

# ONNISTUUKO OPPIMINEN

-oppimistuloksien ja opetuksen laadun  
arviointiperusteita peruskoulussa ja lukiossa

Ritva Jakku-Sihvonen (toim.)

Opetushallitus  
1997

*Sauli Takala*

## VIERAAN KIELEN KEHITTÄMISEN ARVIOINTIPERUSTEITA

### 1. LÄHTÖKOHTA

Tässä artikkelissa keskitytään otsikon mukaisesti käsittelemään kielitaidon arvioinnin perusteita sekä viimeaikaisia kehityslinjoja ja otetaan kantaa siihen, mihin suuntaan kielitaidon arviointia tulisi maassamme kehittää. Tämä merkitsee sitä, että lukijalle ei vastaavasti anneta yksityiskohtaisia ohjeita kielikokeiden laadinnasta ja tulosten arvioinnista. Olen useassa yhteydessä korostanut, kuinka tärkeätä on, että maassamme jatkuu perinne, jonka mukaan opettajalla on keskeinen rooli arvioinnissa. Opettajien tulee olla vaativankin arvioinnin asiantuntijoita, mikä taas edellyttää riittävän hyvää arvioinnin perusteiden tuntemusta. Sen luomiseen pyrkii tämä artikkeli. Tavoitteena on vahvistaa kieltenopettajien arviointitietämystä, jotta he voivat jatkossakin vastata arvioinnin kasvaviin haasteisiin. Samalla vältettäisiin se monessa muussa maassa vallitseva tai voimistuva suuntaus, jossa ulkopuolisilla on keskeinen asema arvioinnissa.

Käytännön vihjeitä arvioinnin perusteiden soveltamiseen käytännössä saa useista julkaisuista, joihin viitataan tuonnempana. Lähivuosina järjestetään KIMMOKE-hankkeen yhteydessä arviointikoulutusta, ja kieltenopettajien arviointitaitojen koulutuksen järjestämisestä kiinnostuneiden käytettävissä on nopeasti kasvava arvioinnin asiantuntijoiden joukko. Suoritusarviointiin voi perehtyä yleisiin kielitutkintoihin liittyvän kirjallisuuden ja mm. tilattavissa olevien esittelytietien avulla. Suoritusarvioinnissa korostuu se, mitä kielellä osataan tehdä, ja siksi se edustaa suuntaa, johon mielestäni peruskoulun ja lukion kieltenopetuksen arvioinninkin tulisi kehittyä.

### 2. ARVIOINNIN HAASTEET JA MAHDOLLISUUDET

Tavoitteet (aikomukset, odotukset, ihanteet, arvot, pyrkimykset) ovat keskeisiä kaikessa ihmisen toiminnassa. Toisaalta toimintamme ohjaamiseksi tarvitsemme palautetietoa siitä, miten edistymme pyrkimyksissämme. Pyrimme siis tiettyihin tavoitteisiin ja arvioimme niiden saavuttamista. Samalla arvioimme itseämme suhteessa omiin tavoitteisiimme ja suhteessa muihin. Arviointi on olennainen osa elämää: arvioimme kaikkea ja kaikkia monista syistä ja monin perustein (Scriven 1980, 1991; Shadish ym. 1991, Takala 1994). Moninainen arviointi on siten tavallista myös koulutuksessa, opetuksessa ja opiskelussa.

Tavoitteita asetetaan instituutioille ja niissä toimiville yksilöille. Esimerkiksi yhteiskunnallisesti tärkeälle organisaatiolle, koululaitokselle, asetetaan tavoitteita laeissa, asetuksissa, sopimuksissa ja ohjeissa. Toisaalta on selvää, että oppilailta on itsellään erilaisia tavoitteita ja odotuksia, jotka voivat osittain

poiketa koulun toiminnalle asetetuista tavoitteista. Koulun, opetuksen ja oppimisen kannalta keskeinen ohjenuora on opetussuunnitelma. Myös koulun, opetuksen ja oppimisen arvioinnille opetussuunnitelma on keskeinen lähtökohta. Se antaa suuntaa erityisesti sille, mitä tulisi arvioida, mutta tietyssä määrin myös sille, miten tulisi arvioida. Se, miten arvioidaan, riippuu kuitenkin osittain myös arvioinnin luonteesta: sen tulee täyttää määrätyt laatuvaatimukset. Kun arviointi vastaa opetussuunnitelmaa ja täyttää lisäksi sille asetettavat laatuvaatimukset, siitä voi olla monenlaista hyötyä kehitettäessä valtakunnallista, koulukohtaista ja oppilaskohtaista opetussuunnitelmaa. Arviointi on siis olennainen osa opetussuunnitelman ja opetuksen kehittämistä.

Sekä kouluyhteisö että yksilöt (opettajat ja oppilaat) voidaan mieltää pääasiassa joko kohteeksi, jota tavoitteiden saavuttamiseksi täytyy ulkoapäin säädellä ja valvoa, tai toimijoiksi, jotka pitkälti säätelevät itse toimintaansa (mm. Darling-Hammond, 1993; García & Pearson 1994; Glaser & Silver 1994; Kohonen 1994; Lieberman 1990). Säätelyn ja itseohjautuvuuden suhde on vaihdellut eri maissa eri aikoina, mutta viime aikoihin saakka useissa maissa, mm. Suomessa, on ulkopuolinen sääntely ollut vallitsevampi suuntaus. Tällä hetkellä todennäköisesti useimmat koulutuksen tutkijat eri puolilla maailmaa ovat sitä mieltä, että itse itseään arvioiva ja säätelevä koulu on ulkoa säädeltyä tuloksellisempi, kehityskykyisempi ja miellyttävämpi elämisen, työn tekemisen ja oppimisen paikka. Siksi on tärkeää, että opettajilla ja kouluilla on pääosa arvioinnissa.

### 3. ARVIOINTIKULTTUURIN KEHITTÄMISEN AJANKOHTAISIA HAASTEITA

Kaikelle arvioinnille voidaan ajankohdasta riippumatta siis asettaa seuraavat vaatimukset: arvioinnin tulee kohdistua oikein (eli antaa vastauksia olennaisiin kysymyksiin), olla luotettavaa ja tuottaa käyttökelpoista tietoa taloudellisesti. Tällä hetkellä korostetaan uusina haasteina seuraavia seikkoja:

- Arvioinnin ja opetuksen tulee olla sopusoinnussa keskenään. Molempien tulee heijastaa samanlaista omakohtaista peruskäytännöstä oppimisesta ja opittavista asioista.
- Arvioinnin tulee vastata mahdollisimman hyvin todellisuutta eli sitä, miten tietoja ja taitoja käytetään myös koulun ulkopuolella.
- Arvioinnin seuraukset yksilöille tulee ottaa aikaisempaa paremmin huomioon. Erityisesti tulee varoa sitä, ettei arviointi ja arvostelu heikennä oppilaiden minäkuvaa. Pikemminkin tulee kaikin keinoin pyrkiä löytämään kunkin oppilaan vahvuuksia ja korostaa oppimisen edistymistä ja näin auttaa rakentamaan positiivista minäkuvaa.
- Arvioinnin tulee tuoda esille entistä paremmin myös oppimisessa käytetyt prosessit ja strategiat.
- Arvioinnissa tulee perinteellisen oppilasarvostelun lisäksi antaa tilaisuus oppilaiden itsearviointille.
- Opettajien ja kouluyhteisön tulee myös aikaisempaa enemmän arvioida oman toimintansa lähtökohtia ja toimintatapoja.

Kaiken kaikkiaan tarvitaan arviointikulttuurin monipuolistamista ja uusien haasteiden mukaisten arviointimenetelmien kehittämistä. Niiden perusteita pohditaan tämän artikkelin eri luvuissa.

#### 4. ARVIOINNIN MONITAHOISUUS

Arviointia käsittelevä tutkimuskirjallisuus tuo selvästi esiin arvioinnin monitahoisuuden. Arvioinnilla tarkoitetaan yleensä toimintaa, jossa määritetään erilaisten asioiden ansiokkuutta tai arvoa eli laatua. Arviointi edellyttää puolestaan lähtökohdakseen näkemystä laadun ulottuvuuksista ja tasoista: millaisista piirteistä laatu käy ilmi ja millaisia tasoja siinä voi olla?

Arvioinnissa on tarpeen kysyä seuraavia asioita (vrt. Leimu 1993; Linnakylä 1994b; Messick 1992):

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

Kun jatkuvasti pohdimme vastauksia näihin kysymyksiin ja kehitämme uusia arviointimenetelmiä, edistämme arviointikulttuurin kehitystä. Kehittyneessä arviointikulttuurissa osataan arvioida tärkeitä asioita niin, että kulloinkin tarvittava arviointitieto saadaan tuotettua luotettavasti ja taloudellisesti, samalla kun tietoisesti pyritään maksimoimaan arvioinnin myönteisiä vaikutuksia ja minimoimaan sen haitallisia sivuvaikutuksia.

Kuten edellä todettiin, arvioinnilla tarkoitetaan tavallisesti toimintaa, jonka tarkoituksena on määrittää, kuinka hyvä (laadukas, ansiokas, arvokas, sopiva jne.) jokin asia, esine, suoritus, toiminta jne. on. Ansiokkuuden määrittelyyn tarvitaan perusteita (kriteereitä), ts. on oltava eritelty käsitys hyvän laadun ominaisuuksista (piirteistä) eli siitä, miten laatu ilmenee. Kielenopetuksessa kuten kaikessa muussakin koulutuksessa opetus suunnitelmallisesti toimivat tavallisesti tällaisena kriteeristönä tai ainakin niiden perustana. Viime aikoina on erityisesti aikuiskoulutuksen puolella yleistynyt toisenlainen arviointitapa: on määritelty kielitaidon tasoasteikkoja, ja henkilöiden kielitaitoa arvioidaan vertaamalla sitä näihin asteikkoihin, joissa kuvataan, mitä henkilö osaa tehdä kielellä.

Arviointia voidaan luonnehtia myös erilaisissa tehtävissä todettavien suoritusten systemaattisena tarkkailuna, jonka tarkoituksena on kerätä tietoa suoritustasosta ja raportoida sitä sopivassa muodossa siitä kiinnostuneille osapuol-

lille. Tiedon vaihto helpottuu, jos arviointimenetelmä on, kuten mm. Euroopan neuvoston toiminnassa on korostettu, avoin ja ymmärrettävä ("läpinäkyvä"), sekä sisäisesti johdonmukainen ja jos sen antama tulos on verrattavissa muihin arviointimenetelmiin ja on niiden kanssa vertailukelpoinen.

Arviointitiedosta ovat kiinnostuneet useat tahot kukin omista syistään. Tällaisia asianosaisia ovat mm.

- yksilöt (oppilaat, opettajat, vanhemmat)
- koulutuspalveluja antavat oppilaitokset
- paikalliset, alueelliset ja valtakunnan tason kouluviranomaiset
- koulutusalan luottamushenkilöt ja poliittiset päätöksentekijät
- erilaiset eturyhmät (työelämän osapuolet, media, suuri yleisö)
- kansainväliset järjestöt (OECD, Euroopan unioni, YK)

Kasvava kiinnostus luotettavaan arviointitietoon merkitsee usein sitä, että halutaan varmistua, että kouluissa tehdään oikeita asioita riittävän tehokkaasti. Äärimmillään tämä kulminoituu tilivelvollisuuden korostumisena: koulujenkin edellytetään pystyvän osoittamaan uskottavasti, että käytetyt resurssit tuottavat hyvää tulosta. Tämä näyttää olevan yleismaailmallinen suuntaus ja se voimistuu todennäköisesti entisestään. Opettajien ja koulujen etujen mukaista on siis itsekkin pyrkiä huolehtimaan siitä, että ne pystyvät tarvittaessa antamaan "läpinäkyvää" ja uskottavaa arviointitietoa. Tämä merkitsee käytännössä mm. sitä, että pitäisi pystyä entistä määrätietoemmin arvioimaan, mitä oppilaat osaavat kielellä tehdä: mitä he pystyvät ymmärtämään ja mitä itse esittämään opiskeltavalla kielellä. Nykyisin oppimääriin perustuva arviointi ja arvosanat eivät valitettavasti ole kovin "läpinäkyviä" koulun ulkopuolisille tahoille.

## 5. ARVIOINNILLE ASETETTAVAT KESKEISET VAATIMUKSET: RELIABELIUS JA VALIDIUS

Koska arvioinnilla voi olla merkittäviä seuraamuksia yksilöille ja oppilaitoksille, sille asetetaan tiukkoja vaatimuksia. Arvioinnin tulee olla luotettava (reliabelia). Koska arviointi on viime kädessä tulkintaa, reliabelius tarkoittaa käytännössä usein nimenomaan tulkinnan luotettavuutta eli sitä, että arvioija on johdonmukainen arvioidessaan samoja suorituksia eri aikoina ja riittävän yksimielinen muiden arvioitsijoiden kanssa. Arvioinnin tulee olla myös validia, mikä tarkoittaa ensijassa tulkintojen ja johtopäätösten pätevyyttä. Mikään arviointimenetelmä ei ole a priori validia tai epävalidia, vaan validius riippuu aina siitä, mitä tulkintoja, johtopäätöksiä tai ratkaisuja halutaan tehdä. Arvioinnin validius joudutaan aina toteamaan tarkastelemalla arviointitilannetta alusta loppuun: mikä oli tavoitteena, miten arviointiaineisto kerättiin ja miten luotettavasti se arvioitiin. Toisin kuin luotettavuustarkastelu, jossa voidaan vedota mm. tulkinnan yksimielisyyttä osoittaviin lukuihin, validiustarkastelu edellyttää huolellista argumentointia, jossa validiusväitettä tuetaan esittämällä, mitä näyttöä sille on. Validiusväite tarvitsee tuekseen sekä teoreettista että empiiristä näyttöä. Siksi usein esitetty käsitys, että esimerkiksi jokin koe on validi,

jos se mittaa, mitä on tarkoitus mitata, on oikeasuuntainen mutta aivan liian kapea näkemys. Koe on vain yksi osatekijä arvioinnissa ja luo pohjaa varsinaiselle tulkinnalle, se ei sinänsä ole vielä validi tai epävalidi.

Validius on keskeinen vaatimus. Riittävä reliabelius on osatekijä validiuuden varmistamisessa. Jos arvioinnissa korostetaan voimakkaasti reliabeliutta, mitataan helposti vain irrallisia asioita ja rajataan arviointitilanne niin kapeaksi, että vain yksi ainoa suppea vastaus on oikea ja kaikki ovat samaa mieltä oikeasta vastauksesta. Hyvin toimiva monivalintakoe on esimerkki tästä: jopa kone voi tuottaa kokeesta saatavan tuloksen. Reliabeliutta yli kaiken korostavan tulisi kuitenkin miettiä, miten hyvin tiukasti rajattu tehtävä vastaa sitä, miten tietoa ja taitoa tulee osata käyttää "todellisessa elämässä". Todellisessa elämässä joudutaan aina tekemään ratkaisuja epätäydellisen ja tiettyä subjektiivisuutta sisältävän tiedon varassa. Esimerkiksi valintatilanteessa osa tulee aina valituksi, vaikka olisi muita yhtä hyviä tai jopa parempia ehdokkaita ja osa tulee hylätyksi vaikka ansaitsisi tulla valituksi. Tälle ei voida mitään, koska mittaus ja arviointi sisältää aina virhettä. Tavoitteena tulee olla vähentää virhettä ja kehittää keinoja, joilla päätöksien osuvuutta voidaan parantaa. Siksi arvioinnissa tulee kehittää sekä arvioinnin etiikkaa että arvioitsijoiden ammattitaitoa edistäviä menettelytapoja. Arvioinnissa tulee olla vahvoilla sekä eettisesti että tiedollisesti.

## 6. ARVIOINTIMENETELMIEN LAAJA KIRJO

Kuten edellä esitetystä on käynyt ilmi, arviointi palvelee niin monia kohderyhmiä, että vuosien varrella on kehittynyt erilaisia arviointimenetelmiä:

- *Normiviitteinen vs kriteeriviitteinen arviointi*. Edellisessä on kysymys siitä, että yksilön saama arvosana riippuu siitä, kuinka hyvin hän on suoriutunut suhteessa muihin kokeeseen osallistuneisiin. Tätä myös suhteelliseksi arvioinniksi nimitettyä menettelyä käytetään mm. ylioppilastutkinnoissa, jossa tietty osa kokelaista saa korkeimman arvosanan, tietty osa seuraavaksi korkeimman arvosanan jne. Kriteeriviitteistä arviointia käytetään puolestaan mm. yleisissä kielitutkinnoissa, joissa tutkintoon osallistuvien suorituksia verrataan taitotason kriteereihin ja saatujen arvosanojen jakauminen on täysin riippuvainen siihen osallistuvien osaamistason jakaantumisesta. Jollakin tutkintokerralla voi olla keskimääräistä parempia osallistujia ja tämä heijastuu hyvien arvosanojen painottumisena. Kehitys näyttää kulkevan kohti kriteeriviitteistä arviointia.
- *Saavutusmittaus vs taitotason arviointi*. Edellisellä tarkoitetaan tavallisesti jonkin kurssin tai oppimäärän tavoitteiden saavuttamisen arviointia. Kriteerinä ovat siis asetetut tavoitteet. Taitotason mittaus perustuu tasojen määrittelyyn, jossa ylintä tasoa edustaa kieltä äidinkielenään puhuvan hyvin koulutetun (ja joskus jopa kielen ammattilaisen) taitotaso ja alinta tasoa kielen alkeet joten hallitsevan kielitaito.
- *Diagnostinen vs summatiivinen vs summatiivinen arviointi*. Diagnostisella arvioinnilla pyritään erityisesti saamaan selville oppimisvaikeuksia aiheuttavia tekijöitä mutta myös esimerkiksi lähtötilannetta, jotta opetus voidaan aloittaa sopivalta tasolta. Laadukas diagnosointi edellyttää vankkaa

tietoa kielitaidon olemuksesta ja sen sisäisestä rakenteesta sekä erityistä vaikeutta aiheuttavista seikoista. Formatiivinen arviointi on opetuksen kuluessa tapahtuvaa arviointia, jotta voidaan saada selville, miten osaaminen edistyy yleensä, ja päättää, voidaanko edetä eteenpäin vai tarvitseeko koko ryhmä tai jotkut yksilöt kertausta. Summatiivinen arviointi on oppimistulosten kartoitusta kurssin tai oppimäärän lopussa ja se vaikuttaa tavallisesti merkittävästi arvosanaan.

- *Standardikokeet vs ei-standardisoidut kokeet.* Standardikokeet edellyttävät huolellisesti valittujen ja laadittujen koetehtävien testaamista edustavalla joukolla ja suorituskriteerien laadintaa tämän tiedon pohjalta. Vieraiskielissä ei maassamme tällaisia kokeita ole koskaan ollut olemassa. Laajamittaisessa käytössä oleva TOEFL-koe on esimerkki standardikokeesta. Maassamme käytettävät kokeet ovat joko opettajien tai esim. oppimateriaalien tekijöiden tekemiä kokeita. Ylioppilastutkinnon kielikokeista voisi jatkokehittelyllä mahdollisesti saada standardisoituja kokeita: osallistuuhan niihin jopa kymmeniä tuhansia kokeilaita (koko perusjoukko, eikä vain suppea ja usein epäedustava otos!).
- *Sisäinen vs ulkoinen arviointi.* Tämä liittyy läheisesti edelliseen, mutta kun edellä oli kyse olennaisesti hyvien normitietojen käytettävyydestä, tässä on kyse siitä, suorittaako arvioinnin jokin oppilaitoksen ulkopuolinen taho vai se itse.
- Sisäisessä arvioinnissa puolestaan voidaan erottaa *itsearviointi, toveriarviointi ja opettajan arviointi.*
- Arvioinnin seuraamusten perusteella erotetaan koetilanteet, joista englannin kielessä käytetään nimitystä *high stakes vs low stakes* -arviointi. Edellisessä ovat "panokset korkeita" ja jälkimmäisessä seuraamukset ovat vähemmän kohtalokkaita. Ylioppilastutkintoa epäilemättä pidetään ratkaisevan tärkeänä (kriittisenä) näytön paikkana.
- Arvioinnin parissa työskentelevät tekevät myös eron *tutkintojen ja ohjelmien arvioinnin* välillä. Tutkinto edellyttää, että kaikki sen suorittavat antavat näytteen osaamisestaan ja useimmiten kaikki suorittavat samat tehtävät. Tutkinnosta saa todistuksen. Ohjelma-arviointi taas ei tuota todistuksia yksilöille. Esimerkiksi peruskoulun lyhyen kielen valtakunnallisen osaamistason selvittämiseksi tarvitaan vain suppeahko oppilasnäyte (n. 100 koulua) mutta suuri joukko tehtäviä. Nämä hajoitetaan useiksi erilaisiksi tehtäväkokonaisuuksiksi, jotka jaetaan satunnaisesti otoksen tulleiden luokkien oppilaille. Ohjelma-arviointia ei voi kunnolla suorittaa, jos sama suppeahko tehtävistö esitetään kaikille oppilaille. Ohjelma-arviointi voi olla myös kansainvälistä (mm. IEA-tutkimukset).

## 7. KOMMUNIKATIIVISUUS KIELITAIIDON ARVIOINNISSA

Kieltenopetuksessa on kommunikatiivisuus painottunut voimakkaasti viime aikoina. Omissa opetussuunnitelmissamme tämä painotus on ollut vallalla 1970-luvun lopusta, ja Suomi tunnetaankin maana, jossa kommunikatiivisuus otettiin viralliseksi tavoitteeksi ensimmäisten joukossa, luultavasti jopa ensimmäisenä Euroopassa. Tosin kommunikatiivisuuden merkityksestä ei liene kos-

kaan vallinnut mitään yksiselitteistä tulkintaa. Alderson (1981) on käsitellyt kriteereitä, joilla voidaan arvioida kommunikatiivisia kielikokeita. Hänen systemaattisesta lähestymistavastaan saa hyvää pohjaa kommunikatiivisuuden pohtimiselle kielitaidon arvioinnissa. Alderson esittää joukon kysymyksiä:

*a. Millainen on kokeen näkemys kielestä?*

- Mitä 'kielen osaamisella' tarkoitetaan kokeessa?
- Pidetäänkö kieltä joukkona erillisiä osioita?
- Missä määrin koesuoritus heijastaa todellisessa elämässä esiintyviä suorituksia?
- Pitääkö kokeeseen osallistuvien tehdä jotakin kielellä?
- Mittaako koe kykyä toimia tietyillä sosiolingvistisillä alueilla?
- Perustuuko koe kommunikaatiomalliin?

*Missä määrin koe mittaa suorittajan kykyä tunnistaa ja tuottaa sosiolingvistisesti asianmukaista kieltä, kun otetaan huomioon*

- viestintätilanne
- viestinnän aihe
- viestinnän tarkoitus
- viestintäkanava (suullinen/kirjallinen)
- viestinnässä ennestään tutuksi oletetut asiat
- viestinnän muodollisuusaste
- viestintään osallistuvien roolit
- viestintään osallistuvien status
- viestintään osallistuvien asennoituminen.

*Missä määrin koe ottaa huomioon, että viestintä*

- perustuu vuorovaikutukseen
- on pitkälti ennustamatonta
- tapahtuu aikapaineessa
- tapahtuu tietyssä tarkoituksessa
- perustuu toimintaan/käyttäytymiseen
- ei välttämättä perustu kokonaan kieleen.

*Toisin sanoen:*

- Ovatko kokeen suorittajien reaktiot ennustettavissa?
- Mitataanko kompleksisia taitoja?
- Onko tilanne realistinen?
- Mitataanko kykyä tulkita alkuperäisiä viestejä?
- Mitataanko kykyä tuottaa omaperäisiä viestejä?
- Edellyttääkö koe luovaa kielenkäyttöä?

*Mikä 'merkitys' on kokeen mukaan?*

- staattinen, sanoihin sisältyvä
- vaihteleva, kontekstista riippuva

- neuvotteluun perustuva, kaikista vuorovaikutustilanteesta vaikuttavista tekijöistä riippuva.
- Ottaako koe huomioon sen, että kieli on redundanttista?  
 Onko kokeen kielinäyte edustava vai epätavallinen, sepitetty?  
 Missä määrin koe kattaa kielitaidon relevantit piirteet?

*b. Mikä on kokeen näkemys oppijasta?*

- Painottuuko kokeessa liikaa taitohierarkian alapää?  
 Edellyttääkö koe yleisiä kognitiivisia taitoja (yleistä taustatietoa, ymmärtämistä, päättelyä)?  
 Edellyttääkö koe suorittajan affektiivista involvoitumista (erityisesti henkilöiden välisessä vuorovaikutuksessa)?  
 Vastaako koe ajateltujen suorittajien lähtötasoa (mm. tietojen, taitojen ja affektiivisten piirteiden osalta)?  
 Ottaako koe huomioon kokeen suorittajan odotukset (esim. kielitaidon tarpeet, käsitys kielitaidosta)?  
 Salliiko koe erilaisten suorittajien osoittaa tasapuolisesti taitojaan vai suosiiko se jotakuta suorittajatyyppejä?  
 Miten syntyperäinen kielenpuhuja suoriutuisi kokeesta?

*c. Mikä on kokeen näkemys kielen oppimisesta?*

- Olettaako koe, että kielen oppiminen merkitsee sitä, että oppii hallitsemaan kielellisiä ongelmia?  
 Vastaako koe kielen oppimisen tavoitteita ja käytänteitä, ts.
- Testataanko kieltä samalla tavalla kuin sitä opiskellaan?
  - Vastaako koe suorittajan tavoitteena olevaa kielitaitoa?
  - Vastaako koe suoritustason taustalla olevaa kompetenssinäkemystä?
  - Vastaako osakokeiden painotus kielen oppimisen painotuksia?

*d. Taustatiedon asema*

- Sisältyykö kokeeseen "ulkopuolisia" tekijöitä (kulttuuri, asiantietoja)? Voidaanko ne sulkea pois?  
 Suosiiko koe tietyn tyyppistä tietoa?  
 Tulisiko kokeessa olla "neutraali" sisältö? Onko se mahdollista?  
 Voidaanko sisältö erottaa kielestä?  
 Entä jos suorittaja tietää, mitä sanoa mutta ei tiedä, miten sen sanoisi?  
 Jos mittaamme myös asianhallintaa edellyttävää viestintätaitoa, pitäisikö alan asiantuntijan osallistua kokeen laadintaan ja suorituksen arviointiin?  
 Kielikokeiden laatijoille olisi erittäin hyödyllistä paneutua Aldersonin kysymyksiin ja tehdä tietoisia ratkaisuja niiden perusteella.

## 8. ITSEARVIOINTI ITSEOHJAUTUVUUDEN VÄLINEENÄ

Oppijoiden itseohjautuvuus on korostunut viime aikoina. Sitä pidetään merkittävänä kasvatuksen kannalta, mutta sen katsotaan myös vaikuttavan keskeisesti oppimisen tehokkuuteen. Itseohjautuvuus oppimisessa on sekä tärkeä



tavoite ja tulos että keino. Oppilaiden valmius itsenäiseen oppimiseen riippuu siitä, miten kehittyneitä ovat heidän (kognitiiviset ja metakognitiiviset) kykynsä a) suunnitella, prosessoida ja valvoa/ohjata oppimistoimintoja ja b) säädellä omia tunteuksiaan ja omaa motivaatiotaan. Itseohjautuvuus edellyttää siis suunnittelu-, toiminta-, itsesäätely- ja arviointitaitoja. Tilanne lienee valitettavasti pitkälti sellainen, että monet oppilaat eivät tiedä tai tiedosta opetuksen tavoitteita. Itsearviointi voisi auttaa tässä.

Itsearviointia voidaan edistää itsekyselyn avulla (Takala 1992, 74–76, 86–100). Itsearviointi on paikallaan ennen toimintaa, sen aikana ja sen jälkeen.

*Oppilaan* itsekysely ja -arviointi *ennen* toimintaa/tehtävien tekemistä (lähtödiagnoosi suunnittelun pohjaksi):

- Mitä osaan (tiedän asiasta jne.) jo tässä vaiheessa?
- Miten suhtaudun opittavaan asiaan? Miten se minua kiinnostaa?
- Mitä mahdollisuuksia minulla tulee olemaan toteuttaa omia valintojani?
- Mitä hankaluuksia minulla tulee mahdollisesti olemaan opiskeluni ohjauksessa ja siihen sitoutumisessa?
- Mitä tulen valitsemaan, miten opiskelemaan, missä järjestyksessä ja missä aikataulussa?

Toiminnan *aikana* (opiskelun tarkkailu ja ohjailu):

- Miten opiskelemiseni sujuu? Mitä vahvuuksia ja heikkouksia osaamisessani on tässä vaiheessa?
- Millainen on kiinnostukseni ja sitoutumiseni?
- Tarvitseanko opettajan, tovereiden tai vanhempieni apua?
- Mihin minun tulee jatkossa kiinnittää huomiota?

Toiminnan *lopussa* (toiminnan ja oppimisen arviointi):

- Miten hyvin olen saavuttanut asetetut tavoitteet ja omat tavoitteeni?
- Mitä opin itsestäni? Mitkä ovat vahvat puoleni oppijana? Missä minulla on parantamisen varaa?
- Mitä tietoja ja taitoja opin?
- Osasinko ottaa riittävästi vastuuta omasta oppimisestani?

Kuten edellä todettiin, myös opettajien ja kouluyhteisön tulisi arvioida omaa toimintaansa sen lisäksi, että huomiota kiinnitetään oppilaiden toiminnan ja edistymisen arviointiin.

*Opettajan* itsekysely ja -arviointi *ennen* toimintaa (lähtödiagnoosi suunnittelun pohjaksi) voi koskea seuraavia asioita:

- Mitä oppilaat todennäköisesti osaavat (tietävät asiasta jne) jo tässä vaiheessa?
- Miten oppilaat todennäköisesti suhtautuvat opittavaan asiaan? Miten se minua itseäni kiinnostaa?
- Mitä mahdollisuuksia minulla tulee olemaan toteuttaa omia valintojani?
- Mitä hankaluuksia minulla tulee mahdollisesti olemaan opiskelun ohjauksessa?
- Mitä tulen valitsemaan, miten organisoimaan opiskelun, missä järjestyksessä ja missä aikataulussa? Mitkä ovat perusteeni? Mitä vaihtoehtoja on?

Toiminnan *aikana* (opiskelun tarkkailu ja ohjailu):

- Miten opiskelu sujuu? Mitä vahvuuksia ja heikkouksia oppilaiden osaamisessa on tässä vaiheessa?
- Millainen on oppilaiden kiinnostus ja sitoutuminen?
- Tarvitseeko joku oppilas/jotkut oppilaat apua?
- Mihinkin minun tulee jatkossa kiinnittää huomiota?

Toiminnan *lopussa* (toiminnan ja oppimisen arviointi):

- Miten hyvin oppilaat ovat saavuttaneet tavoitteet? Miten hyvin olen saavuttanut omat tavoitteeni? Miksi kaikkia tavoitteita ei saavutettu?
- Mitä opin itsestäni? Mitkä ovat vahvat puoleni opettajana? Missä minulla on parantamisen varaa?
- Mitä tietoja ja taitoja opin?
- Annoinko oppilaille riittävästi vastuuta omasta oppimisestaan?
- Mitä olisin voinut tehdä toisin?

Seuraavassa olen referoinut Euroopan neuvoston eri dokumenteissa (erityisesti Brian Northin laatimissa esitelmissä ja muistioissa) esitettyjä itsearviointikuvauksia, jotka on liitetty Euroopan neuvoston omiin taitotasaluonnehdintoihin. Olen merkinnyt myös tason, jota luonnehdinta lähinnä vastaa yleisissä kielitutkinnoissa.

| Taito        | Waystage<br>(YKI 2?)  | Waystage Plus<br>(YKI 3 ?)   | Threshold<br>(YKI 4+?)   | Threshold Plus<br>(YKI 5?)   | Vantage<br>(YKI 6+?)  |
|--------------|---|--|--|--|---|
| Ymmärtäminen | Ymmärrän, mitä minulle henkilökohtaisesti sanotaan selkeästi tavallisessa juttelussa. Ymmärrän, jos juttelukumppani näkee vähän vaivaa. | Pystyn yleensä tunnistamaan, mistä ihmiset juttelevat, kunhan he puhuvat selkeästi. Ymmärrän yleensä minulle kohdistettua selkeää yleiskielistä puhetta tutuista asioista, jos voin pyytää toistoa aika ajoin. | Pystyn yleensä seuraamaan laajemman keskustelun pääkohtia, jos puhe on selkeätä yleiskieltä. Ymmärrän minulle kohdistettua selkeätä puhetta jokapäiväisessä juttelussa, vaikka joskus joudun pyytämään jonkin sanan tai ilmauksen toistoa. | Pystyn pienin ponnistuksin saamaan selville, mistä ihmiset yleensä puhuvat, vaikka en itse pysty tehokkaasti osallistumaan, jos syntyperäiset eivät helpota kieltään. Ymmärrän minulle kohdistetun selkeän puheen, jos se ei sisällä kovin idiomaattista kieltä. | Pystyn seuraamaan usean henkilön käymää keskustelua useimmista yleisistä asioista ja oman alan asioista ja tunnistamaan tarkasti mielipiteiden pääkohdat. Ymmärrän yksityiskohtaisesti, mitä minulle sanotaan, vaikka esiintyisi taustahälyä. |
| Juttelu      | Selviydyn hyvin lyhyistä sosiaalisen vuorovaikutuksen tilanteista mutta en ymmärrä kylliksi, jotta voisin itse ylläpitää juttelua.      | Pystyn osallistumaan lyhyisiin jutteluihin rutiinitilanteissa ja tavallisista aiheista.  | Voin ryhtyä juttelemaan ilman ennalta valmistelua tutuista asioista. Pystyn ylläpitämään juttelua tai keskustelua, mutta minua ei aina ole helppo seurata, kun yritän sanoa tarkasti mitä haluaisin ilmaista.                              | Pystyn toimimaan pätevästi epämuodollisissa tilanteissa. Pystyn ylläpitämään keskustelua vieraiden henkilöiden (mm. vieraiden) kanssa asiaankuuluvista asioista tai omasta alastani tai itseäni kiinnostavista asioista.   | Pystyn osallistumaan aktiivisesti laajempaan keskusteluun useimmista yleisimmistä aiheista, myös hälyisissä oloissa. Minun vuokseni ei syntyperäisten ole tarpeen helpottaa kieltään.   |

| Taito        | Waystage<br>[YKI 2?]   | Waystage Plus<br>[YKI 3 ?]   | Threshold<br>[YKI 4+?]   | Threshold Plus<br>[YKI 5?]  | Vantage<br>[YKI 6+?]  |
|--------------|--|--|--|---|---|
| Asiointi     | Pystyn viestimään yksinkertaisissa ja rutiinitehtävissä, joissa välitetään tiettyä tietoa tutuista ja rutiinimaisista työhön ja vapaa-aikaan liittyvistä asioista. Pystyn asiomaan kaupassa, postissa, pankissa jne. | Pystyn tekemään itseni ymmärretyksi ja vaihtamaan tietoa tavallisista asioista, jotka liittyvät jokapäiväiseen elämään, mm. matkustaessani, ostoksilla ja ravintoloissa.   | Pystyn hoitamaan useimmat matkustamiseen ja ulkomaisen viranomaisten kanssa asioimiseen liittyvät tilanteet. Pystyn hankkimaan ja välittämään selväpiiristä tietoa ja hankkimaan tarvittaessa tarkempaa tietoa.                                  | Pystyn ymmärtämään, vaihtamaan, tarkistamaan ja tiivistämään selväpiiristä tietoa ja hoitamaan myös sellaisia tilanteita, jotka eivät ole aivan rutiinimaisia, vaikka joskus minun täytyy pyytää toistoa, jos puhutaan nopeasti ja pitkään. | Pystyn ymmärtämään ja vaihtamaan yksityiskohtaista tietoa, selittämään ongelmaa tai tekemään selväksi, että erimielisyystilanteissa toisen osapuolen on tehtävä myönnytyksiä.   |
| Keskustelu   | Pystyn keskustelemaan yksinkertaisesti jokapäiväisistä käsitännön kysymyksistä, kun minulle puhutaan suoraan selkeästi ja hitaasti.  | Pystyn sanomaan, mitä ajattelen asioista, kun minulle puhutaan suoraan ja jos voin pyytää tarvittaessa pääkohtien toistoa.   | Pystyn keskustelemaan minua kiinnostavista asioista. Pystyn ilmaisemaan uskomuksia, mielipiteitä sekä kohteliaasti yksi- ja erimielisyyttä.  | Pystyn selittämään, miksi jokin on ongelma ja vertaamaan erilaisia ehdotuksia tai ratkaisuja sekä kommentoimaan muiden näkemyksiä.  | Pystyn osallistumaan aktiivisesti epämuodolliseen ja muodolliseen keskusteluun ja esittämään selkeästi oman näkemyseni. Pystyn kehittämään omaa näkemystäni antamalla asiaankuuluvia selityksiä, perusteluita ja kommentteja. |
| Kuvaileminen | Pystyn kuvailemaan perheittäni, elinolojani, koulustani sekä nykyistä tai viimeisintä työpaikkaani.  | Pystyn kuvailemaan ympäristöni ja taustani jokapäiväisiä piirteitä (ihmisiä, paikkoja, työtä, opiskelua jne.). Pystyn kuvaamaan suunnitelmia ja järjestelyjä, tapoja ja rutiineja, mistä pidän ja en pidä, menneitä tapahtumia ja omakohtaisia kokemuksia. | Pystyn kuvailemaan yksityiskohtaisesti erilaisia tapoja ja todellisia tai kuviteltuja tapahtumia sekä kertomaan tarinoita ja kirjojen/ elokuvien juonia kuvaillen tunteita ja reaktioita. Pystyn myös kuvaamaan unelmia, toiveita ja tavoitteita | Pystyn antamaan selväpiirteisiä kuvauksia erilaisista tutuista aiheista minua kiinnostavilta aloilta. Pystyn kuvaamaan yksityiskohtaisesti ennustamattomia tapahtumia (mm. onnettomuuksia).   | Pystyn antamaan selkeän ja yksityiskohtaisen kuvauksen useista erilaisista asioista minua kiinnostavilta aloilta.   |

Edellä olevassa taulukossa on esitetty taitotasosteikkoja, joiden avulla oppilas (yksin tai yhdessä tovereiden tai opettajan kanssa) voi pohdiskella, millä tasolla hänen kielitaitonsa on. Suomessa on tehty useita tutkimuksia, jotka osoittavat, että aikuiset pystyvät varsin hyvin arvioimaan oman kielitaitonsa tason. Olisi kiinnostavaa selvittää, missä määrin oppilaat pystyvät samaan. Onhan monesti todettu, että suomalaiset oppilaat moniin muihin maihin verrattuna pikemminkin aliarvioivat kuin yliarvioivat omaa osaamistasoaan.

Seuraavassa esitetään muutamia esimerkkejä, joissa keskitytään erilaisten varsin tarkkojen tehtävien arviointiin.

Alla on lueteltu joukko asioita, joita voit joutua ilmaisemaan \_\_\_\_ksi. Arvioi, miten hyvin selviäisit kustakin tehtävästä:

| Ilmaistava asia/viestintätehtävä                                      | Hyvin helposti | Vaatii vähän ponnistelua | Vaatii suuria ponnisteluja | En osaisi |
|---|----------------|--------------------------|----------------------------|-----------|
| Kertoa, mikä viikonpäivä on kyseessä                                  |                |                          |                            |           |
| Ilmaista lukumäärää   |                |                          |                            |           |
| Tilata aterian ravintolassa   |                |                          |                            |           |
| Vaihtaa valuuttaa pankissa  |                |                          |                            |           |
| Ostaa vaatteita kaupasta  |                |                          |                            |           |
| Kertoa taksinkuljettajalle, mihin osoitteeseen haluan ajettavan       |                |                          |                            |           |
| Kertoa omasta taustastani (syntymäaika, asuinpaikka, kotiolot...)     |                |                          |                            |           |
| Kertoa harrastuksistani   |                |                          |                            |           |
| Kertoa, mitä minulle eilen tapahtui                                   |                |                          |                            |           |
| Kertoa tulevaisuuden suunnitelmistani                                 |                |                          |                            |           |
| Pyytää anteeksi myöhästymistäni ja selittää sen syy                   |                |                          |                            |           |
| Kirjoittaa henkilökohtaisen kirjeen                                   |                |                          |                            |           |
| Perustella mielipiteitäni   |                |                          |                            |           |
| Protestoida ja pitää puoleni, kun minua ei kohdella mielestäni oikein |                |                          |                            |           |

Tällaisia väittämiä (osaan tehdä/sanoa....) voisi mielellään kehittää systemaattisesti kattamaan nuorille relevantit kielenkäyttötilanteet ja kielitaidon osa-alueet ja totuttaa oppilaat arvioimaan taitojaan. On todennäköistä, että "osaan tehdä" -osioissa itsearviointi on tarkempaa kuin edellä esitetyn yleisemmän tasokuvauksen avulla.

## 9. ARVIINTIMENETELMÄT: MONIVALINNASTA PORTFOLIOON

Ns. autenttinen (vaihtoehtoinen) arviointi on nostanut uudelleen esille tarpeen eritellä systemaattisesti, miten erilaiset testaus- ja arviointimenetelmät eroavat toisistaan. Tällaista erittelyä on tehnyt mm. John B. Carroll useassa yhteydessä (referoitu Takala 1980, 35–38). Tämän artikkelin kirjoittaja on myös kehittänyt arviointitilanteen typologiaa (Takala 1976). Alla esittelen Snow'n (1993) hiltaintain esittämän alustavan kuvauksen, jossa hän esittelee tuottamiskoetyyppien jatkumoa. Lukija voi luetteloa lukiessaan miettiä, mihin painopiste sijoittuu suomalaisessa arviointikäytännössä ja hänen omassa työssään.

1. Monivalinta
  - a) valitse yksi vaihtoehto
  - b) esittää vaihtoehtojen todennäköisyysjakauma
2. Monivalinta, jota edeltää vastauksen rakentaminen
  - a) palauttaa mieleen/muokata tietoa
  - b) järkeillä tiedon pohjalta

3. Aukon täyttäminen/cloze
  - a) sanan täydentäminen
  - b) fraasin täydentäminen
4. Lyhyt vastaus/laajahko täydentäminen
  - a) tuottaa lause
  - b) tuottaa kokonainen kappale
5. Ongelmatehtävä
  - a) ratkaisun tuottaminen esitettyyn ongelmaan
  - b) ratkaisun selittäminen
6. Omin sanoin esittäminen/tiivistäminen ("teach-back procedure")
  - a) selittää käsite/menettelytapa
  - b) selittää rakenne/järjestelmä
7. Pitkä kirjoitelma/Demonstraatio/Projekti
  - a) tuottaa annetun aiheen mukaisesti
  - b) tuottaa ilman ennalta rajattua aihetta
8. Yllä mainittujen kokoelma, portfolio ym.

Bennett (1993) on esittänyt hieman Snow'n luokittelusta poikkeavan koe-tyyppien luokittelun. Esittelen tämän luokittelun kattavasti, koska se antaa hyvän kuvan mittauksen vaihtoehdoista. Esitän osan esimerkeistä alkuperäisessä muodossa, koska suomenkielisen version kehittäminen vaatisi melkoisesti lisätyötä.

0. *Monivalinta*: Tähän luokkaan kuuluvat osiot edellyttävät, että kokeen suorittaja valitsee vastauksen suppeasta vaihtoehtojoukosta.

Esimerkki: Valitse sana, joka parhaiten sopii lauseen merkitykseen  
 Unable to focus on specific points, he could talk only about \_\_\_\_\_;  
 indeed, his entire lecture was built around vague ideas.  
 (A) personalities  
 (B) statistics  
 (C) vulgarities  
 (D) particulars  
 (E) abstractions

(Oikea vastaus lienee E.)

1. *Valinta/Tunnistaminen*: Tätä luokkaa luonnehtii se, että valitaan yksi tai useampi vastaus suuremmasta joukosta. Monivalintaan verrattuna laaja valintamahdollisuus vähentää suuresti arvaamismahdollisuutta. Ihanteellisessa muodossa tehtävän suorittaja rakentaa mielessään vastauksen eikä pelkästään tunnista sitä. Esimerkkeinä voidaan mainita sanaluettelo, asiaankuulumattoman tekstin yliviivaaminen tai poistaminen ja tietokonenäytöllä suullisesti annettujen suunnistusohjeiden merkitseminen.

Esimerkki: Poista tarpeettomat sanat seuraavasta kappaleesta:  
 Andy Razaf is not *la quickly recognizable/name that is familiar to most people*. Yet Razaf wrote the lyrics to at least 500 or more songs, including

*the words* to the popular "Ain't Misbehain' ", "Honeysuckle Rose", and "Stompin' at the Savoy" *as well*. The American-born son of an *upper class* African nobleman, he still continues to be overshadowed by his composer-collaborators *who worked with him*, Fats Waller and Eubie Blake.

(Todennäköiset ongelmalliset kohdat on merkitty tässä kursiiivilla.)

2. *Uudelleenjärjestäminen*: Myös tässä luokassa vastaukset valitaan laajemmasta luettelosta. Tässä tapauksessa tehtävänä on kuitenkin panna osat oikeaan tai toiseen järjestykseen. Esimerkkeinä voidaan mainita anagramman tekeminen, lauseiden paneminen loogiseen järjestykseen, luettelossa olevien osien luokittelu, kuvien asettaminen järjestykseen ja palapelin kokoaminen.

Esimerkki: Järjestä seuraavat sanat uudelleen niin, että ne muodostavat täydellisen ja mielekkään lauseen. Aloita lauseen ensimmäinen sana isolla kirjaimella ja merkitse sen loppuun piste. Muita välimerkkejä ei tarvita.

a and be both can comedy enlightening entertaining good

(Hyvä vastaus lienee: A good comedy can be both entertaining and enlightening./A good comedy can be both enlightening and entertaining.)

3. *Korvaaminen/korjaaminen*: Tämä osiotyyppi edellyttää, että kokeen suorittaja korvaa esitetyn oikealla vastauksella. Esimerkkeinä voidaan mainita mm. kirjoitusvirheiden tai kielioppivirheiden korjaaminen, sopivampien ilmaisujen keksiminen, lauseiden yhdistäminen tai virheellisten tietokoneohjelmien korjaaminen.

Esimerkki: Yhdistä alla olevat kaksi lausetta yhdeksi kieliopillisesti oikeaksi lauseeksi, joka välittää saman merkityksen kuin alkuperäinen lausepari.

1. Stephen King is the author of numerous horror novels.

2. Many fans of Stephen King assume that he is as crazy as some of his characters.

(Hyvä vastaus lienee: Many fans of Stephen King, who is the author of numerous horror novels, assume that he is as crazy as some of his characters.)

4. *Täydentäminen*: Tässä osiotyypissä tehtävänä on vastata oikein epätäydelliseen ärsykkeeseen. Cloze, lauseen täydennys, vain yhtä oikeaa vastausta edellyttävä matematiikan tehtävä, progressiiviset matriisit ja uuden kohdan merkitseminen graafiin annettujen lukujen perusteella.

Esimerkki: Täydennä seuraavassa lauseessa oleva aukko yhdellä sanalla, joka tekee lauseesta kieliopillisesti ja loogisesti täydellisen.

Melodramas, \_\_\_\_\_ present stark contrasts between good and evil, are popular forms of entertainment because they offer audiences a world where there is moral certainty.

(Hyvä täydennys lienee *which*. Perusteluna on se, että kyseessä aivan ilmeisesti on relatiivilause, ja se ei ole välttämätön vaan voisi muodostaa

oman päälauseen; tällainen relatiivilause erotetaan englannin kielessä pilkulla. Irrallisen relatiivilauseen relatiivipronomini on englannin kielessä which silloin, kun viitataan esineisiin, asioihin tms. – ts. ei henkilöihin.)

5. *Kokonaisen vastauksen tuottaminen (construction)*: Kun edellinen koetyyppi edellyttää annetun ärsyksen täydentämistä, tässä luokassa vaaditaan kokonaisen sisältöyksikön luomista. Esimerkkeinä voidaan mainita kokonaisen graafin piirtäminen annettujen lukujen perusteella; jonkin maan vientituotteiden luetteleminen; kuvaus siitä, miksi vesihöyry tiivistyy ikkunaan; geometrisen todistuksen kirjoittaminen; arkkitehtonisen piirroksen tuottaminen; tietokoneohjelman laatiminen tai kirjoitelman kirjoittaminen.

Esimerkki: Kuvaile jotakin luonnonilmiötä (esim. maanjäristystä, ukkosta, sateenkaarta), joka on aina kiinnostanut sinua ja josta haluaisit tietää enemmän. Mitä haluaisit tietää valisemastasi aiheesta ja miksi? (Sinulla on 30 minuuttia aikaa kirjoittaa teksti.)

6. *Esittäminen*: Tämä koetyyppi edellyttää henkilökohtaista esittämistä tai esiintymistä todellisessa tai simuloitussa tilanteessa, jossa arvioinnin kohteena on melko suuressa määrin esittämis/esitystapa, ei pelkästään sen lopputulos. Esimerkkeinä voidaan mainita mm. auton moottorin korjaaminen, jonkin instrumentin soittaminen, potilaan sairauden diagnosointi, näytetunnin opettaminen, tieteellisen kokeen tekeminen tai koesoiton esittäminen.

Snown mielestä monivalinta- ja tuottamiskoetyyppien välillä ei ole eroa mittauksen tason suhteen: molempia voidaan käyttää arvioimaan sekä syväymmärtämistä ja korkeammantasoista ajattelua että pinnallisten ja irrallisten faktojen ja sääntöjen hallintaa. Mielestäni Snow on pitkälti oikeassa, mutta monivalintahan kyllä mahdollistaa myös pelkän arvauksen: kaikki kokeen suorittajat eivät välttämättä käy läpi tehtävän edellyttämää ajatusprosessia. Snow mainitsee myös, että tuottamiskoetyypit ovat vähemmän reliaabeleita, vähemmän taloudellisia, kattavat mitattavan alueen heikommin ja ovat enemmän alttiita testiharhoille kuin ajallisesti yhtä pitkät monivalintakokeet. Tämä pitänee hyvin pitkälle paikkansa. On kuitenkin syytä huomata, että perusteellisella arviointisijakoulutuksella voidaan tuntuvasti edistää tuottamiskokeiden arvioinnin yhdenmukaisuutta.

## 10. PORTFOLIO OPPIMISEN JA OPPIJAKESKEISEN ARVIOINNIN TUKENA

Oppijakeskeisen oppimisen arviointiin on pyritty etsimään uusia, entistä elämänläheisempiä ja oppimisympäristöön yhdistyviä menetelmiä. Yhteistä näille ovat ainakin seuraavat piirteet (Valencia 1990; Calfee & Perfumo 1993; García & Pearson 1994; Hiebert et al. 1994; Linnakylä 1994):

- ne ovat mahdollisimman lähellä oppilaan opiskelutilannetta ja todellisia, aitoja opiskelutehtäviä
- ne eivät ole pelkästään koetettavia, vaan osa opiskelua ja oppimista
- ne kohdistuvat sekä oppimisprosessiin että sen tuloksiin

- ne painottavat erityisesti tiedon soveltamista, kriittistä ajattelua ja ongelmanratkaisukykyä
- ne painottavat pikemminkin oppilaan omaa tuottamista kuin opettajan tai jonkun ulkopuolisen kokeenlaatijan kysymyksiin vastaamista
- ne sisältävät pikemminkin laajoja, monien oppiaineiden osaamista integroivia ja usein yhteistoiminnallisia projektitehtäviä kuin yksittäisiä koeosioita
- ne kohdistuvat paitsi osaamiseen, myös oppimisstrategioihin ja niiden säätelyyn sekä opiskeluasenteiden kehittymiseen
- ne pyrkivät osoittamaan pikemminkin oppimisen laadun ja vahvuudet kuin määrän ja puutteet.

Portfolioita käyttäneiden opettajien mukaan (Tieney ym. 1991) omien töiden arviointi- ja valikointiprosessi auttaa oppilaita:

- valitsemaan jatkossa kiinnostavia ja merkityksellisiä oppimistehtäviä
- tunnistamaan omia edellytyksiään, vahvuuksiaan ja tarpeitaan
- asettamaan itselleen uusia tavoitteita
- havainnoimaan ja arvioimaan omaa edistymistään
- huomaamaan oppimisen monipuolisuuden
- tarkkailemaan omien tuotosten omaperäisyyttä ja laatua
- ymmärtämään yrittämisen merkityksen
- tunnistamaan erilaisia opiskelutapoja ja -strategioita
- tuntemaan vastuuta ja ylpeyttä omista töistään ja edistymisestään.

Portfolio-toiminnassa keskeistä on mm. Linnakylän (1994) mukaan:

- antaa oppilaalle onnistumisen mahdollisuuksia
- asettaa oppilaan kanssa yhdessä realistisia oppimistavoitteita ja suunnitella oppilaan edellytyksiin ja omiin tavoitteisiin liittyviä oppimistehtäviä
- tukea oppilaan yksilöllisyyttä ja hyväksyä hänen yksilölliset taipumuksensa ja kiinnostuksensa
- auttaa oppilasta arvioimaan itse omaa edistymistään
- kannustaa oppilasta osoittamalla hänen osaamisensa parhaita puolia
- kiittää ja antaa tunnustusta myös hyvästä yrityksestä ja rohkaista perusteluun itsekiitokseen
- unohtaa kyyninen kriittisyys; jos kriittisyys on tarpeen, kohdistaa se joihinkin asioihin tai tietynlaiseen käyttäytymiseen, ei oppilaan persoonallisuuteen
- antaa vastuuta ja päätösvaltaa sekä rohkaista vastuunottoon.

Portfolioarvioinnista on jo varsin paljon kokemusta meillä ja muualla. Sen lähtökohtia ja siitä saatuja kokemuksia on raportoitu mm. kahdessa Kasvatus-tieteiden tutkimuslaitoksen (nytemmin Koulutuksen tutkimuslaitoksen) julkaisussa, jotka on mainittu lähdeluettelossa. Kieltenopettajia kiinnostanevat erityisesti Pirjo Pollarin artikkelit, koska ne liittyvät nimenomaan portfolion käyttöön kieltenopetuksessa. Lisätietoa on saatavissa myös hänen lisensiaattitutkimuksestaan, joka on viimeistelyvaiheessa.



## 11. AUTENTTISEN JA PERINTEELLISEN ARVIOINNIN VERTAILUA

Kun testaus ja arviointi ovat entisestään lisääntyneet opetuksessa ja koulutuksessa, myös niitä arvosteleva kritiikki on voimistunut. Erityisesti on kiinnitetty huomiota niiden mahdollisiin kielteisiin vaikutuksiin. Toisaalta on nähty, että "vaihtoehtoisilla" arviointimuodoilla voidaan tukea oppimista ja saada selville jokaisen oppijan vahvoja puolia ja näin edesauttaa mm. positiivisen minäkuvan kehittymistä. Validiuden käsitteessä ovatkin arvioinnin seuraukset (consequential validity) saaneet huomattavan aseman (Messick 1989, 1994).

Suoritusperustaisen arvioinnin painottuminen johtuu useista syistä. Tärkeiden testien ja tutkintojen takaistusvaikutukseen (WYTIWYG, what you test is what you get) uskovat korostavat, että kyseisen vaikutuksen vuoksi arvioinnin tulee olla sellaista, että siihen kannattaa "prepatakin".

Arvioinnin tulee sisältää tehtäviä, jotka ovat sinänsä arvokkaita opetus- ja opiskelutoimintoja. Arvioinnin edellytetään jopa toimivan mallina opetukselle ja opetussuunnitelmien kehittämiseksi ja virikkeenä opettajien täydennyskoulutukselle.

Wiggins (1990) konkretisoi autenttisen arvioinnin etuja "perinteelliseen" (tässä tarkoittaa amerikkalaista voimakkaasti monivalinnan hallitsemää arviointia) arviointiin verrattuna seuraavasti:

1. Autenttinen arviointi edellyttää, että oppijat käyttävät tehokkaasti hankimaansa tietoa. Perinteellinen arviointi paljastaa lähinnä sen, miten hyvin oppija tunnistaa, muistaa tai osaa yhdistää tietoja, jotka on opittu kontekstista irrallaan. Tämä voi olla yhtä ongelmallista kuin jos päätelisi autolla-ajotaitoa tai opetustaitoa pelkästään kirjallisten kokeiden perusteella.
2. Autenttinen arviointi esittää oppijalle tehtäviä, jotka heijastavat parhaassa opetuksessa esiintyviä painotuksia ja haasteita: tutkielman tekeminen; esitelmän kirjoittaminen, muokkaaminen ja siitä keskusteleminen; äskettäisten poliittisten tapahtumien suullinen analyysi; väittelytilaisuuden valmistelu yhdessä muiden kanssa jne. Perinteiset kokeet ovat yleensä kynä- ja paperikokeita, joissa tehtäviin on yksi oikea vastaus.
3. Autenttinen arviointi kiinnittää huomiota siihen, osaako oppija rakentaa perinpohjaisia, perusteltuja ja viimeisteltyjä vastauksia ja tuotoksia. Perinteiset kokeet edellyttävät yleensä vain, että oppija valitsee tai kirjoittaa oikean vastauksen. On harvoin tilaisuutta suunnitella, laatia ja muokata vastauksia.
4. Autenttinen arviointi toteuttaa validiuden ja reliabiliuden vaatimukset korostamalla ja standardisoimalla asianmukaiset kriteerit, joiden avulla arvioidaan oppijoiden erilaisia vapaita tuotoksia. Perinteinen arviointi vakioi objektiivisia "osioita" ja niiden oikeat vastaukset.
5. "Koevalidiuden" tulisi riippua osin siitä, missä määrin koe vastaa todellisessa elämässä esille tulevia suoritusvaatimuksia. Useimmissa monivalintakokeissa validius perustuu siihen, että tehtävät liittyvät opetussuunnitelmien sisältöön tai että ne korreloivat muiden kokeiden kanssa.

6. Autenttiset tehtävät sisältävät avoimia ja "epämääräisiä" (ill-structured) haasteita ja rooleja, jotka auttavat oppijaa valmistautumaan aikuisten ja työelämän monitahoisiin ja avoimiin tilanteisiin ja vaatimuksiin. Perinteiset kokeet ovat usein enemmän drillien kaltaisia ja arvioivat staattisia ja usein keinotekoisien yksinkertaistettuja toimintoja.

Seuraavaan taulukkomuotoiseen yhteenvedoon olen koonnut autenttisesta (vaihtoehtoisesta, suorituspohjaisesta) arvioinnista ja lähinnä monivalintaan pohjautuvasta arvioinnista esitettyjä näkemyksiä: mihin kussakin tapauksessa pyritään ja toisaalta, mitä etuja ja haittoja erilaisilla piirteillä voi olla.

| Piirre  | Tavoite/pyrkimys  | Mahdollisia vahvuuksia  | Mahdollisia vastaväitteitä   |
|---|---|---|--|
| Autenttinen (vaihtoehtoinen, suorituspohjainen) arviointi   | Arvioinnin tulee heijastaa uutta oppimisenäkemyksiä ja oppimisen luonnollisia käytäntöjä ja -yhteyksiä. | <ul style="list-style-type: none"> <li>- Arvioidaan tärkeitä ja arvokkaina pidettäviä tavoitteita.</li> <li>- Arviointi vastaa opetussuunnitelmaa ja jopa tukee sen toteuttamista.</li> <li>- Arviointi koetaan mielekkääksi ja motivoivaksi.</li> <li>- Arviointi heijastaa yksilöiden vahvoja puolia ja saattaa vahvistaa minäkuvaa.</li> </ul> | <ul style="list-style-type: none"> <li>- Autenttisuus ei ole yksiselitteinen käsite, jolla ei siten myöskään ole yksiselitteisiä kriteereitä.</li> <li>- Arvioinnin autenttisuuden väitetyillä eduilla ei tässä vaiheessa ole vahvaa näyttöä.</li> </ul>                 |
| Perinteellinen (monivalintapohjaisiin standardikokeisiin pohjautuva) arviointi                          | Arvioinnin tulee ennen muuta olla luotettava ja yhteismitallista -käyttökonteksti on toisarvoista.      | <ul style="list-style-type: none"> <li>- Subjektiiivisuus varsin hyvin hallinnassa.</li> <li>- Reliabeelisuus yleensä hyvä.</li> <li>- Mitattava/arvioitava alue saadaan katetyksi hyvin.</li> <li>- Arviointi on kustannuksiin nähden tehokasta.</li> </ul>  | <ul style="list-style-type: none"> <li>- Validius voi olla ongelma.</li> <li>- Takaisustusvaikutus opetukseen voi olla epätoivoittavaa.</li> <li>- Arvioinnissa voi painottua liikaa muistitieto, ja kokonaisuuDET saattavat jäädä vaille riittävää huomiota.</li> </ul> |
| Arvioinnin (mittauksen) suoruus   | Arvioinnin tulee heijastaa sen kohdetta mahdollisimman tarkasti.  | <ul style="list-style-type: none"> <li>- Arvioinnin ilmeisvalidisuus on hyvä.</li> <li>- Tulosten tulkinta on selkeää.</li> </ul>   | <ul style="list-style-type: none"> <li>- Kaikki arviointi on epäsuoraa ja edellyttää aina tulkintaa.</li> </ul>  |
| Arvioinnin rakentaminen tehtävien pohjalle (task-driven)  |   | <ul style="list-style-type: none"> <li>- Arviointi on uskottavaa, koska autenttisissa tehtävissä voi, ja on pakkokin, käyttää hyväksi kaikkia sen tekemisen kannalta tärkeitä tietoja ja taitoja.</li> </ul>  | <ul style="list-style-type: none"> <li>- Erilaisia tehtäviä ei ole helppo määritellä riittävän yksiselitteisesti.</li> <li>- Ei tiedetä paljonkaan siitä, miten yleistämiskelpoista tietoa saadaan tehtäväpohjaisesta arvioinnista.</li> </ul>                           |
| Arvioinnin rakentaminen suorituksen pohjalla olevien tietojen ja taitojen perustalle (construct-driven) |   | <ul style="list-style-type: none"> <li>- Arviointi on yleistämiskelpoista, koska tiedetään, mihin tehtävät perustuvat.</li> </ul>   | <ul style="list-style-type: none"> <li>- Arvioinnin tulkinta ei välttämättä ole yhtä yksiselitteistä kuin puhtaasti tehtäväpohjaisessa arvioinnissa.</li> </ul>  |
| Arvioinnin rakentaminen hyvin avoimelle tilanteelle   |   | <ul style="list-style-type: none"> <li>- Arviointi vastaa hyvin "tavallista elämää", jossa toimintatilanteet ovat usein avoimia, jolloin tulee itse määritellä, mistä on kysymys.</li> </ul>  | <ul style="list-style-type: none"> <li>- Avoimuus voi hämmäntää ja heikentää joidenkin suorituksia.</li> <li>- Avoimuus on suhteellista - myös osittain jäsenneetyt tilanteet voivat olla "elämänläheisiä".</li> </ul>   |

| Piirre   | Tavoite/pyrkimys | Mahdollisia vahvuuksia   | Mahdollisia vastaväitteitä  |
|--|------------------|--|---|
| Arvioinnin rakentaminen pitkäalle jäsenneyille (kontrolloidulle) tilanteelle |                  | <ul style="list-style-type: none"> <li>- Arviointitilanne on hyvin hallinnassa: saadaan diagnostista tietoa halutulla tarkkuudella (vrt. perinteellinen arviointi yllä).</li> <li>- Sidottu arviointitilanne luo turvallisuuden tunnetta.</li> </ul> | <ul style="list-style-type: none"> <li>- Arviointi on keinotekoisia eikä anna kyllin edustavaa kuvaa osaamisesta.</li> <li>- Tilanteen sidonaisuus voidaan kokea liian rajoittavaksi, mikä voi alentaa motivaatiota.</li> </ul> |

Taulukko havainnollistaa sitä arviointikirjallisuudessa todettua tosiasiaa, että ei ole olemassa mitään yksiselitteisesti parasta tai heikkointa arviointimenetelmää. Kullakin menetelmällä on vahvuutensa ja heikkoutensa, hyvät puolensa ja rajoituksensa.

## 12. KYSYMYSTEN TEHTÄVÄ ARVIOINNISSA JA KYSYMYSTAKSONOMIA

Kysymyksillä on keskeinen asema ihmisten välisessä kanssakäymisessä. Näin on myös opetuksessa. Opetuksessa kysymisellä on tavallisesta viestinnästä jossakin määrin poikkeava tehtävä. Kun tavallisessa viestinnässä kysytään, jotta saataisiin puuttuvaa tietoa, opetuksessa kysymisellä paljolti pyritään saamaan selville, mitä kysyttävä tietää ja osaa. Kysyminen on keskeinen väline ymmärtämisen, tiedon ja osaamisen selville saamiseen.

Koetilanne voi tuskin koskaan vastata täysin autenttista viestintää. Koetilanteessa tulee lyhyessä ajassa saada riittävän yleistämiskelpoinen kuva henkilön kielitaidosta, joten koetehtävien tulee olla hyvä otos kaikista kyseeseen tulevista mitattavista asioista. Jokaisen tehtävän tulisi antaa mahdollisimman tehokkaasti tietoa, josta voidaan tehdä mahdollisimman spesifejä tulkintoja. Tulkintojen tulisi myös olla luotettavia, ja samansuuntaisia arvioitsijoista riippumatta, joten on järkevää testata vain sellaista, jonka arviointi on riittävän yhdenmukaista. ("high stakes"-tilanne).

Vieraiden kielten kokeissa kysymällä pyritään saamaan selville, mitä oppilas on ymmärtänyt lukemastaan tai kuulemastaan. Ymmärtäminenhen sinänsä on sisäinen kognitiivinen tapahtuma, josta ulkopuolinen saa otetta vain, jos henkilö havaittavalla toiminnalla tuo ymmärtämisen esille. Tosin ymmärtämistä voi luodata myös kehotuksen avulla, pyytämällä henkilöä esimerkiksi kertomaan, mitkä olivat jonkin tekstin pääajatuksukset. Omin sanoin kertominen, tiivistäminen laajassa mielessä, pohjautuu tavallisesti kehotukseen (tai käsyykn) tehdä niin, ei sisällöstä esitettyihin kysymyksiin.

Myös vieraan kielen tuottamisessa kysymyksillä on tärkeä asema. Haastatteluun perustuvassa puhumisessa kysymyksillä esittämät kysymykset ohjaavat ratkaisevasti tuotettavaa puhetta. Jotta koetilanne olisi kaikille kokeeseen osallistuneille riittävän yhdenvertainen, haastattelija joutuu huolellisesti miettimään esittämiensä kysymysten muotoilua. Haastattelija voi ohjata haastateltavaa myös kehotuksin kertoa lisää jostakin asiasta. Ohjatussa kielistudiokokeessa henkilö voi kuulla kysymyksiä nauhalta ja vastata oman tulkintansa perusteella tai tulkiten annettuja ohjeita tai esim. visuaalista materiaalia. Vapaammassa tilanteessa henkilöä voidaan ohjata sanomaan sitä, mitä hän hänelle esitettyissä erilaisissa tilanteissa katsoisi voivansa sanoa vieraalla kielellä.

Kirjallinen tuottaminen puolestaan voi aidossa viestintätilanteessa perustua henkilön omaehtoiseen tarpeeseen viestiä jollekulle kirjoituksen avulla. Vastaavasti henkilö voi reagoida kirjallisesti saamaansa viestiin. Kirjoittamista käytetään myös kielitaidon osoittamiseen, esim. vastaamalla puheesta tai tekstistä esitettyihin kysymyksiin. Kirjoittaminen perustuu koetilanteessa usein tehtävänantoon, jossa kokeeseen osallistujalle annetaan tehtäväksi kirjoittaa tietyllä tavalla. Tehtävänanto voi olla myös kysymys, johon vastaaminen voi edellyttää pääasiassa kertovaa, kuvailevaa, argumentoivaa, pohdiskelevaa tai ohjailevaa tekstiä.

Kysymysten katsotaan yleensä muodostavan hierarkkisen taksonomian, jossa kysymykset etenevät siten, että alimmalla tasolla kysymykset perustuvat suoraan siihen, mitä esimerkiksi tekstissä on sanottu, tasolle, jossa vastataan tekstiä sivuvaaihiin kysymyksiin. Kysymykset voidaan luokitella seuraavasti sen perusteella mitä vastaukselta edellytetään:

1. Toistaminen: Kysymykset edellyttävät pelkästään tekstissä esitetyn toistamista.
2. Omin sanoin esittäminen: Kysymykset edellyttävät, että oppilas kertoo omin sanoin mitä tekstissä sanottiin (mm. tiivistelmä).
3. Päättelyminen: Kysymykset edellyttävät, että oppilas tekee tekstin pohjalta hyvin perusteltuja johtopäätöksiä.
4. Seuraamusten päättelyminen: Kysymykset edellyttävät tilanteen sallimia johtopäätöksiä.
5. Kokemuksellinen pohdinta: Kysymykset edellyttävät, että oppilas kertoo omia kokemuksiaan tekstissä esitetystä.
6. Arvioiva: Kysymykset edellyttävät oppilaan arvioivan, mitä tekstissä sanotaan ja miten se on kirjoitettu. Voidaan kysyä kirjoittajan lähtöoletuksia, tekstissä esitettyjen näkemysten, perustelujen ja johtopäätösten pitävyyttä.

Kokeenlaatijan tulee pohtia, millaista tietoa tai taitoa erilaiset kysymykset mittaavat ja varmistaa, että kysymykset ovat riittävän monipuolisia. Tutkimus ja käytännön kokemus viittaavat siihen, että "perinteinen" käsitys ymmärtämisestä korostaa liikaa sen alempia tasoja.

### 13. KIELITIEDON MERKITYS KIELITAIIDOLLE JA KIELITIEDON ARVIOINTI

Widdowsoniin (1990) yhtyen kielioppia voidaan pitää keskeisenä myös kommunikatiivisesti suuntautuneessa kielenopetuksessa. Tähän tulokseen on tullut myös Jaakkola (1997) tuoreessa väitöskirjassaan. Tämän näkemyksen mukaan kieliopilla on tärkeä välitystehtävä sanojen ja kontekstin välillä. Jos kontekstin perusteella voi päätellä pitkälti mitä sanotaan, kielioppiin (eli hyvinmuodostettuihin kokonaisiin lauseisiin) perustuva viestintä on tarpeetonta ja joskus jopa vahingollista tai vaarallista. Leikkaussalissa ei ole tarpeen kohteliaasti pyytää erilaisia työkaluja. Varoitettaessa äkillisestä vaarasta ei ole aikaa kuvailla seikkaperäisesti, mikä vaara uhkaa. Varoitushuuto riittää huomion herättämiseksi. Mitä vähemmän tilanteen ja yksityisten sanojen varaan voi uskoa viestin peril-

lemenon, sitä tärkeämmäksi nousee merkityksen täsmentäminen kieliopin tarjoamin keinoin. Mitä vähemmän voidaan ennakoida, mitä sanotaan, sitä suu-remmpi on kieliopillisten keinojen panos viestin ymmärrettävyyden varmistamiseksi.

Edellisen näkemyksen mukaan kieliopilla ei ole ylivaltaista asemaa sanastoon nähden. Viestinnän tehokkuuden kannalta laaja ja tarkka sanavarasto on välttämätöntä. Osa kielen ilmiöistä sijoittuu sanaston ja kieliopin väliin. Kielessä on hyvin suuri määrä vakioilmauksia, jotka koostuvat joukosta sanoja. Näitä tulee ymmärtää ja itse osata käyttää viestinnässä. Joskus ne ovat olennaisen tärkeitä viestinnälle, joskus niiden puutteellinen hallinta vain paljastaa kielenkäyttäjän vielä olevan jokseenkin etäällä idiomaattisesta kielenhallinnasta. Osa tämän tason kielenhallinnasta on tietoa sanojen rektiosta, ts. siitä, millaisia muotoilua sana edellyttää siihen normaalisti liittyviltä muilta sanoilta. Kieliopillisilla ilmiöillä puolestaan ilmaistaan sellaisia yhteyksiä, joita olisi hyvin epätaloudellista ilmaista sanastollisilla keinoilla. Kielioppi on tämän näkemyksen mukaan yksi keskeinen keino viestinnällisen tehokkuuden varmistamiseksi. Se ei ole itsetarkoitus kuten ei kieliopillisuuskaan. Tehokas viestintä edellyttää, että kieliopin keinoja osataan käyttää riittävän tarkasti merkityksen välittämisessä.

Kielioppia ei ole järkevää asettaa sanaston yläpuolelle niin, että sanoja käytetään jossain määrin satunnaisesti kieliopin havainnollistamiseksi. Päinvastoin kielioppi on järkevää alistaa sanaston palvelukseen. Kieliopin opettamisessa peruslähtökohtana on useasti järkevää pitää merkitysten välittämistä ja sanastoa. Tällä on liittymäkohta myös ns. elaborointimenetelmään, jota erityisesti Irene Kristiansen on kehittänyt. Oppilaille voidaan kuvata lyhyesti, millainen merkityskokonaisuus tulisi välittää ja antaa esityksessä tarvittavaa keskeistä sanastoa. Toisessa vaihtoehdossa oppilaat saavat vain joukon sanoja, joiden perusteella heidän tulee kehittää itse yhtenäinen esitys. Molemmissa tapauksissa kyseessä on ongelmanratkaisutilanne kahdessa mielessä: on tuotettava yhtenäinen diskurssi ja esityksessä on pyrittävä käyttämään sanoja niin, että lauseiden sisällä ja niiden välillä osataan käyttää kieliopillisia keinoja kyllin (tai mahdollisimman) hyvin.

Kielitiedon testaamisessa on edellä esitetyn perusteella suositeltavaa, että tunnistamisessa käytetään tekstikontekstia. Monivalinta-aukkokoe vaikuttaisi sopivan hyvin tähän tarkoitukseen. Hyvin laadittu koe edellyttää koko tekstin merkityksen ymmärtämistä ja lähikontekstin vaatimusten huomioonottamista. Tuottamisessa puolestaan voidaan monivalintakohtat jättää aukollisiksi, ja oppilaiden tulee täyttää aukot joko täysin itse päättämällä tai annettujen vihjeiden perusteella. Edellä on kuvattu huomattavasti avoimempaa kielioppi-tiedon testaamista.

#### 14. SANASTON MERKITYS VIESTINNÄSSÄ JA SANASTON TESTAAMINEN

Sanasto on pitkään kiehtonut tämän artikkelin kirjoittajaa ja sanaston testaaminen oli myös väitöskirjani aihe (Takala 1984). Kouluaajan varhaiset kokemukset saivat minut vakuuttumaan laajan sanavaraston tärkeydestä ja omakohtai-

nen tutkimustoiminta on entisestään vahvistanut tätä käsitystä ja tuonut sille tieteellistä perustelua. Ilman riittävää kieliopillista hallintaa on vaikea viestiä tehokkaasti, kuten edellä on esitetty, mutta ilman sanastoa viestinnästä ei tule mitään. Minusta on suurta liiottelua väittää, että viestinnässä non-verbaalinen kanava olisi ratkaiseva. Tätä ei voine ymmärtää muuten kuin siten, että viitataan pääasiassa viestintään sosiaalisten suhteiden hoitamisessa. Kirjallinen viestintä jää kokonaan vaille huomiota. Samoin jäänee kovin vähäiselle huomiolle kielen käyttö asioiden välittämiseen. Miten tehokkaasti non-verbaalisin keinoin pääasiassa luottaen olisi voinut suullisesti esittää sitä kaikkea, mitä edellä on sanottu arvioinnista? Arvelen, että varsin laaja sanasto ja kyky kehittää ajatusrakennelmia johdonmukaisesti ovat paljon tärkeämpiä tässä tapauksessa.

Sanan osaaminen on monitahoinen taito. Seuraavassa Nationin (1990) erittely siitä, mitä sanaston osaamisella voidaan tarkoittaa.

|                  |                |   |
|------------------|----------------|---|
| <i>Muoto</i>     |                |   |
| Puhuttu          | Reseptiivinen  | Miltä sana kuulostaa?   |
|                  | Produktiivinen | Miten sana äännetään?   |
| Kirjoitettu      | Reseptiivinen  | Miltä sana näyttää?   |
|                  | Produktiivinen | Miten sana kirjoitetaan?  |
| <i>Paikka</i>    |                |   |
| Kiellop. rakenne | Reseptiivinen  | Missä rakenteissa sana esiintyy?                                    |
|                  | Produktiivinen | Missä rakenteissa sana esiintyy?                                    |
| Kollokaatio      | Reseptiivinen  | Millaisia sanoja esiintyy tämän sanan edessä/jäljessä?              |
|                  | Produktiivinen | Millaisia sanoja/sanatyyppejä on käytettävä tämän sanan yhteydessä? |
| <i>Funktio</i>   |                |   |
| Yleisyys         | Reseptiivinen  | Kuinka yleinen sana on?   |
|                  | Produktiivinen | Kuinka usein sanaa tulee käyttää?                                   |
| Soveltuvuus      | Reseptiivinen  | Missä voi odottaa tätä sanaa?                                       |
|                  | Produktiivinen | Missä tätä sanaa voi käyttää?                                       |
| <i>Merkitys</i>  |                |   |
| Käsite           | Reseptiivinen  | Mitä sana tarkoittaa?   |
|                  | Produktiivinen | Mitä sanaa käytettävä ilmaisemaan tätä käsitettä?                   |
| Assosiaatiot     | Reseptiivinen  | Mitä muita sanoja tulee mieleen?                                    |
|                  | Produktiivinen | Mitä muita sanoja voidaan käyttää tämän sanan sijasta?              |

Jokaista sanan osaamisen piirrettä voidaan mitata erikseen ja kussakin tapauksessa tulee kysymykseen useampi kuin yksi mittaustapa. Lisätietoa sanaston opettamisesta voi saada mm. tekijän toimittamasta julkaisusta (Takala 1989). Sanastopäätelyä, jota myös mielellään voisi testata diagnostisesti, formatiivisesti ja summatiivisesti, käsittelee Vaurio lisensiaatintutkimuksessaan (1995) ja valmistuvassa olevassa väitöskirjassaan.

## 15. ARVIOINNIN KEHITTÄMISHAASTEITA

Koulussa suoritettavassa arvioinnissa on opetussuunnitelma keskeinen lähtökohta. Nykyisin opetuksen tavoitteena on, jossakin määrin vapaasti muotoillen, että

- oppilas osaa käyttää vierasta kieltä käytännölliseen viestintään ja selviytyy tavallisista viestintätehtävistä ja tilanteista, jotka liittyvät opiskeluun, vapaa-aikaan (ja työelämään), ja hän pystyy myös kehittämään viestintätaitojaan
- oppilaan opiskeleman kielen systemaattinen hallinta edistyy, ja hän pyrkii itse tietoisesti kehittämään viestintävalmiuksien lisäksi myös kielen systemaattista hallintaansa
- oppilaan tietoisuus ja ymmärrys opiskeltavan kielialueen maista, kansoista ja niiden elintavoista laajenee ja syvenee, ja hän oppii asennoitumaan avoimesti ja sympaattisesti kielialueen kulttuuria/kulttuureita kohtaan
- oppilaan kielenopiskelutaidot, itsearviointikyky sekä itsenäinen ja vastuullinen kielenopiskelu kehittyvät
- oppilas kokee opiskeltavan kielen opetuksen tarjoavan kokemuksia, jotka ovat kiinnostavia, tarjoavat mielihyvää, oppimisen kokemista ja tiedollista haastetta.

Katson, että edellä on sanottu kaikki oleellinen: tiedolliset, taidolliset, asenteelliset, motivaatioon ja itseohjautuvuuteen liittyvät sekä myös kokemukselliset tavoitteet. Uskoisin, että tavoitteita ei ole tarpeen pikaisesti muuttaa, vaan ne voivat säilyä pitkään kielenopetuksen yleisenä ohjenuorana. Ne antavat hyvän pohjan myös arvioinnin linjaamiselle.

Ei riitä, että arviointi heijastaa hyvin oppisisältöjä (sisällön edustavuus/ "validius"), kuten perinteisesti korostetaan. Arvioinnin tulee myös (a) heijastaa testattavan asian psykologista olemusta (käsitevalidius) eli kielenopetuksessa erityisesti kulttuurienvälistä viestintätaitoa ja (b) liittyä mielekkäällä tavalla viestintätaidon normaaleihin käyttöyhteyksiin (ekologinen validius). Lisäksi on viime aikoina alettu korostaa niitä seurauksia, joita arvioinnilla on eri yksilöille ("consequential validity"). Perinteisesti on korostettu arvioinnin yhdenmukaisuutta (reliabeliutta), joka onkin tärkeä vaatimus. Olen edellä jo esittänyt, että sitä ei ole kuitenkaan syytä korottaa ylivertaisen tärkeäksi vaatimukseksi. Täydellinen yhdenmukaisuus voidaan saavuttaa vain jokseenkin triviaaleissa seikoissa ja rajaamalla suoritukset valmiiden vaihtoehtojen valintaan. On huomattavasti helpompaa saavuttaa yksimielisyys oikeinkirjoituksen tasosta kuin esim. kirjoitelman ansioista, mutta tuskin tämän vuoksi on järkevää keskittyä (kirjoittamisen) arvioinnissa pääasiassa oikeinkirjoitustaitoon. On mielekäästä arvioida tärkeitä asioita ja kehittää niiden arvioinnin taitoa niin, että se on kohtuullisena pidettävän yhdenmukaista.

Kurssien suunnittelussa, opetuksessa ja opiskelussa on keskeistä, että kielen hallinta ja viestintätaidot ja -strategiat kehittyvät yllä kuvatulla tavalla. Viestintätaitojen ja -strategioiden kehittäminen ei aseta tarkkoja vaatimuksia käsiteltäville aihepiireille, vaan useat aihepiirit sopivat eri vaiheissa.

Viestinnässä on tavallisesti kyse merkityskokonaisuuden välittämisestä. Esimerkiksi tiedottavassa viestinnässä tämä voi olla yhden virkkeen mittainen viesti tai laaja suullinen tai kirjallinen selonteko. Koulussa tulee harjoitella systemaattisesti erilaisten tekstilajien (tiedottavan, kuvailevan, taivuttelevan, erittelevän ja ohjaavan viestin) ymmärtämisen ja tuottamisen kehittämistä, ja kyseisiä taitoja tulee myös arvioida.

Ymmärtämisen – erityisesti tekstin ymmärtämisen – progressiota on syytä pohdiskella. Ainakin A-kielessä tuntuisi perustellulta, että kun kielen perusta alkaa olla valmiina, voitaisiin siirtyä yhä enemmän ekstensiiviseen tekstin lukemiseen, jossa tekstiä on – mahdollisimman usein – autenttisen tekstin mukaisesti jopa useampi sivu. Oppilaiden tehtävänä on löytää tekstin pääajatus, jonka he osoittavat usein – kuten normaalissa viestinnässä – omakohtaisen suullisen/kirjallisen tiivistämisen muodossa, mutta myös vastaamalla pääasi- oita testaaviin avoimiin kysymyksiin ja haluttaessa myös monivalintatehtäviin. Monivalintakokeessa on tärkeätä, että tekstiä on niin runsaasti, että kullekin vaihtoehdolle on tekstissä "lihaa". Tuntuisi myös tärkeältä, että viestintä/tekstilajeihin perustuvan opettamisen lisäksi harjoiteltaisiin ja testattaisiin myös eri tarkoituksiin sopivien ymmärtämisstrategioiden kehittamisestä (ks. esim. Blom ym. 1988; Pitkänen-Huhta 1996; Vaurio 1995).

Puheviestinnän kehittämisen kannalta taas olisi tärkeätä pohtia viestintästrategioiden systemaattista opettamista ja testaamista. Puhetaidon opettamisessa ei riitä, että sen harjoittelua lisätään määrällisesti, vaan on lisättävä harjoittelua, jolla edistetään puhetaidon laadullista kehitystä suunnitelmallisesti. Kielistudiolla olisi edessään uusi tuleminen, jos sitä käytettäisiin tietoisesti ja systemaattisesti puheviestinnän harjoitteluun ja puhumistaidon testaamiseen. Ulko- maiset ja omassa maassamme suoritettut tutkimukset osoittavat, että kielistudiopohjainen puhumiskoe antaa hyvin samanlaisen tuloksen kuin yksilöllinen haastattelu. Kielistudiossa voidaan tehokkaasti reagoida annettuihin kielenkäyttölanteisiin; kertoa kuvista, itsestä ja tapahtumista; esittää oma mielipide jostakin asiasta; jättää puhelinviesti; lukea ääneen lehtiutuinen muille kuuli- joille ja jopa käydä simuloitu ohjattu keskustelu, jossa kuulee keskustelukumpanin repliikit ja vastaa joko annettujen vihjeiden perusteella tai itse vastauksen keksien. Kielistudiokokeella ja haastattelukokeella on kummallakin etunsa ja puutteensa ja molempia voidaan käyttää suullisen kielitaidon testaamiseen. Jos kyseessä on yhteismitallisuuteen pyrkivä testi, vaakakuppi kallistuu aika selvästi kielistudiokokeen puolelle: sehän on kaikille samanlainen, kun haastattelu puolestaan on aina yksilöllinen tapahtuma ja altis erilaisille vaihteluille. Suullisen kielitaidon testaamista käsitellään laajasti tekijän toimittamassa julkaisussa (Takala 1993), Yli-Rengon useissa julkaisuissa (1989, 1992) ja Salevan valmistumassa olevassa väitöskirjassa.

Viestintävalmiuksien kehittymiselle on sanaston kasvuun ja sanastoon liittyvien taitojen systemaattinen kehittäminen tärkeätä. Olisi mietittävä, miten sananmuodostuksen ja sanaston merkityksen päättelyyn liittyviä taitoja voidaan systemaattisesti kehittää. Olisi suotavaa, että kyseisiä taitoja myös arvioitaisiin sopivassa laajuudessa osana arviointia (ks. esim. Takala 1989; Vaurio 1995).



Esitän seuraavassa muutamia kielenopetuksen arviointia koskevia teesejä. Arvioinnin autenttisuutta tulee pyrkiä lisäämään. Tämä merkitsee sitä, että kuullun ymmärtämisessä käytettäisiin mahdollisimman usein autenttista puhetta, jota voitaisiin äänittää esimerkiksi radiolähetyksistä. Ymmärtämistä arvioitaisiin tällaisissa tilanteissa kysymällä vain pääajatuksia. Menettely totuttaisi korvaa kuuntelemaan ekstensiivistä puhetta, etsimään punaista lankaa ja olemaan hermostumatta vaikka kaikkea ei ymmärräkään. Kuuntelemista voisi ohjata myös niin, että ennakolta kehoitettaisiin selvittämään, mitä nauhalla sanotaan tietyistä asioista. Kuullun perusteella oppilaat voisivat esittää suullisen tai kirjallisen yhteenvedon kuulemastaan. Samaa menettelyä voitaisiin käyttää tekstin ymmärtämisen testaamisessa. Sekä luettavat tekstit että lukemisen käyttötavat heijastaisivat todellisessa elämässä esiintyvää tekstien ja niiden lukemisen laajaa kirjoa.

Tulee kehittää kulttuurien välisen viestintätaidon arviointimenetelmiä. Tähän tarvitaan ajankohtaista tietoa kohdemaan (kohdemaiden) nykyisistä viestintänormeista. Anekdoottimaiseen tietoon perustuvat yleistyksyet ja stereotyyppiset mielikuvat tulisi jättää omaan arvoonsa. Luultavasti viestinnässä vallitsevat varsin pitkälle samanlaiset kohteliaisuusnormit ja aina voi jo etukäteen pyytää anteeksi, jos tulee tahtomatta loukanneeksi joitakin normeja. Samoin voinee tiedustella hienovaraisesti, mikä käytäntö vallitsee kussakin tilanteessa.

Tulee kehittää suullisen kielitaidon arviointimenetelmiä. Erityisen suuri haaste on kehittää kielistudiopohjaista testausmenettelyä, koska se tutkimusten mukaan tuottaa luotettavaa tietoa suullisesta kielitaidosta ja sen toimeenpano on käytännön kannalta huomattavasti helpompaa kuin yksilöllisen haastattelun.

Puheen ja kirjoituksen arvioinnissa tulee pyrkiä noudattamaan nykykielen normeja. Puheen tutkimus on osoittanut, että puhetta ei voi läheskään aina luontevasti kuvata perinteisen kielioopin sisältämien kategorioiden avulla. Siksi puheen arvioinnissa tulee noudattaa normaalia puheen tuottamisen mallia. Myös kirjallisen tuotoksen arvioinnissa tulee noudattaa hyväksyttävän nykykielen normeja eikä vanhojen kielioppien joissakin kohdin vanhentuneita sääntöjä.

Seuraavassa esitän puolestaan arviointia yleensä koskevia teesejä. Ensimmäkin opetuksen ja arvioinnin tulisi lähestyä toisiaan niin, että molemmat rakentuvat samalle näkemykselle oppimisesta ja opittavista asioista. Erityisesti tulee varoa sitä, etteivät kokeet kohtuuttomasti kapeuta luokassa toteutettavaa opetussuunnitelmaa. Tämä vaatimus kohdistuu erityisen voimakkaasti ylioppilastutkintoon ja merkitsee jatkuvaa haastetta sen kehittämiseen.

Kaikkien osapuolten tulee päästä osalliseksi arvioinnista. Erityisesti tulee kehittää toistaiseksi liian vähälle jäänyttä oppilaiden itsearviointia. Tämä on tärkeätä sekä kasvatuksellisista että oppimisen tehostamissyistä.

Koska toiminta koulun ulkopuolella paljolti perustuu yhteistoimintaan, on mietittävä, missä määrin ja muodossa arvioinninkin tulisi heijastaa oppilaiden keskinäiseen ja opettajan ja oppilaiden väliseen vuorovaikutukseen (esim. palautteeseen) rakentuvia suorituksia. Miten mm. suullisen kielitaidon arvioinnissa voitaisiin yksilötestauksen ohella käyttää erilaista ryhmätestausta?

Tavoitteiden väljentyessä ja valinnaisuuden lisääntyessä arviointimenetelmiä tulee kehittää niin, että ne mahdollistavat entistä yksilöllisemmän arvioinnin. Portfolioarviointi tarjoaa yhden lupaavan vaihtoehdon.

Edellä on otettu myönteinen kanta autenttisuuden lisäämiseen kielenopetuksessa ja viestintätaidon arvioinnissa. Autenttinen opettaminen ja oppiminen lähtee pitkälti kysymyksistä, joihin etsitään vastauksia. Se korostaa ymmärtämisen merkitystä, kun tehdään haasteellista, mielekästä, yhteistyöhön perustuvaa ja pitkäjänteistä oppimistyötä, jossa suunnittelu, toiminta, arviointi ja pohdiskelu muodostavat jatkuvasti toistuvan ketjun. Kun arviointi myös tehdään tässä hengessä, se on olennainen osa autenttista opetusta ja oppimista.

### *Lähteet*

- Alderson, C. J. 1981. Report of the discussion on communicative language testing. Teoksessa Alderson, C. J. & Hughes, A. Issues in Language Testing. ELT Documents 111. London. The British Council, 55–65.
- Alderson, J. C., Clapham, C. & Wall, D. 1995. Language test construction and evaluation. Cambridge: Cambridge University Press.
- Bachman, L. F. 1990. Fundamental considerations in language testing. Oxford: Oxford University Press.
- Bachman, L. F. & Palmer, A. S. 1996. Language testing in practice. Oxford: Oxford University Press.
- Bennett, R. E. 1993. On the meanings of constructed response. Teoksessa Bennett, R. E. & Ward, W.C. (eds.) Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment. Hillsdale, N.J.: Erlbaum, 1–27.
- Blom, L., Linnakylä, P. & Takala, S. 1988. Tekstien avaaminen luokassa. Jyväskylä: Kasvatustieteiden tutkimuslaitos. Teoriaa ja käytäntöä 10.
- Calfee, R. C. & Perfumo, P. 1993. Student portfolios: opportunities for a revolution in assessment. Journal of Reading, 36, 7, 532–537.
- Darling-Hammond, L. 1993. Reframing the school reform agenda: Developing capacity for school transformation. Phi Delta Kappan, Vol. 74, No. 10, 752–761.
- García, G.A. & Pearson, P. D. 1994. Assessment and diversity. Teoksessa Darling-Hammond, L. (ed.) Review of research in education. Vol. 20, Washington, D.C.: American Educational Research Association, 337–391.
- Glaser, R. & Silver, E. 1994. Assessment, testing and instruction: Retrospect and prospect. Teoksessa Darling-Hammond, L. (ed.) Review of research in education. Vol. 20, Washington, D.C.: American Educational Research Association, 393–419.
- Hiebert, E. A., Valencia, S. W. & Afferbach, P. P. 1994. Definitions and perspectives. Teoksessa Valencia, S. H., Hiebert, E. H. & Afferbach, P. P. (eds.) Authentic reading assessment. Practices and possibilities. Newark: IRA, 6–21.
- Kohonen, V. 1994. Salkkuarviointin mahdollisuuksia opettajan ammatillisen kehittymisen tukena. Teoksessa Linnakylä, P., Pollari, P. & Takala, S. (toim.) Portfolio arvioinnin ja oppimisen tukena. Jyväskylä: Kasvatustieteiden tutkimuslaitos, 33–50.
- Leimu, K. 1993. Johdatusta aiheeseen: Arvioinnin vanhat ulottuvuudet. Esitelmä koulutustutkimuskongressissa, Jyväskylän yliopisto, 21.–22.10.1993.
- Havola, L., Kärkkäinen, K. & Takala, S. 1994. Tiivistäminen tekstin ymmärtämisen tukena. Helsinki: Otava.

- Huhta, A., Määttä, A., Takala, S. & Tarnanen, M. (toim.) 1997. Yleiset kielitutkinnot ja kieltenopetus. Opetushallitus.
- Huhta, A., Sajavaara, K. & Takala, S. (toim.) 1993. Language testing: New openings. Jyväskylä: Kasvatustieteiden tutkimuslaitos.
- Jaakkola, H. 1997. Kielitieto kielitaitoon pyrittäessä. Vieraiden kielten opettajien käsityksiä kielioopin oppimisesta ja opettamisesta. Väitöskirja. Jyväskylän yliopisto: Kasvatustieteen laitos (painossa).
- Linnakylä, P. 1994a. Mikä ihmeen portfolio? Arvioinnin ja opetuksen liitto. Teoksessa Linnakylä, P., Pollari, P. & Takala, S. (toim.) Portfolio arvioinnin ja oppimisen tukena. Jyväskylä: Kasvatustieteiden tutkimuslaitos, 9–31.
- Linnakylä, P., Pollari, P. & Takala, S. (toim.) 1994. Portfolio arvioinnin ja oppimisen tukena. Jyväskylä: Kasvatustieteiden tutkimuslaitos.
- McNamara, T. 1996. Measuring second language performance. London: Longman.
- Messick, S. (1989) Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher* 18 (2), 5–11.
- Messick, S. 1994. The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23, 2, 13–23.
- Nation, I. S. P. 1990. Teaching and learning vocabulary. New York: Newbury House.
- Pitkänen-Huhta, A. 1996. Summary task: A comparison of L2 and L1 performance. *Lisensiaattitutkimus*. Jyväskylän yliopisto: Englannin kielen laitos.
- Pohjala, K. (toim.) 1994. Näkökohtia kielienopetukseen. Opetushallitus.
- Pollari, P., Kankaanranta, M. & Linnakylä, P. (toim.) 1996. Portfolion monet mahdollisuudet. Jyväskylä: Kasvatustieteiden tutkimuslaitos.
- Scriven, M. S. 1980. The logic of evaluation. Inverness, CA: Edgepress.
- Scriven, M. S. 1991. Evaluation thesaurus (4. painos). Newbury Park, CA: Sage.
- Shadish, W. R. Jr., Cook, T. D. & Leviton, L. C. 1991. Foundations of program evaluation: Theories of practice. Newbury Park, CA: Sage.
- Snow, R. E. 1993. Construct validity and constructed tests. Teoksessa Bennett, R. E. & Ward, W.C. (eds.) Construction versus choice in cognitive measurement: Issues in constructed response, performance testing, and portfolio assessment. Hillsdale, N.J. Erlbaum, 45–59.
- Takala, S. 1973. Kielitaidon mittaaminen. POPS-Opas 3B. Kouluhallitus.
- Takala, S. 1975. Oppimistulosten arvioiminen. Julkaisussa Nuotio, I. (toim.) *Kielenopettajien käsikirja*. Helsinki: WSOY, 172–181.
- Takala, S. 1980a. Kielitaidon mittaamisesta. Julkaisussa Sajavaara, K. (toim.) *Soveltava kielitiede*. Helsinki: Gaudeamus, 181–201.
- Takala, S. 1980b. Kriteerimittamisen käsitteestä ja käytännön sovelluksista. Jyväskylän yliopisto. Kasvatustieteiden tutkimuslaitoksen selosteita ja tiedotteita No. 146.
- Takala, S. 1984. Evaluation of students' knowledge of English vocabulary in the Finnish comprehensive school. Kasvatustieteiden tutkimuslaitoksen julkaisuja 350.
- Takala, S. (toim.) 1987. Katsauksia kirjoittamiseen. Kasvatustieteiden tutkimuslaitos. Teoriaa ja käytäntöä 5.
- Takala, S. (toim.) 1989. Sanaston opettaminen ja oppiminen. Kasvatustieteiden tutkimuslaitos. Teoriaa ja käytäntöä 44.
- Takala, S. 1992. Virikkeitä uutta kokeilevaan koulutyöhön. Jyväskylän yliopisto: Kasvatustieteiden tutkimuslaitos.

- Takala, S. (toim.) 1993. Suullinen kielitaito ja sen arviointi. Kasvatustieteiden tutkimuslaitos. Teoriaa ja käytäntöä 77.
- Takala, S. 1995. Arviointikulttuurin kehittäminen. Julkaisussa Salmio, K. & Lindroos-Himberg, K. (toim.) Askelia yleissivistävän koulutuksen arviointiin. Opetushallitus. Arviointi 2, 6–17.
- Takala, S. (toim.) 1995. Arviointi ja koulutuksen laadun kehittäminen. Jyväskylä: Kasvatustieteiden tutkimuslaitos.
- Tierney, R., Carter, M. A. & Desai, L. E. 1991. Portfolio assessment in the reading-writing classroom. Norwood, MA: Christopher-Cordon.
- Vaurio, L. 1995. Lexical inferencing strategies in reading in English on the secondary level. Lisensiaattitutkimus. Jyväskylän yliopisto: Kasvatustieteiden laitos.
- Vähäpassi, A. (toim.) 1987. Todellisuuden kuvaamista, pohdintaa ja arviointia edellyttäviä kirjoitustehtäviä ja niiden arviointi. Kasvatustieteiden tutkimuslaitos. Teoriaa ja käytäntöä 6.
- Vähäpassi, A. (toim.) 1987. Käytännöllisluonteisia ja opiskelussa tarvittavia kirjoitustehtäviä ja niiden arviointi. Kasvatustieteiden tutkimuslaitos. Teoriaa ja käytäntöä 7.
- Weir, C. J. 1988. Communicative language testing. New York: Prentice Hall.
- Widdowson, H. G. 1990. Aspects of language teaching. Oxford: Oxford University Press.
- Wiggins, G. 1989. "A True Test: Toward More Authentic and Equitable Assessment," Phi Delta Kappan, 70, 9 (May).
- Wiggins, G. 1990. The case for authentic assessment. ERIC Digest ED328611.
- Yli-Renko, K. 1989. Suullinen kielitaito ja sen mittaaminen lukion päättövaiheessa. Tutkimuksia 72. Helsinki: Helsingin yliopiston opettajankoulutuslaitos.
- Yli-Renko, K. (toim.) 1992. Vieraiden kielten puheviestinnän kehittäminen lukion kielenopetuksessa. Turun yliopisto. Kasvatustieteiden tiedekunnan julkaisusarja B. 38.