LAMBA: Latvian language in monolingual and bilingual acquisition
What is the project all about?

• Systematic and comprehensive investigation of monolingual and bilingual acquisition of the Latvian language

• Challenges addressed:
  – lack of early language assessment instruments
  – lack of developmental norms and population studies
  – lack of resources: child corpora, experimental data…
  – lack of studies focusing on early bilingual acquisition of Latvian
  – poor visibility of Latvian in the international academic arena
  – lack of international collaboration in language acquisition
Goals of the project

• **Developing tools:**
  - *MacArthur-Bates Communicative Development Inventory (CDI)* – Latvian adaptation: CDI-I (8-16 months) and CDI-II (16-36 months)
  - *Latvian Phoneme Accuracy Test* (3;0-6;0 years) expanded and adapted for use with Russian-dominant children and normed for both languages

• **Developing research resources:**
  - Longitudinal corpora of child language of two Latvian monolingual and two Latvian-Russian bilingual children: recorded, transcribed, coded and annotated
  - Normative dataset of phoneme accuracy of Latvian- and Russian-speaking children
  - Normative dataset of CDI-I and CDI-II for Latvian monolinguals
  - Experimental data on phonological and morphosyntactic development for mono- and bilinguals
Goals of the project

• **Developing knowledge:**
  – Generate new knowledge on early language development of Latvian-speaking children, monolingual and bilingual
  – increase understanding of factors involved in language acquisition in general

• **Sharing research results & increasing awareness:**
  – Dissemination and information targeting different groups to attract attention of researchers, practitioners, parents and the wider community to the importance of child language studies and recent developments

• **Developing expertise:**
  – through training of research staff
  – through developing research-based course modules (Synergy with Scholarship Activity)
Funding & Duration

• Program LV05 “Bilateral Research Cooperation” under the EEA and Norway Grants: **EUR 501 021**
• Self co-financing contributed by the UiT the Arctic University of Norway and the University of Oslo: **EUR 138 673**

• TOTAL BUDGET: **EUR 639 694**

• Duration: **26** months (March 2015 – April 2017)

• Research & technical staff directly involved: **27 people**
The Team

University of Oslo

Center for Multilingualism in Society across the Lifespan

UiO
The Team
Structure of the Project

• **WP1**: Latvian adaptation and norming of MacArthur-Bates Communicative Development Inventories (CDI)

• **WP2**: Development and norming of Latvian and Russian phoneme tests

• **WP3**: Development of 4 dense longitudinal child speech corpora with 2 monolingual Latvian children and 2 Latvian-Russian bilingual children orthographically transcribed and morphologically annotated

• **WP4**: Experimental study of phonological and morphological aspects of monolingual and bilingual acquisition of Latvian
WP1 What is MacArthur-Bates CDI?

- Tool assessing communicative development in children 8 months–3 years:
  - vocabulary
  - communicative gestures
  - grammatical development

- Parental report tool:
  - not affected by performance limitations in the child
  - reliable information on child’s linguistic competence
  - time & cost efficient

- Developed originally for children speaking American English (Fenson et al. 1994; Fenson et al. 2007)

- Adapted to over 60 languages – both spoken and signed, including Norwegian, Danish, Finnish, Estonian, Russian etc.
WP1 What is CDI used for?

• **Typical populations:**
  – norm-referenced assessment tool for mono- and bilingual children
  – cross-linguistic comparisons (Bleses et al. 2008)

• **Atypical populations: evaluation, screening and monitoring tool:**
  – Specific language impairment (Thal et al. 1999)
  – Down syndrome (Berglund, Eriksson & Johansson 2001)
  – Autism (Charman et al. 2003)
  – Cochlear implants (Thal et al. 2007)

• **Populations at risk of developing language impairment: monitoring**
  – otitis media (Feldman et al. 2003)
  – healthy preterm children (Magill-Evans & Harrison 1999)
  – familial risk of dyslexia (Koster et al. 2005)
WP1 Latvian CDI: outputs

For parents & practitioners:
- Two electronic forms, available online:
  - CDI I: Words & Gestures (for children aged 8 to 16 months)
  - CDI II: Words & Sentences (for children aged 16 to 36 months)
  - scores calculated automatically
  - convenient comparison with population norms (based on a sample of 2320 children, assuming 35% response rate)

For scholars:
- Normative data available through open-access Wordbank database (http://wordbank.stanford.edu/): cross-linguistic comparisons, statistics, visualizations, etc.
WP1 Latvian adaptation of CDI

- First language assessment tool normed with Latvian population
- First tool to provide a comprehensive overview of lexical and grammatical development of Latvian children
- First tool that does not require specialized training to administer, score & evaluate
WP1 Timeline

2015
• Develop Latvian adaptation of CDI:WG and CDI:WS
• Pilot adaptations with a sample of parents
• Get approval of the CDI Advisory Board

2016
• Implement a large-scale web-based data collection
• Test Latvian CDI for validity & reliability
• Analyze the results quantitatively and qualitatively
• Create a detailed user manual

2017
• Make all resources freely available: Wordbank, CHILDES, downloads from partners’ websites
WP2 Pronunciation accuracy tests

• **What is it?**
  – a picture-based tool to assess the development of pronunciation in children from 3 to 6 years of age
  – pilot-tested for picture recognizability
  – available in Latvian and in Russian
  – assesses consonants in three positions in the word (initial, medial, final)
  – assesses a range of consonant clusters (initial and medial);
  – modeled after Norsk Fonemtest
  – **complete with developmental norms** (n = 500 in each population)
WP2 Pronunciation accuracy tests

• Why do educators, speech therapists and parents need it?
  – the only norm-referenced tool of its kind for Latvian
  – quick & easy to administer (full test takes 10 to 20 min.)
  – easy to score (comes with user-friendly scoring sheets & detailed manual)
  – easy to interpret the score (comes with developmental norms)
  – two comparable versions – Latvian and Russian – enable evaluation of bilinguals
  – available for download free of charge

• Why do scholars need it?
  – full recorded normative dataset available: wealth of data on phonemic development
  – convenient tool for matching children on phonological skills & pre-selection of children with specific characteristics
WP2 Timeline

2015
• Develop Latvian phoneme test
• Develop an equivalent tool in Russian so that bilingual children can be tested in both languages
• Pilot all picture stimuli to ensure they elicit target words and are easy to identify; & all target words are familiar to the children of target age

2016
• Obtain developmental norms (recording and assessment, n =1000 total)
• Analyze data quantitatively & qualitatively
• Develop detailed user manuals

2017
• Make full normative dataset, tools & manuals freely available through partners’ websites
WP3 Longitudinal Latvian child speech corpora

• What is it?
  – 4 dense recorded longitudinal corpora (naturalistic setting)
  – 2 Latvian monolinguals & 2 Latvian-Russian bilinguals
  – recorded 30 minutes per week, for 16 months (ca. 20 to 36 months of age)
  – total of 192 hours of spoken interaction
  – orthographically transcribed
  – coded using CHAT – the standard transcription system utilized in Child Language Data Exchange System (CHILDES; MacWhinney 1987, 2000, 2014)
  – audio recordings and transcriptions linked and synchronized
  – automatic morphological tagging using MOR program
  – MOR grammar developed for the Latvian language
  – available for free through CHILDES database
**WP3** Longitudinal Latvian child speech corpora

- **Why is it needed?**
  - there are currently no recorded longitudinal data of Latvian child language (diary studies: Rūķe-Draviņa 1982, 1993, Markus 2003)
  - glimpse into bilingual language development as it unfolds
  - wealth of information on language development
    - error rates; type & token frequencies of lexical & grammatical categories; omission errors; developmental tendencies; phonological development
    - child-directed speech; parent-child interactions; conversational turn-taking; etc.
    - cross-linguistic comparisons
    - validation of formal assessment tools
WP3 Timeline

- **2015**
  - Recording & transcription of corpora

- **2016**
  - Recording & transcription of corpora
  - Coding corpora in CHAT format
  - Development of MOR grammar for Latvian & adaptation of existing morphological taggers

- **2017**
  - Analysis of the results
  - Corpora are submitted to CHILDES database and made available to everyone
WP4 Experimental study of (morpho)phonological acquisition of Latvian

• **What is it?**
  – Three experiments targeting Latvian-Russian *bilinguals*:
    • Acquisition of morphosyntactic properties: adjectival gender agreement
    • Acquisition of morphophonological alternations in nominal inflection
    • Acquisition of segmental contrasts: accuracy of production and complexity of inventories

• **Why is it needed?**
  – Unique bilingual situation with a sizeable Russian-speaking minority, and efforts towards bilingual preschool and school education
  – Early bilingual acquisition of Latvian is poorly understood
    • E.g. Kušķe (2013b) lists bilingualism as a reason for language delays in school-aged children
  – Bilingual children – a vulnerable group; bilingualism is often seen as a liability, not an advantage
WP4 Timeline

- **2015**
  - Develop & pilot methodology and stimuli
  - Adapt Utrecht Bilingual Exposure Calculator to Latvian and Russian;

- **2016**
  - Recording sessions in Latvian kindergartens (n = 120 children, 400 sessions)
  - Transcription of experimental data
  - Qualitative & quantitative analysis

- **2017**
  - Publishing the results
  - Sharing experimental datasets via Tromsø Repository of Language and Linguistics
Sharing the Results

• **All** tools, norms & datasets will be **open access**.
• Assessment tools with related manuals and norms - available for free download from the project webpage.
• Longitudinal corpora of child speech - donated to the Child Language Data Exchange System (CHILDES - [http://childes.psy.cmu.edu/](http://childes.psy.cmu.edu/)).
• CDI data shared through Wordbank (web-based cross-linguistic database for lexical data from adaptations of CDI -[http://wordbank.stanford.edu](http://wordbank.stanford.edu)), the full normative dataset - made available via CHILDES.
• Experimental datasets and statistical analyses - made available via the Tromsø Repository of Language and Linguistics (TROLLing - [http://opendata.uit.no/](http://opendata.uit.no/)).
What’s next?

• Massive open-access research resources created = sustainability:
  – joint publications
  – student projects
  – research-based course modules (synergy with Scholarship Activity)
  – new research projects & grant applications

• Comparable resources across countries
  – Basis for cross-linguistic comparisons and joint international projects

• Summer School in linguistics in Latvia?
  – use of new tools & resources;
  – theory & methodology of language acquisition

• Go global?
  – research network uniting scholars working on Baltic linguistics in the Baltic States and beyond
Thank you