

# Speech recognizer for North Sami

Juho Leinonen

Peter Smit\*

August 4, 2015

## **Abstract**

Speech recognizer for Sami built. N-gram tests. OOV-issues. Morphs better.

## **1 Introduction**

Speech recognition [1]. North Sami. Under-resourced languages. (DigiSami & WikiTalk)

## **2 The speech recognition system**

Acoustic modelling, language modelling (more next chapter), lexicon, the decoder.

### **2.1 Acoustic modelling**

### **2.2 Language modelling**

### **2.3 Lexicon / pronunciation dictionary**

### **2.4 Decoding**

## **3 Language modelling**

n-gram, lexicon, words, morphs, morfessor.

---

\*Contact this person, not first author

### **3.1 Words**

### **3.2 Morphs**

#### **3.2.1 Morfessor**

## **4 Experiments**

## **5 Results**

## **6 Discussion**

## **7 Conclusions**

## **8 Acknowledgements**

## **References**

- [1] Xuedong Huang, Alex Acero, Hsiao-Wuen Hon, and Raj Reddy. *Spoken language processing: A guide to theory, algorithm, and system development*. Prentice Hall PTR, 2001.