

# Koodinvaihto viiden videosisällöntuottajan puhutussa kielessä

Julia Hyytiäinen

Pro gradu -tutkielma

Kieli- ja käännöstieteiden laitos

Kielen oppimisen ja opettamisen tutkinto-ohjelma, suomen kieli

Turun yliopisto

Humanistinen tiedekunta

Huhtikuu 2021

*Turun yliopiston laatujärjestelmän mukaisesti tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck -järjestelmällä.*

TURUN YLIOPISTO

Kieli- ja käännöstieteiden laitos / Humanistinen tiedekunta

HYTYÄINEN, JULIA: Koodinvaihto viiden videosisällöntuottajan puhutussa kielessä

Pro gradu -tutkielma, 55s.

Kielen oppimisen ja opettamisen tutkinto-ohjelma, suomen kieli

Huhtikuu 2021

---

Tässä tutkielmassa tarkastellaan viiden sosiaalisen median videosisällöntuottajan puhekielen koodinvaihtoa. Tavoitteena on selvittää, millaisissa tilanteissa tai asiayhteyksissä koodinvaihtoa esiintyy sekä kartoittaa koodinvaihdon mahdollista lisääntymistä sisällöntuottajien puheessa tietyllä aikavälillä.

Prototyypisimmillään koodinvaihto tarkoittaa kahden tai useamman kielen vaihtelua samassa keskustelussa. Vuorovaikutuksen näkökulmasta kielten kohtaaminen kytkeytyy koodinvaihtoon ja laajemmin monikielisyyden tulkintaan kontekstissa.

Käyttäjäkohtaista variaatiota tutkittaessa yksilötasolla puhujat mukauttavat puhetapaansa niiden puhetapojen suuntaan, joiden kanssa yksilö on säännöllisesti tekemisissä. Tällöin yhteiskunnassa toimivasta ryhmästä voi havaita sille tunnistettavia kielellisiä piirteitä.

Sosiolingvistiikka keskittyy kuvaamaan kieltä sosiaalisena ilmiönä ja yrittää etsiä yhteyksiä kielenkäytön ja yhteiskunnallisen todellisuuden väliltä. Nykyaikaisen vuorovaikutussosiolingvistiikan lähtökohtana voidaan pitää metaforisen koodinvaihdon käsitteen perusidea. Metaforinen koodinvaihto tarkoittaa sitä, että saman tilanteen sisällä eri koodeilla herätetään assosiaatioita erilaisiin konteksteihin, joiden näkökulmasta sanottua tulkitaan. Keskusteluanalyysin metodia voidaan painottaa lingvistisesti, jolloin se lukeutuu vuorovaikutuslingvistiikan alaan.

Tutkittavat sisällöntuottajien videot on valittu kvalitatiivisesti ja koodinvaihtokohdat litteroitu aineistoksi. Videot on julkaistu Youtube-alustalla sisällöntuottajien omilla kanavilla vuosina 2013–2021. Aineistoa on tarkasteltu sisällönanalyysin avulla.

Tutkimustuloksista havaitaan jokaisen sisällöntuottajan puheessa ilmenevän koodinvaihtoa metaforisen koodinvaihdon funktioiden mukaan. Koodinvaihtokategorioista eniten sisällöntuottajien puheessa esiintyy yhden tai kahden sanan mittaisia koodinsekoituksia. Lisäksi huomataan, että viidestä sisällöntuottajasta neljällä koodinvaihtotapausten määrä on vähintään kaksinkertaistunut tarkastellulla aikavälillä. Koodia vaihdetaan silloin, kun puheena olevaa asiaa tai sen ominaisuutta halutaan korostaa upotetulla kielellä. Kaksi kieltä toimii rinnakkain sisällöntuottajien puheessa siten, että koodinvaihto näyttäytyy monikielisen yhteisön kommunikaatiomuotona.

Englannin asema *lingua francana* kasvaa jatkuvasti. Tulevaisuudessa jatkotutkimus voisi kohdistua vielä suurempaan aineistoon, josta variaatiota voitaisiin esittää tarkemmin muun muassa alueellisin perustein. Lisäksi voitaisiin kartoittaa, onko koodinvaihto puhekielessä edelleen lisääntyvää, vai voidaanko sitä pitää väliaikaisen internetkielen synnyttämänä ilmiönä.

Asiasanat: *koodinvaihto, monikielisuus, puhekielen variaatio, sosiolingvistiikka*

# Sisällysluettelo

1	Johdanto .....	2
2	Teoriaa ja keskeisiä käsitteitä .....	5
	2.1 Sosiolingvistinen tutkimus .....	5
	2.2 Vuorovaikutussosiolingvistiikka.....	6
	2.3 Keskustelunanalyysi.....	8
	2.4 Puhekielen variaatio .....	10
	2.5 Koodinvaihto ja sen ulottuvuudet .....	11
3	Aineisto .....	14
	3.1 Aineiston lähtökohdat .....	14
	3.2 Videosisällöntuottajat.....	16
4	Tutkimusmenetelmät .....	19
	4.1 Kvalitatiivinen ja kvantitatiivinen tutkimus.....	19
	4.2 Yksilöllisen variaation tarkastelu.....	21
	4.3 Koodinvaihtotapausten tarkastelu .....	23
	4.4 Monikielisyys ja koodinvaihto.....	25
	4.5 Englannin kielen asema ja määrittely .....	28
5	Koodinvaihto puhutussa kielessä .....	31
	5.1 Yksilöllinen variaatio ja koodinvaihtofunktiot .....	31
	5.2 Yksilöllinen variaatio ja koodinvaihtotyypit.....	36
	5.3 Koodinvaihto ja kielenulkoiset tekijät .....	41
	5.4 Kokoavasti koodinvaihtotapauksista.....	46
6	Päätelmät.....	48
	Lähteet.....	51
	Aineslähteet.....	51
	Muut lähteet .....	51

# 1 Johdanto

Tässä tutkielmassa tarkastellaan viiden sosiaalisen median videosisällöntuottajan puhutun kielen koodinvaihtoa. Puhekieltä analysoidaan sisällöntuottajien omista videoista, jotka he ovat itse editoineet ja julkaisseet Youtube-alustalle. Tavoitteena on selvittää, millaisissa tilanteissa ja asiayhteyksissä koodinvaihtoa esiintyy sekä mistä ilmiön esiintyminen tai koodinvaihdon mahdollinen lisääntyminen tietyllä aikavälillä johtuu.

Koodinvaihto (*code-switching*) ei ole täysin yksiselitteinen käsite, sillä tutkijat ovat luonnehtineet käsitettä lukuisin eri tavoin. Yleisesti sillä tarkoitetaan vuorovaikutuksessa ilmaantuvaa kahden tai useamman kielen vuorottelua. Nykyisen koodinvaihdon tutkimuksen painopiste kohdistuu kahden tai useamman kielen rinnakkaiseen käyttöön, tehtäviin ja paikallisiin merkityksiin keskustelutilanteessa. (Pirainen-Marsh 2008, 143). Auerin (1998) mukaan koodinvaihto on keskusteluun liittyvä toimintaa, jolla merkitään keskustelun osia. Koodinvaihtoon palataan tarkemmin luvussa 2.5.

Koska koodinvaihtoa pidetään moniulotteisena käsitteenä, on sitä tutkittu erilaisista näkökulmista suhteellisen runsaasti. Koodinvaihtoa puhekielessä on tutkinut muun muassa Tiina Parjanen (2015) pikaviestikeskusteluista, Satu Rakkolainen saksansuomalaisten puheesta (2002), Henna Tito suomalais-virolaisten lasten puhekielestä (2015) sekä Emilia Korpinen S2-opettajien opetuspuheesta (2006). Koodinvaihdosta on myös tehty runsaasti tutkielmia opetusdiskurssissa ja L1-kielen vaikutuksesta kielenoppimisessa (esim. Pikelner 2018, Hurtig 2012). Pääasiassa koodinvaihtoon liittyvä tutkimus on siis keskittynyt kielen osaamisen ja oppimisen diskurssiin.

Kielipainotteista tutkimusta, jossa aineistona on toiminut Youtube-alusta, on toistaiseksi tehty vähän. Puhekielen variaatiota käsitteleviä tutkimuksia on puolestaan julkaistu useita erityisesti 2000-luvulla (ks. esim. Lappalainen 2004; Paunonen 2005; Kuparinen 2018), mutta edeltävissä tutkimuksissa hyödynnetään harvoin julkaistua videosisältöä. Heidi Muhonen (2019) tutkii pro gradu - tutkielmassaan ”Multimodaalinen analyysi osallistumiskehikon rakentumisesta monikielisessä videoblogissa” vuorovaikutusta Youtube-alustalle julkaistuissa videoblogeissa. Muhosen (2019) mukaan monikieliset keskustelijat hyödyntävät yksikielisiin verrattuna laajemmin multimodaalisia resursseja vuorovaikutuksessaan. Muhonen (2019) keskittyy tutkimuksessaan tarkastelemaan informanttien osallistumiskehikkoa eli vuorovaikutuksen kuulijoiden vastaanottajaroolien monimuotoisuutta.

Karo Junell (2020) tutkii pro gradu -tutkielmassaan englannin kielen käyttämistä erilaisten asenteiden ilmaisemiseen ja positiointiin dialogisella kentällä. Junell (2020) tarkastelee seitsemän tunnetun suomalaisen sisällöntuottajan videoita ja suhteuttaa analyysinsa evaluaatioteoriaan pohjautuvaan diskurssianalyysiin. Junellin (2020) tutkimustulosten mukaan englantia käytetään suomen ohella niin positiivisten kuin negatiivistenkin tunteiden ilmaisuun. Englanninkielinen ilmaisu on usein negatiivisten asenteiden kuvaamisessa rennompaa ja tilapäisempää, eikä positiivinen asenne välttämättä välity yhtä aitona kuin suomeksi. Junell (2020) esittää, kuinka yleisöä loukkaavaa asennetta saatetaan lieventää englannin kielen käytöllä. Hänen mukaansa koodinvaihtoa käytetään usein tilanteissa, joissa halutaan korostaa kontrastia kahden asenteen välillä. (Junell 2020.)

Satu Arvola (2018) puolestaan tutkii pro gradu – tutkielmassaan koodinvaihtoa kirjoitetusta kielestä Facebook-alustalla. Arvola (2018) kuvailee osallistumiskehikon käsitteen avulla ilmiötä, jossa koodinvaihdolla joko laajennetaan tai supistetaan osallistujien määrää. Arvolan (2018) mukaan informantti laajentaa englannin kieleen, jotta myös englanninkieliset ymmärtäisivät halutun viestin. Arvolan (2018) tutkimus käsittelee koodinvaihtoa ja viestin välittymistä sen ymmärrettävyyden näkökulmasta.

Tutkimuskysymykset ovat seuraavat:

1. Millaista koodinvaihtoa sisällöntuottajien puhekielessä esiintyy?
2. Millaisissa tilanteissa tai asiayhteyksissä koodinvaihtoa esiintyy?
3. Onko koodinvaihto on lisääntynyt sisällöntuottajien puheessa?

Tutkittavat sisällöntuottajat ja heidän tarkasteltavat videonsa on valittu tietoisesti, sillä valitut sisällöntuottajat ovat olleet tuttuja tutkielman tekijälle useamman vuoden ajan. Näin ollen mahdollisia ennakkokäsityksiä ja hypoteeseja on muodostunut jo ennen varsinaisia tuloksia. Tutkielman tekijä on kiinnittänyt huomiota sisällöntuottajien puhekielessä ilmenevään koodinvaihtoon, minkä seurauksena tämän pro gradu -työ sai alkunsa.

Hypoteesina on se, että kaikkien tutkittavien sisällöntuottajien puheessa ilmenee koodinvaihtoa ainakin jonkin verran ja että koodinvaihto on lisääntynyt sisällöntuottajien puheessa ajan kuluessa. Hypoteesiin vaikuttaa erityisesti englannin kielen edelleen vahvistuva asema. Englantia käytetään nykyään myös sellaisissa tilanteissa, joissa koodinvaihto ei ymmärrettävyyden kannalta ole tarpeellinen (Leppänen & Nikula 2008, 13). Mantilan (2004) mukaan sellainen puhekielisyyys, joka nykypäivänä on yleistynyt ja osoittaa sosiaalista variaatiota, saattaa tulevaisuudessa olla yleinen ja

sosiaaliselta variaatioltaan neutraali. Yksilöllistä variaatiota tarkastellaan luvussa 4.2 ja englannin kielen asemaa ja määrittelyä luvussa 4.5.

Aikaisemmin on tutkittu koodinvaihtoa ja puhekieltä erityisesti opetusdiskurssissa, osallistumiskehikon rakentajana multimodaalisin resurssein sekä asenteiden ja tunteiden ilmaisijana. Tämä tutkimus luo lisäksi katsauksen nuorten aikuisten puhekieleen ajallisen ulottuvuuden kautta, kun tarkastellaan sitä, miten ja millaisissa asiayhteyksissä koodinvaihto mahdollisesti esiintyy ja lisääntyy sisällöntuottajien puhekielessä.

Jotta koodinvaihdon esiintymisestä voitaisiin tehdä johtopäätöksiä ja jotta tulokset olisivat vertailukelpoisia, on tutkimukseen valikoitunut viisi videosisällöntuottajaa. Tutkittavat sisällöntuottajien videot on valittu kvalitatiivisin menetelmin. Julkaistujen videoiden aikaväli on noin viidestä kahdeksaan (2013–2021) vuotta sisällöntuottajasta riippuen. Videoista on litteroitu yksinkertaisesti sellaiset kohdat, joissa sisällöntuottajan puheessa ilmenee koodinvaihtoa. Litteraateista muodostetaan aineisto, josta tarkoituksena on eritellä koodinvaihdon esiintymisen monimuotoisuutta. Koodinvaihtokohtia tarkastellaan myös kvantitatiivisesti, mutta analyysin pääpaino kohdistuu aineiston kvalitatiivisiin ominaisuuksiin. Aineistosta tarkastellaan koodinvaihtoa lisäksi ajallisessa kontekstissa sekä esimerkiksi asiayhteyksien ja puhetilanteiden näkökulmista. Aineistoon ja sen analyysiin palataan luvuissa 3 ja 5.

Tutkielman toisessa luvussa tarkastellaan tutkimuksen teoreettista viitekehystä ja keskeisiä käsitteitä, kuten *puhekielen variaatio* ja *koodinvaihto*. Kolmannessa luvussa esitellään tarkemmin tutkielman aineisto, eettiset lähtökohdat ja videosisällöntuottajat. Neljännessä luvussa esitellään menetelmät, jotka tässä tutkielmassa keskittyvät sisällönanalyysin, sosiolingvistiikan, vuorovaikutussosiolingvistiikan ja keskusteluanalyysin kehykseen. Neljännessä luvussa lisäksi tarkennetaan koodinvaihdon ja monikielisyyden suhdetta sekä avataan esimerkkien tarkasteluun käytettyjä valintoja. Viidennessä luvussa ammennetaan analyysia aineiston perusteella. Kuudennessa luvussa kootaan päätelmät tutkimustulosten pohjalta ja ehdotetaan aiheita jatkotutkimukselle.

## 2 Teoriaa ja keskeisiä käsitteitä

### 2.1 Sosiolingvistinen tutkimus

Sosiolingvistiikka on kielentutkimuksen ala, joka käsittelee kieltä sen sosiaalisessa ympäristössään ja tarkastelee todellisuutta lingvistiikan keinoin. 1950-luvulta lähtien sosiolingvistinen tutkimus alkoi vakiintua osaksi kielentutkimusta. Erityisesti William Labovin työn ansiosta käsite *sosiolingvistiikka* vakiintui. (Nuolijärvi 2000, 13–14.) Kielitieteen haaran historia ulottuu 1960-luvulle. Vuosikymmenten aikana se on eriytynyt useaksi eri suuntaukseksi, joita yhdistää kielen ja sosiaalisen rakenteen tutkiminen. (Keskimaa 2016, 496.)

Nuolijärven (2000) mukaan sosiolingvistinen tutkimus on jo 1990-luvulle tultaessa monipuolistunut suuresti. Sosiolingvistiikka keskittyy kuvaamaan kieltä sosiaalisena ilmiönä ja yrittää etsiä yhteyksiä kielenkäytön ja yhteiskunnallisen todellisuuden väliltä. Sosiolingvistiikka pohtii siis sitä, mikä vaikuttaa kielen vaihteluun ja muuttumiseen. (Nuolijärvi 2000, 14.)

Nuolijärvi (2000) esittää, kuinka sosiolingvistiikkaa voidaan määritellä useista eri näkökulmista, joten tieteenalana se ei ole täysin yksiselitteinen. Lähtökohtana on kuitenkin kielenkäytön sosiaalinen todellisuus. (Nuolijärvi, 2000, 15.) Sosiolingvistiikan alle kuuluvat esimerkiksi sosiologisesti painottunut kielisosiologia, lingvistisesti painottunut variaatioanalyysi ja vuorovaikutuksen sosiologiasta vaikutteita saaneet suuntaukset. Suomalaisessa puhutun kielen variaatiota kartoittavassa tutkimuksessa tilannevariaatio on jäänyt vähemmälle tarkastelulle kuin käyttäjäkohtainen eli dialektaalinen variaatio, jota on tutkittu erityisesti useissa opinnäytetöissä ja muissa tutkimuksissa (Lappalainen 2004, 13; 609).

Sosiolingvistiikan alue on jaettu karkeasti mikro- ja makrososiolingvistiikkaan. Mikrososiolingvistiikka tarkastelee sosiaalisten rakenteiden vaikutuksia ihmisten puhetapoihin ja esimerkiksi sitä, miten kielen eri varieteetit korreloivat sosiaalisten muuttujien kanssa. Mikrososiolingvistiikka ottaa siis huomioon kielen vaihtelun suhteessa esimerkiksi sukupuolen tai iän kanssa. Makrososiolingvistiikka puolestaan tarkastelee kielen vaihtoa, kielellisiä asenteita, kielen säilymistä ja kuolemaa sekä eri kielimuotojen funktionaalisia tehtäviä kieliyhteisössä. Näillä molemmilla alueilla traditio on pitkälti ollut kvantitatiivinen. (Nuolijärvi 2000, 15.)

Sosiolingvistiikan sisäinen kahtiajako on mahdollinen myös jakamalla käsittely demografiseen ja interaktionistiseen sosiolingvistiikkaan. Demografinen suuntaus tutkii sosiaalisten muuttujien ja kielen ilmiöiden suhdetta, kun taas interaktionistinen tarkastelee puhetilannetta. Usein



sosiolingvistiikan menetelmien avulla on lähestytty puhekieltä ja sen vaihtelun tutkimista ja osoittamista. (Nuolijärvi 2000, 16.) Interaktionaalisessa sosiolingvistiikassa keskeinen käsite on *kontekstualisointi* (*contextualisation*). Termillä tarkoitetaan sitä, että puhuja sijoittaa lausumansa kielellisiin ja ei-kielellisiin keinoihin kontekstiin tai hän osoittaa yhteyden, jossa lausuma on tulkittavissa. Koodinvaihtoa voidaankin pitää puheen kontekstualisoinnin keinona. Siinä keskustelun koodinvaihdot toimivat kontekstivihjeinä, jotka ohjaavat lausuman tulkintaa. (Kalliokoski 1995; Gumperz 1982).

Lappalaisen (2004, 13) mukaan suomalaisessa tutkimuksessa painopiste on ollut variaatioanalyysissä, jonka pioneerinä voidaan pitää William Labovia. 1970-luvulla Suomessa keskusteltiin labovilaisesta ja lomanilaisesta sosiolingvistiikasta. Labovilainen variaatioanalyysi keskittyi kielen äänne- ja muotopiirteisiin, kun taas Lomanin Talsyntax-hankkeen periaatteena oli syntaktisen variaation kuvaus. (Paunonen 2009, 558.) Paunosen (2009) mukaan sosiolingvistiikan pioneerien yhtenä keskeisenä tavoitteena oli tutkia kielessä ilmenevää variaatiota siitä näkökulmasta, miten variaatio auttaisi ymmärtämään kielen rakennetta ja sen muutoksia entistä paremmin (Paunonen 2009, 559).

Lappalaisen (2001b) mukaan 1970-luvulla neljässä kaupungissa toteutettua Nykysuomalaisen puhekielen murros -hanketta voidaan pitää sosiolingvistisen tutkimuksen alkuna Suomessa. Hankkeen pohjalta ilmestyi 1980-luvulla useita raportteja ja pro gradu -tutkielmia. Lappalainen (2001b) esittää, kuinka hanketta seurannut sosiolingvistinen tutkimus on Suomessa pitkälti toteutettu samaan tapaan siten, että eri ikäryhmien kielistä on yleensä tarkasteltu suhteellisen helposti kvantifioitavissa olevia äänne- ja muotopiirteitä. (Lappalainen 2001b, 76.) Variaationtutkimus on pääasiassa keskittynyt äänne- ja muotopiirteiden tarkasteluun, kun taas keskustelunanalyysin tarkastelemat ilmiöt ovat olleet usein toiminnallisia kategorioita. Suomalaisissa tutkimuksissa on esimerkiksi tarkasteltu kannanottoja, nauruja ja läsnäolijaan viittaamista (esim. Priiki 2011, Tikkanen 2011, Haakana 1996).

## 2.2 Vuorovaikutussosiolingvistiikka

1970–1980-luvuilla syntynyt vuorovaikutussosiolingvistiikka tutkii kielellistä variaatiota ja diversiteettiä sosiaalisen kanssakäymisen osana. Keskeistä on se, mitä variaatiolla tehdään vuorovaikutuksessa ja millaisia merkityksiä vuorovaikutustilanteen osallistujat antavat kielelle. (Lehtonen 2015, 31, 33.) Suuntaus rakentuu sosiolingvistiikan pohjalle eli kielellä nähdään olevan

yhteys ympäröivään yhteiskunnalliseen todellisuuteen. Perinteisestä sosiolingvistiikasta keskeinen erottava ajatus vuorovaikutussosiolingvistiikassa on se, että variaatiolla on merkitystä vuorovaikutuksen kannalta. (Priiki 2017, 19.)

Vuorovaikutussosiolingvistiikan pioneerina pidetään John Gumperzia, joka toi esiin, ettei kvantitatiivinen sosiolingvistiikka kykene osoittamaan yksinään kielen sosiaalisten tulkintojen syntymistä (Gumperz 1982; Priiki 2017; Lehtonen 2015, 31). Vuorovaikutussosiolingvistiikan fokus keskittyy siihen, miten sosiokulttuurinen tieto muokkaa kielellistä vuorovaikutusta ja miten se näkyy kielellisissä toiminnoissa (Gumperz 2001).

Vuorovaikutussosiolingvistiikan suuntauksen muodostuminen on tapahtunut aalloittain. Ensimmäinen aalto alkoi William Labovin (1966) tunnetusta tutkimuksesta, jossa hän tarkasteli muun muassa sanaston ja äänteiden vaihtelua esimerkiksi iän, sukupuolen tai sosiaaliluokan mukaan jaotelluissa ryhmissä. Toisessa aallossa tarkastelun keskiöön nousi havainnoitavan yhteisön etnografinen ja sen sosiaalisen todellisuuden tuntemus. Samaan aikaan yksilön kieli alettiin nähdä muuttuvaisena identiteetin ilmaisukeinona lapsuudessa muotoutuvan puhutavan sijasta. Edelleen puhujia kuitenkin jaoteltiin staattisten kategorioiden mukaisesti. Vasta kolmas aalto näkee variaation olennaisena yksilön kieleen ja vuorovaikutukseen kuuluvana ilmiönä. Näin päädytään vuorovaikutussosiolingvistiikassa keskeiseen näkemykseen; tärkeää on se, mitä vuorovaikutuksessa tehdään, ja että puhujan valinnat ovat yhteydessä häntä ympäröivään sosiaaliseen todellisuuteen sekä keskustelun rakentumiseen. (Priiki 2017, 19; Eckert 2012; Gumperz 2001, 218–223.)

Nykyaikaisen vuorovaikutussosiolingvistiikan lähtökohtana voidaan pitää metaforisen koodinvaihdon käsitteen perusidea. Metaforinen koodinvaihto tarkoittaa sitä, että saman tilanteen sisällä eri koodeilla herätetään assosiaatioita erilaisiin konteksteihin, joiden näkökulmasta sanottua tulkitaan. (Gumperz 1982; Lehtonen 2015, 33.)

Myös vuorovaikutussosiolingvistiikan tutkimuskentässä kontekstualisointi on keskeinen käsite. Puhujat voivat eri varianteja käyttämällä esittää omia tulkintojaan ja rakentaa omaa rooliaan kertojana, ilmaista asenteita tai herättää assosiaatioita tiettyihin sosiaalisiin ryhmiin tai tilanteisiin. (Priiki 2017, 21.) Kielessä ilmenevä variaatio toimii tietynlaisena resurssina, josta puhujan tietoiset tai tiedostamattomat valinnat nousevat (Mantila 2004).

Vuorovaikutussosiolingvistiikassa tarkastelun kohteena olevia ilmiöitä on lukuisia, myös yksittäisellä äänne- tai muotopirretasolla. Huomionarvoista onkin se, että niiden analyysillä pyritään selittämään jotain tutkittavan ryhmän sosiaalisesta dynamiikasta tai laajemmin yhteisön voimasuhteista. (Lehtonen 2015, 34.) Priiki (2017, 21) toteaa, että vuorovaikutussosiolingvistiikan näkökulmasta voidaan ajatella, että konteksti ja tilanne vaikuttavat siihen, mikä puhujan valitsemista mahdollisista

varianteista esiintyy. Puhujan valitsema variantti muokkaa puolestaan kontekstia ja muiden osallistujien tulkintaa siitä. Konteksti puolestaan käsitetään vuorovaikutussosiolingvistiikassa laajempaan kuin paikalliseen puhetilanteeseen viittaavana (Mononen 2013, 163).

### 2.3 Keskustelunalyysi

Keskustelunalyysin juuret ovat 1960-luvun Yhdysvalloissa sosiologian piirissä. Etnometodologia syntyi Harold Garfinkelin ideoinnin pohjalta. Tätä lähti kehittämään edelleen Harvey Sacks, jonka luentoja voidaan pitää keskustelunalyysin kulmakivinä. (Lappalainen 2004, 18.) Keskustelunalyysi on tuonut yhteen sosiologeja ja kielitieteilijöitä, joiden yhteisinä tavoitteina on ollut kielen ja toiminnan suhteen tarkastelu. Tällaisella monitieteisellä kentällä joudutaan pohtimaan, miten sovittaa yhteen erilaisia tutkimusmetodeja ja -perinteitä. Yleisesti keskustelunalyysissa on korostettu sitä, että arkikeskustelu on ensisijainen vuorovaikutuksen muoto. (Heritage 1996; Lindholm, Stevanovic & Peräkylä 2016, 9; Halonen & Routarinne 2001.)

Sekä maailmalla että Suomessa keskustelunalyytit keskittyivät aluksi tutkimaan tasavertaisten puhujien keskinäisiä arkikeskusteluja, jotka erityisesti olivat puhelinkeskusteluja. 1990-luvun alusta lähtien myös institutionaaliset vuorovaikutustilanteet keräsivät huomiota keskustelunalyysin kentällä. Etenkin tässä yhteydessä keskustelunalyysi laajeni Suomessa myös kielitieteen ulkopuolelle, kuten sosiologiaan, sosiaalipsykologiaan ja puhetieteisiin. Tutkimus on laajentunut edelleen tarkastelemaan myös kielellisesti epäsymmetrisiä vuorovaikutustilanteita. (Lindholm ym. 2016, 10–11.)

Tutkimussuuntaus laajenee yhä monimuotoisempiin näkökulmiin ja kontekstit lisääntyvät jatkuvasti. Vielä vuosikymmen sitten kaikkea keskustelunalyyttistä tutkimusta ohjasi ymmärrys, että vuorovaikutuksen säännönmukaisuudet esiintyvät alkuperäisimmässä ja monipuolisimmassa muodossaan arkikeskustelussa. Nyt on kuitenkin havaittu, että myös arkikeskustelun sisällä esiintyy suuri määrä vaihtelua. Arkitilanteissakin on vuorovaikutukselle keskeistä se, mikä on osallistujien välinen suhde ja mitä he pyrkivät vuorovaikutuksessaan milloinkin tekemään. (Lindholm ym. 2016, 12.)

Keskustelunalyysi ei siis ole lähtökohtaisesti teoria kielestä, vaan metodi, joka on kehitetty erityisesti sosiaalisen toiminnan tarkasteluun. Keskustelunalyysin metodia voidaan painottaa lingvistisesti, jolloin se lukeutuu vuorovaikutuslingvistiikan (luku 2.2) alaan. Tässä käytännönläheisessä suuntauksessa operoidaan puhujien kategorioilla eli käsitteillä, jotka puhujat

osoittavat itse toiminnan kannalta relevanteiksi. Keskustelunanalyysin ihanteena pidetään sitä, ettei tutkija projisoi aineistoon ennalta oletettuja kategorioita, jotka eivät osoittaudu vuorovaikutuksen etenemisen tai tulkinnan kannalta relevanteiksi. (Etelämäki, Herlin, Jaakola & Visapää 2009, 162.)

Suomessa keskustelunanalyysia on harjoitettu 1980-luvulta lähtien. Ensimmäisen suomalaisen keskustelunanalyttisen tutkimushankkeen tavoitteena oli selvittää syntaksin ja kielellisen toiminnan yhteyksiä (Hakulinen & Sorjonen 1993). Tutkimussuuntaus on kehittynyt elinvoimaiseksi myös fennistiikan piirissä. Lappalaisen (2004, 18) mukaan keskustelunanalyysin keskeinen periaate on ajatus siitä, että vuorovaikutus on järjestynyttä ja jäsentynyttä toimintaa yksityiskohtia myöten. Merkitykset rakentuvat osallistujien yhteistyön tuloksena.

Keskusteluntutkija olettaa yleisesti, että kaikki toiminta on järjestäytynyttä ja merkityksellistä kommunikaation ja sosiaalisen toiminnan rakentumisen kannalta. Puhe on toimintaa, jossa toiminnot suoritetaan kieliopilliselta kannalta määriteltynä sanojen, lausekkeiden tai lauseiden avulla. Keskustelussa tuotetaan lausumia saadakseen aikaan tiettyjä asioita vuorovaikutuksessa toistensa kanssa. Lisäksi keskusteluntutkimuksessa oletetaan, että kukin lausuma on tilanteinen, eli se saa merkityksensä juuri siinä vuorovaikutuskontekstissa, jossa se tuotetaan. Keskustelijoiden välinen yhteisymmärrys syntyy vuorovaikutuksessa sen mukaan, miten he itse ymmärtävät tilanteen. (Siitonen 2016, 4.)

Keskustelunanalyysissa voidaan tutkia esimerkiksi vuorojen järjestyneisyyttä, sekvenssejä eli toisiinsa kytkeytyviä jaksoja sekä vuorojen korjaamista. (Lappalainen 2004, 18.) Koska tämän tutkielman metodina ei toimi yksiselitteisesti keskustelunanalyysi, vaan sen lisäksi näkökulma on vahvasti sosiolingvistinen, ei keskustelunanalyysin metodiin nojata uskollisesti. Perinteisessä sosiolingvistiikassa tarkastellaan esimerkiksi kielellisiä valintoja osallistujien taustamuuttujien näkökulmasta keskustelunanalyysia tarkemmin (Lappalainen 2004, 19). Tämän tutkielman ulkopuolelle jäävät esimerkiksi vuorovaikutuksen multimodaalisuuden tarkasteleminen eli esimerkiksi se, miten puheen lisäksi puhuja hyödyntää prosodiaa tai kasvoniilmeitä, sillä kyseiset piirteet eivät tuo lisäarvoa tutkimuskysymyksiin tai -tuloksiin.

Keskustelunanalyysissa lähtökohtana on se, että konteksti ei ole pysyvä, vaan osallistajat rakentavat sitä jatkuvasti yhteistyössä. Keskustelunanalyysi liitetään tähän tutkielmaan muun muassa sen aineistolähtöisten ominaisuuksien kautta. Periaatteena on tutkia luonnollisesti tapahtuvaa toimintaa. Keskustelunanalyttikko ei haastattele ihmisiä heidän vuorovaikutuskäytänteistään, jotta voidaan saada todellinen kuva tapahtuneesta toiminnasta. Tutkija ei pyri keksimään itse, mitä voisi tapahtua tai mitä keskustelijat mahdollisesti tekevät, vaan hän tutkii sitä, mitä todella tapahtuu. Havaintoja ei perusteta pelkästään tilanteiden havainnointiin ja niiden kirjaamiseen. Keskustelunanalyttisen

näkemyksen mukaan toimintaa tulee ensisijaisesti tarkkailla sen luonnollisessa ympäristössä, jotta siitä saa todenmukaisen kuvan. (Vatanen 2016, 313.)

## 2.4 Puhekielen variaatio

Tarkasteltaessa kielenulkoisten tekijöiden roolia puhujien kielellisten valintojen selittäjinä on syytä erottaa puhetilanteen tapahtumakonteksti ja se, millaisiin erilaisiin ryhmiin yksilö kuuluu taustamuuttujiensa puolesta. Variaationtutkimuksessa on katsottu, että esimerkiksi ikä, sukupuoli ja sosiaaliryhmä korreloivat kielellisen variaation kanssa. Yksilön puhetapa vaihtelee kuitenkin myös tilanteittain. Puhetilanteen kontekstilla tarkoitetaan lähinnä sellaisia kielenulkoisia seikkoja, kuten esimerkiksi tilanteen osallistujia ja heidän taustamuuttujiaan, puheenaiheita ja tapahtumapaikkaa ja -aikaa. (Lappalainen 2001a, 157–159.)

Suomalaisen nykypuhekielen muutosta ja variaatiota kartoittavissa tutkimuksissa on yleisesti tarkasteltu ikää ja sukupuolta kielenulkoisina muuttujina. Vähemmän on kiinnitetty huomiota sosiaaliseen asemaan ja esimerkiksi ammattiin. Myös tilannekohtainen vaihtelu on jäänyt vähemmälle huomiolle käyttäjäkohtaisen eli dialektaalisen variaation sijasta. (Mantila 2004, 322.) Kurjen (2018) mukaan puhuttua suomea koskeva tutkimus on sekä laventunut että kirjavautonut entisestään 2000-luvun kuluessa. Varsinkin tutkielmissa on kuitenkin usein tarkasteltu näkökulma rajattu tarkasti jonkin paikkakunnan, seudun tai alueen ihmisten kielikäsitteisiin.

Variaationtutkimuksen teoreettisena taustana on Suomessa pääsääntöisesti käytetty labovilaista sosiolingvistiikan traditiota, jossa kielen variaatiota tarkastellaan sosiaalisten muuttujien kontekstissa. Informantteja on analysoitu ryhmittäin, jotka useimmiten ovat olleet vanhat, työikäiset ja nuoret henkilöt. Puhekielen ja sen variaation tutkimus ei kuitenkaan ole tuottanut täysin yksiselitteisiä tuloksia. Mantilan (2004, 324) mukaan tuloksia on paljon, mutta ei tiedetä syitä siihen, miksi jotkin kielenpiirteet ovat leimautuneet kielteisesti ja miksi taas jotkin yleistyvät jatkuvasti. Kun erilaisten puhekielen piirteiden variaatiota on tutkittu paljon, on kuitenkin mahdollista alkaa pohtia näihin piirteisiin liittyviä kulttuurisia merkityksiä, joita ovat esimerkiksi kaupunkilaisuus ja korrektius (Paunonen 1982).

Nuorten kieli on ollut sosiolingvistiikassa tutkimuksen kohteena jo tradition alkuajoista lähtien. Variaationtutkimus on painottunut pitkälti äänne- ja muoto-opillisten piirteiden tarkasteluun, mutta nuorten kielestä on tutkittu myös sosiaalisen verkon kielenkäyttöä, sanastoa, keskustelutyylisiä ja

kertomisen tapaa. Suurin osa nuorten kielestä tehdystä tutkimuksesta on tehty variaationtutkimuksen piirissä. (Lappalainen, 2001b.)

## 2.5 Koodinvaihto ja sen ulottuvuudet

Prototyypisimmillään koodinvaihtoa pidetään kahden tai useamman eri kielen vaihtelua samassa keskustelussa (Gumperz 1982; Auer 1995; Myers-Scotton 1998). Gumperz (1982) kuvaa koodinvaihtoa monikielisen vuorovaikutuksen ilmiönä. Vuorovaikutuksen näkökulmasta kielten kohtaaminen kytkeytyy koodinvaihtoon ja laajemmin monikielisyyden tulkintaan kontekstissa (Kalliokoski 2009, 19). Tarkemmin keskusteluanalyttisessä tutkimuksessa koodinvaihtona pidetään sellaisen kahden tai useamman kielen tai kielimuodon vaihtelua, jossa osallistuvat orientoituvat koodien erilaisuuteen. Kahden kielen rinnakkaista käyttöä voi Auerin (1998) mukaan ajatella jatkumona, jonka toisessa päässä on koodinvaihto paikallisena merkitystä kantavana vuorovaikutuskeinona ja toisessa päässä kahden kielimuodon sekoittumisesta aiheutuva fuusio, josta koodinvaihtokohtia ei voi enää erottaa.

Koodinvaihdon käsitettä on keskusteluntutkimuksen näkökulmasta rajattu tarkoittamaan sellaista kahden kielen tai varieteetin vaihtelua, jolla on sosiaalista tai vuorovaikutuksellista merkitystä. Piirainen-Marshin (2008) mukaan koodinvaihto ymmärretään puheen resurssina, joka samanaikaisesti heijastaa puhetilanteen rakenteita sekä jäsentää ja muokkaa kontekstia. (Piirainen-Marsh 2008, 143.) Termiä koodinvaihto (*code-switching*) on kuvailtu monessa merkityksessä eri käsitteillä, kuten *koodinsekoitus (code-mixing)*, *koodinmuutos (code-change)*, *koodien tai kielten vuorottelu (code-/language alternation)* tai *koodin kopiointi (code-copying)* (Kovács 2009, 24). Tässä tutkielmassa käytetään termiä *koodinvaihto* yläkäsitteenä Kovácsin artikkelin (2009) tavoin. Hänen mukaansa koodinvaihtoa on kahden tai useamman kielen tai koodin esiintyminen saman keskustelun aikana. Käsitteeseen sisältyvät niin monilauseiset vaihdot kuin yksittäiset sanatkin. (Kovács 2009, 24.) Tässä tutkielmassa tarkastellaan koodinvaihtotapauksia, joissa kieli vaihtuu matriisikielestä eli suomesta upotettuun kieleen eli englantiin.

Matriisikieli tarkoittaa yleensä keskustelun pääkieltä, joka tavallisesti määrää lauseiden kieliopilliset puitteet. Keskustelussa esiintyviä muita kieliä, joista otetaan aineksia matriisikieleen, kutsutaan joko upotetuksi kieleksi (*embedded language*) tai termillä (*guest language*) vieraskielisistä aineksista puhuttaessa. (Muysken 1995, 198; Myers-Scotton 1997.)

Tilanteinen koodinvaihto liittyy tilanteisiin, joissa vuorovaikuttajat, puhetilanne tai toimintatyyppi ohjaa tiettyjen konventioiden tai normien mukaisesti kielen vaihtamiseen. Käsitteeseen liittyy ajatus siitä, että sosiaalinen tilanne muuttuu kieltä vaihdettaessa esimerkiksi muodollisesta epämuodolliseksi. (Koskela 2015.) Auer (1995) on tutkinut tilanteita, joissa koodinvaihtoa tyypillisesti esiintyy. Auerin (1995) mukaan koodinvaihtoa esiintyy usein sivuhuomautusten yhteydessä ja esimerkiksi silloin, kun puhetoiminnassa tapahtuu muutos, aiheen vaihto tai kielellä leikittelyä. (Auer 1995, 117–120.)

Tilanteisella koodinvaihdolla Blom ja Gumperz (1972) tarkoittavat ilmiötä, jota voi luonnehtia diglossiaksi yksilötasolla. Eri kielimuodot kytkeytyvät eri tilanteisiin, joissa sama henkilö valitsee käyttöönsä jonkin kielellisen varieteetin ulkoisten puitteiden ja vastaanottajien perusteella. (Blom & Gumperz 1972, 410.) Lehtosen (2015, 32–33) mukaan tilanteinen koodinvaihto tarkoittaa myös sitä, että eri koodeja (kuten eri kieliä) käytetään täysin eri tilanteissa, esimerkiksi yhtä koulussa ja toista kotona.

Muhonen (2008) esittää, kuinka koodinvaihto luo kontrastin sen välille, mitä tehdään nyt ja mitä on tehty aikaisemmin. Koodinvaihdolla voidaan siis tarkoittaa diskurssinäkökulmaa vieraan kielen esiintymiseen. *Insertit* (*insertions*) eli *lisäykset* puolestaan tarkoittavat sellaisia rakenteellisia kategorioita, joita nimetään (kuten sana, lause). Lisäys on usein hetkellinen kielen vaihtelu, jonka jälkeen palataan matriisikielen eli viestinnän pääasialliseen kieleen. Prototyyppinen lisäys on yhden sanan pituinen nomini. (A. Muhonen 2008, 171–173.)

Lainaaminen ja koodinvaihto eivät ole toisiaan poissulkevia ilmiöitä, joten niitä on vaikeaa erottaa toisistaan. Sanakirjoihin vakiintuneissa lainasanoissa on samoja kytkentöjä kuin koodinvaihdossa olevilla kielellisillä elementeillä. Pitkällä aikavälillä tilapäisessä käytössä olevat lainat siirtyvät lopulta yksikielisille toistuvasti käytettäessä. Koska sanan vakiintuminen on luonteeltaan asteittainen prosessi, on mahdotonta päättää, onko jokainen laina tilapäinen vai vakiintunut sanastoon. (Auer & Eastman 2010, 86.) Koska ilmiö ei ole yksiselitteinen, tässä tutkielmassa lainaaminen lasketaan koodinvaihdoksi.

Gumperz (1982) puolestaan määrittelee lainaamisen kohteena olevan lähinnä yksittäiset sanat tai idiomaattiset ilmaukset, jotka sitten sisällytetään matriisikielen kieliopilliseen järjestelmään niin, että ne käyttäytyvät kyseisen kielen morfologisten ja syntaktisten sääntöjen ja ominaisuuksien mukaisesti. Lainaus suuntaa ei ole määritelty, vaan se voi tapahtua toisesta kielestä äidinkielen ja päinvastoin. Gumperz erottaa lainaamisen ja koodinvaihdon siten, että koodinvaihto perustuu puhujien tietoiseen tai tiedostamattomaan kielellisen aineksen prosessointiin, jossa operoi samanaikaisesti kaksi erillistä kieliopillista systeemiä. (Gumperz 1982, 66.) Siinä missä Gumperz

käsittää lainaamisen ainoastaan yksittäisten sanojen tai fraasien lainaamisena, Lauttamus toteaa lainaamisen olevan myös kielen rakenteeseen liittyvien elementtien lainaamista. (Lauttamus 1992, 8.)

Koodinvaihdon ja lainan rajanveto on osoittautunut ongelmalliseksi tutkimuksissa, joissa tarkastellaan kaksikielisyyttä ja kielten kohtaamista. Määrittelyssä voidaan todeta ongelmaksi myös sellainen kielenkäyttötilanne, jossa kielten vaihtelu on jatkuvaa eikä yksittäiselle koodinvaihdolle voida välttämättä osoittaa omaa tulkintaa. Kuitenkin vakiintuneen käsityksen mukaan koodinvaihto hahmotetaan ensisijaisesti keskustelun koodien vaihdoksi ja kontekstualisoinnin keinoksi. (Kalliokoski 2009, 15.)

Hiidenmaan (2003) mukaan nykymaailmassa lainasanat eivät ole yleisiä internationalismeja, vaan ne lähtevät yhdestä kielestä, tavallisimmin englannista. Tällaisten lainojen mukauttaminen ei enää ole yhtä helppoa kuin klassisista kielistä lähtevien sanojen, sillä erot suomen ja englannin kielen kirjoitusasussa ja ääntämisessä ovat suuret. Huomionarvoista on, että lainat eivät synny erityisen kirjallisen kielimuodon tarpeeseen, vaan niitä voidaan käyttää arkipäiväisessä puheessakin. Suomen kielen ilmausvarantoa laajennetaan lainaamalla. (Hiidenmaa 2003, 97–98.)

Kalliokoski (1995) lähestyy koodinvaihtoa pragmaattisesta näkökulmasta, jossa hän tarkastelee sitä intertekstuaalisena ilmiönä. Kalliokosken mukaan koodinvaihtoa voidaan pitää yhtenä keinona tuoda kielelliseen vuorovaikutustilanteeseen dialogisuutta ja siten vaikuttaa lausuman tulkintaan. Kalliokoski (1995) esittää, että sosiolingvistinen tutkimus arjen kokemusten ohessa on opettanut, kuinka kielimuodon valinta kertoo puhujasta sen, millaiseksi hän on tulkinnut kielenkäyttötilanteen. Tyylin tai kielen vaihtuessa kesken keskustelun voidaan olettaa, että vuorovaikutustilanteen osallistujien käsitys lausunnan merkityksestä muuttuu samanaikaisesti. Samalla vaikutetaan siihen, miten lausuma tulkitaan. (Kalliokoski 1995, 2.)



## 3 Aineisto

### 3.1 Aineiston lähtökohdat

Tässä tutkielmassa tarkastellaan viiden sosiaalisen median videosisällöntuottajan puhekielen koodinvaihtoa. Puhekieltä analysoidaan sisällöntuottajien omista videoista, jotka he ovat itse editoineet ja julkaisseet Youtube-alustalle tietyllä aikavälillä. Aineistosta pyritään selvittämään, millaisissa tilanteissa ja asiayhteyksissä koodinvaihtoa esiintyy sekä mistä ilmiön esiintyminen tai koodinvaihdon mahdollinen lisääntyminen tietyllä aikavälillä johtuu.

Aineistoksi on valikoitunut viideltä videosisällöntuottajalta aineiston keruuhetkellä neljä varhaisinta ja neljä uusinta videota. Videot on julkaistu vuosien 2013 ja 2021 välillä. Videoiden kestot vaihtelevat sisällöntuottajasta riippuen muutamasta minuutista yli 45 minuuttiin. Videoiden keskimääräiset kestot esitetään luvussa 3.2. Taulukosta ilmenee kunkin informantin vanhimman ja uusimman tuotannon videojulkaisujen keskimääräiset kestot sekä aikaväli, jolloin videot on julkaistu.

Koska aineisto koostuu julkisista videojulkaisuista, jotka ovat kaikkien saatavilla ja nähtävillä, ei tutkimuslupia tarvittu. Sen sijaan on huomioitava, että yksityisyyden kunnioittaminen on yksi tärkeimmistä tutkimuseettisistä normeista. Tutkielman aineiston videoiden tekijät eivät varsinaisesti osallistu tutkielman tekoon tietoisesti, sillä videoita ei ole tuotettu tutkimustarpeeseen. Muun muassa tästä syystä videosisällöntuottajat esitetään anonymyminä, sillä heidän henkilötietonsa eivät tuo lisäarvoa tutkimustuloksille. Informantit on pyritty anonymisoimaan siten, että heistä kerrotaan vain välttävimmät henkilötiedot.

Sosiaalisen median aineistojen helppo tai vapaa saatavuus ei poista eettisen harkinnan tarvetta. Tilanteesta tekee ongelmallisen se, että yleisesti hyväksytyjä suosituksia sosiaalisen median käytöstä tutkimuksessa ei toistaiseksi ole olemassa. Tutkielmaa tehdessä tiedostetaan se mahdollisuus, että informantit voivat olla anonymisoinnista huolimatta tunnistettavissa, jos tunnistamiseen käytetään erityistä vaivaa. Aineistona käytetyistä videoista voidaan tarjota listaus sitä tarvitsevalle, mutta lähdeluetteloon ei videojulkaisuja ole merkitty tunnistettavuuden vuoksi.

Aineiston kerääminen ja sen muodostuminen on luonteeltaan ollut hyvin kvalitatiivista. Ennen aineiston keräämistä on havaittu, että tutkielmaan valikoitujen sisällöntuottajien puhekielessä ilmenee jonkin verran koodinvaihtoa videoiden perusteella. Kaikkia viittä tutkielmaan valikoituja sisällöntuottajia on seurattu sosiaalisen median alustoilla jo pidemmän aikaan tutkielman tekijän toimesta. Näin syntyi kiinnostus tehdä tutkimusta siitä, millaisissa asiayhteyksissä koodinvaihtoa

oikeastaan esiintyy sisällöntuottajien puheessa ja voidaanko ilmiöstä tehdä systemaattisesti johtopäätöksiä. Hypoteesiksi muodostui myös ajatus siitä, että koodinvaihto on mahdollisesti lisääntynyt vuosien varrella sisällöntuottajien puheessa.

Sisällöntuottajien valitsemiseen ovat vaikuttaneet erilaiset kriteerit. Ensinnäkin haluttiin, että tutkielmaan valikoituvat videosisällöntuottajat olisivat tehneet sisältöä Youtube-alustalle vähintään viisi vuotta, jotta tulokset olisivat jokseenkin vertailukelpoisia. Lisäksi sisällöntuottajan tulisi olla aktiivinen myös tutkielman tekohetkellä, jotta vertailua voitaisiin tehdä vanhemman ja mahdollisimman ajankohtaisen tuotannon välillä. Aktiivisia, yli viiden vuoden ajan Youtube-videoita tehneitä videosisällöntuottajia ei Suomessa ole vielä niin runsaasti, että aineistoa olisi voinut rajata esimerkiksi siten, että sisällöntuottajien ikähaitari tai asuinpaikka olisi ollut erityisen poikkeava. Usein rajaava tekijä oli myös se, että sisältö ei ole tarpeeksi aktiivista tai yli viisi vuotta vanhaa, jotta vertailua voitaisiin suorittaa.

Aineiston reliabiliteetti on siinä mielessä epä johdonmukainen, että sisällöntuottaja voi itse päättää, mitkä videojulkaisut ovat milloinkin julkisesti näkyvillä katsojille. Tämä tarkoittaa sitä, että tutkielman tekohetkelläkin jokin tietty sisällöntuottajan video saattaa olla katsottavissa, mutta tulevaisuudessa video voi olla piilotettuna katsojilta. Tutkielmaan on valittu jokaiselta sisällöntuottajalta tarkasteluun neljä varhaisinta ja neljä uusinta videota siltä aikaväliltä, kun aineistoa on kerätty. Käytännössä tämä tarkoittaa sitä, että aineiston keräämisen jälkeen aktiivinen sisällöntuottaja on julkaissut lisää uutta materiaalia. Aineiston keräämisen ajankohta on sijoittunut syksyn 2020 ja alkuvuoden 2021 välille.

Aineistoa on käsitelty siten, että videoilta on kirjattu ylös kaikki ne kohdat sisällöntuottajan puheesta, joissa ilmenee koodinvaihtoa. Nämä kyseiset kohdat on litteroitu yksinkertaisesti, mikä tarkoittaa sitä, että litteraatteihin ei ole merkitty kielenulkoisia tekijöitä, kuten taukoja, rytmiä tai prosodiaa. Tähän on päädytty siksi, että edellä mainitut piirteet eivät vaikuta relevantisti tutkimustuloksiin. Sen sijaan analyysi perustuu pitkälti kielenulkoisten tekijöiden selittämiseen sosiolingvivistisin näkökulmin sekä puhekielen variaation esittämiseen ja tulkitsemiseen koodinvaihdon osalta. Analysoitavana ovat siis ne kohdat, joissa sisällöntuottaja on enemmän tai vähemmän tietoisesti käyttänyt koodinvaihtoa osana puhettaan.

Aineistolähtöistä kvalitatiivista analyysia voidaan tarkastella eri tavoin. Aineistoa on tarkasteltu jo sen keruuvaiheessa muun muassa siitä näkökulmasta, että tutkielman tekijällä on ollut tietty hypoteesi koodinvaihdon esiintymisestä sisällöntuottajien puheessa. Redusointivaiheessa aineistoa on tiivistetty ja pilkottu osiin siten, että jäljelle on jäänyt sisällöntuottajien koodinvaihtotapaukset. Sen jälkeen aineisto on ryhmitelty luokittelukategorioihin luvussa 5, jossa sitä analysoidaan. Tässä

vaiheessa tuloksia tarkastellaan myös kvantitatiivisesti, mutta analyysin pääpaino on kvalitatiivisessa tutkimusmenetelmässä. Tulosten perusteella näkemykset abstrahoidaan ja lopuksi esitetään johtopäätöksiä.

### 3.2 Videosisällöntuottajat

Jotta videosisällöntuottajat voitaisiin esittää anonymisesti, on heidät nimetty kirjaimin A-E. Kaikki informantit ovat nuoria aikuisia eli 24–30-vuotiaita. Informantti A on julkaissut sisältöä eri sosiaalisen median kanaville vuodesta 2014 alkaen. Tutkielman tekohetkellä hänellä on Youtube-alustalla noin 80 000 tilaajaa. Informantti A jakaa omaa elämäänsä avoimesti videoillaan ja on julkaissut kanavalleen lähes 200 videota. Videoiden sisällöt vaihtelevat ruoka- ja sisustusaiheista muun muassa muotiin ja erilaisista aiheista keskustelemiseen. Varhaisimmat videot ovat kestoiltaan keskimäärin yhdeksän minuuttia ja uusimmat 34 minuuttia. Yhteensä informantti A:n videoilla on yli 14 500 000 katselukertaa. Videoiden lisäksi hän muun muassa tekee podcastia ja työskentelee sosiaalisen median parissa.

Informantti B on julkaissut sisältöä Youtube-kanavalleen vuodesta 2016. Ainoastaan informantti B asustaa Varsinais-Suomen alueella, kun taas muut informantit asuvat Uudenmaan sisällä. Tilaa jia informantti B:llä on kanavallaan noin 140 000 ja videoita yli 200 kappaletta. Myös informantti B jakaa sosiaalisessa mediassa avoimesti omaa elämäänsä. Videoillaan hän kuvaa erilaisia haasteita ja testaa uusia tuotteita. Informantti B on maininnut julkaisuissaan, kuinka hän haluaa inspiroida nuoria ja tarjota katsojille humoristisia videoita. Vanhimpien videoiden keskimääräinen kesto on noin kuusi minuuttia ja uusimpien 16,26 minuuttia. Informantti B:n videoilla on yhteensä lähes 30 000 000 katselukertaa.

Informantti C keskustelelee humoristisilla videoillaan erilaisista aiheista. Informantti C:n kanava sisältää yli 300 videota ja hänellä on lähes 150 000 tilaajalla. Vanhimmat videot ovat vuodelta 2013 ja kestoiltaan keskimäärin 3,7 minuuttia, uusimmat puolestaan 42,2 minuuttia. Informantti C työskentelee lisäksi muiden media-alan tehtävien parissa. Informantti C:n videot ovat keränneet lähes 40 000 000 katselukertaa.

Informantti D:n ensimmäinen videojulkaisu on vuodelta 2013. Informantti D:n kanavalla on lähes 240 000 tilaajaa ja yhteensä noin 400 videota. Näyttökertoja hänen videonsa ovat keränneet yli 63 000 000. Vanhimpien videoiden kesto on noin kahdeksan minuuttia ja uusimpien 23,4 minuuttia.

Videoillaan hän kuvaa päivittäistä arkeaan. Myös informantti D työskentelee muissa media-alan tehtävissä videoiden tekemisen lisäksi.

Informantti E:n yli 57 000 tilaajan kanavallaan on lähes 14 000 000 katselukertaa. Kanavan vanhin video on julkaistu vuonna 2015. Yhteensä julkaistuja videoita kanavalla on noin 200. Vanhimpien videoiden keskimääräinen kesto on noin seitsemän minuuttia ja uusimpien noin 28 minuuttia. Informantti E julkaisee kanavallaan muun muassa kosmetiikka-aiheisia videoita. Videoiden teon lisäksi informantti E työskentelee yrittäjänä.

Kaikki tutkielmaan valitut sisällöntuottajat ovat edelleen aktiivisia kanavillaan. Heitä yhdistää muun muassa sama kohderyhmä, joka on nuorista aikuisista yli 30-vuotiaisiin. Kaikkien tutkielmassa tarkasteltavien sisällöntuottajien äidinkieli on suomi. Ikähaarukka sisällöntuottajien välillä on pieni ja heistä neljä viidestä asuu Uudenmaan alueella informantti B:tä lukuun ottamatta. Varhaisin videotuotanto on informanttien C ja D kanavilta vuodelta 2013. Mahdollista kuitenkin on, että sisällöntuottajat ovat julkaisseet videoita myös ennen vuotta 2013, mutta ne on nyt piilotettu katsojilta.

Puhekielessä yleisestä sanotaan sisällöntuottajien tekemän Youtube-alustalle *vlogeja*, joka on yksi tietokonevälitteisen viestinnän muoto. Tässä muodossa tyypillisesti yksi henkilö puhuu suoraan kameralle ilman muiden henkilöiden läsnäoloa, samalla toimintaansa videoiden (Frobenius 2011, 1).

Taulukkoon 1 on koottu sisällöntuottajien videoiden keskimääräiset kestot ja julkaisuvuodet.

<b>Informantit</b>	<b>Keskimääräinen videon kesto (vanhin — uusin tuotanto)</b>	<b>Julkaisuvuodet</b>
<b>Informantti A</b>	7,6 min — 30,7 min	2014–2020
<b>Informantti B</b>	6,0 min — 16,2 min	2016–2021
<b>Informantti C</b>	3,7 min — 42,2 min	2013–2021
<b>Informantti D</b>	8,0 min — 23,4 min	2013–2021
<b>Informantti E</b>	7,0 min — 28,0 min	2015–2021

Taulukko 1. Aineistossa käytetyt videojulkaisut

Seuraavassa luvussa 5 aineistoa analysoidaan esimerkein. Esimerkkiviite muodostuu kunkin videosisällöntuottajaan yhdistetystä kirjaimesta. Lisäksi kirjaimen lopussa oleva numero kertoo videon, josta esimerkki on otettu. Numeroin 1–4 merkityt videot ovat aineiston keruuhetkellä sisällöntuottajan vanhimmasta tuotannosta ja 5–8 uusimmasta. Esimerkiksi viitteissä ”A2” tarkoittaa

informantti A:n esimerkkivideota 2 vanhemmasta tuotannosta. Viitteissä ”C5” tarkoittaa informantti C:n esimerkkivideota 5 uudemmassa tuotannosta. Aineistoon valikoitujen videoiden lista on tarvittaessa saatavilla tutkielman tekijältä.

## 4 Tutkimusmenetelmät

### 4.1 Kvalitatiivinen ja kvantitatiivinen tutkimus

Kvalitatiiviset tutkimusmenetelmät ymmärretään tavallisimmin ei-tilastollisina menetelminä. Käytännössä ero kvalitatiivisten ja kvantitatiivisten menetelmien ero näkyy tavoissa, joilla menetelmien tieteellinen selitysvoima osoitetaan. Kvalitatiivisten menetelmien käyttäjät keskittävät huomionsa erityisesti aineistoon ja sen hankintaan, kun taas kvalitatiiviset menetelmät nojaavat muun muassa matemaattisiin keinoihin tieteellisyyden eksaktiuden osoittamiseksi. Kvalitatiiviset ja kvantitatiiviset menetelmät voi erottaa toisistaan myös se, miten tutkimusaineisto on ensi sijassa hankittu. Kvalitatiivisessa tutkimuksessa tutkijan oma osuus on aineiston keruussa erityisen tärkeä. (Grönfors, 2011 4–5, Rautiainen 2007.)

Kvantitatiivisten menetelmien käyttäjät edustavat usein niin sanottua objektiivista tutkimuslinjaa, jonka tavoitteena on totuus riippumatta käytetyistä menetelmistä tai tutkijasta itsestään. Kvalitatiivisen tutkimuksessa puolestaan tutkijaa pidetään tärkeänä – jopa tärkeimpänä – tutkimusvälineenä. Kyse onkin siitä, että kvantitatiivisten menetelmien hyödyntävät tutkimustulokset ovat useimmiten abstrahoitavissa, kun taas kvalitatiivinen tutkimus näkee usein objektiivisuuden ongelmallisena.

Aineistolähtöistä kvalitatiivista analyysia voidaan tarkastella eri tavoin. Perinteisesti ensin aineisto redusoidaan eli pelkistetään, sitten se ryhmitellään ja lopuksi abstrahoidaan. Redusointivaiheessa informaatiota joko tiivistetään, pilkotaan osiin tai siitä karsitaan pois tutkimuksen kannalta epäolennaiset seikat. Aineiston ryhmittelyvaiheessa tutkitaan kerättyjä ilmaisuja ja etsitään samankaltaisuuksia ja/tai eroavaisuuksia. Tulokset ryhmitellään ja niille pyritään löytämään sisältöä kuvaavat nimet. Luokitteluyksikkönä voi olla esimerkiksi tutkittavan ilmiön piirre tai ominaisuus. Tarkoituksena on, että luokittelulla saadaan tiivistettyä aineiston sisältöä. Ryhmittely luo pohjan tutkimuksen perusrakenteelle. Lopuksi abstrahoinnilla pyritään pääsemään johtopäätöksiin ja teoreettisiin käsitteisiin. (Miles & Hubermann 1994.)

Perusanalyysimenetelmä kvalitatiivisen tutkimuksen perinteissä on sisällönanalyysi. Useimmat eri nimillä kulkevat kvalitatiivisen tutkimuksen analyysimenetelmät perustuvat ainakin osittain sisällönanalyysiin, jos sillä tarkoitetaan kirjoitettujen, kuultujen tai nähtyjen sisältöjen analyysia väljänä teoreettisena kehyksenä. (Tuomi & Sarajärvi 2009, 91.) Sisällönanalyysissa tutkitaan yleisesti kielellistä aineistoa, josta tutkija pyrkii erilaisten sisällöllisten luokittelujen avulla analysoimaan tutkittavaan ilmiöön liittyviä rakenteita ja sisältöjä. Sisällönanalyysi luokitellaan lähinnä

kvalitatiiviseksi aineiston analyysimenetelmäksi, jolla voidaan kuvata kirjoitetun ja puhutun kielen muotoja ja sisältöjä. Tyypillisiä aineistoja sisällönanalyysille ovat auditiiviset, visuaaliset ja kirjalliset dokumentit. Aineistoa tarkastellaan tekstin näkökulmasta, sillä useimmat aineistot ovat saatettavissa tekstin muotoon (esimerkiksi litteroidut haastattelut). (Seitanmaa-Hakkarainen 2014, Chi 1997.)

Kvalitatiivisessa sisällönanalyysissa tutkija on kiinnostunut erityisesti tutkittavaan ilmiöön liittyvistä sisällöllisistä merkityksistä. Painopiste ei välttämättä ole sisältöjen esiintymistiheydessä. Aineiston analyysi ei ole tutkimusprosessin viimeinen vaihe, vaan aineistoa tarkastellaan ja analysoidaan jo sen keruuvaiheessa. Tämä tarkoittaakin sitä, tutkimus itsessään on luonteeltaan syklistä. Luokittelukategoriat kehittyvät ja muuttuvat analyysiprosessin aikana jatkuvasti. (Seitanmaa-Hakkarainen 2014, Chi 1997.) Analyysimenetelmällä pyritään saamaan tutkittavasta ilmiöstä kuvaus tiivistetyksi ja yleisesti ilmaistuna (Tuomi & Sarajarvi 2009, 103).

Tutkimusmetodina sisällönanalyysi on joustava, sillä vaikka analyysiprosessi on luonteeltaan systemaattista ja koko aineiston kattavaa, ei etukäteen määriteltyjä luokitteluja tarvitse noudattaa tiukasti. Uusia analyysiluokkia voi syntyä analyysin edetessä ja aineiston keruun vuorovaikutuksessa. Käytännössä luokittelun toteuttaminen perustuu havaintojen vertailuun. Huomionarvoista on, että luonteeltaan kvalitatiivinen aineiston analyysi on hyvin eklektistä, koska ei ole olemassa yhtä ainoaa oikeaa tapaa tehdä luokittelukategorioita. Toiseen aineistoon kehitetty luokittelujärjestelmä ei välttämättä sovi suoraan toisenlaiseen aineistoon. (Seitanmaa-Hakkarainen 2014; Chi 1997; Tuomi & Sarajarvi 2009, 93.)

Grönforsin (2011, 5) mukaan varsin usein näitä kahta menetelmää käytetään myös yhdessä, kuten tässäkin tutkielmassa. Vaikka tutkielman aineiston keruu ja analysointi on pitkälti kvalitatiivista, tutkielmassa hyödynnetään myös kvantitatiivista menetelmää silloin, kun tarkastellaan ja lasketaan koodinvaihtotapauksien määriä. Kvantitatiivinen lähestyminen jää tässä tutkielmassa suhteellisen vähäiseksi, sillä aineiston määrälliset ominaisuudet ovat pieniä. Tutkimustuloksista ei ole mahdollista saada yksiselitteisesti vertailukelpoisia tai yleistettävissä olevia päätelmiä. Tällöin tuloksia ei voida eksplisiittisesti abstrahoida, mikä nostaa kvalitatiiviset analyysimenetelmät keskiöön.

Tässä tutkielmassa aineiston redusointi tapahtuu siten, että jo sisällöntuottajien videoita litteroidessa tuotos aineistoksi pidetään pelkistettynä. Käytännössä se tarkoittaa sitä, että videoilta litteroidaan vain ne kohdat, joissa sisällöntuottajan puheessa esiintyy koodinvaihtoa. Videoiden sisältämää puhetta ei siis ole litteroitu sanasta sanaan kokonaisuudessaan eikä litteraateista ilmene litteroinnille tyypillisiä taukoja, intonaatioita tai katkoksia. Edellä mainitut piirteet on jätetty litteroinnin ulkopuolelle siksi, että kyseiset piirteet eivät tuo lisäarvoa tämän tutkielman tutkimuskysymyksiin tai tuloksiin. Tutkimuksen kannalta on relevanttia tarkastella sitä, millaisissa asiayhteyksissä koodinvaihtoa

esiintyy, sekä millaisia ilmaisuja sisällöntuottajat toistuvasti käyttävät englannin kielellä ja kuinka usein. Samankaltaisuuksia ja eroavaisuuksia tarkastellaan vertailemalla aineistoa ja eri sisällöntuottajia. Ryhmittelyyn ja luokitteluun palataan tarkemmin luvussa 5 ja päätelmiin luvussa 6.

#### 4.2 Yksilöllisen variaation tarkastelu

Kuten aiemmin on osoitettu, variaation tutkimuksen teoreettisena taustana on Suomessa pääsääntöisesti tarkasteltu sosiaalisten muuttujien kontekstissa. Informantteja on analysoitu ryhmittäin, jotka useimmiten ovat olleet vanhat, työikäiset ja nuoret henkilöt. Usein kielen variaatiota tutkittaessa näkökulmaksi valikoituu esimerkiksi alueellinen, ajallinen, yksilöllinen, sosiaalinen tai tilanteinen vaihtelu (ks. esim. Lappalainen 2004). Tässä tutkielmassa informanttien ikähaarukka on pieni, joten sitä ei käytetä variaatiota selittävänä tekijänä. Lisäksi neljä viidestä sisällöntuottajasta asuu samalla alueella, joten alueellista variaatiota ei ole mielekästä tarkastella variaation pääasiallisena selittäjänä.

Sosiolingvistiikassa keskeisenä kiinnostuksen kohteena on ollut kielessä ilmenevä vaihtelu. Tapoja luokitella kielenulkoisten tekijöiden vaihtelua on kuitenkin olemassa useita. M. A. K. Halliday (1978) erottaa tilannekohtaisen (diatyyppisen) ja käyttäjäkohtaisen (dialektaalisen) vaihtelun toisistaan. Tilannekohtaisessa vaihtelussa huomio kiinnittyy siihen, miten ihminen puhuu erilaisissa tilanteissa. Käyttäjäkohtaista vaihtelua hän perustelee siten, että ihmisten puhuvat eroavat toisistaan sen perusteella, keitä he ovat. Tällaisissa tapauksissa selittävinä tekijöinä pidetään sosiaalisia muuttujia. (Halliday 1978.)

Siinä missä sosiolingvistiikka on karkeasti jaettu mikro- ja makrososiolingvistiikkaan, koskettaa tämä tutkielma molempia suuntauksia. Mikrososiolingvistiikan osalta aineistosta tarkastellaan muun muassa sitä, miten sosiaaliset rakenteet ovat voineet vaikuttaa puhetapoihin. Makrososiolingvistiikka kytkeytyy aineiston analyysin siten, että aineistosta tarkastellaan kielen vaihtoa ja kielellisiä asenteita esimerkiksi silloin, kun pohditaan koodinvaihdon osuutta asiayhteydessään.

Käyttäjäkohtaista variaatiota tutkittaessa perusoletus on, että yhteiskunnassa toimivalle ryhmälle on taipumus kehittyä sille tunnistettavia kielellisiä piirteitä. Tällaiset piirteet voivat erottaa ryhmän muista sosiaalisista ryhmistä. Yksilötasolla tämä tarkoittaa sitä, että puhujat mukauttavat puhetapaansa niiden puhujien puhetapojen suuntaan, joiden kanssa yksilö on säännöllisesti tekemisissä. (Ferguson 1994, 18.) Videoilta analysoitu puhekieli on keskustelututkimuksen metodin



suosimaa siten, ettei aineistoa ole tuotettu tutkimustarpeisiin, joten videoiden sisältöä voidaan kuvata autenttiseksi. Videot ovat siis olemassa ilman tutkimustakin, mihin keskusteluanalyysi pyrkii aineiston muodostamisessa.

Tässä tutkielmassa tarkastellaan sisällöntuottajien puhekielessä esiintyvää koodinvaihtoa yksilön näkökulmasta. Heistä jokainen on julkaissut kanavillaan sisältöä vaihtelevan määrin ja eri aikoina. Sisällöntuottajat ovat itse kuvanneet, editoineet ja julkaisseet videot kanaviinsa. He ovat siis itse päättäneet, mitä lopulliseen julkaistavaan videoon on jäänyt katsojan nähtäväksi ja kuultavaksi, joten videoissa käytetty puhekieli on mitä luultavimmin ollut tietoista. Jotta yksilöllisen variaation merkitystä voisi tulkita, sitä on kuitenkin suhteutettava toisiin yksilöihin (Lappalainen 2004, 350).

Tutkielmassa ajallisen variaation tarkastelu tapahtuu siten, että aineistoa vertaillaan sisällöntuottajan varhaisesta tuotannosta uusimpaan. Aineistosta tarkastellaan sitä, kuinka paljon koodinvaihtoa esiintyy ja onko se mahdollisesti lisääntynyt tai vähentynyt sisällöntuottajien puheessa.

Useissa tutkimuksissa on havaittu ammatin korreloivan parhaiten kielellisen variaation kanssa (Lappalainen 2004, 39). Tässä tutkielmassa sosiaalista variaatiota tarkastellessa otetaan huomioon se, että sisällöntuottajilla on kaikilla enemmän tai vähemmän saman ammatin piiriin kuuluvia työtehtäviä. He tuottavat digitaalista sisältöä Youtuben lisäksi muihin sosiaalisen median kanaviin, kuten Instagramiin. Näiden lisäksi osalla heistä on myös toinen työ. Sisällöntuottajien sosiaaliset taustatekijät esiteltiin luvussa 3.

Tässä tutkielmassa tarkastellaan kielenkäyttäjän koodinvaihdon, identiteetin ja idiolektin suhdetta. Identiteetin muodostuminen on diskursiivista toimintaa. Tämä tarkoittaa sitä, että ihmiset määrittelevät oman ominaislaatunsa ja kuulumisensa johonkin tiettyyn ryhmään diskursseina, teksteinä, puheina ja niiden joukkoina. Identiteetti syntyy siten, että ihmiset omaksuvat tietyt ominaislaatuun määrittelevät seikat muun muassa teksteistä, joita heille opetetaan koulussa ja joita he lukevat mediasta. (Mantila 2004.)

Mantilan (2004) mukaan voidaan ajatella, että erilaiset puhekielen piirteet ovat leimautuneet suomalaisessa puheyhteisössä tietyllä tavalla siten, että niihin voidaan liittää erilaisia kulttuurisia merkityksiä. Näitä piirteitä valitsemalla yksilö voi ilmentää tietynlaista identiteettiä. Identiteetit ovat jatkuvassa muutoksessa ja niitä rakennetaan diskursseissa koko ajan. (Mantila 2004.)

Ihminen identifioituu oman kielenkäyttönsä kautta yhteisöön ja vahva kielellinen identiteetti luo persoonallisuutta. Kielelliset valinnat voivat olla sekä tiedostamattomia että tietoisia. Harmasen (2011) mukaan kieli-identiteetin muotoutuminen alkaa kouluopetuksesta. Tavallista on, että

nykynuori voi oman äidinkiелensä rinnalla käyttää esimerkiksi englantia tunteiden tai ajattelun kielenä, mitä perinteisesti on pidetty äidinkiелisen puhujan kielenkäyttönä. (Harmanen 2011.)

Idiolekti, puhujan henkilökohtainen kielimuoto, vakiintuu aikuisikään mennessä ainakin osittain. Aikoinaan on oletettu, että idiolekti säilyisi pitkälti ennallaan puhujan loppuelämän hänen saavuttaessaan aikuisiän. Kuitenkin tuoreempi tutkimus osoittaa, että idiolektit harvoin pysyvät ennallaan, vaan ne voivat muuttua eri suuntiin. Esimerkiksi toiset puhujat murteellistuvat ja toiset yleiskielistyvät. (Lappalainen, Mustanoja & O'Dell 2019, 550.)

Sosiolingvistiikassa idiolektin käsitteellä tarkoitetaan yksilöllistä yhden puhujayksilön kokonaista, epätarkkarajaista ja muutosaltista kielimuotoa. Idiolektiin kuuluvat kaikki ne kielenkäytön piirteet, jotka ovat saman kielen puhujille ja puheyhteisön jäsenille yhteisiä. Lisäksi persoonallinen variaatio ja ominaisuudet, jotka kuuluvat puhujan henkilökohtaiseen kielelliseen ohjelmistoon, ovat idiolektin osia. (Lappalainen, Mustanoja & O'Dell 2019, 552.)

Identiteettiä ja erityisesti idiolektia tarkastellaan tässä tutkielmassa muuttuvaisina kategorioina yksilöllisestä näkökulmasta. Identiteettiä ja idiolektia tarkastellessa pyritään havainnoimaan koodinvaihdon kuulumista tai sen kuulumattomuutta yksilön henkilökohtaisiin kielenkäytön piirteisiin. Ulkopuolisena tutkijana on yksilön identiteetin kuvaileminen suorastaan mahdotonta, mutta käsite on relevantti silloin, kun tarkastellaan yksilön kieli-identiteettiä ja siihen mahdollisesti vaikuttavia tekijöitä.

#### 4.3 Koodinvaihtotapausten tarkastelu

Tässä tutkielmassa tarkastellaan koodinvaihtotapauksia, joissa kieli vaihtuu matriisikielystä eli suomesta upotettuun kieleen eli englantiin. Koodinvaihtoa tarkastellaan erilaisten luokittelukategorioiden kautta, sillä itse koodinvaihtoakaan ei voi tyhjentävästi selittää yhdenlaisen ilmiön kuvailulla. Tutkielmassa kiinnitetään huomiota muun muassa koodinmuutoksiin ja koodinsekoituksiin. Tarkastelun kohteena ovat lisäksi spontaanisyntyiset lainat eli sellaiset ilmaisut, joissa sana äännetään upotetun kielen fonologian mukaisesti, mutta ilmaisu on muuten matriisikielen rakenteen mukainen. Integroituneita lainoja eli ilmaisuja, jotka ovat mukautuneet matriisikielen myös fonologisesti, on ajoittain kutsuttu myös nimityksellä *finlisch*. Tällaisissa lainoissa sekä kirjoitusasu että ääntäminen toteutuvat suomen kielen mukaisesti. Palmgren ehdottaa pro gradu -tutkimuksessaan, että *finlischin* voitaisiin mahdollisesti suhtautua jonkinlaisena kertakäyttökieleen, tilanteen synnyttämänä ”verkkokiелisenä” viestintäratkaisuna (Palmgren 2007.)

Aineistosta on ensin litteroitu sellaiset vuorot, joissa esiintyy koodinvaihtoa. Aineistoa tarkastellaan yksilöllisen variaation näkökulmasta Gumperzin koodinvaihtofunktioiden sekä Lauttamuksen (1992), että Kovácsin (2009) näkemysten pohjalta muodostettujen luokittelukategorioiden avulla luvussa 5.

Gumperz (1982) jakaa keskustelussa metaforisen koodinvaihdon funktion kuuteen eri kategoriaan:

1. Koodinvaihto siteeratessa tai lainatessa (*quotation*)
2. Koodinvaihto vastaanottajan spesifioinnin apuna (*addressee specification*)
3. Koodinvaihto interjektioiden yhteydessä ja täytesanoina (*interjections, sentence fillers*)
4. Koodinvaihto toiston keinona (*reiteration*)
5. Koodinvaihto viestin modifioinnin keinona (*message qualification*)
6. Koodinvaihto persoonallisen ja objektiivisen näkökulman vaihtelussa (*personalization versus objectivization*)

(Gumperz 1982, 75–82; Kalliokoski 1995, 5.)

Kun tarkastellaan siteeraamista ja lainaamista (1) koodinvaihdon yhteydessä, voidaan puhua myös referoinnista. Vastaanottajan spesifioinnilla (2) tarkoitetaan tilannetta, jossa puhuttelu kohdistetaan tiettyyn henkilöön tai asiaan. Tässä tutkielmassa tarkastellaan tapauksia, joissa vastaanottajaksi kohdistuu videon katsoja. Koodinvaihtoa voi esiintyä myös interjektioiden eli huudahdusten sekä täytesanojen (*niinku, tota, öö*) yhteydessä (3). Toistossa (4) sama ilmaus tai asia esitetään sekä matriisi- että upotetulla kielellä. Viestin modifioinnilla (5) tuodaan lisämerkitystä kieleen esimerkiksi tarkennuksena. Persoonallisen ja objektiivisen näkökulman vaihtelu (6) koodinvaihtotilanteessa on Gumperzin (1982) mukaan vaikeasti selitettävä funktio. Puheenaiheena on usein toiminta tai puhuminen, ja koodinvaihto toimii viestiä vahvistavana tekijänä. (Gumperz 1982: 75–82.)

Koodinvaihto ja lainaaminen voidaan Lauttamuksen (1992) mukaan asettaa neljän kategorian strukturaaliselle jatkumolle. Seuraavia kategorioita käytetään myös tässä tutkielmassa. Koodinmuutoksella (*code-change*) tarkoitetaan täysin upotetun kielen mukaisia ja yleensä monisanaisia keskustelun osia. Koodinsekoituksessa (*code-mix*) on puolestaan kyse yksittäisistä sanoista tai yhdyssanoista, jotka pääasiassa myös operoivat upotetun kielen konventioiden mukaisesti. Hassinen (2002, 407) esittää koodien yhdistymisen erityisesti pienen lapsen kaksikielisyyden kehittymisprosessiin. Spontaanisyntyiset lainat (*noance loans*) ovat ilmaisuja, joissa sana äännetään upotetun kielen fonologian mukaisesti, mutta se on muuten mukautunut matriisikielen rakenteeseen. Integroituneet lainat (*integrated loans*) ovat mukautuneet matriisikielen myös fonologisesti, mutta eri puhujien ääntämisessä voi esiintyä variaatiota. Lauttamuksen mallissa kielellistä ainesta tarkastellaan sen perusteella, kumman puhutun kielen kieliopillinen systeemi operoi kussakin tapauksessa. (Lauttamus 1992: 9–10.)

Tiettyyn keskustelutilanteeseen lainattuja sanoja (tilapäislainoja) on vaikeampi erottaa koodinvaihdosta kuin vakiintuneita lainoja. Tässä tutkielmassa luokitellaan lainaaminen ja koodinvaihto Kovácsin (2009) artikkelin tavoin seuraavasti: matriisikielen ja englannin esiintyessä vuorovaikutuksessa Kovács (2009) mainitsee artikkelissaan esimerkkinä sanan *baker*. Jos suomea ja englantia käytetään puheessa siten, että sana *baker* lausutaan englantilaisittain suomenkielisessä kontekstissa [beikə\*], pidetään sitä koodinvaihtona. Jos taas sanan ääntäminen on mukautettu suomen kielen mukaiseen fonologiaan [pääkäri], pidetään sitä lainana. Vakituksia lainoja ja koodinvaihtoja voidaan kuitenkin pitää saman jatkumon ääripäinä. (Kovács 2009, 24–25, Thomason & Kaufman 1988.) Koska lainaamista ja koodinvaihtoa ei voi yksiselitteisesti erottaa toisistaan, lasketaan tässä tutkielmassa lainaaminen koodinvaihtoilmiöksi.

Luvussa 5 esitetään esimerkkejä koodinvaihtotapauksista analyysin tueksi. Esimerkeissä suomenkieliset osuudet on kirjoitettu auki siten, kuin ne sisällöntuottajan puheessa on tuotettu. Englanninkieliset osuudet puolestaan on kirjoitettu esimerkkeihin englannin kielen kirjoitusasun, morfologian ja kieliopin mukaisesti. Ääntöasu esitetään erikseen silloin, kun esimerkissä fonologisella muodolla on merkitystä analyysin kannalta.

#### 4.4 Monikielisyys ja koodinvaihto

Tässä tutkielmassa koodinvaihtoa puhekielessä tulkittaessa otetaan huomioon monikielisyys-käsitteeseen liittyvät diskurssit, jotka eivät kuitenkaan poissulje toisiaan. Tällaisia samaan aihepiiriin liittyviä käsitteitä ovat muun muassa *moniäänisyys*, *kielten sekoittaminen*, *diglossia*, *kielenylitys* ja *kieleily*.

Suurin osa ihmisistä on monikielisiä ja joka kolmas ihminen puhuu päivittäin kahta tai useampaa kieltä. Vaikka Suomi on aina ollut monikielinen, perinteisesti kielen sekoittamista ja koodinvaihtoa on pidetty osoituksena kielitaidottomuudesta. Monikielisyys ja kielen sekoittuminen vuorovaikutuksessa on siis nähty ja nähdään usein edelleen ongelmana. Kuitenkin useamman kuin yhden kielen käyttöön samassa vuorovaikutustilanteessa suhtaudutaan eri tavoin erilaisissa yhteyksissä. Esimerkiksi sanomalehdissä kielen sekoittumista paheksutaan. (Kalliokoski 2009, 9–10.) Kielenylityksiä, koodin- tai rekisterinvaihtoa ja monikielisyyttä käsittelevät tutkimukset ovat osoittaneet, että erityisen runsaat resurssit kielellä leikittelyyn ja erilaisten identiteettien rakentamiseen tarjoaa monikielisyys (Halonen 2009, 329–330).

Ihmiskunnan enemmistön monikielisyys perustuu siihen, että eri kielten puhujat ovat keskenään vuorovaikutuksessa. Koodinvaihdosta on totuttu yleisimmin puhumaan kahdessa eri kontekstissa, jotka ovat saman kielen varieteetit ja eri kielten välinen vaihtelu. Olennaista kuitenkin on kahden tai useamman kielen kohtaaminen samassa kielenkäyttötilanteessa. (Kalliokoski 2009, 13.)

Koodinvaihtoa voidaan tarkastella myös moniäänisyyden eli *polyfonian* näkökulmasta. Moniäänisyyttä tarkasteltaessa mikään kielellinen ilmaisu ei synny tyhjästä, vaan puheessa on aina kaikuja toisista äänistä. Muhonen (2008) tarkastelee moniäänisyyttä toisen kontekstin eli toisen puhujan tai tahon sanojen, tekstin tai ajatuksien liittämisenä omaan puheeseen. (A. Muhonen 2008, 183.) Ihmisillä on käytössään useita erilaisia kielellisiä resursseja, joista puhuja voi ammentaa erilaisia aineksia vuorovaikutukseensa. Tällaisia resursseja ovat esimerkiksi ensikieli, mahdolliset myöhemmin omaksutut kielet, opiskellut kielet, erilaisten ryhmien ja kulttuurien kielet sekä kielimuotojen kaikki variantit. Näiden lisäksi resursseihin kuuluvat kieliin ja niiden piirteisiin liittyvät asenteet, stereotypiat ja toimintatyyppit. Erityisesti nuoret kielenkäyttäjät hyödyntävät jatkuvasti laajaa repertoaria heidän rakentaessaan identiteettiä ja kielellä leikitellessään. (Halonen 2009, 329.)

Puheen moniäänisyyttä tarkastellessa ja analysoidessa näkökulma on sekä sosiaalisessa identiteetissä ja sen korostamisessa että koodinvaihdon käsittämisenä puheen kontekstualisointikeinona. Yksilöllä on sosiaalisessa repertuaarissaan erilaisia identiteettejä sekä niitä vastaavia kielellisiä repertuaareja. Tilanteen muuttuessa yksilö voi koodia vaihtamalla vaihtaa myös sosiaalista identiteettiään. Jokainen tällainen muutos toimii kontekstivihjeenä, joka puolestaan muovaa puhetilannetta eli kontekstia. (Kalliokoski 1995, 14.) Moniäänisyyttä on tarkastelu kielentutkimuksessa lukuisista eri näkökulmista. Esimerkiksi Lappalainen (2004, 312) tarkastelee lausuman moniäänisyyttä puhujan roolien vaihteluna.

On kuitenkin huomattava, että koodinvaihto ei ole vain monikielisessä tai moniäänisessä kommunikaatiossa esiintyvä piirre. Jokaisessa vuorovaikutustilanteessa on samanaikaisesti mukana useita eri konteksteja. Koodinvaihto voi toimia keinona osoittaa muille osallistujille jokin uusi, relevantti konteksti. Moniäänisyys ja toisten puheisiin tai teksteihin viittaaminen ovat eri tavoilla painottuneena läsnä kaikessa kielenkäytössä. (Kalliokoski 1995, 19.)

Kielten sekoittaminen (*language mixing*) tarkoittaa kahden tai useamman kielen käyttöä yhden diskurssikonan sisällä siten, että ensisijaisesti kielten sekoittaminen on strategia luoda yhteistä kielenkäytön tyyliä. Samalla voidaan osoittaa ja rakentaa ryhmäidentiteettiä sekä sosiaalisia suhteita. Kielen sekoittaminen voi tapahtua niin yksittäisten sanojen kuin lausekkeidenkin tasolla. (Auer 1998; Leppänen 2008.) Koodinvaihdon näkökulmasta sen merkityksellisyys on toisaalta sidoksissa myös sen frekvenssiin. Auer (1998) erottaa koodinvaihdon ja kielten sekoittamisen siten,

että kielten sekoittamisessa kaksi koodia vaihtelee niin tiheästi, ettei yksittäisellä vaihdoksella ole enää merkitystä. Tällöin frekventtiys heikentää koodinvaihdon mahdollisuutta toimia paikallisesti merkitsevänä kontekstualisointivihjeenä. (Auer 1998, 314–321.)

Kielten sekoittamisessa voi olla kyse kielten välisyydestä (*interlingualism*). Siinä, missä kaksikielisyys tarkoittaa siirtymistä yhdestä kielestä toiseen, kielten välisyydessä kielet eivät aina vaihtele ja aktivoidu sekventiaalisesti. Kielet voivat olla kielenkäyttäjälle yhtä aikaa mahdollisia ilmauksia. Kielenkäyttäjä ei välttämättä edes koe käyttävänsä ilmauksia, jotka ovat erikielisiä. Sen sijaan erikieliset ilmaukset ovat käyttäjälleen kielellinen varanto, joka hänen senhetkisessä käyttökontekstissaan koetaan relevanteiksi ja asianmukaisiksi. (Leppänen 2008, 213.)

Gumperz (2001) käsittelee diglossiaa tilanteisena koodinvaihtona. Diglossiassa kielen valinta määräytyy tilanteen mukaan, mutta monikielisten yhteisöjen vuorovaikutuksessa kaksi eri kieltä voivat toimia rinnan ilman, että koodinvaihtoa voisi luonnehtia kommunikaatiostrategiaksi. Poplack, Wheeler ja Westwood (1990, 188–195) puolestaan pitävät diglossiaa monikielisen yhteisön kommunikaation itsenäisenä kielimuotona, jossa kyky vaihtaa kieltä kesken keskustelun on osa puhujien kompetenssia. Kauhasen (2009) mukaan diglossia on yleensä määritelty kielitilanteeksi, jossa kielimuotojen funktiot ovat selkeästi eriytyneet. Varieteetteja käytetään siis omissa kielenkäyttötilanteissaan. Tämän tulkinnan mukaan diglossisessa yhteisössä ei esiintyisi koodien sekoittumista. Kauhanen (2009) tosin kyseenalaistaa edellä mainitun tiukan rajanvedon. (Kauhanen 2009, 97.)

Varsinkin nuorison eri alakulttuurien kielenkäytön tutkijat ovat havainneet, miten koodinvaihto nousee vuorovaikutuksessa huomion kohteeksi ja miten kielten vaihtelu ja tyyllittely palvelevat toisiaan. Kielenylityksessä (*crossing*) korostuu kielenkäytön performatiivinen ulottuvuus ja huomio kohdistuu kielelliseen ilmaukseen sosiaalisen identiteetin rakentajana. (Kalliokoski 2009, 13.)

Monikielisuuden käsitettä ei koodinvaihdon tapaan voi luonnehtia yksiselitteisesti. Lehtonen (2015) on tunnistanut moni-alkuisten käsitteiden ongelmallisuuden. Lehtonen (2015) muistuttaa, ettei ole erikseen olemassa monokulttuuria ja monikulttuuria tai yksikielisyyttä ja monikielisyyttä, sillä ihmiset ovat kaikki osaltaan kulttuurisia ja kielisiä. Sociolinguistiikassa monikielisuuden käsite on yhdistetty *kieleilyyn* (*translanguaging*). Kieleily-ajattelussa lähdetään siinä, että kielenkäyttäjät eivät ensisijaisesti ole esimerkiksi suomen- tai englanninkielisiä. Sen sijaan he ovat kieleilijöitä, jotka käyttävät ympäristönsä tarjoamia kielellisiä resursseja. Tarve kommunikoida on ensisijainen, jonka jälkeen on kehittynyt tieto etnisistä ja kulttuurisista rajoista. Kieleily-ajattelun perusteella ihmiset eivät puhuisi kokonaisia kieliä, vaan ainoastaan osia kielistä. Osat eivät yhdisty vain perinteisesti eri

kieliksi miellettyihin rekistereihin, vaan palasiksi esimerkiksi stadin slangiin tai hip hop -kieleen. (Blommaert 2010; Lehtonen 2015.)

Tässä tutkielmassa kieleilyn käsite jää paitsioon, sillä lähtökohtana pidetään sitä, että tutkittavan sisällöntuottajan matriisikieli on suomi, ja englantia käytetään upotettua kielenä erilaisissa yhteyksissä. Tuloksia muodostettaessa ajatellaan, että informantit ovat äidinkieleltään suomenkielisiä. Toisaalta päätelmiä pohdittaessa huomioidaan, että upotetulla kielellä esitetyt ilmaukset voivat olla tulosta eräänlaisesta internetkielestä. Tämä tukisi kieleily-ajatusta siten, että upotetun kielen mukaiset ilmaukset olisivat ympäristön tarjoaman kielellisten resurssien tulosta. Kokonaisuudessaan kieleily-periaate ei kuitenkaan tässä tutkielmassa ole variaation selittäjä.

#### 4.5 Englannin kielen asema ja määrittely

Tässä tutkielmassa tarkastellaan upotetun kielen eli englannin osuutta suomenkielisessä puheessa. Nuorisomedia on olennainen osa sitä populaarikulttuuria, jossa englannin kielen rooli on näkyvä. Auerin (1995) mukaan kielen vaihtelu (*language alternation*) viittaa kaikentyypisiin kielen esiintymiin, joissa kahta eri kieltä käytetään samanaikaisesti samassa viestintäjaksossa. Kielen vaihtelu voi esiintyä samassa puheenvuorossa, rajoittua yhteen sanaan tai muuttaa koko keskustelun kielen. Kyse on paikallisesta kielen vaihtelusta silloin, kun puhuja ei viittaa toisen kielen vierauteen tai siihen, että käytössä on kaksi eri koodia eli kieltä. (Auer 1995.)

Suomessa vuonna 2012 vähintään yhtä vierasta kieltä ilmoitti ainakin jonkin verran osaavansa noin 94 prosenttia 18–64-vuotiaasta väestöstä. Yli 80 prosenttia suomalaisista ilmoitti osaavansa vähintään kahta ja puolet vähintään kolmea vierasta kieltä. Yleisimmät Suomessa asuvan aikuisväestön hallitsemat kielet suomen lisäksi ovat englanti ja ruotsi. Englannin kielen keskeinen asema sähköisessä mediassa populaari- ja nuorisokulttuurissa nostaa nuorten positiivista asennetta englannin kieltä kohtaan. Englannin kielen taitamisen merkitys on edelleen kasvussa erityisesti työelämässä. (Opetus- ja kulttuuriministeriö 2017.)

Jyväskylän yliopistossa tehdyssä tutkimuksessa (Leppänen ym. 2009) englannin kielen asemasta Suomessa osallistuneilta kysyttiin heidän suhtautumistaan siihen, kuinka todennäköisesti englanti on yksi Suomen virallisista kielistä 20 vuoden kuluttua. Kaikkiaan yli puolet vastaajista piti tätä joko erittäin tai melko todennäköisenä. Enemmistö vastaajista (90 %) oli myös sitä mieltä, että englannin merkitys kasvaa entisestään tulevien 20 vuoden aikana. Lähes puolet (47 %) uskoi, että

tulevaisuudessa kaikkien on osattava englantia. (Leppänen, Pitkänen-Huhta, Nikula, Kytölä, Törmäkangas, Nissinen, Kääntä, Virkkula, Laitinen, Pahta, Koskela, Lähdesmäki, Jousmäki 2009.)

Leppäsen ja Nikulan (2008) mukaan ääriesimerkki englannin vaikutuksesta on se, että tällä hetkellä suomen ja ruotsin kielten ohella käytetään englantia jopa sellaisissa tilanteissa, joissa voisi yhtä hyvin käyttää äidinkieltään. Toisin kuin monien uusien vähemmistökielten kohdalla englannin lisääntynyt käyttö ei johdu englantia puhuvien maahanmuuttajien määrän kasvusta. Merkityksen kasvu on monen tekijän summa. Muun muassa yhteiskunnan rakennemuutos, kaupungistuminen, kansainvälistyminen ja uudet viestintäkanavat ovat vahvistaneet englannin asemaa. (Leppänen & Nikula 2008, 13, 16.)

Leppäsen ja Nikulan (2008) mukaan englannin kielen lisääntynyt vaikutus on jo 1960-luvulta lähtien näkynyt englanninkielisten ilmausten yleistymisenä esimerkiksi nuorisomediassa ja mainonnassa. Myös televisio ja elokuvat ovat tuoneet osansa englannin kielestä suomalaiseen arkeen, koska Suomessa ei ole ollut tapana jälkiäänittää vieraskielisiä ohjelmia. Viime vuosikymmenenä on englannin rooli kasvanut mediassa entisestäänkin, kun sitä on alettu käyttää pääasiallisena kielenä esimerkiksi median konteksteissa.

Englannin kielen merkityksen muotoutumiseen on osaltaan vaikuttanut myös Suomen homogeenisyys kaksikielisyystään huolimatta. Suomalaisille yhteiskunnan kielellinen perusta on suurimmalla osalla ollut yksikielisyys, eli yhdellä kielellä on pärjätty kaikilla elämänalueilla. Leppäsen ja Nikulan (2008) mukaan voidaan siis todeta, että lähes yksikielisellä kansakunnalla ei ole ollut tarvetta *lingua francalle*. (Leppänen & Nikula 2008, 21.)

Suomalaiset eivät siis käytä englantia systemaattisissa tilanteissa, vaan käytön ja tavan ehdot vaihtelevat paljonkin toisistaan. Sitä, miksi englanti jossain tilanteissa valitaan viestinnän kieleksi tai otetaan käyttöön äidinkielen rinnalle, on mahdotonta kuvailla yksitasoisena ilmiönä. Leppänen ja Nikula (2008) kuvailevat keskeiseksi kontaktitilannetyypiksi sellaista, jossa suomen ja englannin kielten vaihtelu on tavallista osoituksena kielenkäyttäjien kaksi- tai monikielisyystään. Ilmiöllä on myös sosiaalinen ulottuvuus. Tällaisella – joskus systemaattisellakin – koodinvaihdolla on monia vaihtelevia tehtäviä. Koodinvaihto voi toimia yhtenä ainesosana sekakielessä, jolla jokin sosiaalinen ryhmä rakentaa yhteistä kielenkäytön tyyliään. (Leppänen & Nikula 2008, 22–23.)

Englannin kielen positiivista tai negatiivista vaikutusta suomen kieleen ei voi ratkaista vetoamalla kielenhuollon yleisiin periaatteisiin. Kysymystä on lähestyttävä yksityisen kielenkäyttäjän näkökulmasta sosiolingvivistisesti. Kielten vuorovaikutus on luonnollista muun muassa siitä syystä, että monet vieraat ainekset omaksutaan kieleen lainaamisella. Lainaminen on helpoin tapa nimetä uusia käsitteitä. Kun nimeämisen tarve on samanaikaista useilla ihmisillä, lainattu aines yleistyy ja vakiintuu nopeasti. Englannin kielen vaikutus on kaksitasoista; kielenkäyttäjä omaksuu tietoisesti tai



tarkoituksellisesti englannin kielestä peräisin olevia sanoja tai ilmausten suoria käännösvastineita. Toisaalta hän omaksuu sellaisia muotoja tai rakenteita, joita hän ei suoraan tunnista vieraan kielen aineksiksi. (Sajavaara & Lehtonen 1981, 289–291.) Tässä tutkielmassa kiinnitetäänkin huomiota englannin kielen yleistyvään asemaan ja sen myötä mahdollisiin vaikutuksiin puhekielessä.

Hiidenmaan (2003) mukaan vieraan sanan mukauttaminen on yleistyvää ilmiötä. Mukauttamisen seurauksena sana taivutetaan ja pitkälti äännetäänkin suomen sanojen kaltaisesti. Kun vieraiden kielten taito paranee, alkuperäinen kirjoitusasu alkaa tuntua luontevammalta. Uudet lainat muodostavat uusia sanavartaloita, joista syntyy suomalaisia johdoksia (*softa – soft ware, longboardaus*). (Hiidenmaa 2003, 90–96.) Tätä ilmiötä tarkastellaan tutkielman aineistossa.

Aikaisemmin on osoitettu, ettei koodinvaihtoa ja lainaamista voi yksiselitteisesti erottaa toisistaan. Tässä tutkielmassa tarkastellaan koodinvaihtoa erilaisten luokittelukategorioiden kautta, koska koodinvaihtoakaan ei voi tyhjentävästi selittää yhdenlaisen ilmiön kuvailulla. Kyseiset luokittelukategoriat pohjautuvat Gumperzin (1892), Lauttamuksen (1992) ja Kovácsin (2009) näkemyksiin, jotka on esitelty luvussa 4.2

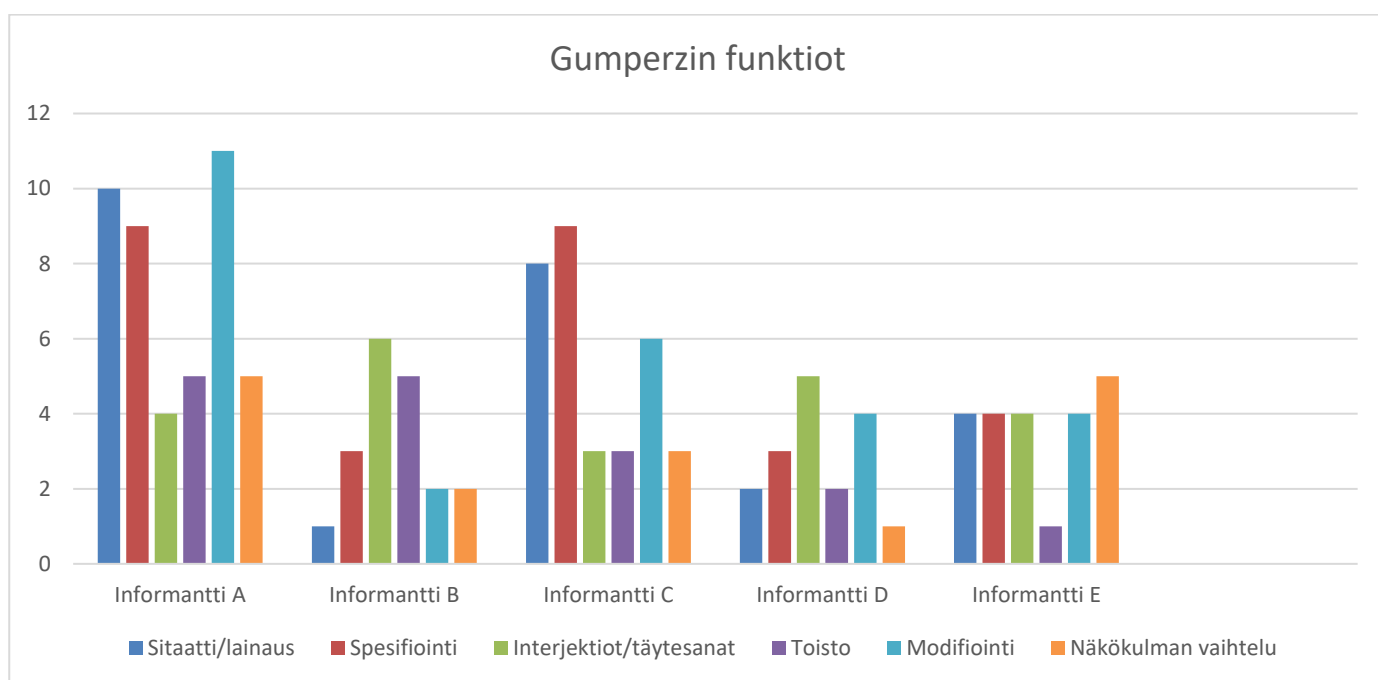
Tutkielmassa kiinnitetään huomiota muun muassa koodinmuutoksiin ja koodinsekoitukseen. Tarkastelun kohteena ovat lisäksi spontaanisyntyiset lainat eli sellaiset ilmaisut, joissa sana äännetään upotetun kielen fonologian mukaisesti, mutta ilmaisu on muuten matriisikielen rakenteen mukainen. Integroituneita lainoja eli ilmaisuja, jotka ovat mukautuneet matriisikielen myös fonologisesti, on ajoittain kutsuttu myös nimityksellä *finenglish*. Tällaisissa lainoissa sekä kirjoitusasu että ääntäminen toteutuvat suomen kielen mukaisesti. Palmgren ehdottaa pro gradu -tutkimuksessaan, että *finenglishiin* voitaisiin mahdollisesti suhtautua jonkinlaisena kertakäyttökielenä, tilanteen synnyttämänä ”verkkokielenä” viestintäratkaisuna (Palmgren 2007.)

## 5 Koodinvaihto puhutussa kielessä

### 5.1 Yksilöllinen variaatio ja koodinvaihtofunktiot

Aineistosta on laskettu koodinvaihtotapaukset siten, että yhden sanan mittaiset koodinvaihdot on laskettu yhdeksi ilmentymäksi. Useampisanaiset koodinvaihdot, jotka aineistossa ovat useimmiten sanontoja, toteamuksia tai kysymyksiä, lasketaan nekin yhdeksi tapaukseksi. Yksisanaiset koodinvaihtotapaukset ovat aineistossa yleisimmin adjektiiveja (*nice*) tai partikkeleita (*anyway*). Yksisanaisina esiintyvänä koodinvaihtoja ilmenee aineistossa kuitenkin verrattain vähän.

Seuraavassa taulukossa esitetään sisällöntuottajien koodinvaihdon ilmeneminen Gumperzin (1982) muodostamien funktioiden avulla. Funktiot on esitelty tarkemmin luvussa 4.3.



Kaavio 1. Koodinvaihto ja Gumperzin funktiot

Kaaviosta 1 ilmenee, että yksilöllisestä näkökulmasta koodinvaihtoa Gumperzin funktioiden mukaan esiintyy määrällisesti eniten informantti A:lla. A:lla modifiointi (11 tapausta) sekä siteeraus ja lainaaminen (10 tapausta) ovat funktioista yleisimmin esiintyvät. Suorasti tiedettävää toisen henkilön siteeraamista ei aineistosta ilmene, vaan informantti A:n kohdalla koodinvaihtoa esiintyy erityisesti

silloin, kun käytetään jotakin kiteytyynyttä sanontaa tai ilmausta. Nämä tapaukset olen luokitellut lainaamiseksi, joka liitetään Gumperzin funktioon ”siteeraaminen ja lainaaminen” (esimerkit 1–3).

- 1) A5: **I wanna thank not only god but jesus** koska katsokaa -- ni must tuntuu et se antaa meille sellasta **peace of mind**.
- 2) A6: **On the other hand** mulla on semmoinen fiilis että mä haluisin jotain silleen uutta inspiroivaa.
- 3) A6: -- mieltä että aina pitää kokeilla kaikkea yrittää rohkeesti ja testata kaikkea uutta, en tiää **you never know, you never know**.

Modifioinniksi on luokiteltu tapaukset, joissa puhuja tuo koodinvaihdossa jonkin lisämerkityksen puhuteltavana olevaan asiaan, kuten esimerkiksi tarkennuksena (4). Informantti A:n puheessa modifiointia esiintyy erityisesti silloin, kun ilmaistaan omaa suhtautumista tai mielipidettä puheena olevaan asiaan. Modifiointia ilmenee ajoittain myös tarkennuksen tehtävissä, kuten esimerkissä 5:

- 4) A6: -- tai ihan muuten vaan haluan olla niin kuin **freshed up** niin näytän et millasil tuotteil mä sen teen, siin -- mulla on siit tosi niin kuin **blessed** ja hyvä fiilis.
- 5) A2: Täs mä heingailin mun **girlfriend** Susannan kanssa, oli pakko laittaa tää yhteiskuva tähän koska me näytetään niin hyviltä.

Spesifiointia koodinvaihdon yhteydessä (9 tapausta) tapahtuu hieman enemmän kuin toistoa ja näkökulman vaihtelua (5 tapausta). Huomattavaa kuitenkin on, että kvantitatiiviset tulokset eivät ole tässä yhteydessä merkittäviä. Spesifioinnilla tarkoitetaan aineiston kannalta sitä, että sisällöntuottaja puhuttelee suoraan katsojaa tai on kohdistanut puheensa katsojalle upotetulla kielellä (esimerkit 6–8):

- 6) A2: -- joten mä päätin että mä haluan esittää teille jotain ihan super privaattia. **You know what I mean**.
- 7) A5: Tää on mun elämän paras päivä koska meille tuli vappumunkkilähetys. **Name a better** lähetys. **I'll wait**.
- 8) A6: -- niin olis myös tarkoitus viedä nää huomenna **wish me luck** kommentteissa mä tulen tarvitsemaan sitä.

Aineistoni tulokset poikkeavat Muhosen (2019) pro gradu -tutkielman tuloksista, joissa hän totesi videoblogiaineistossaan suurimman osan koodinvaihdosta osoitetuksi videoblogin katsojille. Hän havaitsi, että katsojille osoitetut koodinvaihdot ovat yleensä lyhyitä lausekkeita, yksittäisiä sanoja tai tervehdyksiä. (Muhonen 2019, 36.) Yleisesti tässä aineistossa katsojillekin suunnatut

koodinvaihtokohdat ovat usein monisanaisia lausumia. Lisäksi koodinvaihto on sisällöntuottajilla useimmiten kytköksissä esimerkiksi mielipiteen ilmaisemiseen tai johonkin asiaan asennoitumiseen eikä niinkään suoraan osoitettavissa katsojaan.

Toistoa on tarkasteltu tapauksista, joissa sisällöntuottaja ilmaisee saman asian sekä matriisi- että upotetulla kielellä. Informantti A:lla toistoa ilmenee tapauksissa, joissa hän esimerkiksi lukee ääneen englanninkielistä tekstiä ja suomentaa sen myöhemmin (esimerkit 9 ja 10). Toistoa esiintyy myös muutaman kerran silloin, kun jotakin asiaa kuvaillaan adjektiivein sekä suomeksi että englanniksi (esimerkki 11). Esimerkissä 12 havaitaan täytesana, joita interjektioiden lisäksi esiintyy puheessa vähiten (4 tapausta).

9) A6: **Super easy superhelppo** tuote mä usein saatan levittää tämän ihan vaan niin kuin käsin etenkin --.

10) A7: Vau katsokaa miten ihanalt meidän kuusi näyttää tuolla taustalla paitsi se on täysin **crooked** näettekö toi jalka on mennyt rikki nii tää on ihan **vinossa**.

11) A7: Tää [takki] on itseasias H&M:n naisten osastolta ostettu ja tää on koko L. Tästä oli XL se oli ihan sika iso eli tää on tämmönen tosi **oversized**.

12) A1: -- **anyways, niin** jotka nyt on tullut mun ystäviksi vaikka tästä muutaman vuoden säteellä, niin ei välttämättä tiedä sitä että mä oon harrastanut sirkusta.

Kuten aikaisemmin on todettu, persoonallisen ja objektiivisen näkökulman vaihtelu koodinvaihtotilanteessa on vaikeasti selitettävä funktio. Puheenaiheena on usein toiminta tai puhuminen, ja koodinvaihto toimii viestiä vahvistavana tekijänä (esimerkki 13). Näissä tapauksissa on tarkasteltu sitä, tuoko koodinvaihto merkitykseen persoonallisen tai objektiivisen sävyn.

13) A5: Tääl on meidän sima. Paitsi toi pitää eka siivilöidä. Täs menee hermo. Hyvin menee, **you're doing amazing sweetie** [toteaa itselleen].

Määrällisesti toiseksi eniten koodinvaihtoa funktioittain tarkasteltuna esiintyy informantilla C. Myös hänellä kolme eniten toistuvaa koodinvaihtofunktiota ovat modifiointi, siteeraaminen ja lainaus sekä vastaanottajan spesifiointi. Informantti A:sta poiketen informantti C:llä koodinvaihdon suurin ilmaantuvuus on vastaanottajan spesifiointifunktiossa (9 tapausta), toiseksi suurin siteeraamisessa ja lainaamisessa (8 tapausta) ja kolmanneksi suurin viestin modifioinnissa (6 tapausta). Katsojille puhuttaessa informantti C viittaa yleensä yleisöön monikossa, kuten esimerkeistä 14–16 voi havaita:

14) C2: Mä dataan täs samalla mut te ette nää koska mul on koko ajan **poker face**.

15) C6: Tiettekö ku **legit** joskus kutosluokalla mä oon ladannu --.

16) C8: Nonii kattokaa **it's, it's happening**.

Informantti A:n tavoin lainaamista tapahtuu informantti C:lläkin kiteytyneissä ilmauksissa, mutta koodinvaihtoa ilmenee myös suorista sitaateista (esimerkit 17 ja 18.)

17) C2: [hyräillen] **I believe I can fly... I believe I can touch...**

18) C4: **Infoboxissa** se Marilyn Monroe sitaatti se **if you can't handle me at my worst you don't deserve me at my best**.

Modifiointia esiintyy informantti C:n puheessa koodinvaihdon kanssa erityisesti adjektiivien yhteydessä silloin, kun korostetaan omaa suhtautumista kyseiseen aiheeseen (esimerkit 19–21). Informantti A:n tavoin viestiä korostetaan ilmaisemalla suhtautuminen koodinvaihdolla:

19) C4: Ja sit tää on niinku mun ihan **ultimate**-viha.

20) C5: Nyt on vielä **extra level** siihen. Onks tää jotenkin tämä koronakeväästä puhuminen ihan **done**.

21) C8: Ensimmäinen kuva oli löyhästi sanottuna **succee**.

Toistoa (esimerkki 22), täytesanoja ja interjektioita sekä näkökulman vaihtelua esiintyy informantti C:n puheessa verrattain vähän, sillä tällaisia koodinvaihtotapauksia on aineistosta luokiteltavissa vain kolme jokaisesta funktiosta.

22) C5: **Life hack jonka häkkiydestä** en oo ihan varma, mutta siis kun laittaa vissyä tai kivennäisvettä sit sinne jotain tällasta ällömakeeta likööriä. -- No **anyway, kuitenkin**.

Kun tarkastellaan Kaavion 1. mukaisesti informantti B:n, D:n ja E:n koodinvaihtotapauksien määrää funktioittain, voidaan havaita erojen olevan hyvin vähäiset. Informantti B:llä ja D:llä esiintyy eniten koodinvaihtoa interjektioiden ja täytesanojen yhteydessä, E:llä puolestaan persoonallisten ja objektiivisten näkökulmien vaihtelussa. Koska täytesanat ja huudahdukset ovat lähes aina yksi- tai kaksisanaisia osioita, on näiden osio suhteessa matriisikieliseen puheeseen hyvin vähäinen. Informantti B:llä ja D:llä esiintyy informantti E:hen verrattuna enemmän variaatiota koodinvaihtotilanteiden määrässä, sillä E:llä tapauksia on tasaiset 4–5 kappaletta per funktio. Informantti E:llä ilmenee vähiten toistoa, mikä puolestaan on poikkeavaa verrattuna informantti B:hen ja A:han. Toisaalta myös informantti D:llä ja C:llä toistoa esiintyy vähän.

23) E6: Mä odotan et tulee se jakso ku se yks hyppää siit ikkunast, **when did New York get so doll** ja sit se tippuu. [Sitaatti]

24) B4: Kolme jätäkää huus mulle et mikä vittu sä olet sut pitäis tappaa. **Still alive!**  
-- Se tunne kun sä vihdoin tunnet sen et sä oot oma ittes ni se tunne on aivan huikea. **It's amazing!** [Huudahdukset]

25) D6: -- sit tämmönen valokynä tai tämmönen niinku **concealer** --. [Toisto]

Määrällisesti funktioittain tarkastellessa koodinvaihtoa ilmenee kokonaisuudessaan vähiten informantti D:llä eli yhteensä noin 17 koodinvaihtotapausta. A:lla puolestaan esiintyy funktioittain eniten koodinvaihtoa eli yhteensä 44 tapausta. C:llä koodinvaihtoa esiintyy toiseksi eniten 31 tapauksen verran. Informantit B ja E asettuvat määrällisesti koodinvaihtotapausten keskelle siten, että B:llä tapauksia on yhteensä 19 ja E:llä 22. Verrattuna informantti D:n 17 tapaukseen erot ovat hyvin pieniä näiden kolmen sisällöntuottajan välillä. Toisaalta D:n, A:n ja B:n välinen ero on huomattava. Kun tarkastellaan esiintymiä määrällisesti, on otettava huomioon myös videoiden kestot. Keskimäärin voidaan olettaa, että mitä pidempi julkaistu video on, sitä enemmän siinä on puhetta. Kun tarkastellaan videoiden pituuksia, voidaan kuitenkin huomata, ettei pidempi video automaattisesti takaa funktioittain luokiteltujen koodinvaihtojen määrällistä lisääntymistä. Esimerkiksi informantti E:n aineistoon valikoitujen videoiden yhteiskesto on 147,6 minuuttia, B:n 101,4 minuuttia ja D:n 93,87 minuuttia. Silti erot koodinvaihtotapausten määrissä ovat vähäiset. Ilmiö voidaan todeta myös siitä, että informantti C:n videoiden yhteiskesto 85,17 on huomattavasti vähemmän kuin esimerkiksi informantti E:llä, mutta koodinvaihtotapauksia on silti verrattain enemmän. Esiintymien määrällinen tarkastelu on hedelmällisempää myöhemmin, kun tarkastellaan tarkemmin koodinvaihtotapausten tyypittelyä ja niiden määrällistä ulottuvuutta erityisesti Lauttamuksen (1992) näkemysten mukaisesti.

Näyttää siltä, että Junellin (2020) tutkimustuloksia mukaillen voidaan todeta viestin modifioinnin olevan vahvasti yhteydessä puhujaposition. Koodinvaihdolla ja viestin modifioinnilla puhuja esittää asennoitumisensa kyseiseen aiheeseen, ilmaisee tunnetta tai korostaa kontrastia, kuten esimerkeissä 26 ja 27:

26) E5: Se on semmoses hyväns balanssis tietsä siel on välil semmost **edgee** ja välil **I'm so lonely**.  
-- Nää on nii isot pinta-alaltaan **it's ridiculous**.

27) B7: Mut tulevaisuudesta, mitä me oikeesti tiedetään? Mitä me tiedetään? **What the fuck do I know?**

Kokoavasti voidaan todeta, että Gumperzin funktioiden mukaan tyypiteltyä koodinvaihtoa esiintyy kaikilla sisällöntuottajilla. Koska tapauksia on kuitenkin verrattain vähän, ei funktioittain luokitellusta aineistosta kannata vertailla esimerkiksi sitä, onko koodinvaihto lisääntynyt vanhemman ja uudemman tuotannon välillä. Myös yksilöllistä variaatiota on hankalaa suhteuttaa toisiinsa tästä

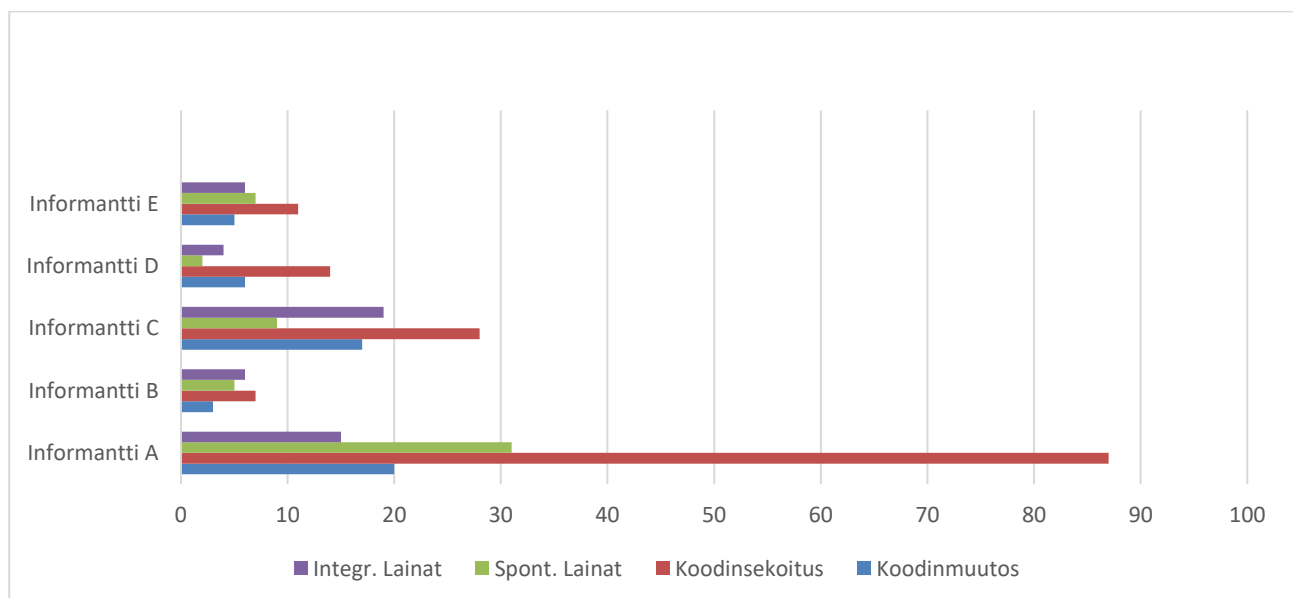
syystä. Näihin palataan myöhemmin seuraavassa alaluvussa, jossa koodinvaihdot luokitellaan eri näkökulmasta ja vertailtavuutta saadaan enemmän.

Aineiston perusteella näyttää kuitenkin siltä, että koodinvaihto on sisällöntuottajien puheessa muutakin kuin esimerkiksi siteeraamista, jossa upotetun kielen käyttö ei välttämättä ole kaikista autenttisimmassa asemassa. Aineistosta pystytään toteamaan, ettei koodinvaihtoa voida pitää yksiselitteisenä ilmiönä, vaan sitä esiintyy monipuolisesti jopa kuudessa eri kategoriassa kaikkien sisällöntuottajien puhekielessä.

## 5.2 Yksilöllinen variaatio ja koodinvaihtotyypit

Yksilöllisen variaation tarkastelussa on mielekästä vertailla myös sitä, millaista koodinvaihtoa sisällöntuottajat puheessaan käyttävät. Edellisen alaluvun perusteella voidaan todeta, että funktioittain luokiteltuna koodinvaihtotapauksissa ei sisällöntuottajien välillä ilmene kovinkaan suurta eroa. Seuraavaksi tarkastellaan sitä, millaista koodinvaihtoa sisällöntuottajien puheesta ilmenee Lauttamuksen, Gumperzin ja Kovácsin näkemyksistä muodostettujen luokittelukategorioiden avulla. Kategoriat on esitelty tarkemmin luvussa 4.3.

Kaavioon 2 on kerätty kaikki koodinvaihtotapaukset Lauttamuksen, Gumperzin ja Kovácsin näkemysten mukaisesti. Seuraavaksi esitellään, miten nämä koodinvaihtotapaukset esiintyvät sisällöntuottajien puheessa.



Kaavio 2. Koodinvaihtotapaukset luokittelukategoriottain

Kaikilla sisällöntuottajilla esiintyy koodinvaihtokategorioista eniten koodinsekoitusta (*code-mix*). Koodinsekoituksessa on kyse yksittäisistä sanoista tai yhdyssanoista, jotka pääasiassa myös operoivat upotetun kielen konventioiden mukaisesti. Kun vuorossa ilmenee koodinsekoitusta, matriisikielisen vuoron joukossa esiintyy yleensä yksi tai kaksi sanaa upotetulla kielellä. Nämä sanat äännetään englannin kielen mukaisesti.

Ylivoimaisesti eniten koodinsekoitusta esiintyy informantti A:lla jopa 87 tapauksen verran. Toiseksi eniten koodinsekoitusta ilmenee informantti C:llä, jolloin määrä on 28 tapausta. Informantti D:llä koodinsekoituksia on 14, E:llä 11 ja B:llä seitsemän. Tästä huomataan, että ero informanttien E ja D välillä on jälleen suhteellisen pieni, mutta esimerkiksi A:lla ja B:llä koodinsekoitusten määrässä on huomattava ero.

Yhden tai kahden sanan mittaisia koodinsekoituksia (esimerkit 28–32) esiintyy siis jokaisella sisällöntuottajalla eniten verrattuna muihin taulukossa 2 esitettyihin luokittelukategorioihin. Yksittäiset sanat ovat useimmiten ilmauksia, jotka kuvailevat puheena olevaa asiaa. Koodinsekoituksella tuodaan lisämerkitys aiheeseen, jolloin useimmissa tapauksissa vuoro ei ole ymmärrettävissä ilman koodinvaihtokohtaa.

28) E7: Sen voi ottaa messiin jos lähtee johonkin, nyt se on näköjään ihan **tangled** tohon.

29) B7: Kun mä tänä vuonna tein sen et mä yritin viikon ajan olla **zero waste** niin mä oppisin siinä viikon aikana niin paljon asioita jotka oli mulle täysin uusii.

30) C5: Halusin tehdä jonkun **comeback** hommelin -- on mun **handle** siellä.

31) D5: Niin siit tulee tämmönen **fried rice**.

32) A5: Joo ja \*\*\*\*\* sano et royal design on **trashy**, olis pitänyt arvata.

Lauttamus (1992) toteaa, kuinka koodinsekoitukset ovat usein erityisesti substantiiveja. Koodinsekoitukselle on tyypillistä, ettei siinä ole havaittavissa morfologista eikä syntaktista mukautumista matriisikieleen, sillä koodinsekoitus operoi pääasiallisesti lähtökielen kieliopin mukaisesti. Tällaisia tapauksia aineistossa esiintyy sisällöntuottajien välillä yhteensä 147 kappaletta. Niistä useimmat ovat substantiiveja ja adjektiiveja. Yli puolet aineistosta kerätyistä koodinsekoitustapahtumista esiintyy informantti A:n puheessa.

Koodinmuutoksella (*code-change*) tarkoitetaan täysin upotetun kielen mukaisia ja yleensä monisanaisia fragmentteja. Lauttamuksen (1992) mukaan koodinmuutokset ovat tyypillisesti monisanaisia lauseita tai kokonaisia lausekkeita, jotka käyttäytyvät leksikaalisesti, syntaktisesti, morfologisesti ja fonologisesti lähtökielen (tässä upotetun kielen) mukaisesti. Koodinmuutokset eivät ole mukautuneet matriisikieleen, sillä niissä operoi lähtökielen kielioppi. (Lauttamus 1992, 9.)



Myös koodinmuutosta esiintyy aineistossa jokaisen sisällöntuottajan puheessa (esimerkit 33–35). Koodinmuutosta eniten ilmenee informantti A:lla, mutta informantti C:n koodinmuutostapaukset ovat määrällisesti A:ta lähellä. Puolestaan informanteilla D, E ja B koodinmuutoksia ilmenee vain kolmesta kuuteen tapausta. Aineistossani olen laskenut koodinmuutokseksi sellaiset vaihdot upotettuun kieleen, jotka ovat pituudeltaan yli kaksi sanaa. Jos koodinmuutosta vertaillaan Gumperzin funktioihin, samankaltaisuuksia löytyy eniten sitaattien ja lainauksien osalta, sillä esimerkiksi sisällöntuottajien käyttämät lainaukset ovat usein monisanaisia kiteytyneitä ilmauksia. Aineiston perusteella koodinmuutoksilla kuvataan eniten omaa toimintaa, mikä puolestaan tukee Gumperzin funktiota näkökulman vaihtelusta (esimerkki 37).

33) E2: **Let's make the magic happen.**

34) B5: **Thanks! I'm blushing.**

35) C5: [Ottaa valokuvan] **Is that the shot?**

36) D5: Sit laitetaan [puuroon] mustikoi **and maybe little bit of peanut butter.**

37) A2: Niinkun **I freaked the fuck out.**

Aineiston perusteella voidaan päätellä, että koodinmuutoksen yhteydessä sisällöntuottaja sananmukaisesti muuttaa kielensä lähes kokonaan upotettuun kieleen koko lausekkeen tai lauseen ajaksi. Tällaisia tapauksia aineistossa on yhteensä 45. Kun tarkastellaan koodinsekoituksen ja koodinmuutoksen määrällisiä eroja sisällöntuottajien välillä, on helpommin havaittavissa, että informanteilla A ja C tapauksia esiintyy huomattavasti enemmän kuin informanteilla B, D ja E. Siinä missä Gumperzin funktioita tarkastellessa erot toisiinsa ovat verrattain pienet, voi koodinmuutoksia ja koodinsekoituksia havainnoidessa huomata suuremman eron yksilöllisessä variaatiossa.

Kuten tässä tutkielmassa on aiemmin mainittu, ei koodinvaihtoa ja lainaamista voi yksiselitteisesti erottaa toisistaan. Aineiston tarkastelussa on yhdistetty Lauttamuksen, Gumperzin ja Kovácsin teorit koodinvaihdon ja lainaamisen välisestä jatkumosta. Kovácsin (2009) mukaan koodinvaihto ja lainaaminen toimivat saman jatkumon ääripäinä siten, että jos matriisikielen joukossa käytetään upotettua kieltä niin, että sana lausutaan upotetun kielen fonologian mukaisesti matriisikielen kontekstissa, pidetään sitä koodinvaihtona. Jos taas sanan ääntäminen on mukautettu matriisikielen mukaiseen fonologiaan, pidetään tapausta lainana. Lauttamuksen (1992) mukaan koodinsekoitus ja koodinmuutos edustavat lähinnä koodinvaihtoprosessin strukturaalisia vastineita, kun taas spontaanit lainat ja integroituneet lainat lainaamisen vastineita. Lisäksi Gumperz (1982) on todennut, ettei lainaamisen suuntaa ole määritelty. Perinteisesti lainaamista ja koodinvaihtoa ei kuitenkaan eroteta tiukasti toisistaan, joten tässä tutkielmassa spontaaneja lainoja sekä integroituneita lainoja tarkastellaan koodinvaihdon käsitteen alla.

Tässä tutkielmassa spontaanit lainat (*noance loans*) ovat ilmaisuja, joissa sanat äännetään upotetun kielen fonologian mukaisesti, mutta ne ovat muuten mukautuneet matriisikielen eli suomen rakenteeseen. Integroituneet lainat (*integrated loans*) puolestaan ovat mukautuneet matriisikieleen myös fonologisesti, mutta eri puhujien ääntämisessä voi esiintyä variaatiota.

Jälleen voidaan huomata, että informanteilla A ja C nämä kaksi kategoriaa ovat kolmea muuta sisällöntuottajaa enemmän edustettuina. Siinä missä informantti A on pitänyt kärkisijaa tähänastisen koodinvaihdon tarkastelun määrällisessä osuudessa, vie B nyt voiton integroituneiden lainojen esiintymismäärässä (19kpl, A:lla 15). Informantilla A puolestaan esiintyy informanttia C huomattavasti enemmän spontaaneja lainoja (31 kpl, C:llä 9kpl). Informantilla D spontaaneja lainoja on vain kaksi ja integroituneita 4. Informantilla B vastaavat luvut ovat viisi spontaania lainaa ja kuusi integroitunutta lainaa. E:llä spontaaneja lainoja esiintyy seitsemän ja integroituneita kuusi. Edelleen huomataan, että kolmen viimeiseksi mainitun sisällöntuottajan yksilöllinen variaatio on pieni. Spontaaneista lainoista esimerkkeinä 39 ja 41, integroituneista 38 ja 41.

38) E3: Mä en osaa käyttää **top coattei**. [*top kouttei*]

39) B1: Bilesnäppääjät on kyl mun niinku **all time favorittei**. [*oll taim feivörittei*]

40) C6: Mä niinku **call outtaan** täs enemmän itteeni. [*kallouttaan*]

41) A2: Sitten seuraavaksi esittelen ehdottomasti tammikuu **highlightin** eli sen kun menin --.  
[*hailaitin*]

Spontaanit lainat ovat aineiston perusteella sisällöntuottajien puheessa yleensä sellaisia, että kiteytynyt ilmaus upotetaan suomen kielioppiin lisäämällä sijapäätte englanninkielisen ilmauksen loppuun. Fonologista integroitumista ei juuri esiinny, vaan lainat säilyttävät lähes poikkeuksetta lähtökielen ääntöasun matriisikielen sijapäätettä lukuun ottamatta.

Integroituneita lainoja esiintyy aineistossa sisällöntuottajien puheessa yhteensä 49 kappaletta. Tällainen laina on mukautunut matriisikieleen spontaania lainaa pidemmälle, jopa siten, että lainasta on muodostunut oma sanansa (esimerkit 42–44). Suomen kieleen integroitunut laina on mukautunut vierasperäisenä sanana, eikä lainasanan käyttö välttämättä johdu sanastollisesta puutteesta. Aineistossa onkin useita tapauksia, joissa integroituneiden lainojen käyttö ei selity sillä, että kyseiselle ilmaukselle ei olisi suomenkielistä vastinetta. Tällaiset tapaukset ovat aineistossa yleensä verbejä:

42) B3: Jos joku **ignora** [*ignore*] sun viestin, niin pommita sitä niin kauan et sen on pakko **blokata** [*blok*] sut.

43) D6: Vaiks aina ennen mä oon sit tehnyt niin et mä **spoilautan** [*spoil*] itseni tiettekö katon jotain Youtube videoit --.

44) A8: Se ottaa tohon sun ihon ylimmän kerroksen pois ja sit iho rupee vähän niinku silleen myös kuoriutuu tai silleen **fleikkaa** [*flake*] --.

Aikaisemmin on mainittu Palmgrenin (2007) ehdotus integroituneiden lainojen ja *finglishin* yhdistämisestä siten, että tällaisiin tapauksiin voitaisiin suhtautua tilanteen synnyttämänä verkkokielisenä viestintäratkaisuna. Ehdotus on mielenkiintoinen, sillä useimmat sisällöntuottajien integroituneista lainoista ovat mitä luultavimmin internetkulttuurin synnyttämiä. Yllä olevien esimerkkien lisäksi tällaisia ilmauksia aineistossa ovat muun muassa *blendautua* > *blend*, *desci* > *description*, *stoori* > *story*, *unfolllaa* > *unfollow* ja *mutettaa* > *mute*.

Integroituneet lainat ovat oikeastaan ainoa koodinvaihdon muoto, joka voi aiheuttaa ymmärrettävyydessä monitulkintaisuutta. Jos vuoro otetaan irti kontekstista, ei kuulija välttämättä yhdistä *stooria* Instagram-tarinoihin tai *desciä* Youtubessa videoiden alla olevaan *description*-laatikkoon.

Määrällistä ulottuvuutta tarkastellessa on palattava vielä videoiden julkaisuajankohtiin ja siihen, miten koodinvaihto on lisääntynyt. Informantti B on sisällöntuottajista ainut, jolla koodinvaihtoa ilmenee hieman enemmän vanhemmassa kuin uudemmassa tuotannossa. Vanhemman tuotannon neljästä videosta koodinvaihtotapauksia esiintyy puhekielessä yhteensä 34 kappaletta ja uudemmassa 31 kappaletta. Kaikilla muilla sisällöntuottajilla koodinvaihtotapausten määrä on vähintään kaksinkertaistunut uudemmissa videoilla. Informantti E:llä vanhemmasta tuotannosta koodinvaihtotapauksia löytyy yhteensä 24 kappaletta ja uusimmalta 54 kappaletta. C:llä koodinvaihdon osuus vanhemman tuotannon videoilla on 17 kappaletta, kun taas uusimmissa videoissa koodinvaihtoa esiintyy jopa 78 tapausta. D:n puhekielessä koodinvaihtoa esiintyy muihin nähden vähiten, mutta tapausten lisääntyminen on suurta; vanhimmassa tuotannossa koodinvaihtoesiintymiä on vain kolme ja uusimmassa 20. A:lla puolestaan esiintyy koodinvaihtoa sisällöntuottajista eniten. Vanhimman tuotannon videoilla koodinvaihtoa ilmenee 57 tapausten verran ja uudemmilta videoilta tapauksia on yhteensä jopa 158. Voidaan kuitenkin todeta, että yksilöllinen koodinvaihto sisällöntuottajien puheessa on lisääntynyt ajan saatossa.

Uutta ja vanhaa vertaillessa on otettava tässäkin huomioon videoiden kestot. Myös tässä näkökulmassa päästään tulokseen, että koodinvaihtotapaukset eivät edelleenkään korreloi suoraan videoiden kestojen kanssa. Esimerkiksi informantti B:llä vanhempien videoiden yhteiskesto on 21,3 minuuttia ja uusien 120 minuuttia, mutta koodinvaihtoa ilmenee silti hieman enemmän vanhemmassa tuotannossa. Toisessa ääripäässä informantti C:n vanhin tuotanto kattaa aineistossa 24,87 minuuttia 17 koodinvaihtotapauksella. Vaikka C:llä uusimman tuotannon yhteiskesto on vähän yli kaksi kertaa

vanhemman (77min), on koodinvaihtotapausten määrä moninkertainen 78 tapausta. Sama ilmiö todetaan myös informantti D:llä ja A:lla.

### 5.3 Koodinvaihto ja kielenulkoiset tekijät

Variaatiota voidaan tarkastella myös kielenulkoisten tekijöiden avulla. Sosiolingvistiikan sisällä kielenulkoista vaihtelua on luokiteltu lukuisin eri tavoin. Seuraavaksi koodinvaihtoa ja kielenulkoisia tekijöitä tarkastellaan tilanteisen ja sosiaalisen vaihtelun näkökulmista.

Tilanteista tai tilannekohtaista vaihtelua tarkastellessa huomio kiinnittyy muun muassa siihen, miten ihminen puhuu tietynlaisissa tilanteissa. Auerin (1995) mukaan sellaiset keskustelutilanteet, joissa koodinvaihtoa yleensä esiintyy, noudattavat pitkälti Gumperzin funktioita. Lisäksi koodinvaihtoa esiintyy silloin, kun puhetoiminnassa tapahtuu muutos, aihe vaihtuu tai kielellä leikitellään. Tilanteista koodinvaihtoa voi luonnehtia myös diglossiaksi yksilötasolla (Blom & Gumperz 1972). Tässä sama henkilö valitsee käyttöönsä varieteetin toisesta kielestä muun muassa vastaanottajien perusteella. Junell (2020) puolestaan esittää, että koodinvaihdolla muutetaan kertomisen tapaa esimerkiksi asenteiden ilmaisemisessa. Koodinvaihdolla joko korostetaan tai laimennetaan sanomaa.

Aineiston perusteella puhetoiminnan muutoksella tai aiheen vaihtumisella ei näytä olevan suurta kytköstä koodinvaihdon esiintymiseen. Sisällöntuottajat käyttävät upotettua kieltä matriisikielen lomassa suorastaan luontevasti, eikä koodin vaihtumiseen kiinnitetä sen suurempaa huomiota. Kun sisällöntuottaja vaihtaa upotettuun kieleen – yleensä hetkellisesti – ei sisällön informaatioarvo nouse matriisikielen yli. Tällä tarkoitetaan sitä, että koodin vaihduttua upotetun kielen mukainen osuus on yleensä lisähuomio puheenaiheeseen, mutta itse sanoma on ymmärrettävissä ilman upotetun kielen lisäystäkin (esimerkit 45–47). Ilmaukset olisi voinut yhtä hyvin esittää suomeksikin. Aineiston perusteella koodia vaihdetaan silloin, kun sisällöntuottaja tuo puheenaiheeseen lisämerkityksen silloin, kun ilmaistaan esimerkiksi asennetta tai korostetaan jonkin asian ominaisuutta:

45) A5: Mut mä olin ehkä **tiny bit** pettynyt nii sen takii mä annan kaksi. Se ei oo semmone niinku hyvä kolmonen vaan ugh.

46) E7: Sun on koko ajan pakko jatkuvasti kehitettävä ittees, **monkey see monkey do**.

47) D5: Mä en muista koska mun tukka olis ollu näin likanen ku siel on niit rasvoja ja hikee, koko **mix**.

Koodinvaihtoa ei siis aineiston perusteella esiinny pääsääntöisesti silloin, kun puheenaihe tai tilanne muuttuu, vaan upotettua kieltä käytetään sen sijaan sellaisissa tilanteissa, kun sanomaan halutaan tuoda lisämerkitys upotetulla kielellä (esimerkki 48). Tällaisissa tapauksissa viestin korostaminen tai laimentaminen on Junellin (2020) tutkimustulosten mukaisesti vahvasti läsnä:

48) B4: Se tunne kun sä vihdoin tunnet sen et sä oot oma ittes ni se tunne on aivan huikea. **It's amazing!**

Aineiston perusteella kielellä leikittelyä tapahtuu eniten spontaanien ja integroituneiden lainojen yhteydessä. Toisaalta kuten Lappalainen (2004) on osoittanut, ei välttämättä ole kannattavaa arvailla sitä, kuinka tietoisesti yksilö toimii kieltä käyttäessään. Aineistosta ei siis voi yksiselitteisesti osoittaa, missä määrin sisällöntuottaja käyttää englantia leikillisessä kontekstissa:

49) D5: [Osoittaa autoa] Hän on niin **pretty**. Hän on niin **pretty!**

Edellisessä esimerkissä 49 informantti D puhuu uudesta autostaan ja kuvailee sen olevan kaunis. Leikittelyä tapahtuu jo matriisikiellelläkin siinä, kun D puhuu elottomasta objektista persoonapronominilla. Kielellä leikittelyä ilmenee myös siten, että D lausuu englanninkielisen sanan *pretty* suomen kielen fonologian mukaisesti eikä sen oikeassa lausumismuodossa [ˈpɪɪr.i].

Tilanteista koodinvaihtoa ja diglossiaa ei voi Gumperzin mukaan erottaa toisistaan. Kuten aiemmin todettiin, aineiston perusteella näyttää siltä, että koodinvaihtoa ei käytetä silloin, kun halutaan kertoa tärkeää informaatiota katsojalle. Tätä tukee Gumperzin näkemys siitä, että diglossiassa kielen valinta määräytyy tilanteen mukaan, mutta monikielisessä yhteisössä vuorovaikutuksessa kaksi eri kieltä voivat toimia rinnakkain ilman, että koodinvaihtoa voisi luonnehtia kommunikaatiostrategiaksi. Sisällöntuottajien koodinvaihto ei olisi tulosta diglossiasta, vaan monikielisen yhteisön itsenäisen kielimuodon kommunikaatiomuodosta. Myös Poplack, Wheeler ja Westwood (1990) ovat todenneet, että kyky vaihtaa kieltä kesken keskustelun on osa puhujien kompetenssia, joka monikielisessä yhteisössä on luonnollisesti vaihteleva.

Vaikka Blom ja Gumperz (1971) toteavat tilanteisen koodinvaihdon ja diglossian tarkoittavan osittain samaa asiaan, ei aineiston perusteella sisällöntuottajien koodinvaihtoa voi pitää suoraan diglossiana. Kielimuotojen funktiot eivät ole selkeästi eriytyneet, eikä varieteetteja käytetä pelkästään englanninkielisissä konteksteissa vaan matriisikielen ohessa. (Kauhanen 2009.) Lisäksi matriisikielen ei voida olettaa vaihtuvan upotettuun kieleen ainakaan vastaanottajan ymmärrettävyyden perusteella, sillä kohdeyleisö on suomenkielistä. Koska koodinvaihtoa kuitenkin ilmenee kaikilla sisällöntuottajilla, voidaan olettaa, että tilanteinen koodinvaihto noudattaisi enemmän monikielisen

yhteisön kontekstia. Vaikka Suomen toinen virallinen kieli on ruotsi, englantia puhuvien osuus on huomattavasti suurempi (4.5).

Halliday (1978) esittää, että ihmisten puhuvat yksinkertaistesti eroavat toisistaan sen perusteella, keitä he ovat, mikä selittäisi vaihtelua. Tällaisissa tapauksissa selittävinä tekijöinä pidetään sosiaalisia muuttujia. Ferguson (1994) puolestaan korostaa ryhmän merkitystä; perusoletuksena on, että yhteiskunnassa toimivalle ryhmälle on taipumus kehittää sille tunnistettavia piirteitä. Puhujat mukauttavat puhetapaansa sellaisten puhetapojen suuntaan, joiden kanssa yksilö on säännöllisesti tekemisissä. Fergusonin ajatusta tukee englannin kielen yleistyminen ja sosiaalisen median saavutettavuus. Sisällöntuottajista kaikki kuuluvat sukupolveen, joka on kasvanut enemmän tai vähemmän internetillä pyörivään maailmaan. Usein käytetyt ilmaukset muuttuvat puhekielessä spontaaniksi lainaksi. Esimerkiksi Instagram-sovelluksen ”tarinat” ovat muuttuneet puhekielessä muodosta *instagram-story* muotoon *stoori* (esimerkit 50–52).

50) B1: -- automaattisen refleksin et sä painat niin nopeesti ne kaikki ohi, et se koko **stoori** on pois sun silmistä alta aikayksikön.

51) C4: Mä oon siis tän vuoden aikana legit mutettanu varmaa 50 prossaa mun seuratuista ja niiden **ig-stoorit**.

52) A7: Mä käytän niitä housui tyylillä joka päivä silloin kun mä näytin mun **stoorissa** niin monet teistä oli kiinnostuneita --.

Leppänen ja Nikula (2008) toteavat, kuinka suomen ja englannin kielen vaihtelulla on myös sosiaalinen ulottuvuus. Tällaisella koodinvaihdolla on monia tehtäviä, joista yhtenä voidaan pitää sosiaalisen ryhmän tapaa rakentaa yhteistä kielenkäytön tyyliään. Ilmiö voisi selittää osittain esimerkiksi sitä, miksi informanttien D, E ja B erot koodinvaihtotapausten määrissä ovat vähäiset.

Sosiaaliset muuttujat ovat sisällöntuottajilla suhteellisen samankaltaiset, kuten on osoitettu luvussa 4.2. Toisaalta juuri tällaiset samankaltaisuudet tukevat Fergusonin kantaa siitä, että ryhmälle kehittyisi tunnistettavia piirteitä myös puhekieleen. Kaikilla sisällöntuottajilla esiintyy koodinvaihtoa ja esiintymien määrä vähintään kaksinkertaistui vuosien kuluessa informanttia B lukuun ottamatta. Vaikka ei voida väittää, että sisällöntuottajat olisivat toistensa kanssa suoraan tekemisissä, voi ilmiölle löytyä selitys. Sosiaalisen median ollessa yhä enemmän saavutettavissa ja englannin kielen yleistyessä yksilö ei voi välttyä päivittäiseltä englannin kielen kuluttamiselta niin lukijana, kuulijana kuin puhujanakaan. Kun kanssakäyminen englannin kielen kanssa yleistyy, myös puhetapa saattaa mukautua englannin kielen suuntaan.

Aineistosta on kuitenkin tulkittavissa se, että vaikka kaikilla sisällöntuottajilla koodinvaihto on yleistynyt, ei englannin kieltä käytetä ymmärrettävyyttä lisäävänä kielenä, vaan pikemminkin

lisämerkityksen tuojana tai asenteen korostamisena. Tulokset eriävät Arvolan (2018) tutkimuksesta, jossa koodinvaihdolla laajennetaan tai supistetaan osallistujien määrää. Toisaalta on huomattava, että Arvola tutkii kirjallista ilmaisua. Lisäksi hänen tutkimuksensa aineiston kohdeyleisö on varioivampi kuin sisällöntuottajien videoiden. Toisaalta, kuten Mantila (2004) on todennut, ei puhekielen ja sen variaation tutkimus ole yleisestikään tuottanut yksiselitteisiä tuloksia.

Koska alueellinen variaatio on sisällöntuottajien välillä lähes olematon, ei sitä pidetä variaation pääasiallisena selittäjänä. Mielenkiintoista kuitenkin on, että sisällöntuottajista ainoastaan informantti B asuu Uudenmaan ulkopuolella ja vain hänen kohdallaan vanhemmilla videoilla esiintyy enemmän koodinvaihtoa kuin uusilla. Kaikilla Uudenmaalla asuvilla sisällöntuottajilla koodinvaihto on aineiston perusteella lisääntynyt. Ilmiötä voi kuitenkin selittää informantti B:n oma idiolekti, sillä koodinvaihtoja ei alkujaankaan esiinny kovinkaan suurta määrää.

Halosen (2009) mukaan monikielisyys tarjoaa erityisen runsaat resurssit erilaisten identiteettien rakentamiseen. Muhonen (2008) tarkastelee moniäänisyyttä eli *polyfoniaa* toisen kontekstin eli toisen puhujan, tekstin tai ajatuksien liittämistä omaan puheeseen (esimerkki 53). Tämän näkökulman mukaan voidaan pitää Gumperzin funktiota siteeraamisen ja lainaamisen ja koodinvaihdon yhteyttä moniäänisyytenä. Koska sosiaaliseen repertuaariin mahtuu erilaisia identiteettejä ja vastaavia kielellisiä repertuaareja, voi yksilö vaihtaa sosiaalista identiteettiään vaihtamalla koodia (A. Muhonen 2008.)

53) E6: Mä odotan et tulee se jakso ku se yks hyppää siit ikkunast, **when did New York get so doll** ja sit se tippuu.

Auerin (1998) ja Leppäsen (2008) mukaan kielten sekoittaminen on strategia luoda yhteistä kielenkäytön tyyliä. Näin voidaan rakentaa ja osoittaa ryhmäidentiteettiä. Kielet voivat puolestaan olla kielenkäyttäjälle yhtä aikaa mahdollisia ilmauksia, jolloin puhuja ei välttämättä koe käyttävänsä vieraskielisiä ilmauksia. Aineistossa tällaisia tapauksia voivat olla esimerkiksi integroituneiden lainojen hyödyntäminen. Erikielisistä ilmauksista on muodostunut kielellinen varanto, joka koetaan käyttökontekstissaan asianmukaiseksi. Aineistosta ilmiötä voi havainnoida siten, että vaikka koodinvaihtoa esiintyy jokaisella sisällöntuottajalla ainakin vähän ja lähes jokaisella videolla, sisällöntuottaja harvemmin suomentaa upotetun kielen mukaisia vuoroja.

Moniäänisyyden ja enemmän tai vähemmän tietoisten koodinvaihdosten avulla on mahdollista, että sisällöntuottajat ovat identifioineet itsensä saman ammattikunnan sisäiseen internet-yhteisöön. Tähän identiteettiin on vaikuttanut muun median kulutus, joka nykyaikaan on päivittäistä ja jatkuvaa. Mantila (2004) onkin esittänyt, että erilaiset puhekielen piirteet ovat leimautuneet suomalaisessa

puheyhteisössä tietyllä tavalla siten, että piirteitä valitsemalla yksilö voi ilmentää tietynlaista identiteettiä. Koska englanninkieliset lainat ja *finenglishin* käyttö on varsinkin nuorison puhekielessä tavallista, koodinvaihto on luultavasti hyväksytty sosiaalisen ryhmän puhekielen repertoariin (Esimerkit 54 ja 55).

Koodinvaihdon valintaa puhekielen piirteeksi voidaan luonnehtia myös kontekstualisoinnin keinoksi. Koodinvaihtoa voidaan siis pitää omien tulkintojen ja sosiaalisen roolin kertojana, asenteiden ilmaisijana tai assosiaatioiden herättäjänä. Sisällöntuottajilla on sosiaalisissa repertuaareissaan erilaisia identiteettejä sekä niitä vastaavia kielellisiä repertuaareja. Kun tilanne muuttuu, voi samalla muuttua myös sosiaalinen identiteetti. Tällöin kielessä ilmenevä variaatio toimii resurssina, josta nousevat puhujan tietoiset tai tiedostamattomat valinnat (Mantila 2004). Harmanen (2011) on todennut, että nykynuori voi käyttää englantia tunteiden tai ajattelun kielenä, mitä perinteisesti on pidetty äidinkielen puhujan kielenkäyttönä. Lisäksi koodinvaihtotapauksia voidaan luonnehtia kielenylityksiksi, joissa huomio kohdistuu kielelliseen ilmaukseen sosiaalisen identiteetin rakentajana, kuten Kalliokoski (2009, 13) esittää.

54) A6: Se tuntuu että ei pitäisi pistää hirveesti **eforttia** kun nää levittyy helposti eikä ei tule mitään **fallouttia** silmän alle.

55) SL5: Hiljasen kesän ja tapetun kevään jälkeen on alkanu tulee erilaisii **eventtei**. Ja sit ku sä meet niihin **eventteihin** ja siel on ihmisii joit sä et oo nähny johonki 8-9 kuukautee.

Lienee siis kiistatonta väittää, että yleistyvä englannin kielen käyttö ja sen vaikutukset ovat taustalla lisääntyvässä koodinvaihdossa. Tämän totesivatkin Leppänen ja Nikula (2008) heidän esittäessään, kuinka englannin kielen lisääntynyt vaikutus on jo 1960-luvulta lähtien näkynyt esimerkiksi nuorisomediassa ja mainonnassa. Koska Suomessa ei ole ollut tapana jälkiäänittää vieraskielisiä ohjelmia, ovat suomalaiset luultavasti tottuneet kuulemaan ja käyttämään englantia. Kun englannin kieltä on alettu käyttämään pääasiallisena kielenä esimerkiksi mediassa, on herätty tilanteeseen, jossa yhdellä kielellä ei enää pärjätä kaikilla elämänalueilla. Siinä missä lähes yksikielisellä kansalla ei ole aikaisemmin ollut tarvetta *lingua francalle*, pystyy enää tuskin elämään päivääkään ilman että on kosketuksissa englannin kieleen. Huomionarvoista kuitenkin on, että sitä, miksi Englanti jossain tilanteissa valitaan viestinnän kieleksi tai otetaan käyttöön äidinkielen rinnalle, on mahdotonta kuvailla yksitasoisena ilmiönä. (Leppänen ja Nikula 2008.)



#### 5.4 Kokoavasti koodinvaihtotapauksista

Kokoavasti aineiston perusteella voidaan päätellä, että Gumperzin funktioiden mukaan luokiteltua koodinvaihtoa esiintyy kaikilla sisällöntuottajilla. Koodinvaihto on sisällöntuottajien puheessa usein kytköksissä muun muassa mielipiteen ilmaisemiseen tai johonkin asiaan asennoitumiseen. Kaikilla sisällöntuottajilla esiintyy koodinvaihtokategorioista eniten yhden tai kahden sanan mittaisia koodinsekoituksia. Yksittäiset sanat ovat useimmiten ilmauksia, jotka kuvailevat puheena olevaa asiaa tuoden siihen lisämerkityksen. Siinä missä Gumperzin funktioita tarkastellessa sisällöntuottajien erot toisiinsa ovat verrattain pienet, voi koodinmuutoksia ja koodinsekoituksia havainnoidessa huomata suuremman eron yksilöllisessä variaatiossa.

Sekä Gumperzin funktioita että Lauttamuksen ja Kovácsin luokittelukategorioita tarkastellessa huomataan, että informantti A:lla ja C:llä esiintyy määrällisesti eniten ja monipuolisemmin koodinvaihtoa. Puolestaan informanteilla D, E ja B erot ovat verrattain vähäiset. Koska koodinvaihdon voidaan olettaa olevan yhteydessä sosiaaliseen identiteettiin, voisi eroa selittää sisällöntuottajien yksilölliset idolektit. Edellä mainitulla kolmella sisällöntuottajalla koodinvaihto ei näytä kuuluvan yhtä vahvasti yksilölliseen idiolektiin kuin informanteilla A ja C. Yksilöiden identiteettiin ei voida vahvasti ottaa kantaa, mutta aineiston perusteella näyttää siltä, että sisällöntuottajien kieli-identiteettiin englannin kielen käyttö ei vaikuta negatiivisesti. Sen sijaan aineiston koodinvaihtokohtiin ei viitata toisen kielen vieraudella tai kiinnitetä huomiota siihen, että käytössä on kaksi eri koodia. Kaikki sisällöntuottajat käyttävät upotettua kieltä luontevasti matriisikielen ohessa.

Sisällöntuottajat näyttäisivät käyttävän upotettua kieltä matriisikielen kanssa hyvinkin luontevasti, eikä katsojakaan välttämättä kiinnitä koodin vaihtumiseen sen suurempaa huomiota. Sisällöntuottajan vaihtaessa hetkellisesti matriisikielestä upotettuun kieleen informaatioarvo säilyy matriisikielessä. Aineiston perusteella koodia näytetään vaihtavan erityisesti silloin, kun puheeseen tuodaan lisämerkitys upotetulla kielellä esimerkiksi jonkin asian ominaisuutta korostettaessa. Ilmiötä voidaan kuvata kielten sekoittumiseen kuuluvaksi kielten välisyydeksi, sillä suomi ja englanti eivät vaihtelevat puheessa sekventiaalisesti.

Näyttää siltä, että varsinkin informanttien A ja C idiolektiin voidaan olettaa kuuluvan englannin kieli siinä mittakaavassa, että englantia pidetään suomen kielen rinnalla mahdollisena vuorovaikutuksen kielenä. Vaikka informanttien D, E ja B puheessa koodinvaihtotapauksia esiintyy vähemmän, ei koodinvaihtoa voi esiintymien määrien perusteella kuitenkaan kuvailla harvinaiseksi. Koska idiolektiin kuuluvat kaikki ne kielenkäytön piirteet, jotka ovat saman kielen puhujille ja puheyhteisön

jäsenille yhteisiä, lukeutuu koodinvaihto enemmänkin A:n ja C:n persoonallisen variaation osaksi. Aineistosta kuitenkin huomataan, että koodinvaihto näyttää lisääntyvän D:n, E:n ja B:n henkilökohtaisessa kielellisessä ohjelmistossa.

Kuten Anu Muhonen (2008) toteaa, voi koodinvaihtoa tarkastella myös moniäänisyyden näkökulmasta. Kun kielellinen ilmaisu ei synny tyhjästä, vaan puheeseen liitetään aina kaikuja toisista äänistä, voidaan koodinvaihto yhdistää esimerkiksi toisen kontekstin liittämisenä puheeseen. Ilmiön voi aineistosta huomata muun muassa lukuisissa sellaisissa koodinvaihtokohdissa, joissa koodinvaihto on joko siteeraamista, kiteytyneiden ilmausten käyttöä tai erityisesti integroituneiden lainojen esiintymistä. Edellä mainitut lainat ovat syntyneet ainakin osittain eräänlaisen internetkielen vaikutuksesta, josta lainat ovat siirtyneet arkipäiväiseen puheeseen, jota sisällöntuottajat videoillaan käyttävät.

Monikielisessä yhteisössä vuorovaikutuksessa kaksi eri kieltä voivat toimia rinnakkain siten, että koodinvaihto voidaan luokitella kommunikaatiostrategiaksi. Tulosten perusteella sisällöntuottajien koodinvaihto ei olisi tulosta diglossiasta, vaan monikielisen yhteisön itsenäisen kielimuodon kommunikaatiomuodosta. Kyky vaihtaa kieltä kesken keskustelun on osa puhujien kompetenssia, joka monikielisessä yhteisössä on luonnollisesti vaihteleva. Kuten Gumperz (1982) on todennut, keskustelun osapuolet eivät kuitenkaan tunnu kiinnittävän huomioita siihen, kumpaa koodia he käyttävät. Puhujille tärkeintä on se, mitä sanotaan eikä se, kumman koodin kautta se ilmaistaan. (Gumperz 1982: 61.)

Ihmisellä on käytössään lukuisia erilaisia kielellisiä resursseja, joista kielenkäyttäjä voi valita erilaisia aineksia vuorovaikutukseensa niin tietoisesti kuin epätietoisestikin. Koodinvaihtoa voidaan pitää esimerkiksi kielellä leikittelynä, identiteetin vahvistajana sekä dialogisena tai moniäänisenä resurssina vuorovaikutuksessa. Koodinvaihdolle on ominaista, ettei sille voi määritellä pysyvää merkitystä tai funktiota, sillä tulkinta syntyy kulloisessakin puhetilanteessa. Usein se on mahdollista tulkita monin eri tavoin. Keskeiseksi piirteeksi on kuitenkin havaittu ilmiö, jossa koodinvaihto toimii keinona määritellä tilannetta uudelleen siten, että se synnyttää kontrastin edellä sanotun ja seuraavan välillä. (Gumperz 1982.)

Kaiken kaikkiaan voidaan todeta, ettei koodinvaihto esiinny sisällöntuottajien puheessa yksiselitteisenä ilmiönä. Aineistosta on osoitettu, ettei määrällistä ulottuvuutta voi yksiselitteisesti kuvailla esimerkiksi videoiden keston kautta. Kaikilla muilla paitsi informantti B:llä koodinvaihtotapaukset ovat vähintään kaksinkertaistuneet uudemmilla videoilla.

## 6 Päätelmät

Tämän tutkielman tavoitteena on ollut selvittää, millaisissa tilanteissa ja asiayhteyksissä viiden videosisällöntuottajan puhutussa kielessä ilmenee koodinvaihtoa. Prototyypisimmillään koodinvaihto tarkoittaa kahden tai useamman kielen vaihtelua samassa keskustelussa. Koodinvaihto on monikielisen vuorovaikutuksen ilmiö, jossa vuorovaikutuksessa kielten kohtaaminen kytkeytyy koodinvaihtoon ja laajemmin monikielisyyden tulkintaan kontekstissa. (Gumperz 1982; Auer 1995; Kalliokoski 1995.) Lisäksi tutkielmassa on tarkasteltu sitä, onko koodinvaihto lisääntynyt tietyllä aikavälillä. Koska kaikki tutkielmaan valikoituneet sisällöntuottajat ovat tutkielman tekijälle enemmän tai vähemmän tuttuja, hypoteesiksi muodostui oletamus siitä, että koodinvaihto on lisääntynyt sisällöntuottajien puheessa.

Kun tarkastellaan kielenulkoisten tekijöiden roolia puhujien kielellisten valintojen takana, erotetaan usein muun muassa puhetilanne ja ryhmät, joihin yksilö taustamuuttujiansa puolesta kuuluu. Variaationtutkimuksessa on katsottu, että esimerkiksi ikä, sukupuoli ja sosiaaliluokka korreloivat kielellisen variaation kanssa. Yksilön puhetapa vaihtelee kuitenkin myös tilanteittain. Yksilötasolla huomioidaan se, että puhujat mukauttavat puhetapaansa niiden puhujien puhetapojen suuntaan, joiden kanssa yksilö on säännöllisesti tekemisissä. (Ferguson 1994, 18.) Useissa tutkimuksissa on havaittu ammatin korreloivan parhaiten kielellisen variaation kanssa (Lappalainen 2004, 39).

Aineistoon on valikoitunut viiden eri sisällöntuottajan videoita vanhasta ja uudesta tuotannosta. Videoilta on litteroitu yksinkertaisin menetelmin sellaiset kohdat, joissa koodinvaihtoa esiintyy. Litteraateista muodostettiin aineisto, joista tarkasteltiin esiintymiä kategorioittain ja määrällisin näkökulmin. Koodinvaihtoa tarkasteltiin niin ajallisessa kontekstissa kuin asiayhteyksienkin näkökulmasta.

Aineiston analyysin lähtökohdat perustuvat pitkälti kielenulkoisten tekijöiden ja yksilöllisen variaation tarkasteluun sosiolingvistisin metodein. Tuloksista voidaan osoittaa, että Gumperzin metaforisen koodinvaihdon funktioita esiintyy jokaisen viiden sisällöntuottajan puheessa videoilla. Tällöin koodinvaihto on usein kytköksissä muun muassa mielipiteen tai asenteen ilmaisemiseen. Lauttamuksen, Gumperzin ja Kovácsin esittämistä näkökulmista luotiin koodinvaihtokategoriat, joista eniten sisällöntuottajien puheessa esiintyy yhden tai kahden sanan mittaisia koodinsekoituksia. Yksittäiset sanat ovat yleensä ilmauksia, jotka tuovat puheenaiheeseen lisämerkityksen. Gumperzin funktioita tarkastellessa sisällöntuottajien erot toisiinsa ovat suhteellisen pienet, mutta koodinmuutoksia ja koodinsekoituksia havainnoidessa yksilöllisen variaation voi huomata olevan suurempi sisällöntuottajien välillä.

Aineiston perusteella näyttää siltä, että englantia käytetään upotettuna kielenä luontevasti matriisikielen eli suomen ohessa. Informaatioarvo säilyy matriisikielessä, mikä osoittaisi, että koodia vaihdetaan silloin, kun puheena olevaa asiaa tai sen ominaisuutta halutaan korostaa. Tällainen kielten välisyys luokitellaan usein kielten sekoittumisen ilmiöksi, sillä kielet eivät vaihtelevat puheessa sekventiaalisesti. Aineistosta voidaan havaita, että koodinvaihtoa esiintyy sekä määrällisesti enemmän että monipuolisemmin informanteilla A ja C. Informanteilla D, E ja B erot jäävät vähäisiksi. Koska koodinvaihdon voidaan olettaa olevan yhteydessä sosiaaliseen identiteettiin, voisi eroa selittää sisällöntuottajien yksilölliset idolektit. Koodinvaihtokohtiin ei kuitenkaan viitata toisen kielen vieraudella yhdenkään sisällöntuottajan kohdalla. Koska koodinvaihtoesiintymät eivät kenenkään sisällöntuottajan kohdalla ole suoraan olemattomat, ei ilmiötä voida kutsua harvinaiseksi.

Tutkimustulosten perusteella koodinvaihto, monikielisyys ja moniäänisyys kulkevat käsi kädessä. Puheeseen liitetään aina kaikuja toisista äänistä, mikä aineistossa ilmenee muun muassa koodinvaihdon ja lainaamisen yhteydessä. Useat lainat ovat syntyneet osittain eräänlaisen internetkielen vaikutuksesta, josta lainat ovat siirtyneet arkipäiväiseen puheeseen, jota sisällöntuottajat videoillaan käyttävät. Näyttää siltä, että kaksi kieltä toimivat rinnakkain sisällöntuottajien puheessa siten, että koodinvaihto näyttäytyy monikielisen yhteisön kommunikaatiomuotona. Keskustelun osapuolet eivät näytä kiinnittävän huomiota siihen, kumpaa koodia käytetään.

Kokoavasti voidaan todeta, että sisällöntuottajat käyttävät koodinvaihtoa yhtenä lukuisista erilaisista kielellisistä resursseista. Koodinvaihto toimii keinona määritellä tilannetta uudelleen siten, että se synnyttää kontrastin edellä sanotun ja seuraavan välillä (Gumperz 1982). Aineistosta huomataankin, ettei koodinvaihtoa voida kuvailla esiintyvän sisällöntuottajien puheessa yksiselitteisenä ilmiönä. Aineistosta on lisäksi osoitettu, ettei koodinvaihtotapausten määrä korreloi suoraan videoiden kestojen kanssa. Kaikilla muilla paitsi informantilla B koodinvaihtotapaukset ovat vähintään kaksinkertaistuneet uudemmilla videoilla.

Sitä, miksi englanti jossain tilanteissa valitaan viestinnän kieleksi tai otetaan käyttöön äidinkielen rinnalle, on mahdotonta kuvailla yksitasoisena ilmiönä. Mantilan (2004, 324) mukaan tuloksia on paljon, mutta ei tiedetä syitä siihen, miksi jotkin kielenpiirteet ovat leimautuneet kielteisesti ja miksi taas jotkin yleistyvät jatkuvasti. Huomioitava on myös tutkielmassa käytettyjen käsitteiden monimuotoisuus. Käsitteiden vaihtaminen eri tieteentekijöiden mukaan voi aiheuttaa variaatiota tutkimustuloksissa. Esimerkiksi kieleily-käsitteen hyödyntäminen yhtenä aineiston analyysin lähtökohtana voi johtaa eriäviin tuloksiin.

Tämä tutkielma on tarjonnut katsauksen siihen, miten ja millaisissa tilanteissa koodinvaihtoa esiintyy nuorten sisällöntuottajien puhekielessä. Englannin asema *lingua francana* kasvaa jatkuvasti ja koodinvaihdon voidaankin todeta lisääntyneen sisällöntuottajien puhekielessä. Tulevaisuudessa jatkotutkimus voisi kohdistua vielä suurempaan aineistoon, jossa variaatiota voitaisiin esittää tarkemmin muun muassa alueellisin perustein. Lisäksi voitaisiin kartoittaa, onko koodinvaihto puhekielessä edelleen lisääntyvää, vai voidaanko sitä pitää väliaikaisen internetkielen synnyttämänä ilmiönä.

## Lähteet

### Aineslähteet

Tutkielman tekijä voi tarvittaessa esittää listauksen aineistona käytetyistä videojulkaisuista. Eettisistä syistä niitä ei ole merkitty lähdeluetteloon.

### Muut lähteet

- Auer, Peter 1998. From codeswitching via language mixing to fused lects: toward a dynamic typology of bilingual speech. *The international Journal of Bilingualism* 3/4, s. 309-332
- Auer, Peter 1995. The pragmatics of code-switching: a sequential approach. *One Speaker, Two Languages: Cross-Disciplinary Perspectives on CodeSwitching*, s. 115–135. Toim. Lesley Milroy ja Pieter Muysken. Cambridge University Press, Cambridge.
- Auer, Peter – Eastman, Carol 2010. Code-switching. *Society and Language Use*. s. 84–112. Toim. Jürgen Jaspers, Jan-Ola Östman ja Jef Verschueren. John Benjamins Publishing Company, Hollanti
- Arvola, Satu 2018. ”Palasia Amerikasta” Koodinvaihtoa Facebook-kirjoituksissa. Pro gradu -tutkielma. Oulun yliopisto.
- Blommaert, Jan 2010. *The Sociolinguistics of Globalization*. Cambridge Approaches to Language Contact. Cambridge University Press, Cambridge.
- Blom, Jan-Peter – Gumperz, John 1971: Social meaning in linguistic structures. *Directions in sociolinguistics*, s. 407–434. Toim. John J. Gumperz ja Dell Hymes. Holt, Rinehart and Winston, New York.
- Chi, Michelene T. H. 1997. Quantifying qualitative analyses of verbal data: A practical Guide. *The Journal of the Learning Sciences*, 6 (3), 271-313. [Verkojulkaisu, viitattu 2.11.2020.] Saatavilla: [https://www.tandfonline.com/doi/abs/10.1207/s15327809jls0603\\_1](https://www.tandfonline.com/doi/abs/10.1207/s15327809jls0603_1)
- Eckert, Penelope 2012. Three waves of variation study. The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology* vol 41, 87–100.
- Etelämäki, Marja – Herlin, Ilona – Jaakola, Minna – Visapää, Laura 2009. Kielioppi käsitteistyksenä ja toimintana: Kognitiivista kielioppia ja keskustelunanalyysia yhdistämässä [Grammar as conceptualization and as action: Combining Cognitive Grammar and Conversation Analysis]. *Virittäjä*, 113 (2). [Verkojulkaisu, viitattu 3.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/4182>
- Ferguson, Charles 1994. Dialect, register, and genre: working assumptions about conventionalization. *Sociolinguistic perspectives on register*, s. 15-30. Toim. Douglas Biber ja Edward Finegan. Oxford University Press, New York.
- Frobenius, Maximiliane 2011. Beginning a monologue. The opening sequence of video blogs. *Journal of Pragmatics* 43 s. 814–827.
- Gumperz, John Jay 1982. *Discourse strategies*. Cambridge University Press. Cambridge.

- Gumperz, John Jay 2001. Interactional sociolinguistics. A personal perspective. *The handbook of discourse analysis*, s. 215–228. Toim. Deborah Schiffrin, Deborah Tannen ja Heidi Hamilton. Blackwell, Oxford.
- Grönfors, Martti 2011. *Laadullisen tutkimuksen kenttätutkimusmenetelmät*. SoFia-Sosiologi-Filosofiapu Vilka, Hämeenlinna.
- Haakana, Markku 1996. Huumori ja vakava keskustelun kategorioina. *Kieli 10. Suomalaisen keskustelun keinoja II*, s. 141–172. Toim. Auli Hakulinen. Helsingin yliopiston suomen kielen laitos, Helsinki
- Hakulinen, Auli – Sorjonen, Marja-Leena 1993. Research on spoken interaction in Finland. *Research on Language and Social Interaction 26:4*, s. 409–426.
- Halliday, M. A. K. 1978. *Language as social semiotic. The social interpretation of language and meaning*. Edward Arnold, London.
- Halonen, Mia 2009. Puhutun kielen variantit resurssina monikielisten koululaisten kirjoitelmissa. *Virittäjä*, 96 (1), 3, s. 329–355. [Verkkajulkaisu, viitattu 6.10.2020] Saatavilla: <https://journal.fi/virittaja/article/view/4201/3918>
- Harmanen, Minna 2011. Kielitiedolla lisää kielitajua – perusopetuksen ja lukion kielitiedon opetus. *Virittäjä*, 115 (3). [Verkkajulkaisu, viitattu 9.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/4554>
- Hassinen, Sirje 2002. Ollako kaksikielinen vai eikö olla. *Virittäjä*, 106 (3), 405. [Verkkajulkaisu, viitattu 3.9.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/40197>
- Heritage, John. 1996. *Harold Garfinkel ja etnometodologia*. Suomentaneet Ilkka Arminen, Outi Paloposki, Anssi Peräkylä, Sanna Vehviläinen & Soile Veijola. Gaudeamus, Helsinki
- Hiidenmaa, Pirjo 2003. *Suomen kieli – who cares?* Otavan Kirjapaino Oy, Keuruu.
- Hurtig, Marjo 2012. *"Heti ku mun silimä välttää" : Opettajan puheen tilanteisen vaihtelun merkitys oppituntidiskurssissa*. Pro gradu -tutkielma. Suomen kielen, suomalais-ugrilaisen ja pohjoismaisten kielten ja kirjallisuuden laitos. Helsingin yliopisto.
- Junell, Karo 2020. *English code-switches in Finnish Youtube vlogs: Playful attitudes and casual evaluation*. Pro gradu -tutkielma. Kielten maisteriohjelma, Tampereen yliopisto.
- Kalliokoski, Jyrki 1995. Koodinvaihto ja keskustelun moniäänisyys. *Virittäjä*, 99 (1), 2. [Verkkajulkaisu, viitattu 4.10.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/38761>
- Kalliokoski, Jyrki 2009. Tutkimuskohteena monikielisyys ja kielten kohtaaminen. *Kielet kohtaavat*. Toim. Jyrki Kalliokoski, Lari Kotilainen ja Päivi Pahta. Gummerus Kirjapaino Oy, Jyväskylä.
- Kauhanen, Irina 2009. Diglossia ja koodinvaihto arabian kielessä. *Kielet kohtaavat*. Toim. Jyrki Kalliokoski, Lari Kotilainen ja Päivi Pahta. Gummerus Kirjapaino Oy, Jyväskylä.
- Keskimaa, Sari 2013. Murteenkäytön funktiot parisuhteen kuvauksessa Kalle Päätalon Iijoki-sarjassa. *Virittäjä*, 117(4). [Verkkajulkaisu, viitattu 22.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/9464>
- Korpinen, Emilia 2006. *Suomi toisena kielenä -opettajan opetuspuheen koodinvaihto: kieliopillinen ja tilanteinen tarkastelu*. Pro gradu -tutkielma. Suomen kieli ja suomalais-ugrilainen kielentutkimus, Turun yliopisto.
- Koskela, Merja 2015. Kielitietoisuus ja kielikäsitys tilanteisen kääntämisen käytänteitä ohjaavina tekijöinä. *Virittäjä*, 119 (3). [Verkkajulkaisu, viitattu 4.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/49639>

- Kovács, Magdolna, 2009. Koodinvaihto ja kielioppi. *Kielet kohtaavat*. Toim. Jyrki Kalliokoski, Lari Kotilainen ja Päivi Pahta. Gummerus Kirjapaino Oy, Jyväskylä.
- Kuparinen, Olli 2018. Infinitiivien variaatio ja muutos Helsingissä. *Virittäjä*, 122 (1). [Verkkojulkaisu, viitattu 3.11.2020.] Saatavissa: <https://doi.org/10.23982/vir.65310>
- Kurki, Tommi 2018. Kielikäsitusten mosaiikki: Havainnot puhekielen näytteistä. *Sananjalka*, 60(60.), 71–95. [Verkkojulkaisu, viitattu 6.11.2020.] Saatavilla: <https://doi.org/10.30673/sja.70049>
- Labov, William 1966. *The social stratification of English in New York City*. Center for Applied Linguistics, Washington D.C.
- Lappalainen, Hanna 2001a. Variaation tutkimuksen ja keskustelunanalyysin näkökulmat toisiaan täydentämässä. *Keskustelunanalyysin näkymiä*, s. 155–184. Toim. Mia Halonen ja Sara Routarinne. Helsingin yliopiston suomen kielen laitos.
- Lappalainen, Hanna. 2001b. Sosiolingvistinen katsaus suomalaisnuorten nykypuhekieleen ja sen tutkimukseen. *Virittäjä*, 105 (1), 74. [Verkkojulkaisu, viitattu 2.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/40068>
- Lappalainen, Hanna 2004. Variaatio ja sen funktiot. Erään sosiaalisen verkoston jäsenten kielellisen variaation ja vuorovaikutuksen tarkastelua. Suomalaisen kirjallisuuden seura, Helsinki.
- Lappalainen, Hanna – Mustanoja, Liisa – O’Dell, Michael 2019. Miten ja milloin yksilön kieli muuttuu? Helsingiläisidioletkien muutos ja muutoksen tutkimuksen menetelmät. *Virittäjä*, 123 (4). [Verkkojulkaisu, viitattu 2.11.2020.] Saatavilla: <https://doi.org/10.23982/vir.67808>
- Lauttamus, Timo 1992. Lainaaminen ja koodinvaihto: havainnot amerikkansuomalaisten kielistä. *Virittäjä*, 96 (1), 3. [Verkkojulkaisu, viitattu 6.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/38495>
- Lehtonen, Heini 2015. Tyylitellen. Nuorten kielelliset resurssit ja kielen sosiaalinen indeksisyys monietnisisessä Helsingissä. Väitöskirja. Suomen kielen, suomalais-ugrialaisten ja pohjoismaisten kielten ja kirjallisuuksien laitos, Helsingin yliopisto. Unigrafia Oy, Helsinki.
- Lehtonen, Jaakko – Sajavaara, Kari 1981. Anglismit nykysuomessa. *Virittäjä*, 85(4), 289. [Verkkojulkaisu, viitattu 14.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/37638>
- Lindholm, Camilla – Stevanovic, Melisa – Peräkyä, Anssi 2016. Johdanto. *Keskustelunanalyysi: Kuinka tutkia sosiaalista toimintaa ja vuorovaikutusta*, s. 312–330. Toim. Melisa Stevanovic ja Camilla Lindholm. Vastapaino, Tampere.
- Leppänen, Sirpa 2008. Kielten valinta, vaihtelu ja sekoittaminen faniuden rakennusaineiksina verkkofiktiossa. *Kolmas kotimainen. Lähikuvia englannin käytöstä Suomessa*. Toim. Sirpa Leppänen, Tarja Nikula ja Leila Kääntä. Hakapaino Oy, Helsinki.
- Leppänen, Sirpa – Nikula, Tarja 2008. Johdanto. *Kolmas kotimainen. Lähikuvia englannin käytöstä Suomessa*. Toim. Sirpa Leppänen, Tarja Nikula ja Leila Kääntä. Hakapaino Oy, Helsinki.
- Leppänen, Sirpa – Pitkänen-Huhta, Anne – Nikula, Tarja – Kytölä, Samu – Törmäkangas, Timo – Nissinen, Kari – Kääntä, Leila – Virkkula, Tiina – Laitinen, Mikko – Pahta, Päivi – Koskela, Heidi – Lähdesmäki, Salla – Jousmäki, Henna 2009. Kansallinen kyselytutkimus englannin kielestä Suomessa. Käyttö, merkitys ja asenteet. [Verkkodokumentti, viitattu 2.2.2021.] Humanistinen tiedekunta, Jyväskylän yliopisto. Saatavilla: <https://jyx.jyu.fi/bitstream/handle/123456789/22892/978-951-39-3815-4.pdf?sequence=1&isAllowed=y>
- Mantila, Harri 2004. Murre ja identiteetti. *Virittäjä* 108 (3), s. 322–346. [Verkkojulkaisu, viitattu 4.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/40346>



- Miles, Matthew – Huberman, Michael 1994. *Qualitative Data Analysis. A methods Sourcebook*. Second Edition. California: Sage.
- Mononen, Kaarina 2013. Inkerinsuomalaisen suomen kielen käyttö Pietarissa ja sen lähialueella. Helsingin yliopiston suomen kielen, suomalais-ugrialaisten ja pohjoismaisten kielten ja kirjallisuuksien laitos.
- Muhonen, Anu 2008. Sheikataan like a polaroid picture: englannin tehtäviä nuorisoradiojuonnoissa. *Kolmas kotimainen. Lähikuvia englannin käytöstä Suomessa*. Toim. Sirpa Leppänen, Tarja Nikula ja Leila Kääntä. Hakapaino Oy, Helsinki.
- Muhonen, Heidi 2019. Multimodaalinen analyysi osallistumiskehikon rakentumisesta monikielisessä videoblogissa. Pro gradu -tutkielma. Oulun yliopisto.
- Myers-Scotton, Carol 1997. *Duelling languages: Grammatical structure in code-switching*. Clarendon Press, Oxford.
- Myers-Scotton, Carol 1998. Code-switching. *The handbook of Sociolinguistics*. Toim. Florian Coulmas. Blackwell Publishing, United Kingdom.
- Muysken, Pieter 1995. Code-Switching and Grammatical Theory. *One Speaker, Two Languages: Cross-Disciplinary Perspectives on Code-Switching*. Toim. Leslie Milroy ja Pieter Muysken. Cambridge University Press, Cambridge.
- Nuolijärvi, Pirkko 2000. Sosiolingvistiikka kielentutkimuksen kentässä. *Kieli, diskurssi & yhteisö*, s. 13–38. Toim. Kari Sajavaara ja Arja Piirainen-Marsh. Soveltavan tutkimuksen keskus, Jyväskylä.
- Opetus- ja kulttuuriministeriö 2017: Monikielisyys vahvuudeksi Selvitys Suomen kielivarannon tilasta ja tasosta. [Verkkodokumentti, viitattu 3.12.2020.] Opetus- ja kulttuuriministeriön julkaisuja 2017:51. Saatavilla:  
<https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160374/okm51.pdf?sequence=1&isAllowed=y>
- Palmgren, Nina 2007. *Ammattislangina fnglish - Teknisten viestijöiden anglohybridi osana globalisaation diskurssia*. Pro gradu -tutkielma. Tampereen yliopiston kieli- ja käännöstieteiden laitos, Tampere.
- Parjanen, Tiina 2015. *Keep your eyes on the sky so you'll see my savumerkit. Koodinvaihto kahden yksikielisen henkilön välisissä pikaviestikeskusteluissa*. Pro gradu -tutkielma. Suomen kieli ja suomalais-ugrilainen kielentutkimus, Turun yliopisto.
- Paunonen, Heikki 2005. Helsingiläisiä puhujaprofileja. *Virittäjä*, 109 (2), 162. [Verkojulkaisu, viitattu 5.11.2020.] Saatavissa: <https://journal.fi/virittaja/article/view/40402>
- Paunonen, Heikki 1982. Suomen kielen sosiolingvistinen vaihtelu tutkimuskohteena. *Sosiolingvistiikan näkymiä*, s 35–110. Toim. Päivikki Suojanen ja Matti Kalevi Suojanen. Gaudeamus, Helsinki
- Piirainen-Marsh, Arja 2008. Koodinvaihto kontekstivihjeenä videopelitalanteissa. *Kolmas kotimainen. Lähikuvia englannin käytöstä Suomessa*. Toim. Sirpa Leppänen, Tarja Nikula ja Leila Kääntä. Hakapaino Oy, Helsinki.
- Pikelner, Ekaterina 2018. *Venäjänkielisten oppilaiden koodinvaihto ja äidinkielen käyttö valmistavan luokan vuorovaikutustilanteissa*. Pro gradu -tutkielma. Suomen kielen, suomalais-ugrialaisten ja pohjoismaisten kielten ja kirjallisuuksien laitos, Helsingin yliopisto.
- Poplack, Shana – Wheeler, Susan – Westwood, Anneli 1990. Distinguishing language contact phenomena: Evidence from Finnish-English bilingualism. *Codeswitching as a Worldwide phenomenon*, s. 185–218. Toim. Rodolfo Jacobson. American University Studies Series XIII. Linguistics Vol 11. Peter Lang, New York.

- Priiki, Katri 2011. *Läsnäolo, tuttuus, tulkinta ja sävy. Ihmisviitteinen hän ja sen tehtävät Harjavallan, Kokemäen ja Huittisten nykypuhekielessä*. Pro gradu -tutkielma. Turun yliopiston kieli- ja käännöstieteiden laitos, Turku.
- Priiki, Katri 2017. *Hän, se, tää vai toi? Vuorovaikutussosiolingvistinen tutkimus henkilöviittauksista Kaakkois-Satakunnan nykypuhekielessä*. Väitöskirja. Humanistinen tiedekunta, kieli- ja käännöstieteiden laitos. Suomen kieli ja suomalais-ugrilainen kielentutkimus, Turun yliopisto.
- Rautiainen, Pirjo 2007. Kenttätöyön lähteillä – katsaus kvalitatiivisen tutkimuksen perusteisiin. Jyväskylän yliopiston historian ja etnologian laitos. [Verkkójulkaisu, viitattu 3.10.2020] Saatavilla: <https://jyx.jyu.fi/bitstream/handle/123456789/20073/3/jargonia13.pdf>
- Seitanmaa-Hakkarainen, Piritta 2014. Kvalitatiivinen sisällönanalyysi. [Verkkodokumentti, viitattu 5.11.2020] Saatavilla: <https://metodix.fi/2014/05/19/seitamaa-hakkarainen-kvalitatiivinen-sisallon-analyysi/>
- Siitonen, Pauliina 2016. Ehdotukseen vastaaminen perheen vuorovaikutustestissä. Keskusteluanalyttinen näkökulma vuorovaikutukseen ja sen arviointiin. *Virittäjä*, 120(2). [Verkkójulkaisu, viitattu 9.11.2020.] Saatavilla: <https://journal.fi/virittaja/article/view/57252>
- Tito, Henna 2015. *Miks sä kaadat isin piima? Koodien yhdistyminen ja koodinvaihto kahden simultaanisesti kaksikielisen suomalais-virolaisen lapsen puheessa*. Pro gradu -tutkielma. Suomen kieli ja suomalais-ugrilainen kielentutkimus, Turun yliopisto.
- Tuomi, Jouni – Sarajärvi, Anneli 2018. Laadullinen tutkimus ja sisällönanalyysi. Kustannusosakeyhtiö Tammi, Helsinki.
- Thomason, Sarah Grey – Terrence Kaufman 1988. *Language contact, creolization, and genetic linguistics*. University of California Press, Berkeley.
- Tikkanen, Nita 2011. ”Emmä voi sitä sille sanoo”: erimielisyyden ilmaisun keinot keski-ikäisten siskosten keskustelussa. Pro gradu -tutkielma. Helsingin yliopiston suomen kielen, suomalais-ugrialaisten ja pohjoismaisten kielten ja kirjallisuuksien laitos, Helsinki.
- Vatanen, Anna Marjatta 2016. Keskusteluanalyttinen tutkimusprosessi. *Keskusteluanalyysi: Kuinka tutkia sosiaalista toimintaa ja vuorovaikutusta*, s. 312–330. Toim. Melisa Stevanovic ja Camilla Lindholm. Vastapaino, Tampere.