin, jokaisessa taulukossa ja vertailussa, muutamia yksittäisiä poikkeuksia lukuun ottamatta.

Puhesynteesin miesääni todettiin kokonaistuloksiltaan ja erityisesti informatiivisuudessa ja kuuntelemisen helppoudessa naisääntä paremmaksi. Esittelen seuraavassa luvussa analyysin tulosten pohjalta tehdyt pohdinnat ja johtopäätökset tuleville askelille.

## 6 Pohdinta

Tutkielman tavoitteena oli mitata ammatti- ja arkikäyttäjien kokemuksia puhesynteestistä teknisellä ohjevideolla. Näitä kokemuksia mitattiin luonnollisuuden ja ymmärrettävyyden pääkäsitteillä sekä näitä tarkentavilla alakäsitteillä. Tavoitteena oli löytää vastauksia myös siihen, mitä eroavaisuuksia ammatti- ja arkikäyttäjien tulosten välillä oli ja siihen, miten muuttajat kuten ikä, sukupuoli, englannin kielen taito tai puhesynteesin mies- tai naisääni vaikuttivat puhesynteesin kokemiseen. Kokemusten perusteella pyrittiin saamaan selville, onko puhesynteesin käyttöönotto ohjevideoille mahdollista.

Tavoitteeseen pyrittiin pääsemään kyselytutkimuksen avulla, jolla ammatti- ja arkikäyttäjien ryhmät arvioivat puhesynteesiä. Kyselyssä vastaajia pyydettiin katsomaan puhesynteesiä sisältävä ohjevideo mies- tai naisäänellä, jonka jälkeen heitä pyydettiin arvioimaan kokemuksensa puhesynteesistä kahdentoista kysymyksen avulla. Näistä tuloksista laskettiin keskiarvoinen tulos kaikille muuttujille, luonnollisuuden ja ymmärrettävyyden pää- ja alakäsitteille sekä erikseen puhesynteesin mies- ja naisäänelle. Tulokset listattiin kuvioilla ja taulukoilla ja niistä suoritettiin varianssianalyysi, jolla pyrittiin löytämään samankaltaisuuksia ja eroavaisuuksia muuttujien ja käsitteiden välillä.

Puhesynteesin luonnollisuus ja ymmärrettävyys koettiin pääosin positiivisesti. Ymmärrettävyys oli selkeästi paremmaksi koettu näistä kahdesta, ja se sai erinomaisia tuloksia informatiivisuudesta sekä kuuntelun helppoudesta. Puhesynteesin luonnollisuus koettiin hyväksyttävänä, mutta sen tuloksissa oli selvästi enemmän vaihtelua ja kritiikkiä. Erityisesti puhesynteesin ihmispuheen kaltaisuus ja oikeakielisyys saivat heikompia arvioita.

Ammattikäyttäjät arvioivat puhesynteesin kokonaisvaltaisesti paremmaksi kuin arkikäyttäjät. Erityisesti informatiivisuus, kuuntelukokemuksen ja kuuntelemisen helppouden käsitteet oli arvioitu paremmaksi ammattikäyttäjien arvioinneissa. Tähän suurimpana syynä on ammattikäyttäjien alan ja sen termistön tuntemus, mikä helpottaa huomattavasti informaation vastaanottoa ja kuuntelemista. Muita selkeitä ja johdonmukaisia eroja ei löytynyt ammattikäyttäjien ja arkikäyttäjien väliltä.

Naiset arvioivat puhesynteesin paremmaksi kuin miehet molemmissa käyttäjäryhmissä. Ammattikäyttäjissä myös vanhemmat ikäluokat kokivat puhesynteesin parempana kuin nuorin ikäluokka. Ammattikäyttäjissä nuorin ikäluokka oli kriittisin arvioinneissaan. Todennäköinen syy tähän kriittisyyteen on nuorten kattava mediakokemus. Alle 35-vuotiaiden ikäluokka on ollut kosketuksessa median ja internetin kanssa käytännössä koko elämänsä ajan ja näin heille on kehittynyt parempi, ja kriittisempi medialukutaito. Arkikäyttäjissä ikäluokkien tulokset olivat hyvin vaihtelevat, ja johdonmukaista yhtäläisyyttä ei löytynyt tuloksista.

Englannin kielen taito vaikutti ymmärrettävyyteen huomattavasti molemmissa käyttäjäryhmissä. Hyvä englannin kielen taito helpotti ymmärtämistä, kun taas huono englannin kielen taito vaikeutti sitä. Puhesynteesin luonnollisuudessa englannin kielen taidolla ei ollut vastaavanlaista merkitystä. Nuorin ikäluokka arvioi englannin kielen taitonsa parhaimmaksi ja vanhin ikäluokka heikoimmaksi.

Puhesynteesin äänellä oli merkittävä ero mies- ja naisäänien välillä. Puhesynteesin miesääni koettiin parempana valtaosissa muuttujia sekä luonnollisuuden että ymmärrettävyyden käsitteissä. Miesääni todettiin erityisesti informaation välitykseltään paremmaksi, kuuntelukokemus oli myös mieluisampi vastaajien mielestä ja sitä oli molempien käyttäjäryhmien vastausten perusteella myös helpompi kuunnella.

Aiemman Deep learning-puhesynteesijärjestelmän tutkimuksen perusteella (ks. Wavenet: A Generative Model for Raw Audio 2016) ammattikäyttäjien tulokset ovat hyvällä tasolla, mutta arkikäyttäjien tulokset jäävät selvästi aiemman tutkimuksen tulosten perään. Tutkielman tulosten perusteella voidaan sanoa, että ammattikäyttäjille puhesynteesin miesääni on varteenotettava vaihtoehto. Arkikäyttäjille puhesynteesi vaikuttaa olevan vielä vaikeasti lähestyttävä.

Tämän tutkielman perusteella suosittelen puhesynteesin käyttöönottoa teknisiin ohjeideoihin ammattikäyttäjille. Arkikäyttäjille suunnattu käyttöönotto on mahdollista, mutta

tarkentava laadullinen jatkotutkimus siitä, miksi sen vastaanotto on vaikeampaa, olisi aiheellinen ennen käyttöönottoa. Jatkotutkimuksessa voisi keskittyä siihen, mitä tunteita ja ajatuksia puhesynteesi tarkalleen herättää vastaanottajassa eri konteksteissa. Vaikuttaako esimerkiksi alusta tai väline, jossa puhesynteesiä käytetään sen vastaanottamiseen. Jatkotutkimus voitaisiin suorittaa haastatteluilla, joilla saataisiin tarkempaa tietoa vastaanottajien tuntemuksista ja ajatuksista puhesynteesin käytöstä. Tutkimukseni tai vastaavanlaisen tutkimuksen suorittaminen suomenkielisellä puhesynteesillä, on myös yksi vartenotettava jatkotutkimuskohde.

## Lähteet

- Aaltonen, J. (2018). *Käsikirjoittajan työkalut. Audiovisuaalisen käsikirjoituksen tekijän opas*. Tampere: Suomen Yliopistopaino Oy.
- Akilandeswari, V., Dinesh Kumar A., Philomin Freeda A., & Niranchan Kumar S. (2015). Elements of Effective Communication. *New Media and Mass Communication*, 37, 44-45.
- Albers, M. J. (2012). *Human-information Interaction and Technical Communication: Concepts and Frameworks*. Pennsylvania: Hershey, PA: IGI-Global.
- Al-Fedaghi, S. (2012). A Conceptual Foundation for the Shannon-Weaver Model of Communication. *International Journal of Soft Computing*, 7 (1), 12-19. DOI: 10.3923/ijscmp.2012.12.19
- Amazon Web Services (2019). *Amazon Polly – Turn text into lifelike speech using deep learning*. Noudettu 2019-04-25 osoitteesta <https://aws.amazon.com/polly/>
- Baccarani, C. & Bonfanti, A. (2015). Effective public speaking: a conceptual framework in the corporate-communication field. *Corporate Communications: An International Journal*, 20 (3), 375-390. <https://doi.org/10.1108/CCIJ-04-2014-0025>
- Bacon, H. (2000). *Audiovisuaalisen kerronnan teoria*. Tampere: Tammer-Paino.
- Bhavik, D., Siddhi D. & Verghese, J. M. (2017). Survey on Methods of Text to Speech Synthesis. *International Journal of Computer Applications*, 165 (6), 26-30. DOI:10.5120/ijca2017913891
- Bunnell, H. T., Jreige, C. & Patel, R. (2009). VocaliD: Personalizing text-to-speech synthesis

for individuals with severe speech impairment. *ASSETS'09 - Proceedings of the 11th International ACM SIGACCESS. Esitetty Conference on Computers and Accessibility*. Pittsburgh: ACM New York, 259-260. DOI: 10.1145/1639642.1639704

Chung, Y., Wang, Y., Hsu, W., Zhang, Y. & Skerry, R. (2018). Semi-Supervised Training for Improving Data Efficiency in End-to-End Speech Synthesis. *Cornell University: arXiv preprint arXiv:1808.10128*, 1-5. DOI: <https://arxiv.org/pdf/1808.10128.pdf>

Evans, M. & Walton, T. (2018). The role of human influence factors on overall listening experience. *Quality and User Experience* 3(1). <https://doi.org/10.1007/s41233-017-0015-4>

De, P. R. (2016). The arithmetic mean - Geometric mean - Harmonic mean: Inequalities and a spectrum of applications. *Resonance*, 21 (12), 1119-1133. <https://doi.org/10.1007/s12045-016-0423-4>

Deng, L. & Yu, D. (2012). Deep Learning: Methods and Applications. *Now Publishers*, 7(3-4). <http://dx.doi.org/10.1561/20000000039>

Deng, L. & Dong, Y. (2014). Deep Learning: Methods and Applications. *Foundations and Trends in Signal Processing*, 7(3-4). DOI: 10.1561/20000000039

Dieleman, S., Van den Oord, A., Graves, A., Senior, A., Zen, H., Simonyan, K., Karukguaciu, K., Kalchbrenner, N. & Vinyals, Oriol (2016). *Wavenet: A Generative Model for Raw Audio*. London: Google DeepMind. DOI: <https://arxiv.org/pdf/1609.03499.pdf>

Fiske, J. (1990). *Introduction to Communication Studies*. Studies in Culture and Communication. (2<sup>nd</sup> edition). London: Routledge.

Flanagan, J. (1972). Speech Synthesis. Teoksessa Flanagan, J. (Toim.), *Speech Analysis*,

*Synthesis, and Perception. Kommunikation und Kybernetik in Einzeldarstellungen, 3,* (204-276). Berlin-Heidelberg-New York: Springer-Verlag.

[https://doi.org/10.1007/978-3-662-01562-9\\_6](https://doi.org/10.1007/978-3-662-01562-9_6)

Hande, S. S. (2014). A Review on Speech Synthesis and Artificial Voice Production. *International Journal of Advanced Research in Electrical, Electronic and Instrumentation Engineering*, 3 (3), 2278-8875. [http://www.ijareeie.com/upload/2014/march/25\\_AReview.pdf](http://www.ijareeie.com/upload/2014/march/25_AReview.pdf)

Heikkilä, T. (2014). *Tilastollinen tutkimus*. Helsinki: Edita Publishing Oy.

Herre, J. & Schoeffler, M. (2014). Towards a Listener Model for Predicting the Overall Listening Experience. *Proceedings of the 9th Audio Mostly: A Conference on Interaction With Sound*, 11. DOI: 10.1145/2636879.2636892

Huang, G. & Lu, Y. (2014). Application of DSP Speech Synthesis System on Service Robots. *Institute of Electrical and Electronics Engineer. Esitetty CACS International Automatic Control Conference (CACS 2014)*. Taiwan.

Hundleby, M. & Allen, J. (2010). *Assessment in Technical and Professional Communication*. Teoksessa *Baywood's Technical Communications Series*. Routledge: Baywood Publishing Company, Inc.

International Telecommunication Union (1994). *ITU-T Recommendation P.85: A method for subjective performance assessment of the quality of speech voice output devices*. Switzerland: International Telecommunication Union. <https://www.itu.int/rec/T-REC-P.85-199406-I/en>

JMC Study Hub (2019). Shannon and Weaver Model of Communication. *JMC Study Hub*:



*Hub of Journalism and Mass Communication*. Noudettu 2019-04-04 <https://jmc-studyhub.com/shannon-and-weaver-model-of-communication-mass-communication/>

Kananen, J. (2014). *Verkkotutkimus opinnäytetyönä. Laadullisen ja määrällisen verkko tutkimuksen opas*. Jyväskylä: Juvenes Print.

Le Maguer, S. (2018). *Speech syntesis evaluation*. Germany: Saarland university.

Noudettu 2019-10-22

[http://www.coli.uni-saarland.de/~slemaguer/teaching/2017-2018/tts\\_evaluation/tts\\_evaluation.pdf](http://www.coli.uni-saarland.de/~slemaguer/teaching/2017-2018/tts_evaluation/tts_evaluation.pdf)

Lehtonen, M. (2007). Ruumis, kieli ja toiminta - Ajatuksia audiovisuaalisten tekstien multimodaalisuudesta. Teoksessa Oittinen, Riitta & Tuominen, Tiina (Toim.) *Olellaisen äärellä. Johdatus audiovisuaaliseen kääntämiseen*. Tampere: Tampereen yliopistopaino Oy.

Mada, S. & Saftoiu, R. (2012). Professional Communication across Languages and Cultures. *Dialogue Studies*, 17 (2012). <https://doi.org/10.1075/ds.17>

Maijala, P., Nordlund, K. & Virolainen, K. (1987). *Kodinkoneiden käyttöohjeet ja varoitus merkinnät*. Helsinki: Elinkeinohallitus.

Medeiros, J. (2018). *Wired UK: How Intel Gave Stephen Hawking a voice*. Noudettu 2019-10-10

<https://www.wired.com/2015/01/intel-gave-stephen-hawking-voice/>

Mehler, A., Romary, L. & Gibbon, D. (2012). *Handbook of Technical Communication*.

France: De Gruyter Mouton. <https://hal.inria.fr/hal-00786386>

Mouton.

Oras Ltd. (2019). Noudettu 2019-06-25 Saatavilla: <https://www.oras.fi>

Pöyhönen, M. & Tiusanen, A. (1991). *Käyttöohje on osa tuotetta. Käyttöohjeen laatijan opas*. Helsinki: Pohjola-yhtiöt.

Rui, R. & Zhenjiang, M. (2008). *Emotional speech synthesis and its application to pervasive E-learning*. Esitetty Ubi-Media Computing, 2008 First IEEE International Conference on Ubi-Media Computing. DOI: <http://dx.doi.org/10.1109/UMEDIA.2008.4570930>

Schnurr, S. (2013). *Exploring Professional Communication: Language in Action*. Milton Park Abingdon, Oxon; New York: Routledge.

Streijl, R. C., Winkler, S. & Hands, D. S. (2016). Mean opinion score (MOS) revisited: methods and applications, limitations and alternatives. *Multimedia Systems*, 22 (2), 213-227. DOI:10.1007/s00530-014-0446-1

Suomen teknisen viestinnän yhdistys (2019). *Mitä on tekninen viestintä?*  
Noudettu 2019-09-20 <https://www.stvy.fi/tekninen-viestinta/>

Taylor, P. (2009). *Text-to-speech-synthesis*. Cambridge: Cambridge University Press.

Tung, L. (2018). Google's human-sounding AI to answer calls at contact centers. *ZDNet*. Noudettu 2019-10-05 <https://www.zdnet.com/article/googles-human-sounding-ai-to-answer-calls-at-contact-centers/>

Tähtinen, J., Laakkonen, E. & Broberg, M. (2011). Tilastollisen aineiston käsittelyn ja tulokinnan perusteita. Teoksessa *Turun yliopiston kasvatustieteiden tiedekunnan julkaisuja C:20*. Turku: Painosalama Oy.

Valli, R. (2015). *Johdatus tilastolliseen tutkimukseen*. Jyväskylä: PS-kustannus.

Vilka, H. (2007). *Tutki ja mittaa. Määrällisen tutkimuksen perusteet*. Helsinki: Kustannusosakeyhtiö Tammi.

Viswanathan, M. & Viswanathan, M. (2005). Measuring speech quality for text-to-speech systems: development and assessment of a modified mean opinion score (MOS) scale. *Computer Speech & Language*, 19 (1), 55-38. DOI: 10.1016/j.csl.2003.12.001

Vehkalahti, K. (2014). *Kyselytutkimuksen mittarit ja menetelmät*. Helsinki: Finn Lectura Ab.

## **Liitteet**

### **Liite 1. Kyselylomake : Arkikäyttäjät**

*Sukupuoli \**

*Mies*

*Nainen*

*Ikä \**

*\_\_ vuotta*

*Englannin kielen taito \**

*Huono*

*Välttävä*

*Tyydyttävä*

*Kiitettävä*

*Erinomainen*

*Onko sinulla aiempaa kokemusta hanojen asentamisesta tai huoltamisesta? \**

*kyllä » jatka sivulle päätä kysely*

*ei*

*Katso video äänen kanssa ja vastaa alla oleviin kysymyksiin puhesynteesistä.*

*1. Miten luokittelisit puhesynteesin äänen miellyttävyyden? \**

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

*2. Minkä arvoasteikon valitsisit puhesynteesin kieliopillisuudesta? \**

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

*3. Miten arvioisit puhesynteesin kuuntelun helppoutta? \**

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

*4. Miten arvioisit puhesynteesin äänen sointuvuutta ihmispuheeseen verrattuna? \**

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

*5. Minkä arvosanan antaisit puhesynteesin äänen ohjeistavuudelle? \**

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

*6. Miten arvioisit puhesynteesin kuuntelukokemuksena? \**

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

*7. Minkä arvosanan antaisit puhesynteesin äänen informatiivisuudelle? \**

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

8. Miten arvioisit puhesynteesin äänen ihmispuheeseen verrattuna? \*

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

9. Minkä arvosanan antaisit puhesynteesin ja huoltovideon yhteensopivuudelle? \*

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

10. Miten arvioisit puhesynteesin puheen rytmityksen? \*

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

11. Miten arvioit puhesynteesin äänenlaatua ja sen selkeyttä? \*

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*

12. Arvioi puhesynteesin puhumisen lauserakennetta ja oikeiden sanojen painotusta? \*

*Erittäin huono*

*Huono*

*Kelvollinen*

*Hyvä*

*Erinomainen*



## Liite 2. Kyselylomake: Ammattikäyttäjät

Täytä alla olevat tiedot.

### 1. Sukupuoli

Mies

Nainen

### 2. Ikä

Enter your answer

### 3. Englannin kielen taito

Huono

Välttävä

Tyydyttävä

Kiitettävä

Erinomainen

Katso video äänen kanssa ja vastaa alla oleviin kysymyksiin puhesynteesistä.

### 4. Miten luokittelisit puhesynteesin äänen miellyttävyyden?

Erittäin huono      Huono    Kelvollinen    Hyvä    Erinomainen

### 5. Minkä arvoasteikon valitsisit puhesynteesin kieliopillisuudesta?

Erittäin huono      Huono    Kelvollinen    Hyvä    Erinomainen

### 6. Miten arvioisit puhesynteesin kuuntelun helppoutta?

Erittäin huono      Huono    Kelvollinen    Hyvä    Erinomainen

### 7. Miten arvioisit puhesynteesin äänen sointuvuutta ihmispuheeseen verrattuna?

Erittäin huono      Huono    Kelvollinen    Hyvä    Erinomainen

### 8. Minkä arvosanan antaisit puhesynteesin äänen ohjeistavuudelle?

Erittäin huono      Huono    Kelvollinen    Hyvä    Erinomainen

### 9. Miten arvioisit puhesynteesin kuuntelukokemuksena?

Erittäin huono      Huono    Kelvollinen    Hyvä    Erinomainen

### 10. Minkä arvosanan antaisit puhesynteesin äänen informatiivisuudelle?

Erittäin huono      Huono    Kelvollinen    Hyvä    Erinomainen

11. Miten arvioisit puhesynteesin äänen ihmispuheeseen verrattuna?

*Erittäin huono      Huono   Kelvollinen   Hyvä   Erinomainen*

12. Minkä arvosanan antaisit puhesynteesin ja huoltovideon yhteensopivuudelle?

*Erittäin huono      Huono   Kelvollinen   Hyvä   Erinomainen*

13. Miten arvioisit puhesynteesin puheen rytmityksen?

*Erittäin huono      Huono   Kelvollinen   Hyvä   Erinomainen*

14. Miten arvioit puhesynteesin äänenlaatua ja sen selkeyttä?

*Erittäin huono      Huono   Kelvollinen   Hyvä   Erinomainen*

15. Arvioi puhesynteesin puhumisen lauserakennetta ja oikeiden sanojen painotusta?

*Erittäin huono      Huono   Kelvollinen   Hyvä   Erinomainen*