

Pohjois-Karjalan ammattikorkeakoulun julkaisuja
A: Tutkimuksia, 5

KIELTENOPETUKSEN HAASTEET AMMATTIKORKEAKOULUISSA

Anneli Airola (toim.)

Pohjois-Karjalan ammattikorkeakoulu
Joensuu 2001

Laadun valvonta keskeistä kielitaidon arvioinnissa

Professori Sauli Takala

Arviointi on aina ollut osa inhimillistä toimintaa: omalla ja muiden toiminnalla on tavoitteita ja olemme tottunut arvioimaan niiden toteutumista. Meillä on myös suuri määrä kriteereitä toiminnan laadulle ja olemme samalla tavalla tottuneet arvioimaan mitä erilaisempien asioiden laatua ja ansiokkuutta (ks. esim. Takala 1997). Arviointi on myös aina ollut olennainen osa opetustoimintaa. Koe- ja arviointityyppien määrä on varsin suuri, ja arviointijärjestelmää laadittaessa ja kehitettäessä on tehtävä tietoisia valintoja koetyyppien suhteen.

Kielitaidon testaaajien (arvioijien) perusvelvollisuus on antaa kokeen (testin, tutkinnon) suorittajille hyvä ja tasapuolinen mahdollisuus osoittaa kielitaitoaan. Opettamisessa on keskeistä hyvien oppimistilaisuuksien luominen. Testaamisessa on vastaavasti keskeistä luoda hyvä tilaisuus osoittaa osaamista. Tämä edellyttää selkeää käsitystä kielitaidosta, jota halutaan arvioida (=mitä testataan). Tässä tulee voimakkaasti esille kieliasiantuntemus. Yhtä tärkeä on käsitys siitä, millainen arviointi on laadukasta. Tässä tulee esille mittaamisen ja testaamisen asiantuntemus (=miten arvioidaan). Kielitaidon laadukas arviointi edellyttää siis kielitaitoon ja testaamiseen liittyvää asiantuntemusta. Kielitaidon arvioinnissa tarvitaan puolestaan laadun valvontaa.

Kielitaito on käsite, joka on varsin abstraktinen ja yleinen termi. Sitä on täsmennettävä opetuksen ja testauksen tarpeisiin. Tällöin siitä tulee ns. konstrukti: kielitaito on kuvattu niin yksityiskohtaisesti, että sen pohjalta voidaan laatia mittavälineitä. Mittavälineet ovat puolestaan käsitteen ja konstruktin operationaalisia määritelmiä. Laadukas testaus edellyttää johdonmukaista kehittelyä, jossa edetään kielitaidon määrittelemisestä testeihin ja suoritusten arvosteluperusteisiin.

Mistä konstruktit tulevat? Mielestäni sekä teoriasta että käytännöstä. Kielitaitokäsitys (-konstrukti) voi mielestäni olla samantapainen erilaisissa, nykyaikaisissa viestintää painottavissa testeissä/kokeissa ja tutkinnoissa, vaikka tiettyjä painotuseroja toki voi esiintyä. Seuraavassa oma alustava luonnokseni kielitaitokäsityksen määritelmäksi. Se perustuu Euroopan Neuvoston viitekehykseen.

1. Kielitaito on pitkälti sosiaalista toimintaa - ihmisten välistä vuorovaikutusta kielen avulla.

2. Kielitaidossa on pitkälti kysymys viestinnällisestä kielellisestä kompetensista, joka ilmenee kykyä kielen tuottamiseen ja ymmärtämiseen, kykyä kielelliseen vuorovaikutukseen ja kykyä "tulkita" kielestä toiseen. Kielitaitoon sisältyy myös leksikaalinen/semanttinen ja kieliopillinen kompetenssi.
3. Viestinnällinen kielellinen kompetenssi ilmenee tarkemmin sanottuna kykyä sekä ymmärtää että tuottaa erilaista tekstiä/diskurssia keskeisiksi miellettyillä kielenkäytön alueilla/ konteksteissa (erityisesti normaali sosiaalinen kanssakäyminen, yleinen asioiden hoitaminen/ asioiminen kielen avulla ja työelämän viestintä). Edellytetään myös kykyä välittää viestejä äidinkielestä vieraaseen kieleen ja päinvastoin.

Vaikka viestintätaitoon vaikuttavat yleinen maailmantieto, sosiokulttuurinen tieto, käytännön tiedot ja taidot, sosiaalinen älykkyys, persoonallisuus, strategiat jne., niitä ei ole helppo testata saatika testata luotettavasti. Koska on hyvin vaikeaa ottaa huomioon kaikkien yksilöllisiä vahvuuksia, laajempien testien tulee eräässä mielessä olla yhtä "epäreiluja" kaikille – ts. sisältöjen tulisi olla yhtä tuttuja tai outoja kaikille.

Viestintätaitoa on mielekästä testata mahdollisimman paljon viestintätehtävien avulla.

Kullakin koe- ja arviointityypillä on omat vahvuutensa ja rajoituksensa. Onkin tarpeen tiedostaa selkeästi, ettei ole olemassa mitään koetyyppiä, joka on aina validi tai aina epäkelpo tilanteesta riippumatta. Täten esim. jonkin koetyypin (esim. monivalinnan tai tiivistelmän) vastustaminen "periaatteesta" ei ole perusteltua. Mikään koe sinänsä ei ole luotettava tai osuva irrallaan, validius ja reliiabelius ovat *toimeenpantujen kokeiden/testien* ominaisuuksia. Validius riippuu arvioinnin tarkoituksesta ja tavoitteista, kohderyhmästä ja tulosten perusteella tehtävistä ratkaisuista ja päätöksistä. *Koe ja arviointi on validi, jos sen perusteella voidaan tehdä oikeanosuvia tulkintoja, päätöksiä tai ratkaisuja.*

Validius tarkoittaa siis yksinkertaisesti sitä, että osataan tehdä *oikeita tulkintoja ja ratkaisuja tulosten perusteella*: arvioidessamme esimerkiksi puhetaitoa teemme sen puhetaidon luonteen mukaisesti; arvioidessamme kuullun ymmärtämistä menettelemme niin, että tulos heijastaa ensisijaisesti juuri kuullun ymmärtämisen olennaisia piirteitä; kieliopin hallinnan "puhtaassa" testaamisessa ei tulokseen vaikuta esim. sanaston tuntemisen ongelmat, jne. *Tuloksissa heijastuu mahdollisimman paljon juuri mitattava taito ja mahdollisimman vähän muut asiainkuulumaton tekijät.*

Yksi validiuteen vaikuttava merkittävä tekijä on reliiabelius. *Reliabelius on validiuden välttämätön mutta ei riittävä edellytys.* Koe sinänsä ei ole reliabeli, vaan kyse on myös tässä tulosten luotettavuudesta. Kyse ei ole yleispätevästä ilmiös-

tä: reliaabelius on tilannekohtaisesti arvioitava ominaisuus, joten se pitää laskea aina omassa aineistoissa, vaikka käyttää muiden tekemiä kokeita. Reliaabelius viittaa ensisijassa arviointitulosten *johdonmukaisuuteen*. Arviointi on luotettavaa, kun tulokset ovat samansuuntaiset, vaikka käytetään esimerkiksi kahta rinnakkaista testikokonaisuutta, vaikka testaus tapahtuu kahtena eri ajankohtana ja vaikka eri henkilöt arvioivat suoritukset.

Entä millainen tilanne on meillä nykyisin? Kieltenopetuksella on meillä pitkät ja vankat perinteet ja maamme kieltenopetus on seurannut varsin hyvin vaatimuksia. Kieltenopettajilla on vankka peruskoulutus ja useimmat osallistuvat aktiivisesti jatkokoulutukseen. *Siksi ei ole syytä epäillä, että kielitaidon validius olisi meillä iso ongelma*. Suurin osa opettajien ja muiden tahojen tekemää testausta on todennäköisesti nykyisin vähintään kohtuullisen validia. Nykyisin eri yhteyksissä pidetyt tekstin ymmärtämisen kokeet mittaavat varsin hyvin tekstin ymmärtämistä, puheen ymmärtämisen kokeet kohtuullisen hyvin puheen ymmärtämistä jne. Parantamisen varaa toki on ja validiuden parantaminen on ikuinen haaste, mutta ei ole syytä mihinkään suureen itsetutkisteluun ja –epäilyyn.

Reliaabelius ja validius ovat läheistä sukua toisilleen. Testiteoriassa on yleisesti hyväksytty periaate, että *reliaabelius on validiuden välttämätön mutta ei riittävä edellytys*. Mikä tahansa lause/virke voidaan usein tulkita eri tavalla. Jotkut ovat olleet taipuvaisia lukemaan tämän lausuman niin, että korostetaan painokkaasti validiuden ensisijaisuutta. Ymmärtääkseni sitä ei ole kuitenkaan tarkoitettu, vaan on haluttu korostaa reliaabeliuden varmistamisen perustavaa laatua olevaa tärkeyttä. Sana "välttämätön" on hyvin voimakas sana. Itse tulkitseen nyrkki-säännön niin, että *pitää aina ehdottomasti varmistaa, että testaus on reliaabelia*. Ylioppilastutkintolautakunnan pitkäaikaisen jäsenen ominaisuudessa olen sisäistänyt tämän voimakkaasti: kaiken tärkeän ("high stakes") arvioinnin tulee olla yhdenmukaista, luotettavaa, johdonmukaista. Kun missä yhteydessä tahansa annetaan yksilöä koskevia tärkeitä arviointeja, on suorittajilla täysi oikeus odottaa luotettavaa, yhdenmukaista kohtelua. Itse olen taipuvainen sanomaan, että kieltenopetuksessa ei tule yksilöille tärkeissä tilanteissa pyrkiä arvioimaan sellaista, jonka arviointiin ei ole kehittynyt pitkälle menevää yksimielisyyttä. Koulutuksen avulla voidaan epäilemättä kehittää jatkuvasti arviointikulttuuria, mutta *saavutettu yksimielisyyden taso on aina selvitettävä ja riittävä taso varmistettava*.

Kokeet ja arviointi ovat aina enemmän tai vähemmän reliaabeleja ja valideja, ja niihin sisältyy aina väistämättä tiettyä virhettä. Arvioinnin seuraamusten merkittävyyden perusteella voidaan ratkaista, kuinka tarkkaan arviointiin olisi pyrittävä ja toisaalta kuinka suuri virhemarginaali voidaan hyväksyä. Mitä useampaan arvosanalokkaan suoritukset tulee voida luokitella, sitä luotettavampaa arvioinnin tulee olla. Jos 3-4 luokkaa edellyttää huomattavan korkeaa luotettavuutta, 6 luokkaa edellyttää lähes täydellistä luotettavuutta.

Tehtävien tulee erotella suorittajat todellisen osaamistason mukaan. Tarvitaan eri vaikeustasoa olevia hyviä tehtäviä. Tehtävien kyky erotella kokeen suorittajia (erottelukyky) on paikallinen ominaisuus: helppo mutta hyvin tehty tehtävä erottelee heikoimmat muista ja hyvin tehty vaikea tehtävä puolestaan erottelee parhaat muista. Keskivaikea osio (ratkaisuprosentti 50) erottelee tehokkaimmin koko taitotasolla. Siksi yleensä on järkevää sijoittaa kokeeseen eniten keskivaikeita tehtäviä ja kohtalainen määrä helppoja ja vaikeita tehtäviä.

Testaamisen laadun valvonnassa on avuksi kysyä seuraavia asioita:

- Miksi arvioidaan? Mihin arviointitietoa tarvitaan? Mitä tulkintoja, ratkaisuja tai päätöksiä on tarkoitus tehdä? Tämä antaa pohjaa validiuden pohdiskelulle.
- Mitä arvioidaan? Mihin eri sisältöihin, prosesseihin, tuotoksiin, vaikutuksiin jne arviointi kohdistuu?
- Kuka arvioi? Ketkä osallistuvat arviointiin?
- Millaisen informaation pohjalla arvioidaan? Millaista tietoa hankitaan ja miten?
- Millaisia laadun kriteereitä käytetään arvioinnissa?
- Miten arviointi toteutetaan? Missä määrin arviointi tapahtuu keskitetysti ja missä määrin paikallisesti, koulukohtaisesti, luokkakohtaisesti, yksilöllisesti?
- Mikä merkitys arvioinnilla on? Millaisia vaikutuksia ja seurauksia sillä on koulutusjärjestelmälle, koululle ja yksilölle?

Hyvä arviointi edellyttää – kuten edellä on todettu – asiantuntemusta. Arvioinnin teorian perusteiden tuntemus on tarpeen kaikille kieltenopettajille. Hyväksi arviointisijaksi ei voi kuitenkaan tulla pelkän ”kirjanoppineisuuden” perusteella. On tarpeen jatkuvasti seurata omien ja muiden laatimien tehtävien toimintakelpoisuutta: mikä toimii ja mikä ei. Teoria antaa testaamiselle hyvää pohjaa mutta käytäntö antaa sille syvyyttä ja ymmärrystä.

Seuraavassa muutama vihje arviointitaidon kehittämiseksi käytännön toiminnassa. 1) Olisi hyvä arvioida ennalta tehtävien vaikeutta, esimerkiksi mitkä ovat todennäköisesti 3-5 helpointa tehtävää ja 3-5 vaikeinta tehtävää. Ennakkokäsitystä verrataan empiiriseen tulokseen ja poikkeamien syitä pohditaan tarkastelemalla tehtäviä yksityiskohtaisesti. 2) Empiirisen tuloksen tarkastelussa yksinkertaiset perusanalyysit kuten koko suorittajaryhmän pistemäärien jakautumat ja yksittäisten tehtävien ratkaisuprosentit ovat jo varsin hyödyllisiä. Enemmän saa irti, jos jakaa suorittajajoukon kolmeen osaan: heikoin neljännes, paras neljännes ja keskitasoiset. Pistemäärien jakaumat ja ratkaisuprosentit lasketaan ryhmittäin. Yksittäinen tehtävä on toiminut hyvin, jos sen ratkaisuprosentti kasvaa johdonmukaisesti tason mukaan. Tällaista osioanalyysia varten on olemassa useita tietokoneohjelmia, jotka antavat runsaasti tietoa kokeista. 3) Monivaikeissa tehtävissä on hyvä tarkastella vaihtoehtojen toimivuutta. Nyrkkisääntönä voi

pitää sitä, että jokaista vaihtoehtoa on valinnut vähintään 5% kokeeseen osallistujista. Hyvien ryhmään kuuluvien tulisi lisäksi valita muita useammin oikea vaihtoehto ja ei juuri lainkaan vääriä vaihtoehtoja. Heikompien ryhmään kuuluvien tulisi puolestaan valita suhteellisesti useammin vääriä vaihtoehtoja.

Maassamme on varsin paljon kielitaidon arvioinnin asiantuntemusta, joten opettajien täydennyskoulutukseen on saatavissa hyviä kouluttajia. Käynnissä oleva AMKKIA-projektissa kehitetään nimenomaan ammattikorkeakoulujen kieltenopetukseen soveltuvia arviointijärjestelmiä. Projektissa on Opetusministeriön tuen avulla tilaisuus kehitellä monia ratkaisuja, jotka tulevat kohottamaan konkreettisesti arvioinnin laatua, mitä tämän esityksen otsikossa peräänkuulutetaan.

Lähteet:

Takala, S. 1997. Vieraan kielen kehittymisen arviointiperusteita. Julkaisussa R. Jaku-Sihvonen (toim.) Onnistuuko oppiminen – oppimistuloksien ja opetuksen laadun arviointiperusteita peruskoulussa ja lukiossa. Opetushallitus: Arviointi 3/1997, 87-114.