

Perusteita kielitaidon arvioinnin laadun kehittämiseen ja valvontaan

Sauli Takala
Jyväskylän yliopisto
Soveltavan kielitutkimuksen keskus

Kielitaidon arviointia tehdään monissa yhteyksissä ja moniin tarkoituksiin. Siksi arvioinnin muodotkin vaihtelevat suuresti. Kun kyseessä on opetukseen läheisesti liittyvä arviointi, arviointi voi olla oleellinen osa opetusta, mistä on monta suurta etua. Toinen ääripää on sellainen arviointi, jolla on tärkeitä – joskus jopa peruuttamattomia – seurauksia yksilölle. Kaikelle arvioinnille voidaan kuitenkin asettaa perusvaatimukseksi, että se antaa arvioitavalle hyvän tilaisuuden osoittaa kielitaitoaan ja että se tuottaa luotettavaa tietoa todellisesta kielitaidosta niin, että arviointiin perustuvat ratkaisut, päätökset ja tulkinnat ovat oikeaanosuvia. Arvioinnin tulee olla myös mahdollisimman taloudellista ja tehokasta. Kaikkeen arviointiin sisältyy virhettä, jonka muodot ja suuruus olisi tiedostettava nykyistä paljon paremmin ja tuotava esille nykyistä yleisemmin ja avoimemmin. Arvioinnissa esiintyy mielestäni liikaa kritiikitöntä luottamusta omaan ja muiden arviointitapaan ja -taitoon. Tarvitaan tietoa huomiota arvioinnin laadun systemaattiseen kehittämiseen ja valvontaan. Edellä sanottu ei liene vaikeasti ymmärrettävää, mutta käytännössä laadukas arviointi on huomattavan vaikeaa.

1. Arvioinnille asetettavat laatuvaatimukset

Tässä lyhyessä artikkelissa käsitelen kaikelle kielitaidon arvioinnille asetettavia laatuvaatimuksia. Yhteisenä haasteena kaikessa kielitaidon arvioinnissa on sen tietoinen ja systemaattinen kehittäminen ja valvonta eli meillä on sosiaalinen tilaisuus yhteisen, pitkäjänteisen Laatu kielitaidon arviointiin –projektin käynnistämiseen.

Hyvä opettaminen (mm. uusien tietojen ja taitojen edistäminen) ja hyvä arviointi ovat yhteydessä toisiinsa mutta ne eivät ole kuitenkaan sama asia. Arviointi on suppeampi toiminta-ala kuin opettaminen. Hyvälle arvioinnilla asetetaan myös aivan spesifejä, juuri nimenomaan arviointia koskevia laatuvaatimuksia. Toisaalta arvioinnilla on erilaisia tehtäviä ja konteksteja, ja siksi eri tyyppiselle arvioinnille asetetaan jossakin määrin erilaisia vaatimuksia. Kuitenkin kaiken arvioinnin tulee olla reliaabelia, validia ja tehokasta. Seuraavassa käsitelen lyhyesti näitä teemoja pyrkien antamaan konkreettista tietoa ja sovellusvihjeitä.

Arviointi on aina ollut osa inhimillistä toimintaa ja se on aina ollut olennainen osa myös opetustoimintaa. Systemaattisen koulutuksen parintuhannen vuoden aikana on kehittynyt monenlaisia tapoja testata ja arvioida opetuksen tavoitteiden saavuttamista. Vaikka eri aikoina on esiintynyt erilaisia arviointitapoja, osa niistä on tullut käyttöön kohtalaisen myöhään. On kuitenkin syytä muistaa, että aikojen kuluessa on kehitetty monia käyttökelpoisia arviointimuotoja. Meillä ei ole mitään syytä väheksyä entisten sukupolvien viisautta ja aikaansaannoksia ja kuvitella, että vain uusia arviointimenetelmiä kehittämällä teemme laadukasta arviointia.

Koe- ja arviointityyppien määrä on kuitenkin jatkuvasti kasvanut, ja samalla on kasvanut tarve harkita, millainen arviointimenetelmä kussakin tilanteessa on sopiva. Arviointijärjestelmää laadittaessa ja kehitettäessä on tehtävä tietoisia valintoja koetyyppien suhteen.



Kullakin koe- ja arviointityypillä on omat vahvuutensa ja rajoituksensa. Siten on syytä tiedostaa selkeästi, ettei ole olemassa mitään yleispätevää koetyyppiä, joka on aina validi tilanteesta riippumatta. Toisaalta testauskäytännöstä on tuskin osoitettavissa koetyyppejä, jotka olisivat jo perusolemukseltaan täysin sopimattomia, epävalideja, kielitaidon arvioimiseen. Validius riippuu arvioinnin tarkoituksesta ja tavoitteista, kohderyhmästä ja tulosten perusteella tehtävistä ratkaisuisista ja päätöksistä. Koe sinänsä ei ole validi vaan validius on pikemminkin toimeenpannun kokeen ominaisuus ("test-in-action"). *Arviointi on validia, jos sen perusteella voidaan tehdä oikeaanosuvia tulkintoja, päätöksiä tai ratkaisuja.* Validiuspäätelmä edellyttää koko arviointijärjestelmän laadun pohtimista.

Yksi käsitevalidiuteen keskeisesti vaikuttava tekijä on reliaabelius. *Reliaabelius on validiuden välttämätön mutta ei riittävä edellytys.* Sama mitä edellä todettiin validiudesta, pitää paikkansa myös reliaabeliudesta. Koe sinänsä ei ole reliabeli, vaan kyse on myös tässä tulosten luotettavuudesta ja – kuten validiuden kohdalla – kyse ei ole yleispätevästä ilmiöstä. Reliaabelius on *tilannekohtaisesti* arvioitava ominaisuus: kunkin arviointitilanteen luotettavuus on selvitettävä erikseen. Reliaabelius viittaa ensisijassa *arviointitulosten johdonmukaisuuteen*. Arviointi on luotettavaa, kun tulokset ovat samansuuntaiset, a) vaikka käytetään esimerkiksi kahta rinnakkaista testikokonaisuutta, b) vaikka testaus tapahtuu kahtena eri ajankohtana (ei kuitenkaan kovin etäällä toisistaan, koska välillä voi tapahtua kypsymistä, oppimista ja/tai unohtamista) ja c) vaikka eri henkilöt arvioivat suoritukset. Toisin kuin validiuden osalta, reliaabelius voidaan ilmaista numeerisesti (0-1; mitä lähempänä ykköstä, sen luotettavampi arviointi). Reliaabeliuden tasoon palataan artikkelissa tuonnempana.

Kokeet ja arviointi ovat siten aina enemmän tai vähemmän reliaableja ja valideja, ja tiettyä virhettä sisältyy väistämättä niihin. Arvioinnin seuraamusten merkittävyyden perusteella voidaan ratkaista, kuinka tarkkaan arviointiin olisi pyrittävä ja toisaalta kuinka suuri virhemarginaali voidaan hyväksyä.

Reliaabelius

- luotettavuus, johdonmukaisuus
- ei oikeastaan voida arvioida yhden mittauksen perusteella (mutta yhden mittauksen mittarin johdonmukaisuus/sisäinen konsistenssi/homogeenisuus on kyllä järkevä likiarvo)
- laskennallinen kerroin mahdollinen
- vaihteluväli normaalisti 0.00 – 1.00
- kyseessä on ensisijaisesti tulosten luotettavuus (ei mittavälineen)
- on tilannekohtainen – ei yleispätevä, yleinen - ominaisuus; arvioitava joka tilanteessa ja aineistossa erikseen
- välttämätön mutta ei riittävä edellytys validiudelle

Validius

- pätevyys, "osuvuus"
- ei yleensä voi ilmaista laskennallisesti (tosin reliaabeliuden neliötä voidaan pitää validiuden maksimaalisena arvona), vaan edellyttää validiusväitettä tukevien perusteiden pohdiskelua; validiusdiskurssi on parhaimmillaan uskottavaa monipuolista argumentointia
- kyseessä on tuloksien pohjalla tehtävien johtopäätösten, ratkaisujen ja päätösten pätevyys; Arviointi on validia, jos sen perusteella voidaan tehdä oikeaanosuvia tulkintoja, päätöksiä tai ratkaisuja. Validiuspäätelmä edellyttää koko arviointijärjestelmän laadun pohtimista.
- validius on reliaabeliuden tapaan kontekstisidonnainen, ei yleispätevä/yleinen ominaisuus; se riippuu arviointitarkoituksesta, kielitaitokäsityksestä, kohderyhmästä ja mittaustavasta



En tässä artikkelissa käsittele laajemmin validiutta. Oletan, että kokeenlaatijalla on selkeä kuva siitä kielitaidosta, jota hän haluaa mitata. Kieltenopettajien perusammattitaitoon kuuluu tällainen tieto. Lähdeluettelossa mainituissa julkaisuissa on runsaasti virikkeitä validiuden pohtimiseen.

Kokeet ja arviointi ovat siten aina enemmän tai vähemmän reliaabeleja ja valideja, ja niihin sisältyy siis väistämättä tiettyä virhettä. Arvioinnin seuraamusten merkittävyyden perusteella voidaan ratkaista, kuinka tarkkaan arviointiin olisi pyrittävä ja toisaalta kuinka suuri virhemarginaali voidaan hyväksyä. Virheen suuruuden havainnollistamiseksi tarvitaan avuksi keskivirheen laskemista. Sen laskemiseksi tarvitaan tietoa koetulosten hajonnasta ja reliaabeliudesta.

Seuraavassa taulukossa on laskettu valmiiksi keskivirheen suuruus erikokoisille hajonnoille ja reliaabeliustasoille.

Hajonta (pisteinä)	Reliaabeliuskerroin					
	.95	.90	.85	.80	.75	.70
30	6.7	9.5	11.6	13.4	15.0	16.4
28	6.3	8.9	10.8	12.5	14.0	15.3
26	5.8	8.2	10.1	11.6	13.0	14.2
24	5.4	7.6	9.3	10.7	12.0	13.1
22	4.9	7.0	8.5	9.8	11.0	12.0
20	4.5	6.3	7.7	8.9	10.0	11.0
18	4.0	5.7	7.0	8.0	9.0	9.9
16	3.6	5.1	6.2	7.2	8.0	8.8
14	3.1	4.4	5.4	6.3	7.0	7.7
12	2.7	3.8	4.6	5.4	6.0	6.6
10	2.2	3.2	3.9	4.5	5.0	5.5
8	1.8	2.5	3.1	3.6	4.0	4.4
6	1.3	1.9	2.3	2.7	3.0	3.3
4	.9	1.3	1.5	1.8	2.0	2.2
2	.4	.6	.8	.9	1.0	1.1

95%:n luottamusväli, eli se alue jonka sisällä todellinen pistemäärä on 95%:n varmuudella, saadaan kertomalla saatu keskivirhe luvulla 1.96 (eli noin kahdella). Pistemäärän alaraja saadaan vähentämällä saatu luku kustakin pistemäärästä ja yläraja puolestaan lisäämällä se kuhunkin pistemäärään. Näin pistemäärän varmuusvälin suuruus riippuu testin reliabiliteetista ja sen hajonnasta. Esimerkiksi jos testin reliabiliteetti on .95 ja sen hajonta 14, testipistemäärän 50 saaneen todellinen pistemäärä (95%:n varmuudella) on siis 50 +/- 6 pistettä (3.1 x 1.96 pyöristettynä) eli se voi vaihdella 44:stä 56:een pisteeseen. Jos saman testin reliabiliteetti olisikin .80, samalla hajonnalla sama pistemäärä eli 50:n todellinen pistemäärä vaihtelisikin lähes 13 pistettä kumpaankin suuntaan eli välillä 37 – 63.

On tärkeää myös tiedostaa, kuinka läheisesti reliaabelius liittyy kykyyn erotella taitotasoja. Mitä useampia taitotasoja halutaan käyttää, sitä luotettavampaa arvioinnin tulee olla. Jos arvioinnin reliaabelius on .70, luotettavien suoritusosien maksimimäärä on 1,5 (.75 – 1,7). Vasta .80 takaa luotettavan luokittelun kahteen tasoon. Seuraavassa lisätietoa asiasta: reliaabeliuskerroin .90 – 3 luokkaa, .94 – 4 luokkaa, .96 – 5 luokkaa, .97 – 6 luokkaa. Luvut osoittavat hyvin konkreettisesti, kuinka vaativa tehtävä on luokitella suorittajat pariinkin luokkaan – puhumattakaan kuuteen tasoluokkaan - yhdellä ja samalla kokeella.



Erottelukyky

- jos kaikkien tulisi osata jokin asia ja näin voidaan kohtuudella odottaakin, erottelukyky ei ole keskeinen vaatimus. Ihannetilanne olisi, että kaikki osaisivat kaiken. Koska tämä on harvoin – tuskin koskaan – realistinen odotus, erottelukyky on käytännössä tärkeä vaatimus.
- erottelukyky tarkoittaa osion/tehtävän kykyä erotella kokeensuorittajia eri suoritustasoilla
- tehtävien kyky erotella kokeen suorittajia (erottelukyky) on paikallinen ominaisuus:
- helppo mutta hyvin tehty tehtävä pystyy erottelemaan 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.
- erotteluindeksi voi vaihdella teoriassa $-1.00/+1.00$. Se perustuu tavallisesti osion ja kokeen summapistemäärän väliseen korrelaatioon, ts. siihen missä määrin osio erottelee samansuuntaisesti kuin kaikki muut tehtävät yhdessä.
- erotteluindeksin tulee olla positiivinen ja mielellään ainakin .20 mutta melko tyytyväisiä voidaan alkaa olla vain osioihin, joiden erotteluindeksi on vähintään .30

Matala erotteluindeksi (osio ei erottele hyviä ja heikkoja tehokkaasti) voi johtua mm. seuraavista tekijöistä:

- Osio on kovin vaikea tai kovin helppo.
- Osio on epäselvä tai harhaanjohtava.
- Osio mittaa jotakin muuta, kuin kokeen useimmat muut osiot.

Osion ratkaisuprosentti	Maksimaalinen erotteluindeksi
95% tai 5%	.19
90% tai 10%	.37
85% tai 15%	.55
80% tai 20%	.74
75% tai 25%	.92

Yllä olevan taulukon luvut (osion maksimaalinen erottelukyky) saavutetaan, jos ja vain jos kaikki hyvät kokelaat (ratkaisuprosentti n. 75 tai parempi) ovat ratkaisseet tehtävän oikein ja kaikki heikot kokelaat (ratkaisuprosentti n. 25 tai alhaisempi) ovat ratkaisseet sen väärin. Taulukko osoittaa siis, että vain erittäin hyvin tehdyt kovin helpot tai vaikeat osiot voivat toimia kohtuullisesti erottelumielessä. Jopa sellaiset tehtävät, jotka neljännes tai kolme neljännestä ovat ratkaisseet oikein, eivät voi erotella maksimaalisesti (1.00). Tämä on mahdollista vain osioille, jotka noin puolet on osannut ratkaista oikein.



Reliaabeliusvaatimus suhteessa tehtävään päätökseen/ratkaisuun/tulkintaan

Korkeaa reliaabeliutta vaaditaan, kun

- päätös on tärkeä
- päätös on lopullinen
- päätöstä ei voi purkaa/perua
- päätöstä ei voi tukea muulla tiedolla
- päätös koskee yksilöitä
- päätöksellä on pysyviä seurauksia

Korkeaa reliaabeliutta tarvitaan esimerkiksi valintatilanteissa ja tärkeissä tasokokeissa.

Matala reliaabelius on siedettävissä, kun

- päätös ei ole tärkeä
- päätös tehdään varhaisessa vaiheessa
- päätös voidaan purkaa/perua
- päätös saa tukea muusta tiedosta
- päätös koskee ryhmiä
- päätöksellä on tilapäisiä vaikutuksia

Esimerkiksi opettajan päätös opetuksen etenemisestä

Reliaabeliuteen vaikuttavia tekijöitä

- tehtävien/osioiden määrä:
⇒ yleensä suurempi tehtävämäärä on eduksi reliaabeliudelle, koska näin saadaan parempi näyte mitattavasta taidosta; satunnaisvaikutus vähenee
- osioiden vaikeustaso:
⇒ koko taitotasoaluetta ajatellen paras tulos saavutetaan, kun noin puolet osaa ratkaista sen
- pistemäärien hajonta:
⇒ yleensä - mitä suurempi hajonta, sitä korkeampi reliaabelius. Jos testi järjestettäisiin uudelleen, vähäisiä eroja sisältävässä testissä yksilöiden järjestys vaihtelee enemmän kuin suurihajontaisessa kokeessa, jossa yksilöiden järjestys säilyy paremmin samana, ja tämä näkyy edellisessä tapauksessa alhaisempana reliaabeliutena.
- objektiivisuus, eli arvioinnin yksimielisyys:
⇒ avoimissa tehtävissä voi objektiivisuutta lisätä tehtävän huolellisella ja selkeällä muotoilulla ja selkeillä arviointiohjeilla; testausolosuhteiden vakiominen.

Mitä johtopäätöksiä edellä esitetystä voidaan tehdä kielikokeen laadinnalle?

- 1) Ratkaisevan tärkeää on yksittäisten osioiden laatu: niiden erottelukyvyn tulisi olla mahdollisimman korkea. Jokaisen osion laadintaan, arviointiin ja muokkaamiseen (ja jos suinkin mahdollista – esitestaamiseen) kannattaa kiinnittää suurta huomiota. Hyvä ja huolellinen esityö kannattaa aina.
- 2) Osioita tulisi olla mahdollisimman paljon, koska pidempi koe on lyhyttä luotettavampi ja virheen osuus pienenee yleensä pituuden ja siis näytön kasvaessa. Kannattaa miettiä



keinoja, jolla saadaan mahdollisimman paljon näytteitä osaamisesta. Sellainen arviointi ei ole järkevää, jossa joudutaan käyttämään paljon aikaa ja kuitenkin saadaan vähän näytettä osaamisesta. Kannattaa harkita (mahdollisimman) lyhyitä ja monia tehtäviä laajojen ja harvojen tehtävien sijasta.

On hyvä olla tietoinen, että

- jos kaksinkertaistamme kokeen suorittajan tuottamaa aineistoa (esim. kaksinkertainen tehtävämäärä), testattavista johtuva virhelähde (mittausvirhe) pienenee tyypillisesti 30%:lla
- jos arvioitsijan on pakko käyttää vain kokonaislukuja (esim. taitotaso 1,2,3,4,5,6 , eikä hienompaa luokitusta (esim. 1+, 1++, 2-) sallita, luokittelun tarkkuus todennäköisesti kärsii. Luokittelu on luotettavinta keskialueella ja epäluotettavampaa luokkien rajoilla. Ennen lopullista tasoarviota on siis suositeltavaa käyttää apuna plussia ja miinuksia. Tästä on erityisesti hyötyä, jos suoritus kaksoisarvioidaan.
- saavutettua arvioinnin yksimielisyyden tasoa voidaan helposti yliarvioida. Esimerkiksi jos käytetään asteikkoa 1-5 tai 1-6 ja jos kaksi arvioitsijaa arvioi täysin samalla tavalla 70% kaikista suorituksista, heidän antamansa arviot eroaisivat harvoin enemmän kuin yhdellä tasolla ja suurempi ero esiintyisi vielä harvemmin. 70% tuntuu varsin hyvältä tulokselta. Jos testin suorittajat kuitenkin edustavat normaalisti jakautuvaa joukkoa, suurin osa suorituksista sijoittuisi skaalan keskivaiheelle. Tämä merkitsee, että jos joku arvioitsijoista ei edes lukisi tai kuuntelisi suorituksia vaan antaisi aina keskitason arvion, arvioinnin yksimielisyys olisi varsin hyvä. Jos noin puolet suorittajista saisi keskiarvosanan, saavutettaisiin tällöinkin 50%:nen täydellinen yksimielisyys. Näin ollen 70%:n yksimielisyys ei itse asiassa olisikaan kovin kehuttava tulos.
- jos oppilaalta vaaditaan useiden osavaatimusten täyttämistä, tämä lisää virhepäätöksen todennäköisyyttä. Tämä on syytä pitää mielessä, kun asetetaan vaatimustasoja. Näin esimerkiksi jokaisessa osataidossa asetettava minimitalo ei liene suositeltavaa.
- monet väärinluokitukset ovat väistämättömiä rajatapauksissa. Luokan keskivaiheilla luokitus on luotettavampaa. Kahden ”askeleen” virhettä ei voida hyväksyä kuin hyvin pienelle määrälle ”oppilaita (esim. 5%)

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?

Tällainen kysely ja itsekysely edesauttaa refleктоivaa toimintaa.



Lopuksi

Laatua kielitaidon arviointiin – projekti tarkoittaa, että arvioinnissa tarvitaan kunnianhimoista laatupyrkimystä mutta samalla nöyrää rehellisyyttä sen suhteen, että parhaasta yrityksestämme huolimatta arviointiin sisältyy aina virhettä. Kunnianhimo voikin realistisesti kohdistua vain virheen minimoimiseen. Se edellyttää tietoisuutta virhelähteistä ja virheen suuruuden selvittämistä sekä taitoa vähentää virheen määrää ennakoita järkevin ratkaisuin. Olen pyrkinyt edellä tuomaan esille eräitä seikkoja, joilla laatutietoisuutta yrittää kohottaa. Kirjallisuusluettelossa on mainittu eräitä muita tähän teemaan liittyviä julkaisuja. AMKKIA-projektissa on myös kehitetty monia menettelytapoja, joilla voidaan edesauttaa yhteismitallista arviointia. Projektin jatkosuunnitelmiin sisältyy laajamittainen kielitaidon arviointikoulutus, jonka avulla saatuja kokemuksia voidaan saattaa laajaan käyttöön.

Hyvä arviointi edellyttää – kuten edellä on todettu – asiantuntemusta. Arvioinnin teorian perusteiden tuntemus on tarpeen kaikille kielenopettajille. Hyväksi arvioitsijaksi 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. Teoria on käytännön paras ystävä ja käytäntö on teorian paras testaaja.

Kirjallisuutta

Huhta, A. & Takala, S. 1999. Kielitaidon arviointi. Julkaisussa K. Sajavaara & A. Piirainen-Marsh (toim.) Kielenoppimisen kysymyksiä. Jyväskylän yliopisto, Soveltavan kielentutkimuksen keskus, 179-228.

Luoma., S. 2001. What does your test measure? Construct definition in language test development and validation. (<http://www.solki.jyu.fi>)

Takala, S. 1994. Arviointi - ongelma ja mahdollisuus Julkaisussa P. Linnakylä, P. Pollari & S. Takala (toim..) Portfolio arvioinnin ja oppimisen tukena: Jyväskylän yliopisto, Kasvatustieteiden tutkimuslaitos, 1-8.

Takala, S. 1994. Arvioinnin uusia tuulia. Julkaisussa S. Takala (toim.) Arviointi ja koulutuksen laadun kehittäminen Jyväskylän yliopisto, Kasvatustieteiden tutkimuslaitos, 1 - 5.

Takala, S. 1995. Arviointikulttuurin kehittäminen. Julkaisussa K. Salmio & K. Lindroos Himberg (toim.) Askelia yleissivistävän koulutuksen arviointiin. Helsinki: Opetushallitus. Arviointi 2/95, pp. 6-17.

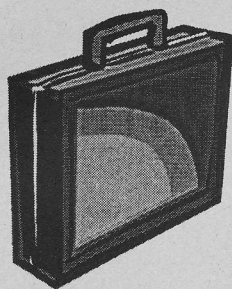
Takala, S. 1996. Suoritusarviointi puntarissa: mahdollisuuksia ja ongelmia valtakunnallisissa kokeissa. Julkaisussa P. Pollari, M. Kankaanranta & P. Linnakylä (toim..) Portfolion monet mahdollisuudet. Jyväskylän yliopisto: Kasvatustieteiden tutkimuslaitos, 207-222

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



AMMATTIKORKEAKOULUJEN SUULLISEN KIELITAIDON KEHITTÄMISPROJEKTI 2000-2001

I ARVIOIJAN SALKKUUN LIITTYVÄT ARTIKKELIT



Anneli Airola	Pohjois-Karjalan ammattikorkeakoulu
Ari Huhta	Jyväskylän yliopisto. Soveltavan kielentutkimuksen keskus
Sari Luoma	Jyväskylän yliopisto. Soveltavan kielentutkimuksen keskus
Sauli Takala	Jyväskylän yliopisto. Soveltavan kielentutkimuksen keskus



Suullisen kielitaidon arvioijan salkku / Ammattikorkeakoulujen suullisen kielitaidon arvioinnin kehittämisprojekti
2000 – 2001.

Anneli Airola (Koordinaattori)

Johanna Lemmikki ammattikorkeakoulun