



KOKEMUKSIA KENTÄLTÄ

DigiTala-tutkimushankkeen
päätösseminaari
13.4.2023 Jyväskylä
anna.vonzansen@helsinki.fi
raili.hilden@helsinki.fi



MITEN OPISKELIJAT, OPETTAJAT JA IHMISARVIOIJAT SUHTAUTUVAT PUHEEN AUTOMAATTISEEN ARVIOINTIIN?



PEDAGOGISET JA TEKNISET LÄHTÖKOHDAT

1. Mitä on tarkoitus mitata?
(konstruktiivin määrittely)
 2. Millaisilla tehtävillä puhumista voidaan mitata? (tehtävän laadinta)
 3. Millä kriteereillä puhe-suorituksia voidaan arvioida? (kriteerit)
- Yksilön puheetta, puheesta tekn. mitattavat ulottuvuudet (audio)
 - Ääneenluku, lyhyet reagoimistehtävät, pidemmät tuottamistehtävät
 - LOPS2003, 2015, 2019: Holistinen (taitotaso) & Analyttiset (tehtävän suorittaminen, sujuvuus, ääntäminen, laajuus, tarkkuus)



OPISKELIJOIDEN ESIIN TUOMIA HUOLIA

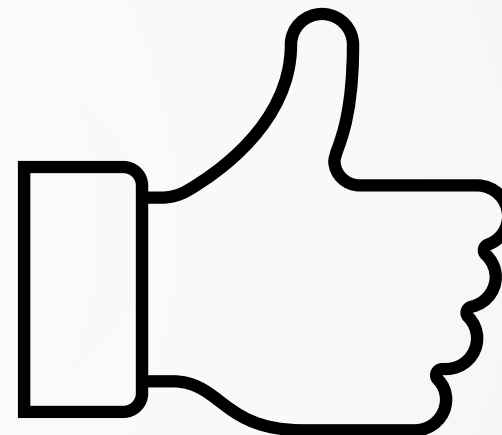
- Entä jos muut opiskelijat kuuntelevat mitä vastaan..
- Entä jos tärkeässä kokeessa muut puhuvat päälle?
- Jännittää, koska en ole tottunut puhumaan tietokoneelle.
- Mihin oma ääneni tallentuu ja menee?
- Antaako koe todellisen kuvan kielitaidostani?
- Ymmärtääkö kone puheeni oikein?
- Miten kone ymmärtää puhekieltä?
- Huomioiko kone reilusti yksilölliset tavat puhua ?





OPIKSELIJOIDEN NÄKEMIÄ HYÖTYJÄ

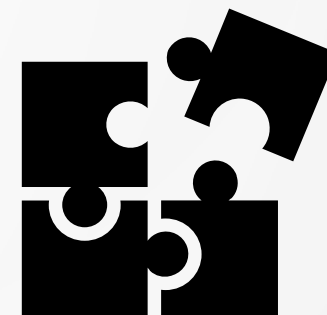
- Kone arvioi 24/7 väsymättä.
- Koneelta saan palautteen heti.
- Voin harjoitella missä haluan milloin haluan.
- Voin harjoitella oma-aloitteisesti.
- Kone arvioi kaikki samalla tavalla.
- Säästää opettajan työtaakkaa.
- Lisää tehokkuutta, säästää kustannuksia ja aikaa.
- Koneelle puhuessa vuorovaikutusta ei tarvitse jännittää.





ARVIOIJEN NÄKEMYKSIÄ

- Kuten opiskelijat: positiivinen suhtautuminen, vastaavia etuja ja uhkia
- Automatiikka ihmisen tueksi: itsenäinen harjoittelu, tason kartoitus, mukautuvuus
- Analyttinen arviointi hitaampaa kuin holistinen
- Tiettyjen ulottuvuuksien arviointi mahdotonta lyhyestä näytteestä
- Arvioinnin tarkoitus > mitkä puheen ulottuvuudet keskeisiä?





OPETTAJIEN ENSIVAIKUTELMAT TYÖKALUSTA

- Palaute kriteerien pohjalta
- Useita käyttötarkoituksia
- Mahdolliset vaikutukset?
- Työkalun räätälöitävyys?
- Opettajaa tarvitaan
- Käsitteelliset erot
- Taitotasojen konkretisoiminen

Deeva et al. 2021

Gu & Davis 2020

von Zansen & Heijala (hyväksytty)

von Zansen & Huhta 2022

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. Koneita on opetettu muiden kielen oppijoiden puheella ja muulla kieliaineistolla.

Tässä tehtävässä ei ole yrityskertojen rajaa.



Puhenäytteesi tekstinä

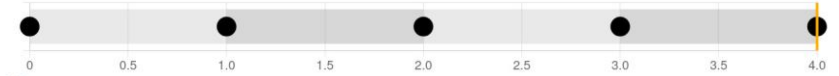
öö mult oli jäi huppari teillä sinne kahvilaan eilen

Analyttinen arvio

Taitotasoarvio

Sujuvuus

Tämä mittari kertoo puhenäytteesi nopeudesta, taukojen määrästä ja empimisestä.

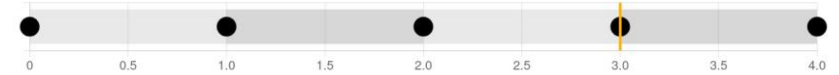


4/4

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, lausuitko kaikki sanat oikein. Tämä mittari kertoo, kuinka hyvin ja varmasti kone tunnistaa puheesi. Tunnistamistarkkuuteen vaikuttavat puhenäytteet, joita kone on aiemmin opetusvaiheessa saanut.



3/4

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

Tehtävänantoon vastaaminen

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



0/3

Automaattisen arvion mukaan vaikuttaa siltä, että valitettavasti 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.



2/3

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

Yritä uudelleen



OPETTAJIEN KÄSITYKSIÄ TÄNÄÄN JA 10 VUOTTA SITTEN – MUUTTUUKO MAAILMA?

- 129 kieltenopettajaa vuonna 2023 ja 202 vuonna 2012 vastasivat verkkokyselyyn

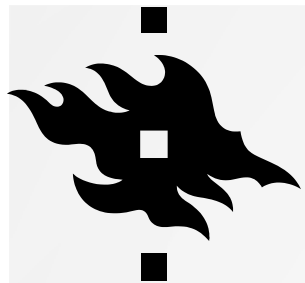
RQ1 Mitä käsityksiä ja kokemuksia kieltenopettajille on puhumisen käsitteestä ja arvioinnin tarkoituksista?

- Sujuvuus, ääntäminen ja vuorovaikutus arviointikriteerit tärkeimmät molemmissa tutkimuksissa
- Kielioppi vähemmän tärkeä
- OPS-taitotasoasteikon soveltaminen suullisen kielitaidon arviointiin
- 2012: 67% samaa mieltä taidostaan soveltaa LOPS 2003-asteikkoa
- 2023: 53% samaa mieltä taidostaan soveltaa LOPS 2015-asteikkoa



PITÄISIKÖ YO-KIELIKOKEISSA OLLA SUULLINEN OSA?

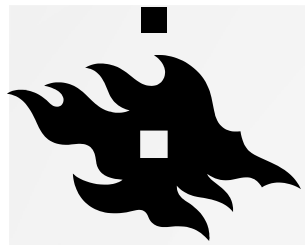
		English A	Finnish / Swedish (B1)	Swedish / Finnish (native-level)	Swedish / Finnish (A)	German / French / Russian / Spanish (A)	B2 and B3 languages	Finnish (S2) / Swedish as a second language and literature
2012	N	218	208	202	203	206	207	-
	Yes	46,3% 101	33,2% 69	41,1% 83	43,8% 89	39,8% 82	26,1% 54	-
2023	N	127	124	121	124	123	124	121
	Yes	89% 113	71% 88	77,7% 94	86,3% 107	81,3% 100	58,9% 73	76,0% 92



OPETTAJIEN KÄSITYKSIÄ TÄNÄÄN JA 10 VUOTTA SITTEEN – MUUTTUUKO MAAILMA?

- RQ2 Mitä käsityksiä ja kokemuksia opettajilla on automaattisesta puhumisen arvioinnista
- Q10 Miten suhtaudut puheen automaattiseen arviointiin?

Question	Very negative	Quite negative	Neutral	Quite positive	Very positive	Mean	Standard deviation
Q10 n=129	1,6 % 2	11,6 % 15	36,4 % 47	34,1 % 44	16,3 % 21	3,52	0,953



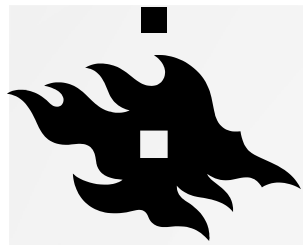
YHTEENVETONA

- Automaattinen arviointi tavoittaa osan kielitaidosta
- Viestinnällisyyttä painottavan opetuksen ja arvioinnin symmetria ei toteudu, jollei puhumista mitata
- Puhe on henkilökohtaista ja henkilötietoa
- Tietokoneelle puhuminen uutta – vaatii harjoittelua
- Tehtävänlaadinnan haasteet
- (Automaattisen) arvioinnin käyttötarkoitus & vaikutukset
- Käyttäjien suhtautuminen positiivista
- Tekniikan käytettävyys ja käyttövarmuus
- Opettajan työn helpottuminen, opiskelijoiden omatoiminen harjoittelu ja palaute
- Kone on virheetön ja väsymätön rajallisella sovellusalallaan



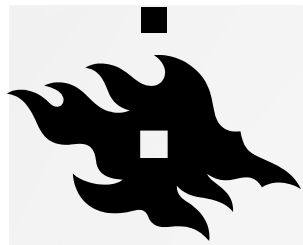
TEKOÄLY JA KONEARVIOINTI SUOMALAISEN KIELIKASVATUKSEN PALVELUKSESSA

- Tekniikan haltuunotto kannattaa, vars. Kielissä, jotka eivät kuulu nuorison vapaa-ajan maailmaan englannin tapaan
- Monet nykyisten opetussuunnitelmaperusteiden periaatteet ja kielikasvatuksen tavoitteet hyötyvät digitaalisista sovelluksista (esim. opiskelun ja oppimisen strategioiden harjoittelu ja seuranta, itseohjautuvuus, koulun ulkopuolella tapahtuva oppiminen)
- Suullisen kielitaidon/puhumisen asema viestinnällisen kielitaidon arvioinnissa
- Digitala-hanke on kehittänyt kotimaisiin kieliin prototyypin, joka helpottaisi puhumisen osakokeen käytännön toimeenpanoa ylioppilastutkinnon kielikokeissa. Prototyyppi on käyttövalmis, mutta sen hienosäätö laajamittaiseen käyttöön vaatii vielä resursseja, samoin sen pitkäaikainen ylläpito.



LÄHTEET 1/2

- ALTE (2020) Principles of Good Practice.
- Deeva, G., D. Bogdanova, E. Serral, M. Snoeck & J. De Weerd 2021. A review of automated feedback systems for learners: Classification framework, challenges and opportunities. *Computers & Education*, 162. <https://doi.org/10.1016/j.compedu.2020.104094>
- Evanini, K., & Zechner, K. (2020). Overview of automated speech scoring. In K. Zechner & K. Evanini (eds.), *Automated Speaking Assessment: Using Language Technologies to Score Spontaneous Speech*. Routledge, 3–20.
- Gu, L., & L. Davis. 2020. Providing SpeechRater feature performance as feedback on spoken responses. Teoksessa K. Zechner & K. Evanini (toim.) *Automated Speaking Assessment: Using language technologies to score spontaneous speech*. New York: Routledge, Taylor & Francis Group, 159–175.
- 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>



LÄHTEET 2/2

- von Zansen, A. & Heijala, M. (accepted). Miten suomen ja ruotsin opettajat käyttäisivät puheen automaattiseen arviointiin kehitettyä työkalua? [How would Finnish and Swedish teachers use a tool developed for automated L2 speaking assessment?].
- 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 ["It was cool and comfortable!" Academic L2 Finnish learners' perceptions of a computer-based speaking test.] S. 72-90. Suomen Ainedidaktinen tutkimusseura. Turun yliopiston opettajankoulutuslaitos. <http://hdl.handle.net/10138/353562>
- von Zansen, A & Huhta, A 2022, Developing Automated Feedback on Spoken Performance : Exploring the Functioning of Five Analytic Rating Scales Using Many-facet Rasch Measurement . in J. H. Jantunen, J. Kalja-Voima, M. Laukkarinen, A. Puupponen, M. Salonen, T. Saresma, J. Tarvainen & S. Ylönen (eds.) *Diversity of Methods and Materials in Digital Human Sciences: Proceedings of the Digital Research Data and Human Sciences DRDHum Conference 2022, December 1-3, Jyväskylä, Finland. University of Jyväskylä, Jyväskylä*, (pp. 211-229), Digital Research Data and Human Sciences (DRDHum), Jyväskylä, Finland, 01/12/2022 <http://urn.fi/URN:ISBN:978-951-39-9450-1>
- von Zansen, A., Kallio, H., Sneck, M., Kuronen, M., Huhta, A., Hilden, R. (2022). Ihmisarvioijien näkemyksiä suullisen kielitaidon automaattisesta arvioinnista, digitaalisesta arviointiprosessista sekä arvioitavista puheen ulottuvuuksista. [Human raters' perceptions of the automated assessment of oral language skills, the digital assessment process and the dimensions to be assessed from speaking performances.] *AFinLA yearbook* <https://doi.org/10.30661/afinlavk.114821>
- von Zansen, A., Sneck, M., Hilden, 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, p.188–219. <http://hdl.handle.net/10138/352128>