



DigiTala – Miten automaattinen palaute voi tukea puhumisen harjoittelua?

Artikkelissa esittelen DigiTala-tutkimushankkeen kehittämän sovelluksen, jonka avulla suomen ja ruotsin oppijat voivat harjoitella puheen tuottamista. Automaattiseen puheentunnistukseen, arviointiin ja palautteeseen perustuvan työkalun kehittämisessä on hyödynnetty monitieteistä tutkimusta.

Julkaistu: 10. toukokuuta 2023 | Kirjoittanut: Anna von Zansen

Puhumisen harjoittelua automaattisen arvioinnin avulla

DigiTala-tutkimushankkeen tavoitteena on tukea suomen ja ruotsin puhumisen harjoittelua ja arviointia. Hankkeen kehittämän työkalun avulla suomen ja ruotsin oppijat voivat harjoitella puheen tuottamista ajasta ja paikasta riippumatta. Lisäksi puheen automaattinen arviointi (*automated scoring*) voisi tukea ihmisen tekemää arviointia ylioppilastutkinnon kaltaisissa suurissa ja tärkeissä kielikokeissa (Kautonen & von Zansen 2020).

Automaattinen arviointi perustuu puheentunnistukseen (*automatic speech recognition, ASR*), joka muuttaa puheen tekstiksi. Lisäksi äänitteestä voidaan automaattisesti mitata esimerkiksi puhenopeutta ja taukoja. DigiTala-työkalun (von Zansen ym. 2022a, ks. [esittelyvideo \(https://www.helsinki.fi/en/unitube/video/8db7bb4e-8612-48f4-92ee-3127ab09816c\)](https://www.helsinki.fi/en/unitube/video/8db7bb4e-8612-48f4-92ee-3127ab09816c)) automaattiset arviointimallit perustuvat ihmisten tekemiin arvioihin, joita koneoppimisen keinoin pyritään ennustamaan. Puheen automaattisesta arvioinnista tarkemmin ks. Evanini & Zechner (2020) yleiskatsaus aiheeseen ja Kautonen & von Zansen (2020) DigiTala-hankkeen lähtökohdat.

Työkalun ensimmäinen versio mahdollistaa ääneenluku- ja tuottamistehtävät suomen ja ruotsin kielillä. Aineistonkeruussa käytetyt tehtävät ja materiaalit löytyvät avoimesta [Zenodo \(https://zenodo.org/communities/digitala?page=1&size=20\)](https://zenodo.org/communities/digitala?page=1&size=20)-tietovarantopalvelusta (n.d.), mutta ajatus on, että opettajat voisivat lisätä työkaluun omia uusia tehtäviään.

Työkalu arvioi tuottamistehtävissä taitotason lisäksi tehtävänannon täyttymistä, sujuvuutta, ääntämistä ja ilmauksen laajuutta – samoja puhumisen ulottuvuuksia, joita hankkeen kouluttamat ihmisarvioijat ovat käyttäneet arviointikriteereinä (von Zansen 2022), tarkkuutta lukuun ottamatta. Tarkkuuden ja laajuuden automaattinen arviointi erillisinä ominaisuuksina oppijoiden puheesta osoittautui vaikeaksi hankkeessa käytetyllä opetusaineistolla. Ääneenlukutehtävissä sen sijaan on mielekästä arvioida vain sujuvuutta ja ääntämistä.

Automaattinen palaute kertoo puhujan vahvuuksien ja kehittämiskohteiden lisäksi taitotason

Puheen äänittämisen jälkeen työkalu esittää arvioinnin tulokset puhujalle automaattisena palautteena. Seuraavissa kuvissa (kuvat 1 ja 2) esitellään lyhyesti palautenäkymä. Lisää kuvia työkalusta löytyy käyttöohjeesta (Alanen ym. 2022).

Puhujalle näytettävä analyttinen palaute (kuva 1) on muotoiltu arviointikriteerien pohjalta. Lisäksi palautenäkymässä puhuja voi kuunnella nauhoittamansa äänitteen ja näkee puheensa muutettuna tekstiksi.

1 Aloitus > 2 Tehtävä > 3 Arviointiraportti

Arviointiraportti

Lähetetty: 10.05.2022 11:21:13

Tämä palaute koskee ainoastaan nauhoittamaasi puhenäytettä, eikä se kuvaa kaikkea suullista kielitaitoa. Automaattinen arvio on koneen tekemä. Koneita on opetettu muiden kielten oppijoiden puheella ja muulla kielilaineistolla.

Tässä tehtävässä ei ole yrityskertojen rajaa.

0:00 / 0:03

Puhenäytteesi tekstinä

öö mult oli jäi huppari tellä sinne kahvilaan eilen

Analyttinen arvio Taitotasoarvio

Sujuvuus

Tämä mittari kertoo puhenäytteesi nopeudesta, taukojen määrästä ja empimisestä.

0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0

4/4

Automaattisen arvion mukaan vaikuttaa siltä, että puheesi on todella sujuvaa ja vaivatonta, ei häiritseviä taukoja, katkoksia tai empimistä.

Ääntäminen

Näet yllä, että kone muunsi puheesi tekstiksi. Voit tarkistaa tekstistä, lausulko kaikki sanat oikein. Tämä mittari kertoo, kuinka hyvin ja varmasti kone tunnistaa puheesi. Tunnistamistarkkuuteen vaikuttavat puhenäytteet, joita kone on aiemmin opetusvaiheessa saanut.

0 0.5 1.0 1.5 2.0 2.5 3.0 3.5 4.0

3/4

Automaattisen arvion mukaan vaikuttaa siltä, että kone ymmärtää puhettasi, ääntämisessä ei vaikuta olevan suurempia ongelmia.

Tehtävänantoon vastaaminen

Tämä mittari perustuu vastauksiin, joilla kone on opetettu arvioimaan tätä tehtävää.

0 0.5 1.0 1.5 2.0 2.5 3.0

0/3

Automaattisen arvion mukaan vaikuttaa siltä, että välitettävästi kone ei ole kuullut tämänkaltaista suoritusta aiemmin, eikä siksi osaa arvioida puhettasi.

Laajuus

Tämä mittari kertoo, kuinka paljon olet puhunut sekä käyttämiesi sanojen ja lauseiden monipuolisuudesta.

0 0.5 1.0 1.5 2.0 2.5 3.0

2/3

Automaattisen arvion mukaan vaikuttaa siltä, että käytät tavallisia sanoja ja osaat tehdä niistä lauseita.

Yritä uudelleen

Anna palautetta ?

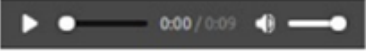
Kuva 1. Automaattinen analyttinen palaute. Ks. kuva isompana pdf-tiedostossa (https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-toukokuu-2023/vonzansen_kuva1_iso.pdf).

Palautenäkymän toisella välilehdellä (kuva 2) näytetään taitotasoarvio asteikolla alle A1–C2. Taitotasokuvaukset pohjautuvat Eurooppalaiseen viitekehukseen, jota suomalaisissa opetussuunnitelman perusteissa on sovellettu vuodesta 2003 lähtien. Työkalun antama taitotasoarvio voisikin konkretisoida taitotasoa kielten opetuksessa ja tuoda näkyväksi puheesta mitattavia ominaisuuksia.

Arviointiraportti
Lähetetty: 10.05.2022 11.21:13

Tämä palaute koskee ainoastaan nauhoittamaasi puhenäytettä, eikä se kuvaa kaikkea suullista kielitaitoasi. Automaattinen arvio on koneen tekemä. Koneetta on opetettu muiden kielen oppijoiden puheella ja muulla kieliaineistolla.


Tässä tehtävässä ei ole yrityskertojen rajaa.



Puhenäytteesi tekstinä
öö mult oli jäi huppari teillä sinne kahvilaan eilen

Analyttinen arvio Taitotasoarvio

Taitotasoarvio



B1
Automaattisen arvion mukaan vaikuttaa siltä, että taitotasosi on B1.

Selviydyt arkielämän tilanteista kohdekielellä. Ääntämisesi on ymmärrettävää, käytät melko laajaa sanastoa ja erilaisia lauseita.

[Yritä uudestaan](#)

Kuva 2. Automaattinen holistinen palaute. Ks. [kuva isompana pdf-tiedostossa \(https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-toukokuu-2023/vonzansen_kuva2_iso.pdf\)](https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-toukokuu-2023/vonzansen_kuva2_iso.pdf).

Opiskelijat näkevät automaattisessa arvioinnissa haittoja ja hyötyjä

DigiTala-työkalu on kehitetty opiskelijatyönä tutkimusprojektissa eikä sen laajamittainen itsenäinen käyttö oppilaitoksissa ole vielä mahdollista. Työkalu on julkaistu vapaiden ohjelmistojen lisenssillä ohjelmistokehitysprojekteille tarkoitetussa Github-palvelussa (von Zansen ym. 2022a), mikä mahdollistaa työkalun jatkokehityksen ja ylläpidon hankkeen päättyessä. Olemme kuitenkin tutkineet kielen oppijoiden, opettajien ja ihmisarvioijien käsityksiä tietokoneella suoritettavista puhetehtävistä ja puheen automaattisesta arvioinnista.

Tutkimuksiimme osallistuneet opiskelijat toivat esiin vastaavia haittoja ja hyötyjä, joita mainitaan alan kirjallisuudessa (Evanini & Zechner 2020). Automaattisen arvioinnin edut liittyvät etenkin lähes reaaliaikaisesti saatavan automaattisen palautteen tuomiin mahdollisuuksiin: puhumista

voi harjoitella oma-aloitteisesti ajasta ja paikasta riippumatta. Kone ei myöskään väsy eikä sillä ole arviointiin mahdollisesti vaikuttavia ennakkokäsityksiä kielenoppijasta. Opiskelijat näkivät automaattisen arvioinnin säästävän opettajien työtaakkaa ja lisäävän tehokkuutta.

Opiskelijoiden näkemät huolet puolestaan liittyvät muun muassa automaattisen arvioinnin reiluuteen ja paikkansapitävyyteen: ymmärtääkö kone erilaisia puhujia, ja antaako arvio kattavan kuvan puhujan suullisesta kielitaidosta? Lisäksi opiskelijoita mietityttivät tietoturvaan ja tietosuojaan liittyvät kysymykset, mikä on täysin aiheellista, sillä ihmisen puheääni on henkilötietoa. Oppijat eivät myöskään olleet tottuneet puhumaan tietokoneelle, joten oman puheen nauhoittaminen vaatii totuttelua ja harjoittelua. DigiTalan kehittämät puhetehtävät (Zenodo n.d.) olivat oppijoiden mielestä pääsääntöisesti onnistuneita, vaikkakin jossain määrin uudenlaisia. (von Zansen, Sneck & Hilden 2022; von Zansen & Hilden 2022)

Opettajat ja arvioijat löytävät automaattiselle arvioinnille useita käyttötarkoituksia

DigiTala-hankkeen haastattelemat opettajat näkivät opiskelijatyönä kehitetylle työkalulle useita käyttötarkoituksia. Automaattiseen arviointiin perustuvaa työkalua voitaisiin esimerkiksi käyttää kotitehtävänä tai osana lukion kieliprofiilia tai suullista koetta. Automaattinen palaute on tervetullutta, sillä kielenopettajilla on rajallisesti aikaa antaa palautetta kullekin oppijalle. Opettajat olivat kuitenkin huolissaan automaattisen palautteen mahdollisista kielteisistä vaikutuksista. Työkalu onkin suunniteltu niin, että opettajalla on mahdollisuus kommentoida ja täydentää koneen antamaa palautetta. (von Zansen & Heijala 2023)

Kuten edellä todettiin, automaattinen arviointi perustuu ihmisarvioijien antamiin arviointeihin opetusaineistona käytetyistä puhenäytteistä. Siksi on tärkeää tutkia arviointien laatua ja arvioijien käsityksiä. Kuten opettajat, myös hankkeen kouluttamat ihmisarvioijat suhtautuivat pääsääntöisesti positiivisesti puheen automaattiseen arviointiin ja toivat esiin vastaavia hyötyjä ja uhkia kuin muut hankkeen tutkimuksiin osallistuneet käyttäjäryhmät.

Ihmisarvioijat, jotka taustaltaan ovat kielenopettajia tai kielitestaajia, hyödyntäisivät automaattista arviointia puheen itsenäisessä harjoittelussa, puhujan tason kartoituksessa ja kielitestien mukauttamisessa. Osallistuneiden arvioijien mielestä valitut puheen ulottuvuudet ovat keskeisiä, joskaan vuorovaikutuksen arviointiin kone ei vielä pysty. Analyttinen arviointi suhteellisen lyhyistä (15 sekuntia – 1,5 minuuttia) puhenäytteistä oli jossain määrin uutta arvioijille, jotka ovat tottuneet arvioimaan oppijan suullista kielitaitoa kokonaisvaltaisemmin. Kaikkien ulottuvuuksien – kuten taitotason ja sujuvuuden – arvioiminen ei myöskään onnistu, jos puhenäyte on liian lyhyt. (von Zansen ym. 2022b)

Miltä puheen automaattisen arvioinnin tulevaisuus näyttää Suomessa?

Tässä artikkelissa esittelin monitieteisen tutkimushankkeen kehittämän prototyypin työkalusta, jonka avulla suomen ja ruotsin oppijat voisivat harjoitella puheen tuottamista ja ääneen lukua. Lisäksi oppijat saavat tuotoksensa perusteella automaattista palautetta taitotasostaan sekä neljästä muusta puheen ulottuvuudesta. Puhujan vahvuuksien ja kehityskohteiden osoittaminen toisi lisää mahdollisuuksia puhumisen itsenäiseen harjoitteluun. Tämän lisäksi automaattista arviointia voitaisiin hyödyntää ihmisen tekemän arvioinnin tukena ylioppilastutkinnon kaltaisissa tärkeissä kielikokeissa, joista puhumisen osakoe vielä puuttuu (Vaarala ym. 2021).

Opettajat ovat suullisen osakokeen kannalla. Alustavien – vielä julkaisemattomien – tutkimustulostemme mukaan opettajien asenteet ovat 10 vuodessa muuttuneet luottavaisempaan suuntaan, sillä suullisen osakokeen kannatus ylioppilastutkinnon kielikokeissa on kaksinkertaistunut oppimäärästä riippumatta (vrt. LukSuS 2013). Sen sijaan alustavien tulostemme mukaan opettajien tärkeinä pitämät puheen ulottuvuudet ovat säilyneet ennallaan – sujuvuuden, ääntämisen ja vuorovaikutuksen arvioiminen on edelleen keskeistä.

Hankkeen tutkimuksiin osallistuneet kielen oppijat, opettajat ja arvioijat suhtautuvat puheen automaattiseen arviointiin pääsääntöisesti positiivisesti ja ymmärtävät tekoälyn käyttöön liittyviä rajoituksia. Keskeisimmät haasteet liittyvät vuorovaikutuksen puutteeseen sekä algoritmisen syrjinnän ennaltaehkäisyyn (ks. Ojanen ym. 2022). Tekoälyä hyödyntävien sovellusten käyttöön liittyy useita eettisiä ja teknisiä kysymyksiä, jotka liittyvät esimerkiksi tasa-arvoon ja saavutettavuuteen. Pidän tärkeänä, että automaattisen arvioinnin toimintaperiaatteista kerrotaan käyttäjille ymmärrettävällä tavalla (*explainable AI, XAI*). Lisää tutkimusta ja kehitystyötä tarvitaan myös vuorovaikutuksen arvioinnista – pidän täysin mahdollisena, että myös vuorovaikutusta voidaan tulevaisuudessa mitata ja arvioida tekoälyn avulla.

Anna von Zansen työskentelee tutkijatohtorina Suomen Akatemian 2019-2023 rahoittamassa DigiTala-hankkeessa, jota johtaa Raili Hilden Helsingin yliopistosta. Hankkeen tutkijat Helsingin yliopistosta (rahoituspäätös 322619), Aalto-yliopistosta (rahoituspäätös 322625) ja Jyväskylän yliopistosta (rahoituspäätös 322965) tekevät kielididaktista, foneettista ja kieliteknologista tutkimusta. Tällä hetkellä von Zansen sijaistaa lukion kielten opetuksesta ja suullisen kielitaidon arvioinnista vastaavaa opetusneuvosta Opetushallituksessa.

Lähteet

Alanen, T., Erkkilä, J., Harjunpää, T., & Heijala, M. (2022). *Digitala Moodle plugin user manual (1.0.0)*. Zenodo. <https://doi.org/10.5281/zenodo.6535377> (<https://doi.org/10.5281/zenodo.6535377>)

Evanini, K. & Zechner, K. (2020). Overview of automated speech scoring. Teoksessa K. Zechner & K. Evanini (toim.), *Automated Speaking Assessment*. New York: Routledge, 3–20.

Kautonen, M. & von Zansen, A. (2020). DigiTala research project: Automatic speech recognition in assessing L2 speaking. *Kieli, koulutus ja yhteiskunta*, 11(4). Saatavilla:

<https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-kesakuu-2020/digitala-research-project-automatic-speech-recognition-in-assessing-l2-speaking>

(<https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-kesakuu-2020/digitala-research-project-automatic-speech-recognition-in-assessing-l2-speaking>)

Ojanen, A., O. Sahlgren, J. Vaiste, A. Björk, J. Mikkonen, K. Kimppa, Laitinen, A. & Oljakka, N. (2022). Algoritminen syrjintä ja yhdenvertaisuuden edistäminen: Arviointikehikko syrjimättömälle tekoälylle. *Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 2022*:54.

<http://urn.fi/URN:ISBN:978-952-383-404-0> (<http://urn.fi/URN:ISBN:978-952-383-404-0>)

LukSuS (2013). Lukion suullisen kielitaidon kurssin seuranta- ja tutkimushanke LukSuS. Loppuraportti opetus- ja kulttuuriministeriölle.

Vaarala, H., Riuttanen, S., Kyckling, E. & Karppinen, S. (2021). *Kielivaranto. Nyt!: Monikielisyys vahvuudeksi -selvityksen (2017) seuranta*. Soveltavan kielentutkimuksen keskus, Jyväskylän yliopisto. https://www.jyu.fi/hytk/fi/laitokset/solki/tutkimus/julkaisut/pdf-julkaisut/kielivaranto-nyt-_julkaisu_sivuittain-1.pdf (https://www.jyu.fi/hytk/fi/laitokset/solki/tutkimus/julkaisut/pdf-julkaisut/kielivaranto-nyt-_julkaisu_sivuittain-1.pdf)

von Zansen, A. 2022. *DigiTala's rating criteria: Holistic and analytic scales for assessing L2 speaking*. Zenodo. <https://doi.org/10.5281/zenodo.6477089> (<https://doi.org/10.5281/zenodo.6477089>)

von Zansen, A., Alanen, Al-Ghezi, R., Erkkilä, J., Harjunpää, T., Heijala, M. & Kallio, H. (2022a). *DigiTala Moodle plugin*. https://github.com/aalto-speech/moodle-mod_digitala (https://github.com/aalto-speech/moodle-mod_digitala)

von Zansen, A. & Heijala, M. (2023). Miten suomen ja ruotsin opettajat käyttäisivät puheen automaattiseen arviointiin kehitettyä työkalua? Teoksessa Mäkipää, T., Hilden, R., Huhta, A. (toim.) *AFinLA-teema*, (15), 124–141. <https://doi.org/10.30660/afinla.124822> (<https://doi.org/10.30660/afinla.124822>)

von Zansen, A., Kallio, H., Sneck, M., Kuronen, M., Huhta, A. & Hildén, R. (2022b). Ihmisarvioijien näkemyksiä suullisen kielitaidon automaattisesta arvioinnista, digitaalisesta arviointiprosessista sekä puhe-suorituksista arvioitavista ulottuvuuksista. Teoksessa Seppälä, T., Lesonen, S., Ilikkanen, P., D'hondt, S. (toim.) *AFinLAn vuosikirja 2022*, Suomen soveltavan kielitieteen yhdistyksen julkaisuja 79, 370–394.

von Zansen, A. & Hilden, R. (2022). "It was cool and comfortable!" Akateemisten alkeistason S2-opiskelijoiden kokemuksia tietokoneella suoritettavasta puhumisen kokeesta. Teoksessa S. Routarinne, P. Heinonen, T. Kärki, A. Roiha, M-L. Rönkkö, & A. Korkeaniemi (Toimittajat),

Ainedidaktiikka ajassa: Laajenevat oppimisympäristöt ja eri-ikäiset oppijat. Suomen Ainedidaktinen tutkimusseura, 72–90 . Turun yliopiston opettajankoulutuslaitos.
<http://hdl.handle.net/10138/353562>

von Zansen, A., Sneck, M. & Hildén, R. (2022). Lukiolaisten käsitykset ja heidän antamansa palaute suullisen kielitaidon automaattisesta arvioinnista. [Language learners' perceptions and their feedback on automated assessment of oral language skills.] *LINGUAPEDA 2021. Conference Proceedings*. Suomen ainedidaktisen tutkimusseuran julkaisuja. Ainedidaktisia tutkimuksia 21,188–219. <http://hdl.handle.net/10138/352128>

Zenodo. n.d. DigiTala research project (2019–2023) <https://zenodo.org/communities/digitala> (<https://zenodo.org/communities/digitala>)

Artikkeliin viittaaminen

von Zansen, A. (2023). DigiTala – Miten automaattinen palaute voi tukea puhumisen harjoittelua? *Kieli, koulutus ja yhteiskunta*, 14(3). Saatavilla: <https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-toukokuu-2023/digitala-miten-automaattinen-palaute-voi-tukea-puhumisen-harjoittelua> (<https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-toukokuu-2023/digitala-miten-automaattinen-palaute-voi-tukea-puhumisen-harjoittelua>)