

JYX



JYVÄSKYLÄN YLIOPISTO
UNIVERSITY OF JYVÄSKYLÄ

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

Author(s):

Title: Kielenoppimista tukeva arviointi

Year: 2023

Version: Published version

Copyright: © Suomen soveltavan kielitieteen yhdistys 2023

Rights: In Copyright

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

Please cite the original version:

Mäkipää, T., Hilden, R., & Huhta, A. (Eds.). (2023). Kielenoppimista tukeva arviointi. Suomen soveltavan kielitieteen yhdistys ry. AFinLA-teema, 15. <https://journal.fi/afinla/issue/view/9173>

Toimittaneet

Toni Mäkipää, Raili Hilden ja Ari Huhta

KIELENOPPIMISTA TUKEVA ARVIOINTI

ASSESSMENT FOR SUPPORTING LANGUAGE LEARNING

AFinLA-teema 2023 / n:o 15

ISBN: 978-951-9388-75-5
ISSN: 2814-6298

AFinLA-teema 2023 / n:o 15

Kielenoppimista tukeva arviointi

Assessment for supporting language learning

Toimittaneet

Toni Mäkipää, Raili Hilden ja Ari Huhta



Suomen soveltavan kielitieteen yhdistys AFinLA, Jyväskylä 2023
© kirjoittajat ja Suomen soveltavan kielitieteen yhdistys AFinLA

Sisältö

Johdanto

- 📖 *Toni Mäkipää, Raili Hilden & Ari Huhta*
Oppimista tukeva arviointi kieltenopetuksessa..... 5
[Assessment for supporting learning in language teaching]

Artikkelit

- 📖 *Toni Mäkipää & Olli-Pekka Salo*
Using self-assessment to enhance self-regulated learning –
an intervention study in a Finnish lower secondary school19
- 📖 *Päivi Virkkunen & Minnaleena Toivola*
Lukio-opiskelijoiden käsityksiä vieraan kielen
ääntämisen opetuksessa saadusta palautteesta40
[High school students' perceptions of feedback on foreign language pronunciation]
- 📖 *Pirjo Pollari & Dmitri Leontjev*
Beyond error correction in EFL writing in a Finnish
upper secondary classroom: a practical approach60
- 📖 *Magdalini Liontou*
How do Finnish and Chinese students' diverse pedagogical
experiences shape feedback interpretation?81
- 📖 *Anssi Roiha, Juuso Nieminen & Katja Mäntylä*
Lukion vieraiden kielten opettajien käytänteitä
ja käsityksiä eriyttävästä arvioinnista.....97
[Foreign language teachers' practices and perceptions of differentiated assessment in
upper secondary schools]

- ❏ *Anna von Zansen & Maikki Heijala*
 Miten suomen ja ruotsin opettajat käyttäisivät puheen
 automaattiseen arviointiin kehitettyä työkalua? 124
 [How would Finnish and Swedish teachers use a tool developed for automated
 L2 speaking assessment?]
- ❏ *Elli Saari & Raili Hilden*
 S2-opettajien käsityksiä formatiivisesta
 arvioinnista oppimisen tukena 142
 [Finnish as a second language teachers' perceptions of supporting learning with formative assessment]
- ❏ *Päivi Laine, Minna Maijala & Katja Mäntylä*
 Kestävän kehityksen periaatteet kielenoppimisen arvioinnissa 162
 [The principles of sustainable development in assessing language learning]
- ❏ *Merja Kauppinen, Mirja Tarnanen, Reeta Neittaanmäki, Eija Aalto, Mari Hankala,
 Johanna Kainulainen, Sanna Mustonen & Päivi Torvelainen*
 Diagnostinen arviointi suomen kielen ja kirjallisuuden
 pedagogiikan opintojakson opiskelua ja opetusta ohjaamassa 184
 [Diagnostic assessment guiding the study and teaching of a Finnish language and
 literature pedagogy course]

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 5–18.

Toni Mäkipää
Helsingin yliopisto

Raili Hilden
Helsingin yliopisto

Ari Huhta
Jyväskylän yliopisto

Oppimista tukeva arviointi kielenopetuksessa

Nostot

- Formatiivisella ja dynaamisella arvioinnilla voidaan tukea kielenoppimista monipuolisilla tavoilla.
- Oppimisen tukeminen arvioinnin avulla on keskeinen osa arviointiosaamista.
- Tyypillisiä formatiivisen arvioinnin muotoja kielenopetuksessa ovat palaute, itsearviointi ja vertaispalaute.

Abstract

The aim of this special issue is to present current research on assessment that supports language learning. The aim of all assessment is to stimulate learning although this customarily applies to formative and dynamic assessment. By using multifaceted assessment practices, language teachers can support students' learning in myriad ways. This is also a key aspect of assessment literacy. However, little is known about assessment that supports language learning in the Finnish context. Therefore, this special issue provides an overview of how assessment can be used to enhance language learning in Finland. The publication starts with an introductory article on formative and dynamic assessment.

Keywords: formative assessment, dynamic assessment, assessment literacy

Asiasanat: formatiivinen arviointi, dynaaminen arviointi, arviointiosaaminen

1 Johdanto

Monipuolisen arvioinnin merkitys korostuu opetussuunnitelman perusteissa (Opetushallitus 2014, 2019) sekä perusopetus- (1998/628) ja lukiolaissa (2018/714). Toisin sanoen arvioinnin ei tule koostua pelkästään kokeista ja muista summatiivisen arvioinnin muodoista, vaan sen rinnalle tarvitaan oppimista tukevaa, formatiivista arviointia. Oppimista voidaan tukea esimerkiksi palautteen, itsearvioinnin ja vertaispalautteen avulla. Nämä tyypilliset formatiivisen arvioinnin välineet ja niiden implementointi opetukseen mainitaan opetussuunnitelman perusteissa (Opetushallitus 2014, 2019). Formatiivisen arvioinnin avulla voidaan lisäksi tukea oppilaiden oppimaan oppimista ja reflektointia, mikä on yksi tärkeimmistä tavoitteista suomalaisessa koulutuksessa (Opetushallitus 2014, 2019). Itsearviointi ja reflektointi mainitaan myös lakitasolla olennaisiksi oppimistavoitteiksi (Lukiolaki 2018/714; Perusopetuslaki 1998/628).

AFinLAN aikaisempi teemajulkaisu arvioinnista käsitteli kielitaidon arviointitutkimusta 2000-luvun Suomessa (Huhta & Hilden 2016). Siinä käsiteltiin summatiivista arviointia. Siksi tässä uudessa teemajulkaisussa on tärkeää pohtia, kuinka formatiivista ja dynaamista arviointia voidaan hyödyntää kieltenopetuksessa tukemaan oppimista ja kielitaidon karttumista. Ensin esittelemme teoriataustan, joka nivoo yhteen teemajulkaisumme yhdeksän artikkelia.

2 Opettajien arviointiosaaminen

Formatiivinen arviointi on osa opettajan arviointiosaamista, joka taas on osa opettajan ammatillista osaamista ja päätöksentekoa. Keskeisiä opettajan ammatillisen kompetenssin rakennuspuita ovat kuvanneet monet kasvatustieteen ja soveltavan kielitieteen tutkijat (Schulman 1987; Caena 2011; Gitomer & Zisk 2015; Borg 2015). Kaikkien kuvausten ytimessä on opetettavan aineen asiantuntijuus yhdistyneenä

yleiseen pedagogiseen tietoon ja taitoon sekä opettajan kykyyn kehittää työtään ja toimijuuttaan reflektion avulla. Sittemmin mukaan on liitetty myös teknologinen kompetenssi (TPACK) (Koehler ym. 2013). Myös arviointiosaamisen teoreettisia malleja on viime vuosikymmenten aikana syntynyt lukuisia. Tässä esittelemme niistä tunnetuimmat ja tutkimuskirjallisuudessa viitatuimmat.

Arviointi nähtiin tutkimuksessa pitkään summatiivisen osa-alueen ja ulkoisten tutkintojen näkökulmasta, eikä opettajia pidetty oleellisina toimijoina muutoin kuin korkeintaan testitulosten tuottamisen edistäjinä. Opetustyön tarpeisiin arviointia analysoi ensimmäisten joukossa Scriven (1967), joka erotti toisistaan summatiivisen ja formatiivisen arvioinnin käsitteet. Pian tämän jälkeen opetukseen nivoutuva formatiivinen arviointi nousi voimakkaasti didaktisen tutkimuksen valokeilaan haastamaan opetuksen ulkopuolista arviointia (Black & Wiliam 1998, 2009; Inbar-Lourie 2008). Sosiokonstruktivismi ja sosiokulttuurinen oppimismäkemys vauhdittivat tätä kehitystä. Samalla formatiivisen arvioinnin käsite hienojakoistui, ja opettajan kompetensseissa korostui taito panna toimeen formatiivisia käytänteitä.

Arvioinnin tavoitteen näkökulmasta formatiivisen arvioinnin perusoletus ja määritelmällinen ydin on, että se sulautuu opetukseen (assessment as learning) ja opiskeluun ja edistää oppimista (assessment for learning), kun taas summatiivinen arviointi perinteisimmillään suuntautuu oppimistuloksiin sivuuttaen niiden saavuttamisprosessin (Cizek ym. 2019). Formatiivista arviointia rasittaa terminologinen samankaltaisuus summatiivisen arvioinnin kanssa. Arviointia terminä sävyttää useimmissa kielissä summatiiviseen arviointiin liittyvä seurauksellisuuden paino. Siksi formatiivisten käytänteiden kuvauksessa on paljolti siirrytty käyttämään palautteen käsitettä ja vastaavaa termiä, joka keventää arvioinnin lopullisuutta. Palautehan on aina kuulunut hyvään opetukseen, ja sen vaikutus oppimistuloksiin on todistettu eri konteksteissa (Hattie 2012; Suomessa Mäkipää 2021a).

Toisaalta käsitteellinen rajankäynti formatiivisen arvioinnin ja palautteen välillä on tuonut uusia pulmia tieteelliseen analyysiin. Vaikka palaute-termillä haetaan etäisyyttä arvottamiseen, ei arvottaminen ole kokonaan suljettavissa pois. Tehokas palaute perustuu tavoitteeseen (Where am I going?), johon oppilaan osaamista verrataan. Tuo osaaminen on todennettava ja arvotettava suhteessa tavoitteeseen. Vasta kun näiden kahden etäisyys on todettu, voidaan antaa neuvoja siihen, mikä on seuraava askel lähemmäksi tavoitetta (Where to next?) (Hattie 2012).

Formatiivisella arvioinnilla on nykyisin tärkeä sija lähes kaikissa opettajan arviointiosaamisen malleissa. Aluksi mallit olivat lähinnä luetteloita opettajan toimenpiteistä, mutta ne syvenivät vähitellen kattamaan kolme pääaluetta: tiedon, taidot ja arvoperustan. Brookhart (2011) katsoi arviointiosaamisen lähtökohdaksi opetuksellista päätöksentekoa tukevien arviointimenetelmien valinnan ja työstämisen, johon kuuluu suorituksen tulosten tulkinta. Tuloksia käytetään opetuksen suunnitteluun ja oppilaiden osaamisen määrittelyyn, ja niistä tiedotetaan oppilaille, huoltajille ja muille asianosaisille. Arviointiosaamiseen kuuluu myös epäeettisten toimintatapojen

tunnistaminen ja välttäminen. Arviointiosaamisen käsitettä laajensi Fulcher (2012) sisällyttäen määritelmään tietojen ja taitojen ohella myös periaatteellisen tason, joka näitä ohjaa. Lisäksi malli huomioi arvioinnin historiallisen ja yhteiskunnallisen kontekstin sekä filosofiset perusteet.

On tiedostettu kauan, että arviointi koskettaa monia eri toimijoita ja ihmisryhmiä (stakeholders) eri tavoin, ja nämä tarvitsevat eri määriä erilaista osaamista. Taylor (2013) määritteli kahdeksan osaamisaluetta ja profiilia eri toimijoille, joiden tarve hallita nämä alueet vaihtelee. Profiilin komponentit ovat teoretieto, käytännön taidot, periaatteet ja käsitteet, kieltenopetus, sosiokulttuuriset arvot, paikalliset käytännöt, henkilökohtaiset uskomukset ja asenteet sekä arvostelu ja päätöksenteko. Opettajien profiilissa painottuu pedagoginen osaaminen (kieltenopetus), ja sen ohella käytännön taidot, sosiokulttuuriset arvot ja henkilökohtaiset uskomukset. Sen sijaan periaatteiden ja teorian tuntemus ja arvostelu ja päätöksenteko eivät ole yhtä keskeisiä kuin ne ovat kokeita ammatikseen laativille tai arviointitutkijoille. Taylorin mallia on sovellettu laajasti arviointitutkimuksessa (esim. Vogt & Tsagari 2014; Kremmel & Harding 2020; Vogt ym. 2020) erityisesti opettajien kokemien täydennyskoulutustarpeiden kartoitukseen, mutta käsitteellinen hyöty on jäänyt jälkeen käytännöllisestä.

Käsitteellisesti edistyneimmät opettajan arviointiosaamisen mallit nivovat yhteen perinteiset perusasiat ja kuvaavat lisäksi niiden välisiä prosesseja ja toimintakontekstin piirteitä. Xun ja Brownin (2016) teoriamalli rakentuu tietoperustalle, johon kuuluvat opettajan aineosaaminen ja pedagoginen sisältötieto, arviointitieto, arvosanan antamiseen liittyvä tieto, palautetieto, tieto vertais- ja itsearviointista, arviointitulosten tulkintaan ja niistä viestimiseen liittyvä tieto sekä tieto arvioinnin etiikasta. Näiden pohjalle rakentuvat työyhteisön tulkintakehyksessä kolme opettajan käsitysnäkökulmaa eli kognitiivinen, affektiivinen ja opetus- ja tietokäsitykset. Nämä siirtyvät opettajan käytännön päätöksentekoon ja toimenpiteisiin, joihin vaikuttavat opettajan sosiokulttuurinen toimintaympäristö ja koulukonteksti.

Käytännön opetustyössä opettaja joutuu tekemään kompromisseja, mutta hänen arviointiosaamisensa ja arvioijaidentiteettinsä kehittyy ja muokkautuu pohdinnan kautta. Arvioijaidentiteetin abstraktia ja dynaamista luonnetta vastakohtana empiristisille ja toimenpidesuuntautuneille näkemyksille korostavat Looney ym. (2018). Arvioijaidentiteetti liittyy opettajan minäpystyvyyteen ja nojaa erityisesti opettajan omiin käsityksiin itsestään arvioijana suhteessa omaan tietoon, rooliin toimijana, itseluottamukseen, uskomuksiin ja tunteisiin (Looney ym. 2018).

Pastore ja Andrade (2019) katsovat Xun ja Brownin (2016) tapaan, että opettajan arviointiosaaminen toteutuu koulutuspoliittisen kontekstin ja kouluympäristön ehdoilla. Itse arviointiosaamisen käsitteeseen he lukevat kolme ulottuvuutta: käsitteellisen tiedon (conceptual knowledge dimension), käytännön arviointitaidot (praxeological dimension) ja sosioemotionaalisen ulottuvuuden (socio-emotional dimension). Sosiokonstruktiivisen opetus- ja oppimisen näkemyksen mukaisesti nämä tiedot, taidot ja uskomukset ovat alati muuttuvia ja kehittyviä. Pastoren ja Andraden malli yksilöi

melko tarkasti eri osaamisalueita ja niiden toimeenpanostrategioita, minkä vuoksi se on hyödyllinen esimerkiksi täydennyskoulutuksen ja käytännöllisesti suuntautuneen tutkimuksen tarpeisiin. Käsitteellisesti se on Xun ja Brownin mallia pelkistetympi.

Ensi sijassa arviointitutkimuksen palvelukseen on tarkoitettu Herppichin ym. (2017) malli, jonka lähtökohtana on opettajan kognitiivinen valmius, joka ammentaa opettajan ammattiosaamisesta ja muista kompetensseista. Toisin kuin edellä esitellyissä malleissa arviointiosaaminen ei sellaisenaan ole opettajakompetenssin osa, vaan yleistä pedagogista osaamista ja pedagogista sisältötietoa sovelletaan arviointitilanteisiin ja niissä tapahtuvaan päätöksentekoon.

3 Oppimista tukevan arvioinnin tapoja

Ajallisesti formatiivinen arviointi edeltää summatiivista. Koska formatiivinen arviointi sisältyy opetukseen, sen muodot mukailevat opetuksen ja opiskelun tapojen moninaista kirjoa. Formatiiiviset käytänteet voidaan yksinkertaisimmin jakaa toimijoiden mukaan opettajan, vertaisten tai opiskelijan itsensä suorittamaan arviointiin. Formatiiivinen arviointi ei ole kuitenkaan synonyymi palautteelle, itsearviointille tai vertaispalautteelle. Jotta formatiiivisella arvioinnilla voidaan tukea oppimista, opettajan tulee ilmoittaa selkeästi oppimistavoitteet, selventää käytössä olevat arviointikriteerit ja suunnitella monipuolisia tehtäviä, joilla oppilaat pääsevät näyttämään osaamistaan ja ymmärtämistään (Black & Wiliam 2018).

Jokaisella oppilaalla on oikeus saada opettajalta seikkaperäistä ja yksilöityä palautetta, joka tukee hänen oppimistaan (Opetushallitus 2014, 2019). Palautteen merkitys oppimiselle on ilmeinen: sen avulla oppilaalle tehdään näkyväksi, missä vaiheessa hän on oppimistaan, mihin hän on suuntaamassa ja mitä pitäisi tehdä, jotta hän saavuttaa tavoitteensa (Hattie & Timperley 2007). Useissa tutkimuksissa on tarkasteltu, millainen palaute tukee oppimista parhaiten. Konsensusta tästä ei ole, mutta palautetta on hyvä tarkastella sisällön ja strategioiden näkökulmasta. Oppimista tukeva palaute on selkeää, yksilöityä ja tarkkaa (Ruiz-Primo & Brookhart 2018; Shute 2008). Palautteen sävyn tulee olla positiivinen, ja vertailukohtaksi palautteessa on hyvä ottaa arviointikriteerit ja kurssin tavoitteet (Brookhart 2017).

Strategioiden näkökulmasta on olennaista huomata ajoitus, määrä, muoto ja yksilö. Palaute tehtävästä tai suorituksesta pitää antaa silloin, kun niiden tekeminen ja muokkaaminen on vielä ajankohtaista oppilaalle, ja palautetta tulee antaa monipuolisesti eri tavoilla mieluiten yksilöllisesti ryhmäpalautteen sijaan (Ruiz-Primo & Brookhart 2018). Palautteen määrässä on hyvä valita suorituksen keskeisimmät seikat, joihin opettaja paneutuu kommentteissaan ja huomioida sekä suorituksen heikkoudet että vahvuudet (Brookhart 2017). Toisin sanoen palautteessa ei ole tarkoitus pelkästään listata suorituksen virheitä ja puutteita, vaan myös tuoda esiin sen hyvät puolet sekä kannustaa oppilasta (Black & Wiliam 2018).

On syytä huomioida, ettei yksittäinen palautteen aspekti automaattisesti takaa oppimisen edistämistä. Toinen huomionarvoinen seikka on, ettei pelkkä palaute (feedback) oppilaan aiemmista suorituksista ole riittävää oppimisen edistämiseksi. Oppilas tarvitsee lisäksi tietoa siitä, miten hän pääsee eteenpäin oppimisprosessissaan kohti tavoitteiden saavuttamista (feedforward) (Black & Wiliam 2018; Hattie & Timperley 2007).

Aiemmissa määritelmissä palaute nähtiin lahjana opettajalta oppilaalle (Askew & Lodge 2000), mutta nykyisen tutkimuksen mukaan myös oppilaan tulee olla aktiivinen palauteprosessissa. Tätä kutsutaan palauteosaamiseksi (Carless & Boud 2018). Palauteosaava oppilas ymmärtää palautteen merkityksen oppimisessa, aktiivisesti pyytää palautetta muilta, osaa säädellä tunteitaan ja hyödyntää sekä itsearviointia että vertaispalautetta monipuolisesti (Carless & Boud 2018; Molloy ym. 2020). On siis olennaista, että kieltenopettaja edistää oppilaiden palauteosaamisen kehittymistä oman palautteensa ja opetuksensa avulla. Toisin sanoen yksittäiset palautteen piirteet eivät ole tärkein ulottuvuus palauteprosessissa. Sen sijaan keskeiset kysymykset ovat: mitä oppilas tekee palautteen avulla ja miten hän käsittelee sitä (Carless & Boud 2018; Molloy ym. 2020).

Usein palautetta antaa opettaja, mutta sitä voi antaa myös vertainen, kuten luokkatoveri. Tällaista palautetta kutsutaan vertaispalautteeksi. Se voi olla joko summatiivista tai formatiivista riippuen sen käyttötarkoituksesta (Topping 2018). Kuitenkin Jackson ja Marks (2016) suosittelevat, että vertaispalautteessa keskitytään enemmän reflektoinnin kehittämiseen kuin arvosanojen antamiseen. Vertaispalautteen hyödyntäminen opetuksessa on tärkeää, sillä sen avulla voidaan tukea sekä itseohjautuvuutta että kriittisen ajattelun kehittymistä (Nawas 2020; Nicol & Macfarlane-Dick 2006). Lisäksi vertaispalautteessa korostuu oppilaan aktiivinen rooli oppimisessa (Klapp 2015).

Vertaispalautteen antaminen voi olla kuitenkin haastavaa monille oppilaille (Ng & Yu 2021), minkä vuoksi he tarvitsevat opetusta siitä, miten vertaispalautetta annetaan (Li ym. 2020). Mikäli opettaja opettaa perusteellisesti, miten vertaispalautetta annetaan luokkatoverille, opettaja edistää samalla oppilaiden myönteistä suhtautumista vertaispalautteeseen (van Zundert ym. 2010). Myös palaute vertaispalautteista saa oppilaat suhtautumaan vertaispalautteeseen vakavasti (Ng & Yu 2021). Opettajan on hyvä myös panostaa luokassa sallivaan ilmapiiriin, jossa virheitä ei nähdä uhkana vaan oppimismahdollisuutena (Klapp 2015).

Vertaispalautteessa oppilas antaa palautetta toiselle oppilaalle, mutta palautetta voidaan antaa myös itselle. Tällöin on kyseessä itsearviointi. Kaiken oppimisen tavoitteena on itseohjautuvuus (Bandura 1993), jonka tärkeä ulottuvuus on kyky arvioida omaa osaamistaan eli itsearvioida (Nicol & Macfarlane-Dick 2006). Mikäli oppilas ei kykene itsearviointiin, hänestä ei voi tulla itseohjautuvaa (Raaijmakers ym. 2018). Itsearvioinnilla on myös muita hyötyjä: se edistää motivaatiota ja tietoisuutta omasta suorituksesta (Dlaska & Krekeler 2008) sekä lisää tietoisuutta omista heikkouksista ja vahvuuksista (Isbell & Lee 2022). Niinpä opettajan tulee hyödyntää itsearviointia

säännöllisesti opetuksessa, jotta hän tukee oppilaiden kykyä itsearviointiin. Opettaja voi esimerkiksi käyttää osan oppitunnista itsearviointitaitojen opettamiseen ja kohtamiseen (Andrade 2019).

Itsearviointi perustuu kriteereihin työn tai tehtävän sisällöstä, ja opettajan on hyvä käydä läpi kriteerit, jotta oppilaat ymmärtävät niiden sisällön (Brooks 2002). Itsearvioinnin tekeminen on haastavaa monelle oppilaalle (Dlaska & Krekeler 2008), sillä erityisesti heikot oppilaat voivat yliarvioida suorituksiaan ja vastaavasti lahjakkaat oppilaat voivat aliarvioida niitä (Isbell & Lee 2022). Suomalaisen tutkimuksen mukaan itsearviointi ei ole yleinen arviointitapa vieraiden kielten opetuksessa (Mäkipää 2021b), minkä vuoksi itsearvioinnin hyödyntämiseen pitäisikin kiinnittää enemmän huomiota.

Yksi oppimista tukevan arvioinnin lupaava muoto on dynaamiseksi arvioinniksi kutsuttu lähestymistapa arviointiin. Tämä Suomessa toistaiseksi vähän käytetty arviointitapa rakentuu erilaiselle epistemologiselle (tietoteoreettiselle) pohjalle kuin edellä kuvatut lähestymistavat oppimista tukevaan arviointiin. Dynaaminen arviointi edustaa lähtökohdiltaan dialektista materialismia, jonka mukaan tietoa ei tule pitää pysyvänä ja muuttumattomana ja joka perustuu teoriatasolla venäläisen kehityspsykologin Lev Vygotskyn jo 1930-luvulla kehittämään sosiokulttuuriseen teoriaan oppimisesta (Vygotsky 1978; dynaamisen arvioinnin perusteista ks. tarkemmin esim. Poehner 2017). Sosiokulttuurisen teorian mukaan lapsen ja yleensäkin ihmisen kehittyminen on prosessi, jossa oppija omaksuu ympäröivän yhteiskunnan arvot, asenteet ja ongelmanratkaisustrategiat vuorovaikutuksessa häntä osaavampien (esim. vanhemmat, opettajat) kanssa. Nämä ohjaavat oppimista askel askeleelta oppijaa tukien; tätä tukea kutsutaan sosiokulttuurisessa lähestymistavassa mediaatioksi.

Keskeinen käsite oppimisen kannalta tässä teoriassa on lähikehityksen vyöhyke (eng. zone of proximal development; ks. esim. Poehner 2008). Tämä vyöhyke kuvaa aluetta, jolla oppija pystyy suoriutumaan opittavista asioista vain toisen henkilön (esim. opettajan) tuella. Vyöhyke sijoittuu siis sellaisten asioiden väliin, jotka oppija osaa itsenäisesti ja joita hän ei osaa edes tuettuna. Sosiokulttuurisen näkemyksen mukaan oppiminen on tehokkainta, kun se tapahtuu lähikehityksen vyöhykkeellä (Poehner 2008). Vyöhykkeen 'pituus' tai 'leveys' ei ole vakio, vaan se riippuu opittavista asioista, oppijan ominaisuuksista ja luultavasti myös mediaation laadusta.

Dynaamista arviointia on länsimaissa käytetty lähinnä psykologiassa ja erityispedagogiikassa (ks. esim. Sternberg & Grigorenko 2002), mutta lähestymistapa on levinnyt jonkin verran myös kielenopetuksen ja -arvioinnin alueelle. Tunnetuimpia sosiokulttuurista teoriaa ja erityisesti dynaamista arviointia kielikoulutukseen soveltaneita tutkijoita ovat amerikkalaiset James Lantolf ja Matthew Poehner (esim. Lantolf & Poehner 2011; 2014; Poehner 2008). Suomessakin on dynaamista kielitaidon arviointia ryhdytty tutkimaan (Leontjev 2016, 2022), mutta sen käyttö ei ole vielä levinnyt kielenopetukseen joitakin mahdollisia kokeiluja lukuun ottamatta.

Vuoden 2022 lopulla Jyväskylän yliopistossa alkoi DD-LANG-tutkimushanke¹, joka selvittää dynaamisen ja diagnostisen arvioinnin yhdistämistä kouluttaen samalla tutkimukseen osallistuvien lukioiden opettajia soveltamaan dynaamista arviointia omassa opetuksessaan. Hanke toimii siis mahdollisena päänavaajana laajemmallekin koulutukselle dynaamisesta arvioinnista. Tällaista koulutusta aloitettiin jo osana laajempaa Opetushallituksen rahoittamaa ARVO-koulutusta (Huhta ym. 2021).

Dynaaminen arviointi voidaan jakaa interaktiiviseen ja interventionistiseen arviointiin (esim. Lantolf & Poehner 2014). Interaktiivinen arviointi on huomattavasti tavallisempi muoto, ja se tapahtuu tyypillisesti arvioitavan ja arvioijan (esim. opettajan) välisessä keskustelussa, jossa arvioija pyrkii selvittämään oppijan osaamista. Kun jokin ongelma ilmenee – oppija tekee esimerkiksi jonkin virheen tai ei pysty suoriutumaan tehtävästä – arvioija pyrkii auttamaan oppijaa, aluksi erilaisilla hyvin yleisillä kysymyksillä tai vinkeillä, joita hän sitten tarkentaa, jos oppija ei niiden avulla osaa korjata suoritustaan tai löytää ratkaisua. Tällainen asteittain tarkentuva ja eksplisiittisemmäksi muuttuva tuki, mediaatio, onkin keskeinen dynaamisen arvioinnin piirre. Kyse ei kuitenkaan ole vain arvioinnista, sillä dynaaminen mediaatioprosessi yhdistää arvioinnin, opetuksen, opiskelun ja oppimisen: tavoitteena on nimittäin, että oppija prosessin aikana myös oppisi ongelmalliseksi osoittautuneen asian. Dynaamisen arviointitapahtuman tuloksena opettaja saa tietoa oppijan ongelmasta ja voi suunnitella opetustaan ja tarvittavia tehtäviä sen pohjalta. Eihän yksi arviointikerta välttämättä johda asian oppimiseen (ainakaan pysyvämmiin), vaikka kyse olisikin dynaamisesta arvioinnista.

Toinen dynaamisen arvioinnin muoto, interventionistinen, nojaa etukäteen hyvin tarkasti määriteltyihin tapoihin antaa oppijalle palautetta, vinkkejä ja tukea arvioinnin aikana (esim. Poehner & Lantolf 2014). Interaktionistisessa arvioinnissa arvioija voi melko vapaasti päättää, millaista tukea hän antaa oppijalle. Ovathan opetus-/oppimiskeskustelut monesti hyvin ennakoimattomia, joten mediaatiota on vaikea etukäteen tarkasti suunnitella. Interventionistista dynaamista arviointia voi yrittää toteuttaa opettajan ja oppijan välisessäkin keskustelussa, mutta tavallisempaa se on yrityksissä luoda tietokonepohjaisia dynaamisia testejä. Tyypilliset dynaamiset tietokonetestit ovat monivalintakysymyksiin perustuvia testejä, joissa tietyt väärät vastauksen johtavat etukäteen suunniteltuihin palautteisiin. Tunnetuin dynaaminen vieraan kielen testi on Lantolfin ja Poehnerin Pennsylvania State -yliopistossa kehittämä puheen ymmärtämisen testi (Poehner & Lantolf 2013). DD-LANG-hankkeessa on myös tarkoitus käyttää verkossa olevia tietotekniikan avulla toteutettuja vieraan kielen lukemis- ja kirjoitustehtäviä.

1 (<https://r.jyu.fi/DDLANG>)

4 Teemajulkaisun sisältö

Tässä teemajulkaisussa on yhdeksän artikkelia, joissa käsitellään moniulotteisesti oppimista tukevaa arviointia kielenopetuksessa. Teemajulkaisun aloittaa Mäkipään ja Salon interventiotutkimus, jossa selvitettiin, vaikuttaako itsearviointia erilaisin tavoin kielenopiskelussa korostava toiminta oppilaiden käsityksiin itseohjautuvuudesta ja itsearvioinnista. Tutkimukseen osallistui 36 yhdeksäsluokkalaista ruotsin opiskelijaa. Aineisto koostui heidän vastauksistaan itseohjautuvuutta kartoittaviin Likert-asteikollisiin väittämiin sekä avoimista vastauksista ruotsiksi puhumisen vahvuuksia ja haasteita koskeviin itsearviointikysymyksiin. Oppilaiden itsearvioinnin laadun havaittiin paranevan intervention aikana, vaikka osaamisen reflektointi jäikin varsin pintapuoliseksi. Myös heidän näkemyksensä itseohjautuvuudesta muuttuivat positiivisemmiksi. Tulokset osoittivat siis, että varsin lyhytkin itsearvioinnin käsittely opetuksessa voi muokata oppilaiden käsityksiä itseohjautuvuudesta myönteisemmiksi ja samalla parantaa heidän itsearviointitaitoaan.

Virkkunen ja Toivola tarkastelivat artikkelissaan lukiolaisten käsityksiä ääntämisestä ja siitä saadusta palautteesta. Aineisto kerättiin kyselyllä, johon vastasi 1953 lukiolaista. Tulosten mukaan lukiolaiset arvostivat hyvää ääntämistä ja toivoivat saavansa siitä opettajalta erityisesti korjaavaa palautetta. Lopuksi Virkkunen ja Toivola toteavat, että ääntämiseen ja siitä saatavaan palautteeseen tulee kiinnittää huomiota laadukkaan kielenopetuksen edistämiseksi.

Pollarin ja Leontjevin artikkeli on suunnattu erityisesti kielenopettajille, mutta se sisältää tutkijoillekin mielenkiintoista tietoa formatiivisesta arvioinnista. Artikkelin on yhteenveto molempien tutkijoiden lukion englannin opetukseen sijoittuvista tutkimuksista, joista yksi tarkasteli arviointia sosiokulttuurisen teorian ja mediaation näkökulmasta. Tutkijat kehittivät formatiivisia arviointitapoja, joilla pyrittiin vaikuttamaan oppimisprosessiin produktin (esim. kirjoitelman) sijasta tai ohella esimerkiksi antamalla muutakin kuin korjaavaa palautetta. Vertaisarviointi ylioppilastutkintoon perustuvan tarkistuslistan avulla oli yksi keskeisistä työskentelytavoista. Tulosten mukaan on tärkeää antaa opiskelijoiden päättää, millaista palautetta he haluavat saada, mutta on sekä tärkeää että mahdollista ohjata heitä hyödyntämään muutakin kuin korjaavaa palautetta, esimerkiksi kirjoitusprosessia koskevaa palautetta.

Liontoun artikkelissa vertaillaan suomalaisten ja kiinalaisten yliopisto-opiskelijoiden käsityksiä opettajan palautteesta ja vertaisarvioinnista. Aineisto kerättiin haastatteluilla. Tulosten perusteella Liontoun toteaa, että opiskelijat kokivat opettajan antaman palautteen erittäin tärkeäksi palautemuodoksi. Kuitenkaan vertaispalautetta ei pidetty yhtä hyödyllisenä kuin opettajan antamaa palautetta. Lisäksi kiinalaiset opiskelijat pitivät tärkeänä sitä, että heidän vanhempansa ovat tietoisia arvosanoista ja palautteista, mutta suomalaiset opiskelijat ovat päinvastaista mieltä. Lopuksi Liontoun toteaa, että tarvitaan lisää tutkimusta, jotta saadaan kartoitettua tarkemmin palautteen merkitys ja käyttö eri koulutusjärjestelmissä.

Roihan, Niemisen ja Mäntylän tutkimus paneutuu lukion kieltenopettajien (n=48) käsityksiin eriytyvästä arvioinnista. Kyselyin ja haastatteluin toteutettu kartoitus antaa uutta tietoa opettajien kokemuksista ja käytänteistä ja arviointiosaamisen ihanteiden mukautumisesta toimintaympäristön ehtoihin. Opettajat tunsivat monia formatiivisen arvioinnin käytänteitä, mutta sovelsivat enimmäkseen kaikille yhtenäisiä arviointijärjestelyjä. Yksilöllinen arviointi kohdistui heikosti menestyvien opiskelijoiden auttamiseen, kun taas hyvin menestyvät helposti sivuutettiin. Opettajat kaipaavat enemmän tietoa ja harjoitusta yksilöllistävään arviointiin työuransa eri vaiheissa.

Puheen automaattinen arviointi digitaalisen arviointivälineen avulla oli von Zansenin ja Heijalan tutkimusaihe. Neljän opettajahaastattelun avulla he kartoittivat kieltenopettajien käsityksiä mahdollisuuksista ja rajoituksista arviointivälineen käytössä sekä kuinka hyödyllisinä he pitivät arviointivälineen antamaa palautetta. Haastattelut osoittivat, että arviointivälinettä pidettiin hyvänä lisäresurssina puhumisen harjoittelussa, mutta tietosuoja koettiin hieman haastavaksi. Arviointivälineen antama palaute koettiin melko hyödylliseksi. Tulosten perusteella von Zansen ja Heijala pohtivat, että arviointivälineen käyttö puheen automaattisessa arvioinnissa voi parhaimmillaan täydentää kieltenopetuksen käytänteitä.

Saari ja Hilden tutkivat suomi toisena kielenä -opettajien näkökulmasta, miten arviointi voi tukea oppimista. Kyselyaineisto koostui 208 S2-opettajan vastauksesta. Tulokset osoittavat, että S2-opettajat tuntevat ja käyttävät useita formatiivisen arvioinnin keinoja, mutta he eivät välttämättä näe niitä oppimista edistävinä. Opettajat kuitenkin korostivat kannustavan palautteen suurta merkitystä tukea tarvitseville oppilaille. Saari ja Hilden tuovat esiin jatkokoulutuksen ja monipuolisen, formatiivista arviointia tukevan opetusmateriaalin tärkeyden.

Laine, Maijala ja Mäntylä tarkastelevat kestävän kehityksen periaatteiden ilmene- mistä kielten opetuksessa ja arvioinnissa. Opettajakyselyn (n=29) vastaukset kuvaavat, miten opettajat ymmärtävät ja panevat toimeen eettisesti kestävä arviointia. Saatujen tulosten mukaan opettajat eivät tunnista selkeästi opetus- ja arviointikäytänteiden yhteyksiä kestävän kehityksen periaatteisiin, mutta osallisuus ja vertaisarviointi tukevat niiden soveltamista kielen opetukseen ja arviointiin. Opettajat tarvitsevat myös enemmän tietoa siitä, miten eettisesti kestävä arviointi edistää kielenoppimisen ohella myös sosiaalista kestävyyttä.

Kauppinen, Tarnanen, Neittaanmäki, Aalto, Hankala, Kainulainen, Mustonen ja Torvelainen käsittelevät luokanopettajiksi opiskelevien (n=127) itsearviointitaidon kehittymistä sekä suomen kielen ja kirjallisuuden ylioppilaskokeen ennustearvoa suhteessa siihen kielitiedon tasoon, joka opiskelijoilla oli kurssin alussa suoritettun diagnostisen kokeen perusteella. Pilottitutkimuksen aineisto käsitti diagnostisen testin tulokset ja opiskelijoiden itsearvioinnit, joita tarkasteltiin määrällisesti ja laadullisesti. Tulokset osoittivat tilastollisesti merkitseviä yhteyksiä diagnostisessa testissä osoitetun kielitiedon ja itsearvioinnin välillä.

Kirjallisuus

- Andrade, H. L. 2019. A critical review of research on student self-assessment. *Frontiers in Education*, 4 (87), 1–13. <https://doi.org/10.3389/feduc.2019.00087>
- Askew, S. & Lodge, C. 2000. Gifts, ping-pong and loops – linking feedback and learning. Teoksessa S. Askew (toim.) *Feedback for learning*. Lontoo: Routledge/Falmer, 1–17.
- Bandura, A. 1993. Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist*, 28 (2), 117–148. https://doi.org/10.1207/s15326985ep2802_3
- Black, P. & Wiliam, D. 1998. Assessment and classroom learning. *Assessment in Education: Principles Policy & Practice*, 5 (1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Black, P. & Wiliam, D. 2009. Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Black, P. & Wiliam, D. 2018. Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25 (6), 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>
- Borg, S. 2015. *Teacher cognition and language education: Research and practice*. Lontoo: Bloomsbury Academic.
- Brookhart, S. M. 2011. Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30 (1), 3–12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Brookhart, S. M. 2017. *How to give effective feedback to your students* (2. p.). Alexandria, VA: ASCD.
- Brooks. V. 2002. *Assessment in secondary schools: the new teacher's guide to monitoring, assessment, recording, reporting and accountability*. Buckingham: Open University Press.
- Caena, F. 2011. *Literature review. Teachers' core competences: requirements and development*. European Commission.
- Carless, D. & Boud, D. 2018. The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, 43 (8), 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>
- Cizek, G. J., Andrade, H. L. & Bennett, R. E. 2019. Formative assessment: History, definition and progress. Teoksessa H.L. Andrade, R.E. Bennet & G. C. Cizek (toim.) *Handbook of formative assessment in the disciplines*. New York: Routledge, 3–19.
- Dlaska, A. & Krekeler, C. 2008. Self-assessment of pronunciation. *System*, 36, 506–516. <https://doi.org/10.1016/j.system.2008.03.003>
- Fulcher, G. 2012. Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9 (2), 113–132. <https://doi.org/10.1080/15434303.2011.642041>
- Gitomer, D. H. & Zisk, R. C. 2015. Knowing what teachers know. *Review of Research in Education*, 39 (1), 1–53. <https://doi.org/10.3102/0091732X14557001>
- Hattie, J. 2012. Visible learning for teachers: Maximizing impact on learning. Lontoo: Routledge. <https://doi.org/10.4324/9780203181522>
- Hattie, J. & Timperley, H. 2007. The power of feedback. *Review of Educational Research*, 77, 81–112. <https://doi.org/10.3102/003465430298487>
- Herpich, S., Praetorius, A.-K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., Behrmann, L., Böhmer, M., Ufer, S., Klug, J., Hetmanek, A., Ohle, A., Böhmer, I., Karing, C., Kaiser, J. & Südkamp, A. (2018). Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher Education*, 76, 181–193. <https://doi.org/10.1016/j.tate.2017.12.001>

- Huhta, A. & Hilden, R. (toim.). 2016. *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFinLA-e. Soveltavan kielitieteen tutkimuksia 9. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA.
- Huhta, A., Ilola, M., Lahti, L., Leontjev, D., Pollari, P., Salo, O-P. & Tarnanen, M. 2021. *ARVO – Kieltenopettajien arviointiosaamisen päivittäminen. Opetushallituksen rahoittama koulutushanke*. <https://aoe.fi/#/materiaali/1982>
- Inbar-Lourie, O. 2008. Constructing a language assessment knowledge base: A focus on language assessment courses. *Language testing* 25. <https://doi.org/10.1177/0265532208090158>
- Isbell, D. R. & Lee, J. 2022. Self-assessment of comprehensibility and accentedness in second language Korean. *Language Learning*, 72 (3), 806–852. <https://doi.org/10.1111/lang.12497>
- Jackson, M. & Marks, L. 2016. Improving the effectiveness of feedback by use of assessed reflections and withholding of grades. *Assessment & Evaluation in Higher Education*, 41 (4), 532–547. <https://doi.org/10.1080/02602938.2015.1030588>
- Klapp, A. 2015. *Bedömning, betyg och lärande*. Lund: Studentlitteratur.
- Koehler, M. J., Mishra, P. & Cain, W. 2013. What is technological pedagogical content knowledge (TPACK)? *Journal of Education*, 193 (3), 13–19. <https://doi.org/10.1177/002205741319300303>
- Kremmel, B. & Harding, L. 2020. Towards a comprehensive, empirical model of language assessment literacy across stakeholder groups: Developing the language assessment literacy survey. *Language Assessment Quarterly*, 17 (1), 100–120. <https://doi.org/10.1080/15434303.2019.1674855>
- Lantolf, J. & Poehner, M. 2011. *Dynamic assessment in the foreign language classroom. A teachers' guide* (2. p.). The Pennsylvania State University.
- Lantolf, J. & Poehner, M. 2014. *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. New York: Routledge.
- Leontjev, D. 2016. *ICAnDoIT: the impact of computerised adaptive corrective feedback on L2 English learners*. Väitöskirja. Jyväskylän yliopisto. <https://jyx.jyu.fi/handle/123456789/49346>
- 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. Ikkänen & Sigurd D'hondt (toim.) *Kieli, muutos ja yhteiskunta*. AFinLAn Vuosikirja. Suomen soveltavan kielitieteen yhdistyksen julkaisuja 79, 133–150. <https://doi.org/10.30661/afinlavk.113656>
- Li, H., Xiong, Y., Hunter, C. V., Guo, X. & Tywoniw, R. 2020. Does peer assessment promote student learning? A meta-analysis. *Assessment & Evaluation in Higher Education*, 45 (2), 193–211. <https://doi.org/10.1080/02602938.2019.1620679>
- Looney, A., Cumming, J., van Der Kleij, F. & Harris, K. 2018. Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy & Practice*, 25 (5), 442–467. <https://doi.org/10.1080/0969594X.2016.1268090>
- Lukiolaki 10.8.2018/714. <https://www.finlex.fi/fi/laki/ajantasa/2018/20180714>
- Molloy, E., Boud, D. & Henderson, M. 2020. Developing a learning-centred framework for feedback literacy. *Assessment & Evaluation in Higher Education*, 45 (4), 527–540. <https://doi.org/10.1080/02602938.2019.1667955>
- Mäkipää, T. 2021a. *"I have never received feedback from my teachers." Students' perceptions of assessment and feedback practices in Finnish general upper secondary education*. Väitöskirja. Helsingin yliopisto. <https://helda.helsinki.fi/handle/10138/336417>

- Mäkipää, T. 2021b. Students' and teachers' perceptions of self-assessment and teacher feedback in foreign language teaching in general upper secondary education – A case study in Finland. *Cogent Education*, 8 (1), 1–19. <https://doi.org/10.1080/2331186X.2021.1978622>
- Nawas, A. 2020. Grading anxiety with self and peer-assessment: A mixed-method study in an Indonesian EFL context. *Issues in Educational Research*, 30 (1), 224–244.
- Ng, W. S. & Yu, G. 2021. Students' attitude to peer assessment process: a critical factor for success. *Interactive Learning Environments*, 1–19. <https://doi.org/10.1080/10494820.2021.1916762>
- Nicol, D. J. & Macfarlane-Dick, D. 2006. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31 (2), 199–218. <https://doi.org/10.1080/03075070600572090>
- Opetushallitus. 2014. *Perusopetuksen opetussuunnitelman perusteet 2014*. Määräykset ja ohjeet 2014:96.
- Opetushallitus. 2019. *Lukion opetussuunnitelman perusteet 2019*. Määräykset ja ohjeet 2019:2a.
- Pastore, S. & H. L. Andrade 2019. Teacher assessment literacy: A three-dimensional model. *Teaching and Teacher Education*, 84, 128–138. <https://doi.org/10.1016/j.tate.2019.05.003>
- Perusopetuslaki 12.8.1998/628. <https://www.finlex.fi/fi/laki/ajantasa/1998/19980628>
- Poehner, M. 2008. *Dynamic assessment: A Vygotskian approach to understanding and promoting second language development*. New York: Springer.
- Poehner, M. 2017. Sociocultural theory and the dialectical-materialist approach to L2 development. *Language and Sociocultural Theory*, 3 (2), 133–152. <https://doi.org/10.1558/lst.v3i2.32869>
- Poehner, M. & Lantolf, J. 2013. Bringing the ZPD into the equation: Capturing L2 development during computerized dynamic assessment (C-DA). *Language Teaching Research*, 17 (3), 323–342. <https://doi.org/10.1177/1362168813482935>
- Raaijmakers, S. F., Baars, M., Paas, F., van Merriënboer, J. J. G. & van Gog, T. 2018. Training self-assessment and task-selection skills to foster self-regulated learning: Do trained skills transfer across domains. *Applied Cognitive Psychology*, 32 (2), 270–277. <https://doi.org/10.1002/acp.3392>
- Ruiz-Primo, M. A. & Brookhart, S. M. 2018. *Using feedback to improve learning*. New York: Routledge.
- Scriven, M. 1967. The methodology of evaluation. Teoksessa R. M. Gagne, R. W. Tyler & M. Scriven (toim.) *Perspectives of curriculum evaluation*. Chicago: Rand McNally, 39–83.
- Shulman, L. S. 1987. Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57 (1), 1–21.
- Shute, V. J. 2008. Focus on formative feedback. *Review of Educational Research*, 78 (1), 153–189. <https://doi.org/10.3102/0034654307313795>
- Sternberg, R. J. & Grigorenko, E. L. 2002. *Dynamic testing: The nature and measurement of learning potential*. Cambridge: Cambridge University Press.
- Taylor, L. 2013. Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30 (3), 403–412. <https://doi.org/10.1177/0265532213480338>
- Topping, K. J. 2018. *Using peer assessment to inspire reflection and learning*. New York: Routledge.
- Vogt, K. & Tsagari, D. 2014. Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly* 11, (4), 374–402. <https://doi.org/10.1080/15434303.2014.960046>

- Vogt, K., Tsagari, D. & Spanoudis, G. 2020. What do teachers think they want? A comparative study of in-service language teachers' beliefs on LAL training needs. *Language Assessment Quarterly*, 17 (4), 386–409. <https://doi.org/10.1080/15434303.2020.1781128>
- Vygotsky, L. S. 1978. *Mind in society: The development of higher psychological processes*. M. Cole, V. John-Steiner, S. Scribner & E. Souberman (toim.). Harvard: Harvard University Press.
- Xu, Y. & Brown, G. T. L. 2016. Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- van Zundert, M., Sluijsmans, D. & van Merriënboer, J. 2010. Effective peer assessment processes: Research findings and future directions. *Learning and Instruction*, 20, 270–279. <https://doi.org/10.1016/j.learninstruc.2009.08.004>

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 19–39.

Toni Mäkipää
University of Helsinki

Olli-Pekka Salo
University of Jyväskylä

Using self-assessment to enhance self-regulated learning – an intervention study in a Finnish lower secondary school

Highlights

- The emphasised use of self-assessment enhances students' perceptions of self-regulated learning.
- The level of reflection in ninth graders' self-assessments is low.
- More attention to supporting reflection and learning to learn is required.

Abstract

The aim of this intervention study was to explore how the emphasised use of self-assessment in Swedish language classes affects ninth graders' perceptions of self-regulated learning and their self-assessments in Finland. The dataset includes the survey responses of 36 students and their written self-assessments. The results suggest that the quality of the self-assessments increased during the intervention, although the level of reflection was still superficial. The students exhibited more positive perceptions of their skills in self-regulated learning at the end of the intervention. The results indicate that incorporating self-assessment into teaching boosts students' perceptions of self-regulated learning and enhances the quality of their self-assessments. Swedish language teachers should consider the extent to which they use self-assessment in courses and whether their practices could be adjusted to promote reflection and self-regulated learning.

Keywords: self-regulated learning, self-assessment, teaching of Swedish

1 Introduction

The purpose of this intervention study was to explore whether the emphasised use of self-assessment in Swedish classes in the ninth grade enhances students' perceptions of self-regulated learning (SRL) and students' written self-assessments. Put differently, the perceptions and the self-assessments were explored from the perspective of learning Swedish as a second language. Self-assessment is a key component of SRL (Hosseini & Nimehchisalem 2021), and self-assessment is a formative assessment practice that basic education (grades 1–9) teachers in Finland are expected to use (FNAE 2014). SRL is a key aspect of learning in Finland as the national curriculum for basic education in Finland (FNAE 2014) accentuates the importance of SRL. At the end of basic education, students are expected to develop their skills in learning to learn, lifelong learning, problem solving, and SRL (FNAE 2014).

A study by Mäkipää (2021) examined students' and teachers' perceptions of self-assessment and teacher feedback in enhancing SRL in general upper secondary education. Otherwise, the relationship between SRL and foreign language teaching and particularly learning Swedish has not been mapped. However, to some extent the role of SRL has been considered in Finnish studies on novel forms of language teaching such as language portfolios (e.g. Kohonen & Korhonen 2007), classroom tandem pedagogy (e.g. Hansell et al. 2013) or distance learning (e.g. Kotilainen 2015). Moreover, Theobald's (2021) meta-analysis showed that SRL training programs foster university students' motivation, SRL strategies and academic performance. However, from teachers' perspectives, it is still unclear how they can stimulate students' SRL effectively (Dignath & Veenman 2020). To map how teachers can enhance SRL, more research is needed in heterogeneous student groups (Dignath & Veenman 2020). Because effective learners are self-regulated (Butler & Winne 1995) and metacogni-

tive processes of SRL correlate notably with academic performance (Dent & Koenka 2016), it is essential to explore SRL. Therefore, we conducted an intervention study to examine the relationship between self-assessment and SRL. We sought answers to two research questions in this paper: (1) How does the emphasised use of self-assessment enhance ninth graders' perceptions of SRL? (2) How does the emphasised use of self-assessment enhance ninth graders' self-assessments?

2 Theoretical background

2.1 Self-regulated learning

According to Hadvin et al. (2018), SRL means that students take metacognitive control of conditions related to emotions, motivation, behaviour, and cognition. They simultaneously plan for actions as well as monitor and evaluate their progress (Hadvin et al. 2018). In essence, SRL refers to working systematically to attain learning goals through sustaining and activating affects and cognitions (Schunk & Greene 2018). After careful evaluation and monitoring, self-regulated learners are cognizant of their motivation, beliefs, and knowledge (Butler & Winne 1995). Self-regulated learners use an array of strategies to attain their goals (Theobald 2021) and are committed to attaining the goals (Zimmerman 1989). In short, self-regulated learners are active and have control over their learning (Zimmerman 1998) as they modify and generate their actions, feelings, and thoughts (Andrade 2010). Self-regulated learners are also proficient users of self-assessment and peer assessment (Nicol & Macfarlane-Dick 2006).

Several models of SRL have been proposed. Although they underscore various components of SRL, scholars acknowledge that SRL involves monitoring and controlling motivation, cognition, behaviour and the environment (Efklides 2011). In other words, becoming a self-regulated learner is not attained solely by personal processes, as students are also affected by environmental and behavioural factors reciprocally (Zimmerman 1989). Therefore, teachers' role in solidifying SRL is indispensable (Norrena 2019). For example, if teachers plan collaborative learning tasks that require student responsibility and provide optimal challenge, teachers can activate SRL (Järvelä et al. 2016). As SRL is related to motivation and therefore is not necessarily a static feature in each student, teachers can also encourage students to use motivational learning strategies and to regulate emotions (Oxford 2017). When planning instruction of strategies, teachers should consider the context and its affordances in order to stimulate engagement (Oxford 2017). This also accentuates the need for context-specific research concerning SRL.

If designed accordingly, SRL interventions can accelerate learning (Panadero 2017). Jansen and colleagues' (2019) meta-analysis indicated that SRL interventions

foster students' achievement and engagement, although the effect sizes were smaller when compared to previous meta-analyses. When planning intervention studies in secondary education, models focusing on metacognitive theories promote learning more compared with primary education (Dignath & Büttner 2008), such as the model proposed by Efklides (2011). Her MASRL model (The Metacognitive and Affective Model of SRL) includes two levels: the Person level and the Task x Person level. The Person level involves traits and characteristics formed in prior learning occasions, such as volitional, affective, motivational, metacognitive and cognitive characteristics. These traits affect decisions on engagement with tasks and are considered to be top-down self-regulation. At the Task x Person level (bottom-up), students process the task. The basic functions of this level are cognition, metacognition, affect, and regulation of affect and effort (Efklides 2011).

2.2 Self-assessment

In SRL, the key catalyst is feedback (Butler & Winne 1995), and the feedback can be provided by a teacher, a peer or the student (Nicol & Macfarlane-Dick 2006). Self-assessment refers to "a descriptive and evaluative act carried out by the student concerning his or her own work and academic abilities" (Brown & Harris 2013: 368). This definition emphasises both performance and development (Hosseini & Nimehchisalem 2021). Self-assessment promotes autonomy and focuses on the process (learning) instead of the product (results) (Dickinson 1987). If used efficiently, self-assessment practices boost learning (Brown & Harris 2013; Li & Zhang 2021) and SRL (Brown & Harris 2013; Nicol & Macfarlane-Dick 2006).

Self-assessment and SRL can be considered to be complementary processes that can improve students' achievement (Andrade 2010). To assist learning with self-assessment, students need a clear rubric and training in how to use the rubric and the criteria (Li & Zhang 2021). Moreover, in her meta-analysis of 76 empirical studies on student self-assessment, Andrade (2019) highlights that self-assessment is most beneficial, in terms of both achievement and self-regulated learning, when it is used formatively and is supported by training. However, Finnish research suggests that training on self-assessment and guidance on using the criteria is inadequate in foreign language teaching (Mäkipää 2021).

In foreign language teaching, the ability to assess one's learning and performance is indispensable (Dickinson 1987). Many recent studies (e.g. Jamrus & Razali 2019; Qasem 2020; Jin 2021) indicate that language learners' self-assessment can successfully develop students' second language learning ability. However, professional training and valid assessment tools are necessary for helping students effectively use self-assessment as an SRL method (Jin 2021).

3 Methodology

3.1 Participants

The participants of this study were 36 ninth graders from two classes in Finland. They were a convenience sample and were studying Swedish as a B1 language. The classes had 39 students but three did not participate in the study. In terms of gender, 19 students were boys, 14 students were girls, and three were non-binary. Thirteen students were 14 years old, and 23 students were 15 years old. The students' previous course grades are displayed in Table 1.

TABLE 1: Students' previous Swedish grades.

Grade	Number of students
6 = fair	2
7 = satisfactory	5
8 = good	5
9 = very good	11
10 = excellent	9
do not remember	4

Note: students in Finnish lower secondary schools are graded from four (fail) to ten (excellent).

As displayed in Table 1, most students' previous Swedish grades were higher than average students' grades in general. Only seven students had grades lower than eight (good).

3.2 Context of the study

Since autumn 2022, the compulsory education in Finland consists of nine years of basic education (grades 1 to 9) and three years of upper secondary or vocational education. In addition, there is pre-primary education for one year, from age six to seven, which has been compulsory since 2015 (see FNAE 2021).

The distribution of lesson hours in basic education totals to a minimum of 224 lesson hours, meaning that the students average 24.9 lesson hours per school week. Studying compulsory languages comprises 24 lesson hours, of which 18 go to the advanced syllabus (A1 language), and six to the intermediate level syllabus (the B1 language).

There are no external examinations in basic education in Finland, as the grades in all school reports are based on teachers' assessments. The only external, high-stakes examination in the Finnish school system is the Matriculation Examination, which the upper secondary students take at the end of their studying (see e.g. Pollari, Salo & Koski 2018).

The students begin their A1 language studies in the first grade and the B1 language in the sixth grade at the latest. In addition, a student may select an optional third language (A2) that begins, at the latest, in the fifth year and an optional fourth language generally starting in the 8th grade of basic education. The compulsory language starting in the 6th grade is the second national language (either Finnish or Swedish) unless the student is already studying this as their A1 or A2 language. In basic education, the intermediate level syllabus in Swedish (B1-Swedish) usually starts in the sixth grade and comprises six lesson hours in four academic years, two of them in grade 6 and four divided between grades 7–9.

3.3 Content of the intervention

The intervention was conducted in autumn 2021 with two groups of ninth graders (n=36). The students had 12 Swedish lessons, 75 minutes each, and self-assessment issues were touched upon either in the form of tasks, exercises, homework or discussions in each lesson (see Appendix 1).

In lesson 1, the students filled in a self-assessment grid which consisted of can-do statements based on the criteria for the final assessment in basic education. They also set the goals for their Swedish lessons for the autumn. In lesson 2, the students answered the self-assessment survey for the first time. In lesson 3, the class discussed self-assessment skills, i.e., what self-assessment is, how and why it is undertaken and how the students can benefit from it when they are studying. After this introduction to the theme of the intervention, the students were immediately to put the theory into practice by doing a self-assessment task after an oral exercise. As homework, the students were given a task in which they were to use the present perfect to describe what they had done in the past week. The students were supposed to record and upload the task on the group's learning platform before the next lesson. The homework task was accompanied by a self-assessment task in which the students were to assess their oral skills and reflect on their study habits.

The students were supposed to submit the homework tasks before lesson 4. Only 20 students out of 36 did that in time, 13 students handed in the tasks late and three did not submit the tasks at all. Nevertheless, the students discussed the tasks in pairs or small groups, followed by a brief discussion led by the teacher. In general, the students regarded them as meaningful. In lesson 5, the teacher gave general oral feedback on how the students had managed both the oral task and the self-assessment task. In Lesson 6, the students did an oral task in which they interviewed each

other in Swedish. The task was preceded by a brief introduction on what communication strategies could be used when executing the task. After the task, the teacher conducted a “thumb vote” on how the students succeeded in the task. About a third put their thumbs up, about a half wagged their thumbs up and down, and the rest put their thumbs down.

In lesson 7, the students did a mid-term assessment of the goals they had set at the beginning of the autumn term and updated them when necessary. After lesson 8, the students were given a read-aloud exercise as homework. They were to read aloud a text that they had previously listened to and studied as a reading comprehension exercise. The homework exercise was accompanied by a self-assessment task in which the students were to assess their oral skills and study habits (a modified version of the previous one, including assessment on the pronunciation of certain words in the text they were to read). The students were supposed to submit the homework tasks before lesson 9 (14/36 students submitted in time, 17/36 were late, and 5/36 did not submit it at all). Also, this time, the class discussed the tasks in pairs or groups, followed by a discussion led by the teacher, and once again, the students regarded the tasks as meaningful.

In lesson 10, the teacher gave general oral feedback on how the students had managed the oral task focusing on how pronunciation affects comprehensibility (the main assessment criterion) and how factors such as sentence stress affect it. In lesson 11, the students did an oral alias task, preceded by a brief introduction to production strategies. Thereafter the teacher conducted a thumb vote on how the students succeeded in the task. About one fifth of the students put their thumbs up, about 60 per cent wagged their thumbs up and down, and one fifth put their thumbs down. In lesson 12, the students answered the self-assessment survey for the second time.

The intervention was based on rigorous research (Dignath & Büttner 2008; Efklides 2011; Nicol & MacFarlane-Dick 2006; Panadero 2017). According to Dignath and Büttner (2008), SRL interventions are more successful if group work is implemented with older students, such as secondary students, who were the target group of our intervention. Metacognitive reflection is also a key issue in SRL interventions, and reflection was present in our research design (Panadero 2017). The teacher also provided feedback on students’ self-assessments and tasks, which is recommended by Nicol & MacFarlane-Dick (2006). Strategy use was also practised diversely (Dignath & Büttner 2008). For ethical reasons, no control group was used in this study. According to the guidelines of the national core curriculum (FNAE 2014), self-assessment practices are a focal aspect of every subject, and they should be incorporated into all teaching. Therefore, using a control group without self-assessments would contradict the requirements of the curriculum and the Basic Education Act.

All the students participated voluntarily and anonymously in this intervention. The students signed a written consent form and received a privacy notice prior to

starting the intervention. Research permission was sought from the headmaster to conduct research at the school.

3.4 Data analysis

The data used in this study comprised students' survey responses (n=36, n=30) and written self-assessments (n=30, n=26), both conducted twice. However, it should be noted that not all students answered the survey or handed in the self-assessments twice, which affects the interpretation of the results. Due to the small number of participants, the non-parametric Mann-Whitney U test was used to compare students' survey responses. The survey included 25 items about SRL (Norrena 2019), and these items had specifically been designed for lower secondary students. However, some of the wordings were altered slightly. The items were divided into five groups: planning, working, choices, assessment, and environment (Norrena 2019). The students responded on a scale from one (I completely disagree) to five (I completely agree). As the number of participants was low, factor analysis was not suitable. Instead, the items were explored one-by-one. However, Cronbach's alpha was used to ensure the reliability of the analysis. The alpha values were as follows: planning ($\alpha = .81$), working ($\alpha = .77$), choices ($\alpha = .80$), assessment ($\alpha = .81$), and environment ($\alpha = .63$). The values were fairly high except for environment.

In terms of the students' self-assessments, deductive content analysis with the SOLO (Structure of the observed learning outcome) taxonomy (Biggs & Tang 2011) was used. The taxonomy encompasses five levels: prestructural, unistructural, multistructural, relational, and extended abstract. These levels are described in Table 2.

TABLE 2: Five levels of the SOLO taxonomy (Biggs & Tang 2011).

Level	Description
prestructural	misses the point, lacks knowledge or understanding
unistructural	is able to mention one issue
multistructural	is able to mention several issues but misses greater connections
relational	is able to integrate concepts, shows understanding
extended abstract	is able to reflect and hypothesise, goes beyond the task

As described in Table 2, at the lowest level, students are not able to perform the task and might try to cover that by using tautology, but gradually, students acquire a more detailed understanding and skills for reflection (Biggs & Tang 2011). The SOLO taxonomy can be regarded as being a hierarchy aimed at increasing knowledge

(unistructural – multistructural) and to deepen understanding (relational – extended abstract) (Biggs & Tang 2011). The minimum age groups for the base stage for the higher levels are 13–15 years (relational) and 16+ years (extended abstract) (Biggs & Collis 1982). It was to be expected that few students from our intervention would reach the extended abstract level, whereas the relational level should be attainable for everyone, as the participants were about 14–15-year-old students.

The self-assessments were assessed using the levels of the SOLO taxonomy. The first author assessed the levels independently, after which the analysis was discussed and confirmed with the second author. The students answered the same questions twice: (1) “Which issues are easy for you in speaking Swedish?”, and (2) “Which issues are difficult for you in speaking Swedish?” Students’ oral proficiency was chosen for the self-assessment because the national core curriculum emphasises oral proficiency (FNAE 2014). Even though the use of oral exercises is prevalent in foreign language teaching in Finland, students are rarely given adequate support and the exercises are fairly often insufficiently intertwined with overall objectives of communicative language proficiency (Bovellan 2020; Hildén et al. 2014). Thirty students handed in the self-assessment at the beginning of the intervention, and 26 at the end of it. According to the teacher who oversaw the intervention, the students who did not hand in the self-assessments have difficulties in submitting course work and often fail to meet deadlines. This needs to be considered in the interpretation of the results as they might be slightly different if the weaker students had submitted the assignments on time.

4 Results

First, the students’ responses to the online survey are shown (RQ 1). Second, the students’ self-assessments at the SOLO taxonomy levels are discussed (RQ 2).

4.1 Students’ perceptions of self-regulated learning

Using the Mann-Whitney U test, the students’ perceptions of SRL were explored. The results are displayed in the five groups employed by Norrena (2019). The means of the responses were primarily higher at the end of the intervention (21 items, 84%), but most of them were not statistically significant. Only statistically significant differences have been reported here.

4.1.1 Planning

The students' perceptions of planning their work are illustrated in Table 3.

TABLE 3: Students' perceptions of planning their work.

	first responses		second responses	
	M	S.D.	M	S.D.
1. The expression "well begun is half done" describes how I work.	3.39	1.05	3.16	1.07
2. I choose a specific tactic or a strategy for doing tasks.	3.11	1.17	3.64	1.04
3. When I start to do a task, I ponder what went wrong previously or at what I succeeded, based on which I plan my work.	3.00	1.12	3.44	1.04
4. When I start to do a task, I ponder what challenges I might encounter and how I can overcome them.	3.03	1.21	3.44	1.23
5. I plan my work before starting to do tasks (in the lesson).	2.81	1.06	2.92	1.00

Note: M = mean, S.D. = standard deviation

As Table 3 illustrates, the students mostly pondered their strategy use, challenges, and accomplishments when they were doing tasks. Nevertheless, the actual planning of working seemed to be incomplete. The difference in the fourth item was almost statistically significant ($p = .064$).

4.1.2 Working

The students' perceptions of working are illustrated in Table 4.

TABLE 4: Students' perceptions of their work.

	first responses		second responses	
	M	S.D.	M	S.D.
1. I focus on the tasks instead of doing something else.	3.47	1.21	3.72	0.98
2. I try to start working as quickly as possible and then try to react accordingly.	3.50	1.08	3.84	1.07
3. In group work, I take the initiative instead of just doing what I am asked to do.	3.47	1.14	4.04	0.98
4. I make use of group members' contributions and strengths.	3.78	0.90	4.20	0.71
5. When I am doing more comprehensive tasks, I stop for a moment to ponder what the goal was.	3.11	0.95	3.72	0.89

As Table 4 highlights, most students had an active role in group work, and they used peers' strengths to deepen learning and working. The students also pondered the goals of tasks and were able to focus on the tasks. Statistically significant differences were detected in the third item, $U(N_{\text{beginning}}=36, N_{\text{end}}=30) = 322.00, z = -1.997, p = .047$, and in the fifth item $U(N_{\text{beginning}}=36, N_{\text{end}}=30) = 293.500, z = -2.430, p = .014$. The effect sizes ($\eta^2 = \text{partial eta squared}$) were medium (.07 and .10, respectively). The fourth item was near the cut-off value of a statistically significant difference ($p = .066$).

4.1.3 Making choices

The students' perceptions of making choices are illustrated in Table 5.

TABLE 5: Students' perceptions of making choices.

	first responses		second responses	
	M	S.D.	M	S.D.
1. I can name another perspective from which I could do the task.	3.78	0.87	3.52	0.77
2. I am intrinsically motivated to do tasks.	3.06	1.19	3.16	0.94

	first responses		second responses	
	M	S.D.	M	S.D.
3. I try various work methods to find the most suitable way to learn.	3.56	0.97	3.84	0.85
4. I want to improve myself during the tasks instead of merely trying to finish the work quickly.	3.36	1.20	3.72	0.84
5. I try various learning tools to find the most suitable one for the situation.	3.22	1.10	3.52	1.05

As shown in Table 5, the students were able to name perspectives and work methods to accommodate learning. They were also interested in developing their skills. In terms of motivation, the students were somewhat intrinsically motivated to learn.

4.1.4 Assessment

The students' perceptions of assessment are illustrated in Table 6.

TABLE 6: Students' perceptions of assessment.

	first responses		second responses	
	M	S.D.	M	S.D.
1. After the task, I usually think about what surprised me or what new things I learned.	3.03	1.00	3.40	0.96
2. After the task, I reserve time for pondering how I succeeded.	2.61	1.05	2.80	1.12
3. I ask others to provide me with feedback on my work.	2.47	1.32	2.24	1.17
4. After the task, I separate the things at which I succeeded and the things that I could have done better.	3.03	1.03	3.32	1.11
5. After the task, I ponder whether my goal was challenging enough.	3.11	1.06	3.36	0.76

Table 6 shows that the students mostly pondered their actions and progress after undertaking the tasks. They also considered the goals. However, the students did not opt for peer feedback, and they did not reserve time for pondering how they had succeeded after the task.

4.1.5 Learning environment

The students' perceptions of the learning environment are illustrated in Table 7.

TABLE 7: Students' perceptions of the learning environment.

	first responses		second responses	
	M	S.D.	M	S.D.
1. We decide the rules for our working principles together.	4.03	0.65	4.00	0.87
2. I ponder how the environment affects my working.	3.11	1.17	3.36	1.04
3. The atmosphere when I am working is usually confidential and safe: I feel good.	3.69	1.09	4.16	1.03
4. I know from where and how I can get help during a task.	4.03	0.88	4.44	0.77
5. I usually have peaceful working conditions.	3.72	1.06	3.80	1.08

As Table 7 depicts, the environment seemed to support student' SRL. The students pondered the effects of the environment on learning, and they received support from the learning environment. The environment also seemed to promote peaceful working conditions. The differences in the third ($p = .076$) and fourth ($p = .057$) items were close to being statistically significant.

4.2 Students' self-assessments

Using the SOLO taxonomy, deductive content analysis was used to analyse the students' self-assessments concerning the second research question. The students' levels are displayed in Table 8.

TABLE 8: The levels of students' self-assessments.

Unchallenging issues in speaking Swedish		
level	at the beginning (n=30)	at the end (n=26)
prestructural	12	4
unistructural	18	22
Challenging issues in speaking Swedish		
level	at the beginning (n=30)	at the end (n=26)
prestructural	7	-
unistructural	14	16
multistructural	9	10

As displayed in Table 8, most students were at the unistructural level regarding unchallenging issues. All the students at the prestructural level answered that they did not know what was unchallenging for them. At the unistructural level, most students found pronunciation, short words, and familiar words to be unchallenging for them in speaking Swedish. Furthermore, the number of students at the prestructural level was lower at the end of the intervention, and the number of students at the unistructural level increased.

In terms of challenging issues, most students were at the unistructural level. At the prestructural level, the students were unable to pinpoint difficult issues in speaking Swedish. At the unistructural and multistructural levels, most students considered pronunciation, long words, and difficult words to be challenging for them. Grammatical issues (e.g. word order and word inflection) were mentioned by a few students. Furthermore, none of the students were at the prestructural level at the end of the intervention, and the number of students at the unistructural and multistructural levels increased.

When comparing unchallenging and challenging issues in speaking Swedish, it is apparent that the students were able to list more issues that were challenging for them. Moreover, pronunciation was one of the most common responses in both unchallenging and challenging issues.

5 Discussion

In the Finnish context, this is the first study focusing on SRL and self-assessment in learning Swedish at the lower secondary level. This is also the first intervention study set in Finland that focuses on second language learning and SRL. Therefore, this study enhances our knowledge of SRL by detailing how the emphasised use of self-assessment contributes to fostering SRL in the context of second language learning. Regarding the first research question, the quantitative results show encouraging results: over 80% of the means increased towards the end of the intervention. To paraphrase, the intervention affected students' perceptions of SRL positively although the intervention was short. However, the number of statistically significant differences was low. Previous research has emphasised that SRL interventions are more successful if group work (Dignath & Büttner 2008) and metacognitive reflection (Panadero 2017) are implemented in the intervention. Our results concur with those recommendations.

The intervention affected students' perceptions of working the most. After the intervention, the students were more active in group work and pondered the goals while working. However, the results also indicate some challenges in fostering SRL. The means of the assessment items were lower than the means of the other items. In particular, it seems that students do not seek peer feedback or reserve time after the task for pondering on it. These results are not surprising as formative assessment practices in foreign language teaching in Finland are not common (Mäkipää 2021). In essence, the results suggest that more attention needs to be paid to enhancing peer feedback practices and reflection after the task.

The second research question focused on students' written self-assessments. As evidenced by the results, the quality of students' self-assessments improved during the intervention. At the end of it, the students were able to pinpoint more challenging and unchallenging issues in speaking Swedish than at the beginning. This result confirms the recommendation in the literature that instruction from the teacher is vital in practising self-assessment (Andrade 2019; Li & Zhang 2021). Put differently, instruction availed many students as a way to attain a higher level in the SOLO taxonomy. Considering that self-assessment promotes SRL (Brown & Harris 2013; Nicol & Macfarlane-Dick 2006) and that SRL interventions are conducive to learning and achievement (Jansen et al. 2019; Panadero 2017), we argue that our intervention supported the students' SRL. This is evident in both the quantitative and the qualitative results. The results are fairly encouraging, as it was not possible to provide the students with enough opportunities to automate strategy use and practice using the strategies (Dignath & Büttner 2008) the intervention being as brief as this. However, SRL skills develop gradually, and thus, they need to be practised accordingly. Nevertheless, it is not yet known how teachers can accelerate SRL efficiently (Dignath & Veenman 2020). This study indicates that emphasising the use of self-assessments could be one solution to that conundrum.

One feature of this study that raises serious questions is the fairly large number of students who did not submit the survey responses ($n=6$, only the second time) or hand in the written self-assessments ($n=6$, $n=10$) and those who decided not to participate in the study at all ($n=3$). In addition, several students handed in the two homework assignments late ($n=13$, $n=17$) or not at all ($n=3$, $n=5$). Partly, the results can be explained by the tight schedule of the intervention and the fact that the group had a two-month break in their Swedish lessons directly after the intervention, and therefore had no regular contact with their teacher. However, more importantly, the students in question lacked the motivation not only to study Swedish, but also for school in general.

As motivation is an extremely important, if not the most important prerequisite of learning (Salmela-Aro 2018; Järvenoja et al. 2018), it also affects SRL. As motivation and learning are intertwined, it is of utmost importance to enhance students' motivation throughout their school years. In their theory of self-determination, Deci et al. (1991) proposed that motivation is enhanced efficiently by giving the students autonomy to decide upon and plan their work. On the other hand, the weaker students in this study seemed to have a fixed mindset of their possibilities to learn new issues, which is characterised by a tendency to avoid challenges and effort as well as ignore useful negative feedback (see Dweck 2006). Therefore, we suggest that teachers critically analyse the tasks they use and contemplate whether the tasks enhance every student's SRL, language learning, and motivation. From a wider perspective, examining the learning goals of education should be a primary concern for all stakeholders. More attention should be paid to shifting the focus on reflection and analytic skills as well as support for students.

6 Conclusions

The following conclusions can be drawn, based on the results. First, we argue that Swedish teachers should consider teaching self-assessment skills more explicitly. In Finnish upper secondary schools, not all language teachers teach self-assessment skills (Mäkipää 2021). As the ability to self-assess is propitious to language learning (Jamrus & Razali 2019; Qasem 2020; Jin 2021), Swedish teachers should consider the extent and amount of time they devote to self-assessment. Teaching explicitly how to self-assess would contribute to both language learning and SRL.

Second, the students were not able to attain the relational level. In other words, the students were not able to reflect on their learning profoundly. To attain that level, students are expected to be 13–15 years old (Biggs & Collis 1982). All the students were in this age group, which is why the result was unanticipated. This suggests that the students did not internalise the core meaning of self-assessment. In other words, when Swedish teachers teach self-assessment skills, they should give more weight

to the actual content of the self-assessment. By providing individual and instructive feedback on students' self-assessments, teachers might help students actualise how they can diversify and develop their self-assessment practices.

Third, the analysis of the self-assessments showed that the students regard pronunciation as challenging, whereas grammatical issues were mentioned by only a few students. The pronunciation of long and unfamiliar words especially seems to cause problems, and some students also said that it is difficult to achieve correct intonation. The results suggest that more time should be allotted to practising oral skills in the classroom which is also entailed in the national core curriculum (FNAE 2014).

One of the major limitations of this study is its design, as the intervention was conducted at one school, and the participants were a convenience sample. The number of students was not high, and not all the students answered the survey or handed in the self-assessments twice. Moreover, this study focused on Swedish lessons and disregarded how teachers in other subjects use self-assessment and whether that usage has provoked any effects on students' perceptions in Swedish courses. Moreover, caution needs to be applied while interpreting the results because most of the differences were not statistically significant.

Given that SRL in foreign language teaching is an unmapped research field in Finland, more research is required to elucidate how language teachers can solidify SRL. Carefully planned intervention studies with a higher number of participants could illuminate how students become self-regulated over time, and which practices enhance SRL the most. Most of the students also had good or excellent grades in Swedish courses, which means that future studies should explore students with low course grades. Subsequent studies also need to explore peer assessment and whether the emphasised use of peer assessment promotes SRL.

Funding

This study was funded by the Swedish Cultural Foundation in Finland.

Literature

- Andrade, H.L. 2010. Students as the definitive source of formative assessment. In H. L. Andrade & G. J. Cizek (eds) *Handbook of formative assessment*. New York: Routledge, 90–105.
- Andrade, H.L. 2019. A critical review of research on student self-assessment. *Frontiers in Education* 4, 1–13. <https://doi.org/10.3389/feduc.2019.00087>
- Biggs, J. B. & Collis, K. F. 1982. *Evaluating the quality of learning: The SOLO taxonomy (structure of the observed learning outcome)*. New York: Academic Press.
- Biggs, J. & Tang, C. 2011. *Teaching for quality learning at university* (4th ed.). Berkshire: Open University Press.
- Bovellan, E. 2020. Seitsemäsluokkalaisten A-englannin oppimistulosten arviointi: Tuloksen pohdintaa englannin oppimisen, opettamisen ja arvioinnin näkökulmasta [Evaluation of 7th graders' learning outcomes in A-level English: Discussing the results from the point of view of learning, teaching and assessing English]. In M. Härmälä & M. Puukko (eds) *Taitotasojen ja suullista kielitaitoa. Näkökulmia englannin kielen oppimistulosten arvioinnin osoittamiin kehittämiskohteisiin*. Helsinki: Kansallinen koulutuksen arviointikeskus, artikkelikokoelma 1:2020, 87–95.
- Brown, G. T. L. & Harris, L. R. 2013. Student self-assessment. In J. H. McMillan (ed.) *SAGE handbook of research on classroom assessment*. Thousand Oaks, California: SAGE, 367–394.
- Butler, D. L. & Winne, P. H. 1995. Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65 (3), 245–281. <https://doi.org/10.3102/00346543065003245>
- Deci, E. L., Vallerand, R. J., Pelletier, L. G. & Ryan, R. M. 1991. Motivation and education: The self-determination perspective. *Educational Psychologist*, 26 (3–4), 325–346. <https://doi.org/10.1080/00461520.1991.9653137>
- Dent, A. L. & Koenka, A. C. 2016. The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, 28, 425–474. 10.1007/s10648-015-9320-8
- Dickinson, L. 1987. *Self-instruction in language learning*. Cambridge: Cambridge University Press.
- Dignath, C. & Büttner, G. 2008. Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning*, 3 (3), 231–264. 10.1007/s11409-008-9029-x
- Dignath, C. & Veenman, M. V. J. 2020. The role of direct strategy instruction and indirect activation of self-regulated learning – evidence from classroom observation studies. *Educational Psychology Review*, 33, 489–533. <https://doi.org/10.1007/s10648-020-09534-0>
- Dweck, C. 2006. *Mindset: The new psychology of success*. New York: Ballantine Books.
- Efklides, A. 2011. Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist*, 46 (1), 6–25. <https://doi.org/10.1080/00461520.2011.538645>
- FNAE 2014 = Finnish National Agency for Education. 2014. *National core curriculum for basic education 2014*. Helsinki: Finnish National Agency for Education.
- FNAE 2021 = Finnish National Agency for Education. 2021. Finnish education system. Retrieved from <https://www.oph.fi/en/education-system> (Accessed on Aug 16, 2022).
- Hadvin, A., Järvelä, S. & Miller, M. 2018. Self-regulation, co-regulation, and shared regulation in collaborative learning environments. In D. H. Schunk & J. A. Greene (eds) *Handbook of self-regulation of learning and performance* (2nd ed.). New York: Routledge, 83–106.

- Hansell, K., Pörn, M., Rusk, F. & Björkskog, L. 2013. Classroom tandem – outlining a model for language learning and instruction. *International Electronic Journal of Elementary Education* 6 (1), 165–184.
- Hildén, R., Härmälä, M., Rautopuro, J., Huhtanen, M., Puukko, M. & Silverström, C. 2014. *Kielten oppimistulokset perusopetuksen päättövaiheessa 2013 [Outcomes of language learning at the end of basic education in 2013]*. Informaatioaineistot 2014:1. Helsinki: Kansallinen koulutuksen arviointikeskus/Opetushallitus.
- Hosseini, M. & Nimehchisalem, V. 2021. Self-assessment in English language teaching and learning in the current decade (2010–2020): A systematic review. *Open Journal of Modern Linguistics*, 11, 854–872. 10.4236/ojml.2021.116066
- Jamrus, M. & Razali, A. 2019. Using self-assessment as a tool for English language learning. *English Language Teaching* 12 (11), 64–73. <https://doi.org/10.5539/elt.v12n11p64>
- Jansen, R. S., van Leeuwen, A., Janssen, J., Jak, S. & Kester, L. 2019. Self-regulated learning partially mediates the effect of self-regulated learning interventions on achievement in higher education: A meta-analysis. *Educational Research Review*, 28, 1–20. <https://doi.org/10.1016/j.edurev.2019.100292>
- Jin, L. 2021. The effects of using self-assessment on L2 speaking. *Frontiers in Educational Research* 4 (8), 85–93. 10.25236/FER.2021.040817
- Järvelä, S., Järvenoja, H., Malmberg, J., Isohäätä, J. & Sobocinski, M. 2016. How do types of interaction and phases of self-regulated learning set a stage for collaborative engagement? *Learning and Instruction*, 43, 39–51. <https://doi.org/10.1016/j.learninstruc.2016.01.005>
- Järvenoja, H., Järvelä, S., Törmänen, T., Näykki, P., Malmberg, J., Kurki, K., Mykkänen, A. & Isohäätä, J. 2018. Capturing motivation and emotion regulation during a learning process. *Frontline Learning Research*, 6 (3), 85–104. <https://doi.org/10.14786/flr.v6i3.369>
- Kohonen, V. & Korhonen, T. 2007. Student perspectives to the ELP: voices from the classrooms. In A. Koskensalo, J. Smeds, P. Kaikkonen & V. Kohonen (eds) *Foreign languages and multicultural perspectives in the European contexts*. Berlin: LIT Verlag, 239–267.
- Kotilainen, M.-R. 2015. *Itseohjautuvuuden tukeminen vieraan kielen etäopetuksessa: design-perustainen oppimisympäristön kehittämistutkimus perusasteen 5.–6. luokilla* [Supporting self-guided learning in distance learning of foreign languages. Development study of design-based learning environment in elementary school grades 5 and 6]. Acta Universitatis Lapponiensis 298. University of Lapland.
- Li, M. & Zhang, X. 2021. A meta-analysis of self-assessment and language performance in language testing and assessment. *Language Testing*, 38 (2), 189–218. <https://doi.org/10.1177/0265532220932481>
- Mäkipää, T. 2021. Students' and teachers' perceptions of self-assessment and teacher feedback in foreign language teaching in general upper secondary education – A case study in Finland. *Cogent Education*, 8 (1), 1–19. <https://doi.org/10.1080/2331186X.2021.1978622>
- Nicol, D. J. & Macfarlane-Dick, D. 2006. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31 (2), 199–218. <https://doi.org/10.1080/03075070600572090>
- Norrena, J. 2019. *Oman oppimisen kapteeni [Captain of one's own learning]*. Jyväskylä: PS-Kustannus.
- Oxford, R. L. 2017. *Teaching and researching language learning strategies: self-regulation in context* (2nd ed.). New York: Routledge.
- Panadero, E. 2017. A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, 8, 1–28. <https://doi.org/10.3389/fpsyg.2017.00422>

- Pollari, P., Salo, O.-P. & Koski, K. 2018. In teachers we trust – the Finnish way to teach and learn. *i.e.: inquiry in education* 10 (1), 1–16.
- Qasem, F. A. A. 2020. The effective role of learners' self-assessment tasks in enhancing learning English as a second language. *Arab World English Journal*, 11 (3), 502–514. <https://dx.doi.org/10.24093/awej/vol11no3.33>
- Salmela-Aro, K. (ed.) 2018. *Motivaatio ja oppiminen* [Motivation and learning]. Jyväskylä: PS-kustannus.
- Schunk, D. H. & Greene, J. A. 2018. Historical, contemporary, and future perspectives on self-regulated learning and performance. In D. H. Schunk & J. A. Greene (eds) *Handbook of self-regulation of learning and performance* (2nd ed.). New York: Routledge, 1–16.
- Theobald, M. 2021. Self-regulated learning training programs enhance university students' academic performance, self-regulated learning strategies, and motivation: A meta-analysis. *Contemporary Educational Psychology*, 66, 1–19. <https://doi.org/10.1016/j.cedpsych.2021.101976>
- Zimmerman, B. J. 1989. A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology*, 81 (3), 329–339. <https://doi.org/10.1037/0022-0663.81.3.329>
- Zimmerman, B. J. 1998. Developing self-fulfilling cycles of academic regulation: an analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman (eds) *Self-regulated learning: From teaching to self-reflective practice*. New York: The Guilford Press, 1–19.

Appendix 1:

Intervention process

Lesson number	Self-assessment related activity for the students
1	self-assessment grid consisting of can do-statements goal setting for Swedish lessons for the autumn
2	self-assessment survey (1st time)
3	pair/group and class discussion on self-assessment skills self-assessment task after an oral exercise homework: oral task using present perfect on what they had done in the past week + self-assessment task on oral skills and study habits
4	pair/group and class discussion on the homework tasks
5	general oral feedback from the teacher
6	oral task in which they interviewed each other in Swedish, preceded by a discussion on communication strategies thumb voting on how the students succeeded in the task
7	mid-term assessment and possible update of goals
8	homework: a read-aloud exercise (a text they had previously listened to and studied as a reading comprehension exercise) + self-assessment task on oral skills
9	pair/group and class discussion on the homework tasks
10	general oral feedback from the teacher
11	oral alias task preceded by a discussion on production strategies thumb voting on how the students succeeded in the task
12	self-assessment survey (2nd time)

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 40–59.

Päivi Virkkunen & Minnaleena Toivola
Helsingin yliopisto

Lukio-opiskelijoiden käsityksiä vieraan kielen ääntämisen opetuksessa saadusta palautteesta

Nostot

- Vieraan kielen ääntämisestä annettava palaute lisää oppijan tietoisuutta ja auttaa oppimisessa.
- Selvä enemmistö lukiolaisista arvostaa hyvää ääntämistä ja toivoo saavuttavansa syntyperäisen puhujan kaltaisen ääntämisen.
- Vastaajat toivovat saavansa opettajalta sekä positiivista että korjaavaa palautetta.
- Tulokset rohkaisevat opettajia antamaan myös korjaavaa ääntämispalautetta opiskelijoilleen.

Abstract

In foreign language teaching, feedback is an essential part of learning pronunciation. It raises the learner's awareness of the strengths and weaknesses of their pronunciation. Yet previous research has shown that teachers often feel uncomfortable when giving feedback on learners' pronunciation. We conducted a survey among Finnish high school students to find out their pronunciation targets and what kind of feedback they would like to receive. The survey revealed that a clear majority (80 %, n=1953) of respondents appreciated good pronunciation and wished for native-like pronunciation. They hoped to receive feedback from the teacher, and they understood the importance of corrective feedback for learning. Based on our findings we encourage teachers to provide positive and corrective pronunciation feedback to their students to promote effective learning.

Keywords: pronunciation, feedback, phonetics

Asiasanat: ääntäminen, palaute, fonetiikka

1 Johdanto

Suomalaiset kieltenopettajat pitävät ääntämistä tärkeänä kielitaidon osana (Tergujeff 2013) ja kokevat sen opettamisen tarpeelliseksi (Virkkunen & Toivola 2020). Palaute on olennainen osa ääntämisen oppimista, koska se lisää oppijan tietoisuutta hänen ääntämisensä vahvuuksista ja heikkouksista. Opettajat saattavat kuitenkin tuntea itsensä epävarmoiksi palautteenantajiksi (Virkkunen & Toivola 2020), minkä seurauksena suomalaiset lukiolaiset kokevat saavansa opettajiltaan hyvin niukasti palautetta (Mäkipää 2021). Toisaalta saatua palautetta ei välttämättä koeta opintojen kannalta hyödylliseksi (Pollari 2017). Opiskelija voi kokea myös opettajan tiedostamattaan antamat tai antamatta jättämät kommentit positiiviseksi tai negatiiviseksi palautteeksi (Huhtala 2015).

Kieltenopetuksessa ääntämismallin valinta ja oppimisen tavoitteen määrittäminen ovat peruskysymyksiä (Lintunen 2014). Lintusen ja Dufvan (2017) mukaan käsitteiden määrittely ei kuitenkaan ole yksinkertaista. He viittaavat mallilla vieraan kielen muotoon, joka on tietynlainen ihanneversio opiskellusta kielestä. Tavoite taas on konkreettinen, tietyn aikavälin toive saavutettavasta kielitaidosta. Opetuksessa ääntämisen malli on yleensä maantieteellisesti ja sosiaalisesti neutraali standardikieli, mutta kielenoppijan ääntämisen mallina voi olla kuka tahansa hänen itsensä valitsema puhuja. Ääntämisen tavoitteet opiskelijoilla vaihtelevat suuresti; osa toivoo voivansa ääntää vierasta kieltä mahdollisimman samalla tavalla kuin äidinkielliset puhujat, kun taas toisille riittää se, että tulee ymmärretyksi (Tergujeff 2013; Ilola 2018). Tietääksemme laajempaa tutkimusta lukio-opiskelijoiden omista ääntämistavoitteista ei ole toistaiseksi tehty.

Teimme suomalaisille lukiolaisille suunnatun kyselyn selvittääksemme, minkälaisia tavoitteita opiskelijoilla on vieraan kielen ääntämisensä suhteen. Tässä tutkimuksessa

myös ruotsin ajatellaan olevan vieras kieli, vaikka sillä onkin Suomessa virallisesti toisen kielen asema. Selvitimme myös, millaista palautetta lukiolaiset haluaisivat saada ääntämisestään, miten he suhtautuvat eri tavoin annettuun palautteeseen ja millaista palautetta he pitävät toimivana.

1.1 Suullisen kielitaidon opetus ja arviointi

Opiskelija on lukion aloittaessaan käynyt läpi peruskoulun aikana vähintään yhden pitkän (A1) ja yhden keskipitkän (B1) kielen oppimäärät, joista toinen on kotimainen kieli (suomi tai ruotsi) ja toinen jokin vieras kieli tai saamen kieli. Lukiossa näiden kielten opiskelu jatkuu, ja opiskelija voi halutessaan aloittaa myös uuden vieraan kielen opiskelun.

Vieraiden kielten opetussisällöt määritellään lukion opetussuunnitelman perusteissa (Opetusministeriö 2019). Niissä tuodaan esille, että suullinen kielitaito on tärkeää ja ääntämiseen tulee kiinnittää huomiota alusta alkaen. Kielten A-oppimäärän keskeisistä sisällöistä löytyvät mm. kohdekielen äänneiden muodostuminen ja puheen tuottaminen sekä kohdekielen variantit ja vertailu muihin kieliin (Opetusministeriö 2019: 186). Ylioppilaskirjoituksissa painotetaan kuitenkin kirjallista kielitaitoa. Koska arviointi ohjaa opetusta (Pohjala 2004), tämä vaikuttaa myös kielten oppimissisältöjen painotukseen opetuksessa. Opettajat ovatkin aikaisemmissa tutkimuksissa (Tergujeff 2013; Virkkunen & Toivola 2020) tuoneet esille, että tunneilla ei ole aikaa keskittyä ääntämisen opettamiseen. Opettajilta saattaa myös puuttua tietoa kielten välisistä foneettisista eroista sekä siitä, miten ääntämistä opetetaan. Ääntämisen opettamista helpottaa tieto fonetiikasta, äänneistä sekä prosodisista piirteistä, kuten intonaatiosta, rytmistä ja kestoista (Virkkunen & Toivola 2020).

Lukion opetussuunnitelman perusteissa (Opetushallitus 2019) kehoitetaan opettajaa tekemään oppiminen näkyväksi arvioimalla opiskelijan edistymistä. Lukiolain (714/2018, 37 § 4 mom.) mukaan ”Kielten opetuksessa opiskelijalta arvioidaan kielitaidon muiden osa-alueiden lisäksi suullinen kielitaito”. Arviointi on oppijalle asetettujen tavoitteiden ja saavutettujen tulosten välistä vertailua. Summatiivinen arviointi viittaa oppimista mittaavaan loppuarviointiin, kun taas formatiivinen arviointi tapahtuu oppimisen aikana, oppimista varten (Bloom ym. 1971). Olennainen osa formatiivista arviointia on palaute, jonka avulla opiskelija oppii suuntaamaan harjoittelunsa oikeisiin asioihin.

1.2 Palaute kielenoppimisessa

Palaute on tärkeä osa kielenoppimisprosessia (Spada & Lightbown 1993). Vieraan kielen opetuksessa annettu palaute voi olla positiivista tai korjaavaa. Pedagogisessa mielessä positiivinen palaute antaa opiskelijalle myönteistä tukea ja lisää motivaatiota

opiskelun jatkamiseen (Harackiewicz 1979; Oroujlou & Vahedi 2011; Thohir 2017). Mäkipää ja Hildénin (2021) tutkimuksessa selvitettiin suomalaisten lukiolaisten käsityksiä kielenopettajien antamasta palautteesta. Tulosten mukaan lukiolaiset (n=282) toivoivat kielenopettajiltaan erityisesti korjaavaa palautetta, mutta palautetta ei nähty tärkeänä osana opettajan arviointityötä.

Yleisimmin tutkimuksissa viitataan viiteen korjaavan palautteen teemaan: 1) pitääkö opiskelijoiden virheitä korjata, 2) mitkä virheet korjataan, 3) kuka korjaa virheet, 4) milloin virheet korjataan ja 5) miten virheet korjataan? (Hendrickson 1978; Lyster & Ranta 1997; Ellis 2009.) Näiden kysymysten lisäksi tutkimus on laajentunut pohtimaan, millainen korjaava palaute on tehokkainta ja mitä vaikutuksia korjaavalla palautteella on esimerkiksi oppijan tunteisiin (Wu ym. 2021).

Lyster ja Rannan (1997, suomenkieliset termit Hakola 2008) mukaan kuusi korjaavan palautteen antamistapaa ovat suora korjaaminen (*explicit correction*), uudelleenmuotoilu (*recast*), selvennyspyyntö (*clarification request*), metakielellinen korjaus (*metalinguistic feedback*), johdattelu (*elicitation*) ja toistaminen (*repetition*). Uudelleenmuotoilu on yleinen korjaamistapa erityisesti kielenopetuksessa. (Panova & Lyster 2002; Sheen 2004; Lyster & Mori 2006; Yoshida 2008.) Opettaja korjaa implisiittisesti oppijan tuottaman virheellisen muodon, mutta ei selitä virhettä tarkemmin. Jos oppija ei saa opettajan korjauksesta selville, missä virhe oli, oppimista ei ehkä tapahdu (Chin 2006; Ahangari 2014; Lee 2016). Metakielellinen korjaus sisältää foneettista tietoa tai muita kommentteja, joiden avulla opettaja johdattelee oppijaa virheen havaitsemisen jälkeen oikeaan kielimuotoon (Lyster & Ranta 1997).

1.3 Vieraan kielen ääntäminen

Vieraan kielen ääntämisen oppimiseen vaikuttavat monet tekijät, kuten esimerkiksi ikä ja motivaatio (ks. esim. Ullakonoja & Dufva 2016). Oppijan tulee olla tietoinen oman ääntämisensä vahvuuksista ja ongelmista, jotta hän tietää, mihin suunnata harjoittelunsa saavuttaakseen tavoittelemansa ääntämistaidon (Tergujeff ym. 2019). Ääntämisen opetuksessa palautteenantaja vahvistaa positiivisella palautteella sitä, että oppija on tuottanut kohdekielen mukaisia äänneitä, intonaatiota ja painoituksia, kun taas korjaavalla palautteella tuodaan esille, että niissä on virheitä. Kielitietoinen opetus, foneettinen tieto ja palautteen mukainen harjoittelu auttavat oppijaa paremmaksi ääntäjäksi (esim. Kennedy & Trofimovich 2010; Couper 2011; Saito 2011; Lintunen 2014; Carey ym. 2015).

Ääntäminen on usein henkilökohtaisimmaksi koettu kielitaidon osa-alue (Lintunen 2014) ja opiskelijat voivat arastella puhumista, koska he pelkäävät tekevänsä virheen tai joutuvansa luokkatovereiden arvioinnin kohteeksi. Ääntämisen arviointi voidaan myös sekoittaa oman äänenlaadun arviointiin (Setter & Jenkins 2005). Ääntäminen ja siitä saatu palaute voi herättää vahvoja tunteita ja jopa hetkellistä osaamattomuuden tunnetta (Baran-Łucarz 2013; Pekrun 2016; Lintunen ym. 2017).

Kielenopetuksessa palaute, joka saadaan opettajalta, koetaan usein hyödylliseksi (Lintunen ym. 2017; Martin & Sippel 2021). Palautetta voi saada opettajan lisäksi myös toisilta oppijoilta (Ketonen 2021; Martin & Sippel 2022). Monet oppijat ovat kokeneet vertaispalautteen auttavan oman ääntämisen oppimisessa, mutta aiheesta on myös ristiriitaisempia tuloksia, sillä onnistunut palautteenanto vaatii oikeanlaista ohjeistusta ja harjoittelua (Martin & Sippel 2021).

1.4 Aiempi tutkimus ääntämispalautteesta

Kielenopetuksessa annettavan korjaavan palautteen laatua ja ajankohtaa on tutkittu jonkin verran (ks. esim. Huang & Jia 2016; Ölmezer-Öztürk & Öztürk 2016; Phuong 2022). Usein tutkimuksissa käsitellään kuitenkin kielitaitoa yleisesti, eikä ääntämistä tarkastella erikseen. Tutkimusta korjaavasta ääntämispalautteesta ja siihen liittyvistä käsityksistä on tehty opettajan näkökulmasta (ks. esim. Baker & Burri 2016; Couper 2019). Oppijan näkökulma jää kuitenkin usein puuttumaan.

Positiivisen palautteen antaminen tuntuu opettajista helpolta, mutta korjaamista vältetään etenkin luokkahuonetilanteissa (Virkkunen & Toivola 2020). Korjaavan palautteen antamisen parhaaseen ajankohtaan ja tapaan ei ole olemassa yhtä oikeaa vastausta (Dlaska & Krekeler 2013). Opettajat kokevatkin usein epävarmuutta siitä, miten ääntämispalautetta tulisi antaa ja millainen vaikutus sillä on oppijoihin (Lintunen & Mäkilähde 2017; Couper 2019). Suomalaiset kielenopettajaopiskelijat tuovat esille toiveensa reilusta, tasapuolisesta ja kannustavasta arvioinnista, mutta toteavat samalla, ettei arviointia tai palautteen antamista harjoitella riittävästi koulutuksen aikana (Huhtala 2015).

Suomalaiset opiskelijat ovat aiemmissa tutkimuksissa toivoneet korjaavaa palautetta ääntämisestään. Ilolan (2018) 9-luokkalaisia oppilaita tutkinut laadullinen tapaustutkimus osoitti, että he halusivat oppia ääntämään paremmin ja toivoivat saavansa myös korjaavaa palautetta. Lintusen ym. (2017) tutkimuksessa englannin yliopisto-opiskelijat kokivat korjaavan palautteen hyödyllisenä ja oppimista edistävänä, ja he toivoivat saavansa korjaavaa palautetta erityisesti opettajilta. Lukio-opetuksessa annettua vieraan kielen ääntämispalautetta ei ole juuri tutkittu. Korjaava palaute näyttäytyy kuitenkin tutkimuksen mukaan oleellisena tekijänä ääntämisen oppimisessa (ks. Saito 2012).

2 Tutkimuskysymykset ja -aineisto

Tässä tutkimuksessa selvitetään suomalaisten lukiolaisten käsityksiä hyvän ääntämisen tärkeydestä ja ääntämisen opetuksessa saatavasta palautteesta.

Tutkimuskysymykset ovat:

1. Miten tärkeänä opiskelijat pitävät hyvää ääntämistä ja haluaisivatko he, että heidän ääntämisensä olisi mahdollisimman samanlaista kuin äidinkielisillä puhujilla?
2. Minkälaista palautetta opiskelijat toivoisivat saavansa ääntämisestään? Keneltä, milloin ja millaisessa muodossa se pitäisi saada?
3. Miten opiskelijat ajattelevat positiivisen ja korjaavan palautteen vaikuttavan heidän ääntämisen oppimiseensa?
4. Miten opettajan tulisi antaa korjaavaa ääntämispalautetta?

2.1 Aineisto ja analyysimenetelmät

Tutkimus toteutettiin verkkokyselynä marras–joulukuussa 2021 Helsingin, Kuopion ja Turun suomenkielisissä päivälukioissa. Kysely lähetettiin kaikkiin kyseisten kaupunkien lukioihin ja koulut saivat itse päättää osallistumisestaan tutkimukseen. Opiskelijoille ja heidän huoltajilleen lähetettiin etukäteen tiedote tutkimuksesta. Anonyymiin kyselyyn vastattiin koulupäivän aikana luokanvalvojan tai kielten tunneilla, joten kyselyn vastaajiksi eivät valikoituneet vain ne opiskelijat, jotka muutenkin pitävät kieltenopiskelusta.

Kyselylomake oli Helsingin yliopiston e-lomake. Sen pohjana oli kysely (Wu 2020), jota muokattiin tutkimukseemme sopivaksi. Lomakkeessamme oli 72 väitettä tai kysymystä, jotka jakautuivat yhdeksään pääteemaan: 1) vastaajan taustatiedot (kuten ikä, sukupuoli, vuosikurssi ja kyselyyn valikoitunut kieli), 2) ääntämisen oppimisprosessi, 3) pitäisikö ääntämistä korjata, 4) mitä virheitä korjataan, 5) kuka korjaa (opettaja vai luokkatoverit) ja milloin korjataan, 6) miten korjataan (eksplisiittinen vai implisiittinen palaute) sekä 7) millaisia tunteita positiivinen ja korjaava palaute herättävät. Lisäksi vastaajat arvioivat 8) opettajalta ja luokkatovereilta saadun positiivisen ja korjaavan palautteen vaikutusta ääntämisen oppimiseensa. Vastaajat myös kertoivat mielipiteensä siitä, 9) miten opettajan tulisi antaa korjaavaa palautetta.

Kyselyn väittämiin tai kysymyksiin vastattiin viisiportaisella Likert-asteikolla. Lomakkeessa oli annettu sanalliset selitykset vain asteikon ääripäille, koska emme halunneet, että annetut termit ohjaisivat vastauksia. Avoimissa kysymyksissä vastaajat saivat kertoa tarkemmin, miten oman ääntämisen oppimiseen vaikuttaa se, että siitä saa positiivista tai korjaavaa palautetta. Avovastausten osalta noin 10 % vastaajista jätti vastauksen tyhjäksi.

Lomakkeen alussa vastaajille kerrottiin, mitä positiivisella ja korjaavalla palautteella tarkoitettiin tässä kyselyssä. Positiivista palautetta kuvattiin seuraavasti: "Positiivinen palaute tarkoittaa sitä, että olet ääntänyt oikein ja opettaja huomioi sen sanomalla esimerkiksi: Hienoa, äänsit sanan juuri oikein! tai Hyvä, tuo on juuri oikeanlainen kysymysintonaatio!". Korjaavaa palautetta kuvattiin seuraavasti: "Korjaava palaute tarkoittaa tässä kyselyssä sitä, että olet tehnyt ääntämisvirheen (esimerkiksi ääntänyt väärän vokaalin, konsonantin tai painottanut väärää sanan osaa, jolloin sanan merkitys

on muuttanut) ja opettaja korjaa virheesi.” Näillä esimerkillä haluttiin myös kiinnittää vastaajien huomio siihen, että ääntämiseen sisältyy sekä äännetason että prosodisen tason ilmiöitä, jotka molemmat vaikuttavat ääntämiseen.

Tässä artikkelissa raportoimme korjaavaan palautteeseen liittyvän kokonaisuuden, jossa käsitellään vastauksia teemoihin 1, 5, 6, 8 ja 9. Tuloksia analysoitiin sekä määrällisiä että laadullisia menetelmiä käyttäen. Määrällisten tulosten kuvailussa käytettiin frekvenssi- ja prosenttijakaumia sekä keski- ja hajontalukuja. Avoimia vastauksia tarkasteltiin sisällönanalyysin keinoin, ja tulokset muodostettiin aineiston perusteella. Analyysimenetelmänä oli määrällinen sisällönerittely (Tuomi & Sarajärvi 2017), jonka avulla selvitettiin, miten toistuvasti tietyt seikat ilmenivät aineiston avoimissa vastauksissa. Vastaukset taulukoitiin teemoittain ja tutkimusraporttiin lainatut esimerkit valikoituivat edustamaan neljää tämän artikkelin kannalta keskeistä teemaa: positiivisen ja korjaavan palautteen saamista, palautteenantotapoja sekä ääntämisen tärkeyttä osana kielitaitoa. Kommentit ovat alkuperäisessä muodossaan, ja niistä on korjattu vain yksittäiset, ilmeiset lyöntivirheet.

2.2 Vastaajien taustatiedot

Kyselyyn vastasi 1953 äidinkieleltään suomenkielistä lukiolaista. Heistä 900 kävi lukiota Kuopiossa, 551 Helsingissä ja 501 Turussa. Yhden vastaajan asuinpaikka ei ole tiedossa. Vastaajista naisia oli 1225 (62,7 %), miehiä 593 (30,4 %) ja muita 63 (3,2 %), 72 vastaajaa (3,7 %) ei ilmoittanut sukupuoltaan.

Vastaajat olivat syntyneet vuosina 1999–2006 (mediaani 2004). Opiskelijat jakautuivat lukion vuosikursseille seuraavasti: 1. vuosikurssi 874 vastaajaa (44,8 %), 2. vuosikurssi 610 vastaajaa (31,2 %), 3. vuosikurssi 453 vastaajaa (23,2 %) ja 4. vuosikurssi 16 vastaajaa (0,8 %).

3 Tulokset

Kerromme ensin vastaajien kyselyssä valitsemista kielistä sekä siitä, kuinka paljon he olivat mielestään saaneet ääntämisopetusta ja palautetta ääntämisestään. Toisessa alaluvussa selvitämme, kuinka tärkeäksi vastaajat kokivat hyvän ääntämisen itselleen ja toivoivatko he sen olevan samanlaista kuin kieltä äidinkielenään puhuvilla. Kolmannessa alaluvussa kerromme, keneltä ja millaista palautetta opiskelijat toivoivat ja milloin he halusivat sitä saada. Neljäs alaluku tarkastelee opiskelijoiden ajatuksia palautteen hyödyllisyydestä sen laadun (korjaava vai positiivinen palaute) suhteen.

Kahdessa viimeisessä alaluvussa käsitellään avovastausten kommentteja palautteen merkityksestä ja hyvästä palautteesta.

3.1 Vastaajien kielivalinnat sekä ääntämisen opetuksen ja saadun palautteen määrä

Opiskelijat saivat itse valita yhden lukiossa opiskelemansa kielen, jonka perusteella he vastasivat kyselyyn (ks. taulukko 1). Valinta kohdistui useimmin englannin kieleen. Koska opiskelija sai itse valita, minkä kielen mukaan hän vastasi, tämä vaikutti samalla välillisesti siihen, miten pitkään hän oli opiskellut kyseistä kieltä. Yli puolet (60,3 %) vastauksista kohdistui ensimmäisenä aloitettuun pitkään kieleen, jota oli lukion alkuun mennessä opiskeltu keskimäärin seitsemän vuotta. Lukiossa alkavan kielen perusteella vastasi noin 12 % opiskelijoista.

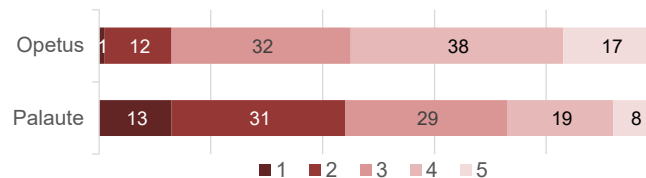
TAULUKKO 1. Kieli, jonka perusteella lukiolaiset (n=1953) vastasivat kyselyyn. Vastausten lukumäärät ja prosenttiosuudet kaikista vastauksista.

Kieli	vastausten lukumäärä	% kaikista vastauksista
englanti	1251	64,1
ruotsi	337	17,3
ranska	121	6,2
saksa	103	5,3
espanja	73	3,7
venäjä	24	1,2
japani	20	1,0
kiina	8	0,4
italia	8	0,4
muu	8	0,4
yhteensä	1953	100

Opiskelijoita pyydettiin arvioimaan, miten paljon heille oli opetettu ääntämistä kyselyyn valitun kielen oppitunneilla. Asteikolla 1 (ei lainkaan) – 5 (täysin riittävästi) kaikkien vastausten keskiarvo oli 3,58 (keskihajonta, jatkossa kh, 0,95). Vastaajista 17,1 % ajatteli, että heille oli opetettu ääntämistä täysin riittävästi ja 1,2 % vastasi, että ääntämistä ei oltu opetettu lainkaan. Opiskelijoita pyydettiin myös arvioimaan, miten

paljon he ovat saaneet palautetta omasta ääntämisestään kyseisen kielen oppitunneilla. Asteikolla 1 (ei lainkaan) – 5 (täysin riittävästi) kaikkien vastausten keskiarvo oli 2,78 (kh 1,13). Alle kolmannes vastaajista (27,1 %) koki, että oli saanut riittävästi (vastaus 4 tai 5) palautetta ääntämisestään. Lähes puolet vastaajista (43,8 %) koki, ettei ollut saanut riittävästi (vastaus 1 tai 2) palautetta ääntämisestään (ks. kuvio 1).

Väitteeseen ”Mielestäni on normaalia, että vieraan kielen opiskelija tekee virheitä oppiessaan ääntämistä” kaikkien vastausten keskiarvo oli 4,79 (kh 0,53). Selkeä enemmistö, 1630 vastaajaa, oli väitteestä täysin samaa mieltä ja vain 11 täysin tai lähes eri mieltä. Opiskelijat siis ymmärtävät, että virheiden tekeminen kuuluu vieraan kielen ääntämisen harjoitteluun.



KUVIO 1. Vastaajien (n=1953) arvio siitä, kuinka paljon heille on opetettu ääntämistä ja kuinka paljon he ovat saaneet palautetta ääntämisestään kyselyyn valitun kielen oppitunneilla. Asteikko 1 (ei lainkaan) – 5 (täysin riittävästi). Kuvion luvut ovat prosenttiosuuksia kaikista vastauksista.

3.2 Oma ääntämistaito ja sen tärkeys

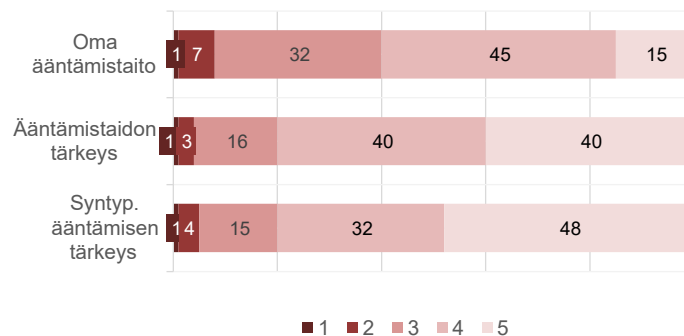
Opiskelijoita pyydettiin arvioimaan omaa ääntämistaitoaan asteikolla 1 (heikko) – 5 (erinomainen). Kaikkien vastausten keskiarvo oli 3,64 (kh 0,87). Lähes puolet vastaajista arvioi ääntämistaitoaan vaihtoehdolla 4 (ks. kuvio 2).

Ensimmäisessä tutkimuskysymyksessä haluttiin selvittää, miten tärkeänä opiskelijat pitivät hyvää ääntämistä. Asteikolla 1 (ei lainkaan tärkeää) – 5 (erittäin tärkeää) kaikkien vastausten keskiarvo oli 4,22 (kh 0,90): *Ääntämisen ja puhumisen pitäisi olla suuremmissa osassa kielten opiskelussa, sillä koen sen olevan vaikein ja tärkein osa kielen osaamista.* (vastaaja 1894) Vastaajista 80 % piti hyvää ääntämistä itselleen tärkeänä (vastaus 4 tai 5, ks. kuvio 2). *Mitä hyötyä on osata lukea ja kirjoittaa täydellisesti, jos ei saa sanaa suustaan, kun on tosi paikka kyseessä?* (vastaaja 1873) Vain 4 % vastaajista ajatteli, että ääntäminen ei ole heille juurikaan tärkeää (vastaus 1 tai 2).

Lisäksi ensimmäisessä tutkimuskysymyksessä haluttiin selvittää opiskelijoiden toiveita ääntämisensä suhteen. Opiskelijoita pyydettiin arvioimaan väitettä ”Haluaisin, että ääntämiseni on mahdollisimman samanlaista kuin kieltä äidinkielenään puhuvilla”, asteikolla 1 (täysin eri mieltä) – 5 (täysin samaa mieltä). Kaikkien vastausten keskiarvo oli 4,22 (kh 0,90). Edellisen kysymyksen hyvä ääntäminen ei vielä määrittele ääntämisen

49 LUKIO-OPISKELIJOIDEN KÄSITYKSIÄ VIERAAN KIELEN
ÄÄNTÄMISEN OPETUKSESSA SAADUSTA PALAUTTEESTA

tasoa, sillä jokaisella on todennäköisesti oma määritelmä sille, mikä on itselle hyvää ääntämistä. Sen sijaan ”mahdollisimman samanlainen ääntäminen kuin äidinkieli-
sellä puhujalla” on objektiivisempi malli, jota kohti omassa harjoittelussa voi pyrkiä. Vastaajista 80 % toivoi, että heidän ääntämisensä olisi mahdollisimman samanlaista kuin kieltä äidinkielenään puhuvilla (vastaus 4 tai 5, ks. kuvio 2).

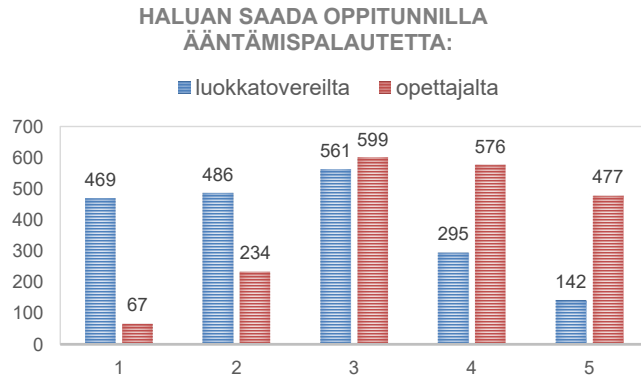


KUVIO 2. Opiskelijoiden (n=1953) arvio omasta ääntämistaidosta asteikolla 1 (heikko) – 5 (erinomainen). Opiskelijoiden arvio hyvän ääntämisen tärkeydestä ja toive siitä, että oma ääntäminen olisi mahdollisimman samanlaista kuin kieltä äidinkielenään puhuvilla on arvioitu asteikolla 1 (ei lainkaan tärkeää) – 5 (erittäin tärkeää). Kuvion luvut ovat prosenttiosuuksia kaikista vastauksista.

3.3 Opiskelijoiden käsityksiä palautteen saamisesta

Toisessa tutkimuskysymyksessä halusimme selvittää opiskelijoiden käsityksiä opettajan ja luokkatovereiden antamasta ääntämispalautteesta, milloin ääntämisen korjaamisen olisi hyvä tapahtua sekä millaista palautetta opiskelijat toivovat saavansa.

Väitteeseen ”Haluan saada oppitunnilla ääntämispalautetta opettajalta.” kaikkien vastausten keskiarvo oli 3,59 (kh 1,08). Luokkatovereilta saadusta ääntämispalautteesta esitettyyn väitteeseen kaikkien vastausten keskiarvo oli 2,57 (kh 1,21). Katso vastausten jakaumat kuviossa 3.



KUVIO 3. Kaikkien vastausten (n=1953) jakaumat väitteisiin ”Haluan saada oppitunnilla ääntämispalautetta opettajalta/luokkatovereilta”, johon vastattiin asteikolla 1 (täysin eri mieltä) – 5 (täysin samaa mieltä). Kuvion luvut ovat vastausten lukumäärät.

Palautteen ajankohtaa koskevat väitteet ja niihin saadut vastaukset ovat taulukossa 2. Eniten opiskelijat toivoivat palautetta, jonka opettaja antaa opiskelijan lähettämästä äänitetystä suullisesta tehtävästä. Luokassa saadusta korjaavasta palautteesta mieluisinta oli koko ryhmälle yleisesti annettu palaute, joka ei yksilöi puhujia. Myös mahdollista henkilökohtaista korjaavaa palautetta toivottiin enemmän oppitunnilla kuin tunnin jälkeen kahden kesken. On hyvä huomata, että opiskelijat eivät vastauksissaan laittaneet eri palautteensaamistapoja paremmuusjärjestykseen, joten kaikki edellä mainitut tavat ovat saaneet kannatusta. Pieni osa vastaajista toivoi, että opettaja ei anna korjaavaa palautetta, vaikka havaitsisi virheen ääntämisessä.

TAULUKKO 2. Palautteen ajankohtaan liittyviä väitteitä, joihin vastattiin asteikolla 1 (täysin eri mieltä) – 5 (täysin samaa mieltä). Kaikkien vastaajien (n=1953) vastausten keskiarvot ja keskihajonnat.

	keskiarvo	keskihajonta
Toivon, että opettaja antaa ääntämisestäni korjaavaa palautetta heti oppitunnilla.	3,29	1,13
Toivon, että opettaja antaa ääntämisestäni korjaavaa palautetta tunnin jälkeen kahden kesken.	2,90	1,31
Toivon, että opettaja antaa ääntämispalautetta, kun olen palauttanut hänelle äänittämäni suullisen tehtävän.	4,08	1,03
Toivon, että opettaja antaa luokassa yhteistä palautetta, johon kokoa yleisimmät ääntämisvirheet.	3,55	1,14

51 LUKIO-OPISKELIJOIDEN KÄSITYKSIÄ VIERAAN KIELEN
ÄÄNTÄMISEN OPETUKSESSA SAADUSTA PALAUTTEESTA

	keskiarvo	keskihajonta
Toivon, että opettaja ei anna minulle korjaavaa palautetta, vaikka havaitsisi virheen ääntämisessäni.	1,89	0,98

Palautteen laatua koskevat väitteet ja niihin saadut vastaukset ovat taulukossa 3. Opiskelijat vastasivat, että omien ääntämisvirheiden korjaamista helpottaa, jos itsellä on tietoa vieraiden kielten ääntämisen oppimisesta tai opiskellun kielen ääntämisestä. Väite, jossa toivottiin, että opettaja kertoo ensin virheen laadusta ja vasta sen jälkeen antaa mallin oikeaan ääntämiseen, sai enemmän kannatusta kuin väite, jossa opettaja korjaa ääntämisvirheen antamalla vain suoraan oikean ääntämissälin.

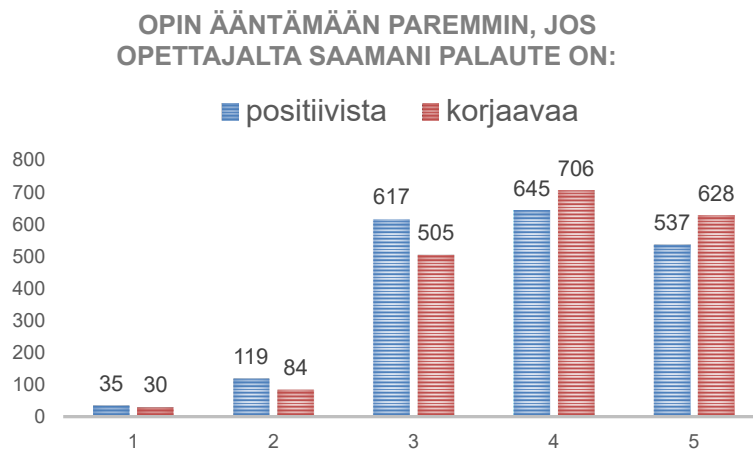
TAULUKKO 3. Palautteen laatuun liittyviä väitteitä, joihin vastattiin asteikolla 1 (täysin eri mieltä) – 5 (täysin samaa mieltä). Kaikkien vastaajien (n=1953) vastausten keskiarvot ja keskihajonnat.

	keskiarvo	keskihajonta
Minun on helpompaa oppia ääntämään oikein, jos ymmärrän, mitkä seikat vaikuttavat vieraan kielen ääntämisen oppimiseen.	3,97	0,96
Minun on helpompaa korjata ääntämistäni, jos ymmärrän, mitkä seikat vaikuttavat opiskelemani kielen ääntämiseen.	4,00	0,94
Toivon, että opettaja korjaa ääntämisvirheeni antamalla suoraan mallin oikeaan ääntämiseen.	3,37	1,01
Toivon, että opettaja korjaa ääntämisvirheeni kertomalla ensin millainen virhe oli kyseessä, ja sen jälkeen antaa mallin oikeaan ääntämiseen.	3,86	0,99
Toivon, että opettaja kertoo millaisen ääntämisvirheen tein (esim. väärä vokaali tai paino väärällä tavulla), mutta antaa minun etsiä itse oikean ääntämissämuodon.	3,21	1,13

3.4 Positiivisen ja korjaavan palautteen vaikutus omaan ääntämiseen

Kolmannessa tutkimuskysymyksessä halusimme selvittää, miten opiskelijat ajattelivat positiivisen ja korjaavan palautteen vaikuttavan heidän ääntämisen oppimiseensa. Väitteiden ”Opin ääntämään paremmin, jos saan positiivista/korjaavaa palautetta

opettajalta” vastausten keskiarvot olivat lähellä toisiaan (positiivinen palaute: ka 3,78, kh 0,98 ja korjaava palaute: ka 3,93, kh 0,94). Väitteeseen ”Opin ääntämään paremmin, jos saan opettajalta korjaavaa palautetta” vastasi vaihtoehdoilla 4 ja 5 yhteensä 1334 opiskelijaa (68,3 % vastauksista). Väitteeseen positiivisen palautteen osalta vaihtoehdot 4 ja 5 saivat yhteensä 1182 vastausta (60,5 % vastauksista). Vastausten jakaumat ovat kuviossa 4.



KUVIO 4. Vastausten (n=1953) jakaumat ja lukumäärät väitteisiin: Opin ääntämään paremmin, jos saan positiivista/korjaavaa palautetta opettajalta. Vastausasteikko 1 (täysin eri mieltä) – 5 (täysin samaa mieltä). Kuvion luvut ovat vastausten lukumäärät.

3.5 Positiivisen ja korjaavan palautteen vaikutus omaan ääntämiseen avovastausten mukaan

Avovastauksissa positiivinen palaute nähtiin yleisesti mielialan kohottajana. Opiskelijat kertoivat, että positiivinen palaute motivoi harjoitteluun, lisää itsevarmuutta ja antaa uskallusta puhua enemmän. Toisaalta opiskelijat ymmärsivät, että positiivisen palautteen lisäksi tarvitaan korjaavaa palautetta antamaan uutta tietoa, joka auttaa ääntämisen oppimisessa. *Verrattuna korjaavaan palautteeseen, positiivinen palaute ei anna paljoa. Sitä on tietysti aina kiva kuulla, mutta minä en ainakaan hyödy siitä.* (vastaaja 879)

Korjaava palaute auttoi harjoittelussa sekä opetti kiinnittämään huomiota omaan ääntämiseen ja ääntämisvirheisiin: *Virheet ovat osa oppimista ja korjaavan palautteen avulla voi oppia paremmin ja helpommin. Korjaava palaute voi auttaa löytämään aiheen jossa tarvitset enemmän harjoitusta.* (vastaaja 2104). Opettajalta saatu korjaava palaute innostaa yrittämään uudelleen, uudella tiedolla (vastaaja 1825) ja motivoi harjoittelemaan enemmän, voittamaan itsensä (vastaaja 148) Korjaavan palautteen saaminen

ei aina ollut mieluista, mutta toisaalta opiskelijat ymmärsivät, että palaute auttaa oppimaan: *Harmittaa, että kieltä on pitänyt korjata, mutta se on tärkeää oppimisen kannalta.* (vastaaja 686)

3.6 Hyvä korjaava palaute

Avovastauksissa opiskelijat pohtivat, miten opettaja voisi parhaiten antaa korjaavaa palautetta. Vastaajien mielestä korjaava palaute olisi hyvä antaa tasapainossa positiivisen palautteen kanssa: *Kertoisi ystävällisesti ja ymmärtäväisesti, mikä virhe ääntämisessä on kannustaen samalla puhumaan lisää, esim. ns. hampurilaismalli (kehu, korjaus, kehu).* (vastaaja 2040) Opiskelijat arvostivat opettajan rehellisyyttä ja suoruutta korjaustilanteessa: *Mieluummin opin ääntämään oikein kuin että en opi koska korjauksia ei tule.* (vastaaja 882)

Opiskelijat kiinnittivät palautteen asiasisällön lisäksi huomiota tapaan, jolla palaute annettiin. Suora ja korjaava palautekaan ei tuntunut pahalta, kun se annettiin asiallisesti ja kannustavalla mielellä, eikä syyttäen tai moittien. *Parannettavista seikoista pitäisi keskustella rohkaisevaan sävyyn.* (vastaaja 2150) Vastauksissa toistuivat hyvän palautteen kuvailut: ystävällinen, lempeä, rakentava, positiivinen. Opettajan äänensävykin vaikutti siihen, miten palaute vastaanotetaan: *Sanoo vaan, että mitä tein väärin, mutta ystävälliseen äänensävyyn.* (vastaaja 1077)

Monet opiskelijat toivoivat henkilökohtaista palautetta, mutta he tiedostivat, että oma palaute voi auttaa myös luokkatovereita, jotka mahdollisesti havaitsevat samoja virheitä omassa puheessaan: *Palautetta saa antaa vaikka koko luokan edessä, sillä muutkin hyötyvät opettajan ohjeistuksesta.* (vastaaja 1906) Oli myös lohduttavaa kuulla, että ääntäminen voi olla muillekin vaikeaa: *Opettaja voisi myös kertoa, että virhe on yleinen ja se tulee monelle kieltä opiskelevalle.* (vastaaja 1820)

Palaute olisi hyvä saada *ohimennen* (useita vastauksia) ja *sulavasti tunnin aikana, tekemättä siitä isoa numeroa* (esim. tunnin jälkeen puhumisesta tulisi helposti fiilis, että olen mokannut tosi isosti). (vastaaja 1062) Henkilökohtainen palaute tuntui turvalliselta, koska luokkatoverit eivät ole kuulemassa omaa puhetta ja siitä saatua palautetta: *Ne [kotona äänitettävät tehtävät] ovat mukavia tehtäviä ja niistä saa myös palautetta. Jos luokan edessä puhuminen jännittää tämä voisi helpottaa sitä.* (vastaaja 896)

4 Pohdinta

Tässä tutkimuksessa selvitimme, miten tärkeänä suomalaiset lukio-opiskelijat pitävät hyvää vieraan kielen ääntämistä ja miltä he haluavat itse kuulostaa sekä selvitimme heidän ajatuksiaan ja toiveitaan ääntämisen opetuksessa saatavasta palautteesta. Selkeä enemmistö (80 %) kyselyyn vastanneista lukiolaisista arvosti hyvää ääntämistä ja toivoi, että heidän ääntämisensä olisi mahdollisimman samanlaista kuin äidinkielisillä

puhujilla. Aiemman tutkimuksen (Virkkunen & Toivola 2020) mukaan kieltenopettajat suhtautuvat ääntämisen korjaamiseen luokkatilanteessa välittelevästi, mutta tämän tutkimukset tulokset osoittivat, että opiskelijat toivoivat saavansa ääntämisestään positiivisen palautteen lisäksi korjaavaa palautetta. Kyselyyn vastanneet lukiolaiset tiedostivat, että vaikka positiivinen palaute antaa rohkeutta puhua ja ylläpitää taitoja, he oppivat ääntämään paremmin, kun he saavat opettajalta myös korjaavaa palautetta.

Vastaajat toivoivat saavansa palautetta enemmän opettajalta kuin luokkatovereilta, samoin kuin Lintusen ym. (2017) englannin yliopisto-opiskelijoita tarkastelleessa tutkimuksessa. Enemmistö vastaajista toivoi palautetta äänitetystä tehtävästä. Tällainen tehtävä on kuitenkin luonteeltaan erilainen kuin luokassa opettajan kanssa käytävä keskustelu, koska kotona luokkatoverit eivät ole kuulemassa omaa ääntämistä tai siitä saatua palautetta. Äänitystä voi lisäksi harjoitella ja lähettää opettajalle vasta sellaisen version, johon on itse tyytyväinen. Vastausten perusteella osa lukiolaisista toivoo palautetta myös luokkahuonetilanteessa. Luokassa korjaava palaute voi olla osoitettu joko yksittäiselle oppijalle tai koko ryhmälle, jolloin siihen kootaan yleiset ääntämisvirheet. Opiskelijat toivoivat, että opettaja korjaa heidän ääntämisvirheensä mieluummin eksplisiittisesti, eli kertomalla myös millainen virhe oli kyseessä, kuin antamalla vain suoraan mallin oikeaan ääntämiseen.

Ääntämisestä saatu vertaispalaute on usein koettu hyödylliseksi (Martin & Sippel 2022), mutta osa tämän tutkimuksen vastaajista koki luokkatovereilta saadun palautteen jopa kielteisenä: *Olen saanut myös korjaavaksi palautteeksi naamioituja haukkuja ääntämisestäni luokkakaverilta, ja se vei minulta kaiken itsevarmuuden puhua kieltä.* (vastaaja 1586) Vertaispalautteen antamista tulisi myös opettaa. Jos opiskelija ei saa riittävää ohjeistusta vertaispalautteen antamiseen, hän voi pyrkiä antamaan pelkkää positiivista palautetta, jolloin arvokas korjaava palaute jää saamatta (Ketonen 2021).

Tämän tutkimuksen tulosten keskiarvot tukevat sitä, että opiskelijat toivovat saavansa myös korjaavaa palautetta. Opiskelijat ovat kuitenkin yksilöitä, ja kyselyn vastauksissa, kuten opetusryhmissäkin, myös ääripäät olivat edustettuina. Pieni osa (1,1 %) tämän tutkimuksen vastaajista kertoi, ettei halua opettajan korjaavan mitään ääntämisvirheitä. Olisi tärkeä selvittää syyt tällaiseen toiveeseen, sillä moniin ääntämispalautteeseen liittyviin asioihin voi vaikuttaa palautteen laadulla tai palautteenantotekniikalla. Jos opettaja tuntee opiskelijoidensa oppimistyylit (Yoshida 2008) tai toiveet palautteen laadusta (Pollari 2017), hänen on mahdollista antaa heille yksilöllistä, oppimista edistävää palautetta. Luokassa kannattaisikin keskustella palautteikäytännöistä yhdessä, kuten myös Mäkipää ja Hildén (2021) ehdottavat. Tällöin opettaja saa käsityksen siitä, millaisia toiveita opetusryhmällä on palautteen suhteen ja opiskelijoiden voi olla helpompi ottaa palautetta vastaan: *Tärkeintä olisi se, että palaute annetaan tyylillä, josta oppilaalle käy selväksi tavoite. Tarkoituksena ei ole murskata oppilaan itsetuntoa, vaan auttaa häntä kehittymään.* (vastaaja 1874)

Miten opettaja voisi parhaiten antaa korjaavaa ääntämispalautetta siten, että se olisi hyödyllistä ja kannustavaa? Avovastauksissa nousi esille, että opiskelijat arvostavat

opettajan suoruutta, ystävällisyyttä ja rehellisyyttä. Opiskelija kestää myös paremmin korjaavaa palautetta, kun se annetaan ystävällisesti ja kannustaen. Palautetta antaessaan opettajan kannattaa erottaa toisistaan asia ja opiskelija. Ääntäminen voi olla herkkä asia, mutta ei ole henkilökohtaista tyhmyyttä tai laiskuutta, jos ääntäminen ei suju. Oppijoiden äidinkielet ja muut opitut kielet vaikuttavat oppimiseen. Opettaja voi pohjustaa palautetta kertomalla äidinkielen siirtovaikutuksesta, jonka seurauksena tietyt virheet ääntämisessä ovat yleisiä. Kielet ovat erilaisia, ja äidinkielemme vaikuttaa siihen, miten kuulemme ja tuotamme opiskelemamme kielen äänneitä ja prosodiaa. Opettaja voi korostaa sen tärkeyttä, että kaikki uskaltaisivat puhua, sillä ääntämistä oppii vain harjoittelemalla.

Foneettinen osaaminen sekä tieto ääntämisestä ja sen oppimiseen vaikuttavista seikoista auttavat opettajaa ääntämisen opettamisessa (Virkkunen & Toivola 2020), ja samoin ne auttavat myös selkeän ja konkreettisen korjaavan palautteen antamisessa. Opetuksen apuna voi käyttää foneettista kirjoitusta (esim. IPA), jonka avulla puhetta voidaan kuvata ilman kielen ortografian mahdollista häiritsevää vaikutusta. Opettaja voi myös tarjota oppijalle artikulatoriseen fonetiikkaan ja opiskeltavan kielen ääntämiseen liittyvää tietoa, jonka perusteella oppija osaa korjata ääntämistään, sillä runsaskaan harjoittelu ei välttämättä auta, jos ei tiedä, miten äänneet pitäisi artikuloida. Metakielellinen keskustelu ja foneettisten termien käyttäminen tekevät palautteesta tarkkaa, mutta termien ja niiden merkityksen täytyy olla tuttuja, että niistä on hyötyä (Couper 2011; Lintunen ym. 2023). Foneettisen tiedon lisäksi myös yleinen kielitietoisuus ja kielenoppimisprosessin ymmärtäminen (Mäkipää 2020) auttavat opiskelijaa ymmärtämään, mitä ääntämistä opittaessa tapahtuu, ja hyväksymään, että oppiminen ei tapahdu hetkessä. Luokassa olisi hyvä käydä myös avointa keskustelua siitä, millaisia malleja ääntämisen oppimisessa käytetään, ja muistuttaa, että opiskelijoilla voi olla keskenään hyvin erilaisia tavoitteita ja että aina tavoitteet eivät harjoittelun jälkeenkään täyty. Tavoitteiden määrittelyn yhteydessä voidaan myös pohtia, millaisia ääntämispiirteitä mallit sisältävät ja miten ko. piirteitä voi harjoitella ja oppia. Tavoitteiden asettaminen vaikuttaa myös siihen, millaisiin piirteisiin opettaja kiinnittää huomiota palautetta antaessaan.

Tämän tutkimuksen laatua arvioitaessa on hyvä huomioida, että opiskelijat saivat itse valita kielen, jonka perusteella he vastasivat kyselyyn. Vastauksissa näkyi englannin asema suosituimpana vieraana kielenä, sillä lähes kaksi kolmasosaa vastauksista kohdistui siihen. Lisäksi yli puolet vastauksista kohdistui pitkään A1-kieleen, kun taas lukiossa alkavasta kielestä kertoi vain hieman yli kymmenen prosenttia vastaajista. Osa ensimmäisen vuosikurssin opiskelijoista tosin kertoi, että heidän lyhyen B3-kielen opintonsa eivät olleet vielä alkaneet kyselyyn vastaamisen aikana. On myös hyvä muistaa, että kyselytutkimuksen väitteet ja kysymykset saattavat ohjata vastauksia, joten kyselymme avoimet vastaukset antoivat vastaajille mahdollisuuden selittää vastauksiaan tarkemmin ja tuoda esille myös muita ajatuksiaan. Kyselyä laadittaessa olisimme voineet pyytää vastaajia asettamaan eri väitteitä tärkeysjärjestykseen, jolloin

olisimme saaneet tarkemman kuvan siitä, miten palautetta halutaan saada. Avoimet kysymykset olisi voitu ryhmitellä tarkemmin kyselyn pääteemojen mukaisesti, jolloin olisimme mahdollisesti saaneet tasaisemmin eri aiheista kertovia vastauksia.

Aihepiiri kaipaa lisää tutkimusta, sillä oppijan näkökulma on ollut heikosti edustettuna aiemmissa tutkimuksissa. Jo kerätystä aineistosta on mahdollista tehdä lisätutkimusta ja selvittää esimerkiksi, miten opiskelijan saama ääntämispalautte eroaa pitkän ja lyhyen kielen opetuksessa tai eri kielten välillä. Lisäksi haluamme selvittää tarkemmin, millaisia tunteita palaute opiskelijoissa herättää.

5 Lopuksi

Olemme tässä tutkimuksessa saaneet kattavan kuvan lukiolaisten käsityksistä ääntämisen opetuksessa saadusta palautteesta. Kyselyyn vastanneiden suomalaisten opiskelijoiden näkemykset osoittavat selkeästi, että vieraiden kielten lukio-opetuksessa on syytä lisätä ääntämisen opetusta ja ääntämisestä annettavaa palautetta. Saimme lukioikäisille suunnattuun kyselyyn runsaasti vastauksia eri puolilta Suomea, joten uskomme, että tulokset ovat yleistettävissä hyvin opetukseen. Toivomme, että tulokset lisäävät keskustelua ääntämistavoitteista ja rohkaisevat opettajia antamaan positiivisen palautteen lisäksi myös opiskelijoiden toivomaa korjaavaa, oppimista edistävää ääntämispalautetta. On myös hyvä muistaa, että ääntämispalautteen antaminen osana opetusta on taito, jota voi oppia ja harjoitella (Lintunen ym. 2023).

Tutkimuksemme tuloksia voidaan käyttää apuna kieltenopettajien koulutuksen ja täydennyskoulutuksen suunnittelussa, ja niiden perusteella voidaan kehittää soveltuvia käytänteitä ääntämisen opetuksen ja formatiivisen arvioinnin tueksi. Tulosten perusteella kielten opetussisältöjen painotuksia olisi hyvä tarkastella myös valtakunnallisella tasolla.

Opiskelijoilla voi olla hyvin erilaisia tavoitteita ääntämisensä suhteen, mutta he eivät välttämättä etene niitä kohti, jos palautetta ei anneta tai se on hyvin yksipuolista. Mielestämme jokaisella tulisi olla oikeus ja mahdollisuus pyrkiä toivomaansa ääntämistavoitteeseen, ja tämä olisi tärkeää ottaa huomioon kaikessa kieltenopetuksessa.

Kiitokset

Kiitämme tutkimukseen osallistuneita kouluja ja opettajia yhteistyöstä sekä opiskelijoita heidän arvokkaista vastauksistaan.

Kirjallisuus

- Ahangari, S. 2014. The effect of self, peer and teacher correction on the pronunciation improvement of Iranian EFL learners. *Advances in Language and Literary Studies*, 5 (1), 81–88. <http://dx.doi.org/10.7575/aiac.all.v.5n.1p.81>
- Baker, A. & Burri, M. 2016. Feedback on Second Language Pronunciation: A Case Study of EAP Teachers' Beliefs and Practices. *Australian Journal of Teacher Education*, 41 (6). <https://doi.org/10.14221/ajte.2016v41n6.1>
- Baran-Łucarz, M. 2013. Phonetics learning anxiety – results of a preliminary study. *Research in Language* 11, 57–79. <https://doi.org/10.2478/v10015-012-0005-9>
- Bloom, B. J., Hastings, T. J. & Madaus, G. F. 1971. *Handbook on formative and summative evaluation of student learning*. McGraw-Hill.
- Carey, M. D., Sweeting, A. & Mannell, R. 2015. An L1 point of reference approach to pronunciation modification: Learner-centred alternatives to 'listen and repeat'. *Journal of Academic Language and Learning*, 9 (1), A18–A30.
- Chin, C. 2006. Classroom interaction in science: Teacher questioning and feedback to students' responses. *International Journal on Science Education* 28 (11), 1315–1346. <https://doi.org/10.1080/09500690600621100>
- Couper, G. 2011. What makes pronunciation teaching work? Testing for the effect of two variables: socially constructed metalanguage and critical listening. *Language Awareness*, 20 (3), 159–182. <https://doi.org/10.1080/09658416.2011.570347>
- Couper, G. 2019. Teachers' cognitions of corrective feedback on pronunciation: Their beliefs, perceptions and practices. *System*, 84, 41–52. <https://doi.org/10.1016/j.system.2019.04.003>
- Dlaska, A. & Krekeler, C. 2013. The short-term effects of individual corrective feedback on L2 pronunciation. *System*, 41 (1), 25–37. <https://doi.org/10.1016/j.system.2013.01.005>
- Ellis, R. 2009. Corrective feedback and teacher development. *L2 Journal* 1, 3–18. <http://dx.doi.org/10.5070/l2.v1i1.9054>
- Hakola, H.-L. 2008. *Opettajan antama korjaava palaute ala- ja yläkoulun laaja-alaisessa erityisopetuksessa. Keskusteluanalyttinen tutkimus*. Pro gradu -tutkielma. Jyväskylän yliopisto. https://jyx.jyu.fi/bitstream/handle/123456789/22642/1/URN_NBN_fi_jyu-200912164557.pdf
- Harackiewicz, J. M. 1979. The effects of reward contingency and performance feedback on intrinsic motivation. *Journal of Personality and Social Psychology*, 37 (8), 1352–1363. <https://doi.org/10.1037/0022-3514.37.8.1352>
- Hendrickson, J. 1978. Error correction in foreign language teaching: Recent theory, research, and practice. *Modern Language Journal*, 62, 387–398. <https://doi.org/10.2307/326176>
- Huang, X. & Jia, X. 2016. Corrective Feedback on Pronunciation: Students' and Teachers' Perceptions. *International Journal of English Linguistics*, 6, 245–254. <https://doi.org/10.5539/ijel.v6n6p245>
- Huhtala, A. 2015. Kokonaisvaltainen arviointi tulevien kieltenopettajien haasteena. *Ainedidaktisia tutkimuksia*, 9, 44–59.
- Ilola, M. 2018. *Oppilaiden käsityksiä englannin suullisesta kielitaidosta, sen oppimisesta ja arvioinnista*. JYU Dissertations 15. Jyväskylä: Jyväskylän yliopisto. <http://urn.fi/URN:ISBN:978-951-39-7548-7>
- Kennedy, S. & Trofimovich, P. 2010. Language awareness and second language pronunciation: A classroom study. *Language Awareness*, 19 (3), 171–185. <https://doi.org/10.1080/09658416.2010.486439>

- Ketonen, L. 2021. *Exploring interconnections between student peer assessment, feedback literacy and agency*. JYU Dissertations 379. Jyväskylä: Jyväskylän yliopisto. <http://urn.fi/URN:ISBN:978-951-39-8635-3>
- Lee, E. J. E. 2016. Reducing international graduate students' language anxiety through oral pronunciation corrections. *System*, 56, 78–95. <https://doi.org/10.1016/j.system.2015.11.006>
- Lintunen, P. 2014. Ääntämisen oppiminen ja opettaminen. Teoksessa P. Pietilä & P. Lintunen (toim.) *Kuinka kieltä opitaan*. Helsinki: Gaudeamus. 165–187.
- Lintunen, P. & Dufva, H. 2019. Suullinen kielitaito: mallit ja tavoitteet. Teoksessa E. Tergujeff & M. Kautonen (toim.) *Suullinen kielitaito: opi, opeta, arvioi*. Helsinki: Otava. 42–58.
- Lintunen, P. & Mäkilähde, A. 2017. Short- and long term effects of pronunciation teaching. Teoksessa J. Volin & R. Skarniz. (toim.) *Pronunciation of English by speakers of other languages*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Lintunen, P., Mäkilähde, A. & Peltonen, P. 2017. Learner perspectives on pronunciation feedback. *PTLC2017*, 67–71. <https://urn.fi/URN:NBN:fi-fe2021042717620>
- Lintunen, P., Mäkilähde, A. & Peltonen, P. 2023. L2 Pronunciation Feedback: Pre-Service Teachers' Beliefs and Practices. Teoksessa V. G. Sardegna & A. Jarosz (toim.) *English Pronunciation Teaching: Theory, Practice and Research Findings*. Bristol, Blue Ridge Summit: Multilingual Matters, 2023. 185–204. <https://doi.org/10.21832/9781800410503>
- Lyster, R. & Mori, H. 2006. Interactional feedback and instructional counterbalance. *SSLA* 28, 269–300.
- Lyster, R. & Ranta, L. 1997. Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in second language acquisition*, 19 (1), 37–66.
- Martin, I. A. & Sippel, L. 2021. Providing vs. receiving peer feedback: Learners' beliefs and experiences. *Language Teaching Research*, 1–22. <https://doi.org/10.1177/13621688211024365>
- Martin, I. A. & Sippel, L. 2022. Do beliefs matter? The relationship between beliefs about peer feedback and peer feedback outcomes on pronunciation. *Language Teaching Research*. <https://doi.org/10.1177/13621688221097307>
- Mäkipää, T. 2020. Feedback practices in language classes in Finnish general upper secondary schools. *Apples: Journal of Applied Language Studies*, 14 (1), 103–123. <http://dx.doi.org/10.17011/apples/urn.202006084002>
- Mäkipää, T. 2021. "I have never received feedback from my teachers." *Students' perceptions of assessment and feedback practices in Finnish general upper secondary education*. Helsinki: Helsingin yliopisto. <http://urn.fi/URN:ISBN:978-951-51-7644-8>
- Mäkipää, T. & Hildén, R. 2021. What Kind of Feedback is Perceived as Encouraging by Finnish General Upper Secondary School Students? *Education Sciences*, 11 (1), 12. <https://doi.org/10.3390/educsci11010012>
- Opetushallitus 2019. *Lukion opetussuunnitelman perusteet 2019*. Määräykset ja ohjeet 2019:2a. Haettu 7.8.2022 osoitteesta https://www.oph.fi/sites/default/files/documents/lukion_opetussuunnitelman_perusteet_2019.pdf.
- Oroujlou, N. & Vahedi, M. 2011. Motivation, attitude, and language learning. *Procedia-Social and Behavioral Sciences*, 29, 994–1000. <https://doi.org/10.1016/j.sbspro.2011.11.333>
- Panova, I. & Lyster, R. 2002. Patterns of corrective feedback and uptake in an adult ESL classroom. *Tesol Quarterly*, 36 (4), 573–595. <https://doi.org/10.2307/3588241>
- Pekrun, R. 2016. Academic emotions. Teoksessa K.R. Wenzel & S.B. Miele (toim.) *Handbook of Motivation at School*. Taylor & Francis Group. 120–144.
- Phuong, T. T. H. 2022. Feedback on Pronunciation: Vietnamese EFL Teachers' Beliefs and Practice. *SAGE Open*, 12 (3). <https://doi.org/10.1177/21582440221126612>

- Pohjala, K. 2004. Suullinen kielitaito ja sen arviointi ylioppilastutkinnossa. Teoksessa R. Jaatinen, P. Kaikkonen & J. Lehtovaara (toim.), *Opettajuudesta ja kielikasvatuksesta – Puheenvuoroja sillanrakentajille*. Tampere: Tampere University Press. 166–171.
- Pollari, P. 2017. (Dis)empowering assessment?: assessment as experienced by students in their upper secondary school EFL studies. *Jyväskylä studies in humanities* 329. Jyväskylä: Jyväskylän yliopisto. <http://urn.fi/URN:ISBN:978-951-39-7178-6>.
- Saito, K. 2011. Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: An instructed SLA approach to L2 phonology. *Language awareness*, 20 (1), 45–59. <https://doi.org/10.1080/09658416.2010.540326>.
- Saito, K. 2012. Effects of instruction on L2 pronunciation development: A synthesis of 15 quasi-experimental intervention studies. *Tesol Quarterly*, 46 (4), 842–854. <https://doi.org/10.1002/tesq.67>
- Setter, J. & Jenkins, J. 2005. State-of-the-Art Review Article. *Language Teaching*, 38 (1), 1–17. doi:10.1017/S026144480500251X.
- Sheen, Y. H. 2004. Corrective feedback and learner uptake in communicative class-rooms across instructional settings. *Language teaching research* 8 (3), 263–300. <https://doi.org/10.1191/1362168804lr1460a>
- Spada, N. & Lightbown, P. M. 1993. Instruction and the development of questions in L2 classrooms. *Studies in second language acquisition*, 15 (2), 205–224.
- Tergujeff, E. 2013. *English pronunciation teaching in Finland*. Jyväskylä Studies in Humanities 207. Jyväskylä: Jyväskylän yliopisto.
- Tergujeff, E., Heinonen, H., Ilola, M., Salo, O.-P. & Kara, H. 2019. Suullisen kielitaidon opetus käytännössä. Teoksessa E. Tergujeff & M. Kautonen (toim.) *Suullinen kielitaito: opi, opeta, arvioi*. Helsinki: Otava, 96–117.
- Thohir, L. 2017. Motivation in a Foreign Language Teaching and Learning. *Vision: Journal for Language and Foreign Language Learning*, 6 (1), 20–29. <https://doi.org/10.21580/vjv6i11580>
- Tuomi, J. & Sarajärvi, A. 2017. *Laadullinen tutkimus ja sisällönanalyysi: Uudistettu laitos*. Tammi.
- Ullakonoja, R. & Dufva, H. 2016. Toisen ja vieraan kielen ääntämisen oppimisen haasteet. *Oppimisen ja oppimisvaikeuksien erityislehti*, 26 (2), 4–18.
- Virkkunen, P. & Toivola, M. 2020. Foneettinen osaaminen helpottaa vieraan kielen ääntämisen opettamista – kyselytutkimus suomalaisten kieltenopettajien käyttämistä ääntämisen opetusmenetelmistä. *Ainedidaktiikka*, 4 (1), 34–57. <https://doi.org/10.23988/ad.85736>
- Wu, L. 2020. Gender and learners' beliefs in corrective feedback in learning Chinese as a foreign language: A network questionnaire survey report. *International Educational Research*, 3 (4), 1–9. <https://doi.org/10.30560/ier.v3n4p1>
- Wu, L., Wu, Y. & Zhang, X. 2021. An On-line Questionnaire Survey on Students' Views and Teachers' Practices in Corrective Feedback in Teaching Chinese to Speakers of Other Languages. *International Linguistics Research*, 4 (1). <https://doi.org/10.30560/ilr.v4n1p1>.
- Yoshida, R. 2008. Teachers' Choice and Learners' Preference of Corrective Feedback Types. *Language awareness*, 17 (1), 78–93. <https://doi.org/10.2167/la429.0>
- Ölmezer-Öztürk, E. & Öztürk, G. 2016. Types and Timing of Oral Corrective Feedback in EFL Classrooms: Voices from Students. *Novitas-ROYAL (Research on Youth and Language)*, 10 (2), 113–133.

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 60–80.

Pirjo Pollari & Dmitri Leontjev
University of Jyväskylä

Beyond error correction in EFL writing in a Finnish upper secondary classroom: a practical approach

Highlights

- According to research, Finnish FL/L2 teachers focus more on summative than formative assessment and feedback, which contrasts with the latest curricula.
- We propose practical assessment and feedback activities that go beyond error correction to support the learning process.
- Learner choice, oral feedback as well as guidance and peer collaboration seemed to foster students' engagement with feedback.
- We hope this article will inspire teachers to develop further ideas to increase the amount of formative feedback in FL/L2 education.

Abstract

In Finland, teachers have considerable autonomy over their assessment practices. Recent studies suggest that FL/L2 teachers, particularly at the upper secondary school, primarily focus on summative assessment rather than on formative assessment and feedback, which is in contrast with the latest Finnish National Core Curricula. Furthermore, while appreciating teacher feedback, learners do not perceive it as an integral part of assessment. In this paper, we propose formative classroom practices to support the learning process and go beyond error correction in FL/L2 writing. These practices are grounded on our earlier research, and they focus mainly on two themes: fostering learner choice in feedback and error correction methodology and supporting peer feedback with the help of a checklist based on the Matriculation Examination rating scale. Drawing from students' comments and texts, we will explore learner experiences with these assessment practices and the changes they may bring to the classroom.

Keywords: corrective feedback, peer feedback, formative assessment, L2 classroom, writing

1 Introduction

Assessment is a major constituent of teachers' and students' work in any school system. In fact, Stiggins (2014) suggests that assessment-related work may take up 25–33 % of the teacher's professional time. It also takes up a substantial amount of students' time. Atjonen (2007) estimates that a Finnish primary school pupil in Year 5 (aged 11) may have over 50 different tests and assessments during the school year, and later the assessments are likely to become both more frequent and more demanding. Therefore, assessment should be designed so that it can impart important information to both guide instruction and enhance learning effectively (Stiggins 2014). Only then would all the time spent on assessment be time truly well spent.

The question is whether teachers are adequately trained and supported for quality assessment both in Finland and elsewhere. Some studies propose that teachers' assessment literacy might not be of a satisfactory standard (Stiggins 2014; Atjonen 2014, 2017; Mäkipää & Hildén 2021).

Compared to many educational contexts, Finnish teachers – also FL and L2 (foreign or second language) teachers – have considerable autonomy over their assessment practices: they decide, design and organise how to gather assessment evidence, they define the assessment criteria and mark the student work, and they also draw conclusions on the assessment results (Sahlberg 2007; Vänttinen 2011). The sole form of assessment teachers cannot influence directly is the Matriculation Examination, the only high-stakes nationwide examination in the Finnish school system.

Perhaps quite surprisingly, despite their assessment autonomy, teachers' assessment and feedback practices in relation to their assessment literacy is a relatively new topic in Finnish FL/L2 research, as are learners' perceptions of classroom as-

assessment practices and engagement with teachers' feedback (Pollari 2017a, 2017b; Atjonen et al. 2019; Mäkipää 2021a; Mäkipää & Hildén 2021).

While relatively recent, these studies have made a substantial contribution, developing the understanding of Finnish FL/L2 classroom assessment practices. For instance, it appears that Finnish teachers – and FL/L2 teachers among them – primarily focus on summative assessment, with different tests as the dominant method of collecting evidence on learners' skills (Väljörvi et al. 2009; Hilden et al. 2015; Atjonen et al. 2019; Mäkipää 2021a; Leontjev 2022). Although Finnish FL/L2 teachers use various tests – such as listening comprehension and oral tests – more than other teacher groups, they all tend to test a student's solo performance (Atjonen et al. 2019).

Finnish FL/L2 teachers' feedback also primarily targets learner performance as a finished product rather than focusing on learning as a continuous process. According to Mäkipää (2021b), Finnish FL/L2 teachers give feedback on learner writing as the finished product, not on drafts in progress. Hence, the feedback acts as *feed back* rather than *feed forward*, as students cannot improve their actual work on the basis of the feedback anymore (e.g., Hattie 2009). Furthermore, according to Atjonen et al. (2019), the primary form of feedback in Finnish schools is still either grades or points. Teachers also mark the errors comprehensively – usually with a correction or an explanation, but occasionally also without any clarification (Atjonen et al. 2019; Mäkipää 2021b). All in all, when compared to the stipulations of the latest Finnish National Core Curricula (FNBE 2016; FNAE 2019), the amount of formative feedback seems low (e.g., Atjonen et al. 2019; Mäkipää & Ouakrim-Soivio 2019).

The recent body of research also sheds light on students' experiences of classroom assessment and feedback. While generally appreciating teacher feedback and often even craving for more (Pollari 2017b; Mäkipää 2021a), learners do not perceive it as an integral part of assessment and largely associate feedback with correction, grading, tests and exams (Mäkipää 2021a).

Furthermore, learners are predominantly more or less passive recipients of assessment and feedback. Although learners are well informed of the assessment practices and criteria at the beginning of each course by the teacher (Väljörvi et al. 2009; Atjonen et al. 2019; Mäkipää & Ouakrim-Soivio 2019), their role in influencing the decisions may be rather limited (Pollari 2017a), and they may not always understand what the criteria mean (Mäkipää 2021a). Self- and peer-assessment are used to a degree but do not appear to have an established role in classroom assessment (Mäkipää 2021a; Leontjev, 2022). Furthermore, they may not be regarded as very relevant or beneficial by some students (Pollari 2017b; Mäkipää 2021a).

This focus on summative assessment, or assessment *of* learning (AoL), is in contrast with the latest Finnish National Core Curricula (FNBE 2016; FNAE 2019), which emphasise the importance of formative assessment or assessment *for* learning (AfL). Thus, the general assumption appears to be that formative assessment or AfL is relatively scarce in Finnish FL/L2 classrooms (Mäkipää 2021a). There may be several rea-

sons for that. First, the focus on summative assessment in Finnish upper secondary FL/L2 classes is not surprising when practically all upper secondary students take the Matriculation Examination at the end of upper secondary school (Mäkipää 2021a; Leontjev 2022). Secondly, FL/L2 teachers' assessment literacy in relation to formative assessment/AfL may be somewhat lacking (Mäkipää & Hildén 2021). Formative assessment may not have been included in the teacher education or training curricula at the time when the majority of the current teachers studied, as assessment in general – even today – seems to get rather little attention in teacher education curricula (Atjonen 2014, 2017). Teachers have not necessarily had further in-service training in formative assessment either. Another reason may well be that formative assessment and its methods are seen as part of teaching and learning by students and teachers alike and is, therefore, not really regarded as part of assessment (Pollari 2017a; Mäkipää & Hildén 2021). Students may also not recognise formative feedback when they receive it (Mäkipää 2022). Finally, the dominant focus on corrective feedback, i.e., correcting language errors, has long and strong traditions in FL/L2 education both in Finland and elsewhere (Alderson et al. 2015).

1.1 The purpose and data of this paper

In this practice-oriented paper, we propose practical classroom assessment and feedback activities whose aim is to shape and support the learning process rather than focus on assessing the product only. These practices originate from our earlier research carried out either collaboratively or individually (e.g., Pollari 2017a; Leontjev & Pollari 2022a, 2022b). The proposed assessment activities will focus on *writing* as, out of all teachers in Finland, FL/L2 teachers appear to use various writing tasks the most for assessment purposes (Atjonen et al. 2019). Furthermore, students seem to perceive the assessment of written texts as one of the most prominent components of teachers' assessment work (Mäkipää & Hildén 2021). The main audience of the present paper is, therefore, FL/L2 teachers whom we invite to adopt and adapt these activities in their classrooms.

The overall goal of the activities discussed in this paper is to foster learner agency in assessment through increasing learner choice, dialogue and peer collaboration. First, we will focus on **individualised corrective feedback** (CF) in FL/L2 writing. We will also discuss the **role of oral feedback**. Secondly, we will report on the construction and practical implementation of a checklist based on the Matriculation Examination rating scale. We will focus on its use in **peer feedback**, yielding insights into learners' areas of struggle as well as areas which they consider important.

As mentioned above, the insights and the data in this paper draw from our earlier studies. Most of these studies have been published elsewhere and had different foci. In this paper, we synthesise the findings of these studies with a practical pur-

pose in mind which we divided into two intertwined questions to create a focus for our synthesis:

1. What formative assessment activities that go beyond error correction emerge from these studies?
2. How can they be implemented in a FL/L2 classroom?

We will explore learners' experiences of these assessment activities and the changes the activities bring to the classroom with the help of learners' texts as well as their answers in questionnaires.

In all these studies, informed consent was obtained from the students. They were informed of the goals of the studies and how the results were to be used. When possible, anonymous data were collected, for example, in questionnaires. Other data, such as learner texts, were anonymised, and learner names were replaced with pseudonyms. Some data, such as learners' comments and responses, and excerpts from the checklist we created, were translated by us for this paper.

2 Feedback, error correction and learner choice

Generally, feedback research quite unanimously suggests that feedback should not be limited to the narrow definition where "the role of feedback is to 'put things right' by taking a corrective action" (Price et al. 2010: 278). The role of grades or scores in conjunction with feedback has also been considered problematic and disruptive for learning (Stobart 2012). Furthermore, feedback should not only feed *back*, that is, address the task at hand, but also feed *forward*, fostering further learning (e.g., Hattie 2009). Neither should feedback be a solely unidirectional activity where the teacher 'gives' feedback and the learner just 'receives' it but a process where the learner is an active participant, and which enables and enhances self-regulated learning and student autonomy (Askew & Lodge 2000; Burke & Pieterick 2010). Feedback can have a strong impact on learning, but no matter how good feedback itself is, its impact depends on how students react to it – and different students may react to it differently (Hattie 2009; Wiliam 2012).

In FL/L2 education and assessment, however, many scholars and teachers alike seem to still focus on *error correction*. Accordingly, FL/L2 education literature has defined feedback mainly as corrective feedback, be it oral or written, and there has been a long and lively debate about the efficacy of CF and its different forms (see, e.g. Truscott 1996; Bitchener & Ferris 2012).

To put it slightly simplistically, second language acquisition scholars tend to advocate *direct corrective feedback* (Bitchener & Ferris 2012): the teacher indicates the error and provides its correct form. Corrections may come with metalinguistic

feedback, a clue or explanation of the language form and its possible 'rule,' which "appeals to learners' explicit knowledge by helping them to understand the nature of the error they have committed" (Ellis et al. 2008: 356). However, many L2 writing studies have promoted *indirect corrective feedback* (Ferris 2010). Indirect feedback may be *uncoded* (the teacher only indicates that an error has been made) or *coded* (the error is marked with a code or clue regarding its nature). Instead of providing the correct forms, the teacher asks the students to correct – and sometimes even to find – the errors themselves. The advocates of indirect CF believe that the self-correction process increases student engagement as well as attention, noticing and reflection, and thus is more effective in the long run than direct correction (Bitchener & Ferris 2012). Effects of focussed and non-focussed CF have also been studied. In focussed CF, the teacher corrects or indicates (one or) a few selected types of errors only; in unfocussed CF, all errors found in the given student text are treated. To date, although CF is an "arguably over-exposed topic" (Ferris et al. 2013: 308), no consensus has been found on which CF method should be the most effective (Guénette 2007; Brown 2012; Bitchener & Storch 2016).

Moreover, both practice and research show that although language teachers spend hours writing corrections and comments on student work, some students do not benefit greatly from corrective feedback, and some do not even pay attention to the feedback they get (e.g., Truscott 1996, 2007; Guénette 2012). Some students do not seem to find feedback useful or motivating, either (Pollari 2017b; Mäkipää 2021b). Why not? Does the feedback students receive not meet their needs (Hattie 2009)? Yet, few studies have considered individual student responses to feedback or attempted to individualise CF based on students' preferences (Storch & Wigglesworth 2010; Ferris et al. 2013). Motivating students to engage in corrective feedback has also been somewhat neglected in most CF research (Brown 2012).

It appears that Finnish FL/L2 teachers usually decide themselves which CF method to use (Mäkipää 2021b). Additional feedback is most likely provided as written comments; oral feedback seems to be lacking in the Finnish FL classrooms (Mäkipää 2021b).

2.1 Feedback to enable learner agency

Not quite satisfied with the utilised CF and feedback practices or their efficiency, we conducted altogether three practically-oriented studies with the aim to find out what kind of feedback and error correction would serve the students best *in their own opinion*, and why.

In the first two small-scale classroom studies (Pollari 2016, 2017a), altogether 46 students were asked to opt for non-focussed direct or indirect CF, as well as getting feedback and metalinguistic explanations either orally or in writing. The students, who were on advanced upper secondary school EFL courses (ENA7 and school-

based ENA9, students aged 17–19), stated at the end of their EFL essays, written in class, whether they wanted feedback on that text orally or in writing, with errors corrected or marked. Later, students were also asked to give reasons for their choices as well as feedback on these small-scale assessment studies in writing, with the help of a small, open-ended questionnaire. Students’ quotes below come from the questionnaires and are coded with respondent numbers (e.g., student30 or, in a later student group, S30).

In the first study, there were 30 students who wrote just one essay; in the second study, the sixteen students wrote two essays, choosing their feedback methods both times. When combined, the students’ choices in these two studies show a range of preferred feedback options (Figure 1).

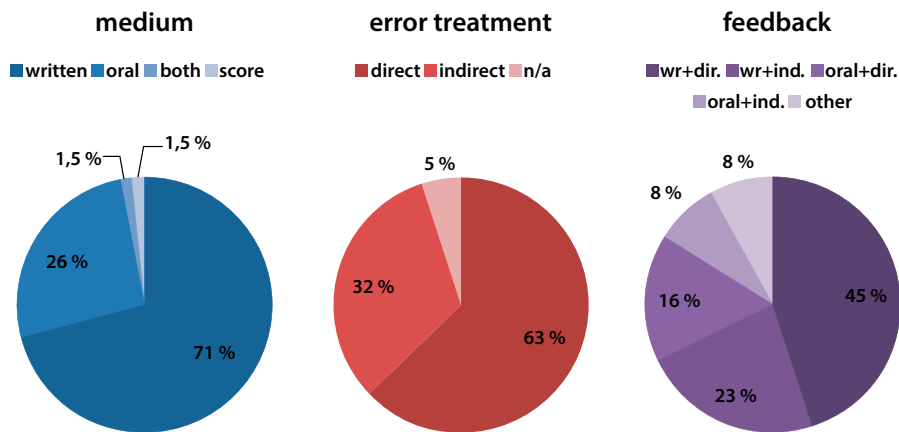


FIGURE 1. Students’ feedback choices.

2.2 Indirect or direct CF

As Figure 1 indicates, the majority of students preferred direct CF. The most common reason for this was the students’ uncertainty in their skills, but there were other reasons as well, as the student comments below indicate.

- I wouldn’t necessarily know what went wrong if the errors were just marked. (student7)
- When the errors are corrected, I learn best from them. (student9)
- I get good examples for my future essays when the sentences are corrected. (student21)
- Correcting the mistakes myself would have taken more time and I still wouldn’t have been sure if I got them right. (student35)

These comments support the views of the proponents of direct CF: direct CF offers students certainty over the correct form. Furthermore, the comments show that direct CF helps students to notice mistakes, reinforces learning and provides them with suitable examples. It is, no doubt, also quicker for the student. On the other hand, the wish to reflect on the errors and their corrections was the primary cause for choosing indirect correction, which was the preferred method for about a third of the students. Moreover, these students believed indirect CF would enhance their learning and reduce future errors. Thus, the advocates of indirect CF seem to be correct, too.

So that I learn to find my errors myself. (student38)

So that I could try and correct my errors and thus learn better. (student2)

I thought I would remember my errors better if I thought about them myself first so that I wouldn't repeat the same errors. (student18)

In a further, collaborative study, only four learners opted for indirect feedback (out of 25 learners, with one learner not stating their preference), as indicated by the results of a questionnaire that the learners took at the end of the course (see Appendix A). This study was conducted with upper secondary school learners in advanced EFL courses. We report other data we collected with these learners elsewhere (Leontjev & Pollari, 2022b); yet, the questionnaire results have not been previously reported. The questionnaire item (the frequency analysis of which is presented in Appendix A) asked the learners to indicate whether they opted for direct or indirect corrective feedback and to elaborate why. We used grounded coding in the analysis, meaning the codes emerged during our engagement with the learners' open-ended responses.

Learners' comments suggest that learners who chose indirect feedback wanted to be responsible for identifying sources for their challenges and how to address them. These learners also seemed to concentrate on their writing processes rather than the product of their writing. To give two examples, the first comment below was written by a learner who opted for indirect feedback and the second by a learner who chose direct feedback:

I want my mistakes to be underlined, for I have enough attention left when I know where the mistake is; I usually understand them myself. In this way, I also learn better about my own weaknesses. (S14)

corrected, then you can immediately see exactly how it should have been done (S15)

2.3 Oral or written feedback

The mode of corrective feedback, that is, written or oral, has not been considered a crucial issue in regard to the efficacy of feedback (Mäkipää 2021a). In general, students are believed to favour written feedback. In the three studies reported above, the students were also asked to select between written and oral feedback.

As Figure 1 above indicates, most students in the first two studies (Pollari 2016, 2017a) preferred written feedback, and only about a quarter of them opted for oral feedback. The students who asked for oral feedback appreciated the chance to ask for clarifications and further advice; those who chose written feedback stated it would be easier for them to revisit their errors as well as the comments and corrections if they were in writing. There was one student who wanted both forms of feedback, and their comment summarises the main reasons for both modes:

Oral + small written (feedback). Because oral is nice as you can also ask and interact, but written, too, because you can go back to it later on. (student28)

There were, however, a few cases of teacher intervention. In the second study, I gave five students oral feedback on their second essays instead of the written feedback they had asked for because there were some aspects that I wanted to show them in their essays and/or found it difficult to explain all my points in writing.

In general, the oral feedback conferences took 5–10 minutes each. Although I, as a teacher myself, recognise the time constraints involved in organising oral feedback sessions, I feel that those minutes were time well spent. The students had the chance to ask for further clarifications and advice, and I could elaborate on my comments and corrections as well as make sure the student had understood my comments. Moreover, I learnt to know some of my students, their goals and needs slightly better as they brought up issues I would not have known of without these discussions.

In addition to increasing student choice and hence also learner agency, the most noteworthy gain in oral feedback was *interaction* – feedback was no longer a unidirectional activity but a dialogue (e.g., Askew & Lodge 2000). This dialogue also facilitated feed forward, as the students often asked for advice for their future writing or language studies in general. My overall impression of these oral feedback sessions – even if they were sometimes rather ad hoc – is that they went further beyond error correction and the task at hand than my written comments. Later, guided and supported by the second author as part of our joint research project, oral feedback sessions became better structured and thought-out and could result in moments of mediation, which impacted the understanding of both the learner and the teacher (for further discussion, see Leontjev & Pollari, 2022b).

Overall, students welcomed the chance to influence the CF methods. In the questionnaires, all the respondents in the first two studies said that the methods they had chosen themselves were either more useful than or equally useful as the previous teacher-chosen methods.

3 Guiding peer feedback on writing: using a checklist

3.1 Constructing the checklist

Section 3.2 discusses how a checklist tool guided the FL/L2 writing process in a Matriculation Examination (ME) preparation course. In this section, we outline how the second author created the checklist together with an ME censor. The first author was consulted and offered some insight into the formulation of the checklist. The motivation to create the checklist emerged from the upper secondary school students' preoccupation with the ME scale: in our experience, students seem to think about the development of their writing in terms of obtaining a higher score rather than considering what features would allow for creating higher quality texts. Our decision that eventually led us to choosing the checklist format for the tool was to remove the scoring scale to encourage learners to focus on features of their texts elicited in the ME writing rating rubric. We then retrofitted the statements in the three broad categories elicited in the ME essay writing rating rubric – *communicativeness, content and organisation, and language breadth and accuracy* (Matriculation examination board 2022).

We synthesised statements focusing on the same aspect of the ME writing construct at different band scores – for instance, *“the writer handles the subject in an especially versatile way”* and *“the writer handles the subject rather one-sidedly”* – into questions that learners can ask themselves during the writing process. In this example, we formed the following question: *If it is an argumentative text, do you present it from more than one point of view?*

Furthermore, we based the checklist on our experience and interpretation of the general points on the ME writing rating scale and on our experience with the challenges learners had with particular aspects of writing. For example, the versatile vocabulary aspect, e.g., *“the author uses a rather limited, usually, (possibly) only partially appropriate range of expression.”*, was transformed into two questions: *Have you tried to use varied words and phrases?* and *Have you tried to use other than just basic words (for example, ‘nice’, ‘good’, ‘very’, ‘thing’, ...)?*

Finally, we based the checklist on the AfL, including self-assessment (Lee 2017), and research on writing processes (Huhta, Harsch, Leontjev, & Nieminen 2023), proposing questions and statements the learners could follow through their planning, text generation and revision stages, e.g., *“Does the text match the intended goals, audience and genre?”* (see Appendix B for the checklist).

3.2 Checklist in peer feedback

The use of checklists in peer feedback is not new (see Seow 2002). The benefit of this checklist lies in that, as detailed in the previous section, it is based on the writing construct as defined in the ME. Thus, it goes beyond writing accuracy and defines strengths and weaknesses of a written text in a rather concrete way while leaving it up to learners how they and their peers could develop their texts. In Leontjev and Pollari (2022a), we reported on using the checklist in peer feedback, focusing on two learners: there, our goal was to explore how the teacher could use the information obtained from the peer interactions in subsequent teaching. Here, we focus on a larger number of students and on how the checklist informed the comments they wrote on one another's drafts.

Nineteen upper secondary school students (in an advanced EFL course) were asked to engage in a peer assessment and feedback session in pairs or groups of three, where they exchanged their draft essays, read them and commented on them with the goal of helping their peers to develop their essays. While reading the essays, the learners were instructed to write comments on their peers' drafts to help them give feedback to their peers. Above all, the goal was to enable learner agency in deciding what is important to address in the text and how. That said, this agency was mediated by the checklist. This, we argue, made the activity balanced in terms of learner agency on the one hand and structure and guidance on the other hand: thus, the activity became a learning activity – not just an activity improving the text. We analysed the data – the initial essay drafts commented by the peers as well as the final essays – in two ways. First, we coded the feedback that the learners wrote on their peers' essays using the points on the checklist (Appendix B) as the codes. Next, we studied the same feedback qualitatively, including its foci, degree of explicitness and its formulations. Furthermore, we analysed what modifications to their essays the students made in response to this feedback. We also compared the feedback comments within the pairs (or groups of three) and across them.

Out of the total of 131 comments the learners wrote of each other's essays, the aspects elicited on the checklist were mentioned in one way or another in 90 comments. The following Figure 2 shows which aspects the learners mentioned in the comments and their frequency. The classification is based on an a priori content analysis of the learner comments, using the checklist as the basis, as we detailed above.

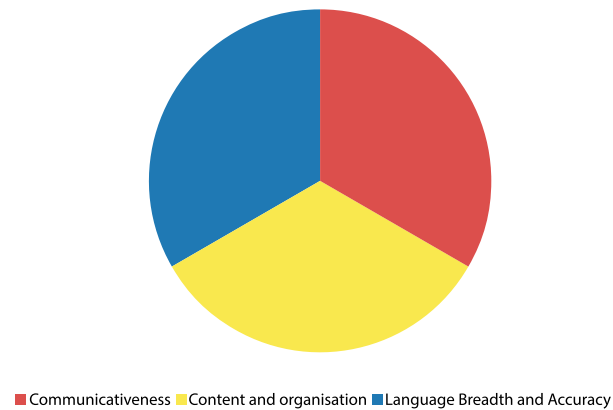


FIGURE 2. Aspects of the checklist (Appendix B) mentioned in learners' comments.

Contrary to our expectation that learners might concentrate on accuracy, the learners took into account all three categories on the checklist. Clarity/readability was mentioned in 23 comments; improving cohesion was mentioned in the further nine notes. Accuracy was also mentioned often, for example, verbs in correct forms in six comments, but less frequently than clarity of expression. These results show that if meaningfully and systematically guided, learners will focus not only, and not as much, on accuracy but also on other aspects of writing. Importantly, studying learner comments with reference to the checklist, such as the one we created, gives the teacher an idea of what learners consider important to address in their drafts – that is, what constitutes development of their writing for them.

We also found that the comments and their foci were rather similar within the pairs/groups of learners but differed across them. We illustrate this with excerpts from the drafts of two pairs of learners (Figures 3 and 4).

4. If I was a scientist...

I want to study information technology and robotics because they're an extremely important part of the future. Technology keeps advancing but the constant progress also creates new problems. Nearly everything can be done on a computer these days so computer security is crucial for all of us. I would like to develop softwares to be more secure against viruses and hacking. Some coders can use their knowledge on taking other people's information or money and I want to prevent that.

Artificial intelligence is one of the latest achievements in technology and it truly fascinates me. Creating a machine that can think for itself is amazing but we need to be careful. The idea of robots rising up against the humankind isn't impossible. They should be very carefully coded and all things should be taken into consideration before giving a machine the power to think. That's something I would also like to work on in the future.

So if I was a scientist I would definitely work on computing and robotics. Computers and robots are the future and I want to make the future safer and better for all of us.

4. If I was a scientist...

I would like to study the genes related to the inheritance of arthritis and the reasons why some people get it even though it doesn't run in their family. I find the subject very intriguing especially because it's still fairly unknown which genes cause this illness.

Arthritis is more commonly known to occur in older people who get the illness because their joints have gotten worse, however, arthritis can also occur in children. This type of arthritis is less known and there are many things no one knows about it. Including the reason why it occurs. There are many things that can trigger it including scratches and impacts, but no one knows why these things can trigger this chronic illness.

Handwritten notes in the top draft include: "Why is robotics extremely important part of the future?", "What?", "What exactly?", and "robots".

Handwritten notes in the bottom draft include: "Could maybe be shorter", "rahana joku", "kassa ok", "what?", and "vähän irrallinen?".

FIGURE 3. Drafts: Learner E (top) commented by learner F, and learner F (bottom) commented by learner E.

73 BEYOND ERROR CORRECTION IN EFL WRITING IN A FINNISH UPPER SECONDARY CLASSROOM: A PRACTICAL APPROACH

When we think of Mars exploration, we should think about what are our goals as humans. I'd argue that a pretty important goal for us is to not go extinct. Given the current state of the Earth's climate and the threat of nuclear weapons, our home planet might not be able to support humans life after a while. Also it is known, that the Sun will eventually die out we will have to leave the solar system if we want to live after that. Inhabiting Mars might not be the solution to these problems but it's the first step us becoming an intergalactic species which is crucial to our evolution and survival.

So in conclusion, inhabiting Mars is definitely not a waste of money but in fact crucial to our evolution. It's a step we have to take, and now is a great time to do it!

Satunnaisia kirjoitusvirheitä, melko hyvä sanastollisesti, monipuolinen
Hyvä, selkeä rakenne. Hicman toistavaa ajoitain. Idea tulee esille
joka kappaleessa.

While robots are more efficient, cheaper and possibly reliable to use as workforce but they also have their downsides. A human being is for example much better at creative thinking and learning in contrast to robots. In functions where these abilities are not required you could use a robot.

However, when people start replacing human workers with their automated counterparts is when we could face some problems. With unemployment already being a somewhat big of a problem robots stealing our jobs could worsen the situation. This is where we need to be careful.

In conclusion, robots are very beneficial to us if used right. However, the transformation should be done carefully and by taking every aspect of it into consideration.

Sanasto on hyvin monipuolista ja vaihtelevaa, mutta tarvitsee kaikkikommentissa olla tarkaisessa lauseessa näin "hieman" sanasto. Muut on mielekkäitä erittäin hyvä!
Kappaleet ovat kohtuullisen lyhyitä. Aloitte oli voinut käsitellä laajam-
masti. Mäkä. Mäkä. tulee toki vastaan.

FIGURE 4. Drafts: Learner N (top) commented by learner O, and learner O (bottom) commented by learner N.

The comments written by the two pairs are markedly different. Learners E and F focused on clarity and coherence. Learners N and O, judging by the underlined parts in the text proper, focused more on accuracy and connective devices. They also summarised the strengths and weaknesses of each other's texts at the bottom of each draft. The way that the learners formulated their guidance was different, too. Learners E and F used prompts and guiding questions. Learners N and O simply underlined issues they identified in the texts, each summarising these points.

Therefore, the teacher can see what learners consider important in their writing and how much guidance they need to develop their texts. We emphasise that in our data, indirect feedback was quite often selected by the learners to help their peers to develop their writing.

4 Discussion

In this paper, we aimed to introduce and discuss classroom activities that can take feedback on FL/L2 student writing beyond traditional error correction and summative assessment. In addition, our goal was to foster learner agency as well as learner collaboration and engagement through these activities. Next, we will discuss how such feedback and assessment activities, emerging from our earlier research, can be implemented in the classroom and what the teacher should pay attention to when implementing them.

The first three studies of this paper focused on making EFL writing feedback more individual and learner-centred. The students were given a choice between direct or indirect CF as well as receiving additional feedback orally or in writing. As the results of these undoubtedly limited and short-term studies showed, students are individuals, with individual needs, wishes and reactions to feedback. Therefore, if the teacher decides the CF method, the chosen method may not cater for all – and not even most – students and their needs and wishes. Having influenced the CF methods themselves, students appeared more invested and interested in the corrections and comments they received. From the teacher's perspective, knowing how each student wanted their feedback made the CF process more focused and meaningful. In other words, the information that the teacher received from the learners about their preferences both added to the teacher's assessment, suggesting how much responsibility learners were willing to take for developing their texts, guided the teacher's feedback to the learners, and made it more likely for the learners to take the feedback into account.

We argue that one reason why learners seem to prefer overt correction coming from the teacher could be because they are used to such feedback, both in Finland (Mäkipää 2021a) and elsewhere (Amrhein & Nassaji 2010). Therefore, asking for learner preferences in feedback shifts control and power of feedback in many respects; in error correction, all the control is with the teacher and error correction is unidirectional. All learners can do is either accept and attempt to act on the feedback, or reject it (Wiliam 2012). Furthermore, mere CF does not necessarily result in learning, although it may improve the accuracy of the text at hand (Zhang & Hyland 2018), as there may be little feed forward to future writing. Nevertheless, our findings also suggest that teachers might try both direct and indirect feedback with their students. Thus, learners could experience both feedback methods and recognise their qualitatively different foci and purposes, allowing them to make more informed decisions when choosing CF methods. This also links to enhancing Finnish upper secondary school students' assessment literacy, as discussed by Mäkipää (2021a). Hence, we suggest that learners should first experience feedback that (a) moves beyond correcting the text and its errors towards developing writing ability, (b) does so indirectly and (c) also involves learners in giving such feedback to their peers.

We illustrated the latter point through a peer review activity guided by a checklist. Due to the introduction of the checklist, the learners got an idea of what to pay attention to in their peers' essays but were free to decide what to focus on and to suggest how their peers could develop their texts. As we discussed, the learners moved away from the sole focus on accuracy and gave feedback in various ways, not focusing on explicit correction. This gave the teacher valuable information regarding what the learners considered important in their own and their peers' writing and whether the degree of directness was enough for the learners to develop their essays. Admittedly, this peer review activity was focused on developing the text at hand. However, the teacher, who is the orchestrator of the teaching/learning process in the classroom, can use the information from such peer activity to guide the subsequent teaching and learning, moving away from the product to the process of writing. In other words, as an assessment activity, the peer assessment guided by the checklist was not just peer-assessment; the learners' original texts, the comments that their peers wrote and the modified texts were all parts of the same teacher-assessment activity, giving the teacher valuable information that could be used to adjust the subsequent teaching.

Before finishing the paper, we would like to list some limitations that further research could address. The first, as we mentioned, is the scale of the studies we synthesised. For further research, this implies larger-scale studies as well as interventions with a larger number of participants. Also, we did not systematically study the effect of the assessment and feedback activities on learner development across texts, although a development from learner drafts to final essays was observed in peer-review study. Hence, pre/post-test intervention designs could be used in future research.

We should highlight, however, that our main readership in this paper are teachers. Addressing this readership, we wish to stress that we do not intend the activities to be used verbatim but rather adapted, informed by the following guidelines based on the research we synthesised:

1. learner choice of feedback is important, but learners should first experience different forms of feedback
2. while time is a constraint, oral feedback that allows more interaction and thus more personalised guidance for learners is often a good investment
3. even if learners are used to focusing on accuracy in peer feedback, when meaningfully guided (e.g., with a checklist), they recognise the importance of other aspects of writing.

Overall, we argue for increased awareness of learners' understanding of and preferences for feedback and for giving learners agency in classroom activities. We hope that the practical activities outlined in this paper can also inspire further ideas for how

this can be done in the FL/L2 classroom and, thus, increase the amount of formative feedback in FL/L2 education.

Literature

- Alderson, J. C., E.-L. Haapakangas, A. Huhta, L. Nieminen, & R. Ullakonoja 2015. *The diagnosis of reading in a second or foreign language*. New York: Routledge.
- Amrhein, H. R. & H. Nassaji 2010. Written corrective feedback: What do students and teachers prefer and why? *Canadian Journal of Applied Linguistics*, 13 (2), 95–127.
- Askew, S. & C. Lodge 2000. Gifts, ping-pong and loops - Linking feedback and learning. In S. Askew (ed) *Feedback for Learning*. London: Routledge/Falmer, 1–17.
- Atjonen, P. 2007. *Hyvä, paha arviointi*. Helsinki: Tammi.
- Atjonen, P. 2014. Teachers' views of their assessment practice. *The Curriculum Journal*, 25 (2), 238–259.
- Atjonen, P. 2017. Arviointiosaamisen kehittäminen yleissivistävän koulun opettajien koulutuksessa – opetussuunnitelmatarkastelun virittämiä näkemyksiä In V. Britschgi, & J. Rautopuro (eds) *Kriteerit puntarissa*. Suomen kasvatustieteellinen seura, 131–169.
- Atjonen, P., H. Laivamaa, A. Levonen, S. Orell, M. Saari, K. Sulonen, M. Tamm, P. Kamppi, N. Rumpu, R. Hietala, & J. Immonen 2019. *"Että tietää missä on menossa"*. Publications of the Finnish Education Evaluation Centre (FINEEC), Finland, 2019:7.
- Bitchener, J., & D. R. Ferris 2012. *Written corrective feedback in second language acquisition and writing*. New York: Routledge.
- Bitchener, J., & N. Storch 2016. *Written corrective feedback for L2 development*. Bristol: Multilingual Matters.
- Bitchener, J., S. Young, & D. Cameron 2005. The effect of different types of corrective feedback on ESL student writing. *Journal of Second Language Writing*, 14 (3), 191–205.
- Brown, D. 2012. The written corrective feedback debate: Next steps for classroom teachers and practitioners. *TESOL Quarterly*, 46 (4), 861–867.
- Burke, D., & J. Pieterick 2010. *Giving students effective written feedback*. Maidenhead: McGraw-Hill Education/Open University Press.
- Ellis, R., Y. Sheen, M. Murakami, & H. Takashima 2008. The effects of focused and unfocused written corrective feedback in an English as a foreign language context. *System*, 36 (3), 353–371.
- Ferris, D. R. 2010. Second language writing research and written corrective feedback in SLA. *Studies in Second Language Acquisition*, 32 (2), 181–201.
- Ferris, D. R., H. Liu, A. Sinha, & M. Senna 2013. Written corrective feedback for individual L2 writers. *Journal of Second Language Writing*, 22 (3), 307–329.
- FNAE 2019. *National core curriculum for general upper secondary education*. Finnish National Agency for Education, Regulations and guidelines 2019:2c.
- FNBE 2016. *National Core Curriculum for general upper secondary schools 2015*. Helsinki: National Board of Education.
- Guénette, D. 2007. Is feedback pedagogically correct? Research design issues in studies of feedback on writing. *Journal of Second Language Writing*, 16 (1), 40–53. <https://doi.org/10.1016/j.jslw.2007.01.001>.
- Guénette, D. 2012. The pedagogy of error correction: Surviving the written corrective feedback challenge. *TESL Canada Journal*, 30 (1), 117–126. <https://doi.org/10.18806/tesl.v30i1.1129>.

- Hattie, J. A. C. 2009. *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Hilden, R., M. Härmälä, J. Rautopuro, M. Huhtanen, M. Puukko, & C. Silverström 2015. *Outcomes of language learning at the end of basic education in 2013*. Finnish Education Evaluation Centre, Finnish National Board of Education. <http://karvi.fi/publication/outcomes-language-learning-end-basic-education-2013/>.
- Huhta, A., C. Harsch, D. Leontjev, & L. Nieminen 2023, forthcoming. *The Diagnosis of Writing in a Second or Foreign Language: European Perspectives*. Routledge.
- Leontjev, D. 2022. Finnish Matriculation Examination and National Curriculum and Teachers' Attitudes, Perspectives, and Practices. *AFinLA Yearbook*, 133–150. <https://doi.org/10.30661/afinlavk.113656>
- Leontjev D., & P. Pollari 2022a. Guiding and assessing development of L2 writing process: the role of peer collaboration, *Innovation in Language Learning and Teaching*. Early online. <https://doi.org/10.1080/17501229.2022.2058514>
- Leontjev, D., & P. Pollari 2022b. Resolving tensions caused by high-stakes assessment in an L2 classroom through mediation. *Apples – Journal of Applied Language Studies*, 16 (1), 155–174. <https://doi.org/10.47862/apples.102950>
- Matriculation Examination Board 2022. Toisen kotimaisen kielen ja vieraiden kielten kokeiden määräykset. https://www.ylioppilastutkinto.fi/images/sivuston_tiedostot/Ohjeet/Koekohtaiset/kielikokeet_maaraykset_fi.pdf?v=270422 [accessed 24 Nov. 2022].
- Mäkipää, T. 2021a. *"I have never received feedback from my teachers": Students' perceptions of assessment and feedback practices in Finnish general upper secondary education*. An unpublished doctoral dissertation. Helsinki: University of Helsinki. <http://urn.fi/URN:ISBN:978-951-51-7644-8>.
- Mäkipää, T. 2021b. Students' and teachers' perceptions of self-assessment and teacher feedback in foreign language teaching in general upper secondary education – A case study in Finland. *Cogent Education*, 8 (1), <https://doi.org/10.1080/2331186X.2021.1978622>.
- Mäkipää, T. 2022. Lukiolaiset eivät tunnista palautetta. *Tempus*, 16–17.
- Mäkipää, T. & R. Hildén 2021. What kind of feedback is perceived as encouraging by Finnish general upper secondary school students? *Education Sciences*, 11 (12), 1–15.
- Mäkipää, T. & N. Ouakrim-Soivio 2019. Perceptions of Finnish Upper Secondary School Students of the Assessment Practices of Their Teachers. *Journal of Teaching and Learning*, 13 (2), 23–42.
- Pollari, P. 2016. A long debate over corrective feedback – why not ask students for a change? A paper presented at Eurosla 26 Conference, Jyväskylä, Finland, 24–28 August 2016.
- Pollari, P. 2017a. *(Dis)empowering assessment? Assessment as experienced by students in their upper secondary school EFL studies*. Jyväskylä Studies in Humanities, 329. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-7178-6>.
- Pollari, P. 2017b. To feed back or to feed forward? Students' experiences of and responses to feedback in a Finnish EFL classroom. *Apples-Journal of Applied Language Studies*, 11 (4), 11–33. <https://doi.org/10.17011/apples/urn.201708073429>.
- Price, M., K. Handley, J. Millar, & B. O'Donovan 2010. Feedback: all that effort, but what is the effect? *Assessment & Evaluation in Higher Education*, 35 (3), 277–289.
- Sahlberg, P. 2007. Education policies for raising student learning: The Finnish approach. *Journal of education policy*, 22 (2), 147–171.
- Stiggins, R. 2014. Improve assessment literacy outside of schools too. *Phi Delta Kappan*, 96 (2), 67–72.

- Storch, N. & G. Wigglesworth 2010. Learners' processing, uptake, and retention of corrective feedback on writing. *Studies in Second Language Acquisition*, 32 (02), 303–334.
- Truscott, J. 1996. The case against grammar correction in L2 writing classes. *Language Learning*, 46 (2), 327–369.
- Truscott, J. 2007. The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16 (4), 255–272.
- Väljjarvi, J., N. Huotari, P. Iivonen, M. Kulp, T. Lehtonen, H. Rönholm, ... S. Ohranen 2009. Lukiopedagogiikka. *Koulutuksen Arviointineuvoston Julkaisuja*, 40. https://karvi.fi/app/uploads/2014/09/KAN_40.pdf.
- Vänttinen, M. 2011. *Oikeasti hyvä numero: oppilaiden arvioinnin totuudet ja totuustuotanto rinnakkaiskoulusta yhteiskouluun*. Joensuu: Itä-Suomen yliopisto.
- Wiliam, D. 2012. Feedback: Part of a system. *Educational Leadership*, 70 (1), 31–34.
- Zhang, Z. & K. Hyland 2018. Student engagement with teacher and automated feedback on L2 writing. *Assessing Writing*, 36, 90–102. <https://doi.org/10.1016/j.asw.2018.02.004>.

APPENDIX A.

Learners' reasons for opting for direct and indirect feedback.

	what to pay attention to in future	understanding reason for mistakes	understanding what mistakes were made	recognising own weaknesses	thinking themselves rather than relying on the teacher
direct n=20	1	1	9	0	0
Indirect n=4	0	0	1	1	2

Continued

	no reason	Learners indicating they learned something incorrectly	learn something when others provide examples	learn how to develop texts to eliminate its drawbacks	Learners indicating they need help from others
direct n=20	1	1	2	7	1
Indirect n=4	0	0	0	1	0

APPENDIX B.

Checklist created from the Finnish Matriculation Examination writing scale.

Communicativeness	Content and organisation	Language breadth and accuracy
<p>Does your text target the target audience that it needs to (check the task)?</p> <p>Do you use the language that is appropriate for the task and the audience? (e.g., appropriately formal / informal)</p> <p>What sections of your text are clear and what paragraphs are still unclear or difficult to read?</p> <p>At which points does your text hold the reader’s interest? Are there points at which it fails to do that? Why?</p> <p>Have you tried to get your reader interested in your text right from the beginning? How did you do that?</p>	<p>What is the central idea of this composition?</p> <p>Which are the ideas work well and which need more elaboration?</p> <p>Where do you have enough details or examples and where should they be added? Why?</p> <p>Where is the organisation clear and where is it confusing?</p> <p>Do you use connecting devices in a versatile way?</p> <p>Did you divide your text into paragraphs? Is there just one main idea in each paragraph?</p> <p>Do you start with each paragraph with a sentence that contains or refers to the main idea of the paragraph?</p> <p>If it is an argumentative text, do you present it from more than one point of view?</p> <p>Do you repeat the same thing several times?</p>	<p>Have you tried to use varied words and phrases?</p> <p>Have you tried to use other than just basic words? (for example, ‘nice’, ‘good’, ‘very’, ‘thing’, ...)</p> <p>Have you used your verbs in the correct tense?</p> <p>Are the verb forms correct?</p> <p>Have you checked for subject—verb agreement?</p> <p>Have you used the correct prepositions?</p> <p>Have you left out the articles where they are required?</p> <p>Have you used all your pronouns correctly?</p> <p>Is your choice of adjectives and adverbs appropriate?</p> <p>Have you written in complete sentences? (narratives may differ in this)</p>

(see Leontjev & Pollari 2022a, 2022b)

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 81–96.

Magdalini Liantou

University of Jyväskylä and University of Oulu

How do Finnish and Chinese students' diverse pedagogical experiences shape feedback interpretation?

Highlights

- Grades are considered the most important form of teacher feedback for Finnish and Chinese students.
- Grades were perceived under the socio-cultural factors of each educational system.
- Finnish students were more experienced in providing peer feedback than the Chinese counterparts.
- The Chinese students included the role of “class monitor” as part of the peer feedback process.

Abstract

Due to the dissemination of joint degree programmes in higher education, more students from different educational backgrounds are exposed to the same teaching and assessment without sharing a common pedagogical culture. Since this is relatively new in Finland, little is known about how students with diverse backgrounds experience assessment compared to their Finnish classmates and how this affects their overall performance. Having as a starting point an English for Specific Purposes course offered in Finland and China, this qualitative study focuses on the role of feedback through seventeen in-depth interviews. Themes such as grades and peer feedback were interpreted based on the educational background to which the students have been previously exposed. These findings indicate that the teacher should be aware of their previous pedagogical experiences and how these affect feedback in the classroom. Additionally, peer feedback needs to be addressed explicitly by the teacher during the lessons and create more scaffolding opportunities to avoid potential misinterpretations.

Keywords: feedback, peer assessment, Classroom-based assessment (CBA), higher education

1 Introduction

Classroom-based assessment (CBA) consists of the collection of assessment results over time through tasks practised in the classroom, the teacher's constructive feedback, and the active participation of students in the assessment process via peer and self-assessment tasks (Davison & Leung 2009). These tasks are also tailored to the given curriculum. Thus, CBA is quite often bound to context and content-based assessment compared to the summative standardised test practices of the psychometric model of assessment (Lewkowicz & Leung 2021). Black and Wiliam (1998:16) point out that "all (classroom) work involves some degree of feedback between those taught and the teacher...the nature of these interactions between teachers and students, and of students with one another, will be key determinants for the outcomes of any changes." Hattie & Timperley (2007) focus on which agents can deliver feedback in the language classroom (e.g., teacher, peers, parent, self). The difference in these definitions portrays that sometimes feedback is given in a strict classroom environment, while in other cases, more stakeholders might be involved in the process. Regardless of who provides the feedback, students should be taught to interpret feedback, link it successfully with their work and utilise it to develop further (Sadler, 1998).

Regarding the feedback students provide to their classmates, peer assessment and feedback are closely associated terms and can occur simultaneously or consecutively. Liu & Carless (2006) distinguish the two terms and consider that peer assessment focuses on performance by grading their peers' work based on given criteria. In peer feedback, students provide detailed comments on each other's work to achieve better performance. However, peer feedback is often incorporated into the peer assessment process. A recent review of 58 studies focusing on the effects of peer feedback on students' learning showed a positive and nontrivial outcome

for the students who participated in peer-related activities compared to their counterparts who did not have that opportunity (Li et al. 2020). Additionally, the most important factor influencing this outcome was peer training, while the same study revealed that the positive results of peer feedback in students' learning were more meaningful than those of teacher assessment.

The context of this study is based on the Finnish and Chinese pedagogical approaches regarding language assessment. Hence, the role of feedback is discussed under the scope of CBA in these two countries. Concerning teacher assessment and feedback, Chinese university teachers tend to grade students holistically in English as a Foreign Language (EFL) classes. Their scores are not often accompanied by feedback in the form of comments due to the large size of their classes (Cheng & Wang 2007). Moreover, they frequently use multiple-choice tasks, standardised tests and translation activities, while tests are the primary assessment form. They are also less likely to engage in peer and ongoing assessment in the form of journals and portfolios if one of their key goals is to prepare students to pass the College English Test (CET) upon completing their undergraduate studies (Cheng 2008; Cheng et al. 2004). On the other hand, one of the core values of the Finnish pedagogical system is the learner-centred approach, in which students are responsible for their learning and are guided by teachers' feedback (FNAE 2019). Regarding Finnish EFL teachers' assessment practices in tertiary education, qualitative and mixed-methods studies indicate that university language centres utilise various forms of assessment through groupwork, portfolios, learning journals, student conferences, peer feedback and self-assessment forms (Jokinen et al. 2018; Liontou & Braidwood 2021). Thus, the difference in assessment tasks and feedback practices between Finland and China raises the question of whether an ESP course and its assessment could be universally applicable and accepted in the Chinese context.

When introducing various assessment types, we also have to consider the pedagogical approaches they are loaded with since they derive from the social reality surrounding a classroom. Specifically, Hu (2004) identifies three main categories for successfully implementing a pedagogical approach. The first category deals with classroom resources, and the second focuses on the impact and value of the target language in the society in which it is taught. The final one is the focal point of this study as it discusses how culture, in other words, the local context of teaching and learning, can affect the effectiveness of a pedagogical approach. Hu (2004) argues that if a pedagogical approach contradicts the local values or endorses learners' and teachers' qualities contrary to the existing ones, the introduced pedagogical approaches could clash and fail to be adopted in the new context. For example, curricula shaped by the northern secular approach often disregard the communal practices of the south (Heugh 2021); in this context, these practices could be considered the role of the parents in students' assessment and the ethical aspect of learning in China (Li 2005; Liontou 2021).

The dissemination of the Finnish educational system, namely education export, has affected higher education (HE) and has been the catalyst for “dual” undergraduate degree programmes in Asia (Schatz 2015). Thus, we sought to investigate how Finnish and Chinese students perceive teacher and peer feedback. The setting of this research involves an ESP course which belongs to the core bachelor studies for first-year engineering students in a Finnish university. Due to a Sino-Finnish joint bachelor's degree, the same course is provided at a Chinese university. The objectives of the course are: to understand the main points in English lectures, carry out necessary interactions in the academic environment in speech and writing, and communicate about engineering topics in professional contexts. Various formative and summative assessment tasks were established to achieve these aims. For the main ongoing activity of the course, the students had to create a poster presentation and present it in a mock exhibition during the final lesson. Formative assessment activities were built around the presentation task, mainly as structured teacher and peer feedback. For example, the students had to fill in self-assessment checklists at every stage of the poster presentation. Additionally, they had to practice their presentations in small groups and give peer feedback through guided prompts before the mock exhibition. Finally, students had to engage with summative assessment tasks through tests and presentations, in which both students and teachers graded students' oral work.

2 Methodology

The focus of this article was the Finnish and Chinese students' perceptions regarding feedback in an ESP course. Hence, two research questions were formed:

1. How do Finnish and Chinese students perceive teacher feedback?
2. How do Finnish and Chinese students perceive peer feedback?

2.1 Participants and data collection

Semi-structured interviews were selected for data collection since this study belongs to a larger research project. During the first stage of the project, the researcher administered two questionnaires investigating students' conceptions of assessment before and after the ESP course. Their previous responses to the questionnaires were used to create a list of predetermined interview questions investigating further teacher and peer feedback. However, follow-up questions were added when participants' answers provided critical areas of interest. King and Horrocks (2010) state that interviews provide unique information about participants' attitudes and reflect assumptions of the world. Thus, interviews could provide a thick description and a deeper understanding of their perceptions than numerical data (Denzin & Lincoln 2017).

A former student was recruited to examine the suitability of these questions in terms of appropriateness and understanding. The student was a suitable candidate to pilot the questions since he knew the course design and had the same proficiency level as the current participants. The final set of questions was created after analysing the data and revising the questions that seemed more complex.

The participants were recruited after the course completion to avoid potential biases and conflicts of interest since the researcher was one of the course instructors. All the instructors sent the information about the research and a consent form via e-mail after the course had ended. Seventeen participants from both universities agreed to be interviewed; eight were Finnish (five males and three females), and the rest were Chinese (seven males and two females). All interviews were in English, most of which took place between December 2019 and January 2020 since the final examinations of the autumn semester differ between the two countries. The interviews were conducted face-to-face in Finland. Due to the Covid-19 restrictions already imposed in China, the interviews with the Chinese participants were online. Using different modes of interviews might have affected the length of students' answers. However, limited options were available to organise the interviews due to the pandemic.

2.2 Data analysis

Thematic analysis was selected as the appropriate method since it is considered a flexible qualitative analysis not bound to any theoretical framework (Braun & Clarke 2006). The flexibility of the method was considered appropriate for this study which belongs to a broader mixed methods project. Each interview was transcribed verbatim, and the themes were created inductively. The steps to interpret the data were based on the general guidelines of thematic analysis (Braun & Clarke 2006; Clarke & Braun 2017; Javadi & Zarea 2016; Kiger & Varpio 2020; Nowell et al. 2017). After generating initial themes, they were discussed with the project supervisors and reviewed based on their feedback. Additional time was given to re-evaluating the candidate themes. Two diagrams were created to refine and name the final themes based on the research questions. The first was a spider web summarising the themes per participant, while the second diagram showed each theme across the participants. Finally, this process allowed further clarification of the level of the coded extracts and either higher-level themes were formed, or themes were separated or removed before summarising the findings and comparing them to the literature.

3 Findings

The participants discussed teacher feedback (comments and grades), its delivery and its implications in their relationship with their close ones e.g., their family. Moreover, they reflected on peer feedback based on their previous experiences and revealed the role of "class monitor" as an intermediary between the teacher and the students. It is worth noting that even though the emerged themes were similar for both groups of students, their perceptions and experiences differed. Due to the limited space of this article, additional extracts are presented in Appendices 1 and 2.

3.1 Teacher feedback

Both groups agreed that teachers delivered the most important form of feedback since they considered them experts. For instance, Finnish Participant 6 compared students' and teacher's feedback and considered the latter to be more professional:

I think when in the classroom, students gave a lot of feedback to the people who were presenting... I think that was great because the people actually listening to the presentation should be the ones to give feedback. But, I think the most qualified person to do that is the teacher. (Fin. Part. 6)

However, the Chinese students attributed an ethical role to their teachers and considered them the academic and moral leaders of the classroom. For example, Participant 7 explained that the role of the teacher is *"to pay attention to student to improve the academic performance. Teachers should also... there's a good example of moral for students. How to do in life."*

Both student groups considered grades as the primary form of teacher feedback. Chinese Participant 9 emphasised that grades instead of comments were the main focus of assessment for his classmates while they tended to disregard and undervalue the teacher's suggestions. The socio-cultural aspect of grades was an exciting finding between the two groups. All the Finnish students' responses were unanimous regarding the privacy of their grades. Participant 1 emphasised that public grades would create an unhealthy form of competition and undermine classroom collaboration:

I don't think in that kind of way it would really be encouraging competition... it would be kinda competing against other students, like not in a good way. Not collaborating as much but just trying to be as good as you possibly can. (Fin. Part. 1)

On the contrary, the Chinese participants provided mixed responses when discussing if the grades should be published. Participants 2, 6, 8, and 9 (Appendix 1) demonstrated a positive stand towards the public display of grades, justifying it as part of

their pedagogical background and culture as this used to be a common practice in their secondary education. They mentioned they could provide external motivation since they could know and compare their academic level. For instance, Chinese Participant 2 said:

I cannot know my place in my class if a teacher doesn't give all the other grades to everyone. It is hard for me to know what platform I'm in. (Chin. Part. 2)

This way of thinking conflicted with what Finnish Participant 1 mentioned about the harmful effects of competition. Additionally, Chinese Participants 1, 3, 5, and 7 (Appendix 1) focused on the adverse effects of publishing students' grades, e.g., privacy issues and bullying. As Chinese Participant 1 noticed:

...it may do harm to some students, and maybe those who get the higher grades may laugh at those who didn't. (Chin. Part. 1)

Teacher feedback, specifically grades, affected the Chinese students in the classroom and outside of it. Hence, the way teacher feedback affects students' lives should be explored holistically. Grades did not affect Finnish students' relationship with their parents (Appendix 1) since they identified themselves as adults. However, Chinese students seemed to be divided regarding grades and parent-child relationships. Their answers were split between the concepts of independence and parental care. The latter group appeared to justify their viewpoint based on the support parents offer and their own culture. For example, Chinese Participant 2 mentioned:

Yes, I think they need to know my grades because parents can help you. In China, especially in senior high school, they give the extra tutor to me because the grade is very important. In my three years in senior high school, I always have the extra tutor in class at my weekend. (Chin. Part. 2)

Most students seemed to think that their parents should be aware of their grades since they are emotionally involved and invest money in their children's education. Understandably, since these were first-year university students, they compared their current situation with their previous experiences. Only two Chinese students had different opinions regarding child-parent relationships, and the difference in their responses seemed to be their self-image. Both Participants 5 and 9 perceived themselves as adults. Nevertheless, Participant 9 still believed that the parents have the right to discuss students' grades with the teacher:

No, everyone doesn't want their parents directly to know their grades from the published scoreboard, but maybe they always say, yeah, you can ask the teachers about my grades, but don't ask me. I don't want to tell you that. (Chin. Part. 9)

3.2 Peer feedback

Apart from teacher feedback, peer feedback and assessment were essential aspects of the course. Even though students gave feedback to their peers (peer feedback) and graded each other's work (peer assessment), they used peer feedback as an umbrella term in their responses. Generally, both groups mentioned the advantages of peer feedback. For example, Finnish Participant 3 highlighted that peer feedback is important, particularly due to the course format, while Finnish Participant 4 described how the students could learn to empathise with someone's work:

It's actually good that I can give feedback because when I give feedback, I think also about how I would be in that situation presenting something and then I would I think (that) I would be very happy if someone gave me feedback. So, when I give feedback, I also think what it could mean to the other person. (Fin. Part. 4)

Moreover, Finnish Participant 7 compared peer feedback to teacher feedback and gave a concrete example of why peer feedback can be less stressful for students. In the same spirit, Chinese participants pointed out that students can learn from each other through peer feedback.

Like, if it's negative feedback, usually, if your friend gives it or a student, like if it doesn't feel as... you've failed, or you've done something wrong. Like, for example, I used to know my friend... when he was speaking or he tried to say some words, I would say, hey, it's this one or you cannot say it this way. And he was, wondering, okay, how do I write this? He wrote it wrong. I would explain to him, okay, do you have to do it this way?... It's different when the teacher says it because I think sometimes students feel a bit of pressure. (Fin. Part. 7)

I think it's good because every team member has shortcomings and advantages. So, I can find my shortcomings, and I can learn from their advantages from them. (Chin. Part. 7)

However, at a closer look, both groups had second thoughts about the effectiveness and willingness to provide peer feedback. They considered students' feedback lenient than the teacher's and usually too positive (see Appendix 2 for further examples of milder peer feedback). For instance, Finnish Participant 8 noticed that in her presentation... *"where I got the feedback on the group, there were mistakes, there were things that I would have probably pointed out to myself, but they didn't."* Some participants would not take it seriously, while Chinese Participant 8 mentioned that students could experience group conflicts.

...If there's a conflict or feedback is too much, we will get... break our relationship... will break, but sometimes if the feedback is hard to fit so hard to accommodate. People like me, (they are) not willing to give feedback. (Chin. Part. 8)

Furthermore, Chinese students mentioned a lack of similar previous experiences and, consequently, a lack of training in this type of feedback. Chinese Participants 1 and 9 discussed that since this type of activity is rare in China, taking the initiative to provide peer feedback might be difficult for some students. Chinese Participant 9 added that peer feedback was a form of leadership since the students had to speak their minds. In his own words:

They are shy to work, and as a result, even though they have something in their mind to say, but they don't dare to say they're afraid because they think they are not good, so the effect of the group is not that good. Because leadership is not the common culture of Chinese cannot change just in two weeks. (Chin. Part. 9)

Chinese students consider the teacher a leadership figure, so peer feedback might be something that opposes this mindset and creates internal conflict. Even though Chinese students were not trained in peer feedback, Participant 9 considered the existence of the "class monitor" as one who provides feedback to their peers. The role of the class monitor is to be a mediator between the teachers and the students and decide what kind of feedback will be delivered to both parties. This role seems to be embedded in the Chinese system and is not similar to the traditional idea of peer feedback, e.g., students give feedback to their classmates during classroom activities. This is an extract of the conversation in which Participant 9 discussed how one should give peer feedback:

Just like there are 50 of us, but there's only one monitor in the classroom. Some students, they do not know manners, they also say some bad words. So, at least this feedback will not directly go into the teacher's head, they should also go through the monitors head that the monitor has it... it just has to go through someone before the letters go directly into that, because, I know, I understand the feeling when it's received the letters full of bad words, criticising you, so this kind of work, maybe you can make it more polite, not just too direct... So, I think there should be someone between the teachers and the students. (Chin. Part. 9)

This new piece of information raises many questions regarding peer assessment in the Chinese context. First, it is crucial to know who guides peer feedback. Usually, this role belongs to the teacher who gives classroom instructions; however, the teacher might not be the only person in this context. This brings us to the next question on how the teachers can ensure that there is no power relationship when receiving or delivering peer feedback to the class monitor. In other words, it raises questions on how the teachers can be sure that the students will not confuse the class monitor's dual role such as the general responsibilities of the class monitor and the feedback this student provides or receives on a specific peer feedback activity. Finally, the role of class monitors is further discussed in the next section.

4 Discussion

In the last twenty years, research on CBA has primarily focused on teachers' beliefs and assessment literacy and various forms of assessment and feedback (Lewkowicz & Leung 2021). Classroom-based feedback encompasses classroom discussion, peer and self-assessment, teacher-led comments and grades (Black & Wiliam 1998). Nevertheless, no matter the nature of feedback, it only becomes meaningful when students consider the given feedback, take action and try to close the gap in learning (Sadler, 1998). Thus, this study aims to discuss Finnish and Chinese tertiary students' perceptions of teacher and peer feedback in an ESP course. The diverse educational culture of students is considered a factor affecting the values resting upon feedback and the implications it has in their lives (Turner & Purpura 2016).

Sadler (1998) has explicitly discussed how students acknowledge teachers' expertise in assessment due to the teacher's superior knowledge of the subject. Both groups in this study also reported a similar finding. Moreover, Chinese participants added a moral hue to teacher feedback since they considered their teachers as moral leaders. Li (2005) and Liantou (2021) discovered similar findings when they investigated Chinese students' perceptions of learning and assessment. Students strived not only for knowledge but also for moral guidance. Finnish and Chinese students' viewpoints regarding teacher feedback echo the concept of northern and southern educational settings, such as individualism versus collectivism (Heugh 2021). The Chinese educational system emphasises social learning through uniformity and leading by example (Zhu & Chang 2019). Social learning transfers valuable knowledge to younger generations while promoting societal stability in contrast with individual learning, which focuses on trial and error (Chang et al. 2011). In comparison, learner agency and various forms of formative assessment, such as self and peer assessment, are promoted in language courses in Finnish tertiary education (Jalkanen et al. 2015; Jokinen et al. 2018). This mismatch in pedagogical traditions could create misconceptions about teacher feedback since Chinese students might expect more from their teachers.

Scoring without incorporating comments is a standard assessment practice in EFL classrooms in China (Cheng & Wang 2007). This practice could explain Chinese students' disinterest in teachers' comments in this study. Students also mentioned the public display of grades and the consequences of this action on them and their close ones. Chinese students' opinions align with previous findings regarding the involvement and expectations of Chinese parents in their children's academic life (Biggs & Watkins 2001; Liantou 2021; Peterson et al. 2013). For Chinese students, learning has moral and social implications and carries a higher status for them, so to improve their learning, students strive for self-perfection (Li 2005). In case of failure, this leads them to feelings of shame and guilt, affecting not only them but also their parents. This can be explained by the filial piety that underpins Chinese culture, in

which family interdependence is highly valued (Zhu & Chang 2019). Unlike Finland, in HE in China, students and parents expect the teacher to have a key role in education in many ways, one manifestation of which is parents' and students' feedback to the teacher.

Chinese students tended to compare grades and/or each other's work. Finnish students showed empathy when provided peer feedback, which had a positive effect as it could help students decompress from teacher feedback and be more open to accepting it from their classmates, especially their friends. Explicit comparison in the classroom triggers internal feedback and can be utilised to enhance higher-order thinking. For example, analogical, e.g., reviewing someone else's work, and analytical comparison, e.g., someone else reviewing their work, could promote different types of internal feedback (Nicol 2021). However, teachers need to be cautious when introducing comparisons to avoid harmful associations between feedback and learning. Therefore, comparisons through structured peer feedback could be a great alternative to stimulate high-order thinking.

Regarding peer feedback, Finnish and Chinese students seem to agree with previous research, which stresses that some of the shortcomings of peer feedback are reliability, perceived expertise, power relations and time (Liu & Carless 2006). Even though the concept of time was not mentioned in this study, the rest of the pitfalls of peer feedback were evident. As Topping (2017) discusses, social processes such as social anxiety, friendships and power relations can negatively affect peer assessment. For example, students could give similar feedback and grades to avoid conflict. Teachers could overcome this obstacle by promoting students' responsibility and the feeling of ownership of their work. They should also consider the students' roles, such as class monitors, and try to mitigate when and if needed. In order to successfully implement various forms of assessment and promote inclusion in the Finnish and Chinese language classroom, students' and teachers' systems of knowledge need to be first acknowledged and then respected mutually. This action could correct the imbalance between the west-south dichotomy of education (Heugh 2021).

Finally, students are affected by the socio-political situation of their pedagogical system to some extent (Turner & Purpura 2016). For example, an unexpected finding of the contextual dimension of feedback was the Chinese concept of "class monitor". The class monitor's function is different from the way peer feedback is typically described in research, e.g., the students giving feedback to their classmates during classroom activities. Generally, the concept of the class monitor was not mentioned in any academic study to the researcher's knowledge, apart from the blog post "Class Monitors in China" (To boldly go... 2017). In this popular article, an English teacher in China describes that the class monitors are usually selected by their headteachers, and their role is to organise the class based on the needs of the university. Various monitors (e.g., behaviour monitor, sports monitor, league secretary) are selected

either by the communist party or the university. As a student in this present study mentioned, the class monitor was responsible for delivering feedback from the students to the teachers and vice versa. This unique role in the Chinese HE is not part of the Finnish HE system. The monitors seem to need to think about the students' views, interpret them and deliver the feedback to either the teacher or the students. Overall, what is required from the class monitor differs from the other students who give feedback to their peers; thus, it raises issues for further research.

5 Conclusion

This paper investigated the perspectives of Finnish and Chinese students participating in an ESP course for engineers. Since the students were previously exposed to different pedagogical systems, there was ambiguity about how they perceived teacher and peer feedback, equally essential pillars of the course. Various implications for similar projects in the Finnish education export sector and methodological limitations are discussed below.

In order to better understand how the teacher and peer feedback are perceived in the classroom, we should acknowledge that students are carriers of different systems of knowledge and understandings of the world since they derive from heterogeneous contexts (Heugh 2021). For example, when introducing peer feedback, the teachers should consider not only the course syllabus but also the educational background of their students to tailor peer feedback accordingly. As previous research has emphasised, introducing specific criteria and training in peer assessment are paramount to achieving positive learning outcomes (Li et al. 2020; Li & Gao 2016). This study confirms that students with little or no experience struggled to understand and practice this assessment form adequately. Extra effort and time are crucial to successfully implement formative feedback, especially in cultures where the opposite, e.g., top-down instructions and competition, are highly valued (Hu 2004; Zhu & Chang 2019). When the feedback is communicated in the classroom, it is also worth considering all the stakeholders, such as class monitors and parents. The pathways of classroom feedback in various educational cultures should also be explored in the future, as it could affect either positive or negative classroom interactions and undermine the validity and process of assessment. As this research showed, there might be various implications regarding feedback in and out of the class, and thus, an emic approach to feedback is worth exploring in the language classroom. Through this approach, we could, for example, better understand the role of the class monitor explicitly and see how it affects the class dynamics and, consequently, the assessment process. Finally, it is worth acknowledging that the different types of interviews, due to the unique circumstances of the pandemic, could have potentially affected the length of the participants' answers.

Literature

- Biggs, J.B. & D.A. Watkins 2001. Insights into teaching the Chinese learner. In D.A. Watkins & J. B. Biggs (eds) *Teaching the Chinese Learner: Psychological and Pedagogical Perspectives*. Hong Kong: University of Hong Kong, Comparative Education Research Centre, 277–300.
- Black, P. & D. Wiliam 1998. Assessment and classroom learning. *International Journal of Phytoremediation*, 21 (1), 7–74. <https://doi.org/10.1080/0969595980050102>
- Braun, V. & V. Clarke 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3 (2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Chang, L., M.C.K., Mak, T., et al. 2011. Cultural Adaptations to Environmental Variability: An Evolutionary Account of East–West Differences. *Educational Psychology Review*, 23, 99–129. <https://doi.org/10.1007/s10648-010-9149-0>
- Cheng, L. 2008. The key to success: English language testing in China. *Language Testing*, 25 (1), 15–37. <https://doi.org/10.1177/0265532207083743>
- Cheng, L., T., Rogers & H. Hu 2004. ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. *Language Testing*, 21 (3), 360–389. <https://doi.org/10.1191/0265532204lt288oa>
- Cheng, L. & X. Wang 2007. Grading, Feedback, and Reporting in ESL/EFL Classrooms. *Language Assessment Quarterly*, 4 (1), 85–107. <https://doi.org/10.1080/15434300701348409>
- Clarke, V., & V. Braun 2017. Thematic analysis. *Journal of Positive Psychology*, 12 (3), 297–298. <https://doi.org/10.1080/17439760.2016.1262613>
- Class Monitors in China*. To boldly go... 2017. <https://toboldlygo.mystrikingly.com/blog/class-monitors-in-china> [accessed 5 August 2022].
- Davison, C. & C. Leung 2009. Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43 (3), 393–415. <https://doi.org/10.1002/j.1545-7249.2009.tb00242.x>
- Denzin, N. K. & Y.S. Lincoln 2017. *The SAGE handbook of qualitative research* (Fifth edition). Los Angeles: Sage.
- FNBE 2019. *National Core Curriculum for General Upper Secondary Schools 2019*. Finnish National Board of Education: Helsinki, Finland.
- Hattie, J. & H. Timperley 2007. The Power of Feedback. *Review of Educational Research*, 77 (1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heugh, K. 2021. Southern multilingualisms, translanguaging and transknowledging in Inclusive and Sustainable Education. In P. Harding-Esch & H. Coleman (eds) *Language and the Sustainable Development Goals*. British Council, 37–47.
- Hu, G. 2004. Pedagogical practices in Chinese EFL classrooms. *Asian Englishes*, 7 (1), 42–59. <https://doi.org/10.1080/13488678.2004.10801130>
- Jalkanen, J., Jokinen, E. & Taalas, Peppi (Eds). 2015. *Voices of pedagogical development – expanding, enhancing and exploring higher education language learning*. Research-publishing.net. <https://doi.org/10.14705/rpnet.2015.9781908416261>
- Javadi, M., & K. Zarea, K. 2016. Understanding Thematic Analysis and its Pitfall. *Journal of Client Care*, 1 (1), 34–40. <https://doi.org/10.15412/JJCC.02010107>
- Jokinen, J., S., Karjalainen & H. Mäkäläinen 2018. *Kielenoppimisen kehittyvät arviointi - ja palautekäytännöt - Developing feedback and assessment practices in language learning*. Helsinki: Helsingin yliopiston kielikeskus.
- Kiger, M. E., & L. Varpio 2020. Thematic analysis of qualitative data: AMEE Guide. *Medical Teacher*, 42 (8), 846–854. <https://doi.org/10.1080/0142159X.2020.1755030>
- King, N. & C. Horrocks 2010. *Interviews in qualitative research*. Los Angeles: Sage.
- Lewkowicz, J. & C. Leung 2021. Classroom-based assessment. *Language Teaching*, 54 (1), 47–57. <https://doi.org/10.1017/S0261444820000506>

- Li, J. 2005. Mind or Virtue. *Current Direction in Psychological Science*, 14 (4), 190–194.
- Li, H., Y., Xiong, C. V., Hunter, X., Guo & R. Tywoniw 2020. Does peer assessment promote student learning? A meta-analysis. *Assessment and Evaluation in Higher Education*, 45 (2), 193–211. <https://doi.org/10.1080/02602938.2019.1620679>
- Liu, N. F. & D., Carless 2006. Peer feedback: The learning element of peer assessment. *Teaching in Higher Education*, 11 (3), 279–290. <https://doi.org/10.1080/13562510600680582>
- Liontou, M. 2021. Conceptions of assessment as an integral part of language learning: A case study of Finnish and Chinese university students. *Languages*, 6 (4), 202–223. <https://doi.org/10.3390/languages6040202>
- Liontou, M., & E. Braidwood 2021. Mediation in practice in an ESAP course: Versions of the Medical English student conference. *CEFR Journal - Research and Practice*, 4, 25–42. <https://doi.org/10.37546/JALTSIG.CEFR4-2>
- Nicol, D. 2021. The power of internal feedback: exploiting natural comparison processes. *Assessment and Evaluation in Higher Education*, 46 (5), 756–778. <https://doi.org/10.1080/02602938.2020.1823314>
- Peterson, E. R., G. T. L., Brown & R. J. Hamilton 2013. Cultural differences in tertiary students' conceptions of learning as a duty and student achievement. *International Journal of Quantitative Research in Education*, 1 (2), 167–181. <https://doi.org/10.1504/ijqr.2013.056462>
- Sadler, D. R. 1998. Formative assessment: Revisiting the territory. *International Journal of Phytoremediation*, 21 (1), 77–84. <https://doi.org/10.1080/0969595980050104>
- Schatz, M. 2015. Toward One of the Leading Education-Based Economies? Investigating Aims, Strategies, and Practices of Finland's Education Export Landscape. *Journal of Studies in International Education*, 19 (4), 327–340. <https://doi.org/10.1177/1028315315572897>
- Topping, K. 2017. Peer Assessment: Learning by Judging and Discussing the Work of Other Learners. *Interdisciplinary Education and Psychology*, 1 (1), 1–17. <https://doi.org/10.31532/interdiscipeducpsychol.1.1.007>
- Turner, C.E. & J. Purpura 2016. Learning-oriented assessment in second and foreign language classrooms. In D. Tsagari & J. Banerjee (eds.), *Handbook of Second Language Assessment*. Berlin: De Gruyter/Mouton, 255–273.
- Zhu, N. & L. Chang 2019. Education and Parenting in China. In E. Sorbring & J. Lansford (eds), *School Systems, Parent Behavior, and Academic Achievement. Young People and Learning Processes in School and Everyday Life*. Springer, 15–28. https://doi.org/10.1007/978-3-030-28277-6_2

APPENDIX 1.

Quotes regarding grades given by the teacher

TGrades	
Fin. Part. 4	<i>no, I don't think my parents need to know my grades (laughing) it's personal and I mean you can always tell the parent if the grade... if you want but... if the parent keeps on watching or knowing the grades it's not right I think ... the parent doesn't have the right almost to an adult man or woman's grade... it is not really their business</i>
Fin. Part. 3	<i>I don't think they need to know anymore because I'm away and I am basically an adult... I would like when you're like a kid and just be like... when you like in high school or definitely like first 10 years or so...</i>
Chin. Part. 2	<i>I cannot know my place in my class if a teacher doesn't give all the other grades to everyone. it is hard for me to know what platform I'm in. Because your grade by Chinese student, it's very important thing from junior school to senior high school. grades is the most important for everyone because it decides whether you cannot go to the university. What or how, how the university is ... if you can join... In China, Senior High School, our exam, after every exam the grades are being displayed. And the teacher will give it to all of us.</i>
Chin. Part. 6	<i>I think that this can help us to know the ability of others ... Yes, I think teachers can get the grades to choose public but not students... I think teachers can place them in our session in Moodle and not put them on the wall or somewhere you choose them for the others who get the lower grades I think what the best is for teacher to put them in Moodle</i>
Chin. Part. 8	<i>I think that someone might want to know about the high the higher score and they should publish them but if they are not very high teachers have better not to...</i>
Chin. Part. 9	<i>if the full class being Chinese because first seeing to the junior school and they were told no school, no junior school will publish the grades we have but actually we just publish the tests that are being valued and see my classroom before used to know and who have bad grades. I want to see is how I got a grade they publish on the school board to see yourself and also review your work hard and get good grades. It is good, but if you are not, it's not that good. So, I believe they should publish but maybe the teachers should also comfort the last one, though, if they want to work harder if the last time, they can be better than they thought.</i>
Chin. Part. 3	<i>Grade is just a measure of what I have started and the relationship with my classmates were may not break because of that... A student's grades are private, and the teacher should not do that.</i>
Chin. Part. 5	<i>I don't want the teachers to show my assessment to others. Because I think it's my assessment and I don't want others to know. I don't want to see others' assessment.</i>
Chin. Part. 7	<i>I think my grades won't affect my relationship with my classroom, because I may pay more efforts to my class so I could get a higher score. I think it's (grade) private. If someone it gets a low grade, he may be ashamed</i>

APPENDIX 2.

Quotes regarding peer feedback

Peer feedback	
Fin. Part. 5	<i>I feel like giving feedback it's hard for me because I don't want to criticise other people's works. But I think it's necessary if I got, because that's how they learn and improve their stuff</i>
Fin. Part. 8	<i>Like, I know, they are not probably going to say as Truthfully, I all those things. So, all I got was basically positive way. So that's always nice to hear about taking that with consideration. I noticed myself that in my presentations, where I got the feedback on the group, there was mistakes, there was things that I would have probably pointed out to myself, but they didn't.</i>
Chin. Part. 2	<i>When I give feedback to my partners, sometimes they will agree sometimes they don't agree. If there's conflict of feedback is too much, we will get... break our relationship... but sometimes if the feedback is hard to fit so hard to accommodate. People like me, not willing to give feedback. We are not willing to share feedback because it hurts</i>
Chin. Part. 1	<i>I think that's great. One can present and then other listen. When you are giving feedback, your partner will know whether they are doing a good job, but there is a problem. Sometimes students in China are not very interested in this form of activity. So, they're not willing to give feedback... I don't think they take it (peer feedback) ... this very seriously. They just give casual feedback. It's honest. Or some of their feedbacks are just sort of like joking.</i>

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 97–123.

Anssi Roiha
Turun yliopisto

Juuso Nieminen
University of Hong Kong

Katja Mäntylä
Jyväskylän yliopisto

Lukion vieraiden kielten opettajien käytänteitä ja käsityksiä eriyttävästä arvioinnista

Nostot

- Tutkimuksessa paneuduttiin lukion vieraiden kielten opettajien käytänteisiin ja käsityksiin eriyttävästä arvioinnista.
- Opettajat käyttivät monipuolisesti erilaisia arviointimenetelmiä ja painottivat formatiivista arviointia.
- Arvioinnin eriyttäminen oli kuitenkin usein vähäistä.
- Opettajat keskittyivät arvioinnissaan enemmän alaspäin kuin ylöspäin eriyttämiseen.

Abstract

This study explores the perceptions and practices of upper secondary school language teachers on differentiated assessment. The data were collected through a closed and open-ended questionnaire (n=48) and individual interviews with four teachers. The results show that the teachers used a variety of assessment methods, with a predominant emphasis on formative assessment. Although the teachers mentioned many differentiated assessment methods, they did not generally differentiate their assessments to a large extent. The teachers also focused mostly on low-achieving students and neglected adversely gifted students. The teachers seemed to follow the practice of having all students demonstrate their learning in the same way and at the same time. Class size was seen as the main barrier to more thorough differentiation. The results of this study suggest that differentiation should be better integrated into assessment and that both pre- and in-service teachers would benefit from training focusing on differentiated assessment.

Keywords: assessment literacy, differentiated assessment, upper secondary school, foreign languages

Asiasanat: arviointiosaaminen, eriyttävä arviointi, lukio, vieraat kielet

1 Johdanto

Viime vuosina ymmärryksemme siitä, mitä onnistunut vuorovaikutus monimediaisessa ympäristössä vaatii, on laajentunut. Tämä muutos on samalla vaikuttanut siihen, miten kielitaitoa arvioidaan (Purpura 2016). Kielenoppijan tulee paitsi ymmärtää ja tuottaa vierasta kieltä myös ratkaista ongelmia, arvioida tietoa ja tehdä päätöksiä kohdekielellä. Näiden tärkeiden taitojen on oltava arvioinnin kohteena, mikä haastaa arviointia monipuolisuuteen ja monimediaisuuteen. Tässä artikkelissa lähestymme lukion kielitaidon arviointia erityisesti eriyttämisen näkökulmasta. Aihe on tärkeä, sillä monipuolistuneiden kielenopetuksen tavoitteiden myötä arvioinnissa on korostunut entisestään yksilöllisyys ja persoonallisuus.

Eriyttäminen on viimeisen reilun kymmenen vuoden aikana noussut Suomessa yhdeksi keskeiseksi opetusta ohjaavaksi käsitteeksi erityisesti perusopetuksessa, jossa se nähdään kaiken opetuksen pedagogisena lähtökohtana (POPS 2014). Lukiotasolla eriyttämiselle ei ole annettu aivan yhtäläistä painoarvoa. Lukion opetussuunnitelman perusteissa (LOPS 2019) sana eriyttäminen esiintyy kolme kertaa. Se tuodaan esiin oppimisen tukea käsittelevässä kohdassa yhtenä mahdollisena tukikeinona muun muassa tukiopetuksen ohella. Lisäksi opetussuunnitelmassa eriyttäminen mainitaan opiskelijan ohjauksen ja liikunnan opetuksen yhteydessä.

Täysipainoinen opetuksen eriyttäminen ulottuu myös arviointiin, mutta Suomessa lukion valtakunnallinen opetussuunnitelma tarjoaa arvioinnin eriyttämiseen vähäisesti työkaluja. LOPS (2019) määrittää arvioinnille kaksi keskeistä tehtävää: 1) oppimisen tukeminen ja ohjaaminen sekä 2) osaamiselle ja oppimiselle asetettujen tavoitteiden saavuttamisen näkyväksi tekeminen. LOPSissa ensimmäinen tavoite on nivottu

formatiiviseen arviointiin ja jälkimmäinen summatiiviseen. Oppimisen tukemisen ja ohjaamisen kuvauksessa keskiöön nostetaan opiskelijan ja opettajan vuorovaikutus ja oppimisen prosessi. Arvioinnin yleisissä tavoitteissa mainitaan arvioinnin monipuolisuus, joskaan eriyttämistä ei tässä yhteydessä käsitellä. (LOPS 2019: 45–46.) Kun puhutaan opintojaksojen arvioinnista, LOPSissa otetaan kantaa myös oppimisen tuen tarpeisiin ja arvioinnin eriyttämisen mahdollisuuksiin muun muassa seuraavasti:

Oppimisen tuen tarpeet, kuten sairaudesta tai vammasta johtuvat haasteet, lukemisen tai kirjoittamisen erityisvaikeus, maahanmuuttajien kielelliset vaikeudet sekä muut syyt, jotka vaikeuttavat osaamisen osoittamista, tulee ottaa arvioinnissa huomioon siten, että opiskelijalla on mahdollisuus erityisjärjestelyihin ja vaihtoehtoisin tapoihin osoittaa osaamisensa. (LOPS 2019: 46.)

Vieraisissa kielissä LOPSin tavoitteiden ja sitä kautta arvioinnin perustana ovat *Eurooppalaiseen viitekehukseen* (Euroopan neuvosto 2001) pohjaavat tavoitetasot ja niiden suomalainen sovellus. Opiskelijoiden yksilöllisyys ei tässäkään kontekstissa nouse vahvasti esiin arvioinnista puhuttaessa. Erikseen mainitaan, että ”Monipuolinen arviointi ja onnistunut, kannustava palaute tukevat opiskelijan minäpystyvyyttä ja vahvistavat opiskelumotivaatiota.” (LOPS 2019: 178.) Vaikka eriyttämisestä ei juurikaan arvioinnin kohdalla puhuta, voisi LOPSissa mainitun kieliprofilin tulkita esimerkiksi oppimisprosessin ja tavoitteiden eriyttämisen välineestä: uuden LOPSin myötä jokaisen lukio-opiskelijan tulee laatia henkilökohtainen kieliprofili, jossa hän pohtii omia kielenoppimisen tavoitteitaan ja reflektoi oppimisprosessiaan.

Suomalaisen lukion vieraiden kielten opetuksen kontekstissa on siis tarpeen selvittää, miten eriyttäminen ymmärretään ja miten sen toteuttamisesta käytännön arviointityössä raportoidaan. Uusi LOPS (2019) ohjaa opettajia varsin monipuoliseen arviointiin, jossa korostuvat oppimisprosessin ohjaaminen ja oppimisen tuen tarpeiden huomiointi. Kuitenkin suomalaiset vieraiden kielten opettajat ovat tyypillisesti toteuttaneet arviointia yksin tehtävien kokeiden avulla, mikä on korostunut erityisesti juuri lukiossa (Atjonen ym. 2019). Suomessa on viime aikoina tutkittu jonkin verran opiskelijoiden käsityksiä ja kokemuksia arvioinnista vieraiden kielten opetuksen kontekstissa (Ilola 2018; Mäkipää & Hildén 2021; Pollari 2017). Samoin on tutkittu vieraiden kielten opettajien arviointikäytänteitä yleisesti (Hildén & Härmälä 2015). Vaikka vieraan kielen opettajien käsityksiä ja näkemyksiä on tutkittu useista eri näkökulmista (Gregersen & Mercer 2021), erityisesti oppimista tukevan ja eriyttävän arvioinnin näkökulma on usein jäänyt paitsioon. Tässä artikkelissa pyrimme osaltamme täyttämään tätä tutkimusaukkoa. Tutkimustehtävämme on perehtyä lukion vieraiden kielten opettajien arviointiosaamiseen (*assessment literacy*) eli siihen, millaisia käsityksiä opettajilla on arvioinnista ja sen eriyttämisestä ja miten he kertovat käytännössä toteuttavansa sitä. Tarkemmat tutkimuskysymykset ovat:

1. Kuinka lukion vieraiden kielten opettajat kertovat toteuttavansa käytännössä arviointia ja sen eriyttämistä?
2. Mitä käsityksiä opettajilla on arvioinnista ja sen eriyttämisestä?

Artikkelimme rakentuu seuraavasti: luvussa 2 esittelemme tutkimuksen teoreettisen taustan. Luvussa 3 kuvaamme tutkimuksessa käytettyjä menetelmiä. Luvussa 4 käsittelemme tutkimuksen tuloksia. Luvussa 5 pohdimme johtopäätöksiä tulosten pohjalta sekä esittelemme tutkimuksesta nousseita jatkotutkimusaiheita.

2 Tutkimuksen teoreettinen tausta

2.1 Mitä on eriyttäminen?

Eriyttämiselle ei löydy yhtä selkeää määritelmää, vaan se ymmärretään hyvin eri tavoin sekä tutkimuksissa että käytännön opetuksessa (Graham ym. 2021). Usein eriyttäminen käsitetään varsin suppeasti, jolloin sen ajatellaan olevan vain valikoima mekaanisia opetuskäytänteitä (esim. Benjamin 2002), ainoastaan reaktiivista toimintaa (esim. Hall 2002) tai sen ymmärretään keskittyvän pelkästään taitotasoihin (esim. Roy ym. 2013; Saloviita 2018). Laajan näkemyksen mukaan eriyttämistä voidaan puolestaan pitää opetuksen lähestymistapana, jossa jo lähtökohtaisesti huomioidaan jokaisen opiskelijan sekä opiskelijaryhmien moninaisuus, jolloin se koskee myös arviointia (Jager ym. 2022; Roiha & Polso 2020; Tomlinson 2014). Lisäksi tämän näkemyksen mukaan eriyttäminen on sekä reaktiivista että proaktiivista toimintaa, jossa keskitytään taitotasojen lisäksi myös esimerkiksi opiskelijoiden kiinnostuksen kohteisiin tai motivaatioon. Käytännössä eriyttämisen ymmärretään kuitenkin usein koskevan vain lahjakkaita opiskelijoita tai niitä, joilla on oppimisvaikeuksia (esim. Brevik ym. 2018; Nepal ym. 2021). Suomalaisessa opetusdiskurssissa puhutaankin yleisesti ylöspäin ja alaspäin eriyttämisestä (Roiha & Polso 2018). Yllä esiteltyjen määritelmien valossa LOPSin (2019) voidaan nähdä nojautuvan jokseenkin suppeaan eriyttämiskäsitykseen: eriyttäminen muun muassa rinnastetaan tukiopetukseen, joka laajan määritelmän mukaan on vain yksi eriyttämiskeino (Roiha & Polso 2021b).

Aiemmissa tutkimuksissa opettajilla on todettu olevan varsin rajoittuneet käsitykset eriyttämisestä (esim. Melesse 2015; Milinga ym. 2022; Santangelo & Tomlinson 2012). Lisäksi opettajien raportoimat eriyttämiskäytänteet ovat olleet usein varsin suppeita (esim. Gaitas ym. 2022; Pozas ym. 2020). Monet opettajat ovat pitäneet eriyttämistä myös haastavana toteuttaa. Tyypillisiä heidän mainitsemiaan haasteita ovat materiaalit ja resurssit, arvioinnin eriyttäminen, aikarajoitteet, luokkakoot, tiedon puute eriyttämiskäytännöistä sekä hallinnollisen tuen puute (esim. Gaitas & Martins 2017; Lavania & Nor 2020; Roiha 2014; Shareefa ym. 2019; Tomlinson & Imbeau 2010; VanTassel-Baska & Stambaugh 2005; Wan 2017).

2.2 Arviointi

2.2.1 Opettajan arviointiosaaminen

Opettajien arviointiosaaminen on tunnistettu pitkään opettajan ammattitaidon tärkeäksi rakennuspalikaksi (esim. Stiggins 1995). Arviointiosaamisen käsitteellä on perinteisesti tarkoitettu opettajan kykyä harjoittaa tarkkaa, luotettavaa ja monipuolista arviointia (Popham 2009). Opettaja, jolla on hyvä arviointiosaaminen, osaa käyttää luotettavasti erilaisia arviointimenetelmiä opiskelijoiden oppimisen tukemiseksi (Lam 2015).

Tässä tutkimuksessa nojautumme Xun ja Brownin (2016) kirjallisuuskatsauksessa esiteltyyn malliin opettajan arviointiosaamisesta. Malli kurkottaa pidemmälle aikaisemmista käsitteellistyksistä, jotka ovat keskittyneet arviointiosaamisen mitaamiseen ja edistämiseen. Xu ja Brown analysoivat työssään tarkemmin sitä, mistä tämä ”osaaminen” (*literacy*) lopulta koostuu. He erittelivät kirjallisuuskatsauksessaan opettajan osaamisen neljään eri tasoon, jotka muodostavat yhdessä kumuloituvan mallin: 1) tietopohja arviointikäytännöistä, 2) arviointikäsitteet, 3) arviointiosaamisen soveltaminen käytännössä sekä 4) arviointi-identiteetti. Huomionarvoista on, että opettajan arviointiin liittyvät käsitykset ja tietoisuus niistä ymmärretään kiinteäksi ja integroituneeksi osaksi mallia, ei siitä erillään olevaksi. Xu ja Brown ovat eritelleet ja ryhmitelleet tarkemmin kahta ensimmäistä arviointiosaamisen tasoa. Arviointikäytäntöihin liittyvää tietoa on esimerkiksi jaoteltu itse- ja vertaisarviointiin sekä arviointitietoihin liittyvään tietoon. Eriyttämiseen liittyviä kysymyksiä ei heidän mallistaan kuitenkaan löydy, joten on paikallaan tarkastella nimenomaan arvioinnin eriyttämiseen liittyviä käytänteitä sekä käsityksiä.

2.2.2 Eriyttävä arviointi

Arviointi on olennainen osa opetuksen eriyttämistä, ja se otetaan huomioon useimmissa eriyttämismalleissa. Tomlinsonin (2014) eriyttämismalli koostuu neljästä ulottuvuudesta: sisällöistä, prosesseista, tuotoksista ja oppimisympäristöstä. Arvioinnin eriyttäminen liittyy tässä mallissa ennen kaikkea tuotoksiin. Tomlinsonin (2014) mukaan eriyttävässä opetuksessa korostuu nimenomaan diagnostinen ja jatkuva arviointi. Rock ym. (2008) ovat luoneet eriyttämiseen viisivaiheisen REACH-mallin, johon yhtenä osana kuuluu opetuksen arviointi. He painottavat niin diagnostista, formatiivista kuin summatiivista arviointia eriyttämisessä sekä opiskelijoiden osallistamista arviointiprosessiin. Niin ikään Prastin ym. (2015) kehämäisessä eriyttämismallissa arvioinnilla on suuri rooli sekä eriyttämistarpeiden tunnistamisessa että eriyttämiskeinojen tehokkuuden arvioinnissa. Myös Roihan ja Polson (2021b) ”Viiden O:n malliin” oppimisen arviointi kuuluu sen viimeisenä eriyttämisen osa-alueena. Ajatuksena tässä mallissa on, että tuen ollessa kunnossa mallin muilla osa-alueilla (opetusjärjestelyt, oppimisympäristö, opetusmenetelmät ja oppimisen tukimateriaali), voidaan opiskelijoiden oppimista arvioida luotettavasti ja yksilöllisesti.

Huolimatta keskeisestä roolista, joka useissa eriyttämismalleissa arvioinnille annetaan, empiiriset tutkimukset eriyttävästä arvioinnista ovat etenkin toisen asteen koulutuksen saralla vähäisiä. Aikaisemmissa tutkimuksissa on tavallisesti kartoitettu opettajien eriyttämistoimintaa yleisemmin, jolloin arviointikäytänteet ovat olleet vain yhtenä osa-alueena. Muun muassa Smitin ja Humpertin (2012) tutkimuksen mukaan opettajat hyödynsivät vaihtoehtoisia arviointikäytänteitä hyvin harvoin. Kotimaisen tutkimuksen kentällä eriyttämistä ei ole vielä tutkittu kattavasti (ks. poikkeuksena mm. Laari ym. 2021; Roiha 2014). Saloviidan (2018) tutkimus tosin paljastaa, että aineenopettajat eriyttävät opetustaan vähemmän kuin erityis- ja luokanopettajat. Saloviita (2018) ei kuitenkaan määrittele eriyttämistä lainkaan: opettajien tuli ainoastaan vastata kyllä tai ei koskien väitettä ”eriytän opetusta säännöllisesti”. Näin ollen hänen tutkimuksestaan ei käy selväksi, mitä opetuksen osa-aluetta eriyttäminen koskee. Erityisesti kieltenopetuksen kentällä eriyttäminen on ylipäättään varsin tutkimaton aihe (Roiha & Polso 2021a; Yavuz 2020). Sen sijaan opinnäytetöiden aiheena eriyttäminen on erittäin yleinen ja suosittu (esim. Härkönen 2020; Kontkanen 2020), mikä kuvastaa sitä, että tulevat opettajat näkevät sen tärkeänä asiana. Tässä tutkimuksessa koetamme osaltamme vastata edellä kuvattuun tutkimusaukkoon. Ensinnäkin keskitymme yksinomaan arvioinnin eriyttämiseen saadaksemme siitä mahdollisimman kattavan ja monipuolisen kuvan. Lisäksi tutkimme suomalaisia lukion vieraiden kielten opettajia, mikä lisää tutkimuksen uutuusarvoa.

3 Menetelmät

3.1 Aineistonkeruu ja osallistujat

Keräsimme tutkimuksen aineiston sekä avoimia että suljettuja kysymyksiä sisältäneen kyselyn ja haastatteluiden avulla. Sekä kyselyyn vastanneet että haastatellut opettajat saivat tietosuojaselosteen luettavakseen, jossa kuvattiin muun muassa se, miten aineistoa tullaan käsittelemään. Lisäksi haastatellut opettajat allekirjoittivat haastattelusuostumuksen.

Kyselyä luodessa otimme huomioon aikaisemman arviointi- ja eriyttämistutkimuksen sekä ajankohtaiset tekijät, kuten LOPS-uudistuksen sekä koronapandemian. Käytettyjä arviointimenetelmiä kartoitimme esittämällä opettajille listan erilaisista arviointimenetelmistä ja pyytämällä heitä vastaamaan 5-portaisen Likert-asteikon avulla, kuinka usein he näitä menetelmiä hyödynsivät (mukaillen Atjonen ym. 2019). Pilotoimme kyselyn kahdella lukion vieraan kielen opettajalla ja muokkasimme sen kommenttien pohjalta lopulliseen muotoonsa (ks. liite 1). Jaoin kyselyä sosiaalisessa mediassa erityisesti lukion vieraiden kielten opetukseen keskittyvissä ryhmissä

alkuvuodesta 2022. Ryhmissä oli yhteensä useita tuhansia jäseniä, joilla oli mahdollisuus nähdä kysely.

Kyselyyn vastasi 48 lukion vieraiden kielten opettajaa, joista 43 opetti englantia. Lisäksi vastaajien joukossa oli espanjan, ranskan, ruotsin ja saksan opettajia. Vastaajien opetuskokemus jakautui varsin tasaisesti muutaman vuoden ja yli 30 vuoden välille. Yleisin opetuskokemus oli 21–25 vuotta (n=10). 76 prosenttia vastaajista kertoi osallistuneensa arviointiin liittyvään täydennyskoulutukseen, kun taas ainoastaan 20 prosenttia ilmoitti saaneensa täydennyskoulutusta eriyttämisestä.

Haastattelimme neljää kyselyyn vastannutta opettajaa tarkemmin arvioinnin teemoista Zoomin välityksellä (ks. taulukko 1; nimet pseudonyymejä). Koska halusimme saada mahdollisimman monipuolisen kuvan tutkittavasta aiheesta, valitsimme neljä profiililtaan erilaista opettajaa niiden joukosta, jotka kyselyyn vastaamisen yhteydessä olivat ilmaisseet halukkuutensa osallistua haastatteluun. Haastateltujen opettajien profiilit eroavat toisistaan muun muassa opetuskokemuksen ja opetettavien kielten osalta.

TAULUKKO 1. Haastatellut opettajat.

	Maria	Johanna	Armi	Mirva
haastattelun kesto	43 minuuttia	41 minuuttia	49 minuuttia	51 minuuttia
opetettavat kielet	englanti, saksa	ruotsi	englanti	englanti, espanja
opetuskokemus	20 vuotta (11 vuotta lukiossa)	5 vuotta	25 vuotta	3 vuotta
lukion koko	750 opiskelijaa	300 opiskelijaa	450 opiskelijaa	1300 opiskelijaa

Suoritimme haastattelut kevään 2022 aikana. Yksi meistä tutkijoista haastatteli kahta opettajaa ja loput kaksi tutkijaa kumpikin yhtä opettajaa. Käytimme haastattelujen pohjana teemarunkoa, joka muodostettiin tutkimuskirjallisuuden ja kyselyvastausten pohjalta (ks. liite 2).

3.2 Analyysimenetelmät

Kyselyaineiston määrällisen osion tuloksista raportoimme tilastolliset tunnusluvut. Näin kuvaamme opettajien raportoimia käytettyjä arviointimenetelmiä yleisellä tasolla. Kyselyaineiston laadullisen osion analysoimme aineistolähtöisen sisällönanalyysin mukaisesti (Schreier 2012). Tässä analyysissä kysymyksen ”Millä tavalla huomioit

oppilaiden yksilöllisyyden arvioinnin eri vaiheissa?” vastaukset analysoimme siten, että jaoimme jokaisen opettajan vastauksen 1–3 analyysiyksikköön vastauksen laajuuden mukaisesti. Näistä analyysiyksiköistä koostimme kategorioita aineistolähtöisesti.

Haastatteluaineiston analysoimme teema-analyysin keinoin (Braun & Clarke 2021). Aluksi litteroimme haastattelut sanatarkasti, jonka jälkeen luimme ne ja teimme niistä alustavia havaintoja. Vertailimme havaintojamme ja etsimme yhtäläisyyksiä ja eroavaisuuksia haastattelujen välillä. Tämän jälkeen lähiluimme ja koodasimme haastattelut. Koodeista muodostimme alustavia teemoja (esim. välitön palaute, havainnointi, yksilökeskeisyys), joista yhdistelimme laajempia yläteemoja suhteessa tutkimuskysymyksiin (esim. alaspäin eriyttäminen, taitotasot, opiskelijatuntemus).

Koko analyysiprosessia taustoittivat tutkimuskysymykset sekä haastattelua varten laaditut teemat, jotka kytkeytyivät ajankohtaisessa tutkimuskirjallisuudessa esitettyihin käsityksiin arvioinnista ja sen eriyttämisestä sekä niihin liittyvistä käytännön keinoista. Näin ollen analyysia voi luonnehtia teoriaohjaavaksi sisällönanalyysiksi (Eskola 2018), sillä arviointi- ja eriyttämistieto sekä omat ennakkokäsityksemme ohjasivat analyysia.

4 Tutkimustulokset

Tässä luvussa raportoimme tutkimuksen tulokset kyselyvastausten ja haastattelujen pohjalta. Olemme jakaneet luvun kahteen alalukuun: 1) arviointi käytännössä ja 2) käsitykset arvioinnista ja sen eriyttämisestä. Tuloksia on havainnollistettu aineisto-oesimerkein, joita on poimittu sekä haastatteluaineistosta että kyselyn avoimista vastauksista.

4.1 Arviointi käytännössä

4.1.1 Vieraiden kielten opettajien yleiset arviointikäytänteet

Kyselyyn vastanneet opettajat hyödynsivät varsin laajasti erilaisia arvioinnin menetelmiä (ks. taulukko 2). Yleisimmin käytetyt menetelmät olivat kirjoitelman tai esseen arviointi, jatkuva arviointi, suljettu koe jakson päätteeksi, kuuntelukoe sekä sanakoe. Vaikka joukossa oli hyvin perinteisiä arviointimenetelmiä, kuten suljettu koe tai sanakoe, positiivista on muun muassa jatkuvan arvioinnin käyttö. Lisäksi esimerkiksi kirjoitelmien arviointiin sisältyy jo itsessään eriyttäviä elementtejä. Harvemmin käytettyjä menetelmiä olivat sen sijaan avoin koe jakson päätteeksi, portfolio, oppimiskeskustelu sekä projekti; näiden kaikkien ajatellaan olevan hyviä eriyttävän arvioinnin menetelmiä (esim. Roiha & Polso 2018; Tomlinson 2014).

TAULUKKO 2. Kyselyyn vastanneiden vieraiden kielten opettajien (n=48) käyttämät arviointimenetelmät niiden yleisyyden mukaan (Likert asteikolla 1–5).

Arviointimenetelmä	Keskiarvo	Keskihajonta	Moodi	Mediaani
Kirjoitelma/essee	4,75	0,53	5	5
Jatkuva arviointi	4,71	0,77	5	5
Suljettu koe jakson päätteeksi	4,67	0,60	5	5
Kuuntelukoe	4,48	0,85	5	5
Sanakoe	4,45	0,85	5	5
Itsearviointi	3,98	1,12	5	4
Suullinen arviointi	3,46	0,99	4	3,5
Vertaisarviointi	3,04	1,15	2	3
Esitelmä	3,00	0,77	3	3
Projekti	2,81	1,14	3	3
Oppimiskeskustelu	2,58	1,01	3	3
Portfolio	2,17	0,99	2	2
Avoin koe jakson päätteeksi	1,89	0,90	1	2

Neljä haastateltua opettajaa toi myös esiin paljon erilaisia arviointimenetelmiä. Lähtökohtaisesti he kaikki kertoivat painottavansa formatiivista arviointia ja kokevansa sen jossain määrin tärkeämmäksi arvioinnin muodoksi kuin summatiivisen arvioinnin. Tämä heijastui myös heidän käyttämiinsä arviointimenetelmiin, joissa korostui jatkuva palautteen anto sekä suullisesti että kirjallisesti. Tästä huolimatta jokainen kertoi kuitenkin pitävänsä opintojakson päätteeksi summatiivisen loppukokeen. Opettajista kaksi (Maria ja Johanna) piti myös välitestejä opintojaksojen aikana.

Kyselyyn vastanneet opettajat vaikuttivat käyttävän vertaisarviointia jonkin verran, joskin selvästi vähemmän kuin itsearviointia (ks. taulukko 2), joka koettiin hyvänä eriyttämisen välineenä. Kaikki haastatellut opettajat kertoivat myös käyttävänsä itsearviointia ja kokivat sen tärkeäksi ja positiiviseksi asiaksi. Sitä toteutettiin sekä suppeasti yksittäisin tehtävin että laajemmin loppukokeen yhteydessä. Haastateltujen opettajien suhtautuminen vertaisarviointiin sen sijaan vaihteli. Kielteisimpänä kaikista

vertaisarvioinnin koki Maria, joka kertoi, ettei juurikaan toteuta sitä. Hänen mukaansa opiskelijat ovat herkkiä vertaispalautteelle eivätkä mielellään anna esimerkiksi kirjoitelmiaan toisten luettavaksi. Samasta syystä Mirva oli toteuttanut vertaisarviointia nimettömänä. Armi puolestaan kertoi laativansa vertaisarviointikysymykset harkiten, jotta opiskelijat eivät antaisi liian negatiivista tai loukkaavaa palautetta toisilleen. Myös Johanna kertoi toteuttavansa vertaisarviointia jonkin verran ja erilaisia tapoja hyödyntäen: Toisinaan kaikki hänen opiskelijansa antoivat kirjallista palautetta toistensa tuotoksista, kuten esitelmistä. Toisinaan taas opiskelijat antoivat vertaispalautetta vain omalle työparilleen, joko suullisesti tai kirjallisesti. Jopa Maria, joka piti vertaisarviointia ongelmallisena, ohjasi joskus tiettyjä opiskelijoita kommentoimaan toistensa tekstejä. Näin hän käytännössä tuli soveltaneeksi eriyttävää vertaisarviointia.

Kolme kyselyyn vastannutta opettajaa mainitsi lukion opetussuunnitelman uusiin perusteisiin (vastedes LOPS 2019) liittyvän kieliprofilin formatiivisen arvioinnin välineenä. Haastateltujen opettajien mielipiteet kieliprofilista erosivat hieman toisistaan. Johanna näki kieliprofilin kaikkein positiivisimpana ja piti sitä hyvänä keinona tehdä opiskelijoiden oppimisprosessin näkyväksi. Hän toi myös esille, ettei kieliprofilin tekoon ole riittävästi aikaa kaiken muun ohella. Myös Armi piti kieliprofilia lähtökohtaisesti positiivisena asiana, mutta kustantajien kieliprofilille tekemiä mallipohjia hän kritisoi liian mekaanisiksi. Mariakin piti kieliprofilia hyvänä asiana, mutta ajatteli sen olevan jokseenkin ”päälleliimattua”, sillä hän katsoi toteuttavansa kieliprofilin tyyppistä työskentelyä jo entuudestaan. Hänen mielestään kieliprofilin oli kuitenkin hyvä eriyttämisen väline, sillä opiskelijat voivat tehdä siihen liittyviä tehtäviä omien kykyjensä ja innostuksensa mukaisesti. Kaikkein kielteisimmin kieliprofilin suhtautui Mirva. Omassa arjessaan hän oli kokenut sen olevan ”kaiken kiireen keskellä pakollinen paha”. Esimerkiksi ensimmäisessä opintojaksossa, jossa kieliprofilia työstettiin, opiskelijat eivät olleet saaneet tarvitsemiaan tietokoneita toimitusvaikeuksien vuoksi. Ajatuksen tasolla myös Mirva näki kuitenkin kieliprofilin positiivisena ja piti sitä muun muassa erinomaisena diagnostisen arvioinnin välineenä.

4.1.2 Vieraiden kielten opettajien eriyttävän arvioinnin käytänteet

Kyselyssä oli avoin kysymys opettajien eriyttävän arvioinnin käytänteistä. Moni kyselyyn vastannut opettaja kertoi eriyttävänsä arviointia varsin vähän, ja kuusi opettajaa jopa mainitsi, ettei eriytä arviointia lainkaan. Eräskin opettaja (OPE18) totesi: ”Tavoitteet ja arviointikriteerit ovat periaatteessa kaikille samat”. Vaikka kyselyyn vastanneet opettajat eriyttivät arviointiaan suhteellisen vähän ja rajoittuneesti, mainitsivat he kuitenkin joitain eriyttävän arvioinnin keinoja (ks. taulukko 3). Yleisimmin käytettyjä niistä olivat yksilöllisten tavoitteiden asettaminen sekä palautteen anto.

TAULUKKO 3. Avoimeen kysymykseen vastanneiden kieltenopettajien (n=48) käyttämät arvioinnin eriyttämismenetelmät.

Kategoria	N	Aineistoesimerkkejä
Eriyttäminen tavoitteita asettamalla ja seuraamalla	16	<p>”Jokainen opiskelija asettaa itse omat tavoitteensa. Opintojakson puolivälissä tsekataan, missä ollaan ja opintojakson lopussa itsearviointi - opettaja kommentoi näitä kaikkia.” (OPE37)</p> <p>”Kurssin alussa kysely tavoitteista, kurssin aikana itsearviointia eri kurssitehtävien päätteeksi ja lopuksi itsearviointi/reflektointi kurssin onnistumisesta oman taitotason mukaan.” (OPE13)</p>
Palaute eriyttämisen välineenä	11	<p>”Opiskelijat määrittelevät itselleen oman tavoitteen opintojakson/kurssin alussa. Annan palautetta sekä suullisesti (useimmiten kuitenkin koko ryhmälle yhteisesti, yksilöllisemmin harvemmin) että kirjallisesti (yksilöllisesti) jakson/kurssien kuluessa säännöllisesti (esim. arviointien suoritusten yhteydessä).” (OPE27)</p> <p>”Jatkuva vuorovaikutus opiskelijoiden kanssa onnistumisista ja mahdollisista epäonnistumisista. Jokaisen opiskelijan tulee kokea itsenä nähdä/kuulla jokaisen oppitunnin aikana.” (OPE38)</p>
Helpotetut arviointitehtävät	7	”Joissain tehtävissä he [opiskelijat] saavat valita vaativan tai helpomman vaihtoehdon (helposta esim. arvosana max 8, vaativassa enemmän työtä).” (OPE20)
Itsearviointi	4	”Itsearviointiin ja itsereflektioon kannustan aktiivisesti ja aina jakson päätteeksi se tehdään kirjallisesti. Opiskelija arvioi omaa oppimisprosessiaan ja oppimistuloksiaan suhteessa omaan tavoitteeseen ja kurssin tavoitteisiin. Itsearviointeihin teen omat yksilölliset kommenttini.” (OPE27)
Opiskelijoiden edistymisen jatkuva seuraaminen	4	”Koska oppikirja on sähköinen, näkee hyvin tekeekö opiskelija tehtäviä ja miten hän niissä onnistuu.” (OPE18)
Tukiopetus	2	”Heikommille oppilaille annan helpotettua lisämateriaalia ja tukiopetusta.” (OPE29)
Arviointiformaatin muutos	2	”Annan mahdollisuuden näyttää osaamistaan eri tavoilla (esim. jos esiintyminen ahdistaa, voi äänittää saman tai joskus tehdä kirjallisesti).” (OPE48)

Yhden kerran mainittuja menetelmiä:

- Arvioinnin pilkkominen pienempiin osiin.
- Käänteinen opetus.
- Opiskelijoiden osallistaminen arvioinnin suunnittelussa.
- Erityisopettaja osallistuu koetilanteeseen.
- Kokeessa saa kysyä kysymyksiä.

Myös haastatellut opettajat toivat esiin monia eriyttäviä arviointikäytänteitä. Maria muun muassa kertoi käyttävänsä toisinaan avointa koetta (esim. oppikirjan tuella tehtävä koe) uusintamahdollisuutena. Johanna puolestaan hyödynsi pari- tai ryhmäkokeita, kun taas Armi antoi joidenkin opiskelijoiden osoittaa osaamisensa suullisesti. Kaikki opettajat huomioivat eriyttämisen jossain määrin myös summatiivisissa kokeissa, muun muassa erilaisten tehtävätyyppien avulla.

4.1.3 Tavoitteiden ja kriteerien asettaminen ja niiden eriyttäminen

Kyselyyn vastanneista opettajista 46 kertoi, että uuden opintojakson alkaessa sen tavoitteet ja arviointikriteerit käytiin läpi yhdessä opiskelijoiden kanssa. 42 opettajista kertoi esittelevänsä tavoitteet jo opintojakson ensimmäisellä tunnilla. Vaikka 23 opettajaa mainitsi, että heillä tavoitteet ovat jatkuvasti näkyvillä kurssialustalla, vain neljä heistä kertoi palaavansa niihin opiskelijoiden kanssa opintojakson aikana. Viisi opettajaa puolestaan kertoi osallistavansa opiskelijat jollain tavoin tavoitteiden asetteluun. Noin kolmasosalla opettajista eriyttäminen kuitenkin ilmeni tavoitteiden asettelussa jollain tavoin (ks. taulukko 3).

Kaikki haastateltavat opettajat kertoivat käyvänsä opiskelijoiden kanssa ensimmäisillä kerroilla läpi opintojakson tavoitteet. Ainoastaan Maria kertoi palaavansa tavoitteisiin jatkuvasti opintojakson aikana, kun taas muut opettajat ottivat ne uudestaan esille vasta opintojakson lopussa. Armi ja Mirva eriyttivät niin, että opiskelijat laativat itselleen omia tavoitteita opintojakson alussa. Sen sijaan Johanna havahtui ajattelemaan tavoitteiden asettelun eriyttämistä vasta haastattelutilanteessa:

Niin mitenäs tuota huomioitko sä jotenkin eriyttämisen siinä tavoitteiden asettelussa.. läpikäynnissä tai niitten käsittelyssä? (haastattelija)

Aika vähän kyllä itseasiassa nyt kun mä oikein mietin asiaa. Eli kyllä se on ehkä sellainen niinku että se ne tavoitteet kuitenkin tietyllä tavalla tulee niinku opetussuunnitelmasta eli se niinku tietyllä tavalla niinku että se olisi sellainen niinku et, jos tähän tavoitteeseen pääsee, niin silloin on se niinku hyvä taso saavutettu. (Johanna)

4.2 Vieraiden kielten opettajien käsitykset arvioinnista ja sen eriyttämisestä

Tässä osiossa tuomme laajemmin esiin opettajien arviointikäytänteiden kirjoa ja paneudumme tarkemmin heidän käsityksiinsä arvioinnista ja sen eriyttämisestä. Temaattisen analyysin tuloksena konstruimme kysely- ja haastatteluaineiston pohjalta neljä teemaa:

1. eriyttävä arviointi dialogina, tukena ja ohjaamisena

2. laajemmat arvioinnin rakenteet
3. taitotasoihin keskittyvä eriyttäminen
4. opiskelijoiden ääni.

Esittelemme seuraavaksi nämä teemat ja nivomme ne yhteen tutkimuskirjallisuuden kanssa.

4.2.1 Eriyttävä arviointi dialogina, tukena ja ohjaamisena

Opettajat käsittivät eriyttävän arvioinnin ideaaliksi eräänlaisen ”inhimillisyyden”, sillä he kuvasivat eriyttävää arviointia usein ihmisten väliseksi, lämpimäksi vuorovaikutukseksi. Arviointi ilmeni tällöin mittaamisen ja testaamisen sijaan dialogina, tukena ja ohjaamisena. Tämä näkemys on linjassa tutkimuskirjallisuuden kanssa, sillä ideaalin eriyttämisen ajatellaan olevan sekä proaktiivista että reaktiivista toimintaa, jossa huomioidaan oppilaiden moninaisuus (Tomlinson 2014). Haastateltujen opettajien arviointi vaikutti olevan yhdistelmä näitä molempia. Formatiivinen arviointi oli enemmän reaktiivista, jolloin opetusta ja toimintaa muokattiin opiskelijoilta saadun palautteen perusteella. Summatiivisessa arvioinnissa opettajat puolestaan olivat jo lähtökohtaisesti koettaneet ottaa eri opiskelijat huomioon ja rakentaneet kokeet sillä ajatuksella, että jokainen pystyisi näyttämään niissä osaamisensa. Näin opettajat kytkivät niin formatiivisen kuin summatiivisen arvioinnin tarkoituksen osaksi opiskelijan ohjaamisen ideaalia.

Haastatellut opettajat painottivat opetuksessaan formatiivista arviointia ja korostivat sen merkitystä oppimisprosessin tukemisessa ja ohjaamisessa. Maria esimerkiksi kuvasi itseään metaforalla ”valmentaja”, joka on varsin yleinen viittaus opettajan rooliin oppimisen mahdollistajana ja fasilitaattorina (esim. Heinonen & Roiha 2022; Lin ym. 2012). Eriyttäminen yhdistettiin vahvasti formatiiviseen arviointiin ja opettajat kokivat, että yksilöllinen palaute jo itsessään eriyttää.

Haastateltujen opettajien näkemykset summatiivisesta arvioinnista vaihtelivat enemmän. Vaikka jokainen heistä pitikin summatiivisia kokeita opintojaksojen lopussa, Johanna ja Maria kokivat, että ideaalitulanteessa ne voisi jättää pois tai niitä voisi kehittää kommunikatiivisempaan suuntaan. Armi puolestaan koki summatiiviset kokeet olennaisena, mutta hänen mukaansa opiskelijat voisivat suorittaa sellaisen vain kerran lukuvuodessa. Mirva näki summatiiviset kokeet hyödyllisimpänä ja ne korostuivat eniten hänen opetuksessaan.

Eriyttävän arvioinnin haasteiksi opettajat nimesivät useita ”laitostuneen” lukion tunnusmerkkejä, jotka estivät vuorovaikutussuhteen syntymistä opettajan ja oppilaiden välille. Esille tuotiin muun muassa ajan rajallisuus sekä eriyttävien arviointikäytänteiden heikko tuntemus. Opettajien esiin tuomat ongelmat ovatkin varsin hyvin linjassa niiden haasteiden kanssa, joita aikaisemmassa tutkimuksessa on nimetty (esim. Lavania & Nor 2020; Roiha 2014; Tomlinson & Imbeau 2010; VanTassel-Baska & Stambaugh 2005; Wan

110 LUKION VIERAIDEN KIELTEN OPETTAJIEN KÄYTÄNTEITÄ JA KÄSITYKSIÄ ERIYTTÄVÄSTÄ ARVIOINNISTA

2017). Opettajien mukaan erityisesti ajan puute ja suuret opetusryhmät vaikeuttivat opiskelijan yksilöllistä kohtaamista ja arvioinnin eriyttämistä:

Enpä oikeastaan vielä ainakaan ehdi [toteuttaa eriyttämistä arvioinnissa]. Lukio-opettajuus on mennyt hyvin paljon työläämmäksi uuden opetussuunnitelman sekä varsinkin oppivelvollisuusmuutoksen myötä (muutosten paljous sekä jatkuva kirjaiminen). (OPE7)

[Ideaalitulanteessa] antaisin enemmän suullista palautetta ja kävisin oppimiskeskusteluja, tälle ei tosiaan ole aikaa. Pystyisin esittämään maamerkkisuorituksia, joihin opiskelijat voisivat verrata omia tuotoksiaan. Lisäksi oppilaat pystyivät jossain määrin valitsemaan osuvammin itselleen sopivan vaikeustason tehtäviä. (OPE4)

Kolme kyselyyn vastannutta opettajaa mainitsi erikseen pienen lukion mahdollistavan opiskelijan opintohistorian perinpohjaisen tuntemisen, mikä edistää eriyttämistä:

Ottamalla huomioon tiedot ko. opiskelijan aiemmasta suoriutumisesta (pienessä koulussa tunnemme opiskelijat melko perinpohjaisesti). (OPE9)

Haastateltavista opettajista Maria, Johanna ja Mirva korostivat opiskelijatuntemusta ja ennakoivan arvioinnin merkitystä omassa arviointitoiminnassaan ja sen eriyttämässä. He toteuttivat ennakoivaa arviointia kuitenkin hieman eri tavoin. Maria kirjoitutti jokaisen opintojakson alussa opiskelijoilla englanninkielisen kirjeen, jossa opiskelijat kertoivat itsestään. Maria kertoi saavansa tällä tavoin tietoa opiskelijoiden mieltymyksistä sekä myös jonkinlaisen käsityksen heidän kielitaidostaan. Johanna puolestaan hyödynsi perinteisempää lähtötasokoetta, mutta ainoastaan lukion alussa. Mirvan lukiossa tehtiin englannissa ja ruotsissa laajamittaiset lähtötason kartoitukset, joiden pohjalta erityisopettajat informoivat kielten opettajia tuen tarpeessa olevista opiskelijoista. Tämän lisäksi Mirva teetätti opiskelijoillaan oman kyselyn, jossa hän kartoitti opiskelijoiden omaa käsitystä kielenosaamisestaan. Mirva näki, että kieliprofili ja siihen liittyvät tehtävät toimivat myös lähtötason kartoituksena. Kaikki kolme opettajaa ottivat ennakoivan arvioinnin huomioon opetuksessaan ainakin jossain määrin. Marian mukaan lähtötason kartoitus paljastaa hänelle heti, ketkä opiskelijat tarvitsevat eniten tukea. Johanna puolestaan sanoi kertaavansa erityisesti sellaisia asioita, jotka ovat osoittautuneet opiskelijoille haasteelliseksi. Hänen mukaansa asiat kuitenkin kerrataan yhteisesti koko luokan kanssa, vaikka ne olisivatkin jo osalla täysin hallussa.

4.2.2 Laajemmat arvioinnin rakenteet eriyttämistä kehystämässä

Sekä kyselyssä että haastatteluissa kysimme opettajilta, kuinka laajemmat suomalaiset arvioinnin kehykset vaikuttavat opettajien eriyttämiskäsitykseen. Ensinnäkin uusi LOPS (2019) ja sen vaikutukset arviointiin jakoi opettajien mielipiteitä. Hieman yli puolet

kyselyyn vastanneista opettajista koki, ettei arviointi ollut juurikaan muuttunut uuden LOPSin myötä. Osa puolestaan näki, että siinä on annettu enemmän painoa formaatiiviselle arvioinnille ja lisätty itse- ja vertaisarviointia. Haastatellut opettajat kokivat uuden LOPSin pääosin positiivisena asiana arvioinnin näkökulmasta ja toivat esiin muun muassa siinä korostuvaa jatkuvan arvioinnin merkitystä. Armin mukaan aiemmin arviointi oli ollut summatiivisempää ja numeroita painottavaa, kun taas uudessa LOPSissa arviointi oli opiskelijälähtöisempää ja otti opiskelijat paremmin huomioon.

Ylioppilastutkinnon kielikokeet olivat toinen instituutio, joka varjosti luokkahuoneissa tapahtuvaa arvioinnin eriyttämistyötä. Eräs kyselyyn vastannut opettaja näki lukion opetuksen päätehtäväksi taata hyvät arvosanat ylioppilaskokeissa:

Hassuahan on oikeastaan se, että lukion päättötodistuksella ei ole mitään merkitystä, joten mitä varten annamme arvosanoja... jotta opiskelijalla olisi jonkinlainen käsitys, miten paljon töitä pitäisi vielä tehdä yo:n eteen? Katson, että lukion opettajana päätehtäväni on saada opiskelijoilleni mahdollisimman hyvät yo-arvosanat ja näin ollen varsinainen arviointi ei ole keskiössä, vaan se, että he oppivat. (OPE28)

Ylioppilaskokeiden vaikutus arviointityöskentelyyn nousi esiin mielenkiintoisesti, ja paikoin ristiriitaisesti, myös kaikissa haastatteluissa. Yhtäältä niiden roolia vähäteltiin lukion opintojaksojen sisältöjen määrittäjänä, toisaalta ne vaikuttivat kuitenkin jossain määrin ohjaavan opettajien opetus- ja arviointikäytänteitä (ns. *washback*-ilmiö, ks. esim. Kuang 2020). Kaikki opettajat olivat periaatteessa sitä mieltä, ettei ylioppilaskokeiden tulisi määrittää lukion opintojaksojen opetusta. Kuten Johanna ja Armi perustelivat, kaikki opiskelijat eivät edes suorita kielten ylioppilaskokeita. Myöhemmin haastattelussaan Johanna kuitenkin mainitsi, että ylioppilaskokeet rajoittavat hänen toteuttamaansa opetusta ja arviointia. Lisäksi hän pohti, onko opiskelijoille haitallista, jos kokeiden kuunteluosuuksia tauotetaan tai kokeen osia jaetaan eri päiville, sillä ylioppilaskokeissa ei näin toimita. Johanna kävi siis tietynlaista sisäistä kamppailua: ajatuksen tasolla hän vastusti ylioppilaskokeiden vaikutusta, mutta muokkasi käytännössä toimintaansa niiden perusteella.

Maria näki ylioppilaskokeet ja lukion oppimäärän erillisinä asioina. Tosin hän kuvasi ylioppilaskokeita metaforalla ”olympialaiset”, joihin mennään näyttämään lukion aikana opittu, mikä antaa ymmärtää, että ne ovat kuitenkin jossain määrin hänenkin opetuksensa taustalla. Mirva ei ollut muodostanut selkeää mielipidettä ylioppilastutkinnon roolista arvioinnissa, sillä hän kuvasi ylioppilaskokeiden ja lukion opintojaksojen arvioinnin välistä jännitettä seuraavasti:

Ehkä semmoinen on vielä hakusessa, että mitenkä niitä painotetaan tavallaan summatiivista ja formaatiivista arviointia tavallaan, koska yo kokeet on se iso kantava ohjaava voima lukiossa, että että voidaanko me jotenkin sitä formaatiivista arviointia ottaa mukaan siihen summatiiviseen opintojakson arviointiin ettei me tehdä itsel-

112 LUKION VIERAIDEN KIELTEN OPETTAJIEN KÄYTÄNTEITÄ JA KÄSITYKSIÄ ERIYTTÄVÄSTÄ ARVIOINNISTA

lemme hallaa siinä että meidän koulun yo tulokset olisi täysin ristiriidassa meidän niin kuin opintojaksojen arvioinnin kanssa. (Mirva)

Sitaatti kuvastaa osuvasti, kuinka voimakas vaikuttaja ylioppilaskokeet ovat lukiossa. Vaikka suurin osa opiskelijoista ei tunnista ylioppilaskokeiden vaikutusta opetuksessa, aiheuttaa se heille paljon stressiä ja ahdistusta lukion aikana (Pollari 2016). Huomionarvoista on, että ylioppilastutkinnoissa arvioidaan vain osaa LOPSin (2019) tavoitteista, kun taas lukion opintojaksojen arviointien tulisi kattaa ne kaikki.

4.2.3 Ylös- vai alaspäin? Taitotasoihin keskittyvä eriyttäminen

Haastateltujen opettajien näkemyksissä eriyttävästä arvioinnista yhdistyi sekä laajoja että suppeita eriyttämiskäsityksiä. Yhtäältä he kaikki toivat esiin monipuolisia arviointitapoja ja erilaisuuden huomioimista arviointi- ja opetuskäytänteissään. Toisaalta heidän puheissaan ilmeni tietynlaisia rajoittuneita ajatuksia eriyttävään arviointiin liittyen, kuten polarisoivaa puhetapaa eriyttämisen ”suunnista”. Tässä puhettavassa oppilaat asetettiin tiukasti akselille ”heikko – lahjakas” heidän jokseenkin yksilölliseen ”tasonsa” perusteella. Opettajat myös keskittyivät eriyttävissä arviointikäytänteissään pääsääntöisesti alaspäin eriyttämiseen (ks. myös Nepal ym. 2021). Kaikki haastatellut opettajat kokivatkin ylöspäin eriyttämisen haasteellisena ja kaipasivat siihen lisää vinkkejä:

Summatiivisessa arvioinnissa niin sitä eriyttämistä tietysti tulee sitten myös tehtyä silloin kun on niinku jotain niinku haasteita, varsinkin.. et ehkä vähemmän ehkä tulee silloin niinku ylöspäin eriytettyä sitten siinä niinku arvioinnissa. (Johanna)

Sitten ne ketkä niinku silmät loistaen siellä sitä kymppiä jahtaa, niin he jää sitten kuitenkin aina vääjäämättä oman onnen nojaan niin niin se että jotenkin saisi sinne niinku työkaluja siihen, että miten niitä opiskelijoita voisi vielä enemmän.. (Armi)

Vahva keskittyminen alaspäin eriyttämiseen tuli Johannan toiminnassa esiin myös siinä, että hän kertoi kohdistavansa kertaavan opetuksensa koko luokalle, mikäli hän huomasi joidenkin opiskelijoiden hyötyvän tietyn asian lisäopetuksesta.

Kyselyyn vastanneet opettajat puolestaan viittasivat vastauksissaan opiskelijoiden ”massaan” ja ”ainekseen” ikään kuin homogeenisenä joukkona, johon on kohdistunut samansuuntaisia yhteiskunnallisia muutoksia: ”tasot” eivät koskeneet opettajien puheissa ainoastaan yksilöitä vaan ennen kaikkea ikäluokkia, kenties jopa sukupolvia. ”Oppilasaineksen heikentyminen” nähtiin esimerkkinä eriyttämisen esteestä:

[Ideaalisti käyttäisin] laajempia kurssitöitä, enemmän ryhmätöitä, videotehtäviä, vertaisarviointia. Ylipäättään monipuolisempia arviointimethodoja. Opiskelija-aines on niin heikkoa, että on vaikeaa saada edes kirjoitelmia kirjoitettua ja palautettua.

Ryhmätöitä on kokeiltu, mutta niistä ei tule mitään, koska osa porukasta ei tee mitään, jolloin muut kärsivät. Siksi perinteiset kokeet toimivat parhaiten. (OPE13)

Mielestäni arviointini on ollut hyvin linjassa opiskelijoiden tason kanssa. Mutta nyt alkaa tulla enemmän ja enemmän opiskelijoita, jotka eivät opiskele. Arvosanat huonoja ja sanoja ei lueta. (OPE17)

Kyselyyn vastanneista opettajista monet lähestyivätkin arvioinnin eriyttämistä lähes yksinomaan taitotasojen näkökulmasta (ks. esim. Roy ym. 2013; Saloviita 2018) eivätkä maininneet eriyttävänsä arviointia esimerkiksi opiskelijoiden työskentelytyylien tai kiinnostuksen kohteiden perusteella.

4.2.4 Opiskelijoiden oman äänen puuttuminen eriyttävässä arvioinnissa

Opettajat eivät juurikaan tuoneet esiin opiskelijoiden omaa näkökulmaa eriyttämisessä. Vaikka esimerkiksi Mirva kertoi antavansa erilaisia kotitehtäviä eri opiskelijoille, teki hän päätöksen itse sen sijaan, että opiskelijat olisivat valinneet haluamansa tasoisia tehtäviä. Opiskelijälähtöisen arvioinnin periaatteisiin kuuluu opiskelijan näkeminen aktiivisena toimijana, joka tulee osallistaa oppimisprosessiin ja siihen liittyviin päätöksiin mahdollisimman kiinteästi (Tomlinson 2014).

Mielenkiintoisena havaintona nousee esiin, että opettajien puheessa ennakoivalla arvioinnilla ymmärretään tavallisesti lähtötasotestiä. Opettajat tarkastelivat ennakoivaa arviointia lähinnä omasta näkökulmastaan eivätkä tuoneet esille sen hyötyjä opiskelijoille itselleen, kuten alla oleva sitaatti kertoo:

Eihän sillä niinku niillä tuloksella ei ole niinku heille [opiskelijoille] mitään merkitystä, mutta et et mulle tietysti antaa sen että tietää vähän niinku mistä lähtee että tuota että että tarvitaanko niinku enemmän kertausta vai vai onko se niinku mitkä asiat on ehkä sitten ihan niinku selviä. (Johanna)

Armi oli haastatelluista ainoa, joka suhtautui ennakoivaan arviointiin negatiivisesti ja totesi, ettei ikinä pidä lähtötason testiä, sillä se voi antaa hänelle vääristyneen kuvan opiskelijan todellisesta osaamisesta. Alla oleva sitaatti kuvastaa Armin tuntemuksia asiasta:

Minä pelkään näissä jos mulle tulee viisikymmentäkaksi uutta ykköstä ja minä pidän niille jonkun semmoisen semmoisen minun mielestä loistavan testin ja joku saa siitä kuus plus ja joku saa siitä ysin niin minä pelkään sitä että mulle tulee siitä semmoinen joku mielikuva siitä. (Armi)

Armin sitaatista voi päätellä, että hänkin tarkastelee ennakoivaa arviointia lähinnä omasta näkökulmastaan eikä pohdi sitä, miten ennakoiva arviointi vaikuttaa opiskeli-

jan käsitykseen itsestään kielenoppijana. Lisäksi Armin vastauksesta saa vaikutelman, että lähtötason arvioinnin tulisi hänen mukaansa olla kertaluontoista eikä jatkuvaa. Mielenkiintoisesti hän kuitenkin käyttää arvioinnissaan perinteisiä summatiivisia kokeita, joiden voidaan ajatella mittaavan arviointia kertaluonteisesti tietyssä hetkessä. Armi myös yhdistää lähtötason kartoituksen yksinomaan testeihin. Toisaalta on huomattava, että vaikka Armi ilmaiseekin painokkaasti, ettei milloinkaan kartoita lähtötasoa, on hänen haastattelussaan tähän liittyen ristiriita. Hän nimittäin kertoo, että opiskelijat tekevät hänelle opintojakson ensimmäisen viikon aikana 2 minuutin videon ja kirjoittavat 100 sanan viestin, joista hän antaa palautetta. Näin ollen Armi vaikuttaa myös tekevän lähtötason kartoitusta, vaikka ei sitä itse sellaiseksi mielläkään.

5 Pohdinta ja loppuksi

Tässä tutkimuksessa paneuduttiin lukion vieraiden kielten opettajien arviointiosaamiseen arviointikäytänteiden ja -käsitysten kautta, ensisijaisesti eriyttämisen näkökulmasta. Moni opettaja koki LOPSin (2019) tuoneen formatiivisen arvioinnin keskeisemmäksi osaksi arviointikäytänteitä, mitä voidaan pitää myönteisenä muutoksena aiempaan (vrt. Atjonen ym. 2019). Vaikuttaa myös siltä, että kieliprofiili on joidenkin opettajien kohdalla löytänyt jo paikkansa formatiivisen ja eriyttävän arvioinnin välineenä.

Useista eriyttämiskeinoista huolimatta opettajien kokonaiskäsitys eriyttävästä arvioinnista vaikutti vielä rajalliselta ja suppealta (ks. myös Melesse 2015; Milinga ym. 2022; Santangelo & Tomlinson 2012). Arvioinnin eriyttämistä tarkasteltiin lähes yksinomaan taitotasojen perusteella, eikä laajan eriyttämisenäkemyksen mukaisia kiinnostuksen kohteita tai työskentelymieltymyksiä tuotu esiin (ks. esim. Tomlinson 2014). Ylipäätään eriyttämistä toteutettiin vielä varsin pienimuotoisesti, ja joidenkin vastaajien kohdalla sitä ei tehty lainkaan. Usealla opettajalla oli myös käytössä samat keinot kaikille opiskelijoille. Yksittäisten keinojen sijaan eriyttävän arvioinnin olisi hyvä lähteä liikkeelle yksilöllisistä tavoitteista, ulottua oppimisprosessin ohjaamiseen ja kulminoitua summatiiviseen arviointiin (Tomlinson & Moon 2013). Toisaalta on hyvä huomata, että osa tutkimukseen osallistuneista opettajista käytti eriyttävän arvioinnin keinoja monipuolisesti.

Opettajat keskittyivät vahvasti alaspäin eriyttämiseen. Ylöspäin eriyttäminen sen sijaan koettiin haasteellisempänä, mikä on linjassa aikaisempien tutkimusten kanssa (esim. Brevik ym. 2018). Eriyttämisen painottaminen alaspäin onkin suomalaisessa pedagogiikassa tyypillistä: opettajat kokevat usein tärkeämpänä tukea opiskelijoita, joilla on oppimisvaikeuksia, kuin huomioida lahjakkaita (Roiha 2014, 2022; Seppälä & Kautto-Knape 2009). Suomalaista koulujärjestelmää on ylipäätään leimannut vahva panostaminen oppimisvaikeuksien tunnistamiseen ja oppimisen tukemiseen (esim. Kivirauma & Ruoho 2007), joskin enemmän perusopetuksen puolella. Kuitenkin laajan käsityksen mukaan eriyttäminen koskee kaikkia opiskelijoita (esim. Tomlinson 2014).

Valtaosa tutkimukseemme osallistuneista opettajista oli saanut täydennyskoulutusta arviointiin, mutta eriyttämiseen liittyvä täydennyskoulutus oli vielä varsin vähäistä. Kuten aikaisempi tutkimus (esim. Wan 2017) on jo osoittanut, mitä laajempi käsitys opettajilla on eriyttämisestä, sitä enemmän he sitä toteuttavat toiminnassaan. Myös oman tutkimuksemme perusteella työssä olevat opettajat hyötyisivät eriyttämiseen liittyvästä täydennyskoulutuksesta, jossa olisi hyvä käsitellä yksityiskohtaisemmin myös lahjakkaiden opiskelijoiden tukemista. Lisäksi opettajaopiskelijoita palvelisi eriyttämisen laajempi käsittely jo perustutkinto-opintojen aikana. Näin eriyttämisestä tulisi helpommin luonnollinen ja kiinteä osa kaikkea opetusta – myös arviointia – eikä sitä nähtäisi erillisenä asiana.

Arvioinnin eriyttäminen koettiin paikoin haasteellisena ja työläänä, ja yhtenä suurimpana syynä tähän mainittiin suuret luokkakoot. Ideaalitulanteessa opettajat olisivat pitäneet enemmän henkilökohtaisia oppimiskeskusteluja opiskelijoiden kanssa. Opettajat näyttävätkin pitävän hyvää opiskelijatuntemusta eriyttämisen perusedellytyksenä (ks. myös Tomlinson 2014). Opiskelijatuntemus on lähtökohtana myös esim. Thousandin ym. (2007) eriyttämismallissa, ja on ymmärrettävää, että pienempien ryhmäkokojen katsotaan mahdollistavan myös syvällisempi opiskelijatuntemus. Ihanteellinen ratkaisu olisikin rajoittaa lukion opetusryhmien kokoja. Suurissa ryhmissä eriyttämistä voi kuitenkin tehdä resurssien sallimissa rajoissa.

Vaikka opetussuunnitelma määrittää kielitaidon tavoitteelliset tasot vieraiden kielten eri oppimäärissä, (arvioinnin) eriyttämisen näkökulmasta yksilöllisempi lähestymistapa oppimistavoitteisiin voisi olla hedelmällisempi. Kun oppilaat ovat yksilöitä erilaisine vahvuuksineen, heikkouksineen ja tarpeineen, taitotasoissakaan yksi koko ei sopine kaikille. Arviointia eriyttämällä oppilaan olisi helpompi havaita omia kehityskohteitaan ja myös oppia asettamaan omia tavoitteitaan. Opettajien mainitsemat ylioppilaskokeet toki määrittävät tavoitteita: kokeethan laaditaan LOPSin tavoitetasoon sovittaen niin, että puolet tehtävistä on tavoitetasolla, neljännes sen ylä- ja neljännes alapuolella. Kaikkien ei kuitenkaan odoteta saavuttavan korkeinta mahdollista arvosanaa, vaan itselleen paras arvosana. Jokaisella pitää olla samat mahdollisuudet, mutta ei ole realistista odottaa, että kaikki saavuttavat saman tason samaan aikaan.

Tutkimuksemme tarjoaa alustavaa tietoa lukion vieraiden kielten opettajien eriyttävästä arviointitoiminnasta. Jatkossa olisi hyödyllistä tutkia eriyttävää arviointia laajemmin, myös opiskelijoiden omasta näkökulmasta. Lisäksi olisi tärkeää tutkia eriyttävän arvioinnin vaikutusta opiskelijoiden kielitaidon kehitykseen sekä affektiivisiin tekijöihin kuten motivaatioon ja kieliminään. Erityisen tärkeää olisi myös vahvistaa eriyttämisen näkökulmaa opettajien arviointiosaamista käsittelevässä kirjallisuudessa.

Kirjallisuus

- Atjonen, P., H. Laivamaa, A. Levonen, S. Orell, M. Saari, K. Sulonen, M. Tamm, P. Kamppi, N. Rumpu, R. Hietala & J. Immonen 2019. *"Että tietää missä on menossa": Oppimisen ja osaamisen arviointi perusopetuksessa ja lukiokoulutuksessa*. Helsinki: Kansallinen Koulutuksen Arviointikeskus.
- Benjamin, A. 2002. *Differentiated instruction. A guide for middle and high school teachers*. Larchmont, NY: Eye on Education.
- Braun, V. & V. Clarke 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18 (3), 328–352. <https://doi.org/10.1080/14780887.2020.1769238>
- Brevik, L. M., A. E. Gunnulfsen & J. S. Renzulli 2018. Student teachers' practice and experience with differentiated instruction for students with higher learning potential. *Teaching and Teacher Education*, 71, 34–45. <https://doi.org/10.1016/j.tate.2017.12.003>
- Eskola, J. 2018. Laadullisen tutkimuksen juhannustaiat: Laadullisen aineiston analyysi vaihe vaiheelta. Teoksessa R. Valli (toim.) *Ikkunoita tutkimusmetodeihin 2. Näkökulmia aloittelevalle tutkijalle tutkimuksen teoreettisiin lähtökohtiin ja analyysimenetelmiin*. Jyväskylä: PS-kustannus, 209–231.
- Euroopan neuvosto. 2001. *Common European framework of reference for languages: learning, teaching, assessment*. Strasbourg: Euroopan neuvosto. <https://rm.coe.int/1680459f97>
- Gaitas, S., C. Carêto, F. Peixoto & J. Castro Silva 2022. Differentiated instruction: 'to be, or not to be, that is the question'. *International Journal of Inclusive Education*. <https://doi.org/10.1080/13603116.2022.2119290>
- Gaitas, S. & M. A. Martins 2017. Teacher perceived difficulty in implementing differentiated instructional strategies in primary school. *International Journal of Inclusive Education*, 21 (5), 544–556.
- Graham, L., K. de Bruin, C. Lassig & I. Spandagou 2021. A scoping review of 20 years of research on differentiation: investigating conceptualisation, characteristics, and methods used. *Review of Education*, 9 (1), 161–198. <https://doi.org/10.1002/rev3.3238>
- Gregersen, T. & S. Mercer (toim.) 2021. *The Routledge handbook of the psychology of language learning and teaching*. New York: Routledge.
- Hall, T. 2002. *Differentiated instruction. Effective classroom practices report*. National Center on Accessing the General Curriculum, Office of Special Education Programs, U.S. Department of Education.
- Heinonen, P. & A. Roiha 2022. Kielten aineenopettajaopiskelijoiden käsityksiä oppimisesta ja opettajuudesta - aineistona metaforat ja visuaaliset narratiivit. Teoksessa S. Routarinne, P. Heinonen, T. Kärki, A. Roiha, M.-L. Rönkkö & A. Korkeaniemi (toim.) *Ainedidaktiikka ajassa: laajenevat oppimisympäristöt ja eri-ikäiset oppijat. Ainedidaktisia tutkimuksia 22*. Helsinki: Suomen ainedidaktinen tutkimusseura, 201–233.
- Hildén, R. & M. Härmälä 2015. *Hyvästä paremmaksi - kehittämisideoita kielten oppimistulosten arviointien osoittamiin haasteisiin*. Helsinki: Kansallinen Koulutuksen Arviointikeskus.
- Härkönen, V. 2020. *Yläkoulun opettajien kokemuksia inklusiosta ja eriyttämisestä äidinkielen ja kirjallisuuden oppitunneilla*. Pro gradu -tutkielma. Itä-Suomen yliopisto.
- Ilola, M. 2018. *Oppilaiden käsityksiä englannin suullisesta kielitaidosta, sen oppimisesta ja arvioinnista*. Jyväskylä: Jyväskylän yliopisto. <https://jyx.jyu.fi/handle/123456789/59522>
- Jager, L., E. Denessen, A. Cillessen & P. C. Meijer 2022. Capturing instructional differentiation in educational research: investigating opportunities and challenges. *Educational Research*, 64 (2), 224–241. <https://doi.org/10.1080/00131881.2022.2063751>

- Kivirauma, J. & K. Ruoho 2007. Excellence through special education? Lessons from the Finnish school reform. *International Review of Education*, 53 (3), 283–302. <https://doi.org/10.1007/s11159-007-9044-1>
- Kontkanen, K. 2020. *Eriyttäminen fysiikan opetuksessa – tapaustutkimus erään yläkoulun opettajien näkemyksistä ja toiminnasta*. Pro gradu -tutkielma. Jyväskylän yliopisto.
- Kuang, Q. 2020. A review of the washback of English language tests on classroom teaching. *English Language Teaching*, 13 (9), 10–17. <https://doi.org/10.5539/elt.v13n9p10>
- Laari, A., S. Lakkala & S. Uusiautti 2021. 'For the whole grade's common good and based on the student's own current situation': differentiated teaching and the choice of methods among Finnish teachers. *Early Child Development and Care*, 191 (4), 598–611. <https://doi.org/10.1080/03004430.2019.1633314>
- Lam, R. 2015. Language assessment training in Hong Kong: implications for language assessment literacy. *Language Testing*, 32 (2), 169–197. <https://doi.org/10.1177/0265532214554321>
- Lavania, M. & F. B. M. Nor 2020. Barriers in differentiated instruction: a systematic review of the literature. *Journal of Critical Reviews*, 7 (6), 293–297. <http://dx.doi.org/10.31838/jcr.07.06.51>
- Lin, W.-C., P. P. Shein & S. C. Yang 2012. Exploring personal EFL teaching metaphors in pre-service teacher education. *English Teaching: Practice and Critique*, 11 (1), 183–199. <https://files.eric.ed.gov/fulltext/EJ970260.pdf>
- LOPS. 2019. *Lukion opetussuunnitelman perusteet*. Helsinki: Opetushallitus. https://www.oph.fi/sites/default/files/documents/lukion_opetussuunnitelman_perusteet_2019.pdf
- Melesse, T. 2015. Differentiated instruction: perceptions, practices and challenges of primary school teachers. *Science, Technology and Arts Research Journal*, 4 (3), 253–264.
- Milinga, J. R., J. Amani & S. E. Lyakurwa 2022. Teachers' perceptions of differentiated instruction for academically high-achieving secondary school students in Tanzania. *Journal of Advanced Academics*. <https://doi.org/10.1177/1932202X221129970>
- Mäkipää, T. & R. Hildén 2021. Kannustava palaute ja opettajan arviointityö lukion kielenopetuksessa. *Kieli, koulutus ja yhteiskunta*, 12 (6). <https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-joulukuu-2021/kannustava-palaute-ja-opettajan-arviointityo-lukion-kieltenopetuksessa>
- Nepal, S., S. Walker & J. Dillon-Wallace 2021. How do Australian pre-service teachers understand differentiated instruction and associated concepts of inclusion and diversity? *International Journal of Inclusive Education*. <https://doi.org/10.1080/13603116.2021.1916111>
- Pollari, P. 2016. Daunting, reliable, important or "trivial nitpicking?" Upper secondary students' expectations and experiences of the English test in the Matriculation Examination. Teoksessa A. Huhta & R. Hildén (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFinLA-e. Soveltavan kielitieteen tutkimuksia 9. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 184–211. <https://journal.fi/afinla/article/view/60854>
- Pollari, P. 2017. *(Dis)empowering assessment? Assessment as experienced by students in their upper secondary school EFL studies*. Jyväskylä Studies in Humanities 329. Jyväskylä: Jyväskylän yliopisto. <https://jyx.jyu.fi/handle/123456789/55478>
- Popham, W. J. 2009. Assessment literacy for teachers: faddish or fundamental? *Theory into practice*, 48 (1), 4–11. <https://doi.org/10.1080/00405840802577536>
- POPS. 2014. *Perusopetuksen opetussuunnitelman perusteet*. Helsinki: Opetushallitus. https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf

- Pozas, M., V. Letzel & C. Schneider 2020. Teachers and differentiated instruction: exploring differentiation practices to address student diversity. *Journal of Research in Special Educational Needs*, 20 (3), 217–230.
- Prast, E. J., E. Van de Weijer-Bergsma, E. H. Kroesbergen & J. E. H. Van Luit 2015. Readiness-based differentiation in primary school mathematics: expert recommendations and teacher self-assessment. *Frontline Learning Research*, 3 (2), 90–116. <https://doi.org/10.14786/flr.v3i2.163>
- Purpura, J. 2016. Second and foreign language assessment. *Modern Language Journal*, 100 (Supplement 2016), 190–208. <https://doi.org/10.1111/modl.12308>
- Rock, M. L., M. Gregg, E. Ellis & R. A. Gable 2008. REACH: a framework for differentiating classroom instruction. *Preventing School Failure: Alternative Education for Children and Youth*, 52 (2), 31–47. <https://doi.org/10.3200/PSFL.52.2.31-47>
- Roiha, A. 2014. Teachers' views on differentiation in content and language integrated learning (CLIL): perceptions, practices and challenges. *Language and Education*, 28 (1), 1–18. <https://doi.org/10.1080/09500782.2012.748061>
- Roiha, A. 2022. Broadening pre-service English language teachers' perceptions of differentiation relying on the 5D model. *The Teacher Educator*. <https://doi.org/10.1080/08878730.2022.2126054>
- Roiha, A. & J. Polso 2018. *Onnistu eriyttämisessä: Toimivan opetuksen opas*. Jyväskylä: PS-kustannus.
- Roiha, A. & J. Polso 2020. Eriyttämisen tarvitaan laajempaa näkökulmaa. *NMI Bulletin -Oppimisen ja oppimisvaikeuksien erityislehti*, 30 (4). https://bulletin.nmi.fi/wp-content/uploads/2021/01/eriyttaminen_laajempaa_nakokulmaa.pdf
- Roiha, A. & J. Polso 2021a. The 5-dimensional model: a Finnish approach to differentiation. Teoksessa D. Luis Banegas, G. Beacon & M. Pérez Berbain (toim.) *International perspectives on diversity in English language teaching*. New York: Palgrave Macmillan, 211–227.
- Roiha, A. & J. Polso 2021b. The 5-dimensional model: a tangible framework for differentiation. *Practical Assessment, Research & Evaluation*, 26 (20). <https://scholarworks.umass.edu/pare/vol26/iss1/20/>
- Roy, A., F. Guay & P. Valois 2013. Teaching to address diverse learning needs: development and validation of a Differentiated Instruction Scale. *International Journal of Inclusive Education*, 17 (11), 1186–1204. <https://doi.org/10.1080/13603116.2012.743604>
- Saloviita, T. 2018. How common are inclusive educational practices among Finnish teachers? *International Journal of Inclusive Education*, 22 (5), 560–575. <https://doi.org/10.1080/13603116.2017.1390001>
- Santangelo, T. & C. A. Tomlinson 2012. Teacher educators' perceptions and use of differentiated instruction practices: an exploratory investigation. *Action in Teacher Education*, 34 (4), 309–327.
- Schreier, M. 2012. *Qualitative content analysis in practice*. London: SAGE.
- Seppälä, R. & E. Kautto-Knape 2009. Eriyttämisen tavat englannin opetuksessa. *Kielikukka*, 4, 13–18.
- Shareefa, M., V. Moosa, R. Mat Zin, N. Z. M. Abdullah & R. Jawawi 2019. Teachers' perceptions on differentiated instruction: do experience, qualification and challenges matter? *International Journal of Learning, Teaching and Educational Research*, 18 (8), 214–226.
- Smit, R. & W. Humpert 2012. Differentiated instruction in small schools. *Teaching and Teacher Education*, 28, 1152–1162. <https://doi.org/10.1016/j.tate.2012.07.003>
- Stiggins, R. J. 1995. Assessment literacy for the 21st century. *Phi Delta Kappan*, 77 (3), 238–245. <https://www.proquest.com/scholarly-journals/assessment-literacy-21st-century/docview/218532914/se-2?accountid=14774>

- Thousand, J. S., R. A. Villa & A. I. Nevin 2007. *Differentiating instruction: Collaborative planning and teaching for universally designed learning*. Thousand Oaks, CA: Corwin Press.
- Tomlinson, C. A. 2014. *The differentiated classroom. Responding to the needs of all learners*. (2. painos). Alexandria, VA: ASCD.
- Tomlinson, C. A. & M. Imbeau 2010. *Leading and managing a differentiated classroom*. Alexandria, VA: ASCD.
- Tomlinson, C. A. & T. R. Moon 2013. *Assessment and student success in a differentiated classroom*. Alexandria, VA: ASCD.
- VanTassel-Baska, J. & T. Stambaugh 2005. Challenges and possibilities for serving gifted learners in the regular classroom. *Theory into Practice*, 44 (3), 211–217. https://doi.org/10.1207/s15430421tip4403_5
- Wan, S. W. Y. 2017. Differentiated instruction: are Hong Kong in-service teachers ready? *Teachers and Teaching*, 23 (3), 284–311. <https://doi.org/10.1080/13540602.2016.1204289>
- Xu, Y. & G. T. Brown 2016. Teacher assessment literacy in practice: a reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Yavuz, A. C. 2020. The effects of differentiated instruction on Turkish students' L2 achievement, and student and teacher perceptions. *Eurasian Journal of Applied Linguistics*, 6 (2), 313–335. <https://orcid.org/0000-0002-3246-952X>

LIITE 1:

Arviontikysely lukion vieraan kielen opettajille

Pakolliset kysymykset merkitty tähdellä (*)

Tässä tutkimuksessa kartoitetaan lukion vieraan kielen opettajien kokemuksia arvioinnista ja eriyttämisestä, sekä pandemian vaikutuksista arviointiin. Vastaajista ei kerätä mitään henkilökohtaisia tunnistetietoja. Voit tarvittaessa tallentaa kyselyn kesken vastaamisen ja jatkaa sen täyttämistä myöhemmin. Tutkimuksen toteuttavat Anssi Roiha (TY), Juuso Nieminen (Hongkongin yliopisto) ja Katja Mäntylä (JY). Lisätietoja: katja.mantyla@jyu.fi

Linkki EU-lainsäädännön mukaiseen tietosuojaselosteeseen:
bit.ly/ArviointikyselyTietosuojaseloste

Taustasi

1. Mitä aineita opetat lukiossa? *
 - englanti
 - espanja
 - italia
 - japani
 - kiina
 - latina
 - portugali
 - ranska
 - ruotsi
 - saksa
 - suomi toisena kielenä
 - muu kieli, mikä?
 - muu kuin kieliaine, mikä?

2. Kuinka monta vuotta olet toiminut opettajana? *
 - 0–5 vuotta
 - 6–10 vuotta
 - 11–15 vuotta
 - 16–20 vuotta
 - 21–25 vuotta
 - 26–30 vuotta
 - yli 30 vuotta

3. Oletko osallistunut täydennyskoulutukseen arviointiin liittyen? *
 - kyllä, mistä aiheesta, miten laaja?
 - en

4. Oletko osallistunut täydennyskoulutukseen eriyttämiseen liittyen? *
 - kyllä, mistä aiheesta, miten laaja?
 - en

5. Kuinka monta erityisen tuen/oppimisen tukea tarvitsevaa oppilasta opettamissasi opetusryhmissä on keskimäärin? *
 - 0–5
 - 6–10
 - 11–5
 - 16–20
 - yli 20

Pohdi alla olevia kysymyksiä vieraiden kielten opetuksen näkökulmasta lukion opetuksessa.

6. Kuvaile, miten arviointityösi on muuttunut uuden LOPSin myötä.

7. Käytkö oppilaiden kanssa kurssin aluksi läpi, mitä arvioidaan?
 - kyllä, miten?
 - en

8. Kuinka usein käytät opetuksessasi seuraavia arviointimenetelmiä opintojakson aikana? (1. en koskaan 2. harvoin 3. joskus 4. usein 5. aina)
 - sanakoe
 - kuuntelukoe
 - suljettu koe jakson päätteeksi (ilman materiaalia)
 - avoin koe jakson päätteeksi (materiaalin kanssa)
 - suullinen arviointi
 - itsearviointi
 - vertaisarviointi
 - projekti
 - portfolio
 - esitelmä
 - kirjoitelma/essee
 - jatkuva arviointi
 - oppimiskeskustelu
 - muu. mikä?

9. Millä tavalla huomioit oppilaiden yksilöllisyyden arvioinnin eri vaiheissa (esim. tavoitteiden asettaminen, oppimisprosessin arviointi, oppimistulosten arviointi)?

10. Kuvaile, kuinka koronapandemia on vaikuttanut arviointityöhösi (sekä formatiiviseen että summatiiviseen)?

11. Kuvittele, että saisit toteuttaa arviointia ”ideaalitulanteessa”, jossa tosielämän realiteetit eivät vaikuttaisi arviointityöhösi. Kuvaile, miten toteuttaisit arviointia.

122 LUKION VIERAIDEN KIELTEN OPETTAJIEN KÄYTÄNTEITÄ
JA KÄSITYKSIÄ ERIYTTÄVÄSTÄ ARVIOINNISTA

12. Jätimmekö kysymättä jotain, mitä sinulta olisi ehdottomasti kannattanut kysyä arviointiin liittyen? Voit halutessasi jatkaa aiheesta tässä.

Olisitko halukas osallistumaan lyhyeen haastatteluun tästä teemasta Zoomin välityksellä? Kyselyn päätteeksi ohjautut uudelle sivulle, jossa kysytään halukkuuttasi osallistua lyhyeen haastatteluun ja sähköpostiosoitettasi. Osoitetta ei voi yhdistää kyselyvastauksiin. Halutessasi jätä yhteystietosi haastattelua varten, muussa tapauksessa sulje selain, kun olet lähettänyt kyselyvastauksesi. Kiitos!

LIITE 2:

Haastattelurunko

Taustatiedot:

- Opetettavat aineet
- Opetuskokemus
- Perus- ja täydennyskoulutus liittyen arviointiin + eriyttämiseen
- Koulun koko

Tieto arvioinnista

- Millaisia ajatuksia herättävät arviointiuudistukset Suomessa?
- Miten sinä toteutat arviointia omassa työssäsi?
- Kuinka toteutat eriyttämistä arvioinnissa?/ Kuinka huomioit yksilöllisyyden arvioinnissa?
- Kuvaile omasta opetuksesta arviointitilanne, joka...
 - » ...oli erityisen onnistunut eriyttämisen näkökulmasta. Miksi?
 - » ...epäonnistui eriyttämisen näkökulmasta. Miksi?
- Mitä haluaisit oppia lisää eriyttämisestä arvioinnista?

Käsitykset arvioinnista

- Miten kommentoit seuraavia väittämiä:
 - » Lukiossa on opiskelijoita niin paljon, että arviointi kannattaa suunnitella "keskiverto-opiskelijalle".
 - » Lukiossa ei tarvitse eriyttää arviointia.
 - » Arviointi on tasavertaista, kun se on kaikille samanlaista.
 - » Eriyttäminen tekee arvioinnista epäreilua.
 - » Arvioinnin eriyttäminen ei ole järkevää lukiossa, koska ylioppilaskirjoitukset ovat kaikille samat.
- Mikä on mielestäsi arvioinnin tärkein tehtävä? Miksi?
- Mikä muuttuisi sinun opetuksessasi, jos arviointi poistettaisiin lukiosta kokonaan? Miksi?

Arviointitiedon käytäntöön siirtäminen

- Millaiset tekijät mahdollistavat mielestäsi sen, että voit tehdä arviointia siten kuin näet tärkeäksi ja hyväksi?
- Millaisia rajoitteita mielestäsi on sille, ettet voi mahdollisesti tehdä arviointia siten kuin näet tärkeäksi ja hyväksi?
- Kuinka koronatilanne on vaikuttanut sinun arviointitoimintaasi?

Lopuksi

- Kuvittele, että saisit täysin omavaltaisesti sanella, kuinka kielten oppimista arvioidaan lukiossa. Miltä arviointi tällöin näyttäisi?
- Haluatko vielä lisätä jotain haastattelun teemoista?

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 124–141.

Anna von Zansen & Maikki Heijala
Helsingin yliopisto

Miten suomen ja ruotsin opettajat käyttäisivät puheen automaattiseen arviointiin kehitettyä työkalua?

Nostot

- Automaattinen palaute tuo uusia mahdollisuuksia puhumisen harjoitteluun.
- DigiTala-työkalun automaattinen palaute konkretisoi taitotasoja ja puheesta mitattavia piirteitä kielten opettajille ja opiskelijoille.
- Analyttinen palaute tuo esiin puhujan vahvuuksia ja kehittämiskohteita.
- Haastatellut opettajat eivät vielä hyödynnä taitotasoasteikkoja suullisen kielitaidon arvioinnissa.

Abstract

In this study, we interviewed four language teachers in order to investigate their perceptions of an automated tool developed for assessing Finnish and Swedish learners' speech. We showed the first version of the Moodle plugin developed by the DigiTala research project to the teachers. In this version, after recording a speech sample, the system provides analytic and holistic feedback to the learner. The automated feedback was developed based on the rating criteria that human raters have used earlier in the project. The data consist of four semi-structured individual interviews conducted in Zoom. The interviews were transcribed and analysed using content analysis. We found that automated feedback has significant implication areas in language learning and teaching. Teachers found the feedback statements understandable and useful, even though they do not cover all dimensions relevant for the speaking construct. The results will benefit those working on automated assessment of oral language skills.

Keywords: automated assessment, language assessment, oral language skills, automated feedback

Asiasanat: automaattinen arviointi, kielitaidon arviointi, suullinen kielitaito, automaattinen palaute

1 Johdanto

Tämä soveltava tutkimus käsittelee automaattista palautetta kielitaidon formatiivisen arvioinnin kontekstissa. Automaattinen palaute on yleistynyt monenlaisissa oppimisympäristöissä (Deeva ym. 2021), joista esimerkkinä ovat kielten opiskelun virtuaaliset tutorit (*tutoring systems*, Golonka ym. 2014). Tutkimus liittyy monitieteiseen DigiTala-hankkeeseen, jossa kehitetään puheentunnistusta (*automatic speech recognition*, ASR) ja automaattista arviointia (*automatic scoring*) hyödyntävää verkkotyökalua suomen ja ruotsin puhumisen harjoitteluun (Kautonen & von Zansen 2020). Puheentunnistuksen ja automaattisen arvioinnin lisäksi työkalun (von Zansen ym. 2022a) kolmas keskeinen ominaisuus liittyy puhujalle näytettävään automaattiseen palautteeseen (*automated feedback*, ks. katsaus automaattiseen palautteeseen oppimiskonteksteissa Deeva ym. 2021, suullisen kielitaidon osalta Evanini & Zechner 2020; Gu & Davis 2020). Työkalun kehityksessä hyödynnetään foneettista (esim. Kallio ym. 2022; Kallio ym. arvioitavana), kielididaktista (esim. von Zansen ym. 2022b; von Zansen & Hildén 2022; von Zansen ym. 2022c) sekä kieliteknologiaan (esim. Al-Ghezi ym. 2021; Al-Ghezi ym. arvioitavana) liittyvää tutkimusta.

Työkalua voitaisiin käyttää osana formaalia koulutusta, mutta myös sen ulkopuolella, kun kielen oppijat harjoittelevat puheen tuottamista itsenäisesti. Formatiivisen arvioinnin lisäksi työkalu voisi toimia ihmisarvioinnin tukena isoissa ja tärkeissä kielikokeissa (ylioppilastutkinnon osalta ks. Vaarala ym. 2021). Työkalun ensisijaisena kohderyhmänä ovat suomalaiset lukiolaiset, jotka opiskelevat suomea tai ruotsia toisena tai toisena kotimaisena kielenä. Myöhemmin työkalun kohderyhmää ja kohdekieliä on tarkoituksena laajentaa – esimerkiksi akateemisten maahanmuuttajataustaisten

aikuisopiskelijoiden (n=115) suhtautuminen tietokoneella suoritettavaan puhumisen kokeeseen on ollut positiivista (von Zansen & Hildén 2022).

Puheen automaattiseen arviointiin perustuvat työkalut pystyvät tunnistamaan vain tiettyjä automaattisesti mitattavissa olevia piirteitä suullisesta kielitaidosta (Evanini & Zechner 2020; Xi 2010), joten tässä artikkelissa käytämme termejä puhuminen ja puheen arvioiminen. Kielten oppimisen kannalta puheen automaattinen arviointi nähdään siihen liittyvistä rajoituksista huolimatta lisäresurssina (Gu & Davies 2020: 159), sillä se lisää suullisen kielitaidon harjoittelumahdollisuuksia tuomalla joustoja ajan ja paikan suhteen. Lisäksi automaattinen arviointi yhdenmukaistaa arviointia ja säästää resursseja – kone arvioi väsymättä kaikki puhujat samalla tavalla (Evanini & Zechner 2020: 3).

Artikkelissa tutkimme kieltenopettajien ensivaikutelmia työkalusta (von Zansen ym. 2022a) haastatteleamalla neljää kieltenopettajaa. Arvioinnin validiustarkastelua on syytä rajata, sillä erilaiset validiusrakenteet ovat keskeisiä riippuen arviointikon-
tekstistä (Xi 2010: 293). Tämän tutkimuksen lähtökohtana on formatiivinen arviointi, jossa oppija saa työkalun antaman analyttisen eli diagnostisen palautteen (*diagnostic feedback*) avulla tietoa vahvuuksistaan ja kehittämiskohteistaan, mikä puolestaan tukee oppijan suullisen kielitaidon kehittymistä ja kieliminän rakentumista. Analyttisen palautteen lisäksi työkalu antaa holistisen arvion puhujan taitotasosta. Jotta palautteesta olisi hyötyä, automaattisen palautteen tulisi kuitenkin olla oppijalle ymmärrettävää, kannustavaa, ja se tulisi saada sopivassa muodossa oikeaan aikaan.

2 Tutkimuksen tausta ja tutkimuskysymykset

Tämän soveltavan kielentutkimuksen alaan kuuluvan artikkelin tavoitteena on tuottaa uutta tietoa automaattisen arvioinnin ja palautteen hyödyntämismahdollisuuksista suullisen kielitaidon harjoittelussa ja opetuksessa (ks. tutkimuskysymykset 1 ja 2 myöhemmin tässä luvussa). Kuten soveltavalle tutkimukselle on luonteenomaista, osallistamme opettajia työkalun suunnitteluun näyttämällä ensimmäistä versiota kehittämästämme työkalusta. Tutkimuksessa yhdistämme sekä laadullisen haastattelututkimuksen (Hyvärinen ym. n.d.) että käytettävyydestauksen (Ovaska ym. 2005) menetelmiä. Käyttäjiltä kerätty palaute kuuluu keskeisesti erilaisten työkalujen ja palvelujen kehitys- ja validointityöhön.

Automaattisia puhumisen arviointityökaluja on olemassa englannin kielelle, mutta näissäkin järjestelmissä automaattinen palaute rajoittuu yleensä ääntämiseen (Golonka ym. 2014). Poikkeuksena voidaan kuitenkin mainita ETS:n SpeechRater (Gu & Davis 2020), Cambridge Englishin Linguaskill (Xu ym. 2020) -työkalut ja sekä Pearson Educationin (2018, 2022) Versant English ja Versant Arabic -testit. Lisäksi automaattista palautetta on hyödynnetty esimerkiksi hollannin sanajärjestyksen harjoittelussa (de Vries ym. 2014). Käyttäjiä haastatteleamalla saadaan konkreettisia kehitysehdotuksia,

joiden perusteella palautenäkymää voidaan muokata (Gu & Davies 2020). Käyttäjien näkemykset voivat myös auttaa tarkastelemaan laajemmin työkalun käyttökokemusta (de Vries 2014) sekä arvioimaan, missä määrin puhumisen tehtävät vastaavat arkielämän viestintätilanteita (Xu ym. 2020).

Suomalaisessa DigiTala-hankkeessa olemme keränneet suomen (n=308) ja ruotsin (n=178) puhujilta puheaineistoa (yli 3500 puhenäytettä, yhteensä noin 20 h puhetta). Puhenäytteiden ja ihmisiltä kerättyjen arviointien avulla olemme kehittäneet automaattista arviointia. Aiemmin olemme tutkineet eri käyttäjäryhmien suhtautumista puheen automaattiseen arviointiin sekä heidän näkemyksiään arviointikriteereistä ja tehtävistä sekä heidän kokemuksiaan puhetehtävien suorittamisesta tietokoneella (von Zansen ym. 2022b, von Zansen ym. 2022c, von Zansen & Hildén 2022). Olemme muun muassa havainneet, että opiskelijat (n=58) tuntevat tekoälyn mahdollisuuksia ja rajoituksia sekä kykenevät pohtimaan automaattiseen arviointiin liittyviä hyötyjä ja uhkia (von Zansen ym. 2022b). Puheen analyttinen arviointi tutkimushankkeen kriteereillä (von Zansen 2022) on osoittautunut arvioijille (n=37) haastavaksi etenkin liian lyhyistä puhenäytteistä (von Zansen ym. 2022c). Kouluttamamme arvioijat ja laatimamme arviointikriteerit ovat kuitenkin osoittautuneet luotettaviksi (von Zansen & Huhta 2022). Sekä opiskelijat että ihmisarvioijat pitävät vuorovaikutuksellisuutta tärkeänä ominaisuutena puhumisessa (von Zansen ym. 2022c, von Zansen & Hildén 2022), vaikka sen mittaaminen automaattisesti on haastavaa (Evanini & Zechner 2020; Xi 2010). Automaattiseen arviointiin liittyvien rajoitteiden takia sen nähdään soveltuvan etenkin ihmisarvioinnin rinnalle (hybridiarvioinnista Evanini & Zechner 2020: 11). Kuten seuraavassa luvussa kerrotaan tarkemmin, saimme kevään 2022 aikana luotua ensimmäisen version automaattisesta puheen arviointijärjestelmästä Moodleen. Toukokuussa 2022 esittelimme työkalua haastatellen sekä opettajia että opiskelijoita, ja tässä artikkelissa käsittelemme ensimmäistä palautenäkymän versiota opettajien näkökulmasta.

Tutkimus perustuu Xin (2010) laajempaan automaattiseen arviointia koskevaan validiusmalliin, jossa automaattisen palautteen (Xi 2010: 296) validiutta voidaan tarkastella sen arviointia (*evaluation*), yleistämistä (*generalization*), selittämistä (*explanation*), sekä käyttöä (*utilization*) koskevien kysymysten avulla. Tämän tutkimuksen kannalta keskeiset validiusnäkökulmat liittyvät kahteen jälkimmäiseen, nimittäin automaattisen palautteen seurauksiin (käyttö ja selittäminen, Xi 2010). Pyrimme tarkastelemaan automaattisen palautteen tuomia hyötyjä oppimisen ja puhumisen kehittymisen kannalta sekä arvioimaan, onko automaattisella palautteella positiivinen vaikutus opetukseen ja oppimiseen (Xi 2010: 296).

Tässä pilottivaiheen artikkelissa keskitymme kieltenopettajien näkemyksiin työkalun ensimmäisestä versiosta. Vastaamme kahteen tutkimuskysymykseen, joista ensimmäisen avulla voimme tehdä johtopäätöksiä automaattisen palautteen käytöstä ja pohtia arvioinnin mahdollisia vaikutuksia. Lisäksi tutkimme, kokevatko opettajat puhe-suorituksista annettavan automaattisen palautteen hyödylliseksi oppimisen

kannalta (tutkimuskysymys 2). Toisin sanoen tutkimme, ovatko hankkeen laatimien arviointikriteerien (von Zansen 2022) pohjalta tuotetut palauteväittämät kieltenopettajien mielestä ymmärrettäviä ja hyödyllisiä. Tutkimuskysymykset ovat:

1. Millaisia mahdollisuuksia ja rajoituksia haastatellut kieltenopettajat näkevät työkalun käytölle?
2. Miltä osin haastatellut kieltenopettajat pitävät työkalun antamaa analyttistä ja holistista palautetta hyödyllisenä?

Tämä laadullinen tutkimus tuottaa uutta tietoa automaattisen palautteen ensimmäisestä versiosta. Kieltenopettajien näkemykset työkalusta ovat keskeisiä oppijoille suunnatun automaattisen palautteen kehityksessä, sillä niiden tutkiminen lisää ymmärrystä siitä, millaista automaattista palautetta puhumisen suorituksista kannattaisi antaa. Tutkimustiedon avulla palautetta voidaan muotoilla tarkoituksenmukaisempaan ja ymmärrettävämpään muotoon, jotta se hyödyttäisi mahdollisimman erilaisia oppijoita ja olisi linjassa opetuksen tavoitteiden ja sisältöjen kanssa.

3 Aineisto ja menetelmät

Tutkimushanke toteutti Moodle-liitännäisen (*plugin*) yhteistyössä Helsingin yliopiston tietojenkäsittelytieteen opiskelijoiden kanssa (von Zansen ym. 2022a). Tässä artikkelissa analysoidaan neljän kieltenopettajan puolistrukturoidut haastattelut järjestettiin toukokuussa 2022 Zoom-videopuhelun avulla. Seuraavissa alaluvuissa esittelemme tutkimuksen aineiston ja menetelmät. Luvussa 3.1 esittelemme Moodlessa toimivaa työkalua, luvussa 3.2 tutkimuksen osallistajat, luvussa 3.3 haastattelujen toteutuksen ja lopulta luvussa 3.4 aineiston analyysimenetelmät.

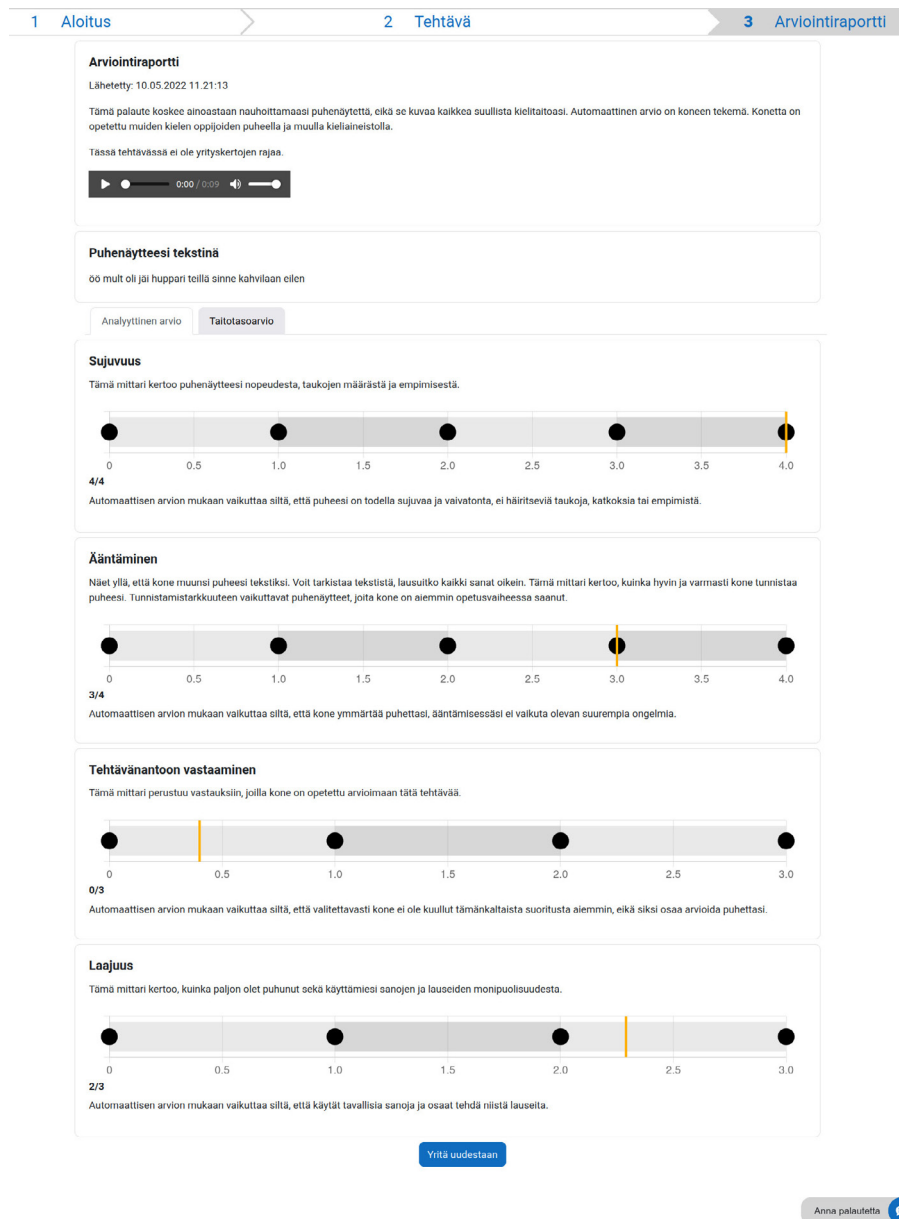
3.1 Palautetyökalu Moodlessa

Selaimessa toimivan Moodle-työkalun (von Zansen ym. 2022a) avulla oppija äänittää ensin puhe-suorituksensa, joka arvioidaan Aalto-yliopiston palvelimilla automaattisesti, minkä jälkeen työkalu esittää automaattisen palautteen oppijalle. Työkalun ominaisuuksia esitellään tarkemmin käyttöohjeessa (Alanen ym. 2022), ja automaattisen puheen tunnistuksen ja arvioinnin kehitystä käsittelevät Al-Ghezi ym. (arvioitavana).

Ensimmäisessä vaiheessa opiskelijoille annettava automaattinen palaute muotoiltiin arviointikriteereistä (von Zansen 2022), joita hankkeeseen osallistuneet ihmisarvioijat ovat käyttäneet automaattisen arvioinnin kehitysvaiheessa. Näitä kriteerejä verrattiin puheesta automaattisesti arvioitaviin piirteisiin (Al-Ghezi ym. arvioitavana), minkä pohjalta palauteväittämät muotoiltiin. Holistisen taitotasoarvion lisäksi automaattista palautetta annetaan opiskelijoille neljästä analyttisestä ulottuvuudesta, jotka

129 MITEN SUOMEN JA RUOTSIN OPETTAJAT KÄYTTÄISIVÄT PUHEEN AUTOMAATTISEEN ARVIOINTIIN KEHITETTYÄ TYÖKALUA?

ovat tehtävän suorittaminen, sujuvuus, ääntäminen sekä ilmaisuuden laajuus. Sanallisen palautteen lisäksi keltaisen viivan avulla visualisoidaan oppijan puhe-suorituksen sijoittuminen kullekin skaalalle (analyttisista skaaloista ks. myös von Zansen & Huhta 2022), ja skaalan portaiden kuvaukset tulevat näkyviin viemällä hiiri kunkin mustan pallon päälle (ks. kuvio 1). Lisäksi työkalun avulla opettaja tai tutkija voi lisätä oman arvionsa, kommentoida äänitettyny puhenäytettä tai automaattista palautetta sekä ladata äänitetyt puhenäytteet ja arviointitiedot (Alanen ym. 2022).



KUVIO 1. DigiTala-työkalun palautenäkymä.

3.2 Osallistujat

Tutkimuksen otantamenetelmänä käytettiin harkinnanvaraista näytettä, sillä tutkimuksen tavoitteena oli selvittää automaattisen palautteen käyttötarkoituksia ja hyödyllisyyttä kielten opetuksessa (ks. luku 2). Kuten laadullisessa tutkimuksessa on tavallista, osallistujat valittiin tarkoituksenmukaisesti edustamaan keskeisiä käyttäjäryhmiä eli sekä suomen että ruotsin opettajia eri kouluasteilta. Haastateltavat valittiin lähettämällä haastattelukutsu 20:lle 60:stä DigiTala-projektiin aiemmin osallistuneesta opettajasta, joiden tiesimme työskentelevän opettajana tällä hetkellä. Heistä tutkimusta varten haastateltiin neljää kieltenopettajaa, joihin viittaamme artikkelissa jatkossa tunnuksilla H1, H2, H3 ja H4. Osallistujista yksi opetti haastatteluhetkellä ruotsia (H2) ja kolme suomea (H1, H3, H4) eri koulutusasteilla yläkoulusta yliopistokursseihin. Haastatteluun osallistuneista opettajista yhdellä (H4) oli opetuskokemusta opetusaineessaan alle viisi vuotta ja muilla noin 20–30 vuotta. Haastatellut opettivat kotimaisten kielten eri oppimääriä: H1 finskan A-, B- ja mofi (*modersmålsinriktad*)-oppimääriä, H2 A- ja B-ruotsia yläkoulussa ja lukiossa, H3 suomi äidinkielenä ja kirjallisuus - ja suomi toisena kielenä -oppimääriä lukiossa sekä IB-lukion Finnish B -oppimäärää kahdella tasolla, H4 suomi toisena ja vieraana kielenä -oppimäärää yliopistossa.

3.3 Haastattelujen toteutus

Haastattelut toteutettiin Helsingin yliopiston Zoom-videoneuvottelutyökalulla toukokuun 2022 loppupuolella. Puolistrukturoidut haastattelut kestivät 50–65 minuuttia, ja haastattelun aikana haastattelija esitteli opettajille DigiTala-työkalua (versio 0.1.40) Moodlella (versio 4.0) näytönjaon avulla. Opettajilta kerättävät taustatiedot minimoitiin, mikä helpotti laadullisen aineiston anonymisointia (Yhteiskuntatieteellinen tietoarkisto n.d.). Haastattelija toi kaikissa haastatteluissa ilmi, että hän on ollut mukana rakentamassa haastattelun keskiössä olevaa Moodle-työkalua, ja pyrki tietoisesti vaikuttamaan haastateltavien vastauksiin mahdollisimman vähän.

Haastattelukysymykset (Heijala & von Zansen 2022) laadittiin kuuden aihepiirin ympärille. Tässä artikkelissa käsittelemme haastattelun osia I taustatiedot, III ensivaikutelmat palautenäkymästä, IV analyttinen palaute, V holistinen palaute ja VI työkalun ja automaattisen arvioinnin käyttömahdollisuudet. Haastattelun osa II käsitteli menetelmiä, joita opettajat tällä hetkellä käyttävät puhumisen opetuksessa ja arvioinnissa, ja rajasimme tämän osan artikkelin ulkopuolelle.

Hyödynsimme haastattelukysymysten laadinnassa aiempaa kokemustamme tutkijoina ja kieltenopettajina sekä tutkimuskirjallisuutta. Kysymysten laadinnassa kiinnitimme huomiota menetelmään liittyviin ongelmiin (ks. esim. Hyvärinen ym. n.d.). Pyrimme esimerkiksi välttämään liian monen asian kysymistä samalla kertaa ja jätimme tilaa haastateltavan vastauksille. Ennen aineistonkeruuta kokeilimme haastattelurunkoa projektitiimin opettajataustaisten tutkijoiden kesken. Pilotoinnin

perusteella muutimme esimerkiksi vaikeaselkoisia kysymyksiä ja järjestimme haastattelurungon kysymykset uudelleen.

Ennen haastattelua osallistujille lähetettiin työkalusta lyhyt esittelyvideo, jonka katsominen oli vapaaehtoista. Haastatteluissa keskityttiin palautenäkymään, jonka oppija näkee nauhoitettuaan äänitteen tuottamistehtävään. Haastattelija oli valmistellut näkymän etukäteen nauhoittamalla lyhyen puhe-suorituksen tehtävään, jossa puhujan tulee soittaa kahvilaan, jonne hän on jättänyt hupparin edellispäivänä. Etukäteen nauhoittamisella vältettiin automaattiseen arviointiin liittyvät odotusajat tai tekniset ongelmat, joita haastattelutilanteessa tehdyssä nauhoituksessa olisi voinut ilmetä. Palautenäkymässä analyttiset arviot näkyvät yhdellä välilehdellä ja holistinen taitotasoarvio toisella (ks. kuvio 1, luku 3.1).

3.4 Analyysimenetelmät

Haastattelut tallennettiin Zoom-videoneuvottelutyökalun avulla tutkijan tietokoneelle, jonka jälkeen litteroimme ne peruslitteroinnin (Yhteiskuntatieteellinen tietoarkisto 2022) periaatteita noudattaen. Esimerkiksi aiheeseen liittymättömät keskustelut jätettiin litteroimatta, ja haastattelijan selostamat käyttöliittymän toiminnot referoitiin. Seuraavaksi yhdistimme neljän haastattelun litteraattit haastattelun kuuden osan perusteella. Taustatietoja lukuun ottamatta kunkin osan (osat II–VI, ks. Heijala & von Zansen 2022) aineisto vietiin tekstitiedostona Atlas.ti -ohjelmaan (versio 22.0), jonka avulla toteutimme sisällönanalyysin (Vuori n.d.). Sisällönanalyysissa hyödynsimme myös aiempaa kokemustamme kieltenopettajina sekä aiempaa tutkimusta. Ensin koodasimme Atlas.ti:n avulla opettajien vastauksia 27 luokkaan, mikä auttoi hahmottamaan vastausten yhtäläisyyksiä.

Opettajahaastattelut olivat sisällöltään melko laajoja (artikkelista pois rajatun osan II litteraatti 16 sivua, osat III–VI yhteensä 26 sivua) ja vastaukset haastattelijan kysymyksiin perusteellisia, joten Atlas.ti:n tarjoama dekonstruoiva tyypittely ei tässä vaiheessa vaikuttanut jatkoanalyysin kannalta enää hedelmälliseltä. Havaitimme, että eri opettajien kommentit rinnakkain kysymyskohtaisesti analysoituna tarjoaisivat paremman kokonaiskuvan heidän näkemyksistään. Seuraavaksi etenimme sisällönanalyysissa tutkimuskysymysten ohjaamina. Nämä uudelleen ryhmitellyt litteraattit tiivistettiin luetteloiksi, jotka sisälsivät vastausten keskeisen sisällön, minkä perusteella voitiin luoda yleiskuva opettajien näkemyksistä. Lopulta jäsensimme sisällönanalyysin tulokset ja valitsimme aineistosta kuvaavia esimerkkejä havaintojemme tueksi.

4 Tulokset

Esittelemme seuraavissa alaluvuissa tulokset tutkimuskysymysten mukaisessa järjestyksessä. Alaluku 4.1 käsittelee haastateltavien tunnistamia mahdollisuuksia ja

rajoituksia työkalun käytössä. Luvussa 4.2 kuvaamme haastateltavien näkemyksiä analyttisen ja holistisen palautteen hyödyllisyydestä.

4.1 Työkalun mahdollisuudet ja rajoitukset

Haastattelujen perusteella automaattinen arviointi tuo lisäresurssin puhumisen harjoitteluun itsenäisesti tai opettajan johdolla. Opettajat käyttäisivät työkalua esimerkiksi kotitehtävissä, osana kieliprofilia sekä arvioinnin tukena (lainaus 1). Lisäksi työkalu soveltuisi osaksi suullisia kokeita. Työkalun pelillisuus voisi myös lisätä oppijan motivaatiota (lainaus 2).

- (1) – – mahdollisuus siis siinä [lukion] suullisessa syventävässä kurssissa hyödyntää tätä. – – Silloinhan sen vois nauhottaa ja käyttää Moodlea ja silloin ne sais palautetta ja opettaja sais nähdä sen palautteen ja tää auttais sitten meitä opettajia siinä arvioinnissa. H1
- (2) – – tietynlainen pelillisuus ja semmonen motivaatiotekijä joillekin eli harjotella sitä ja yrittää päästä vaikka sitten seuraavalle tasolle ja kokeilla mitkä asiat siellä se kone nyt sitten huomaa – – H2

Tunnistetut haasteet puolestaan liittyivät työkalun ja oppimisympäristöjen yhteensopivuuteen, tietosuojaan ja teknologiavastaisuuteen (lainaus 3 ja lainaus 4). Opettajatkaan eivät luottaisi pelkkään automaattiseen arviointiin isoissa ja tärkeissä tehtävissä (lainaus 5). Myös mahdolliset kielteiset vaikutukset opiskelijan itsetuntoon huolettivat opettajia (lainaus 6).

- (3) – – helppo tapa tallentaa – – miten tää on siirrettävissä sitten nää näytteet jonnekin muille alustoille – – H2
- (4) Kaikki opiskelijat ei ehkä pidä tämmösestä, meil on siis tämmösiä tietotekniikkavastaisiakin opiskelijoita, kumma kyllä, mutta mä uskon että useimmat kyllä tekisivät innostuneesti. H3
- (5) – – edut on tietysti tää palaute, jonka opiskelija saa, ja se tietysti auttaa myös opettajaa – – kielteinen voi olla, et ne luottaa liikaa siihen koneeseen – – ei saa jäädä vaan sen koneen varaan se arviointi – – isompi tehtävä tai tärkeämpi tehtävä, silloin myös opettaja joku antaa suullista palautetta. H1
- (6) Ehkä just se että se ei oo sataprosenttisesti luotettavaa tää palaute ja jos se toistuvasti sanoo jotain että sä et oo hyvä – – jos opiskelija yrittää ja yrittää ja saa jotain ei-niin-hyvää palautetta, niin ehkä siitä vois olla haittaakin. H4

133 MITEN SUOMEN JA RUOTSIN OPETTAJAT KÄYTTÄISIVÄT PUHEEN AUTOMAATTISEEN ARVIOINTIIN KEHITETTYÄ TYÖKALUA?

Haastateltavat toivat lisäksi esiin arvioinnin läpinäkyvyyteen liittyvän tarpeen ymmärtää, mistä automaattinen arvio koostuu (lainaus 7). Automaattisen arvioinnin ymmärtäminen on lähtökohta arvioinnin luotettavuuden ja mahdollisten haasteiden pohtimiselle. Esimerkiksi puhenäytteiden lyhyys (lainaus 8) ja erilaisten puhujien huomioiminen (lainaus 9) huolettivat, kuten olemme havainneet muissa tutkimuksissamme (von Zansen ym. 2022b, von Zansen ym. 2022c).

- (7) Mun on vaikee kun mä en tiedä mitä kaikkea se kone osaa [aurua], mitä sinne on syötetty, minkälaisia asioita se pystyy nyt sit sanoo eri tasoille – – H2
- (8) Kun nää näytteet on aika lyhyitä, aika usein tulee sellanen tilanne että se taitotaso-arvio se ei oo ihan oikea – – että siis tää ei oo ehkä luotettavin mittari. H4
- (9) Mä mietin vaan opiskelijoita, joilla on jotain puhevikaa tai jotain, mitä silloin tapahtuu – – miten tarkka tää on tää apuväline nyt sitten niitä ajatellen – – [joilla] on joku ärrävika tai jotain, onks tässä jotain mitä pitäis vielä miettiä? H1

Työkalan ulkonäkö herätti haastateltavissa ajatuksia, ja esimerkiksi värikyyttä toivottiin käyttöliittymään. Opettajat tuntevat kuitenkin saavutettavuusvaatimuksia ja pitivät tärkeänä työkalun soveltuvuutta eri käyttäjäryhmille (lainaus 10). Analyttisissa palauteväittämissä skaalojen eripituisuus kiinnitti huomiota (lainaus 11). Lisäksi haastateltavat toivat esiin tarpeen muuttaa työkalun sisältöä vastaamaan omiin tarpeisiinsa (lainaus 12). Palautteen näyttämistä eri kielillä (suomeksi, ruotsiksi tai englanniksi) pidettiin tärkeänä, ja joitakin palauteväittämiä voisi yksinkertaistaa (lainaus 13).

- (10) – – on aika iso prosentti ihmisiä jotka eivät näe värejä eli ovat värisokeita – – H4
- (11) – – se laajuus, siis oliko se kolme, se on maksimi? Eli näissä ylemmissä se oli nelonen, maksimi oli nelonen, skaala on vähän laajempi näissä. H4
- (12) Pystyykö opettaja vaihtaa tätä asteikkoa mitenkään? H3
- (13) – – semmoselle jolla on pienempi kielitaito nää tekstiosuudet voi olla pikkusen hankalia, esimerkiksi nyt tämmönen sana ku empimisestä. Sitä nyt ei ymmärrä ihan kaikkea. H3

Kaiken kaikkiaan haastateltavat pitivät koneen tarjoamaa analyttistä arviota kattavana. Opettajat toivat kuitenkin esiin joitakin puheen ulottuvuuksia, jotka nyt jäävät automaattisen arvioinnin ulkopuolelle. Esimerkiksi laajuuden lisäksi tulisi arvioida puheen oikeellisuutta (lainaus 14), ja vuorovaikutusta voitaisiin arvioida virtuaalisen keskustelukumppanin kanssa simuloitussa keskustelussa (lainaus 15). Puheen sisällön

ja luovuuden automaattinen arviointi tunnistettiin haastavaksi (lainaukset 16 ja 17). Tällä hetkellä työkalu ei myöskään arvioi nonverbaalista viestintää, kuten kehonkieltä (lainaus 18).

- (14) – – tää ei ota nyt ihan vielä kantaa sitten siihen tavallaan oikeellisuuteen. – – se nyt ei oo se ensimmäinen asia mitä me ruvetaan tarkastelemaan että onks kaikki rakenteet oikein, mutta kyllähän se yks osa tätä on. H2
- (15) – – joku videolla – – sinulta kysyy et hei hei mistäs olet kotoisin ja sitten on tietty aika vastata – – siellä olis tämmönen video, joka antaa sen oman repliikin jälkeen sitten tietyn vastausajan. H2
- (16) – – täällä on tää tehtävänantoon vastaaminen, se sisältöpuoli ei tässä tietenkään näy – – H3
- (17) – – videoissa, joita meillä oli käytössä nyt keväällä, arvioidaan myös kuinka luova opiskelija on. – – Tässä ei saa olla niin luovaa. – – mutta miten kone voi arvioida luovuutta? H4
- (18) – – kun puhutaan kommunikoinnista, niin siinä on muitakin tekijöitä, jotka on tärkeitä. Et opettaja näkee niin paljon muuta sit kuitenkin. Et nehän tulee nää kaikki kehonkieli ja kaikki muut [naurua]. Tää [koneen palaute] on aika raakaa, jos sanotaan näin. H1

Haastateltavat näkivätkin automaattisen arvioinnin soveltuvan etenkin opettajan työn tueksi, eivät sitä kokonaan korvaavaksi. Haastateltavat pitivät tärkeänä, että opettajat voivat kommentoida automaattisia arvioita ja tarjota puhujalle monipuolisempaa palautetta (lainaukset 19 ja 20).

- (19) – – kone antaa omat arvionsa ja sitten vielä kun mä korjaan ja kuuntelen näitä niin sitten mä vielä annan palautetta. H1
- (20) – – mun mielestä se on tärkeitä et opettajalla on se mahdollisuus lisätä tai jotenki laajentaa tätä koneen antamaa palautetta. H4

4.2 Analyyttisen ja holistisen palautteen hyödyllisyys

Automaattinen analyttinen palaute koettiin pääosin hyödylliseksi ja selkeäksi, joskin palautteen kannustavuus mietitytti haastateltavia (lainaukset 21 ja 22). Opettajien oli kuitenkin jossain määrin vaikea ottaa kantaa siihen, missä määrin puheesta annettavat

**135 MITEN SUOMEN JA RUOTSIN OPETTAJAT KÄYTTÄISIVÄT PUHEEN
AUTOMAATTISEEN ARVIOINTIIN KEHITETTYÄ TYÖKALUA?**

analyttiset arviot vastaavat opetusta tai heidän käsitystään suullisen kielitaidosta. Haastateltavat vaikuttivat mieltävän puhumisen osa-alueet eri tavalla. Esimerkiksi lainauksessa 24 opettaja kuvaa sujuvuutta enemmän ymmärrettävyyden kuin puheen taukoisuuden (lainaus 23) kannalta. Lisäksi kuten H1 kertoo, eritasoisten puhujien kohdalla saatetaan kiinnittää huomiota erilaisiin asioihin (lainaus 24). Puheen näyttämistä tekstinä pidettiin hyödyllisenä oppimisen kannalta (lainaus 25).

- (21) – – lähes jokainen opiskelija haluaa tietää yksityiskohdista eli mikä onnistui ja mikä ei onnistunut – – voiko koneelle opettaa että se poimisi sieltä näitä onnistuneimpia kohtia? H4
- (22) – – joku opiskelija ehkä vois – – miettiä, et oonks mä todellakin näin huono vai ja miettiä tätä palautetta. Et silloin voi olla, että opettajan palaute tulee tarpeen. H1
- (23) Mun mielestä niiden [analyttisten palautteiden] ei tarvitse olla pitempiä. Mut se on näitä keskeisiä asioita joita [sujuvuudessa] arvioidaan, et se on sujuvaa, vaivatonta, ei häiritseviä taukoja, ei empimistä. H4
- (24) – – oliko se sujuvaa, – – ymmärsikö kohderyhmä mistä sä puhuit, menikö sanoma perille, sehän on tietysti sitä juuri, et se tehtävänanto ja se sujuvuus, ja tietysti sitten kun et puhutaan, no se laajuus on tietysti aina mukana jossain määrin, et jos mä vaadin tietysti enemmän sitten just esimerkiksi mofeilta eli kakskielisiltä. H1
- (25) Opiskelija itsekin voi [litteraatista] katsoa, mitä mä halusin sanoa ja mitä tästä tuli. H4

Keskustelu holistisesta taitotasoarviosta toi esiin ristiriidan: kielten opetussuunnitelmat (Opetushallitus 2003, 2019) perustuvat taitotasoihin, mutta opettajat eivät vielä hyödynnä niitä opetuksessa tai arvioinnissa (lainaukset 26–28). Sen sijaan opettajat kertovat hyödyntävänsä numeerisia asteikkoja, joten opiskelijat eivät välttämättä tunne taitotasoasteikkoa entuudestaan.

- (26) Tarkotus olis [käyttää taitotasoasteikkoa opetuksessa], jos katsotaan nyt opetussuunnitelmaa taas [naurua] niin siinä on ne päämäärät. Mut mun täytyy kyl myöntää, et en mä hirveesti niitä kato. H1
- (27) Kyllähän ne aina taitotasot esitellään ja etenki tuolla toisella asteella lukiopuolella – – päättöarviointia [tehdessä] eli yhdeksännen luokan päättöarvioinnin kriteerit jotka nyt sit pohjautuu taitotasoihin – – mä kouluarvosanoin edelleen ja näin arviointi koheet ja yleisesti, et en mä nyt taitotasoja tosiaan käytä sitten ku tietyissä tilanteissa vaan. H2

- (28) Ei, opiskelijat ei tiedä näitä [taitotasoja], [eurooppalaisesta] viitekehuksestähän nää tulee, eikö tuukin? Niitä ei kyllä lukiomaailmassa käytetä missään, voi sanoo näin. – Et meillähän on sit numeerinen arvostelu ja IB:llähän on siis ykkösestä seiskaan, siel on semmonen asteikko. H3

Lukion suullisen kielitaidon kokeen uudistuessa (Opetushallitus 2022) opettajien kuitenkin tulee antaa numeroarvosanan lisäksi taitotasoarvio, joten DigiTalan työkalu voisi konkretisoida taitotasoja niin opiskelijoille kuin opettajille (lainaukset 29 ja 30).

- (29) – – nythän [lukion suulliseen kokeeseen] tuli semmonen lisäys, tähän mennessä opiskelija on saanu arvosanan kasin, ysin, ihan tämmösen kouluarvosanan, nyt meidän pitää vielä juuri lisätä siihen vielä juuri tää taitotasoarvio. Eli tulee se kasi ja sitten vielä taitotasoarvio siihen. – – Silloin tää olis todella hyvä apuväline juuri siinä suullisessa. H1
- (30) – – tää olis opettajan työn tuki sitten, arvioinnin tuki, varsinkin sit jos ei oo itsellä kovasti arviointikokemusta – – H2

Haastateltavat eivät kokeneet ongelmallisena taitotasojen näyttämistä päätasolla (B1) välietappien (B1.1) sijaan. He tunnistivat, että välietappien määrittäminen on ihmisarvioijillekin haastavaa, ja toivat esiin opettajien mahdollisuuden täydentää automaattista arviota (lainaukset 31 ja 32).

- (31) No se voi olla ihan hyvä [ettei taitotasoiissa näytetä välietappeja kuten B1.1], koska joskus kun on kattonu noita [taitotasoja], niin vähän miettii itsekin, et [naurua] mikä sitten on se ero. – – nää taitotasoarviot voi olla aika vaikee monesti. H1
- (32) – – tässä voi aatella sitten varmaan että se opettaja määrittelee sen alatasen. Että jos tän on tarkoitus tukea sitä opettajan arviota päätason osalta niin voihan sitä sitten täydennellä. H2

5 Pohdinta

Tämä soveltava tutkimus käsittelee kieltenopettajien näkemyksiä suomen ja ruotsin oppijoille kehitetystä työkalusta, joka antaa oppijalle automaattisesti palautetta hänen puhe-suorituksestaan. Automaattista arviointia ja palautetta hyödynnetään enenevässä määrin etenkin matemaattis-luonnontieteellisten ja teknisten alojen – mutta myös kielten – opetuksessa (Deeva ym. 2021; Golonka ym. 2014). Tutkimusalue on kuitenkin uusi, ja valtaosa puhumisen harjoitteluun suunnitelluista työkaluista rajoittuu ääntä-

137 MITEN SUOMEN JA RUOTSIN OPETTAJAT KÄYTTÄISIVÄT PUHEEN
AUTOMAATTISEEN ARVIOINTIIN KEHITETTYÄ TYÖKALUA?

misestä annettavaan palautteeseen (ks. kuitenkin Gu & Davis 2020, Pearson Education 2018, 2022; de Vries ym. 2014; Xu ym. 2020).

Suomessa automatiikkaa hyödynnetään kielten opetuksessa ja kielitaidon arvioinnissa toistaiseksi hyvin vähän, joten opettajilta kerätyt näkemykset ovat keskeinen osa työkalun kehitys- ja validointityötä. Automaattisen palautteen kehittelyn lisäksi tutkimustulokset hyödyttävät laajemmin kielten opetuksen, kielitaidon arvioinnin ja kieliteknologian parissa työskenteleviä kuten opettajia, opettajien kouluttajia sekä oppimispelien, oppimateriaalien ja kielikokeiden laatijoita.

Tutkimuksessa pyrimme Xin (2010: 296) validiusmalliin nojaten selvittämään automaattisen palautteen vaikutuksia opetukseen ja oppimiseen kartoittamalla työkalun käyttötarkoituksia (tutkimuskysymys 1) sekä automaattisen palautteen hyödyllisyyttä (tutkimuskysymys 2) opettajien näkökulmasta. Tutkimus toi esiin samankaltaisia näkökulmia, joita olemme havainneet muiden käyttäjäryhmien osalta (lukiolaiset ks. von Zansen ym. 2022b, ihmisarvioijat ks. von Zansen ym. 2022c, akateemiset aikuisopiskelijat ks. von Zansen & Hildén 2022). Opettajien näkemät hyödyt koskivat työkalun tarjoamaa lisäresurssia: kouluasteesta riippumatta opiskelijat saavat enemmän myös itsenäisiä harjoittelumahdollisuuksia ja palautetta puhesuorituksestaan (ks. Gu & Davies 2020). Opettajat tiedostivat, että opiskelijoiden olisi hyvä saada enemmän palautetta suullisesta kielitaidostaan, mutta opintojaksojen aikataulujen ja opetusryhmien koon takia opettajat eivät ehdi antaa riittävästi henkilökohtaista palautetta. Näin ollen opettajat hyödyntäisivät työkalua monipuolisesti opintojaksojen aikana esimerkiksi kotitehtävissä ja kokeissa.

Haastateltujen lukio-opettajien mukaan Moodle-työkalusta saatava jatkuva palaute auttaisi opiskelijoita arvioimaan omaa kielitaitoaan sekä sen kehittämiskohteita esimerkiksi osana kieliprofilia, joka otettiin lukioissa käyttöön uuden opetussuunnitelman (Opetushallitus 2019) myötä. Lisäksi työkalu konkretisoisi taitotasoa sekä opiskelijoille että opettajille, mikä on ajankohtaista ja tarpeellista, sillä lukion suullisen kielitaidon kokeen uudistuessa (Opetushallitus 2022) opiskelijat saavat numeroarvosanan lisäksi taitotasoarvion. Tällä hetkellä taitotasojen hyödyntäminen kielten opetuksessa on käytännössä vähäistä, vaikka Suomessa kielten opetussuunnitelma (Opetushallitus 2019) rakentuu Eurooppalaisesta viitekehuksesta (Euroopan neuvosto 2001) peräisin olevalle taitotasoajattelulle.

Taitotasoasteikon sijaan haastateltavat opettajat käyttivät erilaisia asteikkoja ja kriteerejä suullisen kielitaidon arvioimiseen, joista valtaosa oli numeerisia, kuten kouluarvosana. Opettajat vaikuttivat myös pitävän keskenään erilaisia puheen osa-alueita tärkeinä, ja heidän käyttämänsä kriteerit saattoivat vaihdella puhujien taitotason ja tehtävän perusteella. Puheesta automaattisesti mitattavissa olevien ominaisuuksien (Evanini & Zechner 2020; Xi 2010) lisäksi opettajat kertoivat puhetehtäviä arvioidessaan kiinnittävänsä huomiota kehonkieleen ja luovuuteen sekä antavansa myös näistä palautetta opiskelijoille.

Haastattelujen perusteella emme täysin pysty arvioimaan, kuinka hyvin DigiTala-työkalun analyttinen palaute vastaa opetuksessa annettavaa palautetta. Toisaalta opettajat kertoivat pitävänsä DigiTala-työkalussa näkemäänsä palautetta kattavana. Toisaalta havaitsimme, että he ymmärsivät arvioitavan käsitteen (esim. sujuvuus) keskenään eri tavalla, ja paikoitellen myös eri tavalla kuin se määriteltiin automaattisessa palautteessa.

Käsitteelliset erot ovat luonnollisia useastakin syystä. Esimerkiksi kansallisissa oppimistulosten arvioinneissa on havaittu, että arviointikriteereitä tulkitaan jo peruskouluissa eri tavalla yhteneväisistä päättöarvioinnin kriteereistä huolimatta (Kieliverkosto 2022). Tätä tutkimusta varten haastattelemamme kielten opettajat työskentelevät erilaisissa kouluissa, ja suomalaista kielikasvatusta ohjaa pedagoginen vapaus. Lisäksi työkalun palaute on peräisin ihmisarvioijien käyttämistä arviointikriteereistä (von Zansen 2022), jotka puolestaan on laadittu aiemman opetussuunnitelman (Opetushallitus 2003) perusteella. Aiempi opetussuunnitelma sopii paremmin tutkimushankkeen tekniisiin ja analyttisiin lähtökohtiin, sillä automaattinen arviointi rajoittuu puheesta mitattavissa oleviin ominaisuuksiin (Evanini & Zechner 2020; Xi 2010). Opettajat eivät kuitenkaan hyödynnä aiempien tai nykyisten opetussuunnitelmien (Opetushallitus 2003, 2019) taitotasoaasteikkoja suullisen kielitaidon arvioinnissa. Mikäli opettajien käsitystä suullisesta kielitaidosta halutaan yhtenäistää, ehdotamme ratkaisuksi kansallisten puhumisen arviointikriteerien laatimista sekä opettajien täydennyskoulutusta.

Opettajien näkemät rajoitteet työkalun käytössä koskivat arvioinnin oikeudenmukaisuutta ja paikkansapitävyyttä sekä opiskelijoiden mahdollista kriittistä suhtautumista tekniikkaan, mikä on linjassa Xun ym. tutkimuksen (2020) lisäksi aiempien käyttäjäryhmiä koskevien tutkimustemme (von Zansen ym. 2022b, von Zansen ym. 2022c, von Zansen & Hildén 2022) kanssa. Yhtenä uutena uhkana opettajat näkevät opiskelijoiden kyvyn vastaanottaa automaattista palautetta: he pelkäsivät liian suoran palautteen voivan lannistaa kielen oppijaa. Opettajalla nähtiinkin olevan tärkeä rooli myös työkalun antaman palautteen välittämisessä.

Tämän soveltavan tutkimuksen avulla onnistuttiin tuottamaan lisää tietoa puheen automaattisesta arvioinnista, ennustamaan sen mahdollisia vaikutuksia kielten opetuksessa sekä esittämään toimintasuosituksia. Harkinnanvarainen näyte sopi otantamenetelmäksi tähän pilottivaiheen haastattelututkimukseen. Tutkimuksen rajoitukset koskevat tulosten yleistettävyyttä, sillä haastatelimme neljää hankkeeseen aiemmin osallistunutta opettajaa. Aiommekin täydentää opettajien käsitysten tutkimista laajemmalla kyselytutkimuksella, jonka rajaamista tämä tutkimus tukee. Yksi jatkotutkimusmahdollisuus liittyy arviointityökalun erilaisten palauteversioiden kehittämiseen ja kokeilemiseen oppijoiden ja opettajien kanssa.

Yhteenvetona voidaan todeta, että haastateltavat käyttäisivät työkalua jo nyt tukemaan työtään ja täydentämään nykyisiä opetus- ja arviointimenetelmiä, vaikka tiedostivat työkaluun liittyviä kehitystarpeita. Tulosten perusteella automaattisen palautteen tuomat hyödyt ovat merkittäviä oppimisen ja opetuksen kannalta. Lisäksi

hankkeen laatimien arviointikriteerien (von Zansen 2022) pohjalta tuotetut palauteväittämät olivat kieltenopettajien mielestä ymmärrettäviä ja hyödyllisiä, vaikka ne eivät kata kaikkia puheen ulottuvuuksia. Kehitys- ja tutkimustyötä jatketaan yhteistyössä eri toimijoiden kanssa – haastateltavat toivoivat tekoälyn muun muassa edistävän suullisen kielitaidon mittaamista ylioppilastutkinnon kielikokeissa (Vaarala ym. 2021).

Kiitokset

Suomen Akatemia rahoittaa DigiTala-tutkimushanketta (2019–2023). Kiitämme yhteistyöstä projektin tutkijoita Helsingin yliopistosta (rahoituspäätös 322619), Aalto-yliopistosta (rahoituspäätös 322625) ja Jyväskylän yliopistosta (rahoituspäätös 322965).

Kirjallisuus

- Alanen, T., J. Erkkilä, T. Harjunpää & M. Heijala 2022. Digitala Moodle plugin user manual (1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.6535377>
- Al-Ghezi, R., Y. Getman, A. Rouhe, R. Hildén & M. Kurimo 2021. Self-Supervised End-to-End ASR for Low Resource L2 Swedish. Proc. *Interspeech 2021*, 1429–1433. <https://doi.org/10.21437/Interspeech.2021-1710>
- Al-Ghezi, R., K. Voskoboinik, Y. Getman, A. von Zansen, H. Kallio, C. Akiki, M. Kuronen, A. Huhta & R. Hilden (arvioitavana). Automatic speaking assessment of Spontaneous L2 Finnish and Swedish.
- Deeva, G., D. Bogdanova, E. Serral, M. Snoeck & J. De Weerd 2021. A review of automated feedback systems for learners: Classification framework, challenges and opportunities. *Computers & Education*, 162. <https://doi.org/10.1016/j.compedu.2020.104094>
- Euroopan neuvosto (2001). Common European framework of reference for languages: Learning, teaching, assessment. Council of Europe: Cambridge University Press.
- Evanini, K., & K. Zechner 2020. Overview of automated speech scoring. Teoksessa K. Zechner & K. Evanini (toim.) *Automated Speaking Assessment: Using language technologies to score spontaneous speech*. New York: Routledge, Taylor & Francis Group, 3–20.
- Golonka, E. M., A. R. Bowles, V. M. Frank, D. L. Richardson & S. Freynik 2014. Technologies for foreign language learning: A review of technology types and their effectiveness. *Computer assisted language learning*, 27 (1), 70–105. <https://doi.org/10.1080/09588221.2012.700315>
- Gu, L., & L. Davis. 2020. Providing SpeechRater feature performance as feedback on spoken responses. Teoksessa K. Zechner & K. Evanini (toim.) *Automated Speaking Assessment: Using language technologies to score spontaneous speech*. New York: Routledge, Taylor & Francis Group, 159–175.
- Heijala, M., & A. von Zansen 2022. DigiTala’s teacher interview frame for teacher interviews (spring 2022). Zenodo. <https://doi.org/10.5281/zenodo.7082885>

- Hyvärinen, M., E. Suoninen & J. Vuori (n.d.). Haastattelut. Teoksessa J. Vuori (toim.) *Laadullisen tutkimuksen verkkokäsikirja*. Tampere: Yhteiskuntatieteellinen tietoaarkisto. <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus/kvali/laadullisen-tutkimuksen-aineistot/haastattelut/> [luettu 20.06.2020]
- Kieliverkosto 2022. *Tiedote: Yhdeksäsluokkalaisten englannin taitoja arvioitiin keskellä koronaa – osaaminen jäi kouluarvosanaan 7*. <https://www.kieliverkosto.fi/fi/kuulumisia-kentalta/news/tiedote-yhdeksäsluokkalaisten-englannin-taitoja-arvioitiin-keskella-koronaa-2013-osaaminen-jai-kouluarvosanaan-7> [luettu 14.09.2022]
- Kallio, H., R. Suviranta, A. von Zansen & M. Kuronen 2022. Creaky voice and fluency measures in predicting perceived fluency and oral proficiency of spontaneous L2 Finnish. *Proceedings of Speech Prosody 2022*, 777–781. <https://doi.org/10.21437/SpeechProsody.2022-158>
- Kallio, H., M. Kautonen, & M. Kuronen (arvioitavana). Prosody and fluency of Finland Swedish as a second language: investigating global parameters for automated speaking assessment.
- 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>
- Opetushallitus 2003. *Lukion opetussuunnitelman perusteet*. Helsinki: Opetushallitus.
- Opetushallitus 2019. *Lukion opetussuunnitelman perusteet*. Helsinki: Opetushallitus.
- Opetushallitus 2022. *Suullisen kielitaidon arviointiin ja todistuskäytänteisiin muutoksia – koe mahdollista suorittaa myös suullisilla näytöillä*. <https://www.oph.fi/fi/uutiset/2022/suullisen-kielitaidon-arviointiin-ja-todistuskaytanteisiin-muutoksia-koe-mahdollista> [luettu 14.09.2022]
- Ovaska, S., A. Aula & P. Majoranta 2005. *Käytettävyystudkimuksen menetelmät*. Tampere: Tietojenkäsittelytieteen laitos, Tampereen yliopisto.
- Pearson Education 2018. Versant™ Arabic Test. Test Description and Validation Summary. <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/english/SupportingDocs/Versant/ValidationSummary/Versant-Arabic-Test-Description-Validation-Summary.pdf>
- Pearson Education 2022. Versant™ English Test. Test Description and Validation Summary. <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/english/SupportingDocs/Versant/ValidationSummary/Versant-English-Test-Description-Validation-Report.pdf>
- Tietoaarkisto 2022. *Kvalitatiivisen datan käsittely*. <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/kvalitatiivisen-datan-kasittely/> [luettu 7.6.2022]
- Vaarala, H., S. Riuttanen, E. Kyckling & S. Karppinen 2021. *Kielivaranto. Nyt!: Monikielisyyden vahvuudeksi -selvityksen (2017) seuranta*. Jyväskylä: Soveltavan kielen tutkimuksen keskus, Jyväskylän yliopisto. https://www.jyu.fi/hytk/fi/laitokset/solki/tutkimus/julkaisut/pdf-julkaisut/kielivaranto-nyt-_julkaisu_sivuittain-1.pdf
- de Vries, B. P., C. Cucchiari, S. Bodnar, H. Strik & R. van Hout 2015. Spoken grammar practice and feedback in an ASR-based CALL system. *Computer Assisted Language Learning*, 28 (6), 550–576. <https://doi.org/10.1080/09588221.2014.889713>
- Vuori, J. n.d. Laadullinen sisällönanalyysi. Teoksessa J. Vuori (toim.) *Laadullisen tutkimuksen verkkokäsikirja*. Tampere: Yhteiskuntatieteellinen tietoaarkisto. <https://www.fsd.tuni.fi/fi/palvelut/menetelmaopetus> [luettu 17.03.2022]
- Xi, X. (2010). Automated scoring and feedback systems: Where are we and where are we heading?. *Language Testing*, 27 (3), 291–300. <https://doi.org/10.1177/0265532210364643>

- Xi, X., & Y. Sawaki 2017. Methods of test validation. Teoksessa E. Shohamy, I. Or & S. May (toim.) *Language Testing and Assessment*. Cham, Switzerland: Springer, 193–209. <http://doi.org/10.1007/978-3-319-02261-1>
- Xu, J., M. Brenchley, E. Jones, A. Pinnington, T. Benjamin, K. Knill, G. Seal-Coon & A. Geranpayeh 2020. *Linguaskill Building a validity argument for the Speaking test*. Cambridge: Cambridge Assessment English. <https://www.cambridgeenglish.org/Images/589637-linguaskill-building-a-validity-argument-for-the-speaking-test.pdf>
- Yhteiskuntatieteellinen tietokirasto. (n.d.). Tunnisteellisuus ja anonymisointi. <https://www.fsd.tuni.fi/fi/palvelut/aineistonhallinta/tunnisteellisuus-ja-anonymisointi/#minimointi-eli-miten-kerata-aineisto-niin-ettei-se-sisalla-turhia-tunnisteita> [luettu 20.06.2022]
- von Zansen, A. 2022. *DigiTala's rating criteria: Holistic and analytic scales for assessing L2 speaking*. Zenodo. <https://doi.org/10.5281/zenodo.6477089>
- von Zansen, A., T. Alanen, R. Al-Ghezi, J. Erkkilä, T. Harjunpää, M. Heijala & H. Kallio 2022a. DigiTala Moodle plugin. https://github.com/aalto-speech/moodle-mod_digitala
- von Zansen, A., R. Hildén & M. Sneck 2022b. Lukiolaisten käsitykset ja heidän antamansa palaute suullisen kielitaidon automaattisesta arvioinnista: Ainedidaktinen symposium 2021. Teoksessa R. Kantelinen, M. Kautonen & Z. Elgundi (toim.) *LINGUAPEDA 2021*. Konferenssijulkaisu. Ainedidaktisia tutkimuksia 21. Itä-Suomen yliopisto, Filosofinen tiedekunta, 176–205. <http://hdl.handle.net/10138/352128>
- von Zansen, A. & R. Hildén 2022. "It was cool and comfortable!" Akateemisten alkeistason S2-opiskelijoiden kokemuksia tietokoneella suoritettavasta puhumisen kokeesta. Routarinne, S., P. Heinonen, T. Kärki, A. Roiha, M-L. Rönkkö & A. Korkeaniemi (toim.). *Ainedidaktiikka ajassa: Laajenevat oppimisympäristöt ja eri-ikäiset oppijat*. Rauma: Suomen Ainedidaktinen tutkimusseura. Turun yliopiston opettajankoulutuslaitos, 72–90.
- von Zansen, A., H. Kallio, M. Sneck, M. Kuronen, A. Huhta & R. Hildén 2022c. Ihmisarvioijien näkemyksiä suullisen kielitaidon automaattisesta arvioinnista, digitaalisesta arviointiprosessista sekä puhe-suorituksista arvioitavista ulottuvuuksista. Teoksessa T. Seppälä, S. Lesonen, P. Ikkänen & S. D'hondt (toim.) *AFinLAN vuosikirja*.
- von Zansen, A. & A. Huhta 2022. Developing Automated Feedback on Spoken Performance: Exploring the Functioning of Five Analytic Rating Scales Using Many-facet Rasch Measurement. Teoksessa J. H. Jantunen, J. Kalja-Voima, M. Laukkarinen, A. Puupponen, M. Salonen, T. Saresma, J. Tarvainen & S. Ylönen (toim.) *Diversity of Methods and Materials in Digital Human Sciences: Proceedings of the Digital Research Data and Human Sciences DRDHum Conference 2022, December 1-3, Jyväskylä, Finland*, p. 211–229, University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-9450-1>

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 142–161.

Elli Saari & Raili Hilden

Helsingin yliopisto

S2-opettajien käsityksiä formatiivisesta arvioinnista oppimisen tukena

Nostot

- S2-opettajilla on tietoa formatiivisen arvioinnin periaatteista, mutta arvioinnin tarkoitus jää usein epäselväksi.
- Formativisen arvioinnin menetelmiä käytetään, vaikka niitä ei välttämättä koeta hyödyllisinä.
- Opettajat eivät aina miellä formatiivisen arvioinnin keinoja tehokkaiksi opiskelutaitojen edistämiseksi.
- Lisäkoulutusta tarvitaan formatiivisen arvioinnin perusteista, mahdollisuuksista ja integroimisesta muuhun opetustyöhön.

Abstract

Assessment is an important tool in the learning process. This research was aimed to study Finnish as a second language (F2) teachers' perceptions on formative assessment (FA). The data (n=208) was gathered from F2-teachers in 2015. The results show that F2-teachers use many good practices of FA in every-day work. However, many fail to find it effective and meaningful. Perceptions on assessment are quite evenly distributed in the three dimensions of assessment literacy. However, the insight on assessment is thinner in the dimensions of conceptual and praxeological knowledge compared to the socio-emotional dimension. Some assessment components align with multiple dimensions. The core concepts of FA are well known, but seldom considered advantageous. Thus, FA guidelines are not applied as intended. Further teacher education aimed for both clarifying the purpose and good practices of FA should be implemented. More research is needed especially on the role of feedback in F2-context.

Keywords: assessment, Finnish as a second language, basic education

Asiasanat: arviointitutkimus, suomi toisena kielenä, perusopetus

1 Johdanto

Suomea toisena kielenä ja kirjallisuutta (S2) opiskelevien oppilaiden oppimistulokset ovat perusopetuksen päättövaiheessa keskimäärin heikompia kuin suomea äidinkielenä (S1) puhuvien, ja perusopetuksen jälkeen S2-oppilaat ovatkin hakeutuneet muuta väestöä harvemmin toisen ja korkeamman asteen opintoihin (Airas ym. 2019; Jahnukainen ym. 2019; OKM 2022). Vaikka S2-oppilaiden oppimismotivaatio on poikkeuksellisen korkea, motivaation yhteys oppimistuloksiin on olematon, kun taas perusväestön oppilailta tämä korrelaatio on vahva (Harju-Luukkainen ym. 2014; Leino ym. 2019).

S2-oppilaat ovat heterogeeninen ryhmä, jonka oppimiseen vaikuttavat monet taustatekijät, kuten maahanmuuttoikä, koulunkäyntihistoria, lähtökielen vaikutus suomen kielen omaksumiseen ja vanhempien sosioekonominen status (Harju-Luukkainen ym. 2016). Koulu monikielistyy, mikä pakottaa muuttamaan näkemyksiä siitä, mitä on kielen osaaminen (Ahlholm 2020; Piippo 2021). Opettaja joutuu arvioimaan, miten eri kielitaidon tasoilla olevat oppilaat pystyvät osallistumaan opetukseen ja käsittelemään erilaisia tekstejä. Heterogeenisessä ryhmässä kaivataan aiempaa monipuolisempia pedagogisia käytänteitä, joilla voidaan tukea oppilaita yksilöllisesti.

Arviointitutkimus on perusteltu lähtökohta S2-oppilaiden oppimispolkujen parantamiseksi. Arviointi on yksi opettajan työn keskeisimmistä ja vaikuttavimmista pedagogisista keinoista ohjata oppimista, ja hyvin toteutettuna se auttaa oppilasta parantamaan oppimistuloksiaan (Andrade ym. 2019). Black ja Wiliam (1998b, 1998a) osoittavat laajassa kirjallisuuskatsauksessaan formatiivisen arvioinnin parantavan oppilaan sitoutumista ja oppimista. Oppilaan vasta kehittyvän suomen kielen taidon vuoksi opettajan arviointi saattaa kuitenkin kohdistua epätarkasti. Opettajan voi olla vaikea erottaa kehittyvän kielitaidon vaikutusta sisällön osaamisen puutteisiin tai nähdä osaamista kielellisten puutteiden läpi. On mahdollista, että arviointi kohdistuu

tavoitteiden kannalta epäolennaisiin asioihin, jolloin hukataan arvioinnin positiiviset vaikutukset oppimisen ohjaamisen näkökulmasta. Myös perusopetuksen opetussuunnitelman perusteet (OPH 2014) ohjaavat opetusta ja arviointia, ja nykyiset perusteet korostavat erityisesti formatiivista arviointia.

Opettajilla on taipumus omaksua arviointikäytänteitä, jotka ovat yhdenmukaisia heidän senhetkisten käsitystensä ja aiempien kokemustensa kanssa, ja toisaalta hylätä ne, jotka ovat omien käsitysten tai uskomusten kanssa ristiriidassa (Xu & Brown 2016: 156). Käsitukset voivat olla yhtä aikaa pysyviä ja muuttuvia, yksilöllisiä ja jaettuja, rationaalisia ja emotionaalisia sekä samansuuntaisia ja ristiriitaisia (Barcelos & Kalaja 2011; Kalaja ym. 2015). Käsituksiin vaikuttavat sekä tietopohja että henkilökohtaiset kokemukset, asenteet ja mielikuvat, mutta myös ympäröivän yhteisön ajattelu; ne voivat muuttua esimerkiksi opetussuunnitelmien tai koulun arviointikulttuurin muuttuessa. Opetuksen käytänteet puolestaan vaikuttavat oppilaiden opiskelustrategisiin päätöksiin (Cohen 2013). Oppilaat oppivat nopeasti, millaista osaamista kokeissa ja tehtävissä arvostetaan, millaista työskentelyä edellytetään, millaista palautetta on tarjolla ja kuinka virheisiin suhtaudutaan. Tämän perusteella he valitsevat, kuinka syvällisesti tehtävään kannattaa omistautua. Käsitusten ja toiminnan välillä voi olla myös ristiriita niin, että käsitukset eivät aina toteudu tai näy käytänteissä. Arviointiosaamisen kehittämisessä on siksi olennaista tunnistaa arviointikäytänteiden taustalla vaikuttavia käsityksiä.

Tässä artikkelissa selvitetään, miten S2-opettajat hahmottavat formatiivisen arvioinnin eri tapojen merkityksen oppimisen edistämässä ja mitkä ovat opettajien arviointiosaamisen vahvuudet ja mahdolliset täydennyskoulutustarpeet. Aineisto koostuu S2-opettajille suunnatun kyselyn vastauksista perusopetuksen päättövaiheessa, jolloin heidän oppilaansa osallistuivat S2-oppilaiden oppimistulosten kansalliseen arviointiin (Kuukka & Metsämuuronen 2016). Termillä S2-oppilas viitataan oppilaisiin, joille suomen kieli ei ole äidinkieli ja jotka opiskelevat peruskoulussa S2-oppimäärää. Termillä S2-opettaja viitataan äidinkielen ja kirjallisuuden aineenopettajiin, mutta myös muihin perusopetuksen yläluokkien opettajiin, kuten erityisopettajiin, jotka oman työnsä ohella vastaavat myös S2-opetuksesta. Opettajien käsitykset (*perceptions*) nähdään dynaamisina ja kontekstisidonnaisina. Artikkelin tutkimuskysymykset ovat:

1. Mitä käsityksiä S2-opettajilla on formatiivisen arvioinnin keinoista oppimisen edistämässä?
2. Mitä arviointiosaamisen ulottuvuuksia S2-opettajien käsitykset edustavat?

1.1 Formatiivinen arviointi

Formatiivisen arvioinnin tavoitteena on edistää oppimista ja muokata opetusta oppilaan henkilökohtaiset oppimistarpeet huomioiden (Black ym. 2004; Brookhart 2011; Cizek ym. 2019). Keskeistä on tunnistaa, missä oppilas on suhteessa oppimistavoitteisiin, mihin oppilas on menossa ja miten hän voi saavuttaa tavoitteen (William

& Thompson 2007: 14–15). Toimintaa ohjaavat kannustavuus ja osaamisen näkyväksi tekeminen. Käytänteiden tasolla oppilaalle avataan oppimistavoitteet ja arviointikriteerit, oppimistilanteet ja osaamisen näyttö suunnitellaan tehokkaiksi, oppilaalle annetaan eteenpäin ohjaavaa palautetta ja oppilaita aktivoidaan toistensa oppimisen tukemiseen sekä oman oppimisen omistajuuteen (Black & Wiliam 2009).

Oppimistavoitteiden ja arviointikriteerien avaamiseksi opettajalla on oltava selkeä käsitys arvioinnin kohteesta. Oppimistilanteet ja osaamisen näyttö suunnitellaan siten, että niistä saatava tieto kertoo tarkemmin, missä oppilas tarvitsee enemmän ohjausta (Wiliam 2011). Palaute ohjaa oppilasta tehokkaimmin eteenpäin, kun se kohdistuu tehtävätasolle ja päätelmät perustuvat oppimisprosessiin (Hattie & Timperley 2007; Hattie & Yates 2014). Hyvässä palautteessa otetaan huomioon myös oppilaan oppimistausta ja aiempi suoriutuminen (Cizek ym. 2019). Kokonaistilanteen perusteella voidaan arvioida, onko puutteellisessa vastauksessa kysymys huolimattomuudesta, yksittäisen tiedon puuttumisesta vai syvemmästä virheellisestä käsityksestä.

Palautteen antaminen on vuorovaikutteinen prosessi. Tavoitteena on, että oppilas itse etsii aktiivisesti palautetta, mutta myös opettaja saa uutta tietoa opetuksen suunnitteluun (Hattie & Timperley 2007). Opettajan vuorovaikutustyyliin vaikuttaa näkemys oppimisesta ja siitä, onko hän oppimisen kontrolloija vai ohjaaja. Oppilaan vuorovaikutustyyliin vaikuttavat uskomukset omasta oppimiskyvystä, kognitiiviset taidot ja oppimiskonteksti, mutta myös palautteen vastaanottamiseen vaikuttava asenne, joka voi olla joko passiivinen tai aktiivinen (Black & Wiliam 2009; Wiliam 2011). Lisäksi luokan ilmapiirillä on merkitystä palautteenantotilanteessa: virheistä ja negatiivisestakin palautteesta on hyötyä luokassa, jossa palautteen antaminen on tervetullutta ja sen hyödyntämiseen on totuttu (Hattie & Timperley 2007).

Oppimisen omistajuudella viitataan oppilaan kykyyn reflektoida ja rakentaa omaa osaamistaan. Oikein kohdennettu palaute sekä itse- ja vertaisarviointi arvioinnin osana tarjoavat oppilaalle keinoja säädellä oppimista sekä kehittää oppimaan oppimisen taitoja (Panadero ym. 2018).

Formatiivisen arvioinnin käytänteet ilmenevät eri oppiaineissa eri tavoin (Bennett 2011; Wiliam 2019), samoin erilaisille oppijoille eri tavoin. Toisen kielen opetukseen ei ole löydettävissä yhtä oikeaa arvioinnin mallia, vaan opettajaa palvelee parhaiten kuvaus arvioinnin ulottuvuuksista ja oppimista edistävästä strategioista (Turner & Purpura 2016). Toisen kielen kontekstissa vaikuttavat samaan aikaan useat, keskenään jopa ristiriitaiset ulottuvuudet, kuten oppilaan tausta ja oppimisen eri kontekstit, osaamisen näkyväksi tekemisen ja näytön keräämisen eri keinot. Lisäksi on huomioitava kehittyvä kielitaito, oppimiskäsitykset, kielitietoisuuden ja oppiaineen sisällön yhteensovittaminen sekä luokahuoneen vuorovaikutus.

1.2 Opettajan arviointiosaaminen

Opettajan työajasta kuluu jopa puolet erilaisen arvioinnin parissa, minkä vuoksi arvioinnin tärkeimmän tehtävän tulisi olla oppimisen tukeminen (Stiggins 1999: 23). Tutkimukset kuitenkin osoittavat, että opettajilla on puutteita arviointimenetelmien käytössä, eivätkä he osaa välttämättä hyödyntää arviointitietoa opetuksessaan (Atjonen 2017; Hilden & Fröjdendal 2018; Ruiz-Primo & Brookhart 2018: 4).

Viimeaikainen arviointitutkimus on painottanut arviointiosaamisen vuorovai-
kutteista luonnetta (Andrade ym. 2021; Coombs & DeLuca 2022). Xu ja Brown (2016) ovat havainnollistaneet, kuinka erilaiset kontekstit ja käsitykset muokkaavat opettajien arviointikäytänteiden valintaa. Vaikka opettaja tuntisi hyvin arvioinnin tietoperustan, toteuttamista ohjailevat virallisten määräysten lisäksi opettajan henkilökohtainen tai kollegiaalisesti jaettu tulkinta ja käsitykset arvioinnin luonteesta ja tarkoituksesta. Tulkintaan vaikuttavat sekä kognitiiviset että affektiiviset tekijät, jotka esimerkiksi muuttuvissa luokkatilanteissa saattavat olla keskenään ristiriitaisiakin. Pastore ja Andrade (2019) ovat rakentaneet aiempien tutkimusten synteessinä mallin, joka huomioi opetuksen eri kontekstit ja jakaa arviointiosaamisen käsitteelliseen (*conceptual*), käytännölliseen (*praxeological*) ja sosioemotionaaliseen (*socio-emotional*) ulottuvuuteen. Ulottuvuudet ovat osittain limittäisiä. Käsitteellinen ulottuvuus koskee arvioinnin periaatteisiin ja tarkoitukseen liittyvää tietoa, käytännöllinen ulottuvuus taitoa soveltaa tietoa toimintaan ja sosioemotionaalinen ulottuvuus taustalla vaikuttavia arvoja ja uskomuksia. Malli on riittävän väljä ja joustava ja siksi mielekäs lähtökohta, kun halutaan selvittää, miten S2-opettajien käsityksissä arviointiosaaminen ja luokahuonekäytännöt yhdistyvät.

2 Aineisto ja menetelmät

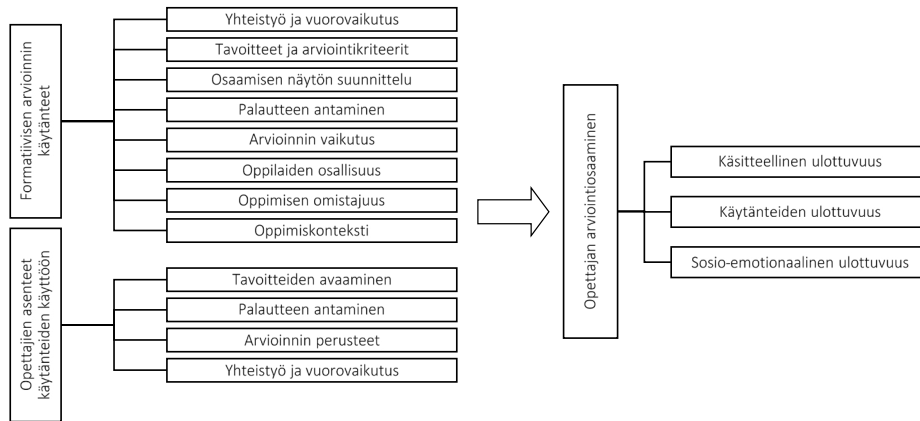
Aineistona (n=208) käytettiin perusopetuksen päättövaiheessa S2-oppilaiden oppimistulosten arvioinnin yhteydessä tehtyä opettajakyselyä, jonka Koulutuksen arviointikeskus (KARVI) suoritti vuonna 2015. Kyselyssä opettajia pyydettiin kuvaamaan oppilaan kielitaitoon, opetukseen, opetusjärjestelyihin, päättöarvioinnin prosessiin ja muihin näkökulmiin liittyviä tekijöitä, jotka edistävät tai vaikeuttavat hyvien oppimistulosten saavuttamista S2-oppimäärässä. Valtaosa opettajista oli äidinkielen ja kirjallisuuden aineenopettajia, kuudesosa vastaajista toimi S2-opetuksen rinnalla myös erityisopettajina tai muissa tehtävissä. Vastaajien taustatiedot kuvataan tarkemmin liitteessä 1 (taulukko 1).

Tässä tutkimuksessa hyödynnettiin sekä määrällisiä että laadullisia osia. Määrällinen aineisto koostui kyselyn strukturoidusta osasta, joka mittasi opettajien asenteita viisiportaisella Likert-asteikolla. Aineistossa formatiivista arviointia käsittelivät väittämät: *Kerron jakson/kurssin/oppitunnin alussa opetuksen tavoitteet (M1)*, *Annan oppilaille*

henkilökohtaista palautetta edistymisestä (M2) ja Selvitän oppilaille arvioinnin perusteita (M3). Määrälliseen aineistoon otettiin mukaan myös strukturoitu kysymys Missä määrin itsearviointi vaikuttaa päättöarviointiin (M4) sekä puolistrukturoidut vastaukset, jotka liittyivät yhteistyöhön ja vuorovaikutukseen (M5). Kysymysten M1-M4 Likert-asteikoiden vastaukset koodattiin lopuksi dikotomiseksi muuttujaksi, jossa vaihtoehdot usein/paljon tai hyvin usein/erittäin paljon merkitsivät tietoa siitä, että käytänne oli opetuksessa käytössä. Vastaavasti muut vaihtoehdot en koskaan/ ei lainkaan, vain harvoin/vähän tai joskus/ jonkin verran ilmaisivat, että arviointikäytänne ei ollut käytössä. Kysymyksestä M5 laskettiin eri vastausvaihtoehtojen jakauma. Laadullinen aineisto koostui vastauksista kahteen avoimeen kysymykseen: Mitkä asiat edistävät oppimista? (L1, n=160) Mitkä asiat vaikeuttavat oppimista? (L2, n=158).

Analyysi käynnistyi poimimalla vastaajat, jotka kysymysten M1-M5 perusteella toteuttivat opetuksessaan formatiivisen arvioinnin keinoja. Tämän jälkeen heidän avovastauksistaan (L1 ja L2) haettiin viittauksia formatiivisen arvioinnin keinoihin temaattisen analyysin menetelmällä (Eskola & Suoranta 1998). Keinot määriteltiin aiempien tutkimusten (William & Thompson 2007; Black & William 2009; Andrade & Heritage 2017; Deane & Sparks 2019) perusteella: tavoitteiden ja arviointikriteerien avaaminen, osaamisen näytön suunnittelu, palautteen antaminen, yhteistyö ja vuorovaikutus, oppilaiden osallisuus, oman oppimisen omistajuus, oppimiskontekstin ja arvioinnin vaikutusten huomioiminen. Temaattisen ryhmittelyn jälkeen avovastaukset luettiin uudelleen sen selvittämiseksi, miten opettajat omin sanoin kuvasivat keinojen hyötyjä, haittoja ja käytännön toteuttamistapoja.

Viimeisessä vaiheessa katsottiin rinnakkain saman aihealueen määrällisiä ja laadullisia vastauksia. Tarkoituksena oli tutkia, kuinka paljon eri keinoja käytetään ja selvittää opettajien käsityksiä käytön perusteista. Erityisesti haettiin tietoa siitä, kokevatko opettajat formatiivisen arvioinnin edistävän oppimista. Lopuksi opettajien arviointikäsitteet sijoitettiin Pastoren ja Andraden (2019) opettajan arviointiosaamisen mallin ulottuvuuksille. Analyysiprosessin kulku esitetään kuviossa 1.



KUVIO 1. Aineiston analyysiprosessi. Käsittely alkoi aineiston laadullisen osion lukemisesta formatiivisen arvioinnin käytännöiden (Andrade & Heritage 2017; Black & Wiliam 2009; Deane & Sparks 2019; Wiliam & Thompson 2007) sekä aineiston strukturoidun osan opettajan asenteiden näkökulmasta. Saatuja tuloksia verrattiin opettajan arviointiosaamisen eri ulottuvuuksiin (Pastore & Andrade 2019).

3 Tulokset

3.1 Opettajien käsitykset formatiivisen arvioinnin keinoista oppimisen edistämisessä

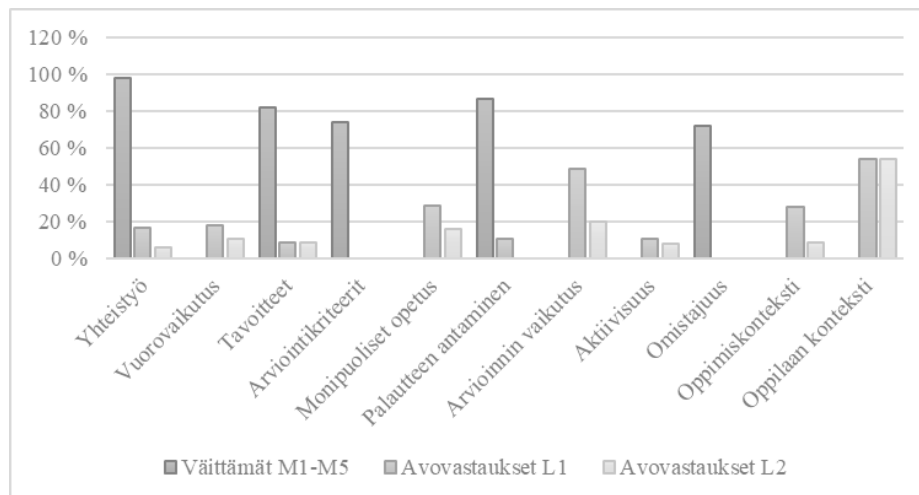
Tulosten perusteella S2-opettajat käyttävät työssään paljon formatiivisen arvioinnin keinoja, mutta eivät kuitenkaan välttämättä näe näiden keinojen edistävän oppimista. Opettajista suurin osa raportoi käyttävänsä usein tai lähes aina kyselyssä kartoitettuja formatiivisen arvioinnin keinoja: tavoitteiden avaamista (82 %), palautteen antamista (87 %) ja arviointikriteerien avaamista (74 %) (kysymykset M1-M31). Ainoastaan 1–3 % kertoi, ettei käytä näitä keinoja lainkaan tai käyttää vain harvoin. Runsaasta käytöstä huolimatta vastaajat kuitenkin mainitsivat nämä keinot vain harvoin avovastauksissa, joissa kysyttiin tärkeimpiä oppimista edistäviä keinoja: vastaajista kaksi (1 %) mainitsi tavoitteiden avaamisen ja 11 % palautteen antamisen. Yksikään vastaaja (0 %) ei maininnut arviointikriteerien avaamista oppimista edistävänä keinona.

Vastaava epäsuhta toistui myös oman oppimisen omistajuuden ja yhteistyön kohdalla. Kolme neljästä (72 %) vastaajasta kertoi itsearviointiin vaikuttavan päät-

1 Karvin kyselyn prosentit on korjattu oikeiksi aiemmin raportissa Kuukka & Metsämuuronen (2016) virheellisesti esitetystä.

töarviointiin vähintään jonkin verran. Avovastauksissa itsearviointi nostettiin oppimista edistävänä tekijänä esiin ainoastaan kerran. Samoin lähes kaikki vastaajat (98 %) ilmoittivat kysymyksen M5 perusteella tekevänsä yhteistyötä muiden opettajien kanssa, mutta avovastauksissa vain noin kuudesosa (15 %) vastaajista ilmaisi sen edistävän oppimista.

Yhteistyötä koskevista vastauksista löydettiin myös alueita, joissa formatiivisen arvioinnin periaatteet ja opettajien käsitykset olivat lähempänä toisiaan. Opintojen aikaisessa arvioinnissa kolme neljästä (74 %) raportoi tekevänsä yhteistyötä. Toisaalta vain 13 % kertoi tekevänsä yhteistyötä koskien kielitaidon arviointia ja sen huomioimista muilla kuin S2-oppitunneilla, vaikka reilu kolmannes vastaajista (36 %) piti nimenomaan kielitaidon huomioimista oppimista edistävänä. Vastausten jakauma esitetään tarkemmin kuviossa 2.



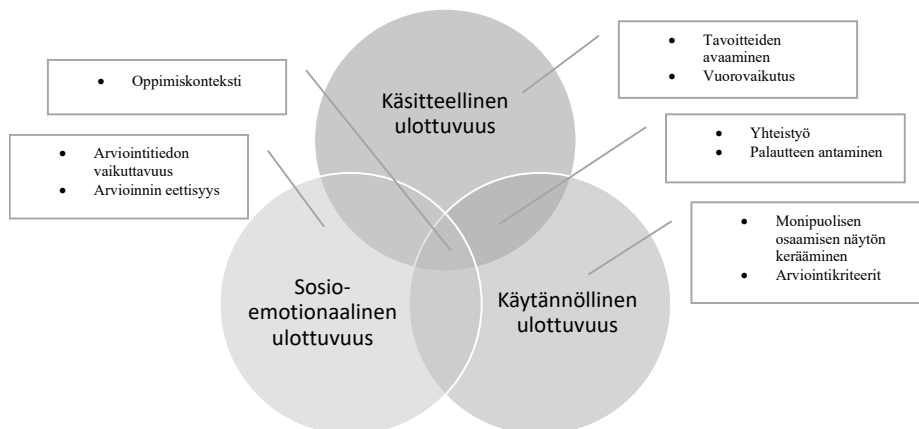
KUVIO 2. Vastausten jakautuminen eri käsitteiden välillä.

Avovastausten L1-L2 perusteella vastaajat raportoivat oppimiskontekstin, vuorovaikutuksen, monipuolisen opetuksen sekä arvioinnin vaikuttavuuden muuttujina, joilla on selkeää vaikutusta oppimiseen. Oppimiskontekstin merkitystä pohdittiin usein (54 %). Siitä nostettiin esiin yksilöllisiä ominaisuuksia, kuten opiskelutaidot (n=26), suomen kielen taso tai lähtökielen vaikutus (n=28) sekä motivaatio (n=118). Opiskelutaitojen puute (10 %) ja vasta kehittyvän suomen kielen tai lähtökielen negatiivinen vaikutus (12 %) nostettiin esiin useammin oppimista vaikeuttavina tekijöinä kuin näiden taitojen hallinta oppimista edistävinä. Motivaation osalta tilanne oli päinvastainen: lähes puolet (42 %) katsoi motivaation edistävän oppimista, kun noin kolmannes (32 %) näki motivaation puutteen vaikeuttavan oppimista.

Yhteisöllisistä tekijöistä nousi esiin luokan ilmapiiri, jonka 23 % vastaajista mainitsi oppimista edistävänä tekijänä. Vastaavasti 9 % vastaajista koki huonon ilmapiirin opiskelua vaikeuttavana. Kolme neljästä vastaajasta (76 %) ei kuitenkaan maininnut ilmapiiriä lainkaan. Vastaajista 18 % kertoi hyvän vuorovaikutuksen luokassa edistävän oppimista, kun 11 % raportoi sen puutteen vaikeuttavan oppimista. Suurin osa (71 %) ei käsitellyt aihetta lainkaan. Monipuolinen opetus edisti noin kolmasosan (29 %) mielestä oppimista, ja noin kuudesosa (16 %) vastaajista mainitsi sen puutteen vaikeuttavan oppimista. Arvioinnin vaikuttavuutta oppimiseen pohti puolet (49 %) vastaajista. Hieman alle puolet (44 %) nosti esiin kielitaidon arvioinnin vaikutuksen ryhmien jakamisessa.

3.2 Opettajien käsitysten suhde arviointiosaamisen ulottuvuuksiin

Avovastauksista (L1-L2) ilmenneet formatiivisen arvioinnin keinoja koskevat käsitykset jakautuivat arviointiosaamisen kolmelle pääulottuvuudelle suhteellisen tasaisesti (ks. kuvio 3).



KUVIO 3. Opettajien käsitysten jakautuminen Pastoren ja Andraden (2019) arviointiosaamisen mallin eri ulottuvuuksille.

Käsitteellisellä ulottuvuudella pohdittiin opetuksen tavoitteita, oppimiskontekstia, yhteistyötä ja vuorovaikutusta sekä palautteen antamista. Tavoitteissa opettajia mietitytti niiden sopivuus ja selkeys, kuten esimerkiksi seuraavassa:

Oppimäärän sisältö on nykyisellään jäsentymätön ja tavoitteet suhteessa tuntimääriin mahdottomia tai ristiriitaisia. (Opettaja 50)

Oppimiskontekstista huomioitiin motivaatio ja ilmapiiri, kielitaidon taso, opiskelutaidot ja oppimisvaikeudet. Näiden merkitys nähtiin useammin niiden puuttumisen kautta oppimista vaikeuttavana, mutta opiskelutaitojen ohjaamisen roolia oppimisen edistämässä ei tuotu esiin.

Yhteistyötä tehtiin pääosin kollegoiden kanssa, mutta myös opettajan ja oppilaan tai oppilaiden keskinäistä vuorovaikutusta käsiteltiin. Hyvä vuorovaikutus nähtiin lähes yhtä usein tärkeänä kuin sen puuttuminen oppimista vaikuttavana. Opettajan ja oppilaan välisen hyvän vuorovaikutuksen piirteitä olivat yksilöllisyys, luottamuksellisuus ja keskusteleva ilmapiiri. Oppilaiden keskinäinen hyvä vuorovaikutus liitettiin kielen kehittymiseen sosiaalisissa suhteissa.

Palaute liitettiin useimmiten kannustavuuteen ja positiivisuuteen. Yksilöllisyyden, yhteistyön ja vuorovaikutuksen näkökulmat palautteen antamisessa olivat harvinaisia, vaikka oppimisen tuesta puhuttaessa yksilöllisyys nähtiinkin tärkeänä, samoin korostuivat opiskelutaidot. Palautteen tai tuen antajana nähtiin yleensä opettaja, vain harvoin mainittiin vertaistuki.

Käytännöllisellä ulottuvuudella käsiteltiin palautetta, ilmapiiriä, monipuolisen näytön keräämistä, oppimiskontekstia ja yhteistyön käytänteitä. Palautteen antamista hankaloittivat vaikeiksi koetut arviointikriteerit: jos opetus suunnitelman tavoitteet olivat liian etäällä oppilaan kielitaidon tasosta, oli haastavaa antaa kannustavaa palautetta. Suuren eron osaamisen ja tavoitteiden välillä koettiin laskevan oppilaan motivaatiota. Ilmapiiri nähtiin kuitenkin keinona lisätä motivaatiota ja oppimista. Moni nimesikin oppimista edistäväksi kannustavan ja turvallisen ilmapiirin, jossa "sallitaan virheet, korjataan oppilasta ystävällisesti ja opettajan ja oppilaan välillä vallitsee luottamus". Myös työrauha ja keskusteleva opetus näkyivät käsityksissä hyvästä ilmapiiristä:

Oppitunneilla pitää olla vapaus kysyä ja ihmetellä, erehtyä ja aloittaa alusta.
(Opettaja 2)

Opetusryhmä ottaa S2-opiskelijan ryhmän jäseneksi ja kaveriksi. Turvallinen ilmapiiri, S2-oppilas rohkaistuu yrittämään ja toimimaan ryhmässä. (Opettaja 166)

Osaamisen näytön keräämisen sijasta opettajat keskittyivät opetusmenetelmien merkitykseen. Oppimisen koettiin parantuvan oppilaantuntemuksen ja yksilöllisten, monipuolisten opetusmenetelmien sekä sopivien materiaalien avulla. Osassa vastauksia käytettyjä opetusmateriaaleja pidettiin kuitenkin liian teoreettisina tai jopa sekavina ja siten oppimista vaikeuttavina. Tällä todennäköisesti viitattiin eri oppiaineissa valittuihin oppikirjoihin ja niiden lisämateriaaleihin, joiden perusteella myös osaamisen näyttö kerättiin. Oppimiskontekstin, kuten kielitaidon, oppimisvaikeuksien ja motivaation vaikutusta pohdittiin runsaasti. Opiskelutaitojen puute vaikeutti myös osaamisen näytön tulkintaa, kuten seuraavasta esimerkistä ilmenee:

Muut oppimisvaikeudet vaikeuttavat kielitaidon arviointia, sillä on vaikea tietää, onko kyseessä heikko kielitaito vai oppimisen ongelma. (Opettaja 55)

Kysymys arviointitiedon hyödyntämisestä oppilasta eri aineissa opettavien opettajien kesken nousi esiin yhteistyöstä puhuttaessa. Taitotasoa koskevan arvioinnin huomioimisen katsottiin toisaalta auttavan eriyttämisessä sekä sopivien materiaalien valinnassa ja opetusryhmien suunnittelussa. Taitotasolla viitattiin sekä kielitaidon että yleisempien opiskelutaitojen tasoon. Toisaalta vastauksista nousi esiin myös joidenkin kollegoiden negatiivinen asenne: S2-opettajat kuvasivat muiden aineiden opettajien suhtautuvan usein haluttomasti eriyttämiseen tähtäävään yhteistyöhön. Toinen yhteistyön nimenomainen haaste oli ajan puute.

Sosioemotionaalisen ulottuvuuden aiheista nousivat esiin arvioinnin eettisyys, oppimiskonteksti ja arvioinnin vaikutus oppilaan osallisuuteen, mutta myös sosiaalisiin suhteisiin. Eettisistä näkökohdista erityisesti arvioinnin oikeudenmukaisuus, tasa-arvoisuus ja yhdenvertainen kohtelu mietityttivät S2-opettajia. Seuraava lainaus kuvaa ajattelua:

Erytisen haastavalta tuntuu arviointi tilanteissa, joissa olen nähnyt kolmen vuoden ajan oppilaan yrittävän kovasti, -- ja tuntuu, että kokeista tulee heikkoja arvosanoja --. Tällöin arviointi tuntuu vaikealta. Samoin toisinaan arviointi tuntuu vaikealta, kun S2-oppilaalla on myös muita oppimisvaikeuksia. Joskus on myös hankala erottaa erityisesti koulukielteisen oppilaan kohdalla, mikä osuus on oppimisvaikeuksilla, mikä kieliongelmillä ja mikä johtuu vain puutteellisesta työnteosta tai asenneongelmista. (Opettaja 51)

Oppimiskontekstin osalta korostuivat opiskelutaidot, motivaatio ja kielitaito sekä lähtökielen vaikutus. Myös sosiaaliset suhteet, henkilökohtainen asenne sekä koulunkäyntihistoria nähtiin merkityksellisinä. Arvioinnin vaikuttavuutta käsiteltiin lähinnä kielitaidon arvioinnin ja opetusjärjestelyiden näkökulmasta. Kielitaidon arviointiin liittyvä yhteistyö katsottiin tärkeäksi keinoksi oppimisen edistämiseksi, kun se ohjasi oppilaiden ryhmittelyä. Vastaajat, jotka pitivät oppilaan taitotason arviointia ryhmien muodostamisen kannalta tärkeänä, kokivat heterogeeniset ryhmät oppimista edistävänä:

Oppilaalla on oltava sopivasti haasteita omaan taitotasoonsa nähden. Opiskelu edistyneempien ryhmässä voi motivoida oppilasta parempiin tuloksiin, mutta toisaalta hänelle tulisi tarjota myös sopivasti omaa tasoaan vastaavia eriyttäviä tehtäviä. (Opettaja 38)

Tärkeimpänä pitäisin sitä, että oppilas otetaan koulussa huomioon kokonaisuutena, eli ei niin, että S2-näkökulma otetaan huomioon vain S2-tunnilla, ja muuten oppilaan on selvittävä muilla oppitunneilla 'muiden mukana'. (Opettaja 60)

Käsitykset ryhmän heterogeenisyydestä kuitenkin jakoivat vastaajia: osa katsoi sen edistävän oppimista, mutta toisaalta tähän ryhmään verrattuna moninkertainen määrä katsoi sen vaikeuttavan oppimista. Vastauksien perusteella ei voitu päätellä tarkemmin, miten vastakkaisiin käsityksiin oli päädytty.

Osoa formatiivisen arvioinnin keinoista käsiteltiin useammalla ulottuvuudella. Oppimiskontekstin ja oppilaan taustan, kuten oppimisvaikeuksien ja opiskelutaitojen huomioimista oppimisessa pohdittiin kaikilla kolmella ulottuvuudella, erityisesti motivaation ja ilmapiirin näkökulmista. Tämän lisäksi yhteistyötä ja palautteen antamiseen liittyviä näkökulmia käsiteltiin sekä käsitteellisellä että käytännöllisellä ulottuvuudella.

4 Pohdinta ja johtopäätökset

Tässä tutkimuksessa selvitettiin S2-opettajien käsityksiä formatiivisen arvioinnin keinoista oppimisen edistämiseksi. Päätulos on, että vaikka opettajat tuntevat monia formatiivisen arvioinnin keinoja ja käyttävät myös niitä työssään, he eivät kuitenkaan välttämättä koe niitä oppimista edistäviksi. Opettajien käsitykset arvioinnin keinoista järjestyivät mielekkäästi Pastoren ja Andraden (2019) mallin kolmelle ulottuvuudelle: käsitteelliselle, käytännölliselle ja sosioemotionaalille. Mallin mukaisesti osa keinoista asettui samanaikaisesti useammalle ulottuvuudelle. Käsitteellisellä ja käytännöllisellä ulottuvuudella teemoja käsiteltiin kapeammin kuin sosioemotionaalilla ulottuvuudella.

Opettajat nostavat esiin tärkeitä formatiivisen arvioinnin keinoja, kuten tavoitteiden avaamisen, kannustavan palautteen, yhteistyön, oppimiskontekstin huomioimisen ja vuorovaikutuksen merkityksen. He eivät kuitenkaan juuri avaa niiden merkitystä eivätkä tuo esiin käytänteitä, joilla periaatteita sovelletaan opetustyössä. Sama huomio on tehty myös kansainvälisissä tutkimuksissa. Vaikka opettajilla on käsitys formatiivisesta arvioinnista, he eivät tunne käytänteitä, joilla se voitaisiin sitoa osaksi luokkahuonetyöskentelyä (Thomas ym. 2011: 394; Pastore ym. 2019: 368).

Tavoitteiden avaaminen vaikuttaa usein opetuksen irralliselta osalta, eikä sen hyötyjä tai käytänteitä juuri käsitellä. Sen sijaan opettajat kiinnittävät huomiota tavoitteiden sopivuuteen ja selkeyteen. Ilmiö näkyy myös kansainvälisissä tutkimuksissa: erityisesti toisen kielen oppijalle on hyödyllistä, että tavoitteet asetetaan ja avataan suhteessa oppilaan taitotasoon (Tsagari & Banerjee 2016: 264–265). On myös huomioitava, että jos tavoitteita lasketaan kielitaidon alkeisvaiheen vuoksi ilman yksilöllistettyä oppimäärää ja keskustelua huoltajien kanssa, vaarannetaan oppilaan oikeus yhdenvertaiseen kohteluun. Tämän aineiston valossa arviointiin liittyvä viestintä painottuu pääasiassa opettajien kesken käytävään keskusteluun, ei huoltajien.

Sekä tämän että aiemman tutkimuksen (Atjonen ym. 2019) perusteella opettajat näkevät kannustavan palautteen vahvistavan erityisesti oppimisessaan tukea tarvitsevia oppilaita. Toisaalta pelkkä kannustavuus ei paranna oppimistuloksia, vaan lisäksi palaute

on kohdennettava tarkasti (Hattie 2012). Opettajat myös toteavat, että kannustavan palautteen antaminen on vaikeaa liian etäällä olevien opetuksen tavoitteiden vuoksi. Tämä voi kertoa vaikeudesta konkretisoida ja avata opetussuunnitelman tavoitteita oppilaille tai kohdentaa palautetta tavoitteiden mukaisesti. Huomio saattaa kertoa myös summatiivisen ja formatiivisen arvioinnin periaatteiden yhteentörmäyksestä vastaajien ajattelussa, kun arvosana halutaan nähdä kannustuskeinona, kuten aiemmissakin tutkimuksissa on huomattu (Riley & Ungerleider 2019). Jos arviointiprosessin eri tehtävät nähtäisiin enemmän toisiaan täydentävinä, opettajat voisivat paremmin antaa sekä kannustavaa että realistista palautetta.

Opettajat eivät kiinnitä vastauksissaan huomiota palautteen antamisen käytänteisiin. Jo aiemmasta tutkimuksesta tiedetään, että palautteen merkitys näyttäytyy oppilaille vähäisenä (Mäkipää & Hilden 2021). Opettajat osoittavat toisaalta huolensa opiskelutaitojen puutteesta. Onkin todennäköistä, että S2-oppilaiden erilaisten koulupolkujen ja sen myötä syntyneiden oppimisen kontekstierojen huomioiminen palautetta annettaessa parantaisi oppimista. Oppimaan oppimisen taitoihin voitaisiin vaikuttaa kohdistamalla palaute oppimisprosessiin ja opiskelustrategioihin (Braund & DeLuca 2018; Wisniewski ym. 2020: 12).

Kuukka ja Metsämuuronen (2016) selvittivät S2-oppilaiden oppimistuloksia käyttämällä samaa aineistoa kuin tässä tutkimuksessa. Heidän mukaansa tavoitteiden kertomisella, arvioinnin perusteiden selvittämisellä tai palautteen annolla ei ollut tilastollisesti merkitsevää yhteyttä oppilaiden osaamiseen (mts. 121), mikä on linjassa tämän tutkimuksen tulosten kanssa.

Joitain formatiivisen arvioinnin olennaisia keinoja ei juuri lainkaan käsitetä oppimista edistäviksi. Itsearviointia käytetään opetuksessa, mutta sen merkitys oppimisprosessissa ja opiskelustrategioiden tukemisessa jää epäselväksi. Itse- ja vertaisarviointi ovat näkyvimpiä käytänteitä oppilaiden aktiivisuuden lisäämisessä (Andrade ym. 2021), mutta opettajat eivät yhdistä näitä keinoja oppilaan osallisuuden lisääntymiseen, vaikka pitävätkin osallisuutta tärkeänä. Tulokset ovat linjassa Atjosen (2014: 255) aiempien tutkimustulosten kanssa itse- ja vertaisarvioinnin vähäisestä arvostuksesta. Onkin mahdollista, että itsearviointia toteutetaan tavalla, joka ei lisää oppilaan ymmärrystä tai vastuunottoa omasta oppimisestaan.

Monipuolisen osaamisen näytön keräämisen sijasta opettajat korostavat monipuolisia opetusmenetelmiä. Kuten tavoitteiden, myös materiaalien yhteydessä viitataan usein niiden liian abstraktiin tasoon ja selkeyden puutteeseen. Tämän tutkimuksen perusteella ei kuitenkaan ole mahdollista määritellä, kohdistuuko kritiikki opetussuunnitelman tavoitteisiin vai itse opetusmateriaaliin. Osittain kyse voi olla siitä, että opetusmateriaalit eivät kovin onnistuneesti kiinnity tavoitteisiin, jotka äidinkielen ja kirjallisuuden osalta ovat hyvin laajat. Tämä on omiaan aiheuttamaan riittämättömyyttä: kehittyvällä kielitaidolla opiskelevien oppilaiden kanssa joudutaan tekemään valintoja, jotta löydetään ydinaines ja tuetaan tasapainoisesti sekä kielitaidon kehittymistä että oppimääräspesifien tavoitteiden saavuttamista. Toisaalta materiaalin puutteet

voivat heijastella myös puutteita arviointimateriaaleissa, jos arviointia tehdään paljon esimerkiksi oppikirjojen arviointimateriaalin varassa.

Vaikka opettajien välistä yhteistyötä pidetään tärkeänä, se koskee vain harvoin kielitaitoa tai arviointikäytänteiden jakamista. S2-opettajien mukaan muiden aineiden opettajat suhtautuvat eriyttämiseen tähtäävään yhteistyöhön usein haluttomasti eivätkä koe sitä hyödylliseksi. On kuitenkin todennäköistä, että yhteistyö ja selkeämpi käsitys oppilaan kielitaidosta auttaisi myös muita opettajia kohdentamaan palautetta paremmin. Arvioinnista saadun tiedon hyödyntäminen muiden oppiaineiden opetuksen suunnitteluun on formatiivisen arvioinnin hyvien käytänteiden mukaista (Turner & Purpura 2016). Yhteistyö tukisi myös kielitietoisuuden opetuksen toteutumista. Kielitietoinen ajattelu ja sen tiimoilta tehtävä yhteistyö on aineiston keruuajankohdan jälkeen luultavasti vahvistunut ja tullut paremmin osaksi koko koulun toimintakulttuuria nykyisten opetussuunnitelman perusteiden (OPH 2014) myötä.

S2-oppilaiden opiskelumotivaatio nähdään pääsääntöisesti hyvänä, kuten aiemmissakin tutkimuksissa (Harju-Luukkainen ym. 2016: 171–173). Tutkimusten perusteella motivaatio ei kuitenkaan aina korreloi hyvien oppimistulosten kanssa, mikä saattaa liittyä opiskelutaitoihin (mts. 179–180). Arvioinnilla voidaan todennäköisesti vaikuttaa siihen, motivoituu oppilas uuden tehtävän edessä kasvattamaan tietämystään vai välttelemään ponnistelua, peläten mahdollista epäonnistumista (Boekaerts 2007). Toisen kielen oppija voidaan ohjata kiinnittämään huomio olennaiseen sisältöön sekä oppimaan myös epäonnistumisista.

Palautteen antamiseen liittyvä luokan ilmapiiri nousee esiin eri näkökulmista. Osa opettajista korostaa luokissa vallitsevan ilmapiirin ongelmia, kuten työrauhan puutetta ja kiusaamista, kun taas osa painottaa hyvän ilmapiirin vaikutuksia oppimisen edistämiseksi. Opettajilla on vahvasti erisuuntaiset käsitykset siitä, tukisiko oppimista paremmin kieli- tai opiskelutaidoiltaan yhtenäinen vai erilainen ryhmä. Kysymys on monitahoinen ja liittyy muun muassa kysymykseen siitä, kuinka monikielisissä luokkayhteisöissä kielet elävät rinnakkain ja limittäin (ks. esim. Piippo ym. 2021).

Opettajien arviointiosaamisessa on huomionarvoista, että käsitteellisellä ulottuvuudella ilmenee epätarkkuutta formatiivisen arvioinnin käsitteiden merkityksissä. Monia tutkitusti hyödyllisiä keinoja ei esimerkiksi nähdä oppimista edistävinä, vaikka ne ovatkin käytössä. Opettajat myös kokevat riittämättömyyttä opetuksen tavoitteiden sekä opetusmateriaalien suhteen, mikä saattaa heijastella formatiivisen arvioinnin periaatteiden ja käytänteiden hallinnan puutteita.

Käytännöllisessä ulottuvuudessa merkittävää on, että juuri mitään formatiivisen arvioinnin periaatteista ei käsitellä käytänteiden kautta. Myös näkemykset arviointitiedon hyödyntämisen käytänteistä puuttuvat kokonaan. Formatiiivinen arviointi hyödyttää tutkitusti oppimista – mutta ainoastaan, jos siitä saatua tietoa käytetään opetuksen ja oppimisprosessin muokkaamisessa. Sosioemotionaalisen ulottuvuuden aiheista opettajat näkevät tärkeänä arvioinnin eettisyyden ja oppilaan osallisuuden

oppimisen edistämässä. Teemat ovat sidoksissa kysymykseen tavoitteiden sopivuudesta, mutta myös palautteen antamisen käytänteisiin.

Tutkimustulosten tulkinnessa ja laajemmassa hyödyntämisessä on huomioitava muutamia rajoitteita. Vastaajien vähäisen lukumäärän ja mahdollisen valikoitumisen vuoksi aineistoa ei voida pitää kattavana S2-opettajien osalta. Aineiston kysymys oppimisen edistämisestä on hyvin avoin, ja vastaajilla on ollut mahdollisuus käsitellä aihetta laajasti, joten sen perusteella arvioinnista tehtäviä päätelmiä voi pitää ainoastaan suuntaa antavina. Lisäksi aineistoa voidaan pitää jo hieman vanhentuneena. Se on vuodelta 2015. Uudet opetussuunnitelman perusteet (OPH 2014) otettiin käyttöön vasta 2016. Niissä formatiiviseen arviointiin kiinnitetään aiempaa enemmän huomiota, ja arviointikulttuuri onkin kehittynyt sen myötä. Vesarannan tuore tutkimus (2022) osoittaa, että S2-opettajien arviointiosaaminen on monipuolistunut ja formatiivisesta arvioinnista on tullut aiempaa säännöllisempää.

Tulosten perusteella ensisijaisena täydennyskoulutuksen ja jatkotutkimuksen kohteena voi pitää palautteen antamisen käytännön taitoja, erityisesti palautteen kohdentamista oppimisprosessiin ja opiskelustrategioihin. Myös itse- ja vertaisarviointikäytänteet toisen kielen opetuksessa kaipaavat erityistä huomiota. Oppimista hyödyttäisi opetusmateriaali, jossa tavoitteet avataan ja pilkotaan opetukseen sopiviksi. Rutiininomaisesti täytettävien itsearviointilomakkeiden sijasta arviointi- ja opetusmateriaalit voisivat ohjata itsearviointia aiempaa enemmän todellisen itsereflektoinnin suuntaan. Myös opettajia tulisi ohjata suhteuttamaan antamaansa palautetta itsearviointin tuloksiin. Opettajat hyötyisivät myös koulutuksesta ja materiaalista, joka osaamisen tason arvioinnin ohella ohjaa tulkitsemaan oppimistuloksia formatiivisesti siten, että ne auttavat valitsemaan oppilaille sopivia opiskelustrategioita. Käytänteiden soveltamista vaikeuttavat todennäköisesti S2-opetusryhmien heterogeenisyys ja opiskelutaitojen epätasainen jakautuminen. Jatkossa on syytä myös selvittää, kuinka S2-oppilaiden erityistarpeet voidaan paremmin huomioida formatiivisen arvioinnin käytänteissä.

Käsitteellisen ulottuvuuden kehittämistarpeet liittyvät formatiivisen arvioinnin keskeisten käsitteiden selventämiseen ja niiden merkityksen ymmärtämiseen. Mikäli opettajalla ei ole selkeää käsitystä formatiivisen arvioinnin keinojen sovellettavuudesta tai kattavaa käsitystä niiden taustalla vaikuttavista periaatteista, arvioinnista saadun tiedon hyöty opetuksessa jää ohueksi.

Sosioemotionaalisen ulottuvuuden osalta opettajia tulisi kannustaa arviointikäytänteiden jakamiseen kollegoiden kanssa. Opettajien yhteistyön muotoja S2-oppijan kielitaidon osalta on syytä lisätä osaksi koulun arviointikulttuuria. Opettajien tulisi myös nykyistä paremmin huomioida luokkahuoneen ilmapiirin vaikutus palautteen antamiseen ja vastaanottoon. Opetusryhmien muodostamisen perusteita on myös syytä selvittää. Arvioinnin vaikuttavuuden näkökulmasta on tärkeää käydä keskustelua siitä, miten oppilaat jaetaan kielitaidon tason perusteella eri opetusryhmiin ja millaiseen kieleen heitä halutaan sosiaalistaa. Jatkotutkimusta tarvitaan myös siitä, miksi

opettajat kokevat opetuksen tavoitteet omaan työhönsä vaikeasti sovellettavina tai epäselvinä, millainen palaute ohjaa S2-oppilaita erityisesti opiskelutaidoissa ja miten itse- ja vertaisarviointikäytänteitä voidaan soveltaa S2-opetuksessa.

Arviointi ei ole muusta opetus kontekstista irrallista toimintaa, vaan palvelee vallassa olevaa oppimiskäsitystä (Black & Wiliam 2009; Wiliam 2019). Nykyiset opetus suunnitelman perusteet painottavat oppilaan aktiivista roolia, osallisuutta, oppimisprosessin vuorovaikutteisuutta, oppimisen tavoitteiden ymmärtämistä ja oppimaan oppimisen metakognitiivisia taitoja, kuten formatiivinen arviointikin. Vaikka periaatteet eivät vielä kaikilta osin toteudu S2-opettajien opetuskäytänteissä, ne tunnetaan melko hyvin, mikä helpottaa jatkossa niiden jalkautumista myös käytännön ulottuvuudelle.

Kirjallisuus

- Ahlholm, M. 2020. Kielitaitoiseksi kasvaminen. Teoksessa L. Tainio, M. Ahlholm, S. Grünthal, S. Happonen, R. Juvonen, U. Karvonen & S. Routarinne (toim.) *Suomen kieli ja kirjallisuus koulussa. Suomen ainedidaktisen tutkimusseuran julkaisuja. Ainedidaktisia tutkimuksia 18*. Helsinki: Suomen ainedidaktinen tutkimusseura, 15–40.
- Airas, M., D., Delahunty, M., Laitinen, M., Saarilampi, T., Sarparanta, G., Shemsedini, H., Stenberg, H., Vuori, H. & Väättäin, H. 2019. *Taustalla on väliä. Ulkomaalaistaustaiset opiskelijat korkeakoulupolulla*. Julkaisut 22:2019. Helsinki: Kansallinen koulutuksen arviointikeskus.
- Andrade, H., Bennett, R. & Cizek, G. 2019. *Handbook of Formative Assessment in the Disciplines*. New York, NY: Routledge.
- Andrade, H., Brookhart, S. & Yu, E. C. 2021. Classroom assessment as co-regulated learning: a systematic review. *Frontiers in Education*, 6, 1–18. <https://doi.org/10.3389/feduc.2021.751168>
- Andrade, H. L. & Heritage, M. 2017. *Using formative assessment to enhance learning, achievement, and academic self-regulation*. Student Assessment for Educators. New York, NY: Routledge. <https://doi.org/10.4324/9781315623856>
- Atjonen, P. 2014. Teachers' views of their assessment practice. *The Curriculum Journal*, 25 (2), 238–259. <https://doi.org/https://doi.org/10.1080/09585176.2013.874952>
- Atjonen, P. 2017. Arviointiosaamisen kehittäminen yleissivistävän koulun opettajien koulutuksessa – opetussuunnitelmatarkastelun virittämiä näkemyksiä. Teoksessa V. Britschgi & J. Rautopuro (toim.) *Kriteerit puntarissa. Kasvatusalan tutkimuksia – Research in Educational Sciences 74*. Jyväskylä: Suomen Kasvatustieteellinen Seura.
- Atjonen, P., Laivamaa, H., Levonen, A., Orell, S., Saari, M., Sulonen, K., Tamm, M., Kamppi, P., Rumpu, N., Hietala, R. & Immonen, J. 2019. *”Että tietää missä on menossa” - Oppimisen ja osaamisen arviointi perusopetuksessa ja lukiokoulutuksessa*. Julkaisut 7:2019. Helsinki: Kansallinen koulutuksen arviointikeskus.
- Barcelos, A. M. F. & Kalaja, P. 2011. Introduction to beliefs about SLA revisited. *System*, 39 (3), 281–289. <https://doi.org/https://doi.org/10.1016/j.system.2011.07.001>
- Bennett, R. 2011. Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18, 5–25. <https://doi.org/10.1080/0969594X.2010.513678>
- Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. 2004. Working inside the black box: assessment for learning in the classroom. *Phi Delta Kappan*, 86 (1), 8–21. <https://doi.org/10.1177/003172170408600105>

- Black, P. & Wiliam, D. 1998a. Assessment and classroom learning. *Assessment in Education*, 5 (1), 7–74. <https://doi.org/http://dx.doi.org/10.1080/0969595980050102>
- Black, P. & Wiliam, D. 1998b. Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, 80 (2), 139–148.
- Black, P. & Wiliam, D. 2009. Developing the theory of formative assessment. *Educational assessment, evaluation and accountability*, 21 (1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Boekaerts, M. 2007. Self-regulation and effort investment. Teoksessa K. A. Renninger, I. E. Sigel, W. Damon & R. M. Lerner (toim.) *Handbook of Child Psychology*. Child Psychology in Practice. Hoboken, New Jersey: John Wiley & Sons Inc. <https://doi.org/10.1002/9780470147658.chpsy0409>
- Braund, H. & DeLuca, C. 2018. Elementary students as active agents in their learning: an empirical study of the connections between assessment practices and student metacognition. *Australian Educational Researcher*, 45 (1), 65–85. <https://doi.org/https://doi.org/10.1007/s13384-018-0265-z>
- Brookhart, S. M. 2011. Educational assessment knowledge and skills for teachers. *Educational measurement: Issues and practice*, 30 (1), 3–12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Cizek, G., Andrade, H. & Bennett, R. 2019. Formative assessment: history, definition, and progress. Teoksessa H. Andrade, R. Bennett & G. Cizek (toim.) *Handbook of Formative Assessment in the Disciplines*. New York, NY: Routledge, 3–19.
- Cohen, A. D. 2013. Using test-wiseness strategy research in task development. Teoksessa *The Companion to Language Assessment*. 893–905. <https://doi.org/https://doi.org/10.1002/9781118411360.wbcla006>
- Coombs, A. & DeLuca, C. 2022. Mapping the constellation of assessment discourses: a scoping review study on assessment competence, literacy, capability, and identity. *Educational assessment, evaluation and accountability*, 34 (3), 279–301. <https://doi.org/10.1007/s11092-022-09389-9>
- Deane, P. & Sparks, J. 2019. Scenario-based formative assessment of key practices in the English language arts. Teoksessa H. L. Andrade, R. E. Bennett; & G. J. Cizek (toim.) *Handbook of Formative Assessment in the Disciplines*. New York, NY: Routledge, 68–96. <https://doi.org/https://doi.org/10.4324/9781315166933>
- Eskola, J. & Suoranta, J. 1998. *Johdatus laadulliseen tutkimukseen*. Tampere: Vastapaino.
- Harju-Luukkainen, H., Nissinen, K., Sulkunen, S., Suni, M. & Vettenranta, J. 2014. *Avaimet osaamiseen ja tulevaisuuteen. Selvitys maahanmuuttajataustaisten nuorten osaamisesta ja siihen liittyvistä taustatekijöistä PISA 2012 -tutkimuksessa*. Jyväskylä: Jyväskylän yliopisto: Koulutuksen tutkimuslaitos.
- Harju-Luukkainen, H., Tarnanen, M. & Nissinen, K. 2016. Monikieliset oppilaat koulussa: eri kieliryhmien sisäinen ja ulkoinen motivaatio sekä sen yhteys matematiikan osaamiseen PISA 2012 -arvioinnissa. Teoksessa A. Huhta & R. Hilden (toim.) *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFinLA-e. Soveltavan kielitieteen tutkimuksia. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 167–183.
- Hattie, J. 2012. *Visible Learning for Teachers. Maximizing Impact on Learning*. <https://doi.org/10.4324/9780203181522>
- Hattie, J. & Timperley, H. 2007. The power of feedback. *Review of Educational Research*, 77 (1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hattie, J. & Yates, G. 2014. *Visible learning and the science of how we learn*. New York, NY: Routledge.

- Hilden, R. & Fröjdendal, B. 2018. The dawn of assessment literacy – exploring the conceptions of Finnish student teachers in foreign languages. *Apples – Journal of Applied Language Studies*, Vol. 12 (1), 1–24.
<https://doi.org/http://dx.doi.org/10.17011/apples/urn.201802201542>
- Jahnukainen, M., Kalalahti, M., Kivirauma, J. & Holmberg, L. 2019. *Oma paikka haussa: maahanmuuttotataustaiset nuoret ja koulutus*. Helsinki: Gaudeamus.
- Kalaja, P., Barcelos, A. M., F., Aro, M. & Ruohotie-Lyhty, M. 2015. *Beliefs, agency and identity in foreign language learning and teaching*. New York: Palgrave Macmillan.
- Kuukka, K. & Metsämuuronen, J. 2016. *Perusopetuksen päättövaiheen suomi toisena kielenä (S2) -oppimäärän oppimistulosten arviointi 2015*. Julkaisut 13:2016. Helsinki: Kansallinen koulutuksen arviointikeskus. <https://doi.org/10.13140/RG.2.1.2569.0005>
- Leino, K., Ahonen, A. K., Hienonen, N., Hiltunen, J., Lintuvuori, M., Lähteinen, S., Lämsä, J., Nissinen, K., Nissinen, V., Puhakka, E., Pulkkinen, J., Rautopuro, J., Sirén, M., Vainikainen, M.-P. & Vettenranta, J. 2019. *PISA18 – ensituloksia*. Opetus- ja kulttuuriministeriön julkaisuja 2019:40. Opetus- ja kulttuuriministeriö.
<http://urn.fi/URN:ISBN:978-952-263-678-2>
- Mäkipää, T. & Hilden, R. 2021. Kannustava palaute ja opettajan arviointityö lukion kieltenopetuksessa. *Kieli, koulutus ja yhteiskunta*, 12(6). Jyväskylä: Soveltavan kielentutkimuksen keskus, Jyväskylän yliopisto; Kielikoulutuspolitiikan verkosto.
<https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-joulukuu-2021/kannustava-palaute-ja-opettajan-arviointityo-lukion-kieltenopetuksessa>
- OKM. 2022. *Suomi/ruotsi toisena kielenä -opetuksen nykytilan arviointi*. Opetus- ja kulttuuriministeriön selvityksiä 2022-03-17. Helsinki: Opetus- ja kulttuuriministeriö.
https://owalgroup.com/wp-content/uploads/2022/03/S2-opetuksen_arviointi_170322.pdf
- OPH. 2014. *Perusopetuksen opetussuunnitelman perusteet*. Määräykset ja ohjeet 2014:96. Helsinki: Opetushallitus. https://www.oph.fi/sites/default/files/documents/perusopetuksen_opetussuunnitelman_perusteet_2014.pdf
- Panadero, E., Andrade, H. & Brookhart, S. 2018. Fusing self-regulated learning and formative assessment: a roadmap of where we are, how we got here, and where we are going. *The Australian Educational Researcher*, 45 (1), 13–31.
<https://doi.org/10.1007/s13384-018-0258-y>
- Pastore, S. & Andrade, H. L. 2019. Teacher assessment literacy: a three-dimensional model. *Teaching and teacher education*, 84, 128–138. <https://doi.org/10.1016/j.tate.2019.05.003>
- Pastore, S., Manuti, A. & Scardigno, A. F. 2019. Formative assessment and teaching practice: the point of view of Italian teachers. *European Journal of Teacher Education*, 42 (3), 359–374. <https://doi.org/10.1080/02619768.2019.1604668>
- Piippo, I. 2021. Muuttuvat näkökulmat monikielisyteen. Teoksessa M. Ahlholm, I. Piippo & P. Portaankorva-Koivisto (toim.) *Koulun monet kielet*. AFinLA-e: Soveltavan kielitieteen tutkimuksia 13. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 21–43.
<https://doi.org/10.30660/afinla.107219>
- Piippo, I., Ahlholm, M. & Portaankorva-Koivisto, P. 2021. Koulun monet kielet: lähtökohtia, kehityskulkuja ja tulevaisuuden näkymiä. Teoksessa M. Ahlholm, I. Piippo & P. Portaankorva-Koivisto (toim.) *Koulun monet kielet*. AFinLA-e: Soveltavan kielitieteen tutkimuksia 13. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 6–20.
<https://doi.org/10.30660/afinla.109285>
- Ruiz-Primo, M. A. & Brookhart, S. M. 2018. *Using feedback to improve learning*. Student assessment for educators. New York, NY: Routledge.
- Stiggins, R. J. 1999. Evaluating classroom assessment training in teacher education programs. *Educational Measurement Issues and Practice* 18 (1), 23–27.

- Thomas, L., Deaudelin, C., Desjardins, J. & Dezutter, O. 2011. Elementary teachers' formative evaluation practices in an era of curricular reform in Quebec, Canada. *Assessment in Education: Principles, Policy and Practice*, 18 (4), 381–398.
<https://doi.org/10.1080/0969594x.2011.590793>
- Tsagari, D. & Banerjee, J. 2016. *Handbook of Second Language Assessment*. Handbooks of applied linguistics Vol. 12. Boston: De Gruyter, Inc.
- Turner, C. E. & Purpura, J. E. 2016. Learning-oriented assessment in second and foreign language classrooms. Teoksessa D. Tsagari & J. Banerjee (toim.) *Handbook of Second Language Assessment*. De Gruyter Mouton, 255–274. <https://doi.org/doi:10.1515/9781614513827-018>
- Vesaranta, H. 2022. *Suomi toisena kielenä ja kirjallisuus -oppimäärän arviointi opettajien käsityksissä*. Tampereen yliopiston väitöskirjat 620. Tampereen yliopisto.
<https://urn.fi/URN:ISBN:978-952-03-2449-0>
- Wiliam, D. 2011. What is assessment for learning? *Studies in educational evaluation*, 37 (1), 3–14. <https://doi.org/10.1016/j.stueduc.2011.03.001>
- Wiliam, D. 2019. Why formative assessment is always both domain-general and domain-specific and what matters is the balance between the two. Teoksessa H. Andrade, R. Bennett & G. Cizek (toim.) *Handbook of Formative Assessment in the Disciplines*. New York, NY: Routledge, 243–264.
- Wiliam, D. & Thompson, M. 2007. Integrating assessment with learning: What will it take to make it work? Teoksessa C. A. Dwyer (toim.). Routledge, 53–82.
<https://doi.org/10.4324/9781315086545-3>
- Wisniewski, B., Zierer, K. & Hattie, J. 2020. The Power of Feedback Revisited: A Meta-Analysis of Educational Feedback Research. *Frontiers in psychology*, 10, 3087–3087.
<https://doi.org/10.3389/fpsyg.2019.03087>
- Xu, Y. & Brown, G. T. L. 2016. Teacher assessment literacy in practice: A reconceptualization. *Teaching and teacher education*, 58, 149–162.
<https://doi.org/10.1016/j.tate.2016.05.010>

Liite 1

TAULUKKO 1. Vastaajien taustatiedot.

	Vastaajat		Ikä				Koulut		Työtehtävä			
	n	%	25–24	35–44	45–54	55–	n	%	S2	S2 ja S1	S2 ja EO	S2 ja muu
AVI-alue												
Etelä-Suomi	106	51,0	26	44	24	12	106	62,7	18	71	9	3
Itä-Suomi	8	3,8	4	0	3	1	8	4,7	0	6	0	2
Lappi	3	1,4	1	1	1	0	3	1,8	0	3	0	0
Lounais-Suomi	34	16,3	7	10	11	6	34	20,1	1	23	7	2
Länsi- ja Sisä-Suomi	34	16,3	13	10	8	3	34	20,1	9	18	2	4
Pohjois-Suomi	12	5,8	5	4	3	0	12	7,1	1	7	1	2
Tieto puuttuu	11	5,3	2	7	2	0	11	6,5	0	9	1	1
Yht.	208	100,0	58	76	52	22	169	100,0	29	137	20	13

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 162–183.

Päivi Laine

Turun yliopisto

Minna Maijala

Turun yliopisto

Katja Mäntylä

Jyväskylän yliopisto

Kestävän kehityksen periaatteet kielenoppimisen arvioinnissa

Nostot

- Kielenopettajat eivät tunnustaneet arvioinnin yhteyttä kestävyyskasvatukseen.
- Kestävän kehityksen periaatteita pidettiin lähinnä aihepankkina, josta voi valita sopivia tekstejä ja sanastoa opetukseen.
- Tutkimus osoitti kuitenkin, että vertaisarvioinnilla ja -palautteella on tärkeä sosiaalinen tehtävä kielenopetuksessa.
- Kielenopettajien koulutuksessa olisi korostettava kestävä kehityksen periaatteiden mukaisesti arvioinnin eettisyyttä ja vaikutusta oppijan tulevaisuuteen.

Abstract

Sustainable development is included in curricula in many countries – from primary to tertiary education, as well as in teacher training. Its principles also encompass language teaching, e.g., through enhancing equality of languages, multilingualism, and multiculturalism. Several ethical issues are linked to sustainable development goals, which can be considered part of language teaching and assessment. Based on questionnaire data, this article addresses language teachers' views on sustainable development and its connection to assessment of language learning. The results show that language teachers (N=29) do not recognize the link between assessment and teaching practices that follow sustainable development principles. Engagement and peer assessment seemed to be a way to support sustainable development principles in language teaching. Further, the results indicate that language teachers need more information about the effects of ethical assessment on supporting language learning and social sustainability.

Keywords: language teaching, language teachers, assessment, sustainable development
Asiasanat: kieltenopetus, kieltenopettajat, arviointi, kestävä kehitys

1 Johdanto

Viime vuosien aikana kestävyyskasvatus ja kestävä kehitys on tullut entistä enemmän kaikkien tietoisuuteen. Kestävyyskasvatuksen perustana ovat Yhdistyneiden kansakuntien (YK) *Agenda 2030 -ohjelma* ja sen tavoitteet (Didham 2018, UNESCO 2018), joiden edistämiseen on sitouduttu maailmanlaajuisesti. Osittain YK:n ohjelman ansiosta kestävyteen liittyvät aihepiirit sisältyvät nykyään monissa maissa koulujen opetussuunnitelmiin (Gericke ym. 2020), yliopistojen ja korkeakoulujen opetusohjelmiin (Weiss ym. 2021) sekä opettajankoulutukseen (Evans ym. 2017). Kestävyydellä tarkoitettiin alun perin yritystä suojella luonnonvaroja ylikuluttamiselta ja säilyttää niitä tuleville sukupolville (Kuhlman & Farrington 2010). Vuonna 1987 julkaistettu Brundtlandin raportti *Our Common Future* (WCED 1987) pyrki selkeyttämään kestävä kehityksen periaatteita muun muassa yhdistämällä sen alle sosiaalisia, ekonomisia, kulttuurisia ja ympäristöllisiä näkökulmia. Käsitteenä ja tulevaisuuden päämääränä kestävä kehitys on ollut tapana jakaa kolmeen ydinalueeseen, joita ovat yhteiskunta, ympäristö ja talous (ks. esim. Giddings ym. 2020). Kieltenopetuksen kannalta merkittävä neljäs alue, kulttuurinen dimensio, katsottiin aikaisemmin osaksi sosiaalista ulottuvuutta, mutta nykyään sitä pidetään omana ydinalueenaan (ks. esim. Burford ym. 2013). Kestävä kehitys aiheilla eli kestävyysaiheilla tarkoitamme tässä artikkelissa YK:n *Agenda 2030* -ohjelmaan perustuvia kestävä kehityksen ekologisia, kulttuurisia, sosiaalisia ja taloudellisia ulottuvuuksia ja niihin liitettyjä aihepiirejä (ks. esim. OKKA-säätiö 2022; Laine ym. 2022). Vaikka kestävä kehityksen käsite usein liitetäänkin fyysiseen ympäristöön, se kattaa laajemmin myös tasa-arvon ja sosiaalisen oikeudenmukaisuuden teemoja, jotka ovat kieltenopetuksen keskiössä (Zygmunt 2016; de la Fuente 2021).

Kestävä kehitys on dynaaminen prosessi, jolla tavoitellaan kestävää tulevaisuutta erilaisilla keinoilla (Sinakou 2019a). UNESCO:n (2021) mukaan kestävyyskasvatus on keino saavuttaa YK:n *Agenda 2030* -ohjelman kestävä kehityksen tavoitteet, eli vuoteen 2030 mennessä kaikilla oppijoilla tulisi olla hallussaan tietoja ja taitoja, joilla kestävä kehitys voidaan edistää. UNESCO:n (2018) edistämän kestävyyskasvatuksen toimintamallin mukaan oppiminen on elinikäinen prosessi. Siinä ei opita pelkästään tiedollisia vaan myös sosiaalisia, emotionaalisia ja vuorovaikutuksellisia taitoja (UNESCO 2018). Kokonaisvaltainen malli kattaa opetuksessa käytettävät sisällöt, menetelmät ja oppimisympäristön. Myös oppimisen arviointiin liittyvät ratkaisut ovat olennainen osa toimintamallia, sillä kestävyyskasvatuksen tavoitteena on lisätä oppijoiden osallistumista sekä koulussa että yhteiskunnassa koulun ulkopuolella. (Sinakou ym. 2019a).

Oppiaineesta riippumatta tulisi opetuksen ja arviointikäytänteiden olla linjassa keskenään. Jos opetuksen tavoitteena on olla kestävä kehityksen periaatteiden mukaista ja valmistaa opiskelijoita tulevaisuutta eikä pelkästään seuraavaa koetta varten, kestävä kehityksen periaatteiden tulisi näkyä myös arvioinnissa (ks. esim. Boud 2000). Kestävyyskasvatus on pohjimmiltaan arvopohjaista (Hofman-Bergholm 2018: 25), mikä tarkoittaa, että sekä opettajien että opiskelijoiden tulee olla tietoisia toimintansa vaikutuksista. Kestävyyskasvatukseen eivät siis kuulu pelkästään eri oppiaineiden aihepiirit ja sisällöt, vaan se käsittää myös opetuksen taustalla vaikuttavat periaatteet ja arvot. Näin ollen kestävä arviointiprosessiin kuuluu tärkeänä osana sen pohtiminen, mitä vaikutusta opetuksella ja arvioinnilla on oppijoiden oppimistaitoihin ja oman oppimisen reflektointiin pitkällä aikavälillä (ks. esim. Beck ym. 2013). Arviointi ei tarkoita vain oppimistulosten saavuttamisen toteamista, vaan tämän lisäksi pohditaan arvioinnin vaikutuksia oppijan tulevaisuuteen ja siinä tarvittaviin taitoihin (ks. esim. Boud & Soler 2016).

Ns. oppimista tukevan arvioinnin periaate (*assessment for learning*), joka usein mainitaan oppimisen arvioinnin yhteydessä (ks. esim. Black 1986), sopii myös kestävyyskasvatukseen. Kieltenopettajat voivat myös arviointikäytänteillään ohjata opiskelijoita näkemään vaihtoehtoisia tulevaisuuksia (vrt. Laherto 2020) ja antaa heille työkaluja siihen, miten he pystyvät itse ratkaisemaan ongelmia ja vaikuttamaan omaan ja muiden tulevaisuuteen. Tässä artikkelissa esittelemme kyselytutkimuksen, jonka tavoitteena oli selvittää, miten kieltenopettajat ymmärtävät kestävä kehityksen ja arvioinnin välisen yhteyden ja miten kieltenopettajat huomioivat kestävä kehityksen periaatteita omassa opetuksessaan. Analysoimme kyselytutkimusaineistosta (N=29) etenkin opettajien huomioita siitä, kuinka arviointi on yhteydessä kestävä kehityksen periaatteisiin ja kestävyyskasvatukseen edistämiseen. Luvussa 2 tarkastelemme, millaista kestävä kehityksen periaatteiden mukainen kieltenopetus voisi olla. Luvussa 3 keskitymme pohtimaan arvioinnin erityispiirteitä osana kieltenopetusta ja kestävyyskasvatusta. Luvussa 4 esittelemme kyselytutkimusaineiston keruu- ja analyysimenetelmät sekä luvussa 5 analyysin tulokset. Lopuksi pohdimme kestävä kehityksen periaatteiden mukaisen arvioinnin toteutumisen mahdollisuuksia ja haasteita kieltenopetuksessa.

2 Kestävä kehitys ja kieltenopetus

Kestävyyskasvatuksen ja kestäväen kehityksen määritelmässä korostuvat usein biodiversiteettiin tai globaaliin terveyteen liittyvät seikat (ks. esim. Murillo-Vargas ym. 2020; Wals 2009). Kestävä kehitys liitetään usein konkreettiseen, fyysiseen ympäristöön, kuten ekologisuuteen tai kierrätykseen, harvemmin taas sosiaaliseen ja kulttuuriin ulottuvuuteen lukeutuviin sisältöihin, kuten ihmisoikeuksiin, demokratiaan ja tasa-arvoon (ks. esim. Inha ym. 2020; Majjala ym. 2023). Koulujen opetuksessa kestäväen kehityksen aihepiirien katsotaan kuuluvan yleensä luonnontieteiden opetukseen (Tani & Aarnio-Linnanvuori 2020) ja ympäristö- ja ilmastoasioihin eikä niinkään ajattelun muutosta mahdollistaviin taitoihin (UNESCO 2021), joita ovat esimerkiksi ihmissuhde- ja sosiaaliset taidot, kriittinen ajattelu, yhteistyötaidot sekä toimintakyky erilaisissa tilanteissa (Sinakou ym. 2019a). Näihin 21. vuosisadan taitoihin liittyy läheisesti kestävyyskasvatuksen alan tutkimuskirjallisuudessa usein esille tuleva systeeminen ajattelu (*systems thinking*), millä tarkoitetaan kykyä ajatella kokonaisvaltaisesti. Systeemisen ajattelun tavoitteena on edistää ymmärrystä järjestelmän toisiinsa liittyvistä osista, niiden keskinäisestä työnjaosta ja vuorovaikutussuhteista sekä myös siitä, miten osat vaikuttavat niitä laajempiin järjestelmiin. (Ks. esim. Lewis ym. 2014.) Systeemisen ajattelun taustalla on ajatus siitä, että oppijoiden pitäisi oppia ymmärtämään kestäväen kehityksen eri ulottuvuuksien välisiä yhteyksiä (Sinakou ym. 2019a). Mielenkiintoista on, että joidenkin tutkimustulosten mukaan opettajista etenkin kieltenopettajat käyttivät systeemisen ajattelun mukaisia työtapoja (Ferguson ym. 2021; Gunina ym. 2021). Tämä on hyvin rohkaisevaa kestävyyskasvatuksen, kieltenopetuksen ja sitä kautta myös kieltenopetuksen kestäväen kehityksen periaatteiden mukaisen arvioinnin kannalta.

Vaikka kestäväen kehityksen merkitystä pidetään yleisesti olennaisena osana kaikkien aineiden opetusta, sen käytännön toteutus vaihtelee esimerkiksi opettajien taidoista riippuen (Monroe ym. 2013). Eri opettajaryhmät huomioivat kestäväen kehityksen opetuksessaan myös eri tavoin. Esimerkiksi aineenopettajia käsittelevissä tutkimuksissa on havaittu, että erot johtuvat oppiaineiden erilaisista opetustraditioista (Sund & Gericke 2020) tai siitä, mitä aiheita opettajat pitävät tärkeinä oman aineensa opetuksessa (Uitto & Saloranta 2017). Kieltenopetuksessa korostuvat käytännön tietojen ja taitojen harjoittelu ja soveltaminen, ja kestäväen kehityksen sisältöjä yhdistetään erityisesti median ja vuorovaikutuksellisten opetusmenetelmien kautta, esimerkiksi roolileikeissä ja suullisissa esityksissä (Sund & Gericke 2020; Uitto & Saloranta 2017). Kieltenopetuksen sisällöt poikkeavat reaaliaineiden sisällöistä (Borg 2006) myös sen vuoksi, että kieltenopetuksessa kieli on sekä opetuksen väline että opetettava sisältö (Hammadou & Bernhardt, 1987; Halliday 1993). Tämän takia kieliaineiden arviointi vastaavasti poikkeaa muiden aineiden arvioinnista.

Kieltenopetusta käsittelevässä tutkimuskirjallisuudessa on tapana erottaa deklariivinen tieto eli tieto kielestä ja proseduraalinen tieto eli tieto siitä, miten kieltä käytetään (ks. esim. O'Malley & Chamot 1990). Kielenoppijat saattavat odottaa dekla-

ratiivisen tiedon arviointia, kuten sanaston ja kieliopin hallintaa, mutteivät esimerkiksi kulttuuritiedon ja -taidon arviointia (ks. esim. Majjala 2013). Kulttuuritaitojen arviointikeinona käytetään usein itsearviointia (ks. esim. Byram 2009), jonka kohteena ovat oppijoiden omat tuotokset, kuten esitelmät, haastattelut, blogit tai portfoliot (Wagner ym. 2018). Kulttuuritaitojen arviointia pidetään yleensä haastavana (ks. esim. Paran & Sercu 2010) mm. kulttuurin käsitteen monimuotoisuuden vuoksi (Liddicoat & Scarino 2010: 52). Mikäli kieltenopetuksessa halutaan edistää kestävän kehityksen arvoja ja periaatteita sekä tukea niiden kautta oppijoiden kasvua kieliyhteisön aktiivisiksi jäseniksi, sosiaalisten ja kulttuuristen taitojen arviointi olisi kuitenkin tärkeää. Arvioinnin oikeudenmukaisuuteen ja eettisyyteen liittyvät ongelmat on viime aikoina tunnistettu myös kielitaidon arvioinnissa. Esimerkiksi International Language Testing Association -järjestön (ILTA) eettisessä koodistossa on ohjattu vuodesta 2000 kielitaidon testaajia toimimaan oikeudenmukaisesti ja eettisten periaatteiden mukaisesti (ILTA Code of Ethics 2000; ks. myös Tossavainen 2016).

Suomalaisissa opetussuunnitelmissa kestävän kehityksen periaatteita huomioidaan alkaen varhaiskasvatuksen opetussuunnitelman perusteista, joiden yhtenä arvoperustana on terveellinen ja kestävä elämäntapa (Opetushallitus 2022). Vaikka perusopetuksen opetussuunnitelma julkaistiin edellisellä vuosikymmenellä, siinä näkyy jo selvästi kestävyyskasvatuksen aseman vahvistuminen (Opetushallitus 2014). Vielä selvemmin kestävyyskasvatuksen osuus näkyy lukion opetussuunnitelmassa, jossa kestävän kehityksen aiheita otetaan esille erityisesti laaja-alaisen osaamisen ja monialaisten oppimiskokonaisuuksien sekä ainekohtaisten tavoitteiden kuvauksissa (Opetushallitus 2019). Näin ollen kestävyyskasvatus kieltenopetuksessa voidaan huomioida paitsi sisällöissä ja opetusmenetelmissä myös arviointikäytännöissä. Koska kieltenopettajien voi olla haasteellista jo pelkästään tunnistaa kestävyysaiheita ja yhdistää niitä käytännön opetukseen (Al Amin & Greenwood 2018; Laine ym. 2022), voi kestävän kehityksen periaatteiden mukainen arviointi olla vielä tunnistamistakin haasteellisempää.

3 Arvioinnin merkitys osana kestävyyskasvatusta

Kieltenopetuksessa esimerkiksi dynaaminen arviointi (Glisan ym. 2007; Lantolf & Poehner 2014) tukee oppilaan osallistamista arviointiin. Myös opetussuunnitelmien perusteet ohjaavat monipuoliseen ja oppimista tukevaan arviointiin sekä arvostavan arviointikulttuurin luomiseen (Opetushallitus 2014; 2020). *Perusopetuksen opetussuunnitelman perusteet* (myöhemmin POPS) erottelee toisistaan formatiivisen ja summatiivisen arvioinnin (Opetushallitus 2014; 2020). Pääpaino on POPSin mukaan formatiivisessa arvioinnissa, jonka tulee yhdistyä ohjaavaan palautteeseen, oppijan oppimisprosessiin sekä oppimista tukevaan arviointikulttuuriin. Summatiivinen arviointi mainitaan, kun puhutaan kurssin tai lukuvuoden päättöarviointista. Kummankin arviointitavan

kohdalla todetaan, että oppilaan on oltava selvillä oppiaineen tavoitteista ja arvioinnin periaatteista (Opetushallitus 2020: 2–3).

Kestävän kehityksen mukaista arviointia on tutkittu erityisesti korkeakoulupedagogiikan kontekstissa (Boud 2000; Carless ym. 2011; Fastré ym. 2013; Boud & Soler 2016). Kielten opetuksen ja oppimisen alalla on eri oppilaitosasteilla tutkittu itsearviointia (esim. Butler 2016; Mäkipää 2020), vertaisarviointia (esim. Melatiadou 2022) sekä dynaamista arviointia (esim. Lantolf & Poehner 2014), jonka mukaan palaute on keino yhdistää arviointia tiiviimmin opetukseen (Glisan ym. 2007). Itse- ja vertaisarvioinnin voidaan katsoa edistävän kestävä kehityksen periaatteita, koska niiden kautta oppija pääsee osallistumaan häntä koskevaan arviointiin (vrt. Sinakou 2019a). Myös tuorein POPS (Opetushallitus 2014; 2020) korostaa oppimista tukevaa arviointikulttuuria ja keskittymistä formatiiviseen, oppijaa ja oppimisen prosesseja tukevaan arviointiin ja palautteeseen. POPS heijastelee siis kestävä kehityksen periaatteiden mukaisia palautekäytäntöjä ja oppijan vastuunottoa omasta oppimisesta (vrt. Boud & Soler 2016). Aikuisten yliopisto-opiskelijoiden keskuudessa tehdyn palautekäytänteisiin keskittyvän tutkimuksen tuloksissa (Carless ym. 2011) korostuivat erityisesti itsesääntelyn ja -arvioinnin merkitys sekä palautteenannon vuorovaikutuksellisuus. Toisaalta kielipedagogiikan alan tutkimuksissa on havaittu, että pyrkimyksistä huolimatta oppijat eivät välttämättä koe saaneensa palautetta vaan kokevat arvioinnin olevan summatiivisesti painottuvaa. Esimerkiksi Mäkipään ja Hildénin (2021) tutkimukseen osallistuneet lukiolaiset (N=160) kokivat kielen oppitunneilla saadun palautteen työnsä sisällöstä hyödylliseksi, mutta oppimisprosessia koskevaa palautetta oppijat eivät juuri arvostaneet. (Mäkipää & Hildén 2021.)

Kuten kaikessa arvioinnissa, myös kielitaidon ja kielenoppimisen arvioinnissa ovat läsnä valtasuhteet. Arvioinnilla on merkitystä yksilön osallisuuden kokemuksen ja yhteiskuntaan osallistumisen kannalta (ks. esim. Shohamy 2017). Omakohtainen osallistuminen arviointiin on tärkeää oppijan itseohjautuvuuden (Benson & Voller 2014) ja metakognitiivisten taitojen kehittymisen kannalta (Oxford 2016). Lisäksi arviointi vaikuttaa kielenoppijan motivaatioon (Dörnyei & Ushioda 2013). Oppimista tukeva arviointi tarkoittaa, että arviointi tehdään oppimisen aikana, esimerkiksi osana luokassa tapahtuvaa opetusta (Hill & McNamara 2012). Siihen kuuluu oleellisena osana vuorovaikutus arvioijan ja arvioitavan välillä – käytännössä usein opettajan ja opiskelijan välillä (Stobart & Eggen 2012). Arvioinnin pitäisi olla osa opetusta eikä erillinen tapahtuma, ja esimerkiksi opetuksen lomassa annettava palaute tulisi tiiviimmin integroida oppijoiden työskentelyprosessiin, mikä mahdollistaisi adaptiivisen, oppimisprosessia hyödyntävän ja hyödyttävän palautteenannon (vrt. Lantolf & Poehner 2014; Glisan ym. 2007; Leontjev 2014). Kieltenopetuksessa on laajasti käytössä monenlaisia kokonaisvaltaisia opetusmenetelmiä (ks. tarkemmin Maijala ym. 2023), joihin voi luonnollisesti yhdistää arviointia. Tällaisia ovat esimerkiksi tehtäväperustainen oppiminen (*task-based learning*; Nunan 2004) ja toimintaperusteinen opettaminen (*action-based teaching*;

van Lier 2007), jotka osaltaan tukevat kestäväen kehityksen periaatteita ja joihin on mahdollista yhdistää itse- ja vertaisarviointia oppimistehtävinä.

Kielten oppimiselle, opettamiselle ja arvioinnille yhteisen eurooppalaisen viitekehityksen (EVK; *The Common European Framework of Reference for Languages*) myötä kielitaidon arvioinnissa on keskitytty enemmän siihen, mitä oppija jo osaa, kuin siihen, missä on vielä puutteita (Euroopan neuvosto 2001). Viitekehitys on tuonut myös suomalaisen kielenopetukseen vahvemmin mukaan itsearvioinnin, joka auttaa kestäväen kehityksen periaatteiden mukaisesti oppijoita tarkastelemaan omaa oppimisprosessiaan (Beck ym. 2013; Cassidy 2007). Joissain kouluissa on käytössä oppimisen ja arvioinnin tukena myös *Eurooppalainen kielisalkku* (Opetushallitus 2014), ja uuden *Lukion opetussuunnitelman perusteiden* (Opetushallitus 2019) myötä jokaisen lukiolaisen edellytetään kokoavan oma kieliprofilinsa, joka ohjaa itsearviointiin ja oman oppimisprosessin suunnitteluun. Oppijat kokevat tällaisen osallistavan arvioinnin yksilöllisesti: Pollari (2017) havaitsi, että lukio-opiskelijoista osa kokee osallistavan arvioinnin hyödyllisenä, kun taas osa kokee sen kautta saamansa palautteen riittämättömäksi. Pollari (2017) korostaakin monipuolista arviointia, jotta jokaisella oppijalla olisi mahdollisuus hyötyä palautteesta.

Suomen kontekstissa esimerkiksi DIALANG-hankkeen yhtenä tavoitteena on ollut lisätä oppijan osallisuutta ja oman oppimisen ohjaamista eurooppalaiseen viitekehitykseen nojaavan kielitaidon itsearvioinnin avulla (Huhta 2019). Luukan ym. (2008) kyselytutkimuksen mukaan yhdeksäsluokkalaisten (N=1720) kokivat, että vertaisarviointia hyödynnettiin harvemmin kuin itsearviointia, eikä itsearviointiaakaan käytetty niin usein kuin opettajat raportoivat. Etenkin heikommin kielissä menestyvät oppilaat pitivät myös itsearviointitaitojaan heikkoina ja kokivat, ettei itsearviointi vaikuttanut päättöarvosanaan. (Luukka ym. 2008.) Lukiolaisten (N=255) palautekäsityksiä tarkastelevassa tutkimuksessaan Mäkipää (2020) toteaa, että ruotsin kielen kursseilla opiskelijat, joiden kielitaito oli hyvä, kokivat hyötyvänsä opettajan palautteesta ja myös saavansa sitä enemmän kuin opiskelijat, joiden kielitaito oli muita heikompi. Englannin kursseilla ei tällaista eroa opiskelijoiden välillä ilmennyt, mutta ruotsin osalta palautteesta näyttivät hyötyvän ne, jotka jo osasivat kieltä hyvin (Mäkipää 2020).

Itsearvioinnin lisäksi kielenoppimisessa on tutkittu viime aikoina dynaamista arviointia, jossa oppija saa arvioinnin edetessä yksilöllistä korjaavaa palautetta ja pystyy muokkaamaan suoritustaan sen perusteella. Tällainen adaptiivinen korjaava palaute heijastelee sosiokulttuurista oppimiskäsitystä ja Vygotskyn ajatusta lähikehityksen vyöhykkeestä: oppijan taitoihin mukautuva palaute tarjoaa riittävästi sekä tukea että haastetta, jotta oppija voi tehdä uudenlaisia valintoja ja kehittyä (Lantolf & Poehner 2014). Leontjev (2014) tarkastelee mukautuvan palautteen hyötyjä tutkimuksessa, johon osallistui 14-vuotiaita äidinkielenään venäjää puhuvia englannin opiskelijoita (N=64). Tulokset osoittivat, että adaptiivista korjaavaa palautetta saaneet oppijat hyötivät palautteesta ja paransivat suoritustaan (Leontjev 2014). Dynaaminen arviointi

näyttäisi tutkimusten mukaan sopivan myös kulttuurienvälisen oppimisen arviointiin (ks. esim. Despoteris & Anada 2018).

Myös lukio-opetuksen yleisissä arviointiperiaatteissa voidaan nähdä kestävyyskasvatuksen periaatteita. Lukiolain mukaan arvioinnin tulisi olla monipuolista ja kannustavaa ja sen pitäisi kehittää opiskelijan itsearviointitaitoja. Tämän lisäksi opiskelijalla tulee olla oikeus saada tietoa arviointiperusteista ja siitä, miten niitä hänen sovelletaan. (Lukiolaki 714/2018, 37 § 1 mom.) Lukiokoulutuksessa arvioinnin tavoitteena on yhtäältä tukea ja ohjata opiskelijana, minkä pitäisi toteutua palautteen sekä itse- ja vertaisarvioinnin avulla. Toisaalta arvioinnin tavoitteena on opiskelijan osaamisen näkyväksi tekeminen. Opiskelijan tulisi olla tietoinen periaatteista, joiden mukaan häntä arvioidaan, ja arvioinnin pitäisi olla osa lukion arvoja, normeja ja käytänteitä. Koulun arviointikulttuurin tulisi olla läpinäkyvää, ja sen periaatteiden ja käytänteiden toteutumista olisi seurattava. Arvioinnin kohteena eivät saa olla opiskelijan arvot, asenteet eivätkä henkilökohtaiset ominaisuudet vaan hänen tietonsa, taitonsa ja työskentelynsä. (Opetushallitus 2019: 45–46.)

Kestävää kehitystä edistetään, kun koulutusjärjestelmä tarjoaa oppijoille mahdollisuuden omaksua taitoja, jotka edistävät kestäväan kehitykseen läheisesti liittyviä moraalisia arvoja, ja siten auttavat heitä myönteisesti vaikuttamaan henkilökohtaisiin ja yhteisöllisiin muutoksiin (ks. esim. Asif ym. 2020). Oppijoiden kouluttaminen muutoksen tekijöiksi on tärkeä tulevaisuuden opettajien taito (Weinberg ym. 2020), eli opettajien pitäisi pystyä tukemaan lasten ja nuorten toimijuutta yhtenä tulevaisuuden keskeisistä taidoista (OECD 2017). Tutkimusten mukaan arviointi on kuitenkin yksi opettajuuden vaikeimmista tehtävistä (Shapiro-Lishchinsky 2011). Kieltenopettajat tarvitsevat koulutusta arviointiin liittyvissä asioissa, erityisesti sellaisessa arvioinnissa, jossa ei käytetä kokeita (Kelly ym. 2004; Tsagari & Vogt 2017). Arviointi saattaa muodostaa opettajalle moraalisen dilemman (Komorowska 2016), ja hän voi esimerkiksi pohtia, arvioiko kaikkia oppijoita oikeudenmukaisesti (vrt. ILTA Code of Ethics 2000). Opettajan moraalikäsitteiden voidaan taas katsoa kuuluvan osaksi opettajan eettistä identiteettiä, mikä on olennainen osa opettajan ammatillista kehittymistä (Miller ym. 2017). Opettajat tekevät opetukseen liittyviä päätöksiä (Borg 2003). He myös päättävät viime kädessä siitä, miten oppijoiden arviointi toteutetaan. Opetuksen arvoperustaa tutkittaessa onkin tärkeää tietää ja ymmärtää, miten opettajat ajattelevat, tietävät tai uskovat. Seuraavaksi tarkastelemme kyselyaineistomme avulla opettajien käsityksiä kestäväan kehityksen, arvioinnin ja kieltenopetuksen yhteyksistä.

4 Aineisto ja menetelmät

Tarkastelemme artikkelissa kestävyyskyselyä, joka teetettiin kieltenopettajille keväällä 2022. Kyselyyn kerättiin vastauksia kieltenopettajien täydennyskoulutusten yhteydessä. Kyselyn 31 vastaajasta 29 antoi yleisen tietosuoja-asetuksen (GDPR) mukaisen luvan

käyttää vastauksiaan tutkimuksessa. Kysely toteutettiin anonymisti, ja vastauksista poistettiin kaikki sellaiset tiedot, joista kävisi ilmi esimerkiksi vastaajien koulu tai kotipaikka. Vastajiin viitataan tässä tutkimuksessa tunnuksilla ID1–19. Aineisto säilytetään yliopiston suojatussa pilvipalvelussa. Vastaajien taustatiedot on koottu taulukkoon 1:

TAULUKKO 1. Opettajien kestävyyskyselyn 2022 taustatiedot vastaajamäärittäin.

Työkokemus opettajana	yli 10 vuotta	5–10 vuotta	alle 5 vuotta		
Vastauksia yht.	21	6	2		
Koulutusaste, jolla opetan	yläkoulu	lukio	alakoulu	vapaa sivistystyö	muu
Vastauksia yht.	13	9	7	4	4
Opetettavat kielet	englanti	ruotsi	saksa	suomi	muu
Vastauksia yht.	18	13	6	3	4

Enemmistö (21/29) kyselyyn vastaajista oli toiminut opettajana jo melko pitkään eli yli 10 vuotta. Vastaajat toimivat opettajina erityisesti yläkouluissa (13) ja lukioissa (9) mutta myös alakouluissa (7) ja muissa oppilaitoksissa (8). Tavallisimmat opetettavat kielet olivat englanti (18) ja ruotsi (13).

Taustatietokartoituksen jälkeen opettajat siirtyivät vastaamaan kysymyksiin, jotka käsittelivät kestävästä kehityksestä neljää ulottuvuutta (ks. esim. Laine ym. 2022). Aluksi heille esiteltiin lyhyesti, mitä ekologinen, sosiaalinen, taloudellinen ja kulttuurinen kestävyys ovat (ks. taulukko 2). Tämän taustatiedon pohjalta opettajia pyydettiin pohtimaan, miten ulottuvuudet ja niiden alle sijoittuvat aiheet, esimerkiksi kierrätyksen suosiminen tai sukupuolten välisen tasa-arvon edistäminen, toteutuvat omassa opetuksessa. Emme kuitenkaan tässä artikkelissa tarkastele kaikkia kyselyn vastauksia vaan keskitymme taulukon 2 kuvaamalla tavalla kahteen avokysymykseen, joista toisessa opettajia pyydettiin pohtimaan, miten arvioinnin kautta voidaan edistää kestävästä kehityksen tavoitteiden toteutumista kieltenopetuksessa (kysymys 14). Taustoitamme vastauksia myös tarkastelemalla edeltävää avokysymystä 13, jossa opettajat kuvailivat omin sanoin, millaiset käytänteet – esimerkiksi opetusmenetelmät tai toimintatavat – edistävät kestävästä kehityksestä periaatteita kielten tunneilla. Huomioimme analyysissä myös kyselyn viimeisen kohdan eli vapaan sanan vastaukset silloin, kun ne täydentävät kysymysten 13 ja 14 vastauksia.

TAULUKKO 2. Analyseissä käytetyt kysymykset taustatehtävineen, kysymystyyppit ja niihin liittyvät tutkimuskysymykset.

<p>Taustatehtävät: Kysymykset 6–9: Merkitse EKOLOGISEN/SOSIAALISEN/KULTTUURISEN/TALOUDELLISEN kestävyiden aihepiireistä ne, joita olet jotenkin voinut huomioida omassa opetuksessasi. Voit valita niin monta aihepiiriä kuin haluat. Kysymys 10: Missä määrin olet samaa tai eri mieltä seuraavien väittämien kanssa? Kysymys 11: Mitkä asiat vaikeuttavat kestävä kehityksen aihepiirien käsittelemistä kieltenopetuksessa?</p>		
	Kysymystyyppi	Kysymykseen liittyvä tutkimuskysymys
A) Kyselyn kysymykset		
13. Millaiset käytänteet (esimerkiksi opetusmenetelmät tai toimintatavat) edistävät kestävä kehityksen periaatteita omilla tunteillasi?	Avoin vastausmahdollisuus, kysymykset eivät olleet pakollisia	Miten opettajat huomioivat kestävä kehityksen periaatteita omassa opetuksessaan?
14. Miten arvioinnin kautta voidaan edistää kestävä kehityksen tavoitteiden toteutumista kieltenopetuksessa?		Miten opettajat tunnistavat arvioinnin yhteyden kestävyteen ja eettisyyteen?

Avokysymysten vastaukset analysoimme laadullisen analyysin NVivo-ohjelmistotyökalu avulla (Bazeley & Jackson 2013) siten, että käytänteitä kuvanneet vastaukset koodattiin neljään ryhmään ja arviointiin liittyvät vastaukset kolmeen ryhmään. Käytimme luokittelussa sekä deduktiivista eli teorialähtöistä sisällönanalyysiä että induktiivista eli aineistolähtöistä analyysiä (esim. Sarajärvi & Tuomi 2018; Vesalainen & Huhtala 2018). Käytänteitä kuvaavat vastaukset luokittelimme deduktiivisesti Heikkolan ym. (2021; ks. myös Majjala ym. tulossa) mallin avulla neljään luokkaan sen mukaan, miten opettajat huomioivat vastauksissaan käytänteiden ja kielen ja kulttuurin opettamisen yhteyden:

1. Vastauksessa on laaja yhteys kieltenopetukseen ja useita käytänteitä (K3).
2. Vastauksessa on aihetta laajempi yhteys kieltenopetukseen ja useita käytänteitä (K2).
3. Vastauksessa on kapea yhteys kieltenopetukseen vain aihetasolla (K1).
4. Vastauksessa ei ole yhteyttä kieltenopetukseen tai käytänteisiin (K0).

Arviointiin liittyvän avokysymyksen 14 vastaukset luokittelimme induktiivisesti kolmeen luokkaan sen mukaan, näkevätkö opettajat arvioinnissa ja kestävän kehityksen tavoitteissa yhteyttä vai eivät ja miten he yhteyttä luonnehtivat:

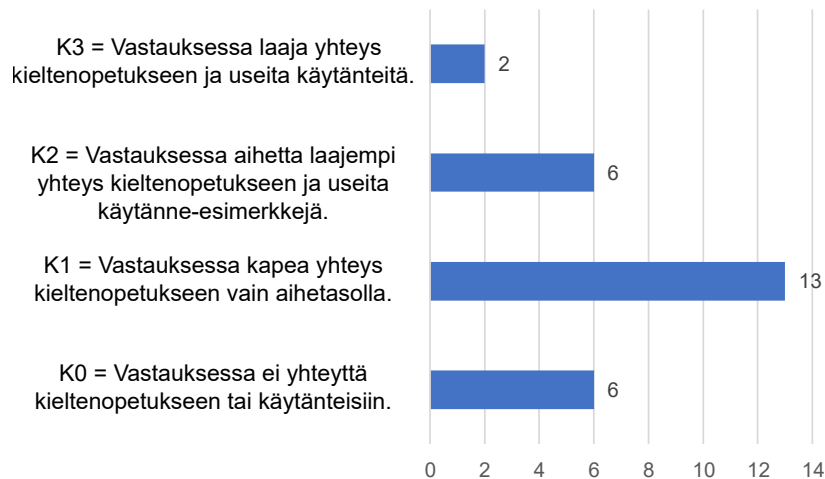
1. Kestävä kehitys on mukana arvioinnissa tehtävien ja sanastojen aiheina (1 AIHE).
2. Kestävä kehitys on mukana arvioinnissa yleisten tavoitteiden kautta (2 TAVOITE).
3. Kestävää kehitystä ja arviointia ei voi yhdistää tai opettajan työhön ei kuulu arviointia (3 EI YHTEYTTÄ).

Yksittäisen opettajan vastaus luokiteltiin pääasiassa yhteen luokkaan, mutta joissakin vastauksissa yhdistyivät esimerkiksi luokat 1 ja 2, jolloin ne koodattiin molempiin luokkiin. Kaksi tutkijaa teki molemmat luokittelut ensin itsenäisesti, ja sen jälkeen vertasimme ja yhtenäistimme luokittelut parantaaksemme tutkimuksen reliabeliutta.

5 Tulokset

5.1 Kestävyyden ulottuvuudet kielten oppitunneilla

Tarkastelemme ensin vastauksia kysymykseen 13: Millaiset käytänteet (esimerkiksi opetusmenetelmät tai toimintatavat) edistävät kestävän kehityksen periaatteita omilla tunneillasi? Kysymys ei ollut pakollinen, ja siihen vastasi 19 opettajaa.



KUVIO 1. Sisällönanalyyssissä käytetty teoriapohjainen luokittelu (Heikkola ym. 2021), sen selitykset sekä vastaajien (n=19) jakautuminen em. ryhmiin.

Kuviosta 1 käy ilmi, että vastaajista (n=19) kuusi ei yhdistänyt kestävästä kehitystä kieltenopetuksen käytänteisiin vaan mainitsi käytänteinä esimerkiksi kierrättämisen tai valojen sammuttamisen eli lähestyi kestävyyttä lähinnä luokkahuoneessa tapahtuvana ulkoisena toimintana, ei suoraan osana opetusta (K0), kuten esimerkeistä 1 ja 2 käy ilmi:

- (1) Oma esimerkki, esim. en roskaa enkä kopioi, käytän digiopetusta (ID02).
- (2) Perustoiminnot: luokassa valot + laitteet päältä pois kun ei tarvita, kouluun ja koulusta ulos liikutaan lihasvoimin joko pyörällä tai jalkaisin. Paperia ja muita välineitä käyteään[!], mutta ei tuhlata, roskat kierrätetään mahdollisuuksien mukaan. (ID08.)

Myös aiemmissa tutkimuksissa on todettu, etteivät kaikki kieltenopettajat tunne kestävyyskasvatusta omakseen eivätkä aina pidä tietopohjaansa riittävänä kestävyysaiheiden opettamiseen (Borg ym. 2012; Räcköläinen 2017). Opettajat ovat aiemmissa tutkimuksissa myös kokeneet, että kestävyysaiheiden käsittely vie aikaa aiheilta, jotka olisivat kieltenopetuksen kannalta keskeisempiä (ks. esim. Haukås ym. 2021). Enemmistö tämän kyselyn vastaajista eli 13 opettajaa kuitenkin kuvasi kestävä kehityksen ja opetuskäytänteiden välistä yhteyttä aiheiden kautta (K1). Kestävä kehitys näyttäytyi näissä vastauksissa oppikirjojen teemoina, opiskeltavina sanastoina tai kirjoitelmien, keskusteluharjoitusten tai projektien aiheina. Aihelähtöinen käsitteleminen on todettu tutuksi sekä opetuksessa että oppikirjoissa myös aiemmissa tutkimuksissa (esim. Mohammadnia & Moghadam 2019). Tähän on liittynyt mahdollisuus valita kieltenopetuksessa esiin nostettavat teemat melko vapaasti, ja siksi opettajilla on myös mahdollisuus ujuttaa opetukseen kestävä kehityksen aihepiirejä. Opetustilanteessa aiheiden valintaa ohjaavat kuitenkin hyvin pitkälle oppikirjojen sisällöt (ks. esim. Hahl 2020) ja niiden taustalla vaikuttava eurooppalainen viitekehys. Aiheena lähestymistä kuvaavat aineistoesimerkit 3–5:

- (3) Monissa kielten oppikirjoissa kestävä kehitys käsitellään jo jonkin verran. Olen myös itse tuonut aihetta esiin esimerkiksi uutisvideoiden, lehtiartikkeleiden ja sarjakuvien muodossa. (ID05.)
- (4) Mielestäni kielten opetuksessa tärkeintä on oppia kestävä kehitykseen liittyvä sanavarasto jotta oppilaat pystyvät keskustelemaan aiheesta ja ymmärtämään tekstiä ja puhetta jotka käsittelevät tätä asiaa. (ID07.)
- (5) Keskustelu on tärkeää ja ajankohtaiset uutiset. Aihe on oppilaille hyvin tärkeä. (ID17.)

Monet vastaajat liittivät kestävä kehityksen opetuskäytänteisiinsä myös laajemmin kuin esimerkiksi käsiteltävien tekstien ja kirjoitustehtävien aiheina: kahdeksassa vastauksessa (K2: 6 vastausta, K3: 2 vastausta) kestävä kehityksellä nähtiin esimerkiksi olevan yhteyksiä toimintakulttuuriin ja kieltenopetuksen rooliin (esimerkki 6), toimin-

nalliseen opetukseen (esimerkit 7 ja 8) ja itseilmaisun ja toimijuuden kehittymiseen (esimerkki 9):

- (6) Kestävä kehitys on tietysti sisältöaihe, toimintakulttuuri ja käytäntö. Mutta vieraat kielet ovat välttämätön keino tehdä kansainvälistä yhteistyötä asiasta. (ID01.)
- (7) Käytännön simulaatiotilanteet, leikki, tiedonhaku, kartan tutkiminen, virtuaalimat-kustaminen, frontaaliopetus asiasta, keskustelut luokassa. Videoiden näyttäminen teeman mukaisesti. (ID06.)
- (8) Luokkatyöskentely: pääosin keskustelemalla ja herättämällä aihepiiriin, tekemällä ryhmätöitä aiheeseen liittyen. Tiedonhakemisharjoituksin etsimällä tietoa eri aiheista ja herättämällä oppilaita ajattelemaan asioita. (ID08.)
- (9) Oman mielipiteen ilmaisun vahvistaminen ja perustelu. Oman toiminnan pohtiminen. Sen tulisi integroitua vielä vahvemmin kaikkeen opetukseen. Ei sen tarvitse silmille hyppiä, mutta voisi sitä korostaakin vielä erikseen materiaaleissa. (ID14.)

Esimerkeissä 7–9 tuodaan esille luokassa käytävät keskustelut, joiden voidaan katsoa olevan kestävän kehityksen sosiaalisen ulottuvuuden mukaisia osallistavia työtapoja (Sinakou ym. 2019a). Vastauksissa näkyvätkin kielenoppimiseen liittyvät erityispiirteet, kuten vuorovaikutuksellisuus (etenkin esimerkki 8) sekä kielten kulttuurinen ja historiallinen rooli yhteisössä (esimerkki 6) (ks. Cowley 2011). Vastauksissa ilmenee myös pyrkimys harjoitella demokratiataitoja vuorovaikutuksessa muiden kanssa käyttämällä kieltä tai kieliä (esimerkki 9). Demokratiataidot ovatkin osa sosiaalista kestävyyttä ja kuuluneet kieltenopetuksen tavoitteisiin jo pitkään (Byram 2014). Opettajat yhdistivät kestävää kehitystä monipuolisesti erilaisiin kielten opetuskäytänteisiin. Laajimmassa vastauksessa (esimerkki 10) kestävän kehityksen ja kielten yhteys nähtiin jopa ekosysteemin kaltaisena:

- (10) Tärkeää olisi pitää esillä kielten ekosysteemin hajoamista ja köyhtymistä muiden ekosysteemien rinnalla ja saada oppilaat ymmärtämään, että nykyinen, jopa vapaaehtoinen siirtyminen paikalliskielistä harvoinhin valtakieleihin on kulttuureille ja ihmisten hyvinvoinnille äärimmäisen tuhoisaa. (ID12.)

Vastauksessa näkyy systeemisen ajattelun (Lewis ym. 2014) ja kielenopetuksen mahdollinen yhteys (*kielten ekosysteemin hajoaminen ja köyhtyminen muiden ekosysteemien rinnalla*) (Ferguson ym. 2020; Gunina ym. 2021), mikä on välttämätöntä eri kestävyys-ulottuvuuksien välisten yhteyksien ymmärtämiseksi (Sinakou ym. 2019b). Tähän tutkimukseen osallistuneiden kieltenopettajien vastauksista ei kuitenkaan käynyt ilmi, että arviointia yhdistettäisiin opetuskäytänteisiin (vrt. Glisan ym. 2007). Kestävän kehityksen yhteyttä opiskelijan arviointiin ei myöskään käytänteiden yhteydessä mainittu, ja esimerkiksi palautteen antaminen puuttui vastauksista kokonaan. Seuraavaksi

tarkastelemme, millaisina opettajat näkivät arvioinnin ja kestävyysvälistä yhteyttä. Vaikka vastaukset eivät suoraan linkity toisiinsa, pohdimme arviointikysymyksen (kysymys 14) vastauksia myös edellä tekemiemme huomioiden pohjalta.

5.2 Opettajien näkemyksiä arvioinnin ja kestävyysvälistä yhteyksistä

Kyselyssä oli mukana myös suoraan arviointiin liittynyt avokysymys 14: Miten arvioinnin kautta voidaan edistää kestävä kehitys tavoitteiden toteutumista kieltenopetuksessa? Kysymys ei ollut pakollinen. Vastaajat (n=16) yhdistivät arvioinnin ja kestävyysvälistä seuraavasti:

1. AIHE, n=7 (kestävä kehitys on mukana arvioinnissa tehtävien ja sanastojen aiheina)
2. TAVOITE, n=3 (kestävä kehitys on mukana arvioinnissa yleisten tavoitteiden kautta)
3. EI YHTEYTTÄ, n=7 (kestävää kehitystä ja arviointia ei voi yhdistää tai opettajan työhön ei kuulu arviointia).

Luvussa 5.1 analysoidun kysymyksen tavoin isossa osassa vastauksista (n=7/16) kestävä kehitys yhdistettiin kielenopetukseen vain aiheiden kautta (1 AIHE) siten, että kestävä kehitys aihepiirit olivat koetehtävien tai esimerkiksi arvioitavien kirjoitelmien aiheita, kuten esimerkeissä 11–13:

- (11) Voitaisiin esim. kirjoittaa arvioitava aine kestävä kehitys teemoista (ID04).
- (12) vaikea opettajien toteuttaa, muuten kuin oppimateriaalien[!] sisältöjen kautta, esim. kokeissa (ID22).
- (13) Esim. tuomalla kestävä kehitys aiheita kurssien sisältöjen arviointiin[!] oppimistehtäviin (ID27).

Koska monet kyselyyn vastanneet opettajat näkivät kestävä kehitys tärkeänä aiheena mutta eivät pohtineet kestävyysvälistä toimintatapana eivätkä siten yhdistäneet sitä laajasti esimerkiksi sosiaaliseen ulottuvuuteen, myös arvioinnin kohdalla vastaukset painoutuivat aiheisiin. Tämä kysymys tai sitä edeltäneet kestävyysvälistä ulottuvuuksia taustoittaneet kysymykset eivät saaneet opettajia pohtimaan esimerkiksi tasapuolista arviointia tai arvioinnin eettisyyden, minäkuvan rakentumisen ja kielenoppimisen välisiä yhteyksiä.

Vaikka kyselyssä oli siis useiden kysymysten kautta luonnehdittu, mitä kestävyysvälistä ulottuvuudet ovat, iso osa vastaajista näki yhteyden vain aiheena ja vähän yllättäenkin yhtä iso osa vastaajista (n=7/16) ei nähnyt ulottuvuuksien pohjalta arvioinnilla ja kestävyysvälistä yhteyttä (esimerkit 14–17) tai ei kokenut, että omaan opettajan rooliin kuuluisi lainkaan arviointia (3 EI YHTEYTTÄ) (esimerkki 17):

- (14) Tähän en keksi mitään sanottavaa (ID3).

- (15) En usko, että tätä aihetta voi liittää arviointiin (ID29).
- (16) Mikä kysymys tämä on (ID13)?
- (17) Vapaassa sivistystyössä opiskelijoita ei arvioida (ID5).

Vastaajat eivät ylipäättään pohtineet vastauksissa arviointia kovinkaan laajasti. Tähän saattoi vaikuttaa se, että kysymys arvioinnista oli sijoitettu kyselyn loppupuolelle. Vastauksista on siksi vaikea päätellä, miten opettajat ymmärsivät arvioinnin. Ne vastaajat, jotka näkivät kestävästä kehityksestä liittyvän arvioinnin osalta vain tehtävien aiheisiin, näkivät arvioinnin erilaisten suoritusten arviointina eivätkä maininneet palautetta tai esimerkiksi arvioinnin oppimiseen kannustavaa merkitystä. Myös opettajat, jotka eivät nähneet arvioinnilla yhteyttä kestäväyteen, lähestyivät kysymystä osin aiheen kautta (esimerkki 15). Kaksi vastaajista totesi, etteivät he arvioi opiskelijoita vapaassa sivistystyössä (esimerkki 17). Näissä vastauksissa arviointi määriteltiin todennäköisesti kokeiksi tai arvosanojen antamiseksi, ei esimerkiksi palautteeksi onnistumisesta, sillä laajasti ymmärrettynä opettaja toimii monissa luokkatilanteissa arvioijana, vaikkei olisikaan kyse esimerkiksi koetilanteesta.

Laajasti kielenopetukseen tavoitteisiin (2 TAVOITE) kestävä kehitys ja arviointi yhdistettiin vain kolmessa vastauksessa. Niissä arviointia ei kuvattu tarkasti vaan todettiin, että opetussuunnitelma velvoittaa huomioimaan kestävästä kehityksestä yleisissä tavoitteissa ja siten se tulee mukaan myös arviointiin, kuten esimerkissä 18:

- (18) Laaja-alaisena osaamistavoitteena. Opsit velvoittaa jo, - - (ID22).

Kyselyn perusteella vaikuttaa siltä, etteivät kielenopettajat näe arvioinnilla ja kestävyyskasvatuksella juurikaan muuta yhteyttä kuin toimia teksti-, sanasto- ja aihepankkina. Vaikka vastauksissa tuotiin käytänteinä esille työskentelytapoja, joissa kestävyysaiheita voidaan käyttää, sekä joitain laajoja toimintatapoihin liittyviä huomioita, niitä ei enää pohdittu arviointikysymyksen yhteydessä. Kiinnostavaa kuitenkin on, että aihetta painottavista vastaajista vain kolme yhdisti kestävästä kehityksestä sekä opetukseen että arviointiin aiheina ja siis painotti aihetta sekä kysymyksessä 13 että kysymyksessä 14. Suurin osa vastaajista mainitsi aiheen käytänteiden kohdalla muttei enää uudelleen arviointikysymyksessä. Arvioinnin eettiset näkökulmat (ILTA Code of Ethics 2000; Miller ym. 2017; Komorowska 2016) tai arvioinnin merkitys oppijan tulevan elämän kannalta (Boud & Soler 2016) eivät tulleet esille tähän tutkimukseen osallistuneiden opettajien vastauksista.

6 Yhteenveto ja pohdintaa

Tämän artikkelin tavoitteena oli löytää yhtymäkohtia kieltenopettajien opetus- ja arviointikäytänteistä ja kestävästä kehityksen tavoitteista. Pienimuotoisen tutkimuksemme tulokset viittaavat siihen, että kieltenopettajat eivät yhdistä arviointia kestävästä kehityksen periaatteiden mukaisiin opetuskäytänteisiin. Mielenkiintoista oli, että osa vastaajista mainitsi, ettei arviointi ylipäättään kuulunut omiin opetuskäytänteisiin. Kyselyn tulosten perusteella näyttää muutenkin siltä, että kieltenopettajat kokevat arvioinnilla ja kestävyyskasvatuksella olevan vain vähän yhtymäkohtia. He näyttävät pitävän kestävästä kehityksestä lähinnä aihepankkina, josta voidaan valita sopivia tekstejä ja sanastoa opetukseen. Opettajat kyllä mainitsivat opetuskäytänteiden yhteydessä myös sellaisia työskentelytapoja, joissa kestävyysaiheita ja kestävyyskasvatuksen mukaisia työskentelytapoja voidaan käyttää holistisesti (vrt. Sinakou 2019a; Sinakou 2019b), mutta niitä ei enää otettu esille arviointikysymyksen yhteydessä. Opettajien vastauksista käy kuitenkin ilmi, että he kiinnittävät huomiota arvioinnin avoimuuteen ja ennakoitavuuteen, mikä taas on kestävästä oppimisesta tukevan arvioinnin periaatteiden mukaista (vrt. Boud & Soler 2016). Opettajien käsitysten mukaan vertaisarvioinnilla ja -palautteella on tärkeä sosiaalinen tehtävä kieltenopetuksessa, minkä voidaan puolestaan katsoa lisäävän kestävyyskasvatuksen mukaista oppijoiden osallistumista sekä koulussa että koulun ulkopuolella yhteiskunnassa (vrt. Sinakou ym. 2019a). Myös tässä korostuu se, että kieltenopettajien vahvuutena kestävästä kehityksestä mukaisessa opetuksessa on nimenomaan sosiaaliseen ulottuvuuteen kuuluvien taitojen edistäminen (vrt. Borg 2006; Mohammadnia & Moghadam 2019).

Tuloksista kävi ennen kaikkea ilmi se, että kestävästä kehityksestä periaatteiden mukaisen arvioinnin toteutumisessa ja tiedostamisessa kieltenopetuksessa on vielä haasteita siksi, etteivät opettajat tiedosta hyvän arvioinnin ja kestävästä kehityksestä yhteyttä. Jotta kestävästä kehityksestä periaatteiden mukainen arviointi voisi toteutua, olisi kiinnitettävä huomiota kielten oppimateriaalien sisältöihin ja myös kokeissa testattaviin tietoihin ja taitoihin. Vaikka tässä tutkimuksessa kiinnitettiin huomiota kestävästä kehityksestä kaikkiin neljään ulottuvuuteen ja niiden yhteyksiin kieltenopetuksessa, kävi selkeästi ilmi, että opettajat ottivat vastauksissaan esille erityisesti kestävästä kehityksestä ”pehmeiden” eli sosiaalisen ja kulttuurisen ulottuvuuksien mukaisia teemoja, kuten demokratia ja tasa-arvo. Pohdinnan arvoista on, mitä opettajat mieltävät arvioinniksi ja miksi arviointi koettiin erillisenä osana eikä kuuluvaksi opetuskäytänteisiin. Arviointitaitojen kehittyminen on haaste kielten opettajien koulutukselle ja myös arviointiin liittyvää kielten opettajien täydennyskoulutusta tarvitaan.

Tutkimus myös osoitti sen, että kieltenopettajien koulutuksessa on entistä enemmän korostettava arvioinnin eettistä näkökulmaa (Miller ym. 2017; ILTA Code of Ethics 2000) ja yhteyttä kestävästä kehityksestä periaatteisiin. Tämän lisäksi koulutuksessa on kiinnitettävä huomiota siihen, millainen vaikutus arvioinnilla on yksilön tulevaisuuteen (Boud & Soler 2016) sekä siihen, miten arviointi edistää yksilön osallisuutta yhteis-

kunnassa (Shohamy 2017). Arviointikäytännöillä, kuten palautteenantamisella, on tässä tärkeä rooli (Glisan ym. 2007; Carless ym. 2011). Opettajat tarvitsevatkin kyselyn perusteella täydennyskoulutusta ja materiaaleja, joiden avulla he voivat reflektoida omaa arvioimistaan ja saada välineitä tehdä arviointia eettisesti ja kestävästi kehityksen tavoitteiden mukaisesti niin, että arviointi edistää oppimista ja antaa oppijalle myös tärkeitä tulevaisuustaitoja.

Kirjallisuus

- Al Amin, M. & J. Greenwood 2018. The UN sustainable development goals and teacher development for effective English teaching in Bangladesh: a gap that needs bridging. *Journal of Teacher Education for Sustainability* 20 (2), 118–138.
- Asif, T., O. Guangming, M. A. Haider, J. Colomer, S. Kayani & N. U. Amin 2020. Moral education for sustainable development: comparison of university teachers' perceptions in China and Pakistan. *Sustainability*, 12 (7), 3014. <https://doi.org/10.3390/su12073014> www.mdpi.com/journal/
- Bazeley, P. & K. Jackson 2013. *Qualitative data analysis with NVivo*. Second Edition. SAGE, London.
- Beck, R. J., W. F. Skinner & L. A. Schwabrow 2013. A study of sustainable assessment theory in higher education tutorials. *Assessment & Evaluation in Higher Education*, 38 (3), 326–348.
- Benson, P. & P. Voller 2014. *Autonomy and independence in language learning*. Abingdon, Oxon: Routledge.
- Black, H. D. 1986. Assessment for learning. Teoksessa D. L. Nuttall (toim.), *Assessing educational achievement*. London: Falmer, 7–18.
- Borg, S. 2003. Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language teaching*, 36 (2), 81–109.
- Borg, S. 2006. The distinctive characteristics of foreign language teachers. *Language Teaching Research*, 10 (1), 3–31.
- Borg, C., N. Gericke, H.-O. Höglund & E. Bergman 2012. The barriers encountered by teachers implementing education for sustainable development: discipline bound differences and teaching traditions. *Research in Science and Technological Education*, 30 (2), 185–207. <https://doi.org/10.1080/02635143.2012.699891>
- Boud, D. 2000. Sustainable assessment: rethinking assessment for the learning society. *Studies in continuing education*, 22 (2), 151–167.
- Boud, D. & R. Soler 2016. Sustainable assessment revisited. *Assessment & Evaluation in Higher Education*, 41 (3), 400–413.
- Butler, Y. G. 2016. Self-assessment of and for young learners' foreign language learning. Teoksessa M. Nikolov (toim.) *Assessing young learners of English: global and local perspectives*. Springer, Cham, 291–315.
- Byram, M. 2009. Evaluation and/or Assessment of Intercultural Competence. Teoksessa A. Hu & M. Byram (toim.) *Interkulturelle Kompetenz und fremdsprachliches Lernen: Modelle, Empirie, Evaluation. Intercultural competence and foreign language learning: models, empiricism, assessment*. Tübingen: Narr, 215–234.
- Byram, M. 2014. Twenty-five years on – from cultural studies to intercultural citizenship. *Language, culture and curriculum*, 27 (3), 209–225.
- Carless, D., D. Salter, M. Yang & J. Lam 2011. Developing sustainable feedback practices. *Studies in higher education*, 36 (4), 395–407.

- Cassidy, S. 2007. Assessing 'inexperienced' students' ability to self-assess: exploring links with learning style and academic personal control. *Assessment & Evaluation in Higher Education*, 32 (3), 313–330.
- Despoteris, J. & K. Ananda 2018. Intercultural Competence: reflecting on daily routines. Teoksessa M. Wagner, D. C. Perugini, & M. Byram (toim.) *Teaching intercultural competence across the age range: from theory to practice*. Bristol: Multilingual Matters, 88–93.
- Didham, R. J. 2018. *Programme on education for sustainable development. Education for sustainable development and the SDGs. Policy brief: advancing ESD policy*. Paris United Nations Educational, Scientific and Cultural Organization.
- Dörnyei, Z. & E. Ushioda 2013. *Teaching and researching motivation*. Abingdon, Oxon: Routledge.
- Euroopan neuvosto 2001. Common European framework of reference for languages: learning, teaching, assessment. Cambridge: Cambridge University Press. <https://rm.coe.int/1680459f97>
- Evans, N. (Snowy), R. B. Stevenson, M. Lasen, J.-A. Ferreira & J. Davis 2017. Approaches to embedding sustainability in teacher education: a synthesis of the literature. *Teaching and Teacher Education*, 63, 405–417.
- Fastré, G. M., M. R. Van der Klink, D. Sluijsmans & J. J. van Merriënboer 2013. Towards an integrated model for developing sustainable assessment skills. *Assessment & Evaluation in Higher Education*, 38 (5), 611–630.
- Ferguson, T., C. Roofe & L. D. Cook 2021. Teachers' perspectives on sustainable development: the implications for education for sustainable development. *Environmental Education Research*, 27 (9), 1–17. <https://doi.org/10.1080/13504622.2021.1921113>
- de la Fuente, M. J. (Ed.) 2021. *Education for sustainable development in foreign language learning: content-based instruction in college-level curricula*. Routledge Research in Language Education. New York: Routledge.
- Gericke, N., L. Huang, M.-C. Knippels, A. Christodoulou, F. Van Dam & S. Gasparovic 2020. Environmental citizenship in secondary formal education: The importance of curriculum and subject teachers. Teoksessa A. Ch. Hadjichambis, P. Reis, D. Paraskeva-Hadjichambi, J. Činčera, J. Boeve-de Pauw, N. Gericke, & M.-C. Knippels (toim.) *Conceptualizing environmental citizenship for 21st century education, vol. 4*. Cham: Springer International Publishing, 193–212.
- Giddings, B., B. Hopwood & G. O'Brien 2002. Environment, economy and society: fitting them together into sustainable development. *Sustainable development*, 10 (4), 187–196.
- Glisan, E. W., D. Uribe & B. Adair-Hauck 2007. Research on integrated performance assessment at the post-secondary level: student performance across the modes of communication. *Canadian Modern Language Review*, 64 (1), 39–67.
- Gunina, N., T. Mordovina & I. Shelenkova 2021. Integrating sustainability issues into English language courses at university. Teoksessa *E3S Web of Conferences* (Vol. 295, p. 05006). EDP Sciences. <https://doi.org/10.1051/e3sconf/202129505006>
- Hahl, K. 2020. Kulttuurin muuttunut olemus opetussuunnitelmassa ja vieraiden kielten oppikirjoissa. Teoksessa R. Hilden & K. Hahl (toim.), *Kielididaktiikan katse tulevaisuuteen: Haasteita, mahdollisuuksia ja uusia avauksia kielten opetukseen*. Ainedidaktisia tutkimuksia 17. Helsinki: Helsingin yliopisto, kasvatustieteellinen tiedekunta, 173–201.
- Halliday, M. 1993. Towards a language-based theory of learning. *Linguistics and Education*, 5, 93–116.
- Hammadou, J. & E. B. Bernhardt 1987. On being and becoming a foreign language teacher. *Theory into practice*, 26 (4), 301–306.

- Haukås, Å., S. Mercer & A. M. L. Svalberg 2021. School teachers' perceptions of similarities and differences between teaching English and a non-language subject. *TESOL quarterly* 56 (2), 441–858. <https://doi.org/10.1002/tesq.3062>
- Heikkola, L. M., E. Repo & N. Kekki 2021. Kielen rooli oppimisessa – tulevien aineenopettajien käsityksiä. Teoksessa L. M. Heikkola, G. Paulsen, K. Wojciechowicz & J. Rosenberg (toim.) *Språkets funktion. Juhlakirja Urpo Nikanteen 60-vuotispäivän kunniaksi. Festschrift i äran av Urpo Nikannes 60 års dag. Festschrift in honor of Urpo Nikannes 60th birthday*. Turku: Åbo Akademis Förlag, 330–351.
- Hill, K. & T. McNamara 2012. Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29 (3), 395–420.
- Hofman-Bergholm, M. 2018. Changes in thoughts and actions as requirements for a sustainable future: a review of recent research on the Finnish educational system and sustainable development. *Journal of Teacher Education for Sustainability*, 20 (2), 19–30.
- Huhta, A. 2019. Understanding self-assessment – what factors might underlie learners' views of their foreign language skills? Teoksessa A. Huhta, G. Erickson & N. Figueras (toim.) *Developments in Language Education: a memorial volume in honour of Sauli Takala*. Jyväskylä: University of Jyväskylä/Centre for Applied Language Studies and EALTA, 131–146.
- Inha, K., S. Laiho, & P. Mattila 2020. Kieli luo osallisuutta ja oikeuksia omaan oppimiseen ja oppijuuteen. *Kieli, koulutus ja yhteiskunta*, 11 (6). <https://www.kieliverkosto.fi/journals/kieli-koulutus-ja-yhteiskunta-lokakuu-2020/kieli-luo-osallisuutta-ja-oikeuksia-omaan-oppimiseen-ja-oppijuuteen>
- ILTA Code of Ethics 2000; <https://www.iltaonline.com/page/CodeofEthics> [luettu 22.6.2022].
- Kelly, M., M. Grenfell, R. Allan, C. Kriza & W. McEvoy 2004. *European profile for language teacher education: a frame of reference. Final report. A report to the European Commission Directorate General for Education and Culture*. Southampton: University of Southampton.
- Komorowska, H. 2016. Dilemmas in language teaching and teacher education. *Glottodidactica. An International Journal of Applied Linguistics*, 43 (1), 81–95.
- Kuhlman, T. & J. Farrington 2010. What is sustainability? *Sustainability*, 2 (11), 3436–3448. <https://doi.org/10.3390/su2113436>
- Laherto, A. 2020. Tiedekasvatuksen muuttuvat tavoitteet: luonnontieteellisestä lukutaidosta kestävyyskasvatukseen, toimijuuteen ja tulevaisuusajatteluun. *Ainedidaktikka*, 4 (3), 41–63.
- Laine, P., S.-R. Kuusalu, M. Majjala & M. Mutta 2022. Kestävä kehitys kieltenopetuksessa: kieliaineiden opettajaopiskelijoiden käsityksiä kestävyysaiheista kielten oppitunneilla. Teoksessa T. Seppälä, S. Lesonen, P. Ikkänen & S. D'hondt (toim.) *Kieli, muutos ja yhteiskunta – language, change and society*. AFinLA:n vuosikirja: Suomen soveltavan kielitieteen yhdistyksen julkaisu n:o 79. Jyväskylä: Suomen soveltavan kielitieteen yhdistys AFinLA, 97–120.
- Lantolf, J. P. & M. E. Poehner 2014. *Sociocultural Theory and the Pedagogical Imperative in L2 Education*. London: Routledge.
- Leontjev, D. 2014. The Effect of Automated Adaptive Corrective Feedback: L2 English questions. *Apples - Journal of Applied Language Studies*, 8 (2), 43–66. <https://apples.journal.fi/article/view/97864>
- Liddicoat, A. J. & A. Scarino 2010. Eliciting the intercultural in foreign language education at school. Teoksessa A. Paran & L. Sercu (toim.) *Testing the untestable in language education*, 52–76. Bristol: Multilingual Matters.
- Lukiolaki 714/2018. Opetus- ja kulttuuriministeriö. <https://www.finlex.fi/fi/laki/smur/2018/20180714>

- Luukka, M.-R., S. Pöyhönen, A. Huhta, P. Taalas, M. Tarnanen & A. Keränen 2008. *Maailma muuttuu - mitä tekee koulu? Äidinkielen ja vieraiden kielten tekstikäytännöt koulussa ja vapaa-ajalla*. Jyväskylä: Jyväskylän yliopisto, Soveltavan kielentutkimuksen keskus.
- Maijala, M. 2013. Merkittävät kulttuuriset sisällöt vieraan kielen aikuisopetuksessa – oppikirjojen ja opiskelijoiden näkökulmat. Teoksessa E. Yli-Panula, A. Virta & K. Merenluoto (toim.) *Oppiminen, opetus ja opettajaksi kasvu ainedidaktisen tutkimuksen valossa*. Turun ainedidaktisen symposiumin esityksiä 11.2.2011, 176–186. Turun yliopisto.
- Maijala, M., N. Gericke, S.-R. Kuusalu, L. M. Heikkola, M. Mutta, K. Mäntylä & J. Rose 2023. Conceptualising transformative language teaching for sustainability and why it is needed. *Environmental Education Research*. <https://doi.org/10.1080/13504622.2023.2167941>
- Maijala, M., L. M. Heikkola, P. Laine, M. Mutta, K. Mäntylä, J. Rose & V. Vaakanainen 2021. Eettisyys ja kestävä kehitys kielten opetuksen ja kielten opettajien koulutuksen tulevaisuuden haasteena. *Kieli, koulutus ja yhteiskunta*, 12 (2). <https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-maaliskuu-2021/eettisyys-ja-kestava-kehitys-kielten-opetuksen-ja-kielten-opettajien-koulutuksen-tulevaisuuden-haasteena>
- Maijala M., L. M. Heikkola, S.-R. Kuusalu, P. Laine, M. Mutta & K. Mäntylä (tulossa). Pre-service language teachers' perceptions of sustainability and its implementation in language teaching.
- Meletiadou, E. 2022. Learners' perceptions of peer assessment: implications for their willingness to write in an EFL context. *International Journal of Teacher Education and Professional Development (IJTEPD)*, 5 (1), 1–14.
- Miller, E. R., B. Morgan & A. L. Medina 2017. Exploring language teacher identity work as ethical self-formation. *The Modern Language Journal*, 101 (S1), 91–105.
- Mohammadnia, Z. & F. D. Moghadam 2019. Textbooks as resources for education for sustainable development: a content analysis. *Journal of teacher education for sustainability*, 21 (1), 103–114.
- Monroe, M. C., A. Oxarart & R. R. Plate 2013. A role for environmental education in climate change for secondary science educators. *Applied Environmental Education & Communication* 12 (1), 4–18.
- Murillo-Vargas, G., C. H. Gonzalez-Campo & D. I. Brath 2020. Mapping the integration of the sustainable development goals in universities: is it a field of study? *Journal of Teacher Education for Sustainability*, 22 (2), 7–25.
- Mäkipää, T. 2020. Feedback practices in language classes in Finnish general upper secondary schools. *Apples - Journal of Applied Language Studies*, 14 (1), 103–123.
- Mäkipää, T. & R. Hildén 2021. What kind of feedback is perceived as encouraging by Finnish general upper secondary school students? *Education Sciences*, 11 (12), 1–15.
- OECD 2017. *Better policies for better lives. Measuring distance to the SDG targets. An assessment of where OECD countries stand*. Paris, France: OECD.
- OKKA-säätiö 2022. *Kestävän kehityksen teemat*. <https://kouluajaymparisto.fi/tukea-kestavan-kehityksen-tyohon/kestavan-kehityksen-teemat/> [luettu 14.8.2022].
- O'Malley, J. M. & A. U. Chamot 1990. *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Opetushallitus 2014. *Eurooppalainen kielisalkku*. <http://kielisalkku.edu.fi/fi> [luettu 14.8.2022].
- Opetushallitus 2014. *Perusopetuksen opetussuunnitelman perusteet*. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/perusopetuksen-opetussuunnitelman-perusteet>.
- Opetushallitus 2019. *Lukion opetussuunnitelman perusteet*. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/lukion-opetussuunnitelmien-perusteet>.

- Opetushallitus 2020. *Oppilaan oppimisen ja osaamisen arviointi perusopetuksessa. Perusopetuksen opetussuunnitelman perusteiden 2014 muutokset*. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/arviointiluku-6-oppimisen-ja-osaamisen-arviointi> [luettu 14.8.2022].
- Opetushallitus 2022. *Varhaiskasvatussuunnitelman perusteet 2022*. <https://www.oph.fi/fi/koulutus-ja-tutkinnot/varhaiskasvatussuunnitelmien-perusteet>.
- Oxford, R. L. 2016. *Teaching and researching language learning strategies: Self-regulation in context*. Abingdon, Oxon: Routledge.
- Paran, A. & L. Sercu (toim.) 2010. *Testing the untestable in language education*. Bristol: Multilingual Matters.
- Pollari, P. 2017. *(Dis)empowering assessment? Assessment as experienced by students in their upper secondary school EFL studies*. Jyväskylä: Jyväskylän yliopisto. <http://urn.fi/URN:ISBN:978-951-39-7178-6>
- Räkköläinen, M., J. Metsämuuronen & J. Kiesi 2017. *Kestävän kehityksen osaaminen, opetus ja koulutuksen järjestäjän toiminta ammatillisissa perustutkinnoissa*. Helsinki: Kansallinen koulutuksen arviointikeskus, julkaisut 12: 2017.
- Sarajärvi, A. & J. Tuomi 2018. *Laadullinen tutkimus ja sisällönanalyysi*. Helsinki: Tammi.
- Shapiro-Lishchinsky, O. 2011. Teachers' critical incidents. Ethical dilemmas in teaching practice. *Teaching and Teacher Education*, 27, 648–656.
- Shohamy E. 2017. Critical language testing. Teoksessa E. Shohamy, I. Or & S. May (toim.) *Language testing and assessment. Encyclopedia of language and education* (3. painos). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-02261-1_26
- Sinakou, E., J. Boeve-de Pauw & P. Van Petegem 2019a. Exploring the concept of sustainable development within education for sustainable development: implications for ESD research and practice. *Environment, Development and Sustainability*, 21 (1), 1–10. <https://doi.org/10.1007/s10668-017-0032-8>
- Sinakou, E., V. Donche, J. Boeve-de Pauw, & P. Van Petegem 2019b. Designing powerful learning environments in education for sustainable development: a conceptual framework. *Sustainability*, 11 (21), 5994. <https://doi.org/10.3390/su11215994>
- Stobart, G., & T. Eggen. 2012. High-stakes testing – value, fairness and consequences. *Assessment in education: principles, policy & practice*, 19 (1), 1–6.
- Sund, P. & N. Gericke 2020. Teaching contributions from secondary school subject areas to education for sustainable development – a comparative study of science, social science and language teachers. *Environmental Education Research*, 26 (6), 772–794. <https://doi.org/10.1080/13504622.2020.1754341>
- Tani, S. & E. Aarnio-Linnanvuori 2020. Ympäristö- ja kestävyyskasvatus osana ainedidaktiikkaa. *Ainedidaktiikka*, 4 (3), 1–3.
- Tossavainen, H. 2016. Kielitestiä eettisyydestä ja oikeudenmukaisuudesta. Huhta, A. & R. Hildén (toim.), *Kielitaidon arviointitutkimus 2000-luvun Suomessa*. AFInLA-e. Soveltavan kielitieteen tutkimuksia 9, 27–43. <http://journal.fi/afinla/article/view/60845>
- 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.
- Uitto, A. & S. Saloranta 2017. Subject teachers as educators for sustainability: a survey study. *Education Sciences*, 7 (1). <https://doi.org/10.3390/educsci7010008>.
- UNESCO 2018. *Programme on Education for Sustainable Development. Education for sustainable development and the SDGs*. Policy brief: Advancing ESD policy. Paris, France: United Nations Educational, Scientific and Cultural Organization.

- UNESCO 2021. *Learn for our Planet. A global review of how environmental issues are integrated in education*. Paris: United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000377362> [luettu 4.7.2021].
- Vesalainen, M. & A. Huhta 2018. Arviointi muutoksessa myös kielenopetuksessa? Opettajien ja opettajaksi opiskelevien käsityksiä arvioinnista. *Yliopistopedagogiikka*, 25 (1). <https://lehti.yliopistopedagogiikka.fi/2018/11/30/arviointi-muutoksessa-myo-kielenopetuksessa-opettajien-ja-opettajaksi-opiskelevien-kasityksia-arvioinnista/>
- Wagner, M., D. C. Perugini & M. Byram (toim.) 2018. *Teaching intercultural competence across the age range: From theory to practice*. Bristol: Multilingual Matters.
- Wals, A. E. 2009. *Review of contexts and structures for education for sustainable development: 2009*. Paris: United Nations Educational, Scientific and Cultural Organization.
- Weiss, M., M. Barth & H. von Wehrden 2021. The patterns of curriculum change processes that embed sustainability in higher education institutions. *Sustainability Science*, 16 (5), 1579–1593.
- WCED = World Commission on Environment and Development 1987. *Our Common Future*. New York: Oxford University Press.
- Zygmunt, T. 2016. Language Education for Sustainable Development. *Discourse and Communication for Sustainable Education*, 7(1), 112–124. <https://doi.org/10.1515/dcse-2016-0008>

Mäkipää, T., R. Hilden & A. Huhta (toim.) 2023. Kielenoppimista tukeva arviointi – Assessment for supporting language learning. AFinLA-teema / n:o 15, 184–201.

**Merja Kauppinen, Mirja Tarnanen, Reeta Neittaanmäki,
Eija Aalto, Mari Hankala, Johanna Kainulainen,
Sanna Mustonen & Päivi Torvelainen**
Jyväskylän yliopisto

Diagnostinen arviointi suomen kielen ja kirjallisuuden pedagogiikan opintojakson opiskelua ja opetusta ohjaamassa

Nostot

- Alkutestin ja itsearvioinnin yhdistelmä auttaa luokanopettajaopiskelijaa hahmottamaan kielellistä osaamistaan.
- Oman kielellisen osaamisen tunnistaminen auttaa luokanopettajaopiskelijaa opiskelunsa suuntaamisessa ja opettajaa oppimisprosessien ohjaamisessa.
- Suomen kielen ja kirjallisuuden pedagogiikkaa pitäisi opettajankoulutuksessa suunnata myös osaamisen ja työskentelyn arviointiin.



VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

Abstract

The article deals with diagnostic assessment conducted in the Finnish language and literature course for classroom teacher students. The research context was a hybrid learning environment, where the students assessed both their linguistic and pedagogical competence using a self-assessment tool. The research questions were 1) what kind of linguistic competence the students showed in the beginning of the course, 2) how the results of linguistic competence and self-assessment were related, and 3) how the final grade of Finnish language and literature and the matriculation examination test of Finnish predicted the results of a diagnostic test. The data of the pilot research consisted of the results of a diagnostic test of 127 students and their self-assessment based on a five-step criterion-referenced scale. The research was quantitative and statistical methods were used in the analysis. The results showed that there were statistically reliable connections between the results of the self-assessment and the diagnostic test. With the help of the test, it is possible to assess different areas of students' linguistic competence to some extent reliably.

Keywords: assessment, teacher education, learning environment,
Finnish language and literature

Asiasanat: arviointimenetelmät, opettajankoulutus, oppimisympäristö,
suomen kieli ja kirjallisuus

1 Johdanto

Opettajankoulutuksella on keskeinen tehtävä koulun arviointikulttuurin ja opettajien arviointiosaamisen kehittämisessä (esim. Xu & Brown 2016; Atjonen ym. 2019). Arviointikulttuurilla tarkoitetaan arviointiin liittyviä uskomuksia, arvoja ja käsityksiä, jotka tiedostettuina tai tiedostamatta vaikuttavat yhteisön arviointikäytänteisiin ja -menetelmiin sekä niiden valintaan (esim. Fuller & Skidmore 2014). Arviointikulttuurin yhtenä haasteena pidetään sitä, että arviointia tehdään opiskelijoille sen sijaan, että sitä kehitettäisiin heidän kanssaan (Wanner & Palmer 2018). Onkin esitetty, että arviointikulttuurin muuttaminen onnistuu parhaiten, jos opiskelijoiden aktiivista, yksilöllistä oppimista ja itsearviointia kehitetään rinnakkain (Boud & Molloy 2013; Wanner & Palmer 2018).

Tällä vuosituhanella arviointiin, kuten myös opetukseen ja oppimiseen, on vaikuttanut oppimisympäristöjen digitalisoituminen (ks. esim. Xu & Brown 2016; Atjonen ym. 2019), mikä on huomioitava myös opettajankoulutuksen kehittämisessä ja opettajaopiskelijoiden oppimisen ohjaamisessa (Timmis ym. 2015). Digitalisoitumisen myötä on syntynyt tarve luoda sellaisia oppimisympäristöjä, jotka mahdollistavat yksilöllisemmän oppimisen tukemisen ja oppimaan oppimisen taitojen omaksumisen teknologian avulla. Arvioinnin näkökulmasta on puolestaan keskeistä, miten diagnostinen, formatiivinen ja summatiivinen arviointi muotoillaan osaksi hybridiä tai digitaalista oppimisympäristöä niin, että se tukee oppimista (Boud 2010; Linder 2017).

Opettajien arviointiosaamisen kehittäminen on keskeinen opettajankoulutuksen haaste. Atjosen ym. (2019: 169) selvityksen mukaan opettajilla on kehitettävää erityisesti arviointitiedon pedagogisessa käytössä. Opettajat pystyivät mielestään tukemaan oppijoiden omien oppimistapojen tunnistamista ja kehittämistä, mutta eivät niinkään auttamaan heitä vertaamaan osaamistaan tavoitteisiin tai arviointiperusteisiin; samoin vain puolet oppijoista oli tyytyväisiä opettajalta saamaansa tukeen omassa arviointiin liittyvässä tavoitteenasettelussaan. (Atjonen ym: 169.)

Tämän tutkimuksen kontekstina ovat luokanopettajaopiskelijat ja heidän opinnoissaan suomen kielen ja kirjallisuuden pedagogiikan opintojakso. Opettajaopiskelijat tarvitsevat osana pedagogista osaamistaan kykyä havainnoida ja arvioida sekä suomea äidinkielenään puhuvien että monikielisten oppijoiden kielitaidon kehittymistä, jotta he voivat tarjota oikea-aikaista tukea oppimiseen. Tämän tavoitteen saavuttaminen edellyttää sitä, että opiskelijat pystyvät kehittämään ja refleктоimaan arviointiosaamistaan tavoitteellisesti koko opettajankoulutuksen ajan. Suomessa on tutkittu pääasiassa luokanopettajaopiskelijoiden omia suomen kielen ja kirjallisuuden eri sisältöalueiden taitoja: kielitiedon ymmärtämistä ja kieliopillista ajattelua (esim. Tainio & Routarinne 2012; Rättyä 2017; Alisaari ym. 2022) ja kirjoittamisen taitoa (esim. Torvelainen ym. 2021). Opiskelijoiden arviointiosaamisesta suomen kielessä ja kirjallisuudessa on sen sijaan niukasti tietoa.

Tarkastelemme tässä pilottitutkimuksessa diagnostista arviointia hybridissä oppimisympäristössä luokanopettajakoulutuksen kontekstissa. Hybridi oppimisympäristö on rakennettu viiden opintopisteen laajuiselle suomen kielen ja kirjallisuuden opintojaksolle, joka on osa perusopetuksessa opettavien aineiden ja aihekokonaisuuksien monialaisia opintoja (60 op). Opintojakson tavoitteena on 1) ymmärtää kielen sosiaalinen ja tilanteinen luonne sekä kielten merkitys identiteettien rakentumisessa, 2) oppia ohjaamaan monimuotoisten tekstien tulkinta- ja tuottamisprosesseja oppijälähtöisesti oppimisen eri vaiheissa sekä 3) osata arvioida ja kehittää aktiivisesti omaa kieliasiantuntijuutta. Opintojakso on jäsennetty Perusopetuksen opetussuunnitelman perusteita (Opetushallitus 2014; Äidinkieli ja kirjallisuus -luku) mukailleen neljään moduuliin: 1) kielet ja kielitietoisuus (kielen tilanteisuus ja variaatio, kielet ja identiteetti), 2) kirjallisuus ja kulttuuri, 3) oppijoiden tekstiympäristöt sekä 4) tekstien tuottaminen ja tulkitseminen ja monilukutaito. Moduuleja sisältöineen opiskellaan kontaktiopetuksessa, itsenäisissä digitaalisissa moduulitehtävissä ja soveltavassa tentissä.

2 Diagnostinen arviointi hybridin oppimisympäristön tutkimuskontekstissa

2.1 Diagnostinen arviointi ja itsearviointi oppimista tukemassa

Diagnostinen arviointi on tärkeä osa kieliopintoja (esim. Alderson ym. 2015), mutta sen hyötyä esimerkiksi pedagogisten käytänteiden kehittämisessä ja suuntaamisessa on tutkittu vähän (esim. Lee 2015). Diagnostisessa arvioinnissa on kyse prosessista, jossa selvitetään oppijan vahvuuksia ja heikkouksia opittavina olevissa taidoissa ja tiedoissa, jotta oppimista voidaan ohjata yksilöidyn palautteen avulla (Lee 2015). On tärkeää havaita, miten vahvuudet esimerkiksi tekstin tulkitsemistaidoissa ja heikkoudet tuottamistaidoissa ovat suhteessa toisiinsa ja miten ne molemmat vaikuttavat kielenkäyttötilanteissa toimimiseen ja pedagogisiin taitoihin.

Lee (2015) jakaa diagnostisen arvioinnin kolmeen vaiheeseen: diagnoosiin, palautteeseen ja oppimisen tukemiseen. Diagnoosissa on kyse diagnostisten tavoitteiden asettamisesta, arviointivälineiden valitsemisesta ja arvioinnin toteuttamisesta. Palautevaiheessa päätetään, miten palautetta annetaan diagnostisen arvioinnin pohjalta: missä määrin palaute on numeerista ja/tai sanallisia kuvauksia ja miten palaute skaalautuu erilaisiksi taitotasoin tai esimerkiksi osaa – ei osaa -dikotomioiksi. Palautteen muotokin voi vaihdella: palaute voi olla esimerkiksi vapaasti muotoiltua tai visuaalista tai se voi vaihdella tehtävittäin. Lopuksi diagnostista arviointitietoa hyödynnetään oppimista tukevien toimintojen suunnittelussa ja kehittämisessä. Arviointitieto auttaa oppimateriaalien valinnassa ja itsenäisen harjoittelun suuntaamisessa oppijoille yksilöllisten oppimistavoitteiden saavuttamiseksi.

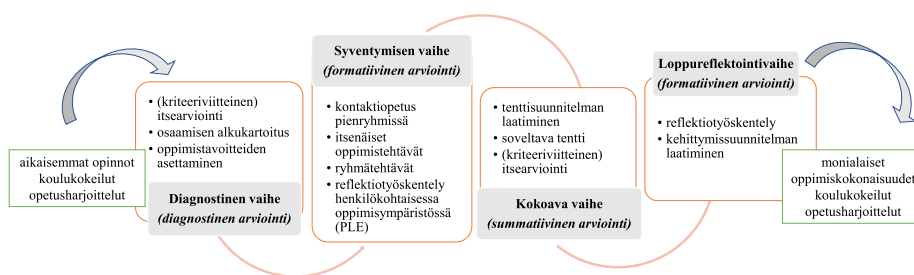
Diagnostiseen arviointiin on hyödyllistä liittää itsearviointi. Näiden arvioinnin muotojen yhdistämisellä opiskelijat saavat valmiuksia kehittää osaamistaan pitkäkestoisesti jatkuvan oppimisen hengessä eikä vain yhden opintojakson tavoitteiden täyttymiseksi (Boud & Molloy 2013; Torvelainen ym. 2019). Andrade (2019) muotoilee itsearvioinnin oppijan kyvyiksi, prosesseiksi ja tuotoksiksi, jotka vaikuttavat siihen, miten oppija rakentaa käsitystä osaamisestaan palautteen avulla. Vaikka itsearvioinnin käyttö on yleistynyt, oppija tarvitsee edelleen opettajan ohjausta siinä, miten itsearviointi toimii oppimisen edistämisessä ja suoriutumisen parantamisessa (Andrade 2019). Itsearviointi kiinnittyykin vahvasti oppimisen säätelyyn: oppimisen monitorointiin joko kriteereihin tukeutuen tai ilman niitä (Andrade & Brookhart 2020). Opettajankoulutuksessa on keskeistä kehittää itsearviointitaitoja laaja-alaisesti koko opintojen ajan sekä ohjauksen että oppimista tukevien arviointikriteerien ja itsearvioinnista annettavan palautteen avulla (Kainulainen ym. 2021).

2.2 Hybridi oppimisympäristö ja arviointi oppimisen tukijana

Suomen kielen ja kirjallisuuden opintojaksossa arvioinnilla on keskeinen pedagoginen tehtävä (Tarnanen ym. 2023). Opintojakson kehittämisen lähtökohtana on ollut opettajaopiskelijoiden aktiivinen osallistaminen arviointiprosessiin, sillä tutkimusten perusteella osallistaminen vahvistaa sitoutumista omaan oppimiseen ja tekee oppimisesta merkityksellisempää (esim. Kearney & Perkins 2014). Lisäksi opiskelijoiden osallistaminen tarjoaa mahdollisuuden tukea itseohjautuvuutta ja valmiuksia tarkastella kriittisesti omaa osaamista niin opettajaopinnoissa kuin tulevassa työssä. Opettajana kehittyminen tapahtuu tällöin jatkuvan oppimisen periaatteella systemaattisesti ja proaktiivisesti (Kearney & Perkins 2014; Korthagen 2016).

Luokanopettajaopiskelijoiden monialaisiin opintoihin sisältyy yhtäältä omien taitojen ja tietojen kehittäminen, toisaalta tämän tiedonalan opettajana toimiminen (Shulman 2004; Tarnanen ym. 2023). Suomen kielen ja kirjallisuuden pedagogisen sisältötiedon opetus sisältää kielitiedon, tekstitaidot, kirjallisuuden, vuorovaikutustaidot ja pedagogisen osaamisen. Näihin liittyvät kiinteästi arviointi ja oppimisen reflektointi. Opettajan arviointiosaaminen pitää täten sisällään myös identiteettityötä, jossa opettaja rakentaa käsitystään arvioinnista paitsi monipuolisen arviointia koskevan tiedon ja pedagogisen sisältötiedon perusteella myös kokemustensa, oppivan asenteensa ja reflektoinnin kautta. (Xu & Brown 2016.)

Opettajaopiskelijoiden suomen kielen ja kirjallisuuden sisältö- ja pedagogista osaamista edistämään olemme kehittäneet hybridin oppimisympäristön (ks. kuvio 1). Siihen kiinnittyvä arviointi kattaa koko oppimisprosessin, jossa diagnostinen, formatiivinen ja summatiivinen arviointi, mukaan lukien itsearviointi, muodostavat jatkumon ja ohjaavat sekä oppimista että opetusta.



KUVIO 1. Arvioinnin muodot osana hybridiä oppimisympäristöä suomen kielen ja kirjallisuuden opintojaksossa.

Hybridi oppimisympäristö koostuu neljästä vaiheesta, joissa oppimisprosessia tuetaan diagnostisen, formatiivisen ja summatiivisen arvioinnin avulla. Ensimmäinen eli

diagnostinen vaihe toteutetaan täysin verkossa, ja siinä opiskelijat suorittavat digitalisoidun alkukartoituksen ja arvioivat itsearviointityökalun avulla kriteeriviitteisesti omaa tekstien tulkinnan ja tuottamisen ja kielitiedon osaamistaan sekä pedagogista osaamistaan. Diagnostisella arvioinnilla kartoitetaan opiskelijoiden oppimisedellytyksiä ja sitä, miten he hallitsevat kieleen, tekstien tuottamiseen ja tulkintaan liittyviä tietoja ja taitoja. Diagnostisessa arvioinnissa tunnistetaan opiskelijan aiemmin hankkimaa osaamista. Arviointi ohjaa opiskelijaa asettamaan omat tavoitteensa opintojaksolle.

Toisessa eli syventymisen vaiheessa lähdetään tutkimaan ja kehittämään pedagogista osaamista pienryhmäopetuksen oppimistehtävien kautta sekä henkilökohtaisessa digitaalisessa oppimisympäristössä (PLE, *personal learning environment*) tehtävien itsenäisten oppimistehtävien avulla. Tällä formatiivisella eli oppimisprosessin aikaisella arvioinnilla (*assessment for learning*) tuetaan ja ohjataan opiskelijoiden oppimista ja annetaan siitä palautetta. Osaamisen itsearviointityökalu kriteereineen on oppimisen perustana ja opastajana opiskelijoiden reflektiossa ja palautteessa sekä PLE:n rakentamisessa.

Kolmannessa eli kokoavassa vaiheessa opiskelijat palaavat osaamisen itsearviointiin, reflektovat oppimistaan ja laativat suunnitelman soveltavaan tenttiin. Itsenäisesti täysin verkossa tehtävässä tentissä opiskelija syventää tai vahvistaa osaamistaan opintojakson sisällöissä aiemman, oppimisen reflektointiin perustuvan suunnitelmansa mukaisesti. Kolmanteen vaiheeseen kuuluu opintojakson summatiivinen arviointi (*assessment of learning*), jonka tehtävänä on kuvata, kuinka hyvin opiskelija on saavuttanut opintojakson tavoitteet. Opintojakson summatiivinen arviointi muodostuu kolmesta osa-alueesta: 1) opiskelijan itsearviointi osaamisensa kehittymisestä, opintojakson oppimistavoitteiden saavuttamisesta ja reflektioivasta työskentelystä, 2) opettaja-arvio opiskelijan kieli- ja viestintäasiantuntijuuden reflektoinnista ja PLE-työskentelystä ja 3) soveltavan tentin opettaja-arviointi. Neljännessä ja samalla viimeisessä eli loppureflektointivaiheessa formatiivinen arviointi jatkuu, kun opiskelija asettaa itselleen kehittymistavoitteita tulevia opintoja ja opettajan työtä silmällä pitäen.

3 Tutkimustehtävä ja -kysymykset

Tarkastelemme artikkelissa pilottitutkimuksen muodossa luokanopettajakoulutuksen suomen kielen ja kirjallisuuden pedagogiikan opintojakson diagnostista arviointia, joka sisältyy opintojaksolle suunniteltuun hybridiin oppimisympäristöön. Arvioinnin välineinä olivat opintojakson alussa tehtävä diagnostinen alkutesti ja kriteeriviitteinen itsearviointi, joiden avulla kartoitettiin tekstien tulkinnan ja tuottamisen taitoa sekä kielitietoa ja pedagogista osaamista. Tutkimuskysymyksemme ovat seuraavat:

1. Millaista suomen kielen ja kirjallisuuden osaamista (tekstien tulkinnan ja tuottamisen taitoa sekä kielitiedon hallintaa) luokanopettajaopiskelijat opintojakson alussa osoittavat?

2. Miten suomen kielen ja kirjallisuuden osaamisen ja itsearvioinnin tulokset suhteutuvat toisiinsa?
3. Miten suomen kielen ja kirjallisuuden lukion päättöarvosana ja ylioppilaskokeen äidinkielen arvosana (aiempi osaaminen) selittävät diagnostisen alkutestin tuloksia?

Tulosten pohjalta tarkastelemme, mitä opettajaopiskelijoiden osoittama osaaminen diagnostisessa alkutestissä sekä alkutestin ja itsearvioinnin väliset yhteydet merkitsevät suomen kielen ja kirjallisuuden pedagogisten opintojen suunnittelun kannalta.

4 Tutkimusaineisto ja -menetelmä

Tutkimus kohdistui diagnostisesta arviointivaiheesta kerättyyn aineistoon, joka koostui 127 luokanopettajaopiskelijan diagnostisen alkutestin tuloksista ja viisiportaisella kriteeriviitteisellä asteikolla tehdystä itsearvioinnista. Digitaalinen alkutesti sisälsi yhden monivalintatehtävän ja kymmenen avotehtävää, joilla mitattiin kielitiedon osaamista sekä tekstien tulkinnan ja tuottamisen taitoja. *Kielitiedon osaamista* mitattiin yhdellä kaksiosaisella tehtävällä. A-kohdassa oli kuusi avotehtävää, ja kysymyksillä *Miksi seuraavat ilmaukset ovat monitulkintaisia? Millä kielen käsitteillä selittäisit niitä?* pyydettiin opiskelijoita lyhyesti kielen käsittein selvittämään kuuden virkkeen monitulkintaisuutta, esimerkiksi *Arkkitehdille sattui vahinko: oli käytävä vessassa*. B-kohdassa oli yksi avotehtävä, jossa kysymyksen *Mitä ja miksi tiedät virkkeen perusteella?* ohjaamana piti tehdä havaintoja epäkielivirkkeestä *Älkköinen hörmöskörkkeli lempautteli piirpäisessä lasmottimessa*.

Tekstien tulkintataitoja mitattiin novellin ja uutisen tulkintatehtävillä. Novellianalyysi koostui opiskelijan omasta tulkinnasta sekä tulkinnan käsitteellistämisestä. Uutisen tulkintatehtävä koostui yhdestä argumentointia vaativasta avotehtävästä (*Miksi uutisen tulkinnassa tarvitaan kriittistä lukutaitoa? Perustelee vastauksesi mahdollisimman monipuolisesti*) ja 14-osaisesta monivalintatehtävästä, jossa oli oikein-väärin-väittämiä, esimerkiksi *Uutinen noudattaa tekstilajille tyypillistä etenemistä (keskeisestä ja tärkeänä pidetystä yksityiskohtiin ja taustoihin)*. *Tuottamistaitoja* mitattiin kahdella avotehtävällä: Ensimmäinen ohjeistettiin laatimaan Wilma-viesti vanhemmille. Toiseksi pyydettiin laatimaan oppilaille ohjeteksti Peda.netin käyttöön sekä refleктоimaan myös ohjeen tuottamisessa tehtyjä valintoja. Diagnostisen alkutestin tehtäviin pyrittiin valitsemaan monipuolisesti sellaisia tekstejä, joita opiskelijat tarvitsevat opettajan työssä. Opiskelijoilta pyydettiin henkilökohtainen, kirjallinen lupa diagnostisen arvioinnin tutkimuskäyttöön Jyväskylän yliopiston tietosuojakäytäntöjen mukaisesti.

Diagnostisen alkutestin toimivuus mittarina varmistettiin sekä klassisen osioanalyysin menetelmin että irt (*item response theory*) -menetelmiin kuuluvalla Raschin mallilla (ks. Törmäkangas & Törmäkangas 2009). Näiden perusteella alkutestin kahdesta tuottamistehtävästä otettiin tutkimukseen mukaan Wilma-viesti vanhemmille.

Lisäksi diagnostisen alkutestin pisterajat muunnettiin testin käytettävyyttä varten raakapisteistä yliopistossa käytössä olevalle 5-portaiselle asteikolle, jolle laadimme kriteerit. Näin varmistettiin, että mittari toimii pilottitutkimuksemme näkökulmasta tarpeeksi luotettavasti.

Itsearviointien arviointikohteet – *kielitiedon osaaminen, tekstien tulkinta- ja tuottamistaidot sekä näiden sisältöalueiden pedagoginen osaaminen* – valikoituivat Perusopetuksen opetussuunnitelman perusteiden keskeisten sisältöalueiden sekä luokanopettajaopintojen monialaisten opintojen pedagogisten painotusten perusteella.

Kohteiden itsearviointien tukena hyödynnettiin digitaalista arviointityökalua, jonne laadittiin kriteeriviitteinen viisiportainen osaamistasoasteikko hatara – kohdalainen – hyvä – vahva – asiantunteva. Asteikko noudattaa opettajankoulutukseen kehitettyä viisiportaista arvioinnin kehystä (Kainulainen ym. 2021), jonka jaottelu perustuu taitojen oppimisen malleihin ja kognitiiviseen näkökulmaan asiantuntijuuden kehittymisessä (ks. esim. Dreyfus & Dreyfus 2005). Osaamistasojen kuvaukset pyrkivät mallintamaan osaamisen kehittämisen ja asiantuntijuuden dynaamista luonnetta ja opettajan tarvetta oppia ja kehittää osaamistaan jatkuvasti.

Tutkimusaineisto analysoitiin määrällisesti tilastollisin menetelmin. Aineistosta tutkittiin ensin kokonaistulokset diagnostisen alkutestin eri sisältöalueilta ja suhteellinen osaaminen eri sisältöalueiden välillä, jotta opiskelijoiden kielellisen osaamisen taso opintojakson alussa kävisi ilmi. Tämän jälkeen diagnostisen alkutestin tuloksia verrattiin itsearviointeihin, jotta nähtiin, missä määrin opiskelijat pystyivät arvioimaan omia taitojaan ja osaamistaan. Luokanopettajaopiskelijat ovat kielellisiltä taidoiltaan valikoitunut kohdejoukko, joten halusimme selvittää myös heidän aiemmin arvioidun ja testatun osaamisensa suomen kielessä ja kirjallisuudessa. Taustamuuttujista kysyttiin täten suomen kielen ja kirjallisuuden oppimäärän arvosanaa lukion päättötodistuksessa ja äidinkielen ylioppilaskokeen arvosanaa.

Tutkimukseen osallistuneista opiskelijoista naisia oli 105 ja miehiä 20, ja ryhmään muut kuului 2 vastaajaa. Yleisimmät arvosanat suomen kielen ja kirjallisuuden oppimäärässä lukion päättötodistuksessa olivat 9 ja 8, jotka kattoivat 87 % vastaajista. Arvosana 10 oli 8 prosentilla vastaajista ja arvosana 7 oli 4 prosentilla. Prosentti vastaajista oli saanut arvosanan 6 tai alle tai ei ollut käynyt lukiota. Suurin osa opiskelijoista oli saanut äidinkielen ylioppilaskokeesta magnan (41 %). Seuraavaksi yleisimmät arvosanat olivat eximia (28 %) ja cum laude (20 %). Laudatur oli 6 prosentilla vastaajista ja approbatur tai lubenter yhteensä 5 prosentilla opiskelijoista. Alimpia arvosanoja selittivät ammattiopintotaustaisten luokanopettajaopiskelijoiden päättö- ja ylioppilaskokeen arvosanat. Ylioppilaskokeen arvosanojen trendi ei vastaa päättöarvosanojen trendiä, vaan päättöarvosanat ovat suhteessa ylempiä kuin ylioppilaskokeen arvosanat. Tätä selittää se, että ylioppilaskokeesta poiketen päättöarvosanoissa on otettu huomioon oppiaineen kaikki sisältöalueet samoin kuin asenne- ja työskentelytavoitteet.

Taustamuuttujien suhdetta diagnostisten alkutestien tuloksiin (mitattuna sekä testien raakapisteinä että 5-portaisella arvosana-asteikolla) analysoitiin muuttujan

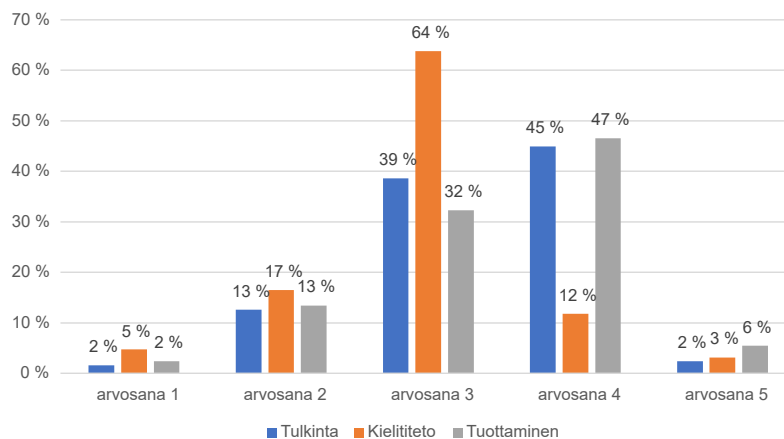
mitta-asteikosta riippuen Kruskall-Wallis-testin ja yksisuuntaisen varianssianalyysin (ANOVA) avulla. Parittaiset vertailut ryhmien välillä tehtiin tarvittaessa Bonferroni-menettelmällä. Taustamuuttujien ja diagnostisen alkutestin tulosten välistä yhteyttä analysoitiin myös lineaarisella regressiomallilla. Lisäksi tarkasteltiin eri alkutestien tulosten yhteyttä toisiinsa Spearmanin ja Pearsonin korrelaatiokertoimen avulla. (Ks. Metsämuuronen 2005.)

5 Tulokset

5.1 Diagnostisen alkutestin tulokset

5.1.1 Kokonaistulokset suomen kielen ja kirjallisuuden eri sisältöalueilta
Valtaosa eli 68 % luokanopettajaopiskelijoista hallitsi diagnostisen alkutestin perusteella kielitiedon sekä tekstien tulkinnan ja tuottamisen taidot arvosanan 3 edestä yliopiston arvosana-asteikolle suhteutettuna. Neljännes opiskelijoista sai kielellisestä osaamisestaan arvosanan 4, 5 % arvosanan 2 ja 2 % arvosanan 5.

Luokanopettajaopiskelijoiden kielellinen osaaminen vaihteli sisältöalueittain. Suhteellisesti parhaiten luokanopettajaopiskelijat hallitsivat diagnostisen alkutestin perusteella eri sisältöalueista tekstien tulkintaa, sillä 86 % opiskelijoista suoriutui tulkintatehtävistä vähintään arvosanan 3 arvoisesti. Vain vähän heikommin opiskelijat selvittivät tuottamistehtävän, sillä 84 % opiskelijoista sai tästä sisältöalueesta vähintään arvosanan 3. Heikoin sisältöalueista oli kielitieto, jossa arvosanan 3 tai enemmän sai 79 prosenttia opiskelijoista. (Ks. kuvio 2.)



KUVIO 2. Luokanopettajaopiskelijoiden suomen kielen ja kirjallisuuden osaaminen sisältöalueittain diagnostisessa alkutestissä.

Kielitiedon osaaminen erottuu muista sisältöalueista myös arvosanojen 3 ja 4 määrissä. Kielitiedossa on suhteellisesti enemmän arvosanaa 3, siinä missä tulkinnassa ja tuottamisessa on arvosanaa 4. Huomionarvoista suhteellisen osaamisen tarkastelussa luotettavuuden näkökulmasta kuitenkin on, että tulkintatehtävät oikein-väärin-väittämiseen eivät erottele osaamista kovin hyvin, vaan kielitietotehtävät erottelevat osaamista paremmin.

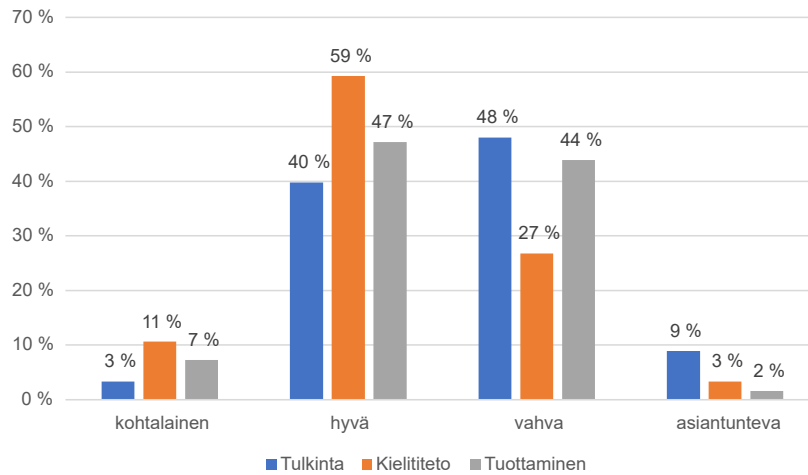
5.1.2 Suomen kielen ja kirjallisuuden eri sisältöalueiden tulosten välinen yhteys diagnostisen alkutestin perusteella

Opettajaopiskelijoiden kielellinen osaaminen eri sisältöalueilla oli diagnostisen alkutestin perusteella jossakin määrin yhteydessä toisiinsa. Tulkinnan testin ja kielitiedon testin tulokset korreloivat kumpikin positiivisesti tuottamisen testin kanssa (Spearmanin korrelaatiokertoimet, kun raakapisteet on käännetty 5-portaiselle arvosana-asteikolle, 0,329 ($p < 0,001$) ja 0,245 ($p=0,005$)); Pearsonin korrelaatiokertoimet testien raakapisteillä 0,374 ($p<0,001$) ja 0,241 ($p=0,006$). Sen sijaan tulkinnan testin ja kielitiedon testin tulokset eivät korreloi keskenään tilastollisesti merkitsevästi (Spearmanin korrelaatiokerroin on 0,158 ($p=0,077$); Pearsonin korrelaatiokerroin on 0,195 ($p=0,028$).

5.2 Itsearviointin tulokset

5.2.1 Opiskelijoiden suomen kielen ja kirjallisuuden eri sisältöalueiden osaaminen itsearviointin perusteella

Osaamisensa suomen kielen ja kirjallisuuden eri sisältöalueilla hyväksi (3) tai vahvaksi (4) arvioineita opiskelijoita oli itsearvioinneissa eniten: tulkinnan taidoissa 88 %, kielitiedossa 86 % ja tuottamisen taidoissa peräti 91 % (ks. kuvio 3). Huomionarvoista on, että 9 % opiskelijoista arvioi osaamisensa tulkinnan taidoissa asiantuntevaksi (5).



KUVIO 3. Suomen kielen ja kirjallisuuden eri sisältöalueiden osaaminen luokanopettaja-opiskelijoiden itsearviointien perusteella.

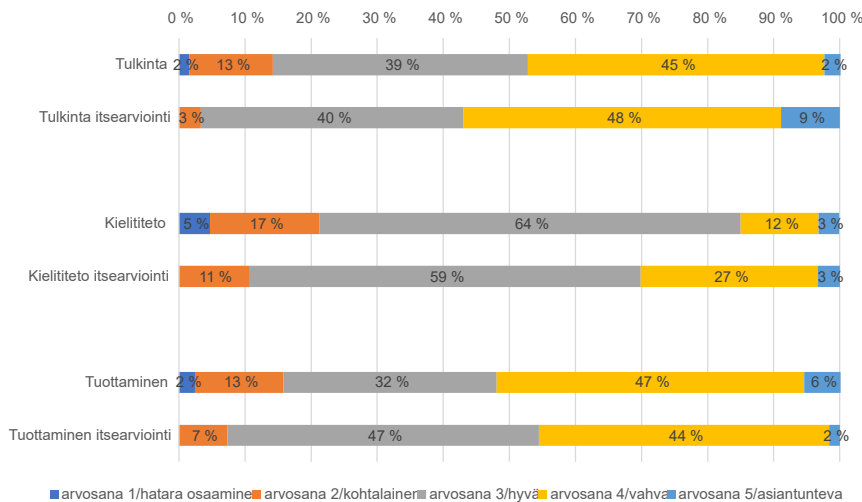
Kielitiedon osaaminen oli sisältöalueista se, jota vain alle kolmasosa opiskelijoista (30 %) piti vahvana tai asiantuntevana oman arvionsa perusteella. Lisäksi joka 10. opiskelija arvioi kielitiedon osaamisensa vain kohtalaiseksi (arvosana 2). Kielitiedon itsearviointi erottui muiden sisältöalueiden arvioinneista myös arvosanojen 3 ja 4 määrissä samaan tapaan kuin diagnostisen alkutestin tuloksissa. Kielitiedon itsearvioinneissa oli suhteellisesti eniten hyvää osaamista siinä missä tulkinnan ja tuottamisen itsearvioinneissa oli vahvaa osaamista.

Eri sisältöalueiden itsearviointit sijoittuvat ylipäätään asteikon keskelle (kuvio 3), mikä on luonnollista, sillä kyseessä on koulutustaustaltaan ja opiskeluintresseiltään valikoitu otos eli opettajaksi opiskelevat.

5.2.2 Diagnostisen alkutestin eri sisältöalueiden tulosten ja niiden itsearviointien välinen yhteys

Diagnostisen alkutestin tulosten ja itsearviointien välillä oli hentoja viitteitä yhteydestä kaikilla kielellisen osaamisen alueilla. Vertailu diagnostisen alkutestin tulosten ja itsearviointien palkkien välillä kuviossa 4 osoittaa trendin eli niiden suhteellisten pituuksien vastaavan toisiaan.

195 DIAGNOSTINEN ARVIOINTI SUOMEN KIELEN JA KIRJALLISUUDEN PEDAGOGIIKAN OPINTOJAKSON OPISKELUA JA OPETUSTA OHJAAMASSA



KUVIO 4. Diagnostisen alkutestin eri sisältöalueiden tulosten ja niiden itsearviointien välinen yhteys.

Opiskelijat pystyivät siis arvioimaan ainakin jossakin määrin omaa kielellistä osaamistaan ja täten tunnistamaan myös kehittämisen kohteitaan. Tarvitaan kuitenkin lisää dataa, jotta tuloksia voidaan luotettavasti yleistää. Korrelaatiokertoimista ainoastaan kielitiedon tulosten ja kielitiedon itsearvioinnin kerroin oli tilastollisesti merkitsevä mutta varsin matala sekin ($r=0,223$, $p=0,013$).

5.3 Luokanopettajaopiskelijoiden arvosanojen yhteys diagnostisen alkutestin eri sisältöalueiden tuloksiin

Lopuksi tuomme luokanopettajaopiskelijoita koskevista taustamuuttujista esiin arvosanojen yhteyksiä suomen kielen ja kirjallisuuden eri sisältöalueiden osaamiseen diagnostisessa alkutestissä. Näin saamme tarkasteltua lukion suomen kielen ja kirjallisuuden päättöarvosanojen ja äidinkielen ylioppilaskokeen arvosanojen ennustusarvoa suhteessa diagnostisessa alkutestissä menestymiseen.

Opiskelijoiden diagnostisen alkutestin tuloksia verrattiin aluksi lukion päättöarvosanoihin. Kruskal-Wallis testin ja yksisuuntaisen varianssianalyysin avulla tarkasteltiin, onko kielitiedon, tulkinnan ja tuottamisen diagnostisen alkutestin tuloksissa tilastollisesti merkitseviä eroja suomen kielen ja kirjallisuuden lukion päättöarvosanan mukaan.

Analyytit tuottivat lukion suomen kielen ja kirjallisuuden päättöarvosanan mu-
kaisten ryhmien välillä merkitseviä eroja ensinnäkin *tuottamisen alkutestin* tuloksissa raakapisteinä mitattuna ($F(4,122)=2,707$, $p=0,033$). Lisäksi havaittiin, että arvosanan 10 saaneet saivat tuottamisen testistä parempia pisteitä kuin arvosanan 7 saaneet

($p=0,018$). Ryhmäkohtaisia tilastollisesti merkitseviä eroja ei kuitenkaan saatu tuottamisen testin tuloksissa 5-portaisella arvosana-asteikolla mitattuna (Kruskal-Wallis $H(4)=7,418$, $p=0,115$).

Myös *kielitiedon osaamisen* tuloksissa (sekä raakapisteinä että 5-portaisella arvosana-asteikolla mitattuna) oli tilastollisesti merkitseviä eroja päättöarvosanan mukaan (Kruskal-Wallis $H(4) = 14,347$, $p= 0,006$ ja/tai ANOVA: $F(4,122)=3,758$, $p=0,006$). Arvosanan 10 saaneet opiskelijat menestyivät paremmin kielitiedon testissä kuin arvosanan 7 ($p=0,036$) tai 8 ($p=0,026$) saaneet. *Tulkinnan taitojen* tuloksella ei sen sijaan ollut yhteyttä suomen kielen ja kirjallisuuden päättöarvosanaan.

Kruskal-Wallis testin ja yksisuuntaisen varianssianalyysin (ANOVA) jatkoksi tarkasteltiin lukion päättöarvosanan yhteyttä kielitiedon, tuottamisen ja tulkinnan alkutestin tuloksiin regressioanalyysin avulla. Sen tehtävänä on kertoa tiivistetysti, ennustaako arvosana ylipäätään alkutestin tuloksia. Regressiokertoimet olivat positiiviset kaikissa malleissa, toisin sanoen päättöarvosanan kasvaessa myös alkutestin raakapisteet kasvoivat. Alkutestin raakapisteitä ennustava regressiomalli selitti tilastollisesti merkitsevästi *kielitiedon* tuloksia ($R^2 = 0,103$, $F(1,125)=14,303$, $p<0,001$). Päättöarvosana selitti 10 % kielitiedon alkutestin raakapisteiden vaihtelusta, ja sen regressiokerroin erosi tilastollisesti merkitsevästi nolasta ($\beta=1,074$, $p<0,001$). Lisäksi regressiomalli selitti kielitiedon alkutestin tuloksia 5-portaisella arvosana-asteikolla mitattuna tilastollisesti merkitsevästi ($R^2 = 0,106$, $F(1,125)=14,880$, $p<0,001$). Päättöarvosana selitti noin 11 % kielitiedon testin tulosten vaihtelusta, ja sen regressiokerroin erosi tilastollisesti merkitsevästi nolasta ($\beta=0,330$, $p= p<0,001$).

Myös *tuottamisen* raakapisteitä ennustava regressiomalli selitti tilastollisesti merkitsevästi alkutestin tuloksia ($R^2 = 0,035$, $F(1,125)=4,559$, $p=0,035$). Päättöarvosana selitti 3,5 % tuottamisen testin pisteiden vaihtelusta. Päättöarvosanan regressiokerroin erosi tilastollisesti merkitsevästi nolasta ($\beta=0,403$, $p=0,035$). Tuottamisen tuloksia 5-portaisella arvosana-asteikolla ennustava regressiomalli selitti samaan tapaan tilastollisesti merkitsevästi tuloksia ($R^2 = 0,034$, $F(1,125)=4,454$, $p=0,037$). Päättöarvosana selitti 3,4 % tuottamisen alkutestin pisteiden vaihtelusta, ja sen regressiokerroin erosi tilastollisesti merkitsevästi nolasta ($\beta=0,212$, $p=0,037$).

Lineaarisella regressiomallilla tutkittiin myös, ennustaako ylioppilaskokeen äidinkielen arvosana kielitiedon, tuottamisen ja tulkinnan diagnostisen alkutestin tuloksia. Vain *kielitiedon* raakapisteitä ennustava regressiomalli selitti tilastollisesti merkitsevästi diagnostisen alkutestin tuloksia ($R^2 = 0,061$, $F(1,125)=8,050$, $p=0,005$).

Ylioppilastutkinnon arvosana selitti noin 6 % kielitiedon diagnostisen alkutestin pisteiden vaihtelusta, ja sen regressiokerroin erosi tilastollisesti merkitsevästi nolasta ($\beta=0,658$, $p=0,005$). Myös kielitiedon tuloksia 5-portaisella arvosana-asteikolla ennustava regressiomalli selitti tilastollisesti merkitsevästi diagnostisen alkutestin tuloksia ($R^2 = 0,046$, $F(1,125)=6,039$, $p=0,015$). Ylioppilastutkinnon arvosana selitti vajaat 5 % kielitiedon diagnostisen alkutestin tuloksien vaihtelusta, ja sen regressiokerroin erosi tilastollisesti merkitsevästi nolasta ($\beta=0,173$, $p=0,015$).

Kootusti luokanopettajaopiskelijoiden päättöarvosana ja ylioppilaskokeen tulos olivat tilastollisesti merkitsevässä yhteydessä kielitiedon osaamiseen ja tuottamisen taitoihin diagnostisessa alkutestissä.

6 Pohdinta

Tarkastelimme artikkelissamme suomen kielen ja kirjallisuuden opintojakson diagnostisesta arvioinnista analysoitua tilastollista dataa. Selvitimme, millaista osaamista opiskelijat osoittivat oppiaineen eri sisältöalueilla, ja pohdimme tässä luvussa, mitä arvioinnin tulokset merkitsevät opetuksen järjestämisen kannalta.

Valtaosa eli noin kaksi kolmasosaa luokanopettajaopiskelijoista osoitti diagnostisessa alkutestissä hyvää kielellistä osaamista tekstien tulkinnan ja tuottamisen taidoissa sekä kielitiedon osaamisessa. Viidellä prosentilla opiskelijoista kielellinen osaaminen jäi puolestaan heikoksi.

Oppiaineen sisältöalueista kielitieto osoittautui diagnostisessa alkutestissä suhteellisesti heikoimmaksi, ja myös itsearvioinneissa opettajaopiskelijat arvioivat kielitiedon eri sisältöalueista heikoimmaksi. Luokanopettajaopiskelijoiden heikko kielitiedon osaaminen on havaittu myös aiemmissa tutkimuksissa (esim. Tainio & Routarinne 2012; Rättyä 2017). Kielitiedon osaamista, siihen liittyviä ajattelumalleja ja niiden siirtämistä opetuksessa eteenpäin onkin pohdittu viime vuosikymmeninä opettajankoulutuksessa lisääntyvässä määrin. Tutkimukset (esim. Tainio & Routarinne 2012; Tainio & Marjokorpi 2014; Rättyä 2017; ks. myös Phipps & Borg 2009) osoittavat omien havaintojemme suuntaisesti, että on tärkeää kartoittaa opettajaopiskelijoiden kielitiedon osaamista ja kielitietoa koskevia ajattelumalleja ja myös käsitellä tuloksia opiskelijoiden kanssa. Voidaankin kysyä, onko siirtymä perinteisestä kielitiedon opetuksesta kielitietoisuusajatteluun ja tekstilajilähtöiseen kielen rakenteiden jäsentämiseen edelleen kesken. Jääkö kielitieto edelleen irralleen tulkinta- ja tuottamistaidoista, ja miten tätä juopaa voidaan opettajankoulutuksessa kuroa umpeen siten, että opiskelijat osaavat tulevassa työssään toteuttaa kielitietoisempaa opetusta? Yhdessä keskustellen ja tuloksia analysoimalla ja havainnollistamalla voidaan pyrkiä saavuttamaan muutoksia opiskelijoiden ajattelussa ja toiminnassa (ks. Phipps & Borg 2009).

Tuottamisen taidot olivat opettajaopiskelijoilla tilastollisesti merkitsevästi yhteydessä tulkinnan ja kielitiedon osaamiseen. Tuottamisen taitojen ja niiden itsearvioinnin välillä oli lisäksi hento yhteys. Opiskelijoiden voikin olla helpompi hahmottaa omia tuottamisen taitojaan, sillä he ovat kaiketi saaneet niistä eniten ja konkreettisinta palautetta kouluopinnoissaan.

Vaikka opettajaopiskelijoiden osaaminen vaihteli suomen kielen ja kirjallisuuden eri sisältöalueilla, nämä alueet kietoutuvat monin tavoin yhteen. On selvää, että hyvät tekstin tulkintataidot edellyttävät tekstilajien tuntemusta, joka on tarpeen myös tekstien tuottamisessa. Kielitieto puolestaan auttaa tilanteeseen ja tekstilajiin sopivien

kielellisten valintojen tekemistä. Tuottamisen taitojen yhteyttä tulkinnan ja kielitiedon osaamiseen voi tarkastella myös testin luotettavuuden kannalta. Tuottamisen tehtävä diagnostisessa alkutestissä on tekstipohjainen ja vaatii näin kattavaa tekstin rakennepiirteiden hallintaa. Näin ollen on luonnollista, että tuottamisessa osoitettu osaaminen on yhteydessä tulkintataitojen ja kielitiedon osaamiseen. Tätä tukee myös Myhillin ja Watsonin (2014) meta-analyysin tulos siitä, miten kieliopin opetus metalingvististä tietämystä lisäämällä eli merkitys- ja kontekstilähtöisesti saattaa tukea kirjoittamisen taitoa. Suomen kielen ja kirjallisuuden eri sisältöalueiden väliset yhteydet mahdollistavatkin sen, että opinnoissa voidaan pohtia käsityksiä kielellisistä taidoista, niiden ilmenemisestä ja tukemisesta oppimisessa.

Tutkimus paljasti hentoja viitteitä myös tekstin tulkinnan ja kielitiedon osaamisen ja niiden itsearviointien välillä. Näin ollen opettajaopiskelijat pystyivät tunnistamaan ainakin jossakin määrin kielellistä osaamistaan näillä sisältöalueilla, mikä auttaa puolestaan opiskelijaa opiskelunsa suuntaamisessa ja suunnittelussa. Tulos auttaa opiskelijaa myös motivoitumaan kielitiedon opiskelussa ja suuntaamaan kielipedagogiikkaansa luokahuoneessa merkityslähtöiseen kielitiedon opetukseen (ks. Rättyä 2017).

Diagnostisen testin pilottivaiheen tulokset osoittivat lisäksi, että testin avulla pystytään arvioimaan suomen kielen ja kirjallisuuden eri sisältöalueiden tuntemusta ainakin jossain määrin luotettavasti, sillä päättöarvosanalla ja ylioppilaskokeen tuloksella oli tilastollista ennustearvoa suhteessa diagnostisen alkutestin tulokseen. Aiemmat arvosanat olivat tilastollisesti merkitsevässä yhteydessä kielitiedon osaamiseen ja tuottamisen taitoihin diagnostisessa alkutestissä. Tässä mielessä diagnostinen alkutesti toimii tehtävässään eli kielellisen osaamisen tunnistamisessa. Diagnostisen alkutestin ja itsearvioinnin yhdistelmän avulla on mahdollista ohjata opettajaopiskelijaa rakentamaan realistista ja tavoiteorientoitunutta käsitystä omasta osaamisestaan (ks. myös Andrade 2019).

Luokanopettajaopiskelijoiden suomen kielen ja kirjallisuuden sisältöalueiden osaamista kehitetään varsin pienin aika- ja opintopisteresurssein, joten resurssien kohdentaminen opiskelijoiden yksilöllisten tarpeiden mukaisesti on erittäin tärkeää. Tarkastellun alkukartoituksen vahvuutena voidaankin pitää toimivaa diagnostisen alkutestin ja itsearvioinnin yhdistelmää oppimistavoitteiden saavuttamiseksi. Siitä saatavan tiedon avulla opettaja pystyy ohjaamaan ja opiskelija säätelemään oppimisprosessiaan oppimista kehittävän arvioinnin hengessä (Atjonen 2015; Lee 2015).

Havaitsimme analyyseissa tilastollisesti merkitseviä yhteyksiä luokanopettajaopiskelijoiden arvosanojen ja diagnostisen alkutestin osoittaman kielellisen osaamisen välillä. Sekä tuottamisen että kielitiedon osalta opiskelijat osoittivat lukion päättöarvosanoihin suhteutettuna tilastollisesti merkitsevää osaamista. Lukion päättöarvosanat myös ennustivat tilastollisesti merkitsevästi näiden sisältöalueiden osaamista. Ylioppilaskokeen arvosana ennusti tilastollisesti merkitsevästi ainoastaan kielitiedon diagnostisen alkutestin tuloksia. Tulokset antavat aihetta monenlaisiin pohdintoihin: Mittaavatko diagnostisen alkutestin tehtävät samantyyppistä osaamista tuottamisessa

ja kielitiedossa kuin lukion oppimäärä? Ylioppilaskoe ainakin poikkeaa tekstien tuottamisen ja tulkinnan määrittelyn osalta alkutestistämme. Kielitiedon diagnostinen alkutesti näyttää sen sijaan olevan vahvassa yhteydessä sekä lukion oppimäärään että ylioppilaskokeeseen. Onko tämä sisältöalueista se, jota on lopulta yksinkertaisin mitata koulukulttuurissa vakiintunein tehtävätyypein? Opettajankoulutuksen kannalta diagnostisen alkutestin yhteyksiä aiempiin arvosanoihin kannattaa tuoda joka tapauksessa esiin ja pohtia yhdessä, mitä tekijöitä tilastollisen ennustettavuuden taustalla on ja miten ne vaikuttavat ylipäänsä oppijoiden käsityksiin kielellisestä osaamisesta ja erityisesti opettajaksi opiskeluun.

Tutkimuksemme tulosten valossa suomen kielen ja kirjallisuuden pedagogiikkaa pitäisi opettajankoulutuksessa suunnata paitsi sisältötiedon ja pedagogiikan menetelmien hallintaan myös osaamisen ja työskentelyn arviointiin ja tätä kautta avautuvaan käsitysten ja kokemusten avaamiseen oppijan roolissa. Kaiken kaikkiaan kyse on oppiaineen taitokäsityksestä: miten hahmotamme kielelliset taidot sekä niiden limittymisen, miten yhdistämme kielellisen toiminnan sekä käsitykset omista kielellisistä taidoistamme ja niiden luonteesta jäsentyneesti suunnitelluksi opintojaksoksi. Kehitämme koulutustamme tavoitteellisesti siihen suuntaan, että kielelliset taidot ja kielitietoisuus ylittävivät suomen kielen ja kirjallisuuden oppiaineen rajat. Kielellisten taitojen rooli kaikessa vuorovaikutuksessa, tiedon rakentamisessa eri oppiaineissa ja oppijan toimijuuden varmistamisessa on koulukontekstissa keskeistä.

Opettajaopiskelijoiden käsitys omasta suomen kielen ja kirjallisuuden eri sisältöalueiden osaamisestaan on korkealla, sillä osaamisensa vahvaksi tai asiantuntevaksi arvioineita opettajaopiskelijoita oli itsearvioinneissa eniten. Itsearviointi voi hyvin pohjautua äidinkielen ylioppilaskokeessa menestymiseen. Kun käsitys omasta kielellisestä osaamisesta on opintojakson alkaessa vahva, koulutuksen tehtävänä lienee auttaa selvittämään realistinen senhetkinen osaaminen ja suhteuttaa se opettajana työskentelyyn. Itsearviointi tuo näkyväksi ja keskusteltavaksi monia keskeisiä opiskelijoiden kieli- ja oppimiskäsityksiin vaikuttavia asioita, kuten kouluhistorian, kielellisen minäpystyvyyden ja sen vaikutuksen sekä osaamisen kehittämiseen että pedagogisiin taitoihin.

Tutkimuksellamme on muutamia rajoituksia. Tulokset perustuvat pilottitutkimukseen, jossa on testattu diagnostista arviointia yhtenä osatekijänä hybridissä oppimisympäristössä. Saatuihin tuloksiin vaikuttaa osaltaan se, että kokonaisuutta ei ole vielä kehitetty eteenpäin tutkimustulosten ja kokemusten pohjalta. Kyseessä on myös esitestausaineisto, joten tapaukset muuttujittain ovat paikoin vähäisiä. Pieni otoskoko voikin vaikuttaa tilastollisten tulosten luotettavuuteen, mikä on huomioitava niitä tulkittaessa. Lisäksi datan muodostamisvaiheessa diagnostisen alkutestin pisterajat muunnettiin käytettävyyden takia raakapisteistä 5-portaiselle asteikolle. Tällaista melko mekaanista pisterajojen asettamista voi pohtia myös tulosten luotettavuuden näkökulmasta.

Kirjallisuus

- Alderson, J. C., Haapakangas, E.-L., Huhta, A., Nieminen, L. & Ullakonoja, R. 2015. *The diagnosis of reading in a second or foreign language*. New York and London: Routledge.
- Alisaari, J., Heikkola, L.-M. & Harju-Autti, R. 2022. "Tää pitäisi piirtää. Tämä tehtävä on kielellisesti vaikea kaikille." Luokanopettajaopiskelijoiden kieliopin hallinta ja käsityksiä kielellisestä tuesta. *Ainedidaktikka* 6 (1), 47–69.
<https://doi.org/10.23988/ad.109885>
- Andrade, H. L. 2019. A Critical review of research on student self-assessment. *Front. Educ.*, 4 (87). <https://doi.org/10.3389/educ.2019.00087>.
- Andrade, H. L. & Brookhart, S. M. 2020. Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy & Practice*, 27 (4), 350–372.
<https://doi.org/10.1080/0969594X.2019.1571992>.
- Atjonen, P. 2015. *Kehittävä arviointi kasvatusalalla*. Joensuu: Kirjokansi.
- Atjonen, P., Laivamaa, H., Levonen, A., Orell, S., Saari, M., Sulonen, K., Tamm, M., Kamppi, P., Rumpu, N., Hietala, R. & Immonen, J. 2019. "Että tietää missä on menossa". *Oppimisen ja osaamisen arviointi perusopetuksessa ja lukiokoulutuksessa*. Helsinki: Kansallinen koulutuksen arviointikeskus. Julkaisut 7:2019.
- Boud, D. 2010. Relocating reflection in the context of practice. Teoksessa H. Bradbury, N. Frost & S. Kilminster (toim.) *Beyond Reflective Practice*. New York: Routledge, 25–36.
- Boud, D. & Molloy, E. 2013. Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education*, 38 (6), 698–712.
<https://doi.org/10.1080/02602938.2012.691462>.
- Dreyfus, H. L. & Dreyfus, S. E. 2005. Peripheral vision. Expertise in real world contexts. *Organization Studies* 26 (5). London: SAGE Publications, 779–792.
<https://doi.org/10.1177/0170840605053102>
- Fuller, M. B. & Skidmore, S. T. 2014. An exploration of factors influencing institutional cultures of assessment. *International Journal of Educational Research*, 65, 9–21.
<https://doi.org/10.1016/j.ijer.2014.01.001>
- Kainulainen, J., Tarnanen, M., & Vihriälä, J. 2021. Arviointiosaamisen kehittäminen digitaalisessa oppimisympäristössä. *Oppimisen ja oppimisvaikeuksien erityislehti NMI-bulletin*, 31 (2), 125–138.
- Kearney, S. P. & Perkins, T. 2014. Engaging students through assessment: the success and limitations of the ASPAL (Authentic Self and Peer assessment for Learning) model. *Journal of University Teaching & Learning Practice*, 11 (3), 2014.
<https://doi.org/10.53761/1.11.3.2>
- Korthagen, F. A. J. 2016. Pedagogy of teacher education. Teoksessa J. Loughran & M. L. Hamilton (toim.) *International handbook of teacher education, Volume 1*. Springer Science, 311–346.
- Lee, Y. W. 2015. Diagnosing diagnostic language assessment. *Language Testing*, 32 (3), 299–316. <https://doi.org/10.1177/0265532214565387>
- Linder, K. E. 2017. Fundamentals of hybrid teaching and learning. *New Directions for Teaching and Learning*, 2017 (149), 11–18. <https://doi.org/10.1002/tl.20222>
- Metsämuuronen, J. 2005. *Tutkimuksen tekemisen perusteet ihmistieteissä*. 3. laitos. Jyväskylä: International Methelp, Gummerus.
- Myhill, D. & Watson, A. 2014. The role of grammar in the writing curriculum: A review of the literature. *Child Language Teaching and Therapy*, 30 (1), 41–62.
<https://doi.org/10.1177/0265659013514070>
- Opetushallitus 2014. *Perusopetuksen opetussuunnitelman perusteet 2014*. Helsinki: Opetushallitus.

- Phipps, S. & Borg, S. 2009. Exploring tensions between teachers' grammar teaching beliefs and practices. *System* 37, 380–390. <https://doi.org/10.1016/j.system.2009.03.002>
- Rättyä, K. 2017: *Kielitiedon didaktiikkaa. Kielentäminen ja visualisointi sanaluokkien ja lauseenjäsenten opetusmenetelminä*. Väitöskirja. Helsingin yliopisto. <http://urn.fi/URN:ISBN:978-951-51-2861-4>
- Shulman, L. S. 2004. *Knowledge and teaching: Foundations of the new reform*. San Francisco: Jossey Bass.
- Tainio, L. & Marjokorpi, J. 2014: Luokanopettajaopiskelijat ja kielioppi. *Kieli, koulutus ja yhteiskunta* 5/2014. <https://jyx.jyu.fi/bitstream/handle/123456789/43543/luokanopettajaopiskelijat-ja-kielioppi.pdf?sequence=1&isAllowed=y>.
- Tainio, L. & Routarinne, S. 2012: Kieliopin ymmärtäminen ja kieliopillinen ajattelu: luokanopettajaopiskelijat lausetta hahmottamassa. Teoksessa P. Atjonen (toim.) *Oppiminen ajassa – kasvatustieteeseen. Joensuun vuoden 2011 kasvatustieteen päivien parhaat esitelmät artikkeleina* (Kasvatusalan tutkimuksia 61). Turku: Suomen kasvatustieteellinen seura, 249–263.
- Tarnanen, M., Kauppinen, M., Kainulainen, J., Hankala, M. & Aalto, E. 2023. Enhancing professional learning of primary student teachers of L1 and L2 through a hybrid learning environment. Teoksessa K. Sadeghi & M. Thomas (toim.) *Second language teacher professional development. Digital education and learning*. Palgrave Macmillan, Cham, 209–227. https://doi.org/10.1007/978-3-031-12070-1_11
- Timmis, S., Broadfoot, P., Sutherland, R. & Oldfield, A. 2015. Rethinking assessment in a digital age: opportunities, challenges and risks. *British Educational Research Journal*, 42 (3), 454–476. <https://doi.org/10.1002/berj.3215>.
- Torvelainen, P., Tarnanen, M., Aalto, E., Hankala, M. & Kainulainen, J. 2021. Luokanopettajaopiskelijat tekstien tuottajina yhden yliopistovuoden aikana. Teoksessa E. Luukka, A. Palomäki, L. Pihkala-Posti & J. Hanska (toim.), *Opetuksen ja oppimisen ytimessä*. Tampere: Tampereen yliopisto. Ainedidaktisia tutkimuksia, 19, 119–141. <http://hdl.handle.net/10138/333969>
- Torvelainen, P., Tarnanen, M., Hankala, M. & Kainulainen, J. 2019. Miten luokanopettajaopiskelijat arvioivat itseään tekstien tuottajina? Teoksessa M. Kallio, H. Krzywacki & S. Poulter (toim.) *Arvot ja arviointi*. Helsinki: Helsingin yliopisto. Ainedidaktisia tutkimuksia 16, 28–50.
- Törmäkangas, K. & Törmäkangas, T. 2009. *Osioanalyysi testien arvioinnissa*. Jyväskylä: Jyväskylän yliopisto, Koulutuksen tutkimuslaitos.
- Xu, Y. & Brown, G. T. L. 2016. Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149–162. <https://doi.org/10.1016/j.tate.2016.05.010>
- Wanner, T. & Palmer, E. 2018. Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback. *Assessment & Evaluation in Higher Education*, 43 (7), 1032–1047. <https://doi.org/10.1080/02602938.2018.1427698>.