

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Huhta, Ari; Leontjev, Dmitri; Honko, Mari; Hildén, Raili

**Title:** Kielitaidon arviointi koulussa

**Year:** 2023

**Version:** Published version

**Copyright:** © kirjoittajat ja Suomen soveltavan kielitieteen yhdistys AFinLA

**Rights:** In Copyright

**Rights url:** <http://rightsstatements.org/page/InC/1.0/?language=en>

**Please cite the original version:**

Huhta, A., Leontjev, D., Honko, M., & Hildén, R. (2023). Kielitaidon arviointi koulussa. In A. Pitkänen-Huhta, K. Mård-Miettinen, & T. Nikula (Eds.), *Kielikoulutus mukana muutoksessa* (pp. 242-265). Suomen soveltavan kielitieteen yhdistys ry. AFinLA-teema, 16.  
<https://doi.org/10.30660/afinla.126733>

*Pitkänen-Huhta, A., K. Mård-Miettinen & T. Nikula (toim.) 2023. Kielikoulutus mukana muutoksessa – Language education engaging in change. AFinLA-teema / n:o 16, 242–265.*

**Ari Huhta, Dmitri Leontjev & Mari Honko**

Jyväskylän yliopisto

**Raili Hildén**

Helsingin yliopisto

## Kielitaidon arviointi koulussa

### Nostot

- Summatiivinen ja formatiivinen arviointi ovat keskeistä kielenopetuksessa, mutta niiden yhdistäminen vaatii huolellista suunnittelua
- Opettajien ja tutkijoiden yhteistyö sekä tutkimuspohjainen opettajankoulutus ovat avaimia opettajien arviointiosaamisen kehittämiseen
- Teknologia tarjoaa arvioinnille uusia mahdollisuuksia, mutta se voi myös rajoittaa sitä, mitä tietoja ja taitoja arvioidaan
- Yhteiskunnan ja koulun monikielistymisen myötä tarvitaan joustavia, erilaiset oppijat ja oppimisen vaiheet huomioivia arviointimenetelmiä



VERTAISARVIOITU  
KOLLEGIALT GRANSKAD  
PEER-REVIEWED  
[www.tsv.fi/tunnus](http://www.tsv.fi/tunnus)

## Abstract

While assessment is related to all levels of the Douglas Fir Group's (2016 hereinafter, DFG) framework, in this paper, we focus on language assessment in schools, that is, institutional meso level. However, we will also discuss micro (individual) and macro (societal) levels of the framework. We will discuss school classroom assessment focusing on four themes: (1) the understanding of quality of assessment in formal education, (2) assessment literacy required from teachers, (3) technology in assessment, and (4) challenges and opportunities that multilingualism in society and schools brings for assessment. We will also discuss future developments in Finnish classroom assessment research and practice, or indeed, a combination of the two.

**Keywords:** language assessment, assessment literacy, technology in assessment, multilingualism in assessment

**Asiasanat:** kielitaidon arviointi, arviointiosaaminen, teknologia arvioinnissa, monikielisyys arvioinnissa

## 1 Johdanto

Kielitaidon arviointi liittyy kaikkiin esimerkiksi Douglas Fir Groupin (DFG, 2016) viitekehysten määrittelemiin tasoihin eli mikro-, meso- ja makrotasoihin. Tämä pätee etenkin formaalissa koulutuksessa tehtävään arviointiin. Arviointi on lähtökohtaisesti yksilöiden arviointia eli mikrotason toimintaa: oppilaiden oppimista tuetaan arvioinnista saatavan tiedon avulla ja heille annetaan arvosanoja sen perusteella, miten hyvin he ovat omaksuneet opetetut asiat. Oppilaitoksissa tehtävä arviointi on samalla vahvasti yhteisöllistä toimintaa (mesotaso): opettajilla on tyypillisesti päävastuu oppilaiden arvioinnista, mutta koulu ja myös kunta paikallisen opetussuunnitelman kautta voivat vaikuttaa oppilaitostason arviointiin. Tiedetään myös, että oppilaitosyhteisössä opettajien arviointilinjat muokkautuvat samansuuntaisiksi, mikä osaltaan selittää koulujen välisiä eroja perusasteen päättöarvioinneissa (Hildén & Rautopuro 2014; Hildén ym 2016). Arviointi on myös yhteiskunnan eli makrotason toimintaa, koska kansalliset koulutusviranomaiset ja poliittiset päättäjät ohjaavat kouluissa tapahtuvaa arviointia erilaisin tavoin. Suomen peruskouluissa ja lukioissa tämä tapahtuu valtakunnallisten opetussuunnitelmien eli opetussuunnitelmien perusteiden kautta (ks. myös Honko & Halonen tässä teemanumerossa). Kielikoulutus on edennyt pitkän kehityskaaren strukturalistisen kielikäsitteilyksen, behavioristisen oppimiskäsityksen ja kielioppi-käännös-menetelmän ajoista nykypäivään, jolloin kieli nähdään viestinnän, vuorovaikutuksen ja kohtaamisen rakentajana. Oppijoiden moninaisuuden tunnustavat oppimiskäsitykset ovat rikastaneet kielen opetuksen ja arvioinnin tapoja, mutta samalla maailman ja yhteiskunnan nopeat kulttuuriset ja sosiaaliset muutokset nostavat esiin uusia haasteita, joihin on vastattava myös kielenopetuksessa ja sen tulosten arvioinnissa (ks. tämän teemanumeron johdantoartikkeli Pitkänen-Huhta ym.).

Tässä artikkelissa pääpaino on kielitaidon arvioinnilla, jota tehdään kouluissa eli mesotasolla, mutta aiheen käsittely kietoutuu väistämättä myös yksilöiden ja yhteiskunnan tason kysymyksiin. Lähestymme arviointia neljän pääteeman kautta. Ensimmäinen käsittelee yleisimpiä arvioinnin tarkoituksia (ja arvioinnin perusteella tehtäviä päätöksiä) koulussa eli summatiivista ja formatiivista arviointia sekä niiden laatuun liittyviä kysymyksiä. Muut kolme teemaa keskittyvät kielikoulutuksessa tärkeiksi nousseisiin kysymyksiin: millaista arviointiosaamista opettajilta nykyään vaaditaan, miten teknologiaa hyödynnetään arvioinnissa ja mitä haasteita ja mahdollisuuksia yhteiskunnan ja koulujen monikielistyminen tarjoaa arvioinnille. Keskitymme mesotasolla eli koulussa tapahtuvaan arviointiin, koska valtaosa kielitaidon arvioinnista tapahtuu perusasteen ja toisen asteen oppilaitoksissa. Yksilötason arviointikin eli opettajien ja oppilaiden toiminta tapahtuu useimmiten oppilaitoskontekstissa, jolloin arviointiin vaikuttavat myös makrotason opetussuunnitelmien perusteet.

## 2 Summatiivinen ja formatiivinen arviointi koulussa

Arviointi on kiinteä osa sekä peruskoulun että lukion toimintaa, ja kuten edellä todettiin, sitä ohjataan opetuksen tavoin lainsäädännöllä sekä valtakunnallisilla ja paikallisilla opetussuunnitelmissa. Perusopetuslaissa (628/1998, 22 § 1 mom) todetaan oppilaan arvioinnista seuraavaa:

Oppilaan arvioinnilla pyritään ohjaamaan ja kannustamaan opiskelua sekä kehittämään oppilaan edellytyksiä itsearviointiin. Oppilaan oppimista, työskentelyä ja käyttäytymistä tulee arvioida monipuolisesti.

Lukiolaissa (714/2018, 37 § 1 mom) puolestaan todetaan:

Opiskelijan työskentelyä, oppimista ja osaamisen kehittymistä tulee arvioida monipuolisesti. Opiskelijan arvioinnilla pyritään ohjaamaan ja kannustamaan opiskelua sekä kehittämään opiskelijan edellytyksiä itsearviointiin. Opiskelijalla on oikeus saada tieto arvioinnissa käytettävistä arviointiperusteista ja niiden soveltamisesta häneen. Opiskelijalle on lukio-opintojen aikana annettava mahdollisuus itsearviointiin.

Arviointi on nykykäsityksen mukaan pohjimmiltaan päätöksentekoa. Päätöksillä on erilaisia seurauksia oppilaalle, hänen lähiyhteisölleen ja lopulta koko yhteiskunnalle. Päätökset taas ohjaavat arviointitapoja. Arviointipäätökset voidaan koulukontekstissa jakaa kahteen lajiin, joita kutsutaan **formatiiviseksi** ja **summatiiviseksi** arvioinniksi (arviointitiedon käytöstä laajemmin ks. esim. Nguyen 2021).

## 2.1 Miten summatiivinen ja formatiivinen arviointi eroavat toisistaan

Summatiiviset arviointipäätökset liittyvät nimensä mukaisesti tyypillisesti opetusjakson tai kokonaisen kouluasteen päättöarviointiin ja arvosanojen antoon. Päättöarvosanat ovat opiskelijalle tärkeitä, koska ne saattavat ratkaista tulevan koulutuspolun tai työpaikan. On tärkeää tiedostaa, että oppilaan arviointi yleissivistävässä koulutuksessa tapahtuu virkavastuulla ja vastuu arvioinnista on viime kädessä aina opettajalla. Opiskelija voi yrittää muuttaa opintojaksojen arvosanoja tai päättöarvosanoja antamalla uusia summatiivisia näyttöjä osaamisestaan, mutta nämä mahdollisuudet ovat usein ajallisesti ja muodollisesti tarkoin säänneltyjä. Summatiivisen arvioinnin tavoitetta vastaa nykyisissä opetussuunnitelman perusteissa (OPH 2014b, 2019) oppimisen arviointi eli sen toteaminen, missä määrin oppija on saavuttanut eri tavoitteet. Summatiivisen arvioinnin kohteiksi sopivat luonteeltaan suhteellisen hitaasti muuttuvat tiedot ja taidot, jotka ovat määriteltävissä pääosin etukäteen.

Formatiivisella arvioinnilla kerätään tietoa oppilaiden osaamisesta opetuksen kuluessa ja se voi kohdistua hyvinkin pieniin kokonaisuuksiin (ks. esim. Wiliam & Leahy 2015). Formatiivisen tiedon avulla pyritään konkreettisesti tukemaan oppimista ja opettamista. Formatiiviseen arviointiin liittyvät päätökset ovat siinä mielessä luonteeltaan summatiivista arviointia kevyempiä, että niiden perustana olevaa näyttöä voidaan helpommin muokata ja arviointi myös voidaan uusia lyhyemmällä aikavälillä kuin summatiiviset näytöt tai se voi olla jatkuvaa. Esimerkkinä formatiivisista päätöksistä mainittakoon tietyn oppiaineen kertauksen tai tukiopetuksen tarpeen toteaminen tai päätös, että opetuksessa voidaan siirtyä eteenpäin ja mahdollisesti käsitellä asiaa syvällisemmin. Formatiivinen arviointi siis auttaa opettajaa suuntaamaan toimintaansa oppilaiden todettujen tarpeiden mukaisesti jo opetuksen aikana. Opetussuunnitelmien perusteissa formatiivisen arvioinnin päätavoitteeksi todetaan oppimisen edistäminen (OPH 2014b). Arviointikohteena on ensi sijassa tietojen ja taitojen opiskelun prosessi ja sen vaiheet. Lisäksi formatiivinen arviointi tavoittaa summatiivista paremmin esimerkiksi laaja-alaiset oppimiskokonaisuudet (esim. monilukutaito ja oppimaan oppiminen) tai oppijan henkilökohtaiset tavoitteet, vahvuudet ja haasteet, koska niiden summatiivinen yhdenmukainen arviointi on vaikeaa (laaja-alaisen oppimisen arvioinnista ks. tarkemmin Korhonen & Mattila 2020). Nykyisissä opetussuunnitelman perusteissa tähdennetään, ettei oppilaan henkilökohtaisia ominaisuuksia, kuten temperamenttia, tule lainkaan arvioida (OPH 2014b). Rajanveto oppilaan henkilökohtaisten ominaisuuksien ja osaamisen välillä on kuitenkin vaikeaa varsinkin kieliaineissa, joissa esimerkiksi suullisen kielitaidon harjoittelu edellyttää oppijalta rohkeutta puhua ääneen.

Kun arvioinnin tarkoitus on todettu, valitaan sen menetelmät. Summatiivisen arvioinnin näytöt on perinteisesti kerätty paperikokein, mutta nykyään myös digitaalisesti, mikä on mahdollistanut monimuotoiset suoritukset ja näin myös monilukutaidon yhä kattavamman arvioinnin. Summatiiviselle arvioinnille toisaalta on

edelleen tyypillistä näytön ajoittuminen tiettyihin ennalta määriteltyihin ajankohtiin (esim. kurssin loppuun) sekä näytön muotojen yhdenmukaisuus (esim. kaikille sama koe). Summatiivinen arviointi ohjautuu opintojakson tavoitteista, joista toiset ovat helpompia arvioida kuin toiset (esim. kielitaidon alueet vs. kielenopiskelutaidot). Perusopetuksen päättöarvioinnin ohjenuoraksi on laadittu arviointikriteeristö, joka astui voimaan 2021. Siinä määritellään sanallisesti, millaista osaamista oppijalta edellytetään kullakin tavoitealueella arvosanoihin 5, 7, 8 ja 9 (OPH 2020; taitotasokuvausten käytöstä kielitaidon arvioinnissa ks. Huhta 2020).

Formatiivisen arvioinnin muoto- ja välinevalikoima on yleensä laaja ja tilanteittain joustava. Tavallisimpia muotoja ovat opiskeluun liittyvät tehtävät ja niistä annettava palaute opettajalta ja vertaisilta sekä itsearviointi eli oman suoritusprosessin ja sen tulosten reflektointi. Tarjolla on myös kansainvälisesti kehitettyjä itsearviointivälineitä kuten eurooppalainen kielisalkku tai kieliprofiili (OPH 2021). Nämä ohjaavat oppijaa tunnistamaan opittavan taidon kehitysvaiheet ja pohtimaan omaa identiteettiään laajemminkin oppijana ja kehittyvänä ihmisenä.

## 2.2 Laatu summatiivisessa ja formatiivisessa arvioinnissa

Summatiivisten ja formatiivisten arviointipäätösten vakavuus ja tarkoitus ohjaavat niiltä edellytettäviä laatupiirteitä eli validiutta. Nykykäsityksen mukaan validius kattaa koko arviointiprosessin tarkoituksen määrittämisestä suoritusnäyttöjen pisteytykseen ja arviointituloksen seurauksiin saakka. Arviointiprosessin laatutekijöitä on hahmoteltu monin tavoin, ja kielitaidon arvioinnissa tunnetuimmat ovat Bachmanin ja Palmerin (1996, 2010), Weirin (2005), Chapellen ym. (2008) sekä Kunnanin (2018) laatimat mallit. Kaikki pohjautuvat Messickin (1989) validiudesta esittämään kokonaismalliin, joka korostaa arvioinnin kokonaisuasetelman sekä seurausten huomioimisen merkitystä validiuden arvioinnissa enemmän kuin aiempi, arvioinnissa käytettyjen mittareiden tilastollisille ominaisuuksille rakentuva psykometrinen näkemys. Messickin näkemys sulautti yhteen varhaisemmat mallit, joissa eriteltiin useita validiuden osa-alueita (esim. sisältö-, konstrukti- ja ennustevalidius) (ks. myös Honko & Halonen tässä teemanumerossa). Useimmat nykyiset mallit tarkastelevat arviointia argumentaatioteorian näkökulmasta (ks. lähemmin Kane 2006), jossa arvioinnin (koe, tutkinto, arviointijärjestelmä) laatua analysoidaan keräämällä perusteluja sen pätevyuden puolesta ja sitä vastaan (Bachman & Palmer 2010). Vaikka mallien välillä on eroja, kaikkien perimmäinen tavoite on toteuttaa reilua, oikeudenmukaista ja eettisesti kestävää arviointia (esim. Kunnan 2018).

Eri validiusmalleissa korostuu yleensä kolme osaa: arvioinnin kohde, arvioinnin autenttisuus ja arvioinnin johdonmukaisuus (reliability). Arvioinnin validiutta/laatua punnittaessa on lisäksi huomioitava arviointiprosessin kaikki vaiheet arviointitarpeen määrittelystä arviointituloksen muodostamiseen. Käytännössä validiuden eri osatekijät kuitenkin kietoutuvat toisiinsa. Validiuden määrittelyä haastaa myös se, että kielitaidon

määrittely itsessään on haastavaa. Lisäksi käsitys kielitaidosta samoin kuin arvioinnin tarpeet ja välineet muuttuvat jatkuvasti esimerkiksi yhteiskunnan monikielistymisen ja digitalisaation myötä. Monet asiat esimerkiksi monikielisyyden arvioinnissa ovat vielä ratkaisematta, ja digitaalisen ja varsinkin automaattisen arvioinnin aiheuttama arvioinnin kohteen kapeutumisen huolettavat monia.

Arvioinnin autenttisuuden tarpeesta ollaan varsin yksimielisiä. Kielitaidon arviointiin käytettävien tehtävien tulisi olla mahdollisimman samankaltaisia kuin tilanteet, joissa kieltä on harjoiteltu ja joissa sitä käytetään. Arvioinnin autenttisuus mukailee kielenkäyttötilanteiden vaihtelua arjessa ja on siksi erityisesti jokapäiväisiin tilanteisiin limittyvän formatiivisen arvioinnin vahvuus. Koska kielenkäyttöympäristöt vaihtelevat ja muuttuvat, opettajan on hyödyllistä reflektoida omia kielitaitokäsityksiään ja tarpeen vaatiessa myös uudistaa omia arviointikäytänteitään.

Reliaabelius viittaa arvioinnin johdonmukaisuuteen, vakauteen, virheettömyyteen ja toistettavuuteen. Reliaabeli arviointi pyrkii myös riippumattomuuteen suorituskontekstista tai arvioijasta. Kaikki mainitut piirteet ovat tyyppillisiä summatiivisen arvioinnin ihanteita, joita etenkin laajoissa ja suorittajan tulevaisuuden kannalta tärkeissä kokeissa on tärkeää tavoitella (Bachman & Palmer 1996). Arvioinnin johdonmukaisuus korostuu laatupiirteenä myös silloin, kun arvioinnilla on portinvartijan rooli esimerkiksi opiskelutai työpaikkojen valinnoissa tai kansalaisuuden myöntämisessä (ks. tarkemmin Nguyen 2021; testiarvioinnista myös Honko & Halonen tässä teemanumerossa). Käytännössä kaikenlainen arviointi on kuitenkin tasapainottelua yhtäältä autenttisuuden ja toisaalta muiden laadukkaan arvioinnin / validiteettia lisäävien piirteiden kuten vakauden ja toistettavuuden kontrolloinnin välillä.

Autenttisuus ja arviointitapojen joustavuus ovat puolestaan tärkeitä (ja yleensä helpommin toteutettavia) laatupiirteitä oppimista tukevassa arvioinnissa ja palautteessa, jotta juuri tietyn oppijan osaamisesta ja oppimisprosessista olisi mahdollista saada tarpeeksi hyvä ja monipuolinen kuva. Vaikka formatiivisessakin arvioinnissa on hyödyllistä kiinnittää huomiota arvioinnin johdonmukaisuuteen, arviointitarpeet voivat eri tilanteissa ja eri oppijoidenkin kohdalla olla erilaisia. Lisäksi yksittäisestä arvioinnista mahdollisesti saatu virheellinen kuva oppilaan osaamisesta todennäköisesti korjautuu seuraavilla arviointikerroilla.

### 2.3 Mitä yhteistä summatiivisella ja formatiivisella arvioinnilla on

Summatiivisen ja formatiivisen arvioinnin välinen raja kuvataan toisinaan hyvin jyrkäksi, mikä ei kuitenkaan vastaa suomalaisen koulukontekstin todellisuutta. Kansainvälisissä testivetoisissa koulukulttuureissa koulun ulkopuolelta ohjattujen tärkeiden arviointien yhteys opetukseen on joskus tarkoituksellisesti katkaistu arvioinnin riippumattomuuden takaamiseksi. Kuitenkin nämä arvioinnit heijastuvat opetukseen takaistusvaikutuksena, jolloin opetus vinoutuu pahimmillaan valmennukseksi ulkoiseen tutkintoon (ks. esim. Cheng & Sultana 2021). Suomessa isoja tutkintoja on hyvin vähän, oikeastaan vain

ylioppilastutkinto lukio-opintojen lopuksi. Varsinkin 2010-luvun opetussuunnitelmien perusteet nostavat formatiivisen arvioinnin tärkeimmäksi arvioinnin muodoksi. Myös eurooppalainen viitekehys, etenkin sen uusi versio Companion Volume (Euroopan neuvosto 2020), tarkastelee kielitaidon arviointia opetuksen ja oppimisen näkökulmasta ja kehottaa nykyisten toimintatapojen kriittiseen tarkasteluun ja pohtimaan vaihtoehtoja standardoiduille testeille (ks. tarkemmin North 2021). Summatiivisessa arvioinnissa tämä tarkoittaa vaihtoehtoisten tai täydentävien arviointitapojen käyttöönottoa, ja esimerkiksi formatiiviselle arvioinnille tyypillisten menetelmien kuten kielisalkun ja jatkuvaan näyttöön perustuvan arvioinnin hyödyntämistä kurssien loppukokeiden rinnalla.

Myös Suomessa on nähty jonkin verran jännitteitä summatiivisen ja formatiivisen arvioinnin välillä, koska summatiivinen arviointi voidaan kokea tärkeämmäksi (esim. Mäkipää 2021; Leontjev 2022). Onkin ensisijaisen tärkeää ymmärtää validiuden erilainen luonne näissä erityyppisiin päätöksiin johtavissa arvioinneissa. Molemmissa arvioinnin kohteena tulisi olla esimerkiksi opetussuunnitelmissa määritellyt tavoitteet (esim. OPH 2014b). Tämän ohella hyvässä summatiivisessa arvioinnissa korostuvat reliaabelius ja yhdenmukaisuus, hyvässä formatiivisessa arvioinnissa puolestaan autenttisuus, joustavuus ja yksilöllisyys. On kuitenkin hyvä muistaa, että käytännössä kaikki arviointimenetelmät sopivat kumpaankin arviointiin, vaikka perinteisesti joitakin niistä käytetäänkin enemmän joko summatiivisessa tai formatiivisessa arvioinnissa (ks. Wiliam & Leahy 2015). Esimerkiksi summatiivisten kokeiden tuloksia tulisi nykyistä enemmän hyödyntää oppimisen tukemisessa ja opetuksen suuntaamisessa – esimerkiksi yli opintojaksojen lukiossa. Toisaalta arvosanoissa on mahdollista ja perusteltua huomioida muitakin näyttöjä kuin loppukokeet.

Tulevaisuudessa lupaavin visio on summatiivisen ja formatiivisen arvioinnin tarkoituksenmukainen yhdistäminen, harjoitetaanhan niitä kumpaakin viime kädessä oppijan tietojen, taitojen ja sivistyksen edistämiseksi. Kahden arviointiperiaatteen vastakohtaisuuden ja arvolatausten ruokkimisen sijasta on hedelmällisempää etsiä keinoja siihen, että arvioinnista muodostuu oppijalle ja opettajille johdonmukainen kasvua tukeva kokonaisuus (Pellegrino 2018). Tarkastelemme erilaisten arviointien yhdistämistä lähemmin seuraavassa luvussa, jossa käsittelemme kielenopettajien arviointiosaamista.

### 3 Kielenopettajien arviointiosaaminen

Edellä kuvattiin, miten Suomen lainsäädännössä ja opetussuunnitelmien perusteissa määrätään, että arvioinnin on tuettava oppimista ja oltava monipuolista: arviointiin täytyy muun muassa sisältyä oppilaiden itsearviointia. Tällainen näkemys arvioinnista vastaa nykyisiä kansainvälisiä arviointikäsitteitä oppimista tukevasta arvioinnista (*assessment for learning*; esim. Wiliam 2017) ja kriittisestä arvioinnista (Shohamy 2001),



joissa oppilaat ovat oppimis-, opetus- ja arviointiprosessin keskiössä ja joissa arvioinnin päätarkoitus on tukea oppimista. Opetussuunnitelmat kehottavat opettajia myös pohtimaan, miten sekä summatiivinen että formatiivinen arviointi yhdessä voisivat auttaa oppimista. Tämä on tärkeä ja mielenkiintoinen kehoitus, sillä kansainvälinen tutkimus on usein osoittanut summatiivisen ja formatiivisen arvioinnin saumattoman yhdistämisen haastavaksi (Davison & Leung 2009; Tsagari & Vogt 2017). Jotkut tutkijat (esim. Wiliam & Leahy 2015) näkevätkin summatiivisen arvioinnin käyttämisen oppimisen tukemisessa vaikeana. Syitä tähän ovat mm. se, että summatiivisesta arvioinnista saatava palaute ei ole yhtä välitöntä ja yksityiskohtaista kuin formatiivisessa arvioinnissa ja se, etteivät summatiivisen arvioinnin tulokset – tyypillisesti arvosanat – tarjoa oppilaalle täsmällistä tietoa omasta osaamisesta tai konkreettisia neuvoja siihen, mitä hänen kannattaisi tehdä seuraavaksi.

Koska näiden kahden arvioinnin tarkoituksen yhdistäminen ei ole aina helppoa, on hyödyllistä kuvata lyhyesti, miten se voisi tapahtua. Pellegrino (2018) sekä Poehner ja Inbar-Lourie (2020) esittävät, että opettaja voi yhdistää nämä arvioinnit esimerkiksi käsittelemällä arviointiperusteita oppilaiden kanssa, käyttämällä eteenpäin suuntaavaa palautetta (oppilaiden suorituskyvyn kehittämiseen suunnattu palaute) ja keskustelemalla palautteesta oppilaiden kanssa sekä kielisalkun avulla. Arviointitapojen yhdistäminen voi tapahtua myös opettajien ja tutkijoiden yhteistyönä (Poehner & Inbar-Lourie 2020). Esimerkiksi Chan ja Davison (2020) tutkivat, kuinka opettajat ja tutkijat yhdessä ratkaisivat haasteita, joita Hongkongissa toteutettu, oppimista tukevaa arviointia korostava opetussuunnitelmaudistus toi luokkahuoneeseen, jossa sitä ennen painotettiin summatiivista arviointia. Yhteistyössä tutkijoiden kanssa opettajat syventyivät oppimista tukevan arvioinnin periaatteiden ymmärtämiseen sekä kehittivät käytäntöjään. Opettajat alkoivat muun muassa kysyä oppilailta, missä asioissa nämä kokivat vaikeuksia, mitkä asiat taas sujuivat hyvin ja mitä oppilaat voisivat itse tehdä oppiakseen. Yhteistyö opettajien kanssa auttoi puolestaan tutkijoita perustelemaan teoreettisia näkemyksiään sen perusteella, mikä toimi luokkahuoneessa.

Opetussuunnitelmissa ilmaistu tavoite kaiken arvioinnin käyttämisestä oppimisen tukemiseen asettaakin opettajien arviointiosaamiselle varsin suuria vaatimuksia, etenkin kun opettajilla on hyvin suuri vapaus päättää, miten he toteuttavat opetuksen ja arvioinnin käytännössä. Toisaalta oppimista tukevan arvioinnin periaatteiden omaksuminen voi tehdä arvioinnista mielekäästä ja vaikuttavaa myös opettajan työn kannalta. Arviointiosaamiselle ei ole yhtä, kaikkien käyttämää määritelmää, vaan määritelmien painopisteet ja yksityiskohtaisuus vaihtelevat. Arviointiosaamisen keskeisten määritelmien (Fulcher 2012; Inbar-Lourie 2017; Xu & Brown 2016 sekä Inbar-Lourien artikkelissaan viittaamat kirjoittajat) yleiskuva on tiivistetty kuvion 1 sanapilveen.



KUVIO 1. Sanapilvi arviointiosaamisen määritelmistä Fulcherin (2012), Inbar-Lourien (2017), Taylorin (2013) sekä Xun ja Brownin (2016) artikkeleissa.

Synteesimme kuviossa 1 osoittaa, että arviointiosaamisen määritelmässä korostuu arviointitieto (*knowledge, principle*), mutta myös käytännöt (*practice*) eli opettaminen (*learning, development, acquisition*), osaaminen (*skill, ability*) sekä käsitykset (*concept, interpretation*) ja opettajan uskomukset ja asenteet (*belief, attitude*) ovat tärkeitä. Määritelmässä on kuitenkin myös eroja. Fulcherin määritelmässä (2012) opettajan arviointiosaamiseen kuuluvat ensinnäkin tiedot ja taidot, joita vaaditaan erilaisten kokeiden laatimisessa, kehittämisessä ja arvioinnissa, mutta myös tietoisuus arvioinnin periaatteista, esimerkiksi eettisistä kysymyksistä. Fulcherin mukaan arviointiosaamiseen kuuluu lisäksi taito soveltaa osaamista kulloiseenkin sosiaaliseen kontekstiin, jotta opettaja ymmärtää arvioinnin mikro-, meso- ja makrotason vaikutukset. Myös Inbar-Lourien (2017) määritelmässä arviointiosaamisesta korostuu kontekstin, periaatteiden ja käytänteiden hallinta. Xu ja Brown (2016) puolestaan painottavat määritelmässään, että opettajien arviointiosaamisen kehittymisen tulisi johtaa heidän arviointikäytänteidensä muuttumiseen, siis esimerkiksi oppimista tukevan arvioinnin periaatteiden siirtymistä tiedosta tai arvostuksista hyviksi käytänteiksi.

Taylor (2013) on pyrkinyt analysoimaan erilaisilta kielitaidon arvioinnin toimijoilta vaadittavaa arviointiosaamista. Hän tarkasteli kahdeksaa arviointiosaamisen ulottuvuutta: teoreettista osaamista, teknisiä taitoja, periaatteiden ja käsitteiden tuntemusta, kielipedagogiikkaa, sosiokulttuurisia arvoja, paikallisten käytänteiden tuntemusta, henkilökohtaisia uskomuksia/asenteita ja arviointitaitoja. Opettajien arviointiosaamisessa painottuu erityisesti pedagoginen osaaminen, kuten kyky mää-

rittää kulloinenkin arviointitarve, valita sopivat arviointimenetelmät ja tarvittaessa mukauttaa näitä oppilaan tarpeita vastaaviksi sekä lisäksi hyödyntää arviointitietoa opetuksessaan.

Suomalainen tutkimus kielitaidon arviointiosaamisesta perustuu pitkälti edellä kuvattuun kansainväliseen perustutkimukseen. Kansallisessa tutkimuksessa ollaan erityisen kiinnostuneita lukion kielenopettajien arviointiosaamisesta ja arviointikäytänteistä, koska lukiossa arviointia ohjaa opetussuunnitelman lisäksi käytännössä myös ylioppilastutkinto. Ylioppilastutkinnon vaikutus on tullut ilmi useissa lukion opettajien arviointikäytänteiden tutkimuksissa (esim. Välijärvi ym. 2009; Härmälä ym. 2019; Mäkipää 2021). Niissä on noussut esille ristiriita opetussuunnitelman arviointisuositusten ja luokkahuonearvioinnin käytänteiden välillä: lukiossa arvioinnissa korostuu kokeiden käyttö oppimisen tulosten, ei niinkään oppimisprosessin, arviointiin. Syynä tähän pidetään ylioppilastutkinnon vaikutusta, mikä näkyy mm. vanhojen ylioppilaskokeiden käyttönä osana oppilasarviointia ja keinona valmistautua kokeisiin (esim. Mäkipää 2021; Leontjev 2022). (Peruskoulun kielenopettajien arviointikäytänteistä ks. tarkemmin Tarnanen & Huhta 2011.)

Mäkipään (2021) väitöskirjatutkimus selvitti lukiolaisten näkemyksiä opettajien arviointi- ja erityisesti palautekäytänteistä. Tutkimuksessa havaittiin, että opiskelijoiden mielestä opettajien arviointi kohdistui oppimistuloksiin ja oli pääosin summatiivista. Opiskelijoiden näkemyksen mukaan opettajat antoivat heille vain vähän palautetta arvioinneista, vaikka opiskelijat olisivat kokeneet tällaisen palautteen tärkeäksi. Tulosten mukaan opiskeltava kieli ja opiskelijan kielitaidon taso olivat yhteydessä palautteen määrään ja hyödyllisyyteen. Opiskelijat kokivat saavansa enemmän palautetta ruotsin kuin englannin tunneilla, ja kielitaidoltaan paremmat opiskelijat pitivät saamaansa palautetta hyödyllisempänä kuin heikommat opiskelijat. Jälkimmäinen tulos nostaa esiin kysymyksen siitä, miten hyvin opettajat pystyvät sovittamaan palautteensa oppijoiden osaamistasolle sopivaksi.

Tutkimustiedon valossa näyttääkin siltä, että ainakin lukiossa arvioinnin painopiste on summatiivisessa arvioinnissa eikä tähän yhdisty riittävästi oppimista ohjaavaa palautetta. Osasy ty tilanteeseen on todennäköisesti ylioppilastutkinto. Kansainvälinen arviointitutkimus viittaa kuitenkin siihen, että tilanne on yleensä monisyinen: mikään yksittäinen tekijä ei täysin selitä arviointikäytänteitä, ei edes ylioppilastutkinnon kaltainen tärkeä päättökoe. Todennäköisesti kyse on monimutkaisesta eri tekijöiden yhteisvaikutuksesta kaikilla DFG:n (2016) viitekehyksen tasoilla. Mikrotasolla oppijan ja opettajan kulloinenkin vuorovaikutustilanne ja aikaisempien vuorovaikutusten historia tekevät aktiviteeteista yhteisiä, mutta samalla ainutlaatuisia. Mesotasolla luokan ja koulun sosiaaliset, taloudelliset ja kulttuuriset tekijät ovat vuorovaikutuksessa, ja opiskelijoiden ja opettajien kuuluminen muihin ryhmiin muokkaa kontekstia. Makrotasolla koulutuspolitiikka, ylioppilastutkinto ja koulutus- ja arviointikulttuuri vaikuttavat kokemuksiin sekä luovat ja rajoittavat oppimis- ja kehittymismahdollisuuksia.

Arviointiosaamisen tutkimus Suomessa (ja muuallakin; Poehner & Inbar-Lourie 2020) on keskittynyt kuvaamaan opettajien arviointiosaamista ja arviointikäytänteitä, mutta osaamisen ja käytänteiden kehittäminen on jäänyt vähemmälle. Tästä on toki poikkeuksia. Hildén & Fröjdendahl (2018) tarkastelivat kielten opettajiksi opiskelevien arviointiosaamista ennen ja jälkeen arviointikoulutuksen. He havaitsivat, että opettajaopiskelijoiden käsitykset arviointikäytännestä muuttuivat jo suhteellisen lyhyen koulutuksen aikana oppijakeskeisemmiksi ja dialogisemmiksi. Leontjev ja Pollari (2022, 2023) tutkivat muutoksia kielenopettajan arviointi- ja ohjauskäytänteissä sen jälkeen, kun opettaja oli tutustunut sellaisiin sosiokulttuurisiin käsitteisiin kuin lähikehityksen vyöhyke ja mediaatio. Lähikehityksen vyöhyke on oppimisen vaihe, jossa oppija ei vielä osaa jotakin tehtävää itsenäisesti, mutta pystyy siihen tuettuna. Mediaatiolla viitataan sosiokulttuurisessa tutkimuksessa merkityksiä välittävään ja rakentavaan vuorovaikutukseen, esimerkiksi tukeen, jota opettaja antaa oppilaalle (ja joka syntyy vuorovaikutuksessa oppilaan kanssa) tai jota oppilaat antavat toisilleen vuorovaikutuksen, kuten ohjaustilanteen tai ryhmätyöskentelyn aikana (Poehner & Leontjev 2020; Vygotsky 1978).

Leontjevin ja Pollarin (2022, 2023) tutkimukset perustuivat Vygotskin praksis-lähestymistapaan (Lantolf & Poehner 2014), jossa opettajan arviointiosaamista pyrittiin kehittämään teoriaa ja käytäntöä dialektisesti yhdistämällä. Tutkimusaineistona oli keskusteluja, joissa opettaja ja opiskelija sekä opiskelijat keskenään käsitelivät opiskelijoiden laatimia kirjoitelmia, joilla he valmistautuivat ylioppilastutkinnon vieraan kielen kokeeseen. Tutkimukset osoittivat, miten opettajan ymmärrys opiskelijan vieraalla kielellä kirjoittamisen ongelmista kehittyi näissä keskusteluissa, minkä ansiosta hän pystyi paremmin ohjaamaan opiskelijan kirjoittamista. Tärkeää on korostaa, että vaikka tavoitteena oli opettajan kehittyminen, lähtökohtana oli vuorovaikutus opettajan ja oppijoiden välillä, mikä antaa oppijoille toimijuutta tässä prosessissa. Kieltenopettajille on myös järjestetty tutkimusperusteista täydennyskoulutusta, jonka avulla opettajia on pyritty auttamaan summatiivisen ja formatiivisen arvioinnin yhdistämisessä johdonmukaiseksi oppimista tukevaksi kokonaisuudeksi (ks. Avointen oppimateriaalien kirjasto 2022).

Toinen aihealue, jolla Suomessa on järjestetty opettajankoulutusta, on dynaaminen arviointi. Se eroaa formatiivisesta arvioinnista siinä, että dynaamisessa arvioinnissa opetus ja arviointi ovat osa samaa toimintaa. Toisin sanoen opetus ei seuraa arviointia, kuten formatiivisessa arvioinnissa, vaan tapahtuu arvioinnin aikana (Poehner 2008). Vaikka toisen tai vieraan kielen dynaaminen arviointi on ollut varsin vakiintunutta esimerkiksi Yhdysvalloissa, se ei ole laajalti tunnettua Suomessa. Silti dynaamisesta arvioinnista on keskusteltu opettajien täydennyskoulutuksessa, kuten Opetushallituksen rahoittamassa kansallisessa ARVO-koulutuksessa (ks. Avointen oppimateriaalien kirjasto 2022). Myös Suomessa on tutkimusta, kuten DD-LANG-hanke (Soveltavan kielentutkimuksen keskus 2022), jonka tavoitteena on kehittää opettajan dynaamisen arvioinnin käytäntöjä luokkahuoneessa opettaja-tutkija-yhteistyössä.

Opettajien ja tutkijoiden yhteistyötä arviointiosaamisen kehittämiseksi tarvitaan Suomessa kuitenkin lisää. Hyviä esimerkkejä siitä, miten tämä voi tapahtua käytännössä, tarjoaa etenkin Poehnerin ja Inbar-Lourien (2020) toimittama teos, joka kuvaa kielenopettajien kokemia haasteita arvioinnissa ja jossa on lukuisia esimerkkejä näihin haasteisiin vastaamisesta yksittäisten opettajien ja tutkijoiden yhteistyön avulla, mutta myös erilaisissa pienemmissä ja isommissa ryhmissä ja kehittämishankkeissa.

## 4 Teknologian hyödyntäminen kielitaidon arvioinnissa

Teknologia ei ole mikään uusi ilmiö kielenopetuksessa eikä myöskään arvioinnissa. Itse asiassa kielenopetus näyttää olleen uusien teknologisten opetusmenetelmien eturintamassa jo 1970-luvulta lähtien, jolloin audiolingvaalinen kielen opetus toi kielistudiot kouluihin (ks. myös Jakonen ym. tässä teemanumerossa). Suomessa kielistudioiden tarvetta lisäsi samalla vuosikymmenellä toteutettu ylioppilastutkinnon uudistus, koska puheen ymmärtämisen koe voitiin suorittaa kielistudiossa helpommin kuin tavallisessa luokassa. Kannettavat tietokoneet ja kuulokkeet ovat toki nyt tehneet kokeiden ja harjoitustehtävien suorittamisen mahdolliseksi missä tahansa.

Ylioppilastutkinto vaikutti siis kielitaidon arvioinnin teknologiseen kehitykseen Suomessa jo puoli vuosisataa sitten, ja tätä roolia on vain vahvistanut tutkinnon digitalisointi 2010-luvun lopulla. Ylioppilaskokeet suoritetaan nykyään digitaalisesti ja niihin myös valmistaudutaan suorittamalla vanhoja digitaalisia kokeita ylioppilastutkintolautakunnan laatiman Abitti-järjestelmän kautta. Lisäksi lukioiden kielenopettajat voivat laatia ja pitää omia kurssikokeitaan Abitin avulla. Vastaavaa järjestelmää ei ole kehitetty peruskouluun, jossa digitaalisuus kielitaidon arvioinnissa liittyy digitaalisissa oppimateriaaleissa oleviin kokeisiin ja muihin arviointitehtäviin, joita voidaan suorittaa joko tietokonealuokissa tai oppilaiden omilla koneilla.

Seuraavassa analysoimme, miten teknologiaa hyödynnetään kielitaidon arvioinnissa kouluissa. Kuvaamme, miten teknologia auttaa arvioinnissa ja mitä rajoituksia se asettaa. Käytämme esimerkkeinä ylioppilastutkinnon (Abittin) ja digitaalisten oppimateriaalien tehtäviä. Lopuksi hyödynnämme kansainvälisiä kielitutkintoja ennakoidaksemme, miten teknologia muuttaa kielitaidon arviointia tulevaisuudessa.

Teknologialla on erilaisia rooleja kielitaidon arvioinnissa. Yksi rooli on toimia alustana, jolla oppilaat ja opettajat käyttävät kieltä ja mahdollisesti samalla arvioivat itseään ja toisiaan. Vuonna 2020 alkanut koronapandemia muutti suuren osan opetuksesta etäopetuksiksi. Tällaiseen viestintään on helppo yhdistää erilaisia itse-, vertais- ja opettaja-arvioinnin muotoja, joilla voi olla sekä formatiivisia että summatiivisia tavoitteita. Näissä tilanteissa teknologia toimii vain viestinnän välineenä samalla tavoin kuin esimerkiksi kielistudiossa.

Teknologialla voi myös olla oppilaiden suorituksia arvioiva rooli. Tyypillisimpiä esimerkkejä tästä ovat erilaiset monivalintakokeet, joita on sekä ylioppilastutkinnossa

että digitaalisissa oppimateriaaleissa. Kokeen laatija määrittelee, mikä vaihtoehto on oikea tai paras kuhunkin kysymykseen ja teknologia (käytännössä tietokoneohjelma, jonka kautta koe pidetään) vertaa oppilaan vastausta laatijan päättämään vastaukseen ja luokittelee sen mukaan vastaukset oikeiksi tai vääriksi. Nykyisissä tehtävissä ja kokeissa voi myös olla lyhyitä, tyypillisesti yhden sanan pituisia oppilaan itse tuottamia kirjallisia vastauksia, koska oppilaan tuottamaa merkkijonoa on helppo verrata tehtävän laatijan antamaan mallivastaukseen.

Yhteenvetona nykyisistä kielitaidon teknologisista arviointitavoista voidaan todeta, että teknologia tarjoaa uusia mahdollisuuksia etenkin suullisen kielitaidon arviointiin, koska arviointi voidaan toteuttaa etänä ja osana normaalia etäopetusta. Suomessa ei kuitenkaan ole vielä käytössä järjestelmiä, jotka pystyisivät automaattisesti arvioimaan puhetta tai kirjoitusta, vaan näiden taitojen arviointi on edelleen täysin ihmisten varassa, joko opettajien tai oppilaiden. Ymmärtämisen arviointiin teknologia on tuonut myös apua, koska tietyt tehtävätyypit sopivat automaattiseen arviointiin, jossa kyse on vastauksen ja mallin mekaanisesta vertailusta. Näin opettaja tai oppilas itse säästyy vastausten pisteyttämisen vaivalta.

Teknologia vaikuttaa siis siihen, miten esimerkiksi kielitaitoa arvioidaan, mutta sillä on vaikutusta myös siihen, mitä arvioidaan. Teknologia tarjoaa mahdollisuuksia laajentaa arvioitavan osaamisen kirjoa ja lisätä arvioinnin autenttisuutta, mutta se voi myös kaventaa arviointia. Esimerkiksi puheen ja kirjoituksen automaattinen arviointi ei ole vielä käytössä Suomessa (ainakaan laajemmassa käytössä olevissa testeissä), joten täysin teknologiaan perustuva arviointi tarkoittaisi vain ymmärtämisen ja esimerkiksi sanaston ja rakenteiden hallinnan arviointia, koska niitä voidaan testata monivalintatehtävien kaltaisten täysin mekaanisesti pisteytettävien tehtävien avulla. Toisaalta nykyinenkin teknologia mahdollistaa mm. monimediaisen puheen ymmärtämisen arvioinnin. Esimerkiksi ylioppilastutkinnon digitaalisissa puheen ymmärtämisen kokeissa käytetään videonauhoitteita perinteisten äänitiedostojen lisäksi, mikä tekee arvioitavista kielenkäyttötilanteista jonkin verran autenttisempia. Vaikka arviointivastuu säilyisi ihmisillä (oppilailla ja opettajilla), voidaan kuitenkin myös puheen ja kirjoittamisen arviointia monipuolistaa teknologian avulla, esimerkiksi niin, että oppilaat tallentavat puhumistehtäviä mobiililaitteiden avulla ja lähettävät ne sitten opettajalle (ja mahdollisesti toisilleen) arviointia ja palautetta varten. Tällaiset tallenteet arviointiin ja palautteeseen voidaan sitten tallentaa vaikkapa digitaaliseen kielisalkkuun.

Kansainvälisen kehityksen perusteella on odotettavissa, että kirjoittamisen ja puheen automaattinen arviointi yleistyy Suomessakin, esimerkiksi ylioppilastutkinnossa ja muissa digitaalisissa testeissä (esim. Karvin kansalliset arviointitutkimukset). Tulevaisuudessa on todennäköisesti käytettävissä myös automaattisia arviointityökaluja, joiden avulla oppilaat voivat saada palautetta digitaalisten oppimateriaalien puhe- ja kirjoitustehtävistä. Esimerkiksi DigiTala-hankkeessa kehitetään tällaista järjestelmää suomen ja ruotsin kielen oppijoille (Kautonen & von Zansen 2020). Niissä kansainvälisissä kielitutkinnoissa, joissa käytetään automaattista arviointia, konearviointi yhdistetään

yleensä ihmisarviointiin, mutta toki täysin automaattisiakin järjestelmiä löytyy (ks. tarkemmin esim. Shermis & Burstein 2013; Zechner & Evanini 2020). Automaattinen arviointi perustuu tekoälyn ja erilaisten kielen analysointimenetelmien (*natural language processing*) käyttöön, ja tyypillinen lähestymistapa on kehittää algoritmeja, jotka simuloivat ihmisarvioijien toimintaa. Yksinkertaisinta tällainen arviointi on ymmärtämistehtävien avointen vastausten sisällön arvioinnissa, vaativampia sovelluksia tarvitaan puheessa ja kirjoituksessa (ks. tarkemmin esim. Zechner & Evanini 2020; suomalaisia tutkimuksia aiheesta mm. Honko ym. 2019; Kautonen & von Zansen 2020). Puheen arviointi edellyttää lisäksi, että oppijan puhe voidaan tunnistaa ja muuttaa tekstiksi, jos siitä on määrä arvioida esimerkiksi sanaston ja rakenteiden hallintaa (ks. DigiTala -hanke esimerkkinä suomalaisesta tutkimuksesta; Kautonen & von Zansen 2020). Vuorovaikutteisen puhumisen arvioinnissa tekoälyä hyödyntävien virtuaalihahmojen (engl. *avatar*) käyttö keskustelukumppanina tai arvioijana/haastattelijana on yksi todennäköinen tulevaisuuden kehityssuunta.

On kuitenkin muistettava, että etenkin vuorovaikutteisen puhumisen automaattinen arviointi on vielä varsin rajallista, koska tietoteknisin keinoin on vaikea arvioida, miten puhuja saavuttaa viestinnälliset tavoitteensa tai suoriutuu tehtävästä. Tietotekniikka mahdollistaa myös yksityiskohtaisemman palautteen summatiivisista testeistä kuten ylioppilaskokeista, mikä voisi olla askel tutkinnon hyödyntämiseen myös formatiivisesti, esimerkiksi osana opiskelijan lukion jälkeisiä kieliopintoja. Teknologia mahdollistaa myös digitaalisessa ympäristössä tapahtuvan opiskelun jatkuvan ja monipuolisen monitoroinnin, mikä häivyttää opetuksen, oppimisen ja arvioinnin rajoja (Cope & Kalantzis 2022). Tämä on erityisen hyödyllistä formatiivisessa arvioinnissa mutta voi myös parantaa summatiivisen arvioinnin validiutta, koska siinä pystytään hankkimaan oppilaan osaamisesta paljon kattavampaa näyttöä kuin yksittäisellä kokeella on mahdollista.

Adaptiivinen, osallistujan osaamiseen mukautuva arviointi on yksi uudistus, jonka teknologia mahdollistaa. Adaptiivisessa testissä oppilaan vastaukset vaikuttavat seuraavien tehtävien vaikeuteen: oikeat vastaukset johtavat hiukan haastavampiin tehtäviin ja väärät vastaukset puolestaan helpompiin, kunnes järjestelmä onnistuu arvioimaan oppilaan osaamistason tarpeeksi tarkasti. Tällaiset testit vaativat tehtävien esitestausta niiden vaikeustason selvittämiseksi ja varsin laajaa tehtävä- tai osiopankkia, joten ne ovat toistaiseksi harvinaisia. Adaptiiviset testit soveltuvat toistaiseksi vain monivalinnan tyyppisiin tehtäviin, mikä rajaa niiden käytön lähinnä ymmärtämisen arviointiin. Suomalainen esimerkki adaptiivisesta kielitestistä on ensikielen lukemisen alkuvaiheisiin ja mahdollisten lukemisongelmien diagnosointiin ja harjoitteluun kehitetty Ekapeli, jonka tehtävät mukautuvat oppilaan osaamistasolle (Richardson & Lyytinen 2014). Tanskan koulujärjestelmässä on käytössä valtakunnallisia adaptiivisia formatiivisia testejä useissa oppiaineissa, mukaan lukien tanskan ja englannin kielissä (ks. Huhta & Boivin 2023 näiden testien haasteista).



## 5 Monikielisyys eri oppiaineiden arvioinnissa

Yksi keskeinen suomalaiseen yhteiskuntaan ja kouluun 2000-luvulla vaikuttavista muutosprosesseista on monikielistyminen (ks. myös Lehtonen ym. tässä teemanumerossa). Monikielinen osaaminen ja kyky toimia monikielisessä ympäristössä ovat Euroopassa asetetuissa yhteisissä tavoitteissa kielten opetuksen, oppimisen ja arvioinnin keskeisiä sisältöjä (Euroopan neuvosto 2001, 2020). Nykyisen, laajan monikielisyysnäkömyksen mukaan jokainen koulu ja luokka on monikielinen, sillä oppilaat osaavat äidinkieltensä lisäksi muita kotona, koulussa ja vapaa-ajalla opittuja kieliä sekä murteita ja muita erilaisia kielimuotoja (esim. OPH 2022; Piccardo ym. 2021). Siksi monikielisyys koskee kaikkia luokka-asteita ja myös varhaisempia vaiheita.

Koulujen kielellinen ja kulttuurinen moninaisuus korostaa kielitietoisuuden opetuksen ja arvioinnin tarvetta. Kielitietoisessa koulussa oppilaita ohjataan arvostamaan eri kieliä ja kulttuureja sekä kehittämään kaksi- ja monikielisyystään, kielitietoisuuttaan ja metalingvistisiä taitojaan (OPH 2014a: 86). Tavoitteena on muun muassa vahvistaa oppilaiden kielellistä itsetuntoa ja rohkaista vähäisenkin osaamisen hyödyntämiseen sekä taitojen käyttämiseen vuorovaikutuksessa ja tukena kaikessa oppimisessa yli oppiainerajojen. Opettajalle kielitietoinen opetus monikielisessä koulussa merkitsee ennen kaikkea sitä, että ymmärretään kielen rooli kaikessa toiminnassa: perehdytään tiedonalan, oppiaineen ja -materiaalien sekä koulun vuorovaikutustilanteiden kielellisiin piirteisiin ja vaativuuteen erilaisten oppilaiden näkökulmista, otetaan selvää oppilaiden kielellisistä resursseista ja mukautetaan omaa toimintaa – myös arviointia – tämän tiedon varassa (mm. OPH 2017; Mustonen & Honko 2018; Aalto ym. 2019).

Opetussuunnitelman perusteissa monikielisyttä on lähestytty erityisesti koulutuksen arvopohjan, hyvän opetuksen piirteiden sekä oppilaan kasvun ja kehityksen kautta. Arviointia koskevia linjauksia sen sijaan on tehty vähemmän. Arvopohjan kannalta monikielisyden huomioimisen voidaan katsoa tarkoittavan muun muassa sitä, että oppilaiden monikieliset resurssit tunnistetaan ja niille annetaan arvoa riippumatta yksittäisten kielten asemasta koulun opetuskielenä (ks. myös García & Wei 2014; Lehtonen ym. tässä teemanumerossa). Näin voidaan toimia esimerkiksi kaikkien luokassa puhuttujen eri äidinkielten osalta, vaikeivat kaikki oppilaat osallistuisi oman kielen opetukseen ja saisi arvosanoin annettua palautetta. Myönteinen asenne monikielisten resurssien ylläpitämistä ja kehittämistä kohtaan ei edellytä opettajalta laajaa kielitaitoa mutta voi lisätä tasa-arvoa ja osallisuutta koulussa sekä yhteiskunnassa laajemmin (García & Kleyn 2016).

Koulussa niin kielten opetus kuin arviointikin ovat itsessään monikielisyttä tukeva ideologinen valinta (ks. myös Liikkanen ym. tässä teemanumerossa). Lisäksi koulun arviointitoiminta oppiaineesta riippumatta voi olla monikielistä tai monikielisyttä muutoin tukevaa. Tällöin kieliä ei pyritä mekaanisesti eristämään toisistaan, vaan kaikki opettajat luokanopettajista kielenopettajiin ja reaaliaineen opettajiin voivat esimerkiksi tukea oppilaita käyttämään eri kieliä tai kielimuotoja joustavasti



osana osaamisen osoittamisen prosessia (esim. Shohamy 2011; Stathopoulou 2020). Oppilaalle tämä voi tarkoittaa vaikkapa mahdollisuutta perehtyä tehtävänantoihin ja vastata kokeessa opittavan kielen tai opetuskielen lisäksi muullakin tarkoituksenmukaisella kielellä yhdessä sovituin järjestelyin. Opettaja voi myös rohkaista oppilasta hyödyntämään tehtävien tekemisessä kaikkia kielellisiä resurssejaan tai osoittamaan mediaatitaitojaan, jolloin huomio siirtyy kielestä sisältöihin ja vuorovaikutukseen.

Jos halutaan kehittää monikielistä arviointikulttuuria syvällisesti, on otettava toisissaan paradigmanmuutos yksikielisestä monikieliseksi (Shohamyn 2013 ja Lopezin ym. 2017 pohjalta Stathopoulou 2020: 43), ei vain tiettyjen ennalta sovittujen menetelmien haltuun ottaminen. Jos oppilaalla on puutteita opetuskielen hallinnassa, opettajan tulee huomioida tämä arviointi- ja näyttötilanteiden suunnittelussa ja toteutuksissa yksilöllistämällä arviointia tarpeen mukaan. Oppilaille voidaan esimerkiksi tarjota mahdollisuus saada lisäaikaa tehtävien tekemiseen, käyttää teknisiä apukeinoja tai vastata tehtäviin kirjoittamisen sijaan suullisesti (opetuksen eriyttämisestä, ks. Roiha ym. tässä teemanumerossa). Kielitietoisuuden opetuksen toteutumisen kannalta tärkeää on, että niin opetus kuin arviointi tehdään oppilaille mahdollisimman hyvin saavutettavaksi ja tasapuoliseksi (esim. OPH 2014a; Aalto ym. 2019).

Käytännössä arvioinnin periaatteet ja käytänteet ovat sekä koulussa että erillisissä testijärjestelmissä kuitenkin pitkään perustuneet essentialististen, kielet erillään pitävien rakenteiden varaan (esim. García & Kleyn 2016; North 2021). Monikielisyyden ideaali ei ole siirtynyt Euroopassa kitkatta koulujen arkiseen toimintaan asti, ja myös monikielinen testiarviointi sekä arviointitiedon hyödyntäminen hakevat vielä muotojaan (North 2021; de Jong 2022). Kuten Leung (2022) toteaa, tarvitaan myös lisää käsitteanalyysia ja kehittämistyötä ennen kuin monikielisyyden rooli ja suhde lähikäsitteisiin, kuten mediaatitaitoihin, voidaan operationalisoida systemaattisesti.

Vaikka Suomessa monikielisyyden tukeminen ja monikieliseksi kasvaminen ovat opetussuunnitelmien perusteissa asetettuja yleisiä tavoitteita ja eri tavoin läsnä kaikkien oppiaineiden arvioinnissa, monikielisyyden mainitaan eksplisiittisesti vain muutamien oppimäärien päättöarvioinnin kriteereissä. Esimerkiksi äidinkielenomaisen ruotsin tavoite T3 vuosiluokilla 3–6 ohjaa oppilaita herättämään uteliaisuutta omaa kaksi- tai monikielisyyttään ja kaksikulttuurisuuttaan kohtaan sekä havaitsemaan ilmiöitä, jotka yhdistävät ja erottavat eri kieliä (OPH 2014a: 211). Kuudennen luokan päättöarvioinnissa monikielisyyden sisältyminen kielellisen päättelyn sisältöalueen arviointiin, jossa arvosanaa hyvä (8) vastaavaa osaamista kuvataan taitona havainnoida ja vertailla kielten ominaispiirteitä. Yhdeksännen vuosiluokan tavoitteisiin puolestaan kuuluu taitojen ja halukkuuden vahvistaminen monikielisessä ja -kulttuurisessa ympäristössä toimimiseen. Tavoitteen saavuttamista arvioidaan painottaen kielellisten oikeuksien ja velvollisuuksien tiedostamista: ”Oppilas tietää, että jokaisella on oikeus käyttää ja kehittää omaa kieltään. Oppilaalla on käsitys Suomen kansalliskieliin liittyvistä oikeuksista ja velvoitteista. Oppilas tuntee ruotsin kielen asemaa Suomessa ja Pohjoismaissa ja maailman kielten joukossa” (Arvosanaa hyvä vastaava osaaminen, mts. 342).

Perusopetusta täydentävän romanikielen tavoitteissa monikielisyys mainitaan vuosiluokkien 7–9 tavoitealueella T1 nimeämällä oppimisen tavoitteiksi monikielisyys ja -kulttuurisissa vuorovaikutustilanteissa toimiminen sekä tähän liittyvä oman toiminnan arvioiminen (OPH 2014a: 460). Lisäksi tavoitealueella T6 tavoitteeksi on asetettu oman kaksi- tai monikielisuuden hyödyntäminen kielen, kirjallisuuden ja kulttuurin ymmärtämisessä. Vieraan kielen B2-oppimäärän tavoitealueen T1 kuvaukseen puolestaan on sisällytetty yleisesti monikielisuuden kehittymisen tuki (mts. 360). Näitä tavoitteita vastaavissa osaamiskriteereissä monikielisyyttä ei mainita. Suomi toisena kielenä ja kirjallisuus -oppimäärän kuvauksissa monikielisyys puolestaan on tavoitealueiden lisäksi nimetty osaksi sekä kuudennen että yhdeksännen vuosiluokan päättöarvioinnin kriteereitä. Hyvään osaamiseen kuuluu, että oppilas osaa kuvailla omaa sekä elinympäristönsä monikielisyyttä ja -kulttuurisuutta (6. vuosiluokka, mts. 189) ja perustella yhteiskunnan monikielisuuden merkitystä (9. luokka, mts. 318).

Perusopetuksen opetussuunnitelman perusteiden luvussa 6 kuvatut arvioinnin yleiset periaatteet ja paikalliset opetussuunnitelmat luonnollisesti toimivat opettajan ohjenuorana kaikkien arviointia koskevien ratkaisujen tekemisessä, vaikka ne eivät kohdistukaan suoraan oppilaiden monikielisyyteen tai monikielisiin arviointikäytänteisiin. Perusteita noudattelevan monikielisyyttä tukevan arvioinnin tulee olla oppimista edistävää ja rohkaisevaa, monipuolista ja vuorovaikutteista sekä oikeudenmukaista ja eettistä. Oppilaan on tärkeää saada tietoa omasta edistymisestä ja tukea myös tavoitteiden asettamiseen sekä niiden saavuttamisen itsenäiseen arviointiin. Vuorovaikutteisuus ja yhteistyö arviointiprosessin eri vaiheissa on erityisen tärkeää monikielisyyskoulussa. Yksittäiset tehtävät, kokeet ja kielitestit keskittyvät tyypillisesti yhteen kieleen kerrallaan. Ne antavat tärkeää tietoa kyseisten kielten osaamisesta mutta samalla pilkkovat ja osin piilottavat oppilaiden kokonaisosaamista.

Yksittäisiä kieliä laajemman arviointitiedon kokoamisessa voidaan opettaja-arvioinnin lisäksi käyttää apuna vertaisarviointia sekä monipuolisia itsearviointimenetelmiä (ks. kuitenkin Saarinen & Huhta 2023 oppikirjojen itsearviointitehtäviin liittyvistä ongelmista): nuorempien oppilaiden kanssa esimerkiksi erilaisia visuaalisia ja narratiivisia menetelmiä (ks. esim. Pietikäinen ym. 2010; Pitkänen-Huhta & Pietikäinen 2013) sekä Eurooppalaista kielisalkkua (OPH 2011). Vanhempien oppilaiden kanssa kielenkäyttötilanteita, funktioita ja taitoja sekä omia tavoitteita voidaan eritellä tarkemmin ja yksityiskohtaisemmin esimerkiksi mukaillen lukion kieliprofilia (OPH 2021).

Lisäksi kielitaidon opetuksen oppimisen ja arvioinnin yhteinen eurooppalainen viitekehys, sekä alkuperäinen että uusi (Euroopan neuvosto 2001, 2020), sisältää lukuisia kuvaimia, joita voidaan soveltaa monikielisyyttä tukevassa opetuksessa ja arvioinnissa (North 2021). Kokonaan uusia kuvaimia ovat uudessa viitekehyksessä monikielistä ja kulttuurista osaamista sekä mediaatitaitoja koskevat kuvaimet. Näissä taitoja kuvataan osaamisperusteisesti ja taitotasoinen etenevinä, kuten muissakin viitekehysten kuvaimissa (myös Honko & Halonen tässä teemanumerossa).

Itsearvioinnissa kuvaimet mahdollistavat pitkäjänteisesti käytettynä oppilaalle omaa monikielisyyttä koskevien tavoitteiden asettamisen ja kehittymisen, esimerkiksi omien vahvuuksien seurannan. Kuvaimia yhdistää huomion kiinnittäminen merkityksiä rakentavaan tekstitoimintaan ja vuorovaikutukseen tilanteissa, joissa ei toimita yksin tai vain yhden kielimuodon varassa (suomeksi OPH 2018). Opettaja voi saada kuvaimista ja niiden suomalaisista sovelluksista virikkeitä esimerkiksi sen havainnointiin, miten oppilaat ymmärtävät ja osaavat työstää suullisia, kirjallisia tai monimodaalisia tekstejä toiseen muotoon tai toiselle kielelle, kuinka he osaavat hyödyntää tässä yhteistyötä ja ongelmanratkaisutaitoja sekä miten he huomioivat toiminnassaan muut osallistujat: tukevatko he muiden osallistumista, osaavatko he huomioida erilaisen kieli- ja kulttuuritaustan jne. (myös Euroopan neuvosto 2020; North 2021).

Kyky viestiä opittavalla kielellä kehittyy tyypillisesti hiljalleen, joskaan ei lineaarisesti konkreettisista havainnoista abstraktien ilmiöiden käsittelyyn, ja kielenkäytöstä tulee yhä monipuolisempaa, sujuvampaa, tarkempaa ja joustavampaa (esim. Aalto ym. 2019). Lisäksi monikielisellä oppijalla on tietoa ja osaamista, jota hän hyödyntää yli kielirajojen. Yksi tapa arvioida monikielisiä resursseja on tarkastella niitä osaa-misperusteisesti kykyinä toimia kielellä erilaisissa ikä- ja kehitysvaiheen mukaisissa tärkeissä tilanteissa. Miten oppilas selviää eri tilanteissa esimerkiksi ruotsiksi, omalla äidinkielellään, suomeksi, englanniksi tai limittäin eri kieliä käyttäen? Kuinka tarkasti hän ymmärtää tilanteissa käytettyjä tekstejä – saako selvän pääasioista tai myös yksityiskohdista – tai kuinka monipuolisesti ja joustavasti hän itse käyttää kieltä erilaisiin tarpeisiin?

Tarkastelemalla oppijan kielellistä toimintaa ja osallistumista vuorovaikutukseen erilaisissa ryhmissä ja tilanteissa opettaja saa tietoa esimerkiksi siitä, mitkä vaikkapa koulun opetuskielessä tärkeät tai vieraan kielen oppimistavoitteisiin kuuluvat asiat jo sujuvat itsenäisesti tai apuvälineiden avulla, missä taas tarvitaan enemmän vuorovaikutuskumppanin tukea ja harjoitusta. Samalla on kuitenkin syytä arvioida, millaisia tuen tarpeita oppijalla on: liittyvätkö ne esimerkiksi kielen hallintaan, kuten ilmaisuvarannon laajuuteen tai kykyyn yhdistellä ja muokata kielen aineksia pidemmiksi ilmauksiksi ja teksteiksi? Vai onko kyse siitä, että tilanne itsessään, käsiteltävät asiat tai vuorovaikutuskumppaneiden odotukset ovat oppijalle jollain tapaa uusia tai vaikeita tai hänen omien mielenkiinnon kohteidensa ulkopuolella?

Osaamisen laajempi kartoittaminen ja keskustelu oppilaan itsensä, huoltajien ja oppilaan oman äidinkielen opettajan kanssa on tärkeää varsinkin silloin, kun oppilas opiskelee koulussa muulla kielellä kuin äidinkielellään. Keskustelun avulla oppilaan monikielisistä resursseista, oppimisen ja koulutyöskentelyn sujumisesta sekä mahdollisista erityisen tuen tarpeista saadaan mahdollisimman kattava ja ajantasainen käsitys. Konkreettisia välineitä varhaislapsuudesta alkavaan monikielisen kielitaidon arviointiin ja kielitietoiseen yhteistyöhön on kehitetty useissa työryhmissä ja hankkeissa (ks. esim. Espoon kaupunki 2010, Turun yliopisto 2020; Kangasvieri ym. 2021). Välineitä voi hyödyntää kertaluonteisesti, mutta pitkittäin ja systemaattisesti kerätystä

arviointitiedosta on eniten ja pitkäkestoisinta hyötyä (myös OPH 2014a). Tieto opiskelun sujumisesta eri oppiaineissa sekä eri kielten taidoista ja rooleista oppilaan arjessa auttaa opettajaa esimerkiksi sopivien arviointitapojen valinnassa ja tarvittavan tuen antamisessa oikeaan aikaan (mm. Aalto ym. 2019).

## 6 Lopuksi

Laadukkaan arvioinnin suunnittelu ja toteutus on osa opettajan ammattitaitoa. Arvioinnin on kohdistuttava opetuksen tavoitteisiin ja oltava oppijoiden kehitysvaiheeseen ja tarpeisiin nähden sopivaa. Lisäksi arvioinnin tulee perustua riittävään näyttöön ja olla toistettavissa eri ryhmien kanssa. Sekä opettajan että oppilaan työn mielekkyyden kannalta olennaista on myös arvioinnin toteutettavuus käytännössä sekä arviointitiedon todenperäisyys ja soveltuvuus siihen käyttöön, mihin tietoa kulloinkin tarvitaan. Oppijan näkökulmasta arvioinnin kohdetta, reliabiliteettia ja käytettävyyttäkin tärkeämpää voi kuitenkin olla arvioinnin eettisyys, reiluus, läpinäkyvyys ja kannustavuus eli se, kohteleeko (etenkin summatiivinen) arviointi kaikkia oppilaita tasapuolisesti ja oikeudenmukaisesti, onko arviointi avointa ja oppilasta arvostavaa sekä kannustavaa ja perustuuko se kaikkien, myös oppilaan itsensä, tuntemiin periaatteisiin.

Opettajilla on oman työnsä asiantuntijoina paras ja ajankohtaisin näkymä juuri omassa työssään ajankohtaisiin arviointitarpeisiin. Hyvien arviointikäytänteiden tuntemus on tärkeää, mutta vähintään yhtä tärkeää on viedä laadukkaan arvioinnin keskeiset periaatteet arjen käytäntöihin asti sekä sovittaa arviointi paitsi arviointitarpeisiin myös omaan ajankäyttöön ja jaksamiseen. Arvokkaita ja edelleen kehittämisen arvoisia keinoja tähän tarjoavat muun muassa arviointivastuun jakaminen ja arvioinnin vaiheistaminen oppimisprosessin eri vaiheisiin, arvioinnin vuorovaikutteisuus, teknologian tarkoituksenmukainen käyttö ja monimodaalisten ja monikielisten resurssien hyödyntäminen arvioinnissa. Useimmiten arviointitaitojen kehittäminen ei edellytä opettajalta niinkään kokonaan uuden oppimista tai uutta tempuvalikoimaa, vaan tärkeintä on tunnistaa oppimista tukevan arvioinnin kulloisetkin tarpeet sekä omat vahvuudet ja mielenkiinnon kohteet esimerkiksi eri menetelmien käytössä ja palautteen annossa. Näiden pohjalta, oppilaita kuullen ja yhdessä muiden opettajien kanssa, jo olemassa olevaa arviointiosaamista on mahdollista kehittää osana koulun tavallista arkea.

## Kirjallisuus

- Aalto, E., S. Mustonen, M. Järvenoja, & J. Saario 2019. *Monikielisen oppijan matkassa. Verkkosivusto opettajankoulutukseen*. Erityisesti osio *Kielitaidon kehitys ja arviointi ja sen alla kohdat Osallistuminen ja vuorovaikutustaidot sekä Monikielisyys resurssina*. Jyväskylä: Jyväskylän yliopisto.  
<https://monikielisenoppijanmatkassa.fi/kielitaidon-kehitys-ja-arviointi/>
- Avointen oppimateriaalien kirjasto 2022. *ARVO – Kieltenopettajien arviointiosaamisen päivittäminen -koulutuksen osallistujilta koottuja ideoita*.  
[https://urn.fi/urn:nbn:fi:oyerfi-202209\\_00025072\\_1](https://urn.fi/urn:nbn:fi:oyerfi-202209_00025072_1)
- Bachman, L. F. & A. S. Palmer 1996. *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F. & A. S. Palmer 2010. *Language assessment in practice*. Oxford: Oxford University Press.
- Chan, C. & C. Davison 2020. Learning from each other: school-university collaborative action research as praxis. Teoksessa M. E. Poehner & O. Inbar-Lourie (toim.) *Toward a reconceptualization of second language classroom assessment*. Cham: Springer, 129–149.
- Chapelle, C., M. Enright, & J. Jamieson (toim.) 2008. *Building a validity argument for the test of English as a foreign language*. New York: Routledge.
- Cheng, L. & N. Sultana 2021. Washback. Looking backward and forward. Teoksessa G. Fulcher & L. Harding (toim.) *The Routledge handbook of language testing* (2. painos). London: Routledge, 136–152.
- Cope, B. & M. Kalantzis 2022. The changing dynamics of online education: five theses on the future of learning. Teoksessa C. Lütge (toim.) *Foreign language learning in the digital age. Theory and pedagogy for developing literacies*. New York: Routledge, 9–33.
- Davison, C. & C. Leung 2009. Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43 (3), 393–415.
- Douglas Fir Group (DFG) 2016. A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal*, 100 (S1), 19–47. <https://doi.org/10.1111/modl.12301>
- Espoon kaupunki 2010. *Pienten kielireppu*. [https://peda.net/pori/perusopetus/opiskelun-tuki/mmo/arviointi/nimet%C3%B6n-4176/p.file/download/21adc25c4cbbad4d8e53621e33d0a824b165db73/Pienten\\_kielireppu.pdf](https://peda.net/pori/perusopetus/opiskelun-tuki/mmo/arviointi/nimet%C3%B6n-4176/p.file/download/21adc25c4cbbad4d8e53621e33d0a824b165db73/Pienten_kielireppu.pdf)
- Euroopan neuvosto 2001. *Common European Framework of Reference for Languages*. Strasbourg: Council of Europe.
- Euroopan neuvosto 2020 [2018]. *Common European Framework of Reference for Languages. Companion volume with new descriptors*. Strasbourg: Council of Europe.
- Fulcher, G. 2012. Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9 (2), 113–132.
- García, O. & T. Kleyn 2016. Translanguaging theory in education. Teoksessa O. García & T. Kleyn (toim.) *Translanguaging with multilingual students: learning from classroom moments*. New York, London: Routledge, 9–33.
- García, O. & L. Wei 2014. *Translanguaging: language, bilingualism, and education*. New York, NY: Palgrave MacMillan.
- Hildén, R. & Rautopuro, J. 2014. *Ruotsin kielen A-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013*. (Kansallisen koulutuksen arviointikeskuksen julkaisuja; No. 2014:1). Opetushallitus, Kansallisen koulutuksen arviointikeskus.  
<https://karvi.fi/publication/ruotsin-kielen-oppimaaran-oppimistulokset-perusopetuksen-paattovaiheessa-2013/>
- Hildén, R., Ouakrim-Soivio, N. & Rautopuro, J. 2016. Kaikille ansionsa mukaan? Perusopetuksen päättöarvioinnin yhdenvertaisuus Suomessa. *Kasvatus*, 47(4), 342–357.

- Hildén, R., & Fröjdendahl, B. 2018. The dawn of assessment literacy – exploring the conceptions of Finnish student teachers in foreign languages. *Apples: Journal of Applied Language Studies*, 12(1), 1–24. <https://doi.org/10.17011/apples/urn.201802201542>
- Honko, M., Jarvis, S. & Vainio, S. 2019. Lukijat sanaston monimuotoisuutta määrittämässä: leksikaalisen diversiteetin tarkastelua määrällisen ja laadullisen tutkimuksen rajapinnassa. *Virittäjä* 123 (1); 44–78. <https://doi.org/10.23982/vir.59025>
- Huhta, A. 2020. Taitotasojen hyödyntämisestä opetuksessa ja arvioinnissa. Teoksessa M. Härmälä & M. Puukko (toim.) *Taitotasojen ja suullista kielitaitoa. Näkökulmia englannin oppimistulosten arvioinnin osoittamiin kehittämiskohteisiin*. Helsinki: Kansallinen koulutuksen arviointikeskus, 25–47. [https://karvi.fi/wp-content/uploads/2020/05/KARVI\\_Artikkelikokoelma\\_1\\_2020.pdf](https://karvi.fi/wp-content/uploads/2020/05/KARVI_Artikkelikokoelma_1_2020.pdf)
- Huhta, A. & N. Boivin 2023. Changes in language assessment through the lens of new materialism. Teoksessa J. Ennsner-Kananen & T. Saarinen (toim.) *New materialist explorations into language education*. Cham: Springer, 39–56. <https://link.springer.com/book/10.1007/978-3-031-13847-8>
- Härmälä, M., M. Huhtanen, M. Puukko, & J. Marjanen 2019. *A-englannin oppimistulokset 7. vuosiluokan alussa 2018*. Helsinki: Kansallinen koulutuksen arviointikeskus.
- Inbar-Lourie, O. 2017. Language assessment literacy. Teoksessa E. Shohamy, E., I. Or, & S. May (toim.) *Language testing and assessment. Encyclopedia of language and education*. Cham: Springer, 257–270
- 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, 53–63.
- Kane, M.T. 2006. Content-related validity evidence in test development. Teoksessa S. M. Downing & T. M. Haladyna (toim.) *Handbook of test development*. Mahwah, NJ: Lawrence Erlbaum, 131–153.
- Kangasvieri, T., L. Lempel, A. Palojärvi, & J. Moate (toim.) 2021. *Innovatiivisen kielikasvatuksen kartta & kompassi: askeleita kehitystyöhön*. Jyväskylän yliopisto. <http://urn.fi/URN:ISBN:978-951-39-8644-5>
- Kautonen, M. & A. von Zansen 2020. DigiTala research project: automatic speech recognition in assessing L2 speaking. *Kieli, koulutus ja yhteiskunta*, 11 (4). <https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-kesakuu-2020/digitala-research-project-automatic-speech-recognition-in-assessing-l2-speaking>
- Korhonen, A. & P. Mattila 2020. Laaja-alainen osaaminen osana opintojakson arviointia. Teoksessa Opetushallitus (toim.) *Uutta LOPSia rakentamassa*. <https://opetushallitus.mobiezone.fi/zine/9/article-1391>
- Kunnan, A. J. 2018. *Evaluating language assessments*. New York: Routledge.
- Lantolf, J. P. & M. E. Poehner 2014. *Sociocultural theory and the pedagogical imperative in L2 education*. New York: Routledge.
- Leontjev, D. 2022. Finnish Matriculation Examination, National Curriculum, and teachers' attitudes, perspectives, and practices: when the two assessment cultures meet. Teoksessa T. Seppälä, S. Lesonen, P. Iikkanen, & S. D'hondt (toim.) *Kieli, Muutos ja Yhteiskunta - language, change and society*. AFinLAN vuosikirja 2022. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 133–150.
- Leontjev, D. & P. Pollari 2022. Resolving tensions caused by high-stakes assessment in an L2 classroom through mediation. *Apples - Journal of Applied Language Studies*, 16 (1), 155–174.
- Leontjev, D. & P. Pollari 2023. Guiding and assessing development of L2 writing process: the role of peer collaboration. *Innovation in Language Learning and Teaching*, 17 (2), 440–452.

- Leung, C. 2022. Action-oriented plurilingual mediation: a search for fluid fountains. Teoksessa D. Little & N. Figueras (toim.) *Reflecting on the Common European Framework of Reference for Languages and its companion volume*. Luku 6. Multilingual Matters.
- Lopez, A., T. Sultan, & D. Guzman-Orth 2017. Assessing multilingual competence. Teoksessa E. Shohamy, I. Or & S. May (toim.) *Language testing and assessment. Encyclopedia of language and education*. Cham: Springer, 91–102.
- Messick, S. 1989. Validity. Teoksessa R. L. Linn (toim.) *Educational measurement*. New York: Macmillan, 13–103.
- Mustonen, S. & Honko, M. 2018. Monikielisyttä tukeva pedagogiikka. Teoksessa M. Honko & S. Mustonen (toim.) *Tunne kieli: matka maailman kieliin ja kielitietoisuuteen*. Helsinki: Finn Lectura, 118–141.
- Mäkipää, T. 2021. "I have never received feedback from my teachers." *Students' perceptions of assessment and feedback practices in Finnish general upper secondary education*. Helsingin yliopisto. <http://urn.fi/URN:ISBN:978-951-51-7644-8>
- Nguyen, P. 2021. Uses of language assessment. Teoksessa C. Chapelle (toim.) *The encyclopedia of applied linguistics*. Wiley. <https://doi.org/10.1002/9781405198431>
- 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, 5–23. <https://doi.org/10.37546/JALTSIG.CEFR4-1>
- OPH (Opetushallitus) 2011. *Eurooppalainen kielisalkku*. <http://kielisalkku.edu.fi/fi/opettajille/>
- OPH (Opetushallitus) 2014a. *Esiopetuksen opetussuunnitelman perusteet*. Määräykset ja ohjeet 2014: 1. 3. muutettu painos. Helsinki: Opetushallitus. [https://www.oph.fi/sites/default/files/documents/esiopetuksen\\_opetussuunnitelman\\_perusteet\\_2014.pdf](https://www.oph.fi/sites/default/files/documents/esiopetuksen_opetussuunnitelman_perusteet_2014.pdf)
- OPH (Opetushallitus) 2014b. *Perusopetuksen opetussuunnitelman perusteet*. Määräykset ja ohjeet 2014: 96. Helsinki: Opetushallitus. [https://www.oph.fi/sites/default/files/documents/perusopetuksen\\_opetussuunnitelman\\_perusteet\\_2014.pdf](https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf)
- OPH (Opetushallitus) 2017. *Kielitietoinen opetus - kielitietoinen koulu*. Helsinki: Opetushallitus. <https://www.oph.fi/fi/tilastot-ja-julkaisut/julkaisut/kielitietoinen-opetus-kielitietoinen-koulu>
- OPH (Opetushallitus) 2018. *Kielten eurooppalainen viitekehys uudistui. Monikielisyys, kulttuurinen osaaminen ja merkitysneuvottelut osaksi EVK:ta*. [https://www.oph.fi/sites/default/files/documents/cefr\\_companion-volume.pdf](https://www.oph.fi/sites/default/files/documents/cefr_companion-volume.pdf)
- OPH (Opetushallitus) 2019. *Lukiokoulutuksen opetussuunnitelman perusteet*. Määräykset ja ohjeet 2019: 2a. [https://www.oph.fi/sites/default/files/documents/lukion\\_opetussuunnitelman\\_perusteet\\_2019.pdf](https://www.oph.fi/sites/default/files/documents/lukion_opetussuunnitelman_perusteet_2019.pdf)
- OPH (Opetushallitus) 2020. *Perusopetuksen päättöarvioinnin kriteerit*. Opetushallituksen määräys OPH-5042-2020. <https://www.oph.fi/sites/default/files/documents/Perusopetuksen%20p%C3%A4%C3%A4tt%C3%B6arvioinnin%20kriteerit%2031.12.2020.pdf>
- OPH (Opetushallitus) 2021. *Kieliprofiili*. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/kieliprofiili>
- OPH (Opetushallitus) 2022. *Koulun monikielisydestä*. Helsinki: Opetushallitus. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/koulun-monikielisyudesta>
- Pellegrino, J. 2018. Assessment of and for learning. Teoksessa F. Fisher, C. E. Hmelo-Silver, S. R. Goldman, & P. Reimann (toim.) *International handbook of the learning sciences*. London: Routledge, 410–421.
- Piccardo, E., A. Germain-Rutherford, & G. Lawrence 2021 (toim.) *The Routledge handbook of plurilingual language education*. Lontoo: Routledge.
- Pietikäinen, S., M. Dufva, & K. Mäntylä 2010. Monikielisyys liikkeessä: monikielisen perheen kielelliset resurssit. Teoksessa M. Garant & M. Kinnunen (toim.) *AFinLA-e Soveltavan kielitieteen tutkimuksia* 2010 n:o 2, 17–30. <https://journal.fi/afinla/article/view/3873>



- Pitkänen-Huhta, A. & S. Pietikäinen 2013. Toiminnallisuus kielenoppimisessa – pedagogisia kokeiluja saamen luokissa. *Kieli, koulutus ja yhteiskunta*, 4 (4).  
<https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-lokakuu-2013/toiminnallisuus-kielenoppimisessa-pedagogisia-kokeiluja-saamen-luokissa>
- Poehner, M. E. 2008. *Dynamic assessment: a Vygotskian approach to understanding and promoting second language development*. Berliini: Springer.
- Poehner, M. E. & O. Inbar-Lourie (toim.) 2020. *Toward a reconceptualization of second language classroom assessment*. Cham: Springer.
- Poehner, M. E. & D. Leontjev 2020. To correct or to cooperate: mediational processes and L2 development. *Language Teaching Research*, 24 (3), 295–316.
- Richardson, U. & H. Lyytinen 2014. The GraphoGame method: the theoretical and methodological background of the technology-enhanced learning environment for learning to read. *Human Technology*, 10 (1), 39–60.  
<http://dx.doi.org/10.17011/ht/urn.201405281859>
- Saarinen, T. & A. Huhta 2023. The ideal learner as envisioned by Can Do statements and grammar revisions: how textbook agency is constructed. Teoksessa J. Ennsner-Kananen & T. Saarinen (toim.) *New materialist explorations into language education*. Cham: Springer, 151–174. <https://link.springer.com/book/10.1007/978-3-031-13847-8>
- Shermis, M. & J. Burstein (toim.) 2013. *Handbook of automated essay evaluation: current applications and future directions*. New York: Routledge.
- Shohamy, E. 2001. *The power of tests. a critical perspective on the uses of language tests*. London: Pearson Education.
- Shohamy, E. 2011. Assessing multilingual competencies: adopting construct valid assessment policies. *The Modern Language Journal*, 95 (3), 418–421.
- Shohamy, E. 2013. Expanding the construct of language testing with regards to language varieties and multilingualism. Teoksessa D. Tsagari, S. Papadima-Sophocleous, & S. Ioannou-Georgiou (toim.) *International experience in language testing and assessment: selected papers in memory of Pavlos Pavlou*. Language Testing and Evaluation Series. Vol. 28. Frankfurt am Main: Peter Lang, 17–32.
- Soveltavan kielentutkimuksen keskus 2022. *Dynamic diagnostic language assessment – conceptual and practical innovation in foreign language education and assessment*. <https://r.jyu.fi/DDLANG>
- Stathopoulou, M. 2020. The new CEFR descriptors for the assessment of written mediation: exploring their applicability in a local context in an effort towards multilingual testing. *CEFR Journal - Research and Practice*, 2, 40–78. <https://doi.org/10.37546/JALTSIG.CEFR2-3>
- Tarnanen, M. & A. Huhta 2011. Foreign language assessment practices in the comprehensive school in Finland. Teoksessa D. Tsagari & I. Csepes (toim.) *Classroom-based language assessment*. Language Testing and Evaluation Series. Vol. 25. Frankfurt am Main: Peter Lang, 129–146.
- Taylor, L. 2013. Communicating the theory, practice and principles of language testing to test stakeholders: some reflections. *Language Testing*, 30 (3), 403–412.
- Tsagari, D. & K. Vogt 2017. Assessment literacy of foreign language teachers around Europe: research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6 (1), 41–63. [http://www.altanz.org/uploads/5/9/0/8/5908292/5\\_si3tsagarivogt\\_final\\_formatted\\_proofed.pdf](http://www.altanz.org/uploads/5/9/0/8/5908292/5_si3tsagarivogt_final_formatted_proofed.pdf)
- Turun yliopisto 2020. *KieliPeda-työväline*.  
[https://www.utu.fi/sites/default/files/media/Kielipeda\\_tyovaline.pdf](https://www.utu.fi/sites/default/files/media/Kielipeda_tyovaline.pdf)
- Vygotsky L. S. 1978. *Mind in society: the development of higher psychological processes*. Cambridge, MA: Harvard University Press.



- Väljärvi, J., N. Huotari, P. Iivonen, M. Kulp, T. Lehtonen, H. Rönholm, G. Knubb-Manninen, J. Mehtäläinen, & S. Ohranen 2009. *Lukiopedagogiikka*. Koulutuksen Arviointineuvoston Julkaisuja, 40. <http://urn.fi/URN:ISBN:978-951-39-3731-7>
- Wei, L. 2014. Negotiating funds of knowledge and symbolic competence in the complementary school classrooms. *Language and Education*, 28 (2), 161–180. <https://doi.org/10.1080/09500782.2013.800549>
- Weir, C. J. 2005. *Language testing and validation*. Hampshire: Palgrave MacMillan.
- William, D. 2017. Learning and assessment: a long and winding road? *Assessment in Education: Principles, Policy & Practice*, 24 (3), 309–316.
- William, D. & S. Leahy 2015. *Embedding formative assessment: practical techniques for K-12 classrooms*. West Palm Beach, FL: Learning Sciences International.
- Xu, Y. & G. Brown 2016. Teacher assessment literacy in practice: a reconceptualization. *Teaching and Teacher Education*, 58, 149–162.
- Zechner, K. & K. Evanini 2020. *Automated speaking assessment: Using language technologies to score spontaneous speech*. New York: Routledge.