

Kautonen, Maria och Kuronen, Mikko 2021: Kvantitativa perspektiv på L2-tal på olika färdighetsnivåer. *Folkmålsstudier*, nr 59. S. 11–39.

Maria Kautonen och Mikko Kuronen

KVANTITATIVA PERSPEKTIV PÅ L2-TAL PÅ OLIKA FÄRDIGHETS-NIVÅER

Quantitative perspectives on L2 speech at different proficiency levels

The aim of this study is to investigate differences in the amount (the occurrence rate) of second language (L2) traits at different proficiency levels in spoken L2 Finland-Swedish. In the analysis, we focus on L2 features in *segments*, *segment lengths*, *word stress* and *sentence stress*. Further, we analyze the ability to produce speech and the complexity of production at different proficiency levels: *the amount of speech*, *token length*, *token structure* and *speech rate*. Through previous research, we know that pronunciation plays an important role for comprehensibility and foreign accent but also for the level of oral proficiency. Although pronunciation and different pronunciation features are usually more target-like at higher proficiency levels, knowledge on the amount and type of L2 traits at different levels in Swedish is scarce. The results augment our understanding of L2 speaking construct and have the potential to enhance development of assessment criteria for oral proficiency. Quantitative knowledge of pronunciation skills at different proficiency levels is also needed to develop automated (digital) assessment of and feedback on L2 speaking.

The speakers in the study ($n = 24$) were assessed at proficiency levels A1–C2 (CEFR) in oral proficiency tests. We examined their speech in monologues of 40–60 seconds. The analysis was carried out by carefully listening and coding the L2 features in the selected phonetic categories. The results show that the speakers' pronunciation of segments improves almost linearly from A1 to C2, but this is not the case in sentence prosody. On the contrary, the amount of L2 traits in sentence prosody increases at level B2. The ability and complexity of speech production also increase at level B2, and there seems to be a connection between the increase of prosodic L2 traits and the increase of ability and complexity of speech production.

1 Inledning

Utveckling av datateknik och dess tillämpningar har stor potential att betjäna andra- och främmandespråksinlärning¹ (L2-inlärning) och bedömning av språkfärdighet. Automatisk, dvs. datorbaserad eller maskinell, bedömning av skriftlig språkfärdighet har tillämpats inom andraspråksinlärning i över ett

¹ I artikeln gör vi ingen skillnad mellan begreppen andraspråk (L2) och främmande språk (jfr Abrahamsson 2009:14) utan syftar med andraspråk och L2 till alla sådana språk som inte är talarens förstaspråk (L1).

decennium, och den bedömning som datorn numera kan ge av L2-texter anses tillförlitlig, dvs. den kan ge lika bra resultat som professionell mänsklig bedömning om programvaran är av hög kvalitet (jfr t.ex. *e-rater*² som utvecklats av Educational Testing Service, ETS, se Monaghan & Bridgeman 2005, Burnstein m.fl. 1998). Vid automatisk bedömning av skriftlig färdighet brukar de bedömda egenskaperna (eng. *features of assessment*) gälla olika aspekter av ordförråd (t.ex. storlek, variation och precision) och innehåll (kohesion och grammatik). Under det senaste decenniet har också metoder för automatisk bedömning av uttal och muntlig färdighet i L2 utvecklats (Kyle & Crossley 2015, Isaacs 2018). Vid bedömning av L2-tal används automatisk taligenkänning (eng. *automatic speech recognition, ASR*; t.ex. Eskenazi 1999, Zechner m.fl. 2009). Utgångspunkten för automatisk bedömning av tal är att tal först konverteras till text (tal-till-text, eng. *speech-to-text, STT*), och sedan jämförs texten med L1-talares och/eller L2-talares prestationer i en likadan uppgift. Automatisk bedömning av språkfärdighet har hittills utvecklats huvudsakligen för engelska, men det pågår också forskning om andra språk. Bland annat har system som ger feedback på språkfärdighet och/eller uttal i tyska, holländska och svenska utvecklats (Gretter m.fl. 2019, Cucchiariini m.fl. 2007 respektive Wik 2011).

Vad gäller automatisk bedömning av tal har bedömning av uttal och flyt visat sig lika tillförlitlig som vid mänsklig bedömning, dvs. bedömningarna uppnår ungefär samma precision. I bästa fall är automatisk bedömning mer objektiv i betydelsen att den uppvisar mindre variation vid bedömning av enskilda prov än mänsklig bedömning. Vid automatisk bedömning av uttal har följande egenskaper visat sig fungera bra i L2-engelska: *akustisk korrekthet av segment, antalet betonade stavelser i procent, avståndet mellan betonade stavelser, avståndet mellan stavelser med intonationsförändring, andelen tal som består av vokaler samt genomsnittlig durationsskillnad mellan konsekutiva fonetiska enheter, till exempel mellan vokaler och stavelser* (t.ex. Hsieh, Zechner & Xi 2020). Dessa egenskaper fungerar förmodligen också i många andra språk, t.ex. i svenska, eftersom många av egenskaperna är språkoberoende – eller mer precist kan egenskaperna vara desamma även om de kan realiseras annorlunda i olika språk. Vid automatisk bedömning av flyt brukar de bedömda egenskaperna vara följande: *artikulationshastighet, taltempo, antalet tysta och fyllda pauser, genomsnittlig pauslängd, antalet långa pauser, antalet pauser inne i fraser, genomsnittlig längd på yttranden mellan två pauser samt korrigeringar* (t.ex. Cucchiariini, Strik & Boves 2002). Vad gäller bedömning av ordförråd och innehåll möter automatisk bedömning större utmaningar än bedömning av

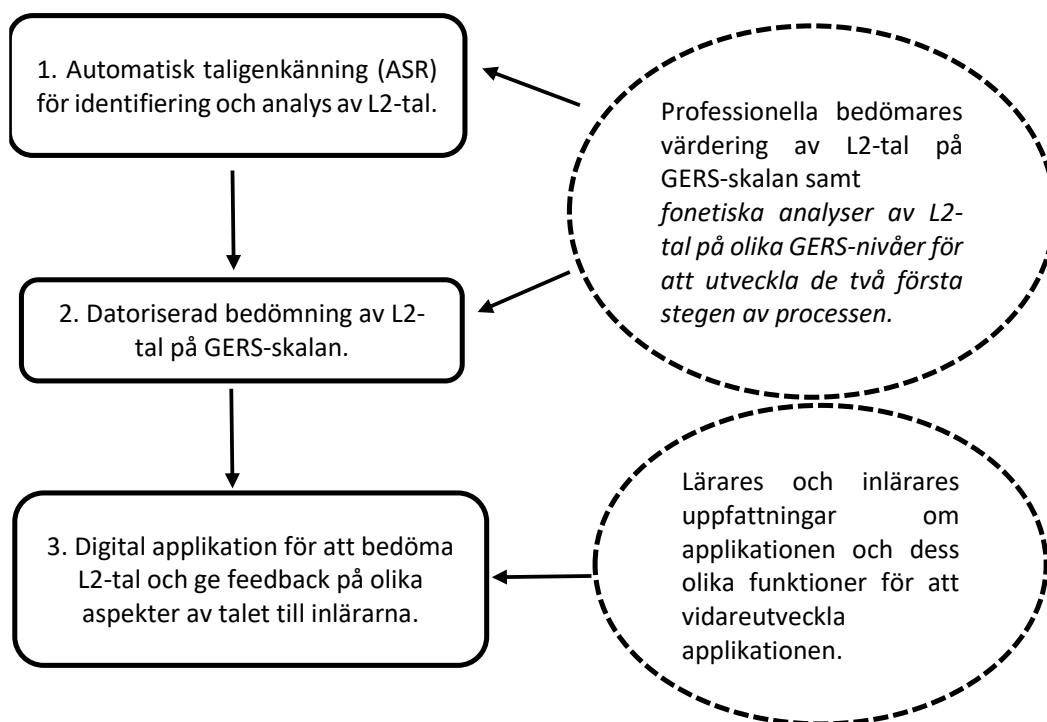
² <https://www.ets.org/erater/about>

uttal och flyt, och därför bedöms speciellt innehåll ännu oftast av människan eller av människan och datorn i samarbete (s.k. hybridbedömning, eng. *hybrid scoring*, se t.ex. Yoon & Zechner 2017). Hybridbedömning kan öka tillförlitligheten vid bedömning av L2-tal genom att det bästa hos datorn och människan kombineras till exempel så att datorn bedömer uttal och flyt, medan människan fokuserar på bedömning av innehåll.

Automatisk bedömning av L2-tal är en viktig utveckling, eftersom man med hjälp av den har potential att bedöma ett stort antal talare tidseffektivt och tillförlitligt (jfr Isaacs 2018). Datorn blir inte trött och den behandlar alla talare likadant. I framtiden kan muntlig färdighet t.ex. i den finländska studentexamen bedömas automatiskt genom att de största utmaningarna – objektiv och tidseffektiv bedömning av ett stort antal talare – kan lösas. Dessa svårigheter har varit en viktig anledning till att muntliga test ännu inte införts i studentexamen i Finland (Undervisningsministeriet 2006), även om behovet av ett muntligt delprov erkänts (Undervisnings- och kulturministeriet 2017). Förutom att bedöma talet kan datorn också ge feedback till L2-talaren på de aspekter av talet som han/hon kan utvecklas i (se t.ex. Gu & Davis 2020), vilket möjliggör övande utanför undervisningskontexter.

I Finland arbetar forskningsprojektet DigiTala³ kring frågor om automatisk bedömning av språkfärdighet. Syftet med projektet är att utveckla en digital applikation för bedömning av uttal och muntlig färdighet i L2-svenska och L2-finska. Applikationen ska kunna användas i språktester såsom studentexamen, och den ska också kunna ge feedback till inläraren på dennes språkfärdighet och därmed bistå honom/henne vid inläringen (figur 1, se också Karhila m.fl. 2016).

³ DigiTala är ett forskningsprojekt mellan Helsingfors universitet, Aalto-universitetet och Jyväskylä universitet. Projektet finansieras av Finlands Akademi 2019–2023. Närmare om projektet se Kautonen & von Zansen 2020 och <https://www.helsinki.fi/sv/digitala>.



Figur 1. En beskrivning av hur automatisk taligenkänning ska utvecklas och är tänkt att användas för bedömning och feedback av L2-tal inom projektet DigiTala. Förkortningen GERS i fas 2 syftar på färdighetsnivåerna enligt den gemensamma europeiska referensramen för språk (Europarådet 2001, 2018). Föreliggande undersökning bidrar till att skapa kunskap om den kursiverade delen till höger i figuren.

En av utmaningarna vid automatisk bedömning av tal är att det saknas kvantitativ kunskap, dvs. kunskap om förekomstfrekvens, om fonetiska L2-drag i segment (vokaler och konsonanter) och prosodi (ordbetoning, satsbetoning [rytm] och satsintonation) hos talare på olika färdighetsnivåer. Kvantitativ kunskap behövs, eftersom den hjälper oss att identifiera och beskriva de fonetiska drag (eng. *phonetic features*) som datorn kan basera sin bedömning av talet på. Med hjälp av kunskap om de fonetiska dragen kan datorn läras att i sin bedömning imitera mänsklig bedömning av L2-tal. Kunskap om L2-dragen kan också behövas för att man ska kunna ge automatisk feedback till talaren på dennes prestation. Vad gäller feedbacken är en av svårigheterna att kunna leverera den i en sådan form att L2-talaren kan förstå och utnyttja den.

Vid automatisk bedömning är det relativt lätt att bedöma uttal i uppläst tal, medan bedömning av fritt (spontant) tal är svårare (Zechner m.fl. 2009). Det beror på att ASR har lättare att predicera vad talaren förväntas säga i uppläst

tal, och därmed är också bedömning av korrektheten i uttalet lättare. Bedömning av fritt tal är svårare också på grund av att fritt tal i högre grad präglas av situationell, individuell och regional variation, i synnerhet prosodiskt (van Santen m.fl. 2009). Ett sätt att komma till rätta med detta problem är att ASR bedömer talet baserat på möjligast robusta talegenskaper och jämför analysen gentemot en korpus som består av ett stort antal talprov bedömda på olika färdighetsnivåer. Då kan det ASR-baserade systemet söka likheter och skillnader i talproven på de olika färdighetsnivåerna. Det är allra svårast att bedöma prosodin i fritt tal, men inte heller det är numera omöjligt, eftersom många av de fonetiska egenskaper som kan användas vid bedömning av uppläst tal har visat sig fungera relativt väl också vid bedömning av fritt tal (van Santen m.fl. 2009, Chen m.fl. 2018).

Syftet med föreliggande undersökning är att analysera huruvida andelen segmentella och prosodiska L2-drag relaterar till färdighetsnivån. Därtill undersöks centrala aspekter av produktionsförmågan på färdighetsnivåerna: *ord per talprov*, *ord per yttrande*, *yttrandets komplexitet* och *taltempo*. Undersökningen kan bidra till utveckling av automatisk bedömning av L2-tal, i synnerhet bedömning av prosodin i fritt tal med svenska som målspråk. Resultaten ökar också vår förståelse av L2-talets fonetiska och strukturella uppbyggnad och dessas relation till bedömning av L2-talet. Därmed kan studien bidra till utveckling av kriterierna för bedömning av L2-tal. Vidare avser undersökningen att ge insikter i möjliga inlärningsgångar vid L2-uttalsinläring genom att visa på typiska uttalsdrag i olika stadier av inläringen.

2 Tidigare studier om L2-tal på olika färdighetsnivåer

Vid bedömning av språkfärdighet utgår många testarrangörer och utbildare numera från färdighetsnivåtänkande i stället för nationella betyg och vitsord. Den gemensamma europeiska referensramen för språk (GERS⁴; Europarådet 2001, 2018) är ett av de mest använda ramverken som också de finländska läroplanerna baserar sig på. Nivåerna i GERS beskriver utvecklingen från nybörjare (nivåerna A1–A2) och självständig språkanvändare (nivåerna B1–B2) till avancerad språkanvändare (nivåerna C1–C2). I den uppdaterade referensramen från 2018 beskrivs uttalsfärdigheterna närmare i en separat matris för fonologisk kontroll, vilket preciserar den rätt bristfälliga beskrivningen av uttalet i den tidigare referensramen. Nyckelbegreppen vad gäller

⁴ Också beteckningarna CEF och CEFR används för den gemensamma europeiska referensramen för språk.

uttalsfärdigheterna är begriplighet (hur lätt det är för samtalspartnern att förstå talaren), graden av påverkan från andra språk, uttalet av enskilda ljud (hur noggrant/med hur bra precision talaren kan uttala enskilda ljud) och kontroll över prosodiska drag (i vilken grad talaren kan använda betoning, rytm och intonation för att uttrycka finare betydelsenyanser). Kriterierna är inte språk-specifika utan de är gemensamma för flera språk, vilket motiverar en analys av hur väl kriterierna beskriver uttalet på olika färdighetsnivåer i olika målspråk.

Baserat på tidigare forskning har vi rätt bra kunskap om hurdana uttalsdrag som är viktiga för muntlig färdighet, upplevelse av främmande accent och begriplighet av talet (om L2-svenska se t.ex. Kautonen 2019, Kuronen & Zetterholm 2017 och Heinonen 2020). Däremot har vi bristande kunskap om förekomstfrekvensen av fonetiska L2-drag på olika färdighetsnivåer. Det är med andra ord ovisst om L2-dragen minskar linjärt och om det i så fall gäller alla drag, eller om vissa av L2-dragen kan vara mer frekventa och/eller av fler typer på de högre än på de lägre färdighetsnivåerna (vilket t.ex. skulle kunna bero på U-formad inlärning som kan påverka frekvensen och typen av L2-drag i tillägnandefasen av ett andraspråk, jfr Abrahamsson 2009:272).

Uttal på olika färdighetsnivåer har analyserats i L2-finlandssvenska av Kautonen (2019). Resultaten visar förbättring av uttalet från A1 till C2 i såväl segment, ord- som satsprosodi. Bortsett från den här tendensen verkar svårigheterna i uttalet vara likadana på de olika färdighetsnivåerna – det tycks således snarare vara fråga om en gradskillnad än en artskillnad mellan nivåerna. Ljud som inte förekommer i talarens L1 eller som skrivs på ett annat sätt i talarens L1 (jfr ortografisk interferens) vållar problem på alla nivåer, likaså målspråklig satsbetoning och satsintonation. På de högre nivåerna behärskar talarna dock de satsprosodiska dragen bättre än på de lägre nivåerna: betonade ord tycks inte vara lika frekventa på de högre nivåerna, vilket också leder till en mer målspråklig satsintonation, eftersom ett mindre antal betonade ord oftast resulterar i ett mindre antal intonationstoppar.

Kautonens (2019) undersökning baserar sig i huvudsak på lyssnarvärderingar och lyssnarnas upplevelse av uttalet, och därför ger den relativt lite kvantitativ kunskap om skillnaderna mellan färdighetsnivåerna. Några studier visar emellertid att andelen betonade ord tenderar att vara mindre på de högre nivåerna. Andelen betonade ord verkar minska från nivån B1 till B2 och från nivån C1 till C2 (Kang & Yan 2018 och Kang 2013 om L2-engelska), och likaså tenderar flertoppiga intonationskonturer att vara mindre frekventa på nivån B2 än på nivån B1 (Kautonen 2017 om L2-svenska). Också finlandssvenska L1-talare och finskspråkiga L2-talare av svenska på nivån C1 hade tydligare skillnader mellan betonade och obetonade stavelser än L2-talare på de lägre nivåerna (Kallio m.fl. 2020). Vidare kom Kallio m.fl. (2017) fram till

att finskspråkiga L2-talare av svenska behärskar drag som taltempo och pauser bättre på de högre färdighetsnivåerna; snabbare taltempo och en lägre andel fyllda pauser indikerar en högre färdighetsnivå (för liknande resultat om taltempo se Hildén 2000). Också studier om L2-engelska visar att talare på de högre färdighetsnivåerna tenderar att ha snabbare taltempo och kortare och färre pauser (Kang & Yan 2018, Kang 2013). Taltempo och pausering anses ofta höra till flyt, men de är nära kopplade till uttal genom att flyt påverkar rytm och intonation.

Sammantaget visar tidigare forskning att uttalet generellt är bättre på högre än på lägre nivåer av muntlig färdighet, men en heltäckande, kvantitativ analys av segmentella och prosodiska drag i L2-svenska på alla färdighetsnivåer saknas. Bortsett från studier med engelska som målspråk finns det få undersökningar om frågan.

3 Mål, material och metod

Mot bakgrund av det beskrivna forskningsläget undersöker vi i föreliggande studie andelen (= förekomstfrekvensen) av fonetiska L2-drag i svenska på färdighetsnivåerna A1–C2. Det primära målet med studien är att kunna besvara forskningsfrågan huruvida andelen segmentella och prosodiska L2-drag relaterar till färdighetsnivån. Därtill undersöks produktionsförmågan hos talare på olika färdighetsnivåer med hjälp av följande parametrar: *ord per talprov*, *ord per yttrande*, *andelen bisatser av alla satser*, *flerstaviga ord per talprov* och *taltempo*. Produktionsförmågan analyseras, eftersom den kan hänga ihop med andelen fonetiska L2-drag på färdighetsnivån. Studien är en fördjupning av Kautonen (2019), där finskspråkiga talares uttal av finlandssvenska i huvudsak analyserades kvalitativt med hjälp av lyssnaruppfattningar om uttalet.

Talmaterialet kommer från två språktester: (i) ett utvärderingstest av finskspråkiga grundskoleelevers inlärningsresultat i främmande språk och i det andra inhemska språket svenska från år 2013 som gjorts av Utbildningsstyrelsen i Finland (Hildén & Rautopuro 2014) och (ii) de allmänna språkexamina i Finland (Utbildningsstyrelsen 2011). Talarna i utvärderingstestet är finskspråkiga skolelever i årskurs 9 i grundskolan (i 15–16 års ålder) och talarna i den allmänna språkexamen i svenska är vuxna talare, i vårt material 27–49 år. Skoleleverna studerade lång lärokurs i svenska (A-språk), dvs. den lärokurs som börjar i årskurs 1–6 i den grundläggande utbildningen, medan talarna i den allmänna språkexamen hade studerat svenska i formell utbildning mellan fyra och cirka tio år. Majoriteten av de vuxna talarna hade utbildning från universitet, medan andra hade studerat i yrkesskola eller på

yrkeshögskola. Vi analyserar emellertid inte talarnas tal i förhållande till utbildningsnivån eller åldern.

Talarna har tilldelats färdighetsnivåerna A1–C2; grundskoleeleverna enligt lärarens bedömning⁵ och talarna i den allmänna språkexamen enligt bedömningar gjorda av testarrangörens egna bedömare som utbildats för uppgiften. Bedömningen baserar sig på generell muntlig färdighet som enligt kriterierna i båda testerna formas av ordförråd, flyt, strukturer och uttal. Eftersom bedömningarna kommer från autentiska språktester kan de anses ge tillförlitlig information om talarnas färdighetsnivåer.

Talarmaterialet består av 24 talare (16 kvinnliga talare och 8 manliga talare) så att varje färdighetsnivå representeras av fyra talare⁶. Talarna på nivåerna A1–B1 är elever som deltog i utvärderingstestet, medan talarna på nivåerna B2–C2 deltog i den allmänna språkexamen. Det finns 2–3 kvinnliga talare och 1–2 manliga talare på varje färdighetsnivå, bortsett från C2 där alla fyra talare är kvinnor. Talarna har ett uttal som enligt författarnas bedömning närmast kan beskrivas som finlandssvenska, och deras uttal analyserades därmed i förhållande till uttalet i finlandssvenska. Finlandssvenska är en varietet av svenska som i viss mån skiljer sig från sverigesvenska både vad gäller prosodin (främst rytm, ord- och satsintonation) och uttalet av enskilda ljud (Kuronen & Leinonen 2010).

Talmaterialet är en monolog där talarna talar fritt om ett givet tema. Temana varierar beroende på talarens färdighetsnivå: grundskoleeleverna fick berätta om sig själva och sin närmiljö, medan de vuxna talarna hade lite mer krävande teman. Eftersom fokus i analysen ligger på talarnas uttalsfärdigheter och allmänna produktionsförmåga, påverkar de olika temana inte analysen. Längden på uppgiften var högst 1–2 minuter, och talprov på 40–60 sekunder analyseras här. Grundskoleelevernas prestationer var så pass korta att de analyseras i sin helhet, medan ett utdrag på 50–60 sekunder från mitten av testet analyseras hos talarna i den allmänna språkexamen för att motsvara grundskoleelevernas prestationer. Ett exempel på ett talprov på färdighetsnivån A2 ges nedan. Inom parentes anges innehåll som utelämnades i analysen och/eller forskarnas kommentarer till stället med kursiv stil.

⁵ En tiondel av alla prestationer i testet bedömdes också av en censor vid Utbildningsstyrelsen för att kontrollera reliabiliteten av bedömningarna.

⁶ Dessa talare har analyserats tidigare med en annan metod i Kautonen (2019). Talarna valdes så att de kan anses representera en typisk prestation på färdighetsnivån vad gäller innehåll, och därför kan antalet talare anses vara tillräckligt stort för syftet i undersökningen.

Exempel på ett talprov på nivån A2:

Hej jag heter (namn). Jag är femton år. Jag bor i (namn). Jag har familj pappa mamma och storasyster. Min storasyster är åttande år. Jag går till (namnet på skolan). Och vi (*otydligt tal*) och. (*ett ofullständigt yttrande*) Ofta på sommarlovet vi går till (exem-; *omtagning*) exempel till Italia (or; *ord på ett annat språk*) någon var är varmt. Och jag tycker jag tycker. (*ett ofullständigt yttrande*) Jag gillar musik och jag spelar trummor. Och jag också spelar tennis. Jag har spelat tennis för fem år.

I analysen delades varje talprov (dvs. varje talares monolog) i yttranden. Vi definierade ett yttrande som en syntaktisk helhet som ofta är omgiven av längre pauser och som förmedlar en tanke, en åsikt eller en kommentar som är semantiskt koherent, till exempel < jag har ett sommarjobb >, < också vi har planerat att resa till Sverige på sommaren > samt < och inte tycker jag att det var något något problem >⁷. I yttrandeanalysen betraktades enskilda ord som användes för talplanering liksom < och > och < men > inte som yttranden. Exempel på längre fraser som räknades som talplanering är < och vi och >, < att att det är > och < jag tycker jag tycker >. I det ovanstående exemplet anger punkterna i transkriptionen yttrandegränserna. Kommentaren ”ett ofullständigt yttrande” medför att yttrandet inte beaktades i analysen.

Talproven analyserades auditivt av de två författarna, vilket skedde i följande steg. Först transkriberade vi talproven med hjälp av programmet Praat (Boersma & Weenink 2019). Därefter genomförde vi en gemensam analys av *en* talares talprov, varvid vi diskuterade de fonetiska analyskategorierna (se nedan), bedömde kategoriernas analyserbarhet och kontrollerade att vi uppfattade kategorierna på samma sätt. Resten av talarna analyserades i enskilda sessioner så att vi inte påverkades av varandras analyser. Vi lyssnade på varje yttrande i varje talprov 4–5 gånger, vid behov i mindre delar än som ett helt yttrande. Efter detta jämförde vi våra analyser. Om analyserna skilde sig åt, lyssnade vi på nytt på yttrandet för att nå en samsyn av analysen. Skillnaderna gällde en minoritet av yttrandena, och också i de yttranden där skillnader fanns gällde de en minoritet av de fonetiska kategorierna. Vi var endast två personer som lyssnade på yttrandena, eftersom uppgiften krävde expertis på fonetik och finlandssvenskt uttal. Att vi som lyssnare var så pass eniga om analyserna som vi var tyder på att utfallet av analysen kan anses tillförlitligt också med bara två lyssnare.

⁷ Vi använder vinkelparentes < > för grafem, snedstreck // för fonem och hakparentes [] för foner och allofoner. Vid fonetisk skrift används tecken från det internationella fonetiska alfabetet IPA (2015). Vid grafematisk skrift anges betoning med understrykning: < jag har en hund >.

Vid lyssnandet identifierade vi de delar som talarna uttalade på ett från målspråket⁸ avvikande sätt gällande *segment*, *segmentlängder*, *ordbetoning* och *satsbetoning*. Avvikelserna kategoriserades i olika typer enligt följande kriterier (i stort i enlighet med Isaacs & Trofimovich 2012, se också Trofimovich & Isaacs 2012 och Munro & Derwing 1995):

- **Segment.** Som segmentella L2-drag räknades tre typer: (i) reduktion, dvs. bortfall av segment (< strand > uttalas ['srand]), (ii) epentes, dvs. tillägg av segment (< stol > uttalas ['stu:lɑ]) och (iii) fonemiska avvikelser, dvs. att talaren använder ett felaktigt fonem (< bil > uttalas ['pi:l] eller < hund > uttalas ['hund] eller ['hynd]). Som fonemiska avvikelser betraktades uttal som avviker från målspråket med minst ett tecken i IPA, medan mindre, svårigen hörbara avvikelser bortsågs från; t.ex. ett möjligt L2-drag i slutenhetsgraden av [e:] eller [y:]. Det sistnämnda draget påverkar inte begripligheten, även om det kan bidra till förnimmelse av främmande accent i finskspråkiga talares uttal av finlandssvenska. För att beskriva segmentuttalet generellt analyserade vi andelen segmentella L2-drag per uttalat ord hos varje talare.
- **Segmentlängder.** Som L2-drag räknades felaktig längd, dvs. ett kort ljud uttalas som långt (< papper > uttalas [pa 'pe:r]) eller ett långt ljud uttalas som kort (< ljus > uttalas [jus]). För att beskriva segmentlängderna generellt analyserade vi andelen L2-drag i segmentlängder per uttalat ord hos varje talare.
- **Ordbetoning.** Som L2-drag betraktades (i) avsaknad av betoning, (ii) för många betonade stavelser, (iii) betoningen ligger på fel stavelse i ordet eller (iv) betoningen produceras för svagt eller kraftigt. För att beskriva ordbetoningen generellt analyserade vi andelen L2-drag i ordbetoning per uttalade flerstaviga ord hos alla talare.
- **Satsbetoning** (distribution av betonade och obetonade stavelser i yttrandet, dvs. rytmen i yttrandet). Som L2-drag betraktades (i) avsaknad av betoning, (ii) för många ord betonas i satsen, (iii) betoningen ligger på fel ord i satsen eller (iv) betoningen produceras för svagt eller kraftigt (jfr kategorierna i Heinonen & Kautonen 2020). För att beskriva satsbetoning generellt analyserade vi andelen yttranden med L2-satsbetoning av alla yttranden hos varje talare. Därtill

⁸ Vi analyserar inlärarnas uttal i förhållande till finlandssvenskt standarduttal. Standarduttalet i finlandssvenska är något mer svårdefinierat än standarduttalet i sverigesvenska (Ivars 2015, Bruce 2010), men också finlandssvenska anses ha ett överregionalt uttal som kan kallas standarduttal eller standardspråkligt uttal (Reuter 2015). Uttalet i Mellannyländ anses ha flest likheter med standarduttalet (Ivars 2015:354).

analyserade vi andelen betonade ord av alla uttalade ord hos varje talare (jfr Kang 2013).

Betoningens fonetiska korrelat i svenska är längd, intonation och intensitet (Leinonen m.fl. 1990). I vår studie baserar sig analysen av ord- och satsbetoning på förnimmelsen av betoning som i sin tur grundar sig på en kombination av korrelaten ökad längd, högre intonation och större intensitet i den betonade stavelsen i förhållande till de omgivande, obetonade stavelserna. Kriteriet för förekomst respektive avsaknad av betoning var med andra ord auditiv. Frågan vi ställde oss var om yttrandet i sin kontext kunde ha uttalats med detta betoningsmönster vid L1-uttal. I de flesta fall finns inte bara ett enda möjligt uttal av ett yttrande utan yttrandet kan uttalas på olika sätt i målspråket. Därför accepterade vi alla sådana uttal som vi bedömde som möjliga i målspråket. Att analysera yttrandets betoningsmönster i fritt tal är med nödvändighet i viss mån subjektivt, men analysen här kan motiveras med att vi förberedde den noga med en pilotstudie och vi fokuserade inte på fina nyanser av yttrandets rytm utan använde oss av relativt lättidentifierbara rytmiska L2-drag. Vid analys av betoning opererade vi med en binär kategori betonat–obetonat, även om också bibetonade stavelser kan påverka förnimmelsen av rytm. Att bedöma olika grader av bibetoning har dock den nackdelen att analysen kan leda till fler och mer subjektiva värderingar än analys med en binär kategori.

Grammatiska fel såsom fel artikel (< en > i stället för < ett >), fel pronomen (< hon > i stället för < han >) och fel i kongruens (< vår > i stället för < vårt >) betraktades inte som L2-uttalsdrag såvida det inte ansågs motiverat av något speciellt skäl. Däremot betraktades bortfall av presensändelse som uttalsfel, även om också det kan orsakas av morfologiska eller lexikala svårigheter. Denna typ av gränsfall kommenteras närmare i avsnitt 4.2 som handlar om segmentella L2-drag.

Produktionsförmågan och komplexiteten analyserades med hjälp av parametrarna *ord per talprov*, *ord per yttrande*, *andelen bisatser av alla satser*⁹, *flerstaviga ord per talprov* och *taltempo* (jfr Isaacs & Trofimovich 2012). Eftersom talproven inte har samma genomsnittliga längd på de olika

⁹ Vi har definierat sats som en grammatisk enhet vars konstitutiva led är subjekt och predikatsled (jfr Svenska Akademiens grammatik, SAG 1999, Del IV:3). I talspråk, i synnerhet i dialoger, är det inte alltid lätt att kategorisera satser som huvud- eller bisatser, eftersom talaren kan stanna av mitt i en sats och övergå till en annan struktur eller bli avbruten av samtalspartnern. I vårt material var emellertid en klar majoritet av satserna relativt entydigt definierbara som huvud- eller bisatser. Ett yttrande i materialet kan bestå av (flera) huvud- och/eller bisatser, t.ex. < när du kommer här vi ska gå och simma på sommaren >.

färdighetsnivåerna (se avsnitt 4.1), kan ord per talprov och flerstaviga ord per talprov inte jämföras direkt mellan nivåerna. Vi räknade ändå båda parametrarna därför att (i) vi anser att jämförelsen de facto inte är missvisande, eftersom talarna på nivåerna A1–A2 inte kunde producera mer tal om det givna temat än de 41–47 sekunder som de genomsnittligt gjorde, (ii) de förhållandevis små skillnaderna i längden på talprov mellan färdighetsnivåerna (jfr avsnitt 4.1., tabell 1) inte ens tillnärmelsevis förklarar de stora skillnader i ord per talprov och flerstaviga ord per talprov som finns mellan nivåerna samt (iii) parametrarna ord per talprov och flerstaviga ord per talprov korrelerar med de tre övriga parametrarna för produktionsförmågan (se avsnitt 4.1). Namn, nonsensord och ord på andra språk än svenska beaktades inte i det totala antalet ord och inte heller i andelen betonade ord av alla uttalade ord (jfr beskrivningen av satsbetoning ovan). Nonsensord och ord på andra språk än svenska förekom endast på nivåerna A1–A2. Närmare analys av ordförrådet föll utanför ramen för den här studien, även om ordförrådet i viss mån kommenteras i analysen.

Taltempo analyserades genom parametern ord per sekund pauser medräknade. Däremot analyserade vi inte paustyper och pauserinsgmönster, även om dessa kan vara viktiga för upplevelse av rytm. Pauser har emellertid studerats i L2-svenska på olika färdighetsnivåer av Kallio m.fl. (2017) och analyserades därför inte här. Vidare kan pauser snarare anses tillhöra flyt än uttal (se t.ex. Hsieh, Zechner & Xi 2020, Kallio m.fl. 2017, Derwing m.fl. 2004). I perception av löpande tal är det svårt att skilja brister i flyt (t.ex. pauser) från brister i prosodin, men vi anser att det kan göras vid analytiskt lyssnande med fokus på prosodin såsom var fallet i vår analys.

Korrelationer mellan olika variabler i materialet räknades med Spearmans rangkorrelation som är avsedd för kontinuerliga variabler i non-parametriska data som inte förutsätter att variablerna är normalfördelade. Korrelationerna $r = .50-1.0$ och $r = -.50-(-1.0)$ anses vara starka, korrelationerna $r = .30-.49$ och $r = -.30-(-.4.9)$ är medelstarka och korrelationerna $r = .10-.29$ och $r = -.10-(-.29)$ är svaga (Pallant 2005:126). Signifikansnivån i studien är $p \leq 0.01$, dvs. skillnaden är signifikant på 1 %-nivån.

4 Resultat

I detta kapitel presenteras först generella aspekter av talproduktion på de olika färdighetsnivåerna (avsnitt 4.1). Också de sammantagna resultaten om fonetiska L2-drag presenteras. Därefter behandlar vi mer detaljerat segmentella L2-drag (avsnitt 4.2) och prosodiska L2-drag (avsnitt 4.3).

4.1 Talet på olika färdighetsnivåer: talproduktion och förekomst av L2-drag i uttalet

Talarna talar mer (räknat i antalet ord per talprov) och har fler ord per yttrande på de högre än på de lägre färdighetsnivåerna (tabell 1). Antalet ord per talprov ökar linjärt från A1 till C2, och likaså ökar taltempot närmast linjärt från A1 till C2¹⁰ (tabell 1). Däremot ökar antalet ord per yttrande inte linjärt, eftersom det finns en tydlig skillnad mellan nivåerna A1–B1 och nivåerna B2–C2: antalet ord per yttrande är 2–3 gånger större på nivåerna B2–C2 än på nivåerna A1–B1. Skillnaden i antalet ord per yttrande mellan B1 och B2 är således en vattendelare i materialet. Antalet flerstaviga ord per talprov är större på nivåerna B2–C2 än på nivåerna A1–B1, och också andelen bisatser av alla satser är högre på nivåerna B2–C2 än på nivåerna A1–B1 (tabell 1). På nivåerna A1–B1 använder talarna nästan enbart huvudsatser.

¹⁰ Taltempo varierar situationellt och individuellt, och därför är det svårt att jämföra olika studier med annorlunda talmaterial. Det taltempo som talarna på nivån C2 har i vår studie (2,37 ord per sekund; tabell 1) kan ändå anses vara nära tempot i L1-tal. I sverigesvenska var taltempot vid nyhetsuppläsning genomsnittligt 2,79 ord per sekund (pauser längre än 150 millisekunder inte medräknade, 18 informanter) enligt Hilton, Gooskens och Schüppert (2011:211). Detta tempo är ca 20 % högre än talarna på nivån C2 har här, men eftersom Hilton m.fl. utelämnade pauser längre än 150 millisekunder, kan ett tempo på 2,37 ord per sekund anses vara nära tempot hos L1-talarna i Hilton m.fl. (2011).

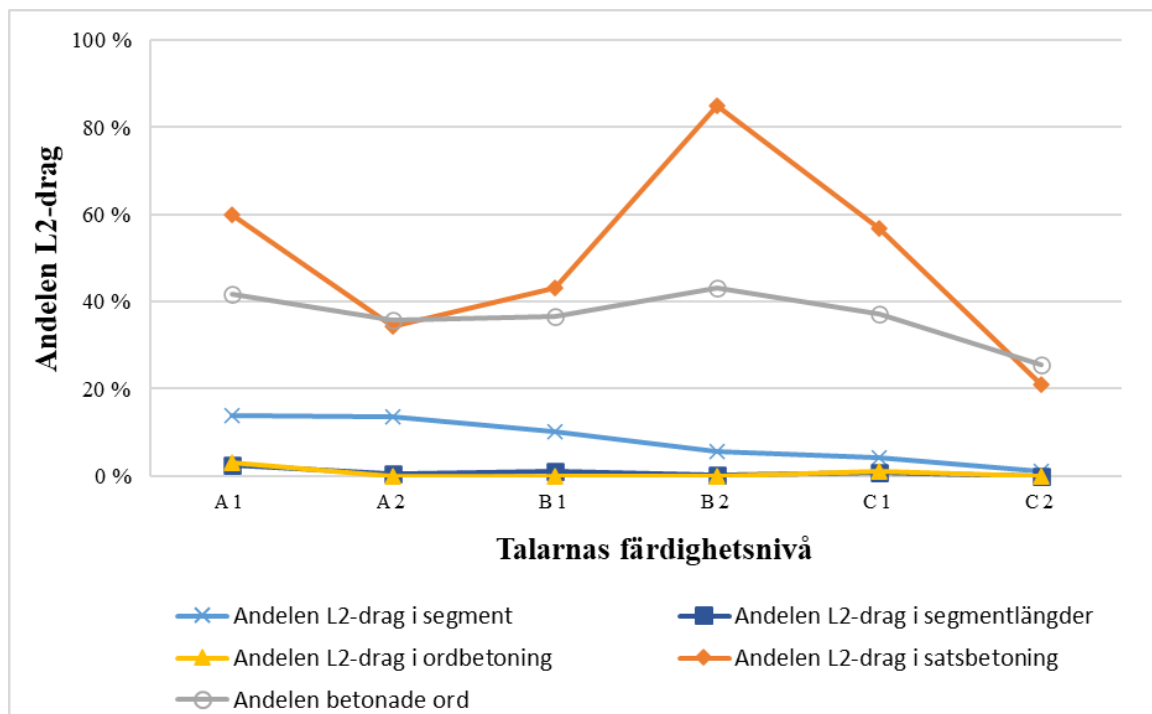
Tabell 1. Genomsnittlig(t) längd av talprovet, taltempo, antalet ord per talprov, antalet flerstaviga ord per talprov, antalet yttranden per talprov, antalet ord per yttrande samt andelen bisatser av alla satser på färdighetsnivåerna A1–C2. Variationen mellan talare på samma färdighetsnivå anges inom parentes.

Nivå	Längden av talprovet (sek.)	Taltempo (ord per sek. pauser medräknade)	Antalet ord per talprov	Antalet flerstaviga ord per talprov	Antalet yttranden per talprov	Antalet ord per yttrande	Andelen bisatser av alla satser
A1	41 (35–48)	0,77 (0,42–1,00)	31 (20–38)	8 (5–11)	6 (4–8)	4,9 (4,8–5,0)	0 % (0 %)
A2	47 (30–58)	1,24 (1,10–1,50)	57 (44–72)	21 (16–28)	10 (6–13)	6,0 (4,9–7,5)	2 % (0–13 %)
B1	55 (47–61)	1,24 (0,98–1,67)	67 (57–87)	21 (14–27)	11 (10–13)	6,1 (5,7–6,7)	8 % (0–19 %)
B2	60 (58–61)	1,65 (1,49–1,86)	98 (91–108)	43 (38–47)	8 (7–10)	11,9 (10,1–15,4)	46 % (33–50 %)
C1	61 (59–64)	1,80 (1,57–2,10)	110 (94–128)	44 (38–50)	9 (6–12)	11,8 (10,7–15,7)	47 % (40–54 %)
C2	60 (58–61)	2,37 (2,16–2,84)	141 (131–165)	53 (49–57)	10 (8–12)	14,8 (10,9–17,0)	43 % (33–57 %)

Färdighetsnivån korrelerar starkt med antalet ord per talprov ($r = .957$, $p \leq 0.01$), antalet flerstaviga ord per talprov ($r = .926$, $p \leq 0.01$), antalet ord per yttrande ($r = .912$, $p \leq 0.01$), andelen bisatser av alla satser ($r = .814$, $p \leq 0.01$) samt taltempo ($r = .903$, $p \leq 0.01$). Parametrarna korrelerar också starkt med varandra: antalet ord per talprov med taltempo ($r = .958$, $p \leq 0.01$), antalet flerstaviga ord per talprov med taltempo ($r = .889$, $p \leq 0.01$), antalet ord per talprov med antalet ord per yttrande ($r = .901$, $p \leq 0.01$) och antalet ord per yttrande med taltempo ($r = .865$, $p \leq 0.01$). Komplexiteten i strukturerna kan påverka uttalet, främst satsbetoningen, vilket vi återkommer till senare i avsnittet. Resultaten tyder sammantaget på att talarna har tilldelats rätt färdighetsnivå, eftersom produktionsförmågan (antalet ord per talprov och taltempo) ökar linjärt från lägre till högre nivå och eftersom komplexiteten (antalet ord per yttrande, andelen bisatser av alla satser och antalet flerstaviga ord per talprov) är störst hos talarna på nivåerna B2–C2.

Andelen fonetiska L2-drag på nivåerna A1–C2 i segment, segmentlängder, ordbetoning och satsbetoning anges i figur 2 och tabell 2. Andelen L2-drag i segment minskar linjärt från A1 (14 %) till C2 (1 %) (figur 2, tabell 2).

Procentsiffrorna i tabell 2 betyder att ett yttrande på tio ord genomsnittligt innehåller 1,4 segmentfel på nivåerna A1 och A2, 0,6–1,0 segmentfel på nivåerna B1 och B2 samt 0,1–0,4 segmentfel på nivåerna C1 och C2. Skillnaderna mellan nivåerna låter kanske inte så stora, men om de upprepas systematiskt i yttrandena i talproven är deras effekt på lyssnaren sannolikt inte oväsentlig. Korrelationen mellan färdighetsnivån och andelen L2-drag i segment är stark ($r = -.783$, $p \leq 0.01$). Samtidigt är andelen L2-drag i segment inte speciellt stor på nivåerna A1–A2 heller, vilket kan bero på de relativt små fonetiska skillnaderna i segmenttuttatet mellan finska och finlandssvenska (Kuronen & Leinonen 2010, Kautonen 2019). Andelen L2-drag i segment skulle antagligen ha varit högre speciellt på nivåerna A1–A2, om målspråket t.ex. hade varit engelska, franska eller ryska.



Figur 2. Andelen fonetiska L2-drag genomsnittligt på färdighetsnivåerna A1–C2.

Tabell 2. Genomsnittlig andel fonetiska L2-drag på färdighetsnivåerna A1–C2. Variationen mellan talare på samma färdighetsnivå anges inom parentes.

Nivå	Segment	Segmentlängder	Ordbetoning	Satsbetoning	
	Andelen L2-drag per ord:	Andelen L2-drag per ord:	Andelen L2-drag per ord:	Andelen yttranden med L2-drag av alla yttranden:	Andelen betonade ord av alla uttalade ord:
A1	14 % (10–17 %)	2 % (0–6 %)	3 % (0–11 %)	60 % (38–86 %)	42 % (28–60 %)
A2	14 % (7–22 %)	0 % (0–1 %)	0 % (0 %)	34 % (31–40 %)	36 % (30–43 %)
B1	10 % (7–12 %)	1 % (0–4 %)	0 % (0 %)	43 % (27–70 %)	37 % (29–44 %)
B2	6 % (1–11 %)	0 % (0–1 %)	0 % (0 %)	85 % (71–100 %)	43 % (37–48 %)
C1	4 % (1–8 %)	1 % (0–4 %)	1 % (0–4 %)	57 % (33–78 %)	37 % (34–42 %)
C2	1 % (0–3 %)	0 % (0 %)	0 % (0 %)	21 % (10–33 %)	26 % (22–33 %)

Andelen L2-drag i ordprosodi (segmentlängder och ordbetoning) är mycket liten på alla färdighetsnivåer: genomsnittligt innehåller 0–3 % av alla flerstaviga ord L2-drag i ordbetoning, och 0–2 % av alla ord innehåller L2-drag i segmentlängder på de olika färdighetsnivåerna (figur 2, tabell 2). Med andra ord uttalar talarna i mycket hög grad de ord som förekommer i deras yttranden med rätt segmentlängd(er) och rätt ordbetoning, dvs. rätt betonad stavelse i ordet. Resultatet kan tyckas överraskande med tanke på att svenska har morfologiskt betingad ordbetoning och därmed skiljer sig från finska med fast initialbetoning. Låg andel L2-drag i ordprosodi kan bero på att talarna på de lägre nivåerna i sina yttranden har ord som tillhör de vanligaste i svenska och som är morfologiskt enkla (t.ex. < jag >, < har >, < är >, < gammal >, < år >, < heter >, < bor >, < mamma >, < pappa > och < hund >). Talarna på nivåerna B2–C2 har ett mer utvecklat ordförråd och deras yttranden innehåller oftare ord som är mindre frekventa och morfologiskt mer komplexa (t.ex. < specialaffär >, < ljustake >, < djupfryst >, < uppfostra >, < grymmare >, < oändlighet > och < effektivitet >). Denna typ av ord skulle säkerligen ha orsakat prosodiska problem för talarna på nivåerna A1–A2. Iakttagelsen gäller möjligen generellt ordförrådet i L2-tal: inlärare tenderar att använda sådana ord vars uttal de behärskar och undviker ord vars uttal de inte känner till eller är osäkra på, vilket leder till att L2-dragen i ordprosodin är färre än man kanske skulle förvänta sig baserat på L2-talarens färdighetsnivå.

Vad gäller satsbetoning visar analysen att andelen L2-drag inte minskar linjärt från lägre till högre nivå (figur 2, tabell 2). Endast på nivån C2 är andelen L2-drag i satsbetoning och andelen betonade ord av alla ord betydligt lägre än på de övriga nivåerna (figur 2, tabell 2). Det är inte helt lätt att förstå varför andelen L2-drag i satsbetoning inte är lägre på nivåerna B2–C1 än på nivåerna A1–B1. Kanske beror det på att talarna på nivåerna B2–C1

producerar mer tal och framför allt på att de har längre yttranden med större grammatisk komplexitet än talarna på nivåerna A1–B1. I längre yttranden är sannolikheten för L2-drag i satsbetoning större än i korta yttranden: om man t.ex. säger < jag har mamma, pappa och lillasyster > jämfört med < när du kommer här vi ska gå och simma på sommaren > (exemplen kommer från materialet) är risken för L2-drag i satsbetoning större i det senare yttrandet. Det finns en tydlig ökning av L2-drag i satsbetoning på nivån B2, och just på den nivån ökar också såväl förmågan att producera tal som komplexiteten i talet. Kanske hänger produktionsförmågan, komplexiteten och L2-dragen i satsbetoning ihop. Ökad andel L2-drag i satsbetoning på nivån B2 skulle därmed kunna förstås i ljuset av U-formad inläring av andraspråket enligt vilken antalet felaktiga former kan öka i tillägnandefasen (Abrahamsson 2009). Vad gäller L2-drag i satsbetoning ökar feltyperna inom kategorin inte per se i vårt material (vilket kan hända vid inläring av segment i L2; Kautonen & Kuronen, manus) utan samma feltyp, dvs. för många betonade stavelser i yttrandet, ökar i frekvens (se avsnitt 4.3).

Under analysen funderade vi mycket på om våra resultat om L2-drag i satsbetoning kan bero på perspektivfel, dvs. att fokus i analysen snarare legat på L2-dragen än det talarna kan uttala. Det är möjligt att det inom ett längre yttrande (som på de högsta nivåerna) fanns bra sekvenser i prosodin som sedan inte nödvändigtvis visade sig i analysen av hela yttrandet. Vi tror ändå inte att perspektivet förklarar resultaten, eftersom också andelen betonade ord av alla uttalade ord tydligt minskar först på nivån C2 (tabell 2). Andelen L2-drag i satsbetoning och andelen betonade ord av alla uttalade ord hänger ihop som fonetiska parametrar, eftersom det överlägset vanligaste L2-draget i satsbetoning på alla nivåer är för många betonade ord i yttrandet (se avsnitt 4.3, tabell 5). Andelen betonade ord av alla uttalade ord korrelerar också signifikant med andelen L2-drag i satsbetoning ($r = .667$, $p \leq 0.01$).

Sammantaget kan vi om satsbetoningen på färdighetsnivåerna konstatera att (i) den till skillnad från segmentuttalet inte förbättras linjärt, (ii) tvärtom tycks L2-dragen i satsbetoning öka då talproduktionen och dess komplexitet ökar samt (iii) satsbetoning tycks vara svårare att tillägna sig i andraspråket än segment – åtminstone i språkparet finska–finlandssvenska.

Variationen inom nivåerna är störst för andelen L2-drag i satsbetoning, vilket innebär att några talare mycket oftare hade svårigheter med satsbetoning än andra trots att de hade tilldelats samma färdighetsnivå (tabell 2). Detta kan bero på att flyt (främst pauser och korrigeringar), kvalitativa aspekter av ordförråd och kohesion inte beaktades i vår analys, men de kan ha påverkat bedömningen av färdighetsnivån i testet.

4.2 Andelen L2-drag i segment på färdighetsnivåerna

De flesta L2-drag i segment i orsakades av (i) fonetiska svårigheter att uttala segment främst beroende på inverkan från inlärnarnas L1, (ii) ortografisk interferens från L1 eller (iii) inflytande från andra L2-språk, främst engelska. Fonemiska avvikelser var den vanligaste typen av segmentella L2-drag på alla färdighetsnivåer (tabell 3). Med fonemisk avvikelse avses här att ett fonem i målspråket uttalas som ett annat och felaktigt ljud. Exempel på fonemiska avvikelser i materialet var:

- < bror > uttalades ['bro:r]
- < färg > uttalades ['vɛrjɛ] (också tillägg av fonem ordfinalt)
- < finns > uttalades ['vɪns]
- < beslut > uttalades [pɛ'slu:t] (två stycken fonemiska avvikelser) samt
- < njut > och < hund > uttalades ['nju:t] och ['hund].

I obetonade stavelser var fonemiska avvikelser inte lätta att höra (jfr < tycker > ['tyk:ɛr] och < heter > ['he:tɛr]), men om båda forskarna hörde en fonemisk avvikelse i obetonad stavelse, betraktades också den som avvikelse. En klar majoritet av avvikelserna gällde dock ljud i betonade stavelser.

Tabell 3. Andelen olika typer av segmentella L2-drag av alla segmentella L2-drag. Variationen mellan talare på samma färdighetsnivå anges inom parentes.

Nivå	Bortfall av fonem	Tillägg av fonem	Fonemisk avvikelse	Alla segmentella L2-drag sammantaget
	Andelen L2-drag:	Andelen L2-drag:	Andelen L2-drag:	Andelen L2-drag per ord:
A1	6 % (0–25 %)	29 % (0–100 %)	65 % (0–100 %)	14 % (10–17 %)
A2	10 % (0–14 %)	39 % (20–67 %)	52 % (33–80 %)	14 % (7–22 %)
B1	30 % (13–43 %)	11 % (0–25 %)	59 % (50–67 %)	10 % (7–12 %)
B2	5 % (0–11 %)	9 % (0–50 %)	86 % (50–100 %)	6 % (1–11 %)
C1	6 % (0–25 %)	0 % (0 %)	94 % (75–100 %)	4 % (1–8 %)
C2	17 % (0–50 %)	0 % (0 %)	83 % (0–100 %)	1 % (0–3 %)

Fonemtilläggen var få i materialet (t.ex. < är > uttalades ['ɛ:r]¹¹ och < universitet > uttalades [junɪvɛrsɪ'te:t]). När fonemtillägg förekom, gällde det mestadels nivåerna A1–A2 (tabell 3). På nivåerna C1–C2 förekom inga

¹¹ Vi anser att uttalet ['ɛ:r] har ett fonemtillägg /r/ jämfört med de målspråkliga uttalen ['ɛ:] eller [ɛ]. Uttalet ['ɛ:r] kunde också ses som ett vårdat, tydligt uttal av < är >, men då är vokalkvaliteten avvikande.

fonemtillägg. Vad gäller bortfall av fonem gällde majoriteten presensändelsen -r (< badar >, < bor >), vilket också kan förekomma i L1-tal (Wikner 2019). I de fall då bortfallet var tydligt hörbart betraktades det ändå som L2-drag. Det förekom inga bortfall av fonem i konsonantklustren, även om det uppgetts vara typiskt för finskspråkiga inlärare av svenska (Bannert 2004:117).

4.3 Andelen L2-drag i prosodin på färdighetsnivåerna

L2-dragen i segmentlängder och ordbetoning var mycket få i materialet. Inlärarnas L1 finska med ett utvecklat kvantitetssystem (jfr Suomi, Toivanen & Ylitalo 2008) kan ha bidragit till att segmentlängderna i svenska var lätta för dem. Det fanns sammanlagt 12 stycken L2-drag i segmentlängder. Båda typer av L2-drag i segmentlängder förekom, dvs. ett långt ljud uttalades kort och ett kort ljud uttalades långt (tabell 4). Det förstnämnda var något vanligare, t.ex. i ord som < alkohol >, < moped >, < fotboll > och < modern > (adj.) där de stavelser som ska ha betoning uttalades kort. Korta ljud uttalades också ibland som långa: < rum > och < alkoholist > kunde uttalas [ˈru:m] och [alkohɔˈli:st].

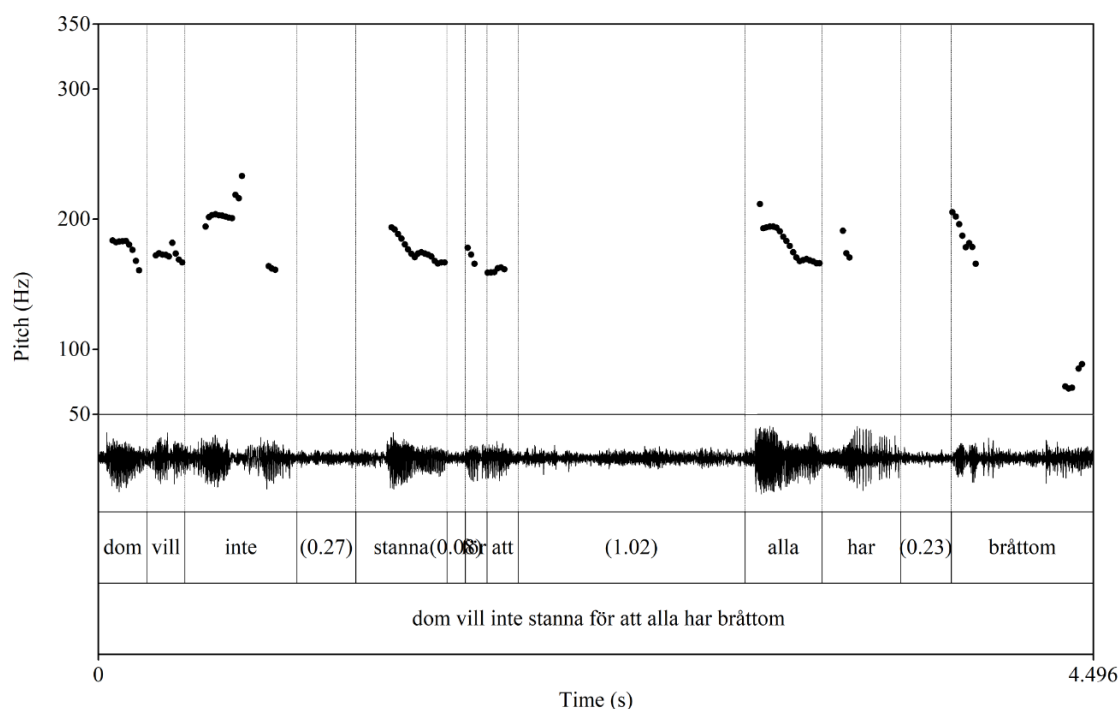
Tabell 4. Andelen olika typer av L2-drag i segmentlängder av alla L2-drag i segmentlängder. Variationen mellan talare på samma färdighetsnivå anges inom parentes.

Nivå	Kort ljud blir långt	Långt ljud blir kort	Alla L2-drag i segmentlängder
	Andelen L2-drag:	Andelen L2-drag:	Andelen L2-drag per ord:
A1	0 % (0 %)	100 % (0–100 %)	2 % (0–6 %)
A2	0 % (0 %)	100 % (0–100 %)	0 % (0–1 %)
B1	67 % (0–100 %)	33 % (0–50 %)	1 % (0–4 %)
B2	0 % (0 %)	100 % (0–100 %)	0 % (0–1 %)
C1	25 % (0–25 %)	75 % (0–75 %)	1 % (0–4 %)
C2	0 % (0 %)	0 % (0 %)	0 % (0 %)

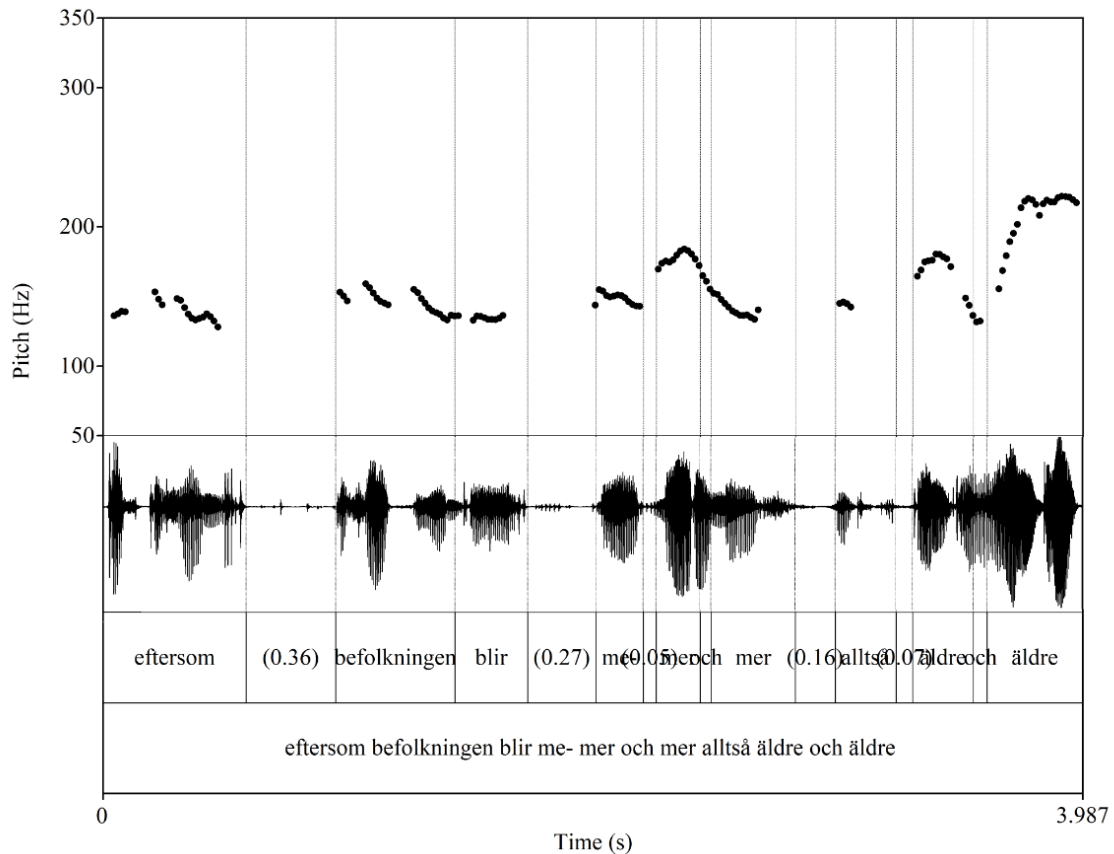
Vad gäller ordbetoning är resultaten i linje med segmentlängder: L2-dragen var mycket få, sammanlagt fem stycken. Detta är väntat eftersom längd är en viktig parameter för ordbetoning, och därmed hänger betoning och segmentlängder ihop också vid inläringen. På nivån A1 saknades ordbetoning en gång (< moped >), medan betoningen var på fel stavelse på nivån C1 (2 st., < alkoholist > och < speciellt >). Talarna hade dock ibland relativt långa pauser mellan leden i sammansatta ord, vilket kan påverka förnimmelsen av

ordbetoning, men eftersom pauser inte beaktades i vår analys framgår detta L2-drag inte av analysen.

Den vanligaste typen av L2-drag i satsbetoning var *för många betoningar*: typen utgjorde 86–100 % av alla L2-drag i satsbetoning på färdighetsnivåerna (tabell 5). Typiskt för draget är att talarna betonade ord som normalt inte får betoning i L1-uttal: < jag har en hund >, < och jag ska också hänga med mina kompisarna > och < dom vill inte stanna för att alla har bråttom > (figur 3). Även om andelen betonade ord av alla uttalade ord minskar något på de högsta nivåerna, korrelerar parametern inte direkt med målspråklig satsbetoning. Till exempel i yttranden där någonting räknas upp är det vanligt även i L1-tal att ha flera betonade ord: < jag bor med min mamma, pappa och lillasyster och en hund >, < min familj är mamma, pappa, storbror och storsyster > och < eftersom befolkningen blir mer och mer, alltså äldre och äldre > (figur 4). Därför måste man också beakta vad som sägs i vilken kontext, vilket vi gjorde i analysen.



Figur 3. Exempel på ett yttrande som ansågs ha för många betonade ord. Yttrandet är < dom vill inte stanna för att alla har bråttom > med betoning i orden < dom >, < inte >, < stanna >, < alla > och < bråttom >. Betoningen syns i analysen som lång relativ duration och/eller stigning av grundtonen i de betonade stavelserna. Analysen har gjorts med Praat (Boersma & Weenink 2019). Grundtonskurvan (f0-kurvan) har kontrollerats manuellt.



Figur 4. Exempel på ett yttrande som inte ansågs ha några L2-drag i satsbetoning trots att yttrandet innehöll flera betonade ord. Yttrandet är < eftersom befolkningen blir mer och mer alltså äldre och äldre > med betoning i orden < befolkningen >, < mer >, < mer >, < äldre > och < äldre >. Analysen har gjorts med Praat (Boersma & Weenink 2019). Grundtonskurvan (f0-kurvan) har kontrollerats manuellt.

L2-draget *betoning saknas* förekom endast på nivån A1 i två yttranden (< hej jag heter xx > och < jag bor i xx >). Också L2-drag av typerna *betoningen ligger på fel ord* och *betoningen produceras för kraftigt eller svagt* förekom sällan, mest på nivåerna B2–C2 (tabell 5). De sistnämnda L2-dragen kan överlappa med L2-draget *för mycket betoning*: om för många ord betonas, ligger åtminstone någon av betoningarna på ett i kontexten oväntat ord och/eller betoning på dessa ord markeras för kraftigt.

Tabell 5. Andelen olika typer av L2-drag i satsbetoning av alla L2-drag i satsbetoning samt andelen betonade ord av alla uttalade ord. Variationen mellan talare på samma färdighetsnivå anges inom parentes.

Nivå	Betoning saknas	För många betoningar	Betoningen ligger på fel ord	Betoningen produceras för kraftigt/svagt	Alla L2-drag i satsbetoning	
					Andelen yttranden med L2-drag i av alla yttranden:	Andelen betonade ord av alla uttalade ord:
A1	13 % (0–33 %)	87 % (67–100 %)	0 % (0 %)	0 % (0 %)	60 % (38–86 %)	42 % (28–60 %)
A2	0 % (0 %)	100 % (100 %)	0 % (0 %)	0 % (0 %)	34 % (31–40 %)	36 % (30–43 %)
B1	0 % (0 %)	100 % (100 %)	0 % (0 %)	0 % (0 %)	43 % (27–70 %)	37 % (29–44 %)
B2	0 % (0 %)	96 % (88–100 %)	0 % (0 %)	4 % (0–13 %)	85 % (71–100 %)	43 % (37–48 %)
C1	0 % (0 %)	86 % (50–100 %)	5 % (0–17 %)	10 % (0–33 %)	57 % (33–78 %)	37 % (34–42 %)
C2	0 % (0 %)	100 % (100 %)	0 % (0 %)	0 % (0 %)	21 % (10–33 %)	26 % (22–33 %)

5 Sammanfattande diskussion med kritiska anmärkningar

Syftet med föreliggande undersökning var att besvara frågan huruvida andelen segmentella och prosodiska L2-drag relaterar till färdighetsnivån. Också centrala aspekter av produktionsförmågan på färdighetsnivåerna analyserades: *ord per talprov*, *ord per yttrande*, *andelen bisatser av alla satser*, *flerstaviga ord per talprov* och *taltempo*. Materialet bestod av 24 talare, fyra på varje färdighetsnivå (A1–C2) enligt GERS. Talproven – monologer på 40–60 sekunder – analyserades auditivt av forskarna. Vid lyssnandet identifierades L2-drag i *segment*, *segmentlängder*, *ordbetoning* och *satsbetoning*. Undersökningen ger en mångsidig bild av talet på färdighetsnivåerna A1–C2 genom att såväl uttal, produktionsförmåga som komplexitet i talet analyserades. Oss veterligt finns det sedan tidigare ingen liknande undersökning om L2-svenska.

Resultaten kan sammanfattas i följande fynd:

- (i) Produktionsförmågan ökade närmast linjärt från A1 till C2. Det här syntes i större antal ord per talprov och högre taltempo på nivåerna B2–C2 än på nivåerna A1–B1.
- (ii) Komplexiteten i talet ökade också med högre färdighetsnivå, dock inte lika linjärt som produktionsförmågan. På nivåerna B2–C2 var antalet ord per yttrande större, andelen bisatser av alla satser högre och antalet flerstaviga ord per talprov större än på nivåerna A1–B1.
- (iii) Andelen L2-drag i segment minskade närmast linjärt från A1 till C2. Fonemfel var den vanligaste typen av L2-drag i segment. L2-drag i segment var inte heller särskilt frekventa på nivåerna A1–A2, vilket delvis kan bero på att finska och finlandssvenska har vissa fonetiska likheter i segmentuttalet.
- (iv) L2-drag i segmentlängder och ordbetoning var relativt få. Att dragen behärskades redan på nivåerna A1–A2 kan bero på att det aktiva ordförrådet i början av L2-inläring i huvudsak består av högfrekventa och morfologiskt enkla ord. Att inlärarna hade finska med ett utvecklat kvantitetssystem som L1 kan också ha bidragit till att segmentlängderna i svenska var lätta för dem.
- (v) Andelen L2-drag i satsbetoning minskade inte linjärt, tvärtom ökade den på nivån B2. Först på nivån C2 skedde det en tydlig minskning av L2-drag i satsbetoning. Att L2-dragen i satsbetoning inte var färre på nivåerna B2–C1 än på nivåerna A1–B1 kan bero på att talarna på nivåerna B2–C1 producerade mer tal och hade mer komplexa yttranden än talarna på nivåerna A1–B1. Det är troligt att produktionsförmågan, komplexiteten och L2-dragen i satsbetoning har ett orsakssamband.
- (vi) Det dominerande L2-draget i satsbetoning var för många betonade ord i yttrandet, vilket tillhör de vanligaste dragen i L2-prosodin oavsett målspråk (t.ex. Rasier & Hiligsmann 2007). Andelen betonade ord av alla uttalade ord minskade något på de högre nivåerna (jfr en liknande iakttagelse om L2-engelska i Kang & Yan 2018 och Kang 2013).

Resultaten ökar vår kunskap om L2-talets fonetiska och strukturella uppbyggnad samt dessas förhållande till bedömning av L2-tal. Vad gäller inlärningsgångarna visar de att utveckling av satsprosodin inte tycks kunna beskrivas linjärt i samma grad som utveckling av segment, ordprosodi och produktionsförmåga (eller vissa grammatiska strukturer, jfr Abrahamsson 2009:63ff.). Kanske kan inlärningsgången i satsprosodin snarare sägas likna U-formen, möjligen på grund av att den påverkas av utveckling av lexikon

och grammatik. Vad gäller diskussionen om förhållandet mellan komplexitet och korrekthet i L2-produktionen (se t.ex. Housen, Kuiken & Vedder 2012), stöder våra resultat iakttagelsen om att ökad komplexitet kan försämra korrektheten av L2-talet. Detta gällde emellertid endast satsprosodin i vår studie, inte andra aspekter av uttalet.

Om resultaten jämförs med beskrivningarna i GERS (Europarådet 2018:136, 144), stämmer de rätt väl vad gäller produktionsförmågan och komplexiteten. Segmenten behärskades däremot bättre på nivåerna A1–A2 än förväntat, medan satsprosodin på nivåerna B2–C1 inte till alla delar var så bra som den borde vara enligt GERS (2018:136) där det t.ex. konstateras om nivån C1 (kursiveringen i citatet är vår):

Can produce smooth, intelligible spoken discourse with only occasional lapses in control of stress, rhythm and/or intonation, which do not affect intelligibility or effectiveness. Can vary intonation and place stress correctly in order to express precisely what he/she means to say.

Resultaten bör dock inte tolkas strikt i förhållande till GERS. För det första beaktade vi inte flyt annat än genom taltempo. Det kan vara att talarna på nivån B2 skiljer sig från talarna på nivån B1 snarare i flyt än uttal – det snabbare taltempot hos talarna på nivån B2 tyder på det. För det andra baserar sig färdighetsnivåutdelningen på en bedömning av talarens allmänna muntliga färdighet där uttalet endast är ett av kriterierna. Därför är det möjligt att uttalet varierar mellan olika talare på samma färdighetsnivå. För det tredje var antalet talare i vår studie relativt litet, och därför måste resultaten bekräftas i andra studier. Även om vi beaktar dessa osäkerhetsfaktorer, anser vi att vår studie pekar på att L2-satsprosodin inte utvecklas linjärt, åtminstone inte i språkparet finska–finlandssvenska, och att prosodin därmed tycks skilja sig från några andra egenskaper i L2-talet. Man kan därför förhålla sig kritiskt till att utveckling av prosodin beskrivs som en tydlig inlärningsgång i GERS. Detta bör också beaktas vid bedömning av uttal: bedömare bör vara medvetna om att uttalsbeskrivningen i GERS inte bör tillämpas strikt.

Tidigare forskning visar att begriplighet av L2-tal i betydande grad påverkas av prosodin, främst ord- och satsbetoningen (t.ex. Munro & Derwing 1995 och Hahn 2004 om L2-engelska; Heinonen 2020 om L2-svenska). Eftersom GERS-skalan prioriterar begriplighet framför inföddlighet, kunde man kanske förvänta sig att också satsprosodin utvecklas såsom det beskrivs i GERS. Att så ändå inte tycks vara fallet beror kanske på att bedömare tar hänsyn till andra faktorer än uttal vid bedömning av L2-talarens färdighetsnivå. Ökad förekomst av L2-drag i satsbetoning kan då kanske accepteras av bedömaren som konsekvens av till exempel rikare innehåll och

ökad komplexitet, eller så är de L2-drag som finns kvar i satsbetoningen inte så viktiga för begripligheten.

Resultaten är värdefulla för projektet DigiTala som avser att skapa en digital applikation för automatisk bedömning av L2-svenska och L2-finska. För att kunna ge ändamålsenlig feedback ska applikationen lära sig att fokusera på relevanta uttalsdrag. Ett viktigt syfte med studien var därför att försöka identifiera och beskriva de fonetiska drag som kan vara relevanta vid bedömning av muntlig språkfärdighet. Nästa steg i projektet är att försöka identifiera akustiska parametrar som motsvarar mänsklig bedömning av individuella uttalsdrag och som automatisk bedömning kan basera sig på. Genom att jämföra dessa uttalsdrag och akustiska parametrar i svenska med parametrar i andra språk (om engelska se t.ex. Hsieh, Zechner & Xi 2020) får vi värdefull information om hur väl samma bedömningskriterier fungerar i olika målspråk. Eftersom satsbetoning inte visade någon tydlig utveckling mellan nivåerna i den här studien, ska draget testas med andra parametrar än de som vi använde i vår auditiva analys. Främst ska de parametrar som använts vid automatisk bedömning av satsprosodin i L2-engelska (jfr inledningen) testas i L2-svenska med samma eller liknande talmaterial som användes här.¹²

Den centrala frågan är då om automatisk bedömning som utnyttjar ASR bättre kan skilja satsprosodin åt på nivåerna A1–B1 och B2–C1 än vad som var fallet här. En följdfråga är om automatisk bedömning i så fall också bättre motsvarar en holistisk mänsklig bedömning av talproven (jfr Kautonen 2019). Om så är fallet, blir konsekvensen att det inte är lönt att försöka beskriva uttalet analytiskt när datorn lärs att imitera mänsklig bedömning, eftersom det de facto kan leda till en mindre tillförlitlig bedömning av uttalet och därmed sämre input till ASR. Här kommer vi in på det som diskuterades i inledningen av artikeln, nämligen att automatisk bedömning i bästa fall kan vara mer tillförlitlig vid bedömning av uttal än människan, eftersom datorn också kan beakta flytet och produktionsförmågan. Också människan kan göra det vid holistisk bedömning, men det är svårt för oss att skilja olika egenskaper åt vid bedömningen och att bedöma ett stort antal talare med samma kriterier. Samtidigt är det viktigt att kunna beskriva uttal och olika uttalsdrag på ett konkret och för inlärare förståeligt sätt för att man ska kunna ge användbar feedback till L2-talare på deras muntliga färdighet.

¹² Detta bör dessutom testas med såväl finlandssvenska som sverigesvenska, vilka satsprosodiskt skiljer sig en del från varandra (Kuronen & Leinonen 2010). Det kan emellertid vara så att samma parametrar fungerar med en likartad precision vid analys av L2-tal i båda varieteterna.

Vi vill tacka Finlands Akademi för finansiering av projektet DigiTala (finansieringsnummer 322965) samt de allmänna språkexamina och Utbildningsstyrelsen för det material som studien baserar sig på.

Litteratur

- Abrahamsson, Niclas (2009). *Andraspråksinläring*. Lund: Studentlitteratur.
- Bannert, Robert (2004). *På väg mot svenskt uttal*. Lund: Studentlitteratur.
- Boersma, Paul & Weenink, David (2019). Praat: doing phonetics by computer [Ett dataprogram för akustisk analys av tal]. Amsterdam: University of Amsterdam. <http://www.fon.hum.uva.nl/praat/>
- Bruce, Gösta (2010). *Vår fonetiska geografi: Om svenskans accenter, melodi och uttal*. Lund: Studentlitteratur.
- Burnstein, Jill, Kukich, Karen, Wolff, Susanne, Lu, Chi, Chodorow, Martin, Braden-Harder, Lisa & Harris, Mary Dee (1998). Automated Scoring Using A Hybrid Feature Identification Technique. *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics. Montreal, Canada: Association for Computational Linguistics*, s. 206–210. <https://doi.org/10.3115/980845.980879>
- Chen, Lei, Zechner, Klaus, Yoon, Su-Youn, Evanini, Keelan, Wang, Xinhao, Loukina, Anastassia, Tao, Jidong, Davis, Lawrence, Lee, Chong Min, Ma, Min, Mundkowsky, Robert, Lu, Chi, Leong, Chee Wee & Gyawali, Binod (2018). *Automated Scoring of Nonnative Speech Using the SpeechRaterSM v. 5.0 Engine*. Research Report ETS, RR-18-10. <https://doi.org/10.1002/ets2.12198>
- Cucchiari, Catia, Neri, Ambra, de Wet, Febe & Strik, Helmer (2007). ASR-based pronunciation training: Scoring accuracy and pedagogical effectiveness of a system for Dutch L2 learners. *Interspeech 2007*, s. 2181–2184.
- Cucchiari, Catia, Strik, Helmer & Boves, Lou (2002). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *The Journal of the Acoustical Society of America*, 111/6, s. 2862–2873. <https://doi.org/10.1121/1.1471894>
- Derwing, Tracey & Munro, Murray (1997). Accent, intelligibility, and comprehensibility: Evidence from Four L1s. *Studies in Second Language Acquisition*, 19(1), s. 1–16. <https://doi.org/10.1017/S0272263197001010>
- Derwing, Tracey, Rossiter, Marian, Munro, Murray & Thomson, Ron (2004). Second Language Fluency: Judgments on Different Tasks. *Language Learning*, 54(4), s. 655–679. <https://doi.org/10.1111/j.1467-9922.2004.00282.x>
- Eskenazi, Maxine (1999). Using automatic speech processing for foreign language pronunciation tutoring: Some issues and a prototype. *Language Learning & Technology*, 2(2), s. 62–76.
- Europarådet. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge, Storbritannien: Cambridge University Press. <https://rm.coe.int/1680459f97> (Hämtad 6.5.2020)

- Europarådet. (2018). *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume with new descriptors. Provisional edition. September 2017.* <https://rm.coe.int/cefr-companionvolume-with-new-descriptors-2018/1680787989> (Hämtad 6.5.2020)
- Gretter, Roberto, Matassoni, Marco, Allgaier, Katharina, Tchistiakova, Svetlana & Falavigna, Daniele (2019). Automatic Assessment of Spoken Language Proficiency of Non-native Children. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. <https://doi.org/10.1109/ICASSP.2019.8683268>
- Gu, Lin & Davis, Larry (2020). Providing SpeechRater Feature Performance as Feedback on Spoken Responses. I: Zechner, Klaus & Evanini, Keelan (red.). *Automated speaking assessment: Using language technologies to score spontaneous speech*. New York: Routledge. S. 159–175. <https://doi.org/10.4324/9781315165103-1>
- Hahn, Laura D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), s. 201–223. <https://doi.org/10.2307/3588378>
- Heinonen, Henna (2020). *Uttal och dess begriplighet i finskspråkiga gymnasisters L2-svenska*. JYU Dissertations, 296. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-8325-3>
- Heinonen, Henna & Kautonen, Maria (2020). L2-ruotsin lausepainon sanallinen kuvaaminen – menetelmällisiä käännteitä ääntämisen arviointiin. I: Grasz, Sabine, Keisanen, Tiina, Oloff, Florence, Rauniomaa, Mirka, Rautiainen, Ira & Siromaa, Maarit (red.). *AFinLAN vuosikirja 2020*. Jyväskylä: Suomen soveltavan kielitieteen yhdistys. S. 71–94. <https://doi.org/10.30661/afinlavk.89312>
- Hildén, Raili (2000). *Att tala bra, bättre och bäst: Suomenkielisten abiturienttien ruotsin kielen taito testisuoritusten valossa*. Doktorsavhandling. Helsingfors: Helsingfors universitet. <http://hdl.handle.net/10138/229501>
- Hildén, Raili & Rautopuro, Juhani (2014). *Ruotsin kielen A-oppimäärän oppimistulokset perusopetuksen päättövaiheessa 2013*. Helsingfors: Nationella centret för utbildningsutvärdering; Utbildningsstyrelsen. http://karvi.fi/app/uploads/2014/09/KARVI_0114.pdf (Hämtad 6.5.2020)
- Hilton, Nanna Haug, Gooskens, Charlotte & Schüppert, Anja (2011). Artikulasjonshastighet i norske, svenske og danske radionyheter. *Norsk Lingvistisk Tidsskrift*, 29, s. 205–220.
- Housen, Alex, Kuiken, Folkert & Vedder, Ineke (2012). Complexity, accuracy and fluency: Definitions, measurement and research. I: Housen, Alex, Kuiken, Folkert & Vedder, Ineke (red.). *Dimensions of L2 Performance and Proficiency. Complexity, Accuracy and Fluency in SLA*. Amsterdam: John Benjamins. S. 1–20. <https://doi.org/10.1075/llt.32.01hou>
- Hsieh, Ching-Ni, Zechner, Klaus & Xi, Xiaoming (2020). Features Measuring Fluency and Pronunciation. I: Zechner, Klaus & Evanini, Keelan (red.). *Automated speaking assessment: Using language technologies to score spontaneous speech*. New York: Routledge. S. 101–122. <https://doi.org/10.4324/9781315165103-1>
- IPA, The International Phonetic Alphabet (2015). International Phonetic Association. https://www.internationalphoneticassociation.org/sites/default/files/IPA_Kiel_2015.pdf
- Isaacs, Talia (2018). Fully automated speaking assessment: Changes to proficiency testing and the role of pronunciation. I: Kang, Okim, Thomson, Ron I. & Murphy, John M. (red.). *The Routledge handbook of contemporary English pronunciation*. Abingdon, UK: Routledge. S. 570–584. <https://doi.org/10.4324/9781315145006-36>

- Isaacs, Talia & Trofimovich, Pavel (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, 34, s. 475–505. <https://doi.org/10.1017/s0272263112000150>
- Ivars, Ann-Marie (2015). *Dialekter och småstadsspråk. Svenskan i Finland – i dag och i går I:1*. Helsingfors: Svenska litteratursällskapet i Finland. <http://urn.fi/URN:NBN:fi:sls-978-951-583-496-6>
- Kallio, Heini, Šimko, Juraj, Huhta, Ari, Karhila, Reima, Vainio, Martti, Lindroos, Erik, Hildén, Raili & Kurimo, Mikko (2017). Towards the phonetic basis of spoken second language assessment: temporal features as indicators of perceived proficiency level. I: Kuronen, Mikko, Lintunen, Pekka & Nieminen, Tommi (red.). *AFinLA-E: Soveltavan kielitieteen tutkimuksia*, 10. S. 193–213. <https://doi.org/10.30660/afinla.73137>
- Kallio, Heini, Suni, Antti, Šimko, Juraj & Vainio, Martti (2020). Analyzing second language proficiency using wavelet-based prominence estimates. *Journal of Phonetics*, 80, 100966. <https://doi.org/10.1016/j.wocn.2020.100966>
- Kang, Okim (2013). Linguistic analysis of speaking features distinguishing general English exams at CEFR levels. *Research notes* 52, s. 40–48.
- Kang, Okim & Yan, Xun (2018). Linguistic features distinguishing examinees' speaking performances at different proficiency levels. *Journal of Language Testing and Assessment*, 1, s. 24–39. <https://doi.org/10.23977/langta.2018.11003>
- Karhila, Reima, Rouhe, Aku, Smit, Peter, Mansikkaniemi, André, Kallio, Heini, Lindroos, Erik, Hildén, Raili, Vainio, Martti & Kurimo, Mikko (2016). Digitala: An augmented test and review process prototype for high-stakes spoken foreign language examination. *Interspeech 2016*, s. 784–785. <http://urn.fi/URN:NBN:fi:aalto-201701191309>
- Kautonen, Maria (2017). Finskspråkiga talarers intonation av finlandssvenska i påståendeyttranden i fritt tal. *Folkmålsstudier*, 55, s. 31–60. <http://urn.fi/URN:NBN:fi:juu-201710314099>
- Kautonen, Maria (2019). *Finskspråkiga inlärares uttal av finlandssvenska i fritt tal på olika färdighetsnivåer*. JYU Dissertations, 90. Jyväskylä: University of Jyväskylä. <http://urn.fi/URN:ISBN:978-951-39-7778-8>
- Kautonen, Maria & Kuronen, Mikko. Uttalsinläring med fokus på svenska som andraspråk. Manuskript till en bok.
- Kautonen, Maria & von Zansen, Anna (2020). DigiTala research project: Automatic speech recognition in assessing L2 speaking. *Kieli, koulutus ja yhteiskunta*, 11(4). <https://www.kieliverkosto.fi/fi/journals/kieli-koulutus-ja-yhteiskunta-kesakuu-2020/digitala-research-project-automatic-speech-recognition-in-assessing-l2-speaking> (Hämtad 21.8.2020)
- Kuronen, Mikko & Leinonen, Kari (2010). *Svenskt uttal för finskspråkiga*. Tammerfors: Tampereen yliopistopaino.
- Kuronen, Mikko & Zetterholm, Elisabeth (2017). Olika fonetiska dragens relativa betydelse för upplevd inföddlighet i svenska. *Nordand*, 12(2), s. 134–156. <https://doi.org/10.18261/issn.2535-3381-2017-02-03>
- Kyle, Kristopher & Crossley, Scott A. (2015). Automatically assessing lexical sophistication: Indices, tools, findings, and application. *Tesol Quarterly*, 49(4), s. 757–786. <https://doi.org/10.1002/tesq.194>
- Leinonen, Kari, Vihanta, Veijo & Pitkänen, Antti (1990). Om rytmen i finlandssvenska och sverigesvenska: Ett bidrag till forskningen om finlandssvenskt uttal. I: Nikula,

- Kristina & Pitkänen, Antti (red.). *Svenskan i Finland: SiF rapport 1. seminariet i Tammerfors 12–13 okt. 1989*. Tammerfors: Tampereen yliopistopaino. S. 49–100.
- Monaghan, William & Bridgeman, Brent (2005). E-rater as a quality control of human scores. *R&D Connections*, 2, April 2005. Princeton, NJ: Educational Testing Service. https://www.ets.org/Media/Research/pdf/RD_Connections2.pdf (Hämtad 6.5.2020)
- Munro, Murray J. & Derwing, Tracey M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), s. 73–97. <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x>
- Pallant, Julie (2005). *SPSS survival manual: A step by step guide to data analysis using SPSS*. Sydney, Australien: Allen & Unwin.
- Rasier, Laurent & Hiligsmann, Philippe (2007). Prosodic transfer from L1 to L2. Theoretical and methodological issues. *Nouveaux cahiers de linguistique française* 28. Genève, Schweiz: Université de Genève. S. 41–66.
- Reuter, Mikael (2015). Finlandssvenskt uttal. I: Tandefelt, Marika (red.). *Gruppspråk, samspråk, två språk. Svenskan i Finland i dag och i går I:2*. Helsingfors: Svenska litteratursällskapet i Finland. S. 19–34. <http://urn.fi/URN:NBN:fi:sls-978-951-583-492-8>
- SAG = Teleman, Ulf, Hellberg, Staffan och Andersson, Erik (1999). *Svenska Akademiens grammatik*. Stockholm: Svenska Akademien. <https://svenska.se/grammatik/>
- Suomi, Kari, Toivanen, Juhani & Ylitalo, Riikka (2008). *Finnish sound structure. Phonetics, phonology, phonotactics and prosody*. Uleåborg: Uleåborgs universitet.
- Trofimovich, Pavel & Isaacs, Talia (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15(4), s. 905–916. <https://doi.org/10.1017/s1366728912000168>
- Undervisningsministeriet. (2006). *Lukiokoulutuksen suullisen kielitaidon arviointiryhmän muistio*. Opetusministeriön työryhmämuistioita ja selvityksiä 2006: 26. Helsingfors: Undervisningsministeriet. <http://urn.fi/URN:ISBN:952-485-213-6>
- Undervisnings- och kulturministeriet. (2017). *Gaudeamus igitur – ylioppilastutkinnon kehittäminen*. Opetus- ja kulttuuriministeriön julkaisuja 2017: 16. Helsingfors: Undervisnings- och kulturministeriet. <http://urn.fi/URN:ISBN:978-952-263-462-7>
- Utbildningsstyrelsen. (2011). *Grunderna för allmänna språkexamina 2011. Föreskrift 24/011/2011*. Föreskrifter och anvisningar 2011: 33. Tammerfors: Juvenes Print.
- van Santen, Jan P. H., Prud'hommeaux, Emily Tucker & Black, Lois M. (2009). Automated assessment of prosody production. *Speech Communication*, 51(11), s. 1082–1097. <https://doi.org/10.1016/j.specom.2009.04.007>
- Wik, Preben (2011). *The Virtual Language Teacher: Models and applications for language learning using embodied conversational agents*. Doctoral Thesis. Stockholm: Kungliga Tekniska högskolan. urn:nbn:se:kth:diva-33579
- Wikner, Sarah (2019). *Svenskan i Helsingfors: Uppfattningar, perception och variation*. Doktorsavhandling. Åbo Akademi: Åbo. <http://urn.fi/URN:ISBN:978-952-12-3873-4>
- Yoon, Su-Youn & Zechner, Klaus (2017). Combining human and automated scores for the improved assessment of non-native speech. *Speech Communication*, 93, s. 43–52. <https://doi.org/10.1016/j.specom.2017.08.001>
- Zechner, Klaus, Higgins, Derrick, Xi, Xiaoming & Williamson, David M. (2009). Automatic scoring of non-native spontaneous speech in tests of spoken English. *Speech Communication*, 51(10), s. 883–895. <https://doi.org/10.1016/j.specom.2009.04.009>