

Pitkänen-Huhta, A., K. Mård-Miettinen & T. Nikula (toim.) 2023. Kielikoulutus mukana muutoksessa – Language education engaging in change. AFinLA-teema / n:o 16, 376–400.

Mari Honko & Mia Halonen

Jyväskylän yliopisto

Kielitaidon arviointi on ideologisesti latautunutta toimintaa

Nostot

- Kielitaidon arvioinnin kriteerit ja kynnystasot heijastavat ja rakentavat käsityksiä kielitaidosta ja kielenkäyttäjistä.
- Koulussa arviointi ohjaa opetusta ja vaikuttaa oppijaan sekä oppimiseen, minkä vuoksi opettajan on tärkeä tiedostua arvioinnin ideologisuudesta ja hyvän/eettisen arvioinnin periaatteista.
- Virallisissa arviointijärjestelmissä testiarvioinnin ideologisuudella on yhteiskunnallista suoria vaikutuksia osallistujien elämään.
- Arvioinnin kohteena olevien oikeudenmukainen kohtelu edellyttää ideologisuuden reflektointia ja tähän perustuvaa arvioinnin kehittämistyötä.

Abstract

Language assessment is an activity guided by prevailing language ideologies as well as more general ideas that permeate the entire society. In this chapter, we will examine the ideologies of language proficiency descriptors and assessment, based on the theoretical perspectives of critical language testing (e.g. Lynch 2001; Shohamy 2001). More specifically, we will present some of the key concepts of the Common European Framework of Reference for Languages (CEFR, Council of Europe 2001, 2020a) and assessment based on the framework, especially on test assessment. With the help of concrete examples, we will also show that a deep understanding of the effects of language assessment and the development of assessment require recognition of both its ideology, its origin, and the mechanisms of its effects.

Keywords: language assessment, test assessment, ideology, ethics

Avainsanat: kielitaidon arviointi, testiarviointi, ideologisuus, eettisyys

1 Arvioinnin ideologisuus kertoo siitä, mitä pidetään arvokkaana

Ideologiat eivät koskaan ole neutraaleja eikä mikään toiminta ole ideologioista vapaata. Tätä nyky-ymmärrystä vasten on ironista, että Antoine Destutt de Tracy loi ideologian käsitteen vuonna 1796 Ranskan vallankumouksen keskellä ”ideoiden tieteenksi” ja tieteenä nimenomaan neutraaliksi. Pian siitä kuitenkin muodostui käsite, jolla rakennettiin ja osoitettiin erityisesti aatteellisia vastakkainasetteluja. (Stråth 2013.) Nykyään ideologialla viitataan aate- ja uskomusjärjestelmiin, jotka ovat joidenkin tahojen määrittelemiä näkemyksiä ”todellisuudesta” ja samalla oikeasta ja väärästä. Tällaiset uskomukset ulottuvat kaikkialle, myös kieleen (Schieffelin, Woolard & Kroskirty 1998), ja niin kielikoulutukseen, kielenoppimiseen kuin kielitaidon arviointiin.

Arvioimisen tarve on jo itsessään ideologista: jokin taho on päättänyt, millaista on oikeanlainen kieli ja sen taito, päättänyt, että taitoa tulee arvioida, sekä määrittänyt, miten tämä tehdään. Kielitaidon arviointia vaaditaan esimerkiksi kansalaisuushakemuksiin useissa Euroopan valtioissa, myös Suomessa (Migri 2023), mikä on yhteiskunnan rakenteiden tasolla tehty kansalliseksi valinta sen hallinnointiin, millaisilla taidoilla ja resursseilla muualla syntyneet ihmiset voivat hakea kansalaisuutta. Kielitaito on Suomessa asetettu niin ikään korkeakoulujen yleisen hakukelpoisuuden ehdoksi (OPH 2022a), ja lisäksi kielitaitoehdot voi paikallisesti tai organisaatiotasolla sisällyttää esimerkiksi toisen asteen oppilaitosten hakumenettelyyn tai työpaikan hakuun. Näidenkin ehtojen taustalla on ideologiseen keskusteluun pohjautuvia sopimuksia riittävästä kielitaidosta.

Tässä artikkelissa rajaamme tarkastelun kielitaidon arviointiin, jolla on suoria vaikutuksia niin arvioinnin kohteena oleviin henkilöihin kuin yhteiskuntaankin mutta jonka ideologisuutta suomenkielisessä kirjallisuudessa on toistaiseksi käsitelty niukasti. Käsitteily on kuitenkin tärkeää, sillä arviointi kertoo kieli-ideologioista eli siitä,

mitä kielessä, kielitaidossa ja kielenpuhujissa arvostetaan. Artikkelimme tuo yhteen kolme 2000-luvun nousevaa ilmiötä, joilla kaikilla on vaikutuksia kieltä ja kielitaitoa koskeviin arvostuksiin ja kielitaidon arviointiin globaalisti ja siten myös Suomessa: 1) lisääntyvä kielitaitoarvioihin tukeutuminen yhteiskunnallisessa päätöksenteossa ja tästä seuraava arvioinnin vastuullisuuden tarve, 2) kriittisten ja oppijalähtöisten näkökulmien nousu kielitaidon arvioinnissa sekä 3) eurooppalaisen viitekehyksen laajeneva käyttö ja painoarvon lisääntyminen. Artikkelin alussa esittelemme kielitaidon arvioinnin ideologisuutta erityisesti suhteessa näihin ilmiöihin, minkä jälkeen syvennymme käytännön esimerkkeihin.

Douglas Fir Groupin (2016: 33) mukaan ideologiat lävistävät kielitaidon oppimisen kaikilla tasoilla. Näin on myös kielitaidon arvioinnissa. Yhteiskunnan makrotasolla ideologiat ohjaavat esimerkiksi arviointia koskevaa lainsäädäntöä, poliittista päätöksentekoa sekä opetussuunnitelmien ja testijärjestelmien valtakunnallisten perusteiden laadintaa. Yhteisöjä koskevalla mesotasolla ideologiat vaikuttavat perusteiden pohjalta tehtäviin paikallisiin ratkaisuihin, kuten siihen, mitä arvioinnista kirjataan paikallisiin opetussuunnitelmiin tai testijärjestelmän ohjeisiin ja miten arviointia ohjailaan koulussa opettajien kesken tai testijärjestelmän arvioijakoulutuksissa. Keskustellaanko esimerkiksi arvioinnin periaatteista yhdessä? Kuullaanko keskustelussa eri näkökulmia, kuten sekä vasta aloittaneita että kokeneempia arvioijia? Mikrotasolla ideologisuus on läsnä muun muassa siinä, miten yksittäisen arvioijan asenteet, uskomukset, mielipiteet ja teoriat kielestä ja kielitaidosta vaikuttavat hänen tekemiinsä ratkaisuihin arvioinnissa ja palautteenannossa. Millaisia tehtäviä opettaja esimerkiksi päättää arvioinnissa käyttäen tai millaisiin näyttöihin nojata arvionsa?

Vaikka kielitaidon arvioinnin tulos koskettaa suoraan yksilön elämää ja usein myös hänen lähipiiriään, arvioinnin tulos ei siis koskaan perustu pelkästään kielenoppijan suoriutumiseen tai opetuksen onnistuneisuuteen, vaan arviointia koskevilla, ideologisesti latautuneilla sopimuksilla ja toteutuneilla arviointikäytännöillä on olennainen rooli arviointitulosten muodostumisessa. Näitä on tärkeää oppia tunnistamaan ja tarkastelemaan kriittisesti, jotta ideologisuuden vaikutuksia esimerkiksi omaan opetukseen ja arviointikäytäntöihin voi kontrolloida.

Koulussa toimiva opettaja tai testijärjestelmässä työskentelevä arvioija voi esimerkiksi pysähtyä tarkastelemaan, nojautuuko hän omassa arvioinnissaan arviointikriteereihin vai myös näiden ulkopuolisiin tekijöihin tai painottaako hän kielitaidon arvioinnissa tiettyjä seikkoja. Molemmissa ympäristöissä on myös tärkeää pohtia, millaista ideologisuutta arvioinnin suunnitteluun liittyviin valintoihin sisältyy. Millaisia ”annettuja” ideologioita esimerkiksi oppimateriaaleihin, testitehtäviin ja arviointiohjeisiin sisältyy ja miten näitä ideologioita samalla käytön kautta vahvistetaan? Millaisia vaihtoehtoja on käytettävissä tai tarvittaisiin?

Käytännön esimerkkeinä kielitaidon arvioinnin ideologisuudesta tarkastelemme tässä artikkelissa ensinnäkin makrotason määrittelyjä: millaisia eksplisiittisiä tai implisiittisiä taidon ihanteita eurooppalaiseen viitekehykseen on rakentunut (luku 4).

Toiseksi tarkastelemme mikrotason arviointia, perusteluja kielitaitotasoarvioille *Rikkinäistä suomea* -hankkeen (Suomen Akatemia 315581) aineistossa, joka taas pohjaa viitekehyksenkin ihanteisiin nojaavaan suomalaiseen Yleiset kielitutkinnot -testijärjestelmään (luku 5). Lopuksi esittelemme ja pohdimme eettisesti kestävän kielitaidon arvioinnin periaatteita.

2 Kielitaidon arvioinnin ideologisuus ja roolit yhteiskunnassa

Kieli-ideologiat eivät ole irrallaan muista ideologioista. 1960- ja 1970-luvuilla amerikkalaisten kieliantropologien piirissä nousi kuitenkin omaksi alakseen kieli-ideologioiden tutkimus. Se lähti erityisesti kieleen liittyvien ideologioiden kritiikistä: tutkimus syntyi reaktiona kielten ja niiden käyttäjien arvottamista ja hierarkisoimista vastaan. Silverstein (1979: 193) ja Irvine (1989: 255) määrittelevät kieli-ideologiat kieliyhteisön jäsenten eksplisiittisiksi ja implisiittisiksi jaetuiksi käsityksiksi kielestä, sen käyttäjistä ja näiden välisistä suhteista (kieli-ideologioiden perusteista, ks. Mäntynen ym. 2012). Kieli-ideologioiden tutkimuksen perusteoksessa *Regimes of language* (2000) Kroskirty esittää, että kieli-ideologioiden ytimessä on oikeanlaisen kielen määrittely ja näiden määrittelijöiden intressien suojele. Niin kielitaidon kouluarvioinnissa kuin kielitaitotesteissä ollaan siis (kieli-)ideologioiden ytimessä: jonkun kielitaito on arvioitavana yhteisön määrittelemää oikeanlaista vasten.

Kielitaidon arvioinnin ideologisuus ilmenee muun muassa siinä, keiden kielitaitoa arvioidaan, mitkä kielet arvioitavaksi on valikoitu, millaista kielitaitoa arviointikriteereissä arvostetaan ja mihin tuloksia käytetään. Osaltaan se näkyy siinä, millaiseksi vaadittava kielitaito on määritelty: mille taitotasolle se esimerkiksi on asetettu ja miten tuota tasoa on kuvattu. Kuten kaikki tasot, myös esimerkiksi kouluarvioinnissa määritetyt arvosanakohtaiset kriteerit, kokeiden mallivastaukset ja Suomen kansalaisuushakemukseen vaadittava tyydyttävä kielitaito eli eurooppalaisen viitekehyksen mukainen B1-taso ovat abstraktioita tietynlaisesta osaamisesta. Arvioinnin aikaanakaan osaamista ei voida kuvata kokonaisesti tai aukottomasti, vaan se tavallisesti kiinnitetään joihinkin olennaisimpina pidettyihin piirteisiin. Kieli-ideologiat näkyvät siten arvioinnin kaikilla tasoilla erityisesti tavoissa hierarkisoida osaamista.

Kieli-ideologioita voisi kuvata myös systeemiksi, jolla kielenkäyttäjät yhdistävät toistensa merkeiksi kokemuksensa kielistä, kielitaidosta ja kielenkäyttäjistä ja itse kielelliset piirteet, vaikkapa jonkin ääntämisen tavan (ks. Silverstein, esim. 2003). Näin jokin yksittäinen ominaisuus voi, aluksi rajatun tietyissä kielenkäyttäjäyhteisöissä ja vähitellen laajemminkin, nousta tärkeimmäksi osaamisen merkiksi. Esimerkiksi ääntämisen alueella puhenopeutta ja sen suhdetta taukoihin pidetään yhtenä osaamisen vahvimista yksittäisistä merkeistä (esim. Cucchiari ym. 2002; Riggenbach 1991). Tällainen merkkiytyminen tapahtuu Irvinen ja Galin (2000) käsitteellistämässä

kieli-ideologisissa prosesseissa. Ensinnäkin tapahtuu samastamista (*iconization*), jossa osaaminen liitetään puheen sujuvuuteen. Samastamisprosessi edellyttää kuitenkin samalla joidenkin muiden piirteiden sivuuttamista (*erasure*). Esimerkiksi jokin normin vastainen ääntämisen tapa voidaan sivuuttaa sujuvaksi koetun puhenopeuden vuoksi.

Ideologisuus ei rajaudu vain kieliin tai kielenkäyttäjiin, vaan kielten opetus ja arviointi uusintavat ja rakentavat eetosta ihmisten välisistä ihanteellisista vuorovaikutussuhteista. Tätä tehdään muun muassa rakentamalla opetukseen ja testitettäviin halutunlaisia vuorovaikutustilanteita osallistujarooleineen sekä kuvaamalla arviointikriteereissä tavoittelemisen arvoisena pidettyjä vuorovaikutustaitoja ja jäsentämällä taitoja eri tasoille suhteessa toisiinsa. Vuorovaikutusta ja toimintaa korostavan kielitaitokäsityksen vahvistumisen myötä on nähtävissä, että niin opetuksessa kuin arvioinnissakin kiinnitetään yhä enemmän huomiota merkityksiä välittävään ja rakentavaan vuorovaikutukseen. (Esim. Euroopan neuvosto 2020a; OPH 2022b; ks. myös Huhta ym. tässä teemanumerossa.) Arvostettua ei ole pelkästään virheetön kielenhallinta, valmiiden tekstien omaäänisyys tai genrenmukaisuus, vaan painoarvoa saa myös se, miten kielenoppija osallistuu yhteiseen toimintaan: onko hän aktiivinen merkitysneuvotteluissa ja sensitiivinen vuorovaikutuskumppaneitaan kohtaan tai pyrkiikö hän ylläpitämään ja luomaan myönteistä ilmapiiriä.

Kielitaidon arviointi on osa poliittista päätöksentekoa, ja useissa maissa keskeinen osa esimerkiksi koulutus- ja maahanmuuttopolitiikkaa (Shohamy 2007: 141; Bruzos ym. 2018: 433; Euroopan Neuvosto 2020b). Eryteisesti valtakunnallisilla päätösarvioinneilla sekä ns. high-stakes-testiarvioinnilla, jonka tuloksella on suuri vaikutus osallistujan elämään, on monissa maissa sekä yksilöllistä että yhteiskunnallista merkitystä, sillä tuloksia voidaan käyttää ehtona esimerkiksi opiskelupaikan tai työpaikan saamisessa ja oleskeluluvan tai kansalaisuuden myöntämisessä. Testiarvioinnin merkitys koulutusta sekä maahanmuuttoa koskevassa päätöksenteossa kasvoi Pohjois-Amerikassa sekä Australiassa 1980-luvun alkupuolelta lähtien (McNamara 1998: 312). Nykyään myös Euroopassa yhä useampi maa, Suomi mukaan lukien, edellyttää riittävän kielitaidon todentamista joko testiarvioinnilla tai muulla keinolla osana kansalaisuuden hakemista (Euroopan neuvosto 2020b).

Kansalaisuuteen liitettyjä kielitaitovaatimuksia on perusteltu etenkin kielitaidon mukanaan tuomilla integroitumisen mahdollisuuksilla: kielitaitoisena on helpompaa osallistua yhteiskuntaan ja toimia sen itsenäisenä jäsenenä. Kritiikkinä puolestaan on esitetty muun muassa, ettei tällaisista myönteisistä vaikutuksista ole riittävästi tutkimusnäyttöä. Lisäksi on esitetty, että kielitaitoehto toimii lähinnä yhteiskunnan olemassa olevien valtarakenteiden ja yksikielisyyden ideologian pönkittäjänä ja on haitallinen polkiessaan tästä normista poikkeavien yksilöiden kielellisiä oikeuksia ja monikielisyyttä. (Shohamy 2007; Milani 2008; Bruzos ym. 2018.) Kritiikille löytyy historiasta myös tukea, sillä paikallisesti, esimerkiksi Australiassa Dictation Test -testissä 1900-luvun alusta vuoteen 1958 saakka, yhteiskunnallisina portinvartijoina toimineita kielitestejä on manipuloitu niin, että niiden läpäiseminen on ollut käytän-

nössä mahdotonta ja niiden tavoitteena onkin ollut lähinnä karsia ei-toivotut tulijat (ks. Davies 2004: 101; Fulcher & Davidson 2010: 152–153; Deygers 2019).

Toisaalta on liian suoraviivaista tulkita, että kielitaitovaatimukset itsessään olisivat osoitus valtioiden yksikielisyyspolitiikasta ja monikielisyysideologian hylkäämisestä. Esimerkiksi Suomessa valtion virallisiin kansalliskieliin, suomeen tai ruotsiin, kytkettyjen kielitaitoedellytysten asettaminen ei automaattisesti tarkoita yksikielisyteen kannustamista tai muiden kielten syrjäyttämistä. Voidaan ajatella, että demokratiassa elävän yksilön oikeudet toteutuvat parhaiten silloin, kun hän pystyy toimimaan kaikilla keskeisten yhteisöjensä kielillä, siis myös yhteiskunnan valtakielellä tai -kielillä. Näiden taito on lukuisten tutkimusten perusteella myös yhteydessä sekä yhteiskunnassa pärjäämiseen että pärjäämisen kokemuksiin (esim. Tarnanen & Pöyhönen 2011; Martelin ym. 2020).

Kansalaisille asetetusta kielitaitoehdosta huolimatta myös monikielisyden tukemiseen on löytynyt poliittista tahtoa, mikä näkyy esimerkiksi lainsäädännössä ja eri koulutusasteiden opetussuunnitelmien perusteissa (esim. OPH 2014, 2022). Jokaisen yksilön oikeus omaan kieleen ja kulttuuriin nähdään tärkeänä (Suomen perustuslaki 11.6.1999/731, 17 §). Monikielisyttä arvostetaan ja monikielisyden ylläpitämistä sekä kehittämistä tuetaan (esim. OPH 2014, 2022). Virallisista kielitutkinnoista sekä ylioppilastutkinnot että Yleiset kielitutkinnot on mahdollista suorittaa useilla kielillä lähellä monen oppijan kotipaikkaa, vaikkei se aina ole testien järjestämisen näkökulmasta helppoa tai taloudellisesti kannattavaa. Tämäkin on yhdenlainen tuki Suomen monikielisyydelle.

Monet arvioinnin peruskysymyksistä yhdistävät kielitaidon arviointia koulutuksessa ja testijärjestelmissä. Nämä kuitenkin myös poikkeavat toisistaan monella tavalla (ks. esim. Tarnanen 2014). Koulussa kielitaidon arviointi on tärkeä osa kielitaito-opetusta ja vähintään implisiittisesti läsnä kaikessa vuorovaikutuksessa. Testijärjestelmissä arviointitilaisuudet puolestaan ovat tarkasti säädeltyjä ja arviointi perustuu standardoidun menettelyn kautta kerättyihin suorituksiin. Kouluarvioinnille tyypillinen oppilaantuntemus puuttuu, mutta yhdenmukainen ja tarkasti määritettyjä laatukriteereitä noudattava arviointiprosessi puolestaan lisää arviointitulosten vertailukelpoisuutta. Tämä tukee testiarvioinnin luotettavuutta ja oikeudenmukaisuutta, mutta tekee siitä samalla kouluarviointiin verrattuna jäykkää.

Testiarviointia ei esimerkiksi voida ennalta sovittuja erityisjärjestelyjä lukuun ottamatta mukauttaa testattavien tarpeisiin, soveltaa vapaasti tai uudistaa nopeasti, vaan se edellyttää pitkäjänteistä suunnittelu- ja kehittämistyötä ja perustuu lukuisille pitkäkestoisille ja melko pysyväisluonteisille valinnoille. Tässä mielessä opettaja-arvioijalla on enemmän vapauksia kuin testijärjestelmässä toimivalla arvioijalla mutta myös enemmän suoraa vastuuta arjen arviointikäytänteistään ja sen myötä arvioinnin kautta mahdollisesti kierrättämistään ideologioista. On tärkeää kuitenkin samalla muistaa, että koulussa arvioinnin tehtävänä on arvioida kielitaidon lisäksi muun muassa sitä, miten hyvin oppija on oppinut myös kyseisen kurssin muita sisältöjä.

Vaikka ideologisuus on läsnä kaikessa arviointitoiminnassa ja sitä koskevassa keskustelussa, se ei usein ole eksplisiittistä vaan ilmenee toiminnassa ja puhumisen tavoissa (ks. myös lukua 5). Toisinaan ideologisuus on kuitenkin näkyvää, julistuksellistakin, kun ideologisoituneella puheella esimerkiksi vahvistetaan tai luodaan oletuksia arvioinnin kohteesta ("Hyvä puhuja on hyvä ihminen") tai tarkennetaan arvioinnin sidoksia ja pyrkimyksiä ("Hyvä kielitaito takaa potilasturvallisuuden", "Oppimista tukevan arvioinnin tarkoituksena on osoittaa kielitaidon laatu ja vahvuudet").¹ Ideologisen puheen välittämät käsitykset voivat olla tosia tai epätosia. Ideologioiden vaikuttavuuden näkökulmasta käsitysten totuusarvoa keskeisempää on kuitenkin niiden jaettuus yhteisö(i)ssä.

Koska ideologiat eivät ole suoraan riippuvaisia reaalimaailman asiantiloista, eri yhteisöjen kieli-ideologiat voivat poiketa toisistaan huomattavasti. Kielitaidon arvioijat niin koulussa kuin arviointijärjestelmissä kantavat kokemuksiaan esimerkiksi omilta kouluajoiltaan, opiskeluvuosiltaan ja (työ)yhteisöistään (Ruohotie-Lyhty & Kaikkonen 2009; Honko 2018) niihin ehkä tiedostamattaan mukautuen, niitä vaalien tai vastustaen. Siksi samassakin työyhteisössä toimivien asiantuntijoiden arviointiin liittyvät näkemykset ja toimintatavat voivat poiketa toisistaan (Eckes 2008; Tarnanen 2014). Lisäksi tilanteiset ja yksilölliset tekijät, kuten kiire, tietämättömyys ja välinpitämättömyys tai tutuksi ja turvalliseksi koettujen ajattelutapojen suosiminen, voivat jarruttaa uusien periaatteiden siirtymistä hyväksi käytänteiksi. Säännöllinen keskustelu arviointilinjauksista ja konkreettisista arvioinneista puolestaan rakentaa yhteistä arviointikulttuuria ja tukee siihen sitoutumista (ks. myös Tarnanen 2014; Tossavainen & Ahola 2014; Ahola 2022).

3 Kriittisen kielitaidon arvioinnin perinne

Arvioinnin periaatteiden avaaminen ei automaattisesti takaa niiden toteutumista tai totuudenmukaisuutta. Tämä johtuu ensinnäkin siitä, että useimmiten arviointiympäristössä – koulussa, testijärjestelmässä, yhteiskunnassa – vaikuttaa yhtä aikaa lukuisia toisiinsa limittyviä ristiriitaisiakin ideologioita, ja kielitaitoon liittyvät ihanteet ovat sidoksissa kulttuuriseen, sosiaaliseen ja poliittiseen ympäristöön ja perinteisiin. Vaikka myös vallitsevat ihanteet muuttuvat ajan myötä, niiden tietoinen ohjailu ja muuttaminen on hidasta. Jos kielitaidon arvioinnin ideologisuutta ei tunnisteta, ymmärrys arvioinnin vaikutuksista jää vajaaksi. Tällöin voi jäädä näkemättä, millaiset asiat arvioinnissa painottuvat tai jäävät huomiotta ja miten tämä heijastuu arviointitulosten luotettavuuteen ja kielitaidon opetukseen. Ideologisuuden syntymekanismien tunnistaminen puolestaan auttaa korjaamaan mahdollisia epäeettisiä arviointikäytänteitä tai tulosten epäeettistä käyttöä.

¹ Lainaukset julkisista verkkolähteistä.

Kielitaidon arvioinnin laatua on perinteisesti kehitetty etenkin positivismin hengessä, eksaktia ja mitattavissa olevaa tietoa korostaen. Tällöin arvioinnissa ovat painottuneet erityisesti itse mittarin ominaisuudet, kuten kyky tuottaa yksiselitteisiä, usein kvantifioitavissa olevia tuloksia, joiden avulla arvioinnin kohteena olevat henkilöt on mahdollista luokitella tai asettaa keskinäiseen järjestykseen. Viimeistään 1970-luvulta alkaen rinnalle alkoi kuitenkin laajemman sosiaalitieteiden voimistumisen yhteydessä nousta uudenlaisia kysymyksiä, jotka kohdistuivat erityisesti arvioinnin eettisyyteen ja mahdollisiin seurauksiin.²

Kielitaidon arvioinnin viitekehyksessä sosiaalisia ja poliittisia näkökulmia tärkeänä pitävät julkiset kannanotot ja keskustelut yleistyivät 1990-luvun taitteessa (esim. Messick 1989; Pennycook 1990; Shohamy 1993). Hyvän (testi)arvioinnin ominaisuuksina alettiin korostaa oikeudenmukaisuutta (*fairness, validity, access, justice*) sekä myönteisiä vaikutuksia opetukseen ja oppimiseen (*washback*) ja yhteiskuntaan laajemmin (*impact*).³ Arvioinnin laadukkuuden ja tasapuolisuuden lisäksi esille nostettiin yksilön oikeus saada arvioinnista ja arviointitiedon hyödyntämisestä asianmukaista tietoa sekä oikeus kunnioittavaan kohteluun (Lynchin 1997 pohjalta McNamara 1998: 307–308). Sosiaalisen oikeudenmukaisuuden (ks. Ilikkanen ym. ja Ennsler-Kananen ym. tässä teemanumerossa) mukaan arvioinnin eettisyys merkitsee muun muassa sitä, että testaaja kantaa vastuuta testien seurauksista ja pyrkii kaikkiin käytettävissä olevin keinoin lisäämään testaamisen oikeudenmukaisuutta (Tossavainen 2016; Deygers 2019). Nämä testiarviointia koskevat periaatteet ovat yleistettävissä myös arviointiin koulussa.

Kriittisen suuntauksen tavoitteena ei ollut halvaannuttaa tai lopettaa kielitaidon arviointia, mutta arvioinnin eettisyyden varmistamiseksi kaivattiin selkeämpiä periaatteita (Daviesin 1997 pohjalta McNamara 1998). Sitten arvioinnin yhteisiin periaatteisiin tähtääviä suosituksia ja sitoumuksia ovat erityisesti testiarvioinnin käyttöön julkaisseet muun muassa kansainväliset arviointijärjestöt ILTA (the International Language Testing Association), ALTE (the Association of Language Testers in Europe) sekä EALTA (European Association for Language Testing and Assessment). Koulussa arvioinnin eettisyyttä ohjataan ensisijaisesti opetussuunnitelmien perusteilla, joilla puolestaan on perustansa kunkin koulutusasteen lainsäädännössä. Lisäksi Yleisten kielitutkintojen ja ylioppilastutkintojen kaltaisia virallisia kansallisia testejä ohjataan lainsäädännöllä ja testijärjestelmän omilla kirjallisilla perusteilla, standardeilla ja tarkentavilla menettelytapojen kuvauksilla, kuten tehtävien laadinta- ja arviointiprosessien kuvauksilla (ks. Tossavainen 2016: 36).

2 Ensimmäisten julkaistujen puheenvuorojen joukossa mm. Spolsky 1981 ja Stevenson 1981, ks. tarkemmin esim. McNamara 1998, suomeksi Tossavainen 2016.

3 Ks. tarkemmin Hamp-Lyons 1997; McNamara 1998: 308–309; Kunnan 2000: 10; Lynch 2001; Shohamy 2001. Käsitettä *impact* käytetään usein kattamaan laajasti kaikki testaamisen vaikutukset tai vaikuttavuus, kun taas *washback* viittaa tyypillisesti siihen, miten testaaminen vaikuttaa opetukseen ja oppimiseen luokkahuoneympäristössä (Tsagari & Cheng 2017).

Eettisten normien on katsottu yhtenäistävän ammattieettistä arvomaailmaa ja turvaavan arviointitoiminnan lainmukaisuutta (esim. Fulcher & Davidson 2010; Deygers 2019). Kriittikinä etenkin varhaisimmille eettisille koodistoille puolestaan on esitetty, että käytänteisiin keskittyminen on jättänyt varjoonsa kielitaidon arvioinnin laajemat sosiaaliset ja poliittiset kytkökset (McNamara 1998: 310). Yleisempien, arvioinnin moraalialia, ihanteita ja arvoja käsittelevän etiikan haasteena on pidetty yhtäältä sitä, että sen muuntaminen käytännön toimenpiteiksi on vaikeaa, ja toisaalta sitä, että kielitaidon arvioinnin eettiset kysymykset eivät ole luonteeltaan eksakteja. Tulkintoja on tehtävä myös paikallisella ja yksilötasolla, kuten alueellisissa ja koulukohtaisissa opetussuunnitelmissa, testijärjestelmissä ja yksittäisten opettajien ja arviointiasiantuntijoiden työpöydän ääressä. (Laine ym. 2023; Tossavainen 2016: 35.)

Omaantuntoon, etiikkaan ja moraaliiin nojautuvista ratkaisuista ei aina, jos koskaan, saavuteta täyttä yksimielisyyttä. Täysin ongelmaton kielitaidon arviointi onkin lähinnä hyödyllinen illuusio, joka kannustaa refleктоimaan ja kehittämään arvioinnin periaatteita ja käytänteitä (ks. myös Tarnanen 2014). Koska kielen oppimisessa on kyse sekä tiedon että taidon karttumisesta, kielitaidon mittaamisessa on aina mukana myös subjektiivisuutta, ja kielitaidon arviointia on pidetty erityisen alttiina eettisille haasteille (Davies 2004: 97; Ahola ym. 2016). Samoista syistä etenkin standardoitujen kielitestien arvioinnissa ei toisaalta myöskään pidä nojautua yksittäisen arvioijan näkemukseen (myös Atjonen 2007; Tossavainen 2016). Vaikka kielitaito-opetuksesta erillisiä testijärjestelmiä on kritisoitu paljon, niillä on myös tärkeitä tasa-arvoa tuovia etuja, kuten standardoitu, yhtenäinen menettely, arvioijien koulutukset ja seuranta sekä mahdollisuus useiden arvioijien käyttämiseen (myös Tossavainen 2016).

Kriittisen kielitaidon arvioinnin näkökulma lähtee siitä, ettei arviointi ole neutraalia tai irrotettavissa yhteiskunnallisista sidoksistaan. Esimerkiksi testiarviointiin katsotaan liittyvän monenlaista kontrollointia ja voimasuhteita myös varsinaisen arviointitapah-tuman ulkopuolella, minkä vuoksi on esitetty, että sitä tulisi tulkita, arvioida ja kehittää laajassa kehityksessään, osana sosiaalisia, koulutuksellisia ja poliittisia kytköksiään (esim. Spolsky 1997; Shohamy 2007). Tieteen kentällä kriittinen kielitaidon arviointi on silti ennemminkin kudelma toisiinsa löyhästi linkittyviä ajattelumalleja tai tieteenfilosofisia suuntauksia kuin yksittäinen malli tai teoria.

McNamaran (1998: 313) siteeraaman Pennycookin (1990) mukaan kriittinen kielitaidon testaamisen paradigma (*critical language testing*) voidaan ymmärtää ainakin neljällä tavalla: 1) tiedonalana, joka on luonteeltaan itseensä suuntautuvaa, sosiaalisesti tiedostavaa ja itseään korjaavaa, 2) oppirakennelmana, joka ylläpitää kriittisyyttä kielitaidon testiarviointia kohtaan "foucault'laisessa hengessä" eli tarkastelee muun muassa tiedon, eettisyyden, vallan ja toiminnan rakentumista, 3) suuntauksena, joka pyrkii vastaamaan kritiikkiin ja etsimään vaihtoehtoisia testaamisen ja arvioinnin tapoja sekä 4) ajattelumallina, joka näkee testiarvioinnin itsessään voivan muuttaa yhteiskuntaa.

Osa näkemyksistä on varsin kriittisiä ja kyseenalaistaa koko kielitaidon testaamisen mielekkyyden tai suhtautuu kielteisesti sen tiettyihin käyttötarkoituksiin, kuten käyttämiseen koulukontekstissa summatiivisessa arvioinnissa tai oleskeluluvan tai kansalaisuuden myöntämisen ehtona (ks. Shohamy 2001, 2007; Weir 2005; Bruzos ym. 2018). Yhteisinä ja laajemmin hyväksytyinä piirteinä suuntauksille voidaan pitää tutkimuksellista otetta arviointiin ja kriittisen tarkastelun kohdistamista arvioinnin koko elinkaareen. Elinkaari kattaa vaiheet testien suunnittelusta suoritusten arviointiin sekä arvioinnin luotettavuuden monitorointiin ja kehittämiseen. Käytännön pyrkimyksenä on lisätä arvioinnin oikeudenmukaisuutta arvioinnin kaikissa vaiheissa ja yhteiskunnan kaikilla tasoilla (Lynch 2001; McNamara 1998).

4 Eurooppalaisen viitekehysten osaamisen ideologia

Kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys (CEFR,⁴ Euroopan neuvosto 2001) ja sen kielitaitotasokuvaimet ovat jo yli 20 vuoden ajan vaikuttaneet laajamittaisesti sekä Euroopassa että muualla maailmassa erityisesti englantia toisena ja vieraana kielenä -alalla. Kuvaimista on muodostunut keskeinen kielitaidon opettamista ja arviointia globaalisti yli kielirajojen ohjaava ideologinen kehys. Tällaisena sillä on merkittävä valta-asema kaikessa kielitaidon arvioinnissa, mistä syystä sitä on voitava tarkastella erityisen kriittisesti.

Suomessa eurooppalainen viitekehys ja sen paikalliset sovellukset, kuten opetussuunnitelmiin tuodut kuvaimet (alun perin OPH 2003 lukiokoulutuksesta, 2004 perusopetuksesta), Täydentävä osa ja lisäkuvaimet (OPH 2018) sekä Kehittyvän kielitaidon tasojen kuvausasteikko (OPH 2019), ovat vaikuttaneet etenkin summatiiviseen arviointiin toisen ja vieraan kielen opetuksessa 2000-luvun alusta saakka (ks. myös OPH 2014 perusopetuksesta, 2015 lukiokoulutuksesta ja OPH 2022b kotoutumiskoulutuksesta). Lisäksi Yleisten kielitutkintojen testiarvioinnissa on vuodesta 2003 saakka käytetty viitekehukseen pohjautuvia puhumisen ja kirjoittamisen taitotasokuvaimia (OPH 2011; Laki Yleisistä kielitutkinnoista 12.11.2004/964). Koska eurooppalaista viitekehystä ei ole luotu erityisesti arviointiin, sen soveltaminen niin testiarviointiin kuin kielitaidon arviointiin koulussa tai koulutuksessa perustuu viitekehysten yleisille periaatteille, joista keskeisimpiä käsittelemme seuraavaksi.

Yhteisessä eurooppalaisessa viitekehyksessä painotetaan kielen merkitystä toiminnan ja vuorovaikutuksen välineenä, taitoa tiedon sijaan. Viitekehys ei ole sitoutunut yksittäisiin kielitieteen tai koulutuksen alan teorioihin tai käytänteisiin, muttei se toimi myöskään irrallaan esimerkiksi toisen kielen oppimisen teoreettisista suuntauksista. (Euroopan neuvosto 2001, 2020a; North 2021.) Brian North (2021: 9–10) esittelee viitekehysten (2001, 2020) teoreettista taustaa kuvaamalla siihen vaikuttaneita ja toisiinsa

4 *The Common European Framework of Reference for Languages: Learning, teaching, assessment*

kietoutuneita teoreettisia suuntauksia ainakin sosiolingvistiikan, diskurssianalyysin, toisen kielen oppimisen sekä kielitaidon arvioinnin aloilta. Näiden pohjalta keskeistä eurooppalaisen viitekehyksen kielikäsitteilyssä on kielen näkeminen kontekstisidonnaisena toimintana ja vuorovaikutuksena.

Uudistustyöhön ovat sittemmin vaikuttaneet sekä yhteiskunnalliset ja globaalit muutokset että tieteen paradigmuutokset. Viitekehyksen uudistuksen yhteydessä korostetaan entistä voimakkaammin esimerkiksi monikielisyyttä ja -kulttuurisuutta, mediaatio- ja verkkovuorovaikutustaitoja sekä kielellisten taitojen kokonaistarkastelun merkitystä (myös Piccardo & North 2019; North 2021). Uudistuksen taustalla onkin nähtävissä ainakin kielenoppimisen dialogisuuden, sosiaalisen osallisuuden ja toimijuuden painotuksia.

Koska kuvainten on haluttu edistävän tavoitteellista oppimista eikä niinkään oppijoiden vertailua, kuvaimissa on päädytty osaamisperusteiseen esittämistapaan (*can dos*) eli sen kuvailuun, mitä kielenoppija kullakin taitotasolla jo osaa tehdä – sen sijaan, että kuvailtaisiin kielitaidon puutteita (ks. esim. Euroopan neuvosto 2001). Näin on tehty, vaikka onkin tiedostettu, että etenkin kielenoppimisen alkuvaiheessa osaamista on vaikeampi kuvata kuin sen puuttumista (mts. 205–206).

Lisäksi kuvaimissa on pyritty selkeyteen ja mahdollisimman suureen yksiselitteisyyteen sekä tiiviiseen esittämistapaan ja siihen, että yksittäiset kuvaimet toimivat myös itsenäisesti. Kuvaimissa suositaankin lyhyitä, myöntömuotoon kirjattuja väitelauseita, joissa kielellisiä taitoja kuvataan kykynä suorittaa konkreettisia toimintoja (”Can relate the plot of a book or film and describe his/her reactions.”; Euroopan neuvosto 2001: 59). Kuvainten tekijät ovat tietoisia, että nämä valinnat, kuten myös kuuteen taitotasoon päätyminen, ovat yksi monien joukosta ja että nämäkään eivät ole täydellisiä. Viitekehyksen käyttäjiä onkin haastettu paitsi refleктоimaan toimintaansa myös kriittisyyteen kuvainten käytössä. (Euroopan neuvosto 2001: xiii-xiv.) Alusta saakka on myös ollut selvää, että viitekehys on avoin muutoksille ja jatkuvalla kehittämiselle (esim. North 2021: 8).

Viitekehystä ja sen keskeistä asemaa käsittelevän kirjansa esipuheessa Little ja Figueras (2022: xv) kuvaavat, kuinka eurooppalainen viitekehys on vakiinnuttanut asemansa korvaamattomana vertailukohtana toisen ja vieraiden kielten koulutuksen kaikilla osa-alueilla. Siitä on muodostunut kielitaidon opetussuunnitelmatyön, opettamisen ja arvioinnin selkänöjä, jonka tehokasta ja systemaattista käyttämistä pidetään paitsi mahdollisuutena myös Euroopan neuvoston jäsenmaiden velvollisuutena (Euroopan neuvosto 2008; Goullier 2007). Toisaalta käyttöä ei ohjata yksityiskohtaisesti, ja kriittistä keskustelua on herätetty muun muassa taitotasokuvainten alkuperäistä tarkoitusta laajemmasta soveltamisesta.

Eurooppalaisen viitekehyksen käyttöä maahanmuuttokonteksteissa on moitittu esimerkiksi siitä, että viitekehys on alun perin kehitetty ”perinteisempiin” (*more orthodox*) kielen oppimisen tarpeisiin kuin mitä esimerkiksi maahanmuuttaneiden kielitaitotarpeisiin olisi tarvittu (Bruzos ym. 2018: 433 Krummiin 2007 viitaten).

Viitekehys ei myöskään itsessään ole sisäisesti täysin eheä ja ongelmaton. Kuvaimiin tulee aina sisältymään sellaista julkaisuaikaansa liittyvää ideologista painolastia, jonka havaitseminen mahdollistuu ja ymmärtäminen syvenee vasta myöhemmin, uudessa tulkintakehyksessä.

Varsinkin taitotasoajattelu on omaksuttu laajalti käyttöön, mutta monien muiden eurooppalaisen viitekehysten keskeisten sisällöllisten pyrkimysten siirtyminen kielitaidon opetuksen ja arvioinnin arkeen on ollut huomattavasti hitaampaa (de Jong 2022; Levy & Figueras 2022). 2020-luvun alussa (*CEFR Companion Volume*) keskeiseksi nousseiden mediaatiotaitojen eli merkityksiä välittävän ja rakentavan vuorovaikutuksen (suomeksi ks. Salo ym. 2020) sekä monikielisen ja monikulttuurisen kompetenssin suhde testiarviointiin vasta hakee muotojaan.

Lisäksi on olennaista huomata, että kuvaimiin sisältyy piirteitä, joiden käyttö esimerkiksi yleisessä kielitestissä voi paitsi lisätä arvioinnin tasa-arvoa myös heikentää sitä. Kuvainten moniulotteisuuden haasteita sekä tunneällyn ja monikielisyys- ja mediaatiotaitojen johdonmukaisen operationalisoinnin haasteita ovat tarkemmin käsitelleet mm. Leung ja Jenkins (2020), de Jong (2022) sekä Leung (2022). Leung ja Jenkins (2020) kirjoittavat näiden taitojen olevan luonteeltaan sellaisia, että niiden asettaminen asteikolle ei ole yksiselitteistä, koska vuorovaikutustilanteissa kaikkien osallistujien toiminnalla on olennainen rooli siinä, millaiseksi vuorovaikutus muodostuu ja millainen yksittäisen osallistujan taito tilanteessa riittää tai ei riitä.

Vaikka kuvaimet tunnetaan ehkä parhaiten arviointivälineinä, on tärkeää muistaa, ettei viitekehystä alun perin luotu ainoastaan tai edes ensisijaisesti kielitaidon arvioinnin tarpeisiin, vaan tavoitteena oli tarjota välineitä laadukkaaseen opetus suunnitelmatyöhön, yhdenmukaisemman opetuksen ja oppimateriaalien kehittämiseen taitokuvauksia hyödyntäen ja ylipäättään koulutuksellisen tasa-arvon lisäämiseen (Euroopan neuvosto 2001, 2020a; North 2021, 2022). Siksi kuvainten käyttäjiä muistutetaan siitä, että CEFR:iä ei ole tarkoitettu oikeuttamaan arviointityökalujen käyttämistä portinvartijan tehtävissä (Euroopan neuvosto 2020a: 11; North 2021: 6). Viitekehysten operationalisointi poliittiseksi työkaluksi ei myöskään saisi olla automaattista. Sen sijaan keskeinen tavoite on ollut tehdä näkyväksi kielitaidon (*language proficiency*) kompleksinen luonne ja tarjota välineitä, metakieli, sen käsittelemiseen.

5 Osaamisen ideologia keinona irtautua syntyperäisen kielenkäyttäjän ihanteesta

Eurooppalaisen viitekehysten ensimmäisessä versiossa (2001: 17) syntyperäinen kielenkäyttäjä esitettiin useissa kielitaitokuvaimissa oppimisen tavoitteena, vaikkakin samalla korostettiin, että ihmisten väliset kompetenssierot ovat suuret riippumatta siitä, ovatko he kielen syntyperäisiä vai oppineita käyttäjiä. Näin siis jo yli 20 vuotta sitten tavoitteeksi asetettiin muun muassa oppijan aktiivisen toimijuuden ja monikie-

lisyiden tukeminen sekä tietopainotuksesta ja natiivi-ihanteesta luopuminen, mutta käytännössä muutokset olivat ainakin paikallisesti jääneet vähäisiksi (North 2021: 10).

Uusimmassa versiossa, Companion Volumessa (CV), syntyperäisen puhujan ihanteesta on irrottauduttu tietoisesti ja läpinäkyvästi. CV esittelee viitekehyksen uudistamisprosessissa tehdyt muutokset eksplisiittisesti. Liitteen 7 taulukkoon (Euroopan neuvosto 2020a: 257) on koottu näkyviin kaikki kuvainmuutokset tasoille A1–C1. Yksi muutoksista on juuri natiivi(puhuja)viittausten poisto, jota perustellaan tekstiosion alussa:

- (1) *Changes are also proposed to certain descriptors that refer to linguistic accommodation (or not) by “native speakers”, because this **term has become controversial** since the CEFR was first published.*

(Euroopan neuvosto 2020a: 24; lihavointi tämän artikkelin tekijöiden).

Tekijät kuvaavat nimenomaan termin, ei käsitteen, *natiivi* muuttuneen kiistanalaiseksi. Muotoilu kuvastaa kehittämisvaihetta, jossa pyrkimyksistä huolimatta ollaan vielä kiinni syntyperäisen kielenkäyttäjän mallissa. Esimerkiksi kuvaimissa edelleen toistuvien luonnehdintojen *luonnollinen*, *normaali* tai *idiomaattinen* voidaan kuulla kaiuttavan syntyperäisyyttä. Natiivin käsite on ollut keskeinen niin kauan kuin on ajateltu, että kielenoppija on ikään kuin kompetentti tai tasa-arvoinen kielenkäyttäjä aina vasta ”joskus myöhemmin”. Sittemmin on avauduttu ajatukselle, että kieltä voi ensinnäkin käyttää jo vasta matkalla ihanteeseen ja että syntyperäisille kielenkäyttäjille yhteinen kielitaito on illuusio.

Paikeday (1985) julisti jo lähes neljäkymmentä vuotta sitten ”natiivin olevan kuollut”. Julistuksen jälkeen muiden muassa kielenoppimisen tutkija Cook (1999: 189; 2003) on huomauttanut, että natiivipuhujan käsite on juurrutettu niin syvään, että siitä on vaikea päästä eroon, vaikkei oppijoita pitäisikään käsitellä natiivipuhujien imitaattoreina. Cookin tapaan kielenoppimisen ja arvioinnin tutkijat Isaacs ja Rose (2022) huomauttavat, että itse käsitteestä eroon pääsemiseen, ”post-natiiviin aika-kauteen”, on vielä pitkä matka. Tähän pääsemiseksi tarvitaan heidän mukaansa uutta läpinäkyvää terminologiaa, käsitteitä ja arviointikriteereitä.

Kokonaan oma kysymyksensä ja tälläkin hetkellä kiistelty aihe puolestaan on, onko syntyperäisen puhujan mallista syytäkään kokonaan luopua – tai miksi se koetaan ongelmalliseksi. Yksi peruste luopumiselle on, että natiiviutta on erittäin vaikea operationaalistaa, sillä ihmisten ”osaaminen” ei ole samanlaista esimerkiksi kansalaisuuden, äidinkielen tai asuinalueen perusteella eivätkä ”natiivitkaan” siis keskenään samanlaisia tai täydellisiä kielenkäyttäjiä (Davies 2013). Syntyperäisen kielenkäyttäjän ihanne istuu kuitenkin tiukassa muun muassa siksi, että se juontaa kolonialistisen ajan voimasuhteiden perinnöstä, josta luopuminen edellyttäisi globaalien valtasuhteiden uudelleenjärjestäytymistä (ks. esim. Tupas 2022).

Yksi keino, jolla täydellisen kielentaitajan ihanteesta on pyritty irrottautumaan, on kielen ymmärrettävyyden (*intelligibility*) ja käsitettävyyden (*comprehensibility*) erottaminen. Ymmärrettävyydellä on viitattu tässä keskustelussa siihen, mitä vastaanottaja todella ymmärtää tuotoksesta, käsitettävyydellä (huom. vapaa käänös) puolestaan siihen, kuinka työlästä ymmärtäminen on eli minkä verran vastaanottaja tietoisesti pyrkii ymmärtämään. (Ks. Derwing & Munro, esim. 1997.) CV:ssä onkin korostettu juuri käsitettävyyttä esimerkiksi puhumisen taitojen, fonologisen kontrollin, kuvaimissa, mutta myös kirjoittamisen konventioiden yhteydessä (Euroopan neuvosto 2020a: 133–135, 136).

Useat tutkimukset (ks. esim. Fayer & Krasinski 1987; Kang & Rubin 2009; Lindemann 2006) ovat osoittaneet, että oppijan kielen vastaanottamisessa vaikuttavat usein varsinaiseen ymmärtämiseen liittymättömät seikat, esimerkiksi oppijan oletettuun taustaan liittyvät asenteet ja stereotyyppit. Ymmärtämisen sijaan käsitettävyyden huomioiminen kielenoppimisessa ja arvioinnissa olisikin tärkeää, koska se siirtää vastuun vuorovaikutuksen onnistumisesta myös vastaanottajalle.

Kriteerinä käsitettävyyks on kuitenkin vaikeasti operationaalistettava, erityisesti, koska se perustuu subjektiiviseen kokemukseen. Kielitaidon arvioijat, kuten muutkin ihmiset, poikkeavat toisistaan esimerkiksi siinä, kuinka harjaantuneita he ovat kuuntelemaan erilaisia varieteetteja ja aksentteja. Arvioinnin laadun kannalta on tärkeää, että arvioija ei juutu yksittäisiin piirteisiin, kuten johonkin aksenttiin ja sen herättämiin mielikuviin, vaan noudattaa arviointikriteereitä, reflektoi omaa arviointiaan ja pyrkii kuuntelemaan puhetta ”aksentin läpi”. (Ks. Ahola 2022: esim. 78–79.)

Siirtyminen natiivi-ihanteesta kohti ”osaamisen” ihannetta näkyy niin makrotasolla yleiseurooppalaisessa kielitaidon kuvaamisessa, mesotasolla arviointikriteerien kuvauksissa kuin mikrotasolla taitoarvioiden perusteluissa. Osaamisen painottaminen näkyy esimerkiksi käytännön arvioinnissa Yleisten kielitutkintojen (=YKI) kontekstissa Soveltavan kielentutkimuksen keskuksessa toteutetun *Rikkinäistä suomea* -hankkeen aineistossa.

YKIn taitotasojen kuvaimissa kielitaito kuvataan pääasiassa osaamisena (*selviää, pystyy, osaa*), mutta myös jonkin osataidon puuttumisena, mikä havainnollistaa Euroopan neuvoston (2001: xiii - xiv, 205–206) ennakoimaa vaikeutta kuvata pelkästään osaamista. (Ks. myös edellistä lukua 4.) Tasolla 3, kynnystasolla (joka vastaa EVK:n tasoa B1), puuttumista kuvataan joustavuuden, koherenssin ja sidoksisuuden sekä ääntämisen kuvaimissa:

- (2) *Selviytyy arkipäivän viestintätilanteista hyvin, mutta muodollisen ja epämuodollisen rekisterin vaihtelu on puutteellista* (joustavuus)
- (3) *Pystyy yhdistämään ilmauksia yhtenäiseksi ja johdonmukaiseksi puheeksi, vaikka sidoskeinojen käyttö voi olla puutteellista ja toistavaa.* (koherenssi ja sidoksisuus)
- (4) *Ääntäminen on selvästi ymmärrettävää, vaikka siinä on puutteita* (ääntäminen/ fonologinen tarkkuus)

Tarkastelimme osaamisen ja puuttumisen suhdetta empiirisesti *Rikkinäistä suomea*-hankkeen aineistosta. Hanke perustuu Yleisten kielitutkintojen suomen kielen puhutun kielen osakokeen kertomisen tehtävästä saatuihin autenttisiin suorituksiin. Hankkeessa 44 koulutettua YKI-arvioijaa arvioi puhumisen suorituksia YKI-kriteereillä eli perustuen kokonaisarvionsa holologista tarkastelua edellyttävään yleiskriteeriin ja kuuteen analyttiseen kriteeriin: sujuvuuteen, joustavuuteen, koherenssiin ja sidoksisuuteen, sanaston laajuuteen ja idiomaattisuuteen sekä ääntämiseen. Arvioitavia suorittajia oli kaikkiaan 49.

Kriteerien kuvaamat osataidot arvioijat sijoittivat asteikolle 1–6, jossa 3 vastaa eurooppalaisen viitekehyksen tasoa B1. Lisäksi arvioijat kuvasivat puhujaa ja/tai tämän suoritusta avovastauksissaan sanallisesti. Näistä vastauksista tarkastelimme sellaisia, joissa arvioijat kielensivät joko puhujan *osaamista, pystymistä, selviämistä* tai näiden yhdistelmiä, esimerkiksi *pystyy selviämään*. Viitekehyksen ytimessä olevan osaamisen painotuksesta huolimatta sitä kuvailtiin vähän: *osata-tapauksia* oli aineistossa 57, *pystyä-tapauksia* 91 ja *selvitä-tapauksia* 120. Tällaista osaamisen kuvausta sisältyi siis vain noin joka kymmenenteen (12 %) arvioon. Osassa kuvauksista nostettiin esiin nimenomaan osaamisen *puute* eli 'ei osaaminen', 'ei pystyminen' tai 'ei selviytyminen'. Puutteen kuvauksia esiintyi aineistossa vähemmän kuin osaamisen kuvauksia mutta silti lähes kymmenesosassa tapauksia (n. 8 %). Osaamisen ja puutteiden rinnakkaista kuvailua havainnollistavat esimerkiksi seuraavat tapaukset:

- (5) *yllättävästi asiasta saa selvää, vaikka sekä rakenteissa, sanastossa että ääntämisessä on selvästi puutteita.*
- (6) *rakenteissa paljon puutteita, mutta pystyy ilmaisemaan sanottavansa ymmärrettävästi ja melko laajasti.*
- (7) *Ymmärrettävää, viesti käy selväksi, taukoja ja puutteita rakenteissa on, mutta eivät vaikuta ymmärtämiseen.*

Osaamisen puuttumiseen keskittymisessä on vaarana, ettei arviointi perustu kaikkien kriteerien tasapainoiseen käyttöön ja suoritukseen kokonaisuudessaan, vaan yksittäiset kriteerit, esimerkiksi rakenteiden hallinta, saavat liikaa painoarvoa (ks. myös Tossavainen & Ahola 2014). Suomen kaltaisessa synteettisessä kielessä, jossa taivutusta ja siihen liittyvää morfofonologista vaihtelua on erityisen paljon, on odotuksenmukaistakin, että korkeammillakin kielitaidon tasoilla näissä esiintyy puutteita. Samasta syystä rakenteiden hallintaa tai sen puutetta olisi tarpeetonta kuvata arviointikriteereissä vain tiettyyn taitotasoon liittyvänä ilmiönä.

Käytännössä ylipäätään missä tahansa kehittyvässä taidossa on erittäin vaikea täsmällisesti kuvata, mitata ja arvioida osaamista suhteuttamatta sitä taidon seuraavaan vaiheeseen. Kielitaidon taitotasoarvioinnin tapaan esimerkiksi peräkkäisten uintimerkkien suorittaminen perustuu samantyyppiseen kriteeriperusteiseen osaamisen

arviointiin, jossa osaamisen arvioimisessa väistämättä joudutaan tarkastelemaan myös sitä, mihin uimataito mahdollisesti ei vielä riitä. Luonteeltaan moniulotteisemman kielitaidon ”osaamisesta” on vielä vaikeampi saada otetta: mistä osaaminen alkaa?

Toiminta on käsitteenä osaamista yksiselitteisempi sikäli, että sen onnistuminen voidaan todeta toiminnan lopputuloksesta, jonkin tapahtumisesta. CEFRin ihanteet ovat menossa siihen suuntaan, että osaamisen sijaan ja lisäksi kielitaitoa kuvataan niinä toimintoina, mihin oppija kielellä pystyy. Jos jo osaamisen kuvaaminen on kuitenkin vaativaa, yksinomaan toiminnan onnistumisen operationaalistaminen tasapuolisesti ja oikeudenmukaisesti arvioitavaksi kohteeksi olisi sekin äärimmäisen vaikea, ellei mahdoton, tehtävä.

6 Kohti mahdollisimman oikeudenmukaista arviointia

Ideologisuuden merkitys korostuu testijärjestelmissä, joista saatuja tuloksia käytetään yhteiskuntien portinvartijoina määrittämässä kielitaidon riittävyttä esimerkiksi opiskelupaikan, ammatillisen kelpoisuuden tai kansalaisuuden saamisessa. Vaikka testijärjestelmät itsessään eivät voi säädellä testaamisen vaikutuksia järjestelmän ulkopuolella, viralliset arviointijärjestelmät, kuten Yleiset kielitutkinnot, vaikuttavat sekä tutkintoihin osallistuvien elämään, heidän lähiyhteisöönsä että yhteiskuntaan.

Lisäksi testijärjestelmien arviointikäytänteet vaikuttavat suoraan myös kielikoulutukseen (esim. Huhta & Hildén 2016; Huhta ym. tässä teemanumerossa). Jos oppimisen tavoitteena on kielitaitotestin läpäisy, testien kynnystasot ja testikäytänteet ohjaavat opetusta. (Ks. *washback*-ilmiöstä tarkemmin Shohamy ym. 1996; Wall 2000.) Äärimmillään vietynä testijärjestelmään valmentautuminen voi muodostua liian keskeiseksi osaksi koulutusta ja muu kuin testikielissä kynnystasoille arvioitu kielitaito voi jäädä tunnistamatta. Jos yhteiskunnassa ylipäätään annetaan liian suuri painoarvo nimenomaan kielitaidon arvioinnille, on lisäksi vaarana, että tämä tehdään muun osaamisen ja yksilöllisten tarpeiden kustannuksella.

Kielitaidon arvioinnilla on suomalaisessa yhteiskunnassa kuitenkin useita vaikiintuneita tehtäviä (Leblay ym. 2014), ja kielitaitovaatimuksia on myös yleisesti pidetty perusteltuna esimerkiksi potilas- tai työturvallisuuden takaamiseksi esimerkiksi terveydenhuollossa ja ilmailualalla (Valvira 2021; TRAFICOM 2021). Siksi ei ole näköpiirissä, että arvioinnin merkitys muuttuisi ainakaan nopeasti tai että arvioinnista osana koulutusjärjestelmää tai erillisinä testijärjestelminä luovuttaisiin (ks. myös esim. Shohamy 2001: 160–162; Fulcher & Davidson 2010: 138). Arvioinnista luopumisen skenaarioiden sijaan keskitymmekin seuraavaksi siihen, miten tulevaisuudessa voidaan pyrkiä varmistamaan arvioinnin oikeudenmukaisuus.

Tietoisuus on muutoksen ensimmäinen askel (Pennycook 1999: 336), minkä vuoksi arvioinnin ideologisuuden ja etiikan tunnistaminen on olennainen osa oikeudenmukaisemman kielitaidon arvioinnin tavoittelua. Se, että yhteiskuntaa valvotaan

asettamalla kielitaitoehtoja ja mittaamalla kielitaidon edistymistä, on jo itsessään ideologinen valinta. Kielitaitoehdoilla ja niiden arvioinnilla on riskejä, sillä väärin mitoitettujen kielitaitovaatimukset tai virheellistä tietoa antava arviointi voivat molemmat johtaa monenlaisiin haittoihin sekä yksilön elämässä, hänen lähiyhteisöissään että yhteiskunnassa.

Toisaalta kielitaitoehdoilla voi olla myös positiivisia vaikutuksia niin kielikoulutukseen kuin laajemmin yhteiskuntaan (esim. Wall 2000). Arvioinnilla esimerkiksi voidaan tietoisesti edistää myönteistä kehitystä, kuten viestinnällisyyden lisäämistä kielenopetuksessa (Huhta & Hildén 2016). Taitotasoihin sidotut osaamisvaatimukset voivat myös toimia yhtenä kannustimena kielen tavoitteelliseen omaehtoiseen opiskeluun ja edistymisen seurantaan. Lisäksi ne voivat auttaa esimerkiksi sopivien koulutus- ja urapolkujen valinnassa.

Oikeudenmukaisuuteen ja tasa-arvoiseen kohteluun pyrkivässä yhteiskunnassa kielitaitoehtojen asettaminen tai kielitaidon testaaminen ei kuitenkaan saisi olla automaatio, vaan kaikkien keskeisten sidosryhmien on pohdittava ja keskusteltava, milloin testaamista todella tarvitaan ja millainen kielitaidon arviointi on eettisesti kestävä (Davies 2004; Shohamy 2007). Shohamy (2001) ja Deygers (2019) painottavat kielitaidon arviointiin liittyvän vastuun jakamista juuri sidosryhmien kanssa. Esimerkiksi Yleisissä kielitutkinnoissa vastuuta jakavat Opetushallitus, Opetus- ja kulttuuriministeriö sekä Opetushallituksen viisivuotiskausiksi asettama ja eri alojen asiantuntijoista koostuva tutkintotoimikunta, jonka tehtävänä on muun muassa käsitellä tutkintojen kehittämiseen liittyviä periaatteellisia asioita ja tehdä aloitteita tutkintojen kehittämiseksi (Laki yleisistä kielitutkinnoista 12.11.2004/964).

Arvioinnin käyttötarkoituksista riippumatta on huolehdittava arvioinnin oikeudenmukaisuudesta ja sovittujen eettisten periaatteiden noudattamisesta. Kriittisen kielitaidon opetuksen ja arvioinnin perinteessä pidetään hyveenä ajattelutapaa, jossa vallitsevia käytänteitä sekä omaa osaamista kyseenalaistetaan ja koetellaan jatkuvasti (Pennycook 1999). Kielitaidon arvioinnin kannalta tämä tarkoittaa muun muassa sitä, että arviointituloksiin ei uskota kriittikittävästi sen enempää koulussa kuin testijärjestelmissäkään, vaan varmistetaan, että kaikki keskeiset näkökulmat on huomioitu arviointitulosten syntymisessä ja tulosten perusteella tehtävissä tulkinnoissa (Fulcherin 2015: 172 pohjalta Tossavainen 2016: 38).

Koska tehtävä on vaativa, sen suorittaminen edellyttää yhteistyötä ja avointa keskustelua arvioinnista. Lisäksi tärkeää on huolehtia siitä, että arvioinnin kohde ja tavat vastaavat todellisia kielenkäyttötarpeita (mm. Shohamy 2007). Kielitaidon arvioinnin on oltava avointa kielikäsitteiden uudistumiselle ja sen pohtimiselle, millainen kielitaito – kuten minkä kielten osaaminen – kulloinkin aidosti on tarpeen esimerkiksi arjen vuorovaikutustilanteissa, opinnoissa ja työssä. Opettaja hyötyy erityisesti omien kieli- ja kielitaitokäsitteensä sekä arviointikäytänteidensä reflektoinnista (ks. myös Huhta ym. tässä teemanumerossa).

Käytännön tasolla vastuu eettisyydestä kuuluu ensisijaisesti testijärjestelmille ja koulutuksen osalta kielikoulutusta järjestäville organisaatioille, minkä vuoksi toiminnan on oltava ammattimaista (Davies 2004: 98; Fulcher & Davidson 2010: 141). Näissä on huolehdittava henkilöstön ammattitaidosta ja säännöllisesti kouluttamisesta sekä varsinkin testijärjestelmissä myös arvioijakäyttäjätymisen jatkuvasta seurannasta (myös Huhta 2014). Lisäksi vastuu tarkoittaa jokaisen ammattilaisen vastuuta oman osaamisensa käytöistä (esim. Davies 2004: 102; ILTA 2018[2000]: Code Principle 4), koulussa siis opettajan vastuuta omasta arviointitoiminnastaan.

Toisaalta eettisyyden vaatimus koskee prosessin kaikkia muitakin vaiheita ja osapuolia, kuten päättäjiä kielipoliittisen järjestelmän eri tasoilla, testien järjestäjiä, arviointitulosten vastuullisia tulkitsijoita sekä arvioitavia itseään niiltä osin kuin myös he voivat vaikuttaa esimerkiksi arviointitulosten luotettavuuteen ja käyttöön. Tässä toimijoiden verkostossa on tehtävä kaikki mahdollinen, jotta kielitaidon arviointi olisi läpinäkyvää eikä ajaisi kyseenalaisia piiloagendoja tai ylläpitäisi tai synnyttäisi epätasa-arvoa. (Shohamy 2007; Fulcher & Davidson 2010; Tossavainen 2016; Deygers 2019.) Juuri läpinäkyvyydestä huolehtiminen on keskeistä arvioinnin laadun, oikeudenmukaisuuden ja reiluuden varmistamisessa. Sitä voidaan lisätä tarjoamalla tietoa arvioinnin perusteista ja käytännön toteutuksesta sekä erityisesti erillisen testiarvioinnin osalta myös arvioinnin hallinnollisista ja taloudellista kytköksistä julkisesti, saavutettavassa muodossa.

Arvioinnin laadukkuus ja eettisyys eivät perustu yksittäisten tehtäväpakettien tai arviointikertojen ominaisuuksiin, vaan pitkäjänteiseen kehittämistyöhön. Siksi yhteiskunnassa tarvitaan kielikoulutukseen liittyvän arvioinnin ja testijärjestelmien säännöllistä ulkoista arviointia ja arviointitutkimusta sekä kielipoliittista keskustelua. Lisäksi arviointia suorittavissa yhteisöissä – kouluissa ja arviointijärjestelmissä – tarvitaan ajantasaista tietoa kielenoppimisesta ja arvioinnista sekä tähän nojautuvaa ja eteenpäin suuntautuvaa itsearviointia. Tällaiseen itseä uudistavaan toimintatapaan kuuluu sen hyväksyminen, että käsitykset hyvästä kielitaidosta ja arvioinnista ovat muuttuvia.

Arviointia tulee siis tehdä yhteisten sitoumusten hengessä ja tarpeen mukaan myös päivittää tutkimuksen ohjaamana. (Ks. myös Davies 2004: 104; Fulcher & Davidson 2010: 139; Xi 2010: 148.) Esimerkiksi monikielisten ja monimodaalisten resurssien johdonmukaisempi huomioiminen on noussut yhdeksi kehittämisen kohteeksi kielitaidon arvioinnissa. Kehittäminen edellyttää asiantuntemusta, halua ja resursseja yhtäältä olemassa olevien käytänteiden tarkasteluun, toisaalta niiden uudistamiseen (esim. Huhta & Hildén 2016; Tossavainen 2016).

Tutkimuspohjaista tietoa kielitaidon arvioinnin vaikutuksista tarvitaan edelleen lisää, mutta tärkeää on myös hyödyntää jo olemassa olevaa tietoa ja pysähtyä refleктоimaan arviointia uudenslaisista näkökulmista. Tossavainen (esim. 2014, 2016: 38) huomauttaa, että kielitaidon arvioinnin kentällä tulisi entistä paremmin kuulla arvioitavien ääni ja olla avoimempi kertomaan arvioinnin tarkoituksesta, seurauksis-

ta ja käytetyistä kriteereistä myös arvioitaville itselleen. Vaikka Tossavaisen artikkeli koskee erityisesti testijärjestelmiä, vastaava läpinäkyvyys on tärkeää myös koulussa. Shohamy (2007: 147) puolestaan on peräänkuuluttanut tutkimukseen perustuvaa työskentelyä, jossa kielitaidon arvioinnin potentiaalisia seurauksia, väärinkäyttöä, eettisyyttä, valta-asetelmia, vinoumia ja syrjintää johdonmukaisesti tarkasteltaisiin kriittisen kielitaidon teorian näkökulmasta.

Suomessakin kaivataan tällaista ns. impact-tutkimusta, joka fokusoi kielitaidon arvioinnin seurauksiin oppilaan ja kielitaitotestin osallistujan näkökulmasta. Myös tutkimuksen menetelmällinen kehittäminen olisi tärkeää. Arviointi on aiheena merkittäväyhtensä vuoksi kuitenkin arkaluonteista niin arvioitavien kuin arvioijien kannalta. Tämän vuoksi on vaikeaa saada tutkittavaksi luotettavaa empiiristä aineistoa todellisista arvioinneista, ja erityisesti pitkäaikaistutkimusten tekeminen arvioinnin yksilöllisistä vaikutuksista on vaikeaa.

Kriittinen ajattelu ja tutkiva ote arviointityöhön sopivat erillisten testijärjestelmien lisäksi erityisen hyvin kouluun, jossa arviointi ohjaa opetusta ja monin tavoin vaikuttaa sekä oppimiseen että oppijoihin. Esimerkiksi arviointitapojen ja -kriteerien valinnassa ja arviointiin liittyvässä palautteenannossa on tärkeä tiedostua näihin sisältyvistä ideologisista kytköksistä. Opettajalle tämä tarkoittaa aika ajoin vaikkapa sen selvittämistä, mitä kielitaidon tai vuorovaikutuksen piirteitä oma arviointi korostaa ja mitä taas jättää syrjään ja miten hyvin oppilaat ovat perillä siitä, mitä heiltä odotetaan. Jos arviointia ei pysty perustelemaan, siinä voi olla tarkistamisen varaa. Selkeä, jaettu näkemys oppimisen ja arvioinnin tavoitteista puolestaan tukee sekä opettamista, arvioimista että oppimista.

Kirjallisuus

- Ahola, S. 2022. *Rimaa hipoen selviää tilanteesta. Yleisten kielitutkintojen suomen kielen arvioijien käsityksiä kielitaidon arvioinnista ja suullisesta kielitaidosta*. Jyväskylä: Jyväskylän yliopisto.
- Ahola, S., R. Neittaanmäki & T. Hirvelä 2016. Puheen ymmärtämisen tehtävien taitotasolle asettaminen Yleisissä kielitutkinnoissa. Teoksessa A. Huhta & R. Hildén (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa. AFinLA-e: soveltavan kielitieteen tutkimuksia 9*. Jyväskylä: Suomen soveltavan kielitieteen yhdistys, 68–88. <https://journal.fi/afinla/issue/view/4267>
- Atjonen, P. 2007. Eettinen näkökulma arviointiin: miten ja kenen hyvää etsitään. *Didacta Varia* 12 (2), 31–41. URI: <http://hdl.handle.net/10224/4597>
- Bruzos, A., I. Erdocia. & K. Khan 2018. The path to naturalization in Spain: old ideologies, new language testing regimes and the problem of test use. *Language Policy*, 17, 419–441. <https://doi.org/10.1007/s10993-017-9452-4>
- Cook, V. J. 1999. Going beyond the native speaker in language teaching. *TESOL Quarterly*, 33 (2), 185–209. <https://doi.org/10.2307/3587717>
- Cook, V. J. 2003. *Effects of the second language on the first*. Clevedon: Multilingual Matters. <https://doi.org/10.21832/9781853596346>
- Cucchiari, C., H. Strik & L. Boves 2002. Quantitative assessment of second language learners' fluency: comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America*, 111 (6), 2862–2873. <https://doi.org/10.1121/1.1471894>
- Davies, A. 1997. Demands of being professional in language testing. *Language Testing*, 14 (3), 328–339. <https://doi.org/10.1177/026553229701400309>
- Davies, A. 2004. Introduction: Language testing and the golden rule. *Language Assessment Quarterly*, 1 (2–3), 97–107. <https://doi.org/10.1080/15434303.2004.9671778>
- Davies, A. 2013. Is the native speaker dead? *Histoire Épistémologie Langage* 35–2, 17–28.
- Derwing, T. & M. Munro 1997. Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition*, 19 (1), 1–16. <https://doi.org/10.1017/S0272263197001010>
- Deygers, B. 2019. Fairness and social justice in English language assessment. Teoksessa X. Gao (toim.) *Second handbook of English language teaching*. New York: Springer, 1–29. https://doi.org/10.1007/978-3-319-58542-0_30-1
- Douglas Fir Group 2016. A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal*, 100, Supplement Issue, 19–47. <https://doi.org/10.1111/modl.12301>
- Eckes, T. 2008. Rater types in writing performance assessments: a classification approach to rater variability. *Language Testing*, 25 (2), 155–185.
- Euroopan neuvosto 2001. *Common European Framework of Reference for Languages*. Strasbourg: Council of Europe.
- Euroopan neuvosto 2008. *Recommendation CM/Rec (2008) 7 to member states on the use of the Council of Europe's 'Common European Framework of Reference for Languages' (CEFR) and the promotion of plurilingualism*. Strasbourg: Council of Europe. http://search.coe.int/cm/Pages/result_details.aspx?ObjectId=09000016805d2fb1 (Luettu 13.4.2022)
- Euroopan neuvosto 2020a [2018]. *Common European Framework of Reference for Languages. Companion volume*. Strasbourg: Council of Europe.
- Euroopan neuvosto 2020b. *Linguistic integration of adult migrants: requirements and learning opportunities. Report on the Council of Europe and ALTE survey on language and knowledge of society policies for migrants*. <https://www.alte.org/LAMI-SIG> (Luettu 30.4.2022)

- Fayer, J. M. & E. Krasinski, 1987. Native and nonnative judgments of intelligibility and irritation. *Language Learning*, 37 (3), 313–326.
<https://doi.org/10.1111/j.1467-1770.1987.tb00573.x>
- Fulcher, G. 2015. *Re-examining language testing: a philosophical and social inquiry*. London & New York: Routledge. <https://doi.org/10.4324/9781315695518>
- Fulcher, G. & F. Davidson 2010. *Language testing and assessment: an advanced resource book*. London & New York: Routledge. 2. painos vuonna 2007 ilmestyneestä alkuperäisteoksesta.
- Goullier, F. 2007. *The Common European Framework for Languages (CEFR) and the development of language policies: challenges and responsibilities*. Policy Forum presentation. Strasbourg: Council of Europe. <https://rm.coe.int/impact-of-the-common-european-framework-of-reference-for-languages-and/16805c28c7> (Luettu 12.9.2022)
- Hamp-Lyons, L. 1997. Washback, impact and validity: ethical concerns. *Language Testing*, 14 (3), 295–303. <https://doi.org/10.1177/026553229701400306>
- Honko, M. 2018. Kieli- ja kielitaitokäsitykset tutkivan opettajan kenttäpäiväkirjamerkinnoissa. *Puhe ja Kieli*, 37(4), 215–238. <https://doi.org/10.23997/pk.63203>
- Huhta, A. 2014. Yleiset kielitutkinnot kansainvälisessä testauskentässä. Teoksessa T. Leblay, T. Lammervo & M. Tarnanen (toim.) *Yleiset kielitutkinnot 20 vuotta*. Helsinki: Opetushallitus, 159–168. <https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/yleiset-kielitutkinnot-20-vuotta> (Luettu 1.11.2022)
- Huhta, A. & R. Hildén 2016. Kielitutkinnot ja muu laajamittainen kielitaidon arviointi Suomessa. Teoksessa A. Huhta & R. Hildén (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa. AFInLA-e: soveltavan kielitieteen tutkimuksia 9*. Jyväskylä: Suomen soveltavan kielitieteen yhdistys, 3–26.
<http://journal.fi/afinla/article/view/60844> (10.10.2022)
- ILTA 2018[2000]: *ILTA code of ethics*. Marietta, GA: International Language Testing Association. <https://www.iltaonline.com/page/CodeofEthics> (Luettu 6.10.2023)
- Irvine, J. T. 1989. When talk isn't cheap: language and political economy. *American Ethnologist*, 16 (2), 248–267. <https://doi.org/10.1525/ae.1989.16.2.02a00040>
- Irvine, J. & S. Gal 2000. Language ideology and linguistic differentiation. Teoksessa P. V. Kroskrity (toim.) *Regimes of language. Ideologies, politics, and identities*. Santa Fe, New Mexico: School of American Research Press, 35–84.
- Isaacs, T. & H. Rose 2022. Redressing the balance in the native speaker debate: assessment standards, standard language, and exposing double standards. *TESOL Quarterly*, 56 (1), 401–412. <https://doi.org/10.1002/tesq.3041>
- de Jong, J. H.A.L. 2022. The action-oriented approach and language testing: a critical view. Teoksessa D. Little & N. Figueras (toim.) *Reflecting on the Common European Framework of Reference for Languages and its Companion Volume*. Luku 4. <https://doi.org/10.21832/9781800410206-009>
- Kang, O. & D. Rubin 2009. Reverse linguistic stereotyping: measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*, 28 (4), 441–456. <https://doi.org/10.1177/0261927X09341950>
- Kroskrity, P. V. 2000. Regimenting languages: language ideological perspectives. Teoksessa P. V. Kroskrity (toim.) *Regimes of language. Ideologies, politics, and identities*. Santa Fe, New Mexico: School of American Research Press, 1–34.
- Krumm, H. J. 2007. Profiles instead of levels: the CEFR and its (mis)abuses in the context of immigration. *The Modern Language Journal*, 91, 667–669.
https://doi.org/10.1111/j.1540-4781.2007.00627_6.x

- Kunnan, A. 2000. Fairness and justice for all. Teoksessa A. Kunnan (toim.) *Fairness and validation in language assessment: selected papers from the 19th Language Testing Research Colloquium, Orlando, Florida*. Studies in Language Testing 9. Cambridge: Cambridge University Press, 1–14.
- Laine, P, M. Maijala & K. Mäntylä 2023. Kestävän kehityksen periaatteet kielenoppimisen arvioinnissa. T. Mäkipää, R. Hilden & A. Huhta (toim.) 2023. *Kielenoppimista tukeva arviointi – Assessment for supporting language learning*. AFinLA-teema 15, 162–183.
- Laki Yleisistä kielitutkinnoista 12.11.2004/964.
<https://www.finlex.fi/fi/laki/ajantasa/2004/20040964> (Luettu 20.4.2022)
- Leblay, T., T. Lammervo & M. Tarnanen 2014 (toim.) *Yleiset kielitutkinnot 20 vuotta*. Helsinki: Opetushallitus. <https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/yleiset-kielitutkinnot-20-vuotta> (Luettu 1.11.2022)
- Leung, C. 2022. Action-oriented plurilingual mediation: a search for fluid foundations. Teoksessa D. Little & N. Figueras (toim.) *Reflecting on the Common European Framework of Reference for Languages and its Companion Volume*. Luku 6. <https://doi.org/10.21832/9781800410206-011>
- Leung, C. & J. Jenkins 2020. Mediating communication – ELF and flexible multilingualism perspectives on the Common European Framework of Reference for Languages. *Australian Journal of Applied Linguistics*, 3 (1), 26–41. <https://doi.org/10.29140/ajal.v3n1.285>
- Levy, M. & N. Figueras 2022. The action-oriented approach in the CEFR and the CEFR Companion Volume: a change of paradigm(s)? A case study from Spain. Teoksessa D. Little & N. Figueras (toim.) *Reflecting on the Common European Framework of Reference for Languages and its Companion Volume*. Luku 5. <https://doi.org/10.21832/9781800410206-010>
- Lindemann, S. 2006. What the other half gives: the interlocutor's role in non-native speaker performance. Teoksessa R. Hughes (toim.) *Spoken English, TESOL and applied linguistics: challenges for theory and practice*. London: Palgrave MacMillan, 23–49. https://doi.org/10.1057/9780230584587_2
- Little, D. & N. Figueras 2022. Introduction. Teoksessa D. Little & N. Figueras (toim.) *Reflecting on the Common European Framework of Reference for Languages and its Companion Volume*, xv–xviii. <https://doi.org/10.21832/9781800410206-003>
- Lynch, B. K. 1997. In search of the ethical test. *Language Testing*, 14 (3), 315–27. <https://doi.org/10.1177/026553229701400308>
- Lynch, B. K. 2001. Rethinking assessment from a critical perspective. *Language Testing*, 18 (4), 351–372. <https://doi.org/10.1177/026553220101800403>
- McNamara, T. 1998. Policy and social considerations in language assessment. *Annual Review of Applied Linguistics*, 18 (3), 304–319. <https://doi.org/10.1017/S0267190500003603>
- Martelin, T., T. Nieminen, A. Väänänen & M. Toivanen 2020. Työ ja työllistymisen esteet. Teoksessa H. Kuusio, A. Seppänen, S. Jokela, L. Somersalo & E. Lilja. (toim.) *Ulkomaalaistaustaisten terveys ja hyvinvointi Suomessa – FinMonik-tutkimus 2018–2019*. Raportti 1/2020- Helsinki: THL, 49–60. <https://urn.fi/URN:ISBN:978-952-343-034-1>
- McNamara, T. F. 1998. Policy and social considerations in language assessment. *Annual Review of Applied Linguistics*, 18 (3), 304–319. <https://doi.org/10.1017/S0267190500003603>
- Messick, S. 1989. Validity. Teoksessa R. L. Linn (toim.) *Educational measurement*. New York: Macmillan, 13–103.
- Migri 2023. Kielitaito. Suomen kansalaisuus. <https://migri.fi/kielitaito> (15.5.2023)
- Milani, T. M. 2008. Language testing and citizenship: a language ideological debate in Sweden. *Language in Society*, 37 (1), 27–59. <https://doi.org/10.1017/S0047404508080020>

- Mäntynen, A., M. Halonen, S. Pietikäinen & A. Solin 2012. Kieli-ideologioiden teoriaa ja käytäntöä. *Virittäjä*, 116 (3), 325–348.
- North, B. 2021. The CEFR Companion Volume – what's new and what might it imply for teaching/learning and for assessment? *CEFR Journal – Research and Practice*, 4 (December 2021). <https://doi.org/10.37546/JALTSIG.CEFR>
- North, B. 2022. The CEFR Companion Volume Project: What has been achieved. Teoksessa D. Little & N. Figueras (toim.) *Reflecting on the Common European Framework of Reference for Languages and its Companion Volume*. Luku 3. <https://doi.org/10.21832/9781800410206-007>
- OPH 2003. *Lukiokoulutuksen opetussuunnitelman perusteet 2003*. Helsinki: Opetushallitus.
- OPH 2004. *Perusopetuksen opetussuunnitelman perusteet 2004*. Helsinki: Opetushallitus.
- OPH 2011. *Yleisten kielitutkintojen perusteet*. Helsinki: Opetushallitus.
- OPH 2014. *Perusopetuksen opetussuunnitelman perusteet 2014*. Helsinki: Opetushallitus.
- OPH 2015. *Lukiokoulutuksen opetussuunnitelman perusteet 2015*. Helsinki: Opetushallitus.
- OPH 2018. *Kielten oppimisen, opettamisen ja arvioinnin yhteinen eurooppalainen viitekehys. Täydentävä osa ja lisäkuvaimet*. Helsinki: Opetushallitus.
- OPH 2019. *Kehittyvän kielitaidon tasojen kuvausasteikko (liite 2, lukion opetussuunnitelman perusteet 2019) sekä Kehittyvän kielitaidon tasojen kuvausasteikko, suomi ja ruotsi toisena kielenä ja kirjallisuus (Perusopetuksen opetussuunnitelman perusteiden tukiaineisto, liite)*. Helsinki: Opetushallitus.
- OPH 2022a. Opintopolku-portaali. Hakukelpoisuusvaatimukset. Helsinki: Opetushallitus. opintopolku.fi (10.11.2022)
- OPH 2022b. *Kotoutumiskoulutuksen opetussuunnitelman perusteet 2022*. Helsinki: Opetushallitus.
- Paikeday, T. M. 1985. *The native speaker is dead!: an informal discussion of a linguistic myth with Noam Chomsky and other linguists, philosophers, psychologists, and lexicographers*. Toronto: Paikeday Publishing.
- Pennycook, A. 1990. Towards a critical applied linguistics for the 1990s. *Issues in Applied Linguistics*, 1 (1), 9–29. <https://doi.org/10.5070/L411004991>
- Pennycook, A. 1999. Introduction: critical approaches to TESOL. *TESOL Quarterly*, 33 (3), 329–348. <https://doi.org/10.2307/3587668>
- Piccardo, E. & B. North 2019. *The action-oriented approach: a dynamic vision of language education*. Bristol: Multilingual Matters. <https://doi.org/10.21832/9781788924351>
- Riggenbach, H. 1991. Toward and understanding of fluency: a microanalysis of nonnative speaker conversations. *Discourse Processes*, 14 (4), 423–441. <https://doi.org/10.1080/01638539109544795>
- Ruohotie-Lyhty, M. & P. Kaikkonen 2009. The difficulty of change – the impact of personal school experience and teacher education on the work of beginning language teachers. *Scandinavian Journal of Educational Research*, 53 (3), 295–309. <https://doi.org/10.1080/00313830902917378>
- Salo, O.-P., K. Kajander, R. Kantelinen & K. Wallinheimo 2020. Mediaatio: mitä se on? *Tempus*, 55 (6), 12–13. <http://urn.fi/URN:NBN:fi:jyu-202012107042>
- Schieffelin, B., K. A. Woolard & P. V. Kroskity 1998 (toim.), *Language ideologies. Practice and theory*. New York: Oxford University Press.
- Shohamy, E. 1993. *The power of tests: the impact of language tests on teaching and learning*. Washington, DCC: The National Foreign Language Center at Johns Hopkins University.
- Shohamy, E. 2001. *The power of tests: a critical perspective on the uses of language tests*. Harlow: Pearson Education.

- Shohamy, E. 2007. Tests as power tools: looking back, looking forward. Teoksessa J. Fox, M. Wesche, D. Bayliss, L. Cheng, C.E. Turner & C. Doe (toim.) *Language testing reconsidered*. Ottawa: University of Ottawa Press, 141–152.
<https://doi.org/10.2307/j.ctt1ckpccf.14>
- Shohamy, E., S. Donitsa-Schmidt. & L. Ferman 1996. Test impact revisited: washback-effect over time. *Language Testing*, 13 (3), 298–317. <https://doi.org/10.1177/026553229601300305>
- Silverstein, M. 1979. Language structure and linguistic ideology. Teoksessa P. R. Cline, W. F. Hanks & C. L. Hofbauer (toim.) *The elements. A parsession on linguistic units and levels*, 193–247. Chicago: Chicago Linguistic Society.
- Silverstein, M. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, 23 (3–4), 193–229. [https://doi.org/10.1016/S0271-5309\(03\)00013-2](https://doi.org/10.1016/S0271-5309(03)00013-2)
- Spolsky, B 1981. Some ethical questions about language testing. Teoksessa C. Klein-Braley & D. K. Stevenson (toim.) *Practice and problems in language testing*. Frankfurt: Peter Lang, 5–21.
- Spolsky, B. 1997. The ethics of gatekeeping tests: what have we learned in a hundred years? *Language Testing*, 14 (3), 242–247. <https://doi.org/10.1177/026553229701400302>
- Stevenson, D. K. 1981. Language testing and academic accountability: on redefining the role of language testing in language teaching. *International Review of Applied Linguistics*, 19 (1), 15–130. <https://doi.org/10.1515/iral.1981.19.1-4.15>
- Stråth, B. 2013. Ideology and conceptual history. Teoksessa M. Freeden & M. Stears (toim.) *The Oxford Handbook of Political Ideologies*. Oxford: Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199585977.013.0013>
- Suomen perustuslaki 11.6.1999/731. <https://www.finlex.fi/fi/laki/ajantasa/1999/19990731> (Luettu 17.9.2022)
- Tarnanen, M. 2014. Arvioija taidon arvottajana. Teoksessa T. Leblay, T. Lammervo & M. Tarnanen (toim.) *Yleiset kielitutkinnot 20 vuotta*. Helsinki: Opetushallitus, 115–124.
<https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/yleiset-kielitutkinnot-20-vuotta> (Luettu 1.11.2022)
- Tarnanen, M. & S. Pöyhönen 2011. Maahanmuuttajien suomen kielen taidon riittävyys ja työllistymisen mahdollisuudet. *Puhe ja kieli*, 31 (4), 139–152.
- Tossavainen, H. 2016. Kielitestiä eettisyydestä ja oikeudenmukaisuudesta. Teoksessa A. Huhta & R. Hildén (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa. AFInLA-e: soveltavan kielitieteen tutkimuksia 9*. Jyväskylä: Suomen soveltavan kielitieteen yhdistys, 27–43. <https://journal.fi/afinla/issue/view/4267> (Luettu 3.11.2022)
- Tossavainen, H. & S. Ahola 2014. Yleisten kielitutkintojen arvioijakoulutus. Teoksessa T. Leblay, T. Lammervo & M. Tarnanen (toim.) *Yleiset kielitutkinnot 20 vuotta*. Helsinki: Opetushallitus, 125–134.
<https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/yleiset-kielitutkinnot-20-vuotta> (Luettu 1.11.2022)
- TRAFICOM 2021. *Ilmailun kielitaitovaatimukset ja kielitaidon arviointi*. Luettavissa: https://www.traficom.fi/sites/default/files/media/regulation/PEL%20M2-92_2021_lausunolle.pdf (Luettu 25.4.2022)
- Tsagari, D. & L. Cheng 2017. Washback, impact, and consequences revisited. Teoksessa E. Shohamy, I. Or & S. May (toim.) *Language testing and assessment. Encyclopedia of language and education*. Springer, 359–372. https://doi.org/10.1007/978-3-319-02261-1_24
- Tupas, R. 2022. The coloniality of native speakerism. *Asian Englishes*, 24 (2), 147–159.
<https://doi.org/10.1080/13488678.2022.2056797>
- Valvira 2021. *Kielitaito*. <https://www.valvira.fi/terveydenhuolto/ammattioikeudet/kielitaito> (Luettu 25.4.2022)

- Wall, D. 2000. The impact of high-stakes testing on teaching and learning: can this be predicted or controlled. *System*, 28 (4), 499–509. [https://doi.org/10.1016/S0346-251X\(00\)00035-X](https://doi.org/10.1016/S0346-251X(00)00035-X)
- Weir, C. J. 2005. *Language Testing and validation: an evidence-based approach*. New York: Palgrave Macmillan.
- Xi, X. 2010. How do we go about investigating test fairness? *Language Testing*, 27 (2), 147–170. <https://doi.org/10.1177/0265532209349465>