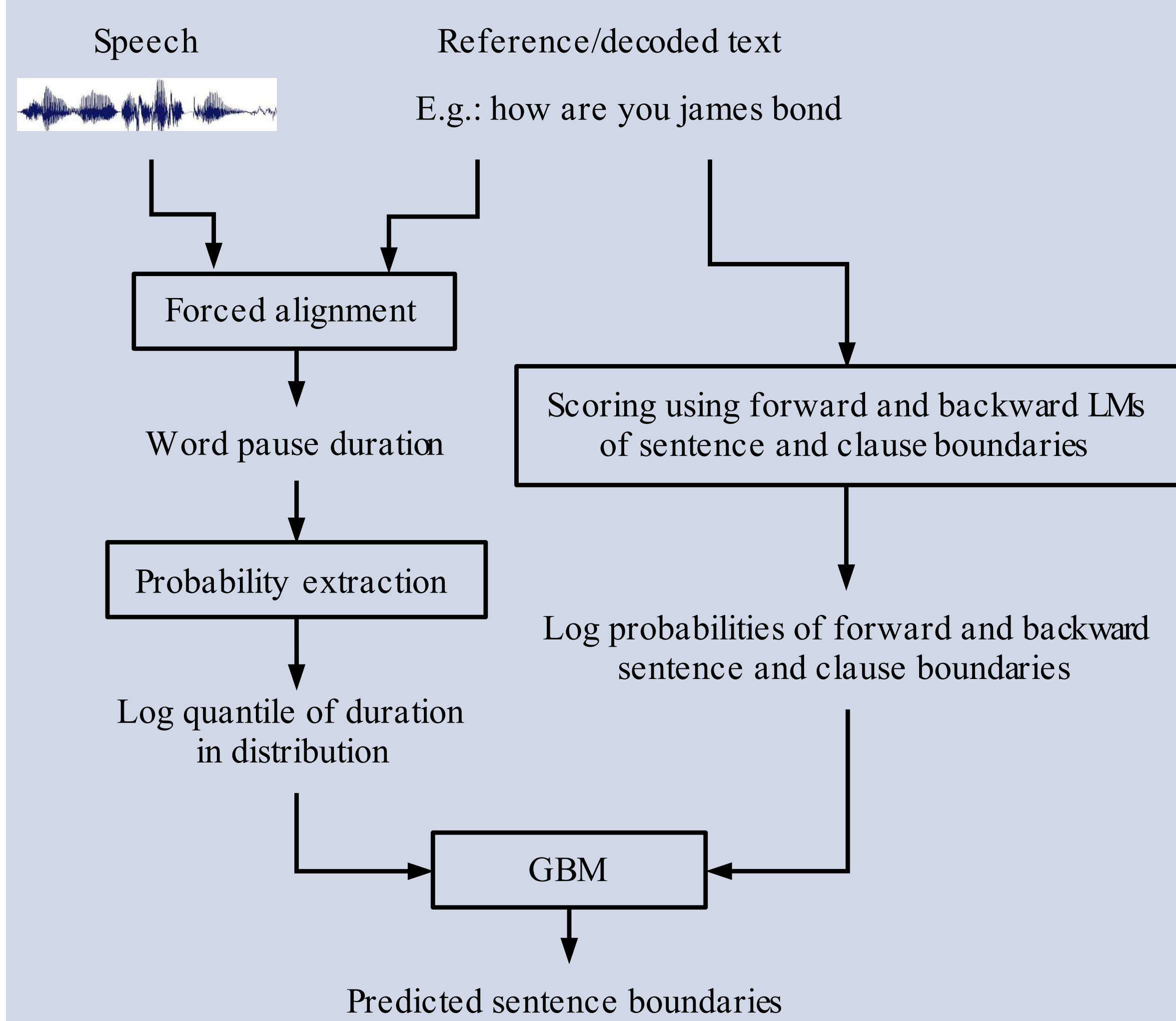


1. Overview

- ASR systems typically output transcripts as sequence of words lacking sentence boundaries, punctuation marks and case differentiation. However, the missing information is vitally necessary to downstream NLP tasks.
- The goal of punctuation prediction is to restore sentence boundary and punctuation marks for ASR transcripts.
- Punctuation prediction systems have been developed for English, German, and Russian for SUMMA project up to now, while Arabic, Spanish, and Latvian will be available soon.

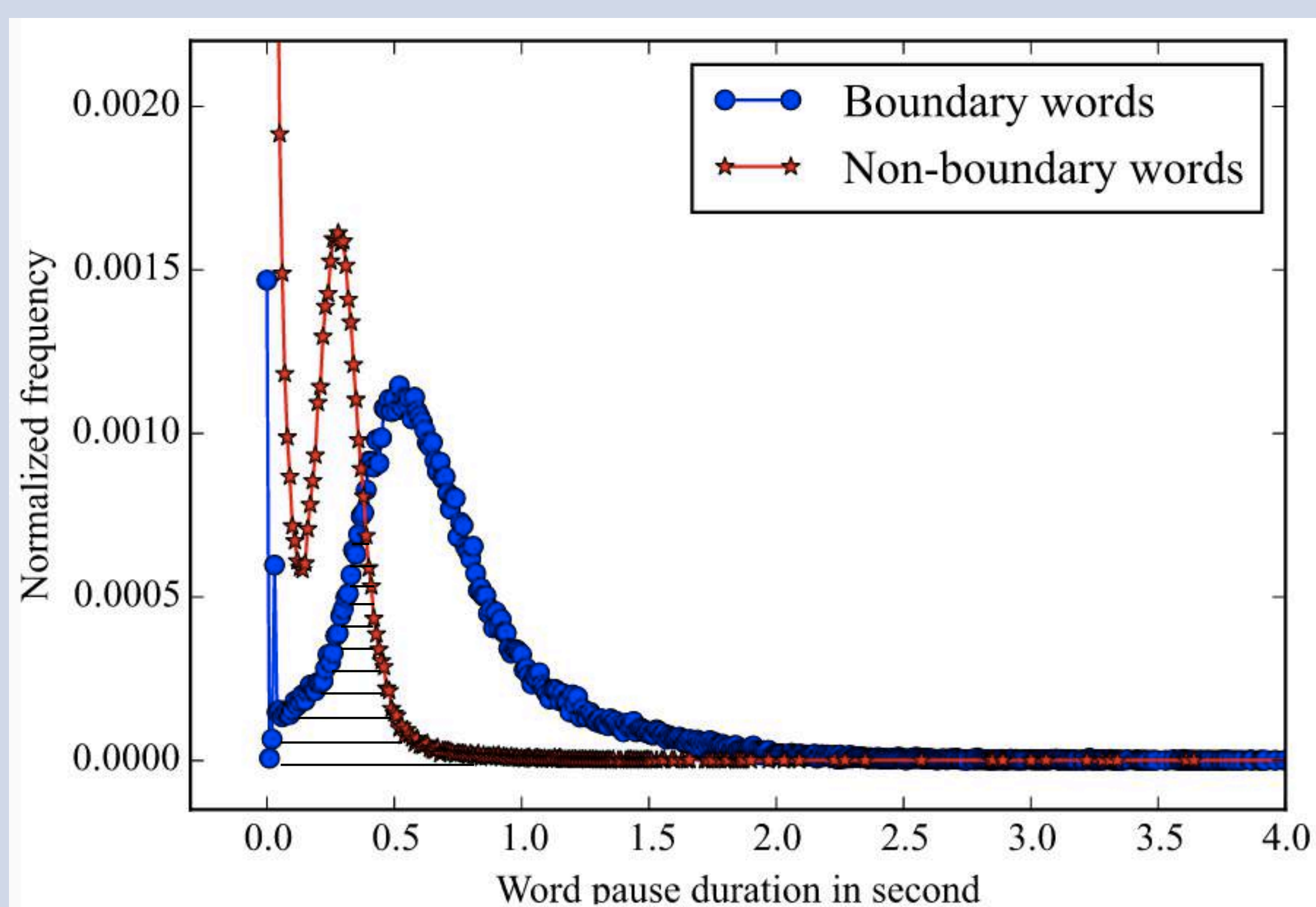
2. Boosting method

This method was developed at Idiap.



3. Experiments using boosting method

The German BCN corpus (WER: 7.2%)

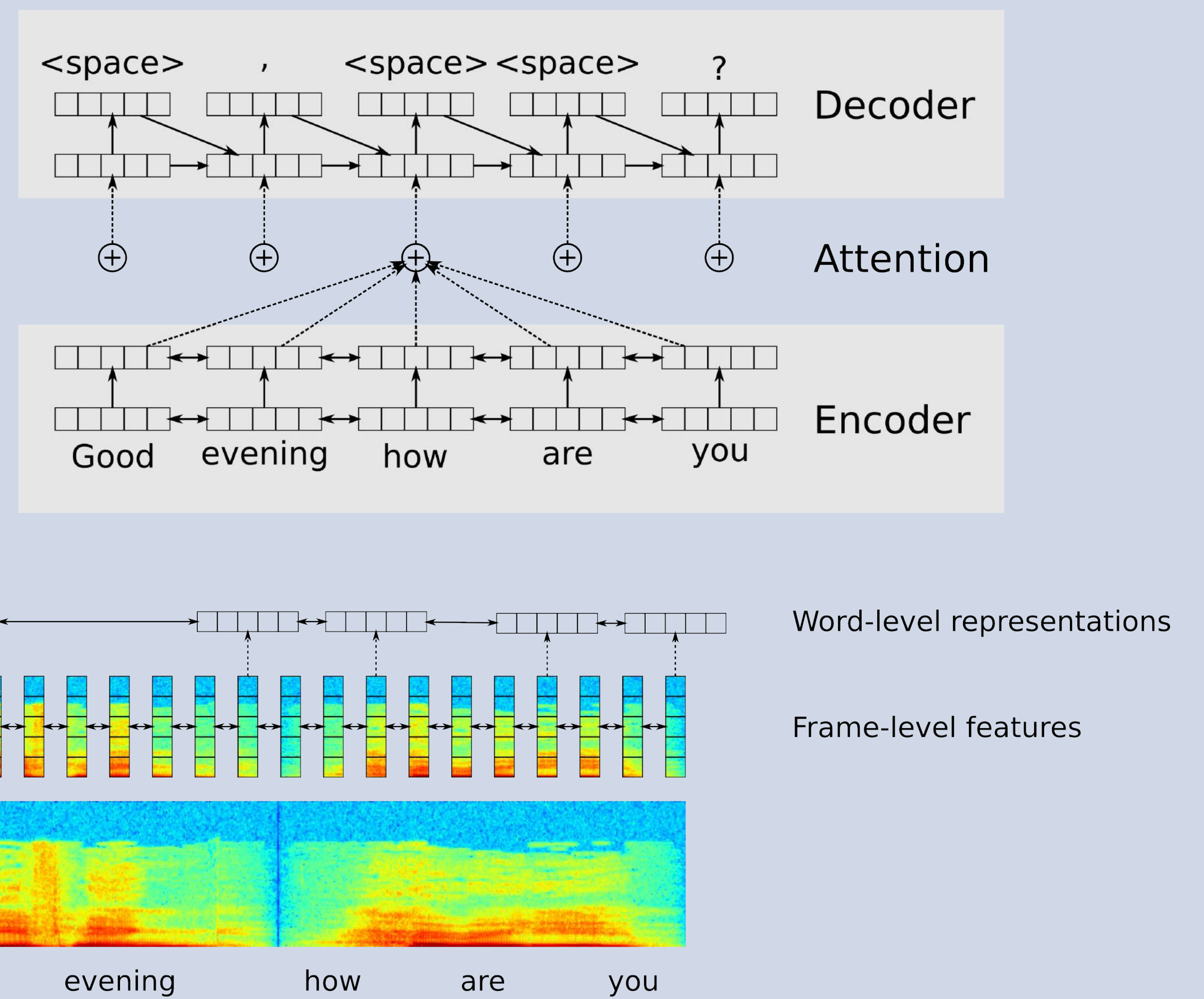


Features	Reference text (%)			ASR transcript (%)		
	P.	R.	F.	P.	R.	F.
UP	100.00	18.55	31.29	100.00	18.55	31.29
Dur	83.75	86.30	85.01	83.01	87.69	85.29
LM	78.74	58.06	66.84	77.75	56.82	65.66
Dur+LM	91.47	88.99	90.21	91.35	88.12	89.71
Dur+UP	87.68	87.04	87.36	88.11	86.93	87.52
LM+UP	82.52	64.33	72.30	81.64	63.20	71.25
Dur+LM+UP	91.91	89.25	90.56	91.67	88.18	89.89

4. Neural Machine Translation (NMT) method

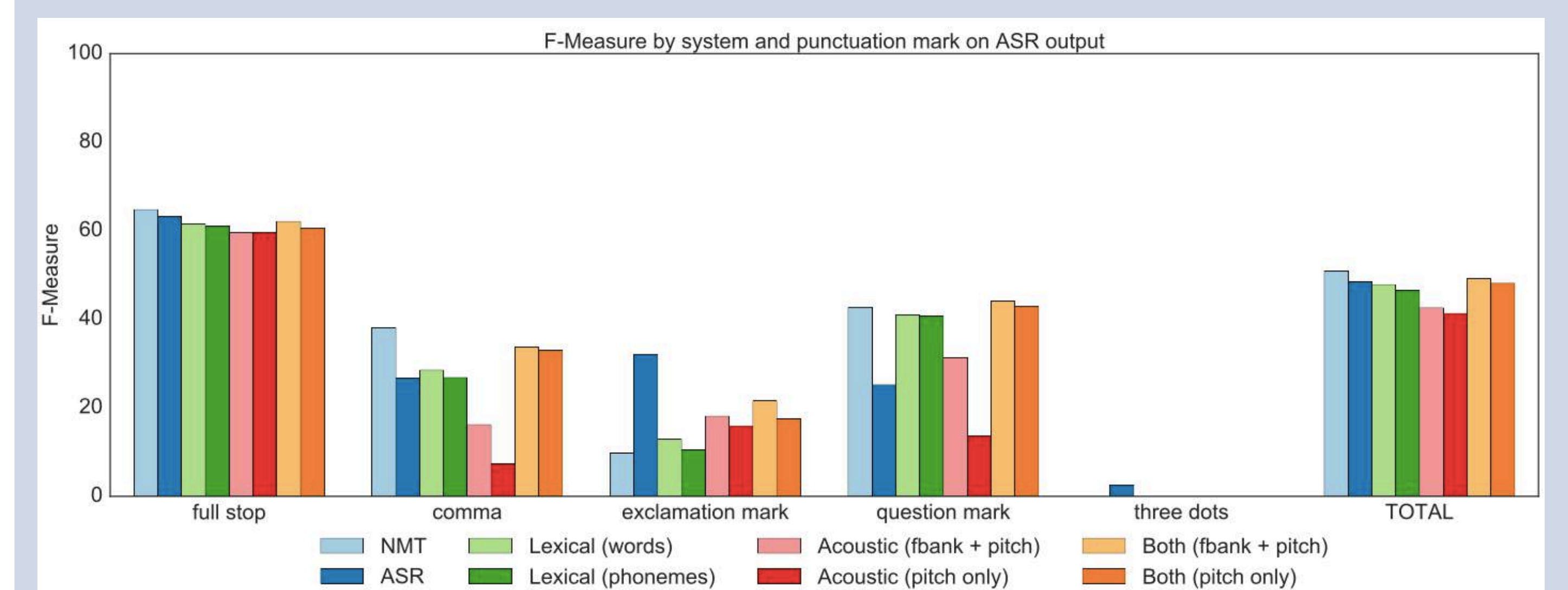
This method was developed at UEDin.

- NMT based system inserts punctuation marks by translating unpunctuated text to sequence of punctuation marks.
- Extended system uses both text and audio modalities. This system uses a hierarchical encoder to map frame level acoustic features to word level acoustic representations.
- Deployed systems are text-based and they are trained with open-source toolkit Nematus. AmunMT is used in production for fast decoding.



5. Experiments using NMT method

English MGB corpus (WER: 29%)



6. Future work

- PP systems will be developed for all SUMMA languages
- Idiap will generalize their system to predict more types of punctuation marks while trying to keep high F-measure

Publications

1. Wang Y., Nanchen A., Lazaridis A., Imseng D., Garner P. Comparative Study on Sentence Boundary Prediction for German and English Broadcast News. [Report] Idiap, 2017
2. Klejch O., Bell P., Renals S. Punctuated transcription of multi-genre broadcasts using acoustic and lexical approaches. SLT, 2016
3. Klejch O., Bell P., Renals S. Sequence-to-sequence models for punctuated transcription combining lexical and acoustic features. ICASSP, 2017