Morphological Segmentation for Keyword Spotting

Karthik Narasimhan[1], Damianos Karakos[2], Richard Schwartz[2], Stavros Tsakalidis[2], Regina Barzilay[1]

[1]CSAIL, Massachusetts Institute of Technology [2]Raytheon BBN Technologies

{karthikn,regina}@csail.mit.edu {dkarakos,schwartz,stavros}@bbn.com

1. Overview

- We explore the impact of morphological segmentation on Keyword Spotting (KWS).
- Handling out-of-vocabulary (OOV) words is a major challenge in KWS - we aim to alleviate this problem by utilizing sub-word units.
- We augment a state-of-the-art KWS system with sub-word units derived from supervised and unsupervised morphological segmentations, and compare with phonetic and syllabic segmentations.
- Our experiments demonstrate that morphemes improve overall KWS performance, both individually and in combination with other sub-word units.

2. Keyword Spotting

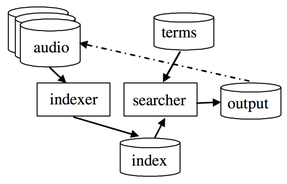
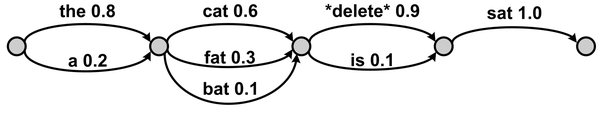


Fig 1: KWS system architecture (Fiscus et. al., Interspeech 2007)

System Components:

- Language model (bigram, trigram) over text
- Word confusion network (acoustic)



Morphologically rich languages have high OOV rate, which presents a severe problem for KWS systems.

Keyword Set	OOV rate
Turkish Dev	27.03%
Turkish Eval	36.99%
Turkish Dev+Eval	36.06%

Table 1: OOV rates for keywords in Turkish (Babel Limited Lexicon babel105b-v0.4)

3. Segmentation Methods

- Supervised morphological
- Unsupervised morphological (Morfessor)
- Random
- Phone-based
- Syllable-based

Supervised Segmentation

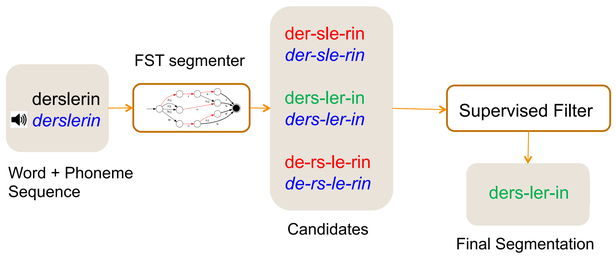


Fig 2: Supervised segmentation pipeline

Feature	Example
morpheme unigrams	tak, acak
morpheme bigram	⟨tak, acak⟩
phonetic seq. unigrams	t.a.k., 1v.dZ.a.k.
phonetic seq. bigram	⟨t.a.k., 1v.dZ.a.k.⟩
number of morphemes	2
morpheme lengths	3, 4

Table 2: Features used in supervised filters

We can encode features that go beyond individual boundaries - this global view distinguishes our classifier/ranker from traditional sequence tagging approaches to morphological segmentation.

5. Results

- Using sub-word units improves overall KWS performance (Fig 3)
- Syllabic units rival the performance of morphological units (Fig 3)
- Improving morphological accuracy beyond a certain level does not translate into improved KWS performance (Figs 3 & 4)
- Adding phonetic information improves morphological segmentation (Fig 4)
- Combining morphological, phonetic and syllabic segmentations provides better results than either in isolation (Fig 5)
- Morphological segmentation helps KWS across different languages (Table 3)

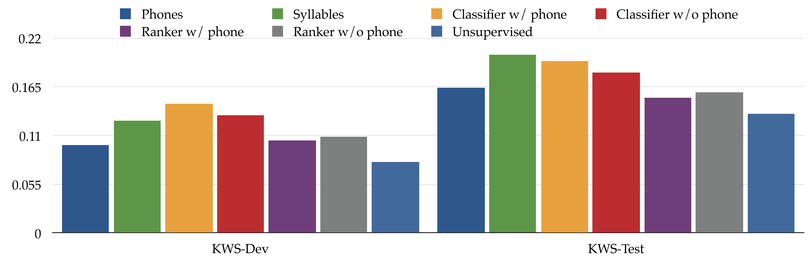


Fig 3: ATWV scores on Babel Turkish data

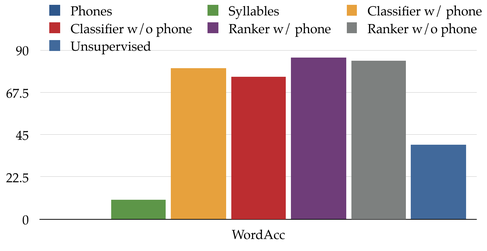


Fig 4: Segmentation Accuracy on MC2010

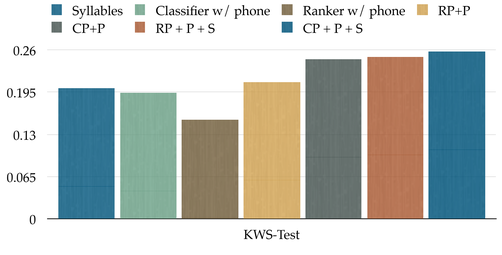


Fig 5: ATWV scores for combination systems

	Assamese		Bengali		Haitian		Lao		Tamil		Zulu	
	Dev	Test	Dev	Test	Dev	Test	Dev	Test	Dev	Test	Dev	Test
P + S	0.213	0.230	0.277	0.296	0.371	0.342	0.228	0.139	0.349	0.267	0.279	0.215
P + S + U	0.214	**0.263**	**0.294**	**0.328**	**0.393**	0.342	0.237	0.146	**0.395**	**0.284**	0.275	0.218

Table 3: ATWV scores for Phone+Syllable (P+S) and Phone+Syllable+Unsupervised Morphemes (P+S+U)

4. Modified KWS Pipeline

Morphemes are added to words in vocabulary → Training lexicon is segmented based on chosen morphemes → Decoder produces lattices containing both morphemes and words → Lattices converted to confusion networks consisting of morphemes only → Keyword search is done by breaking up keywords into morphemes

This work was supported by the Intelligence Advanced Research Projects Activity (IARPA). Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements of IARPA, DoD/ARL, or the U.S. Government