

# Bedömning av muntlig språkfärdighet

Bedömning av muntlig språkfärdighet, dvs. förmågan att tala ett främmande eller ett andraspråk, var temat för Raili Hildens föredrag i SRO:s sommarkurs på Hanaholmen. Företeelsen är mångfasetterad och den gäller allt från förmågan att använda språkliga enheter och former i monologiskt tal till att kunna interagera i en dialog.

**Raili Hilden**, Helsingfors universitet  
**Mikko Kuronen**, Jyväskylä universitet

Det tal som brukar bedömas varierar från diskussion mellan bedömaren och inläraren (Oral proficiency interview, OPI) till inspelade taluppgifter som simulerar interaktion. Det färskaste bedömningsunderlaget utgörs av portfolioliknande sammandrag av autentiska eller simulerade taluppgifter, som visar hur muntlig färdighet utvecklats hos inlärare under en given tidsperiod.

Bedömning av tal påverkas både av den tilldelade uppgiften och den bedömningsskala som används. Skalan kan vara verbal eller numerisk och den kan ha flera nivåer. De flesta bedömningsskalor i Finland och Europa grundar sig på den europeiska referensramens skala med sex färdighetsnivåer. Den skala som används i grunderna för läroplanerna för den grundläggande utbildningen och gymnasiet specificerade i den föregående versionen (2003, 2004) följande drag i talet: hur talaren klarar av språkbrukssituationen, flyt, uttal, uttrycksförmåga och språklig korrekthet. Dessa kvalitativa drag förekommer ofta i bedömningsskalor på olika håll i världen. I de nuvarande läroplansgrunderna används en bredare och mer integrerad skala, där muntlig färdighet delas in i förmågan att producera texter och förmågan att kommunicera.

Bedömningen påverkas alltid av sitt syfte. Formativ bedömning stöder studierna och lärandet och kan sägas ingå i dem, och på grundval av formativ bedömning fattas i allmänhet inga beslut som på något påtagligt sätt påverkar inlärarens liv. Utöver läraren kan en kompis eller inläraren själv granska sina framsteg med hjälp av formativ bedömning och feedback. Summativ bedömning syftar i sin tur till att mäta och visa den uppnådda kompetensen hos inlärare vid en viss tidpunkt på ett så enhetligt sätt som möjligt. Enhetlighet krävs också när summativ bedömning rapporteras på grund av bedömningens selektiva funktion: resultatet är ju ofta viktigt till exempel när studieplatser tilldelas på nästa utbildningsstadium.

Ju fler bedömare det finns och ju fler olika miljöer de arbetar i, desto svårare är summativ bedömning. I omfattande och viktiga examina, till exempel den finländska studentexamen, medför också lagstiftningen utmaningar i och med att den förutsätter att en talprestation genomgår dubbel bedömning av människor (lärare och censorer). Människor behöver lön, och deras tolkning av kriterierna kan variera även om de har en grundlig utbildning. De blir också trötta och måste lägga ner



mycket tid på att försöka se till att de gör en noggrann bedömning. För att minska det mänskliga arbetet och effektivisera bedömningen utvecklas numera automatiska bedömningssystem både i Finland och utomlands.

I huvuddrag följer utvecklingen av ett automatiskt bedömningssystem följande steg: Studenten genomför en muntlig uppgift och prestationen sparas. Systemet tränas att identifiera talet så exakt som möjligt genom att det jämför mänskliga transkriberingar med den text som det automatiska taligenkänningsystemet producerar. Utbildade bedömare ger en bedömning av allmän färdighetsnivå och en profilbedömning av olika delkriterier (t.ex. flyt och uttal) hos den prestation de lyssnar på. Mänskliga bedömares enighet analyseras, och de talprov som bedömts enigast införs i systemet tillsammans med bedömningsuppgifterna. Modellerna för maskininlärning kombinerar dessa data, identifierar kritiska drag för bedömningen och lär sig att ge en bedömning av nya talprov på samma sätt som mänskliga bedömare skulle bedöma dem. Så här gick man till exempel till väga i projektet DigiTala. Mer information om projektet finns på adressen <https://www.helsinki.fi/sv/projekt/digitalt-stod-ovning-och-bedomning-av-andraspraktal>

Fördelarna med automatisk bedömning jämfört med mänsklig bedömning är uppenbara. Maskinen är snabb, den tröttnar inte, den tar inga pauser och begär ingen lön, och algoritmen behandlar alla talare på samma sätt. AI-baserade verktyg är också mångsidiga, eftersom inlärare på egen hand kan öva talet genom att till exempel göra liknande uppgifter som ingår i ett kommande prov eller test. Instrumenten ger också feedback på olika aspekter av talet och de kan identifiera styrkor och utmaningar hos inläraren.

Automatisk bedömning har naturligtvis också sina begränsningar. Tills vidare bedömer algoritmerna noggrannast monologiskt tal och dess kvantifierbara drag, såsom fonologi, flyt, uttrycksförmåga och i viss mån språklig korrekthet. Identifieringen av viktiga drag i dialog och dessa drags koppling till färdighetsnivåerna är däremot fortfarande mycket utmanande, men även detta undersöks av ett konsortium mellan Helsingfors, Aalto och Jyväskylä universitet, se projektet AASIS <https://www.helsinki.fi/fi/projektit/vieraskielisen-suullisen-vuorovaikutuksen-automaattinen-arviointi>