Leveraging supplemental representations for sequential transduction

Aditya Bhargava\textsuperscript{1} \quad Grzegorz Kondrak\textsuperscript{2}

\textsuperscript{1}Department of Computer Science
University of Toronto

\textsuperscript{2}Department of Computing Science
University of Alberta

NAACL-HLT 2012
Pronunciation-based tasks

orthography
Dickens

transliterations

dिकेंस
ディケンズ
Диккенс
Ντίκενς

transcriptions
/dɪkɪnz/
dIkInz
D IH K AH N Z
dIk@nz
d I k x n z
Pronunciation-based tasks

Orthography

Dickens

Transliterations

ディケンズ
Диккенс
Ντίκενς

Transcriptions

/dɪkɪnz/
D IH K AH N Z

dIkInz
dI k@nz
d I k x n z
Pronunciation-based tasks
Pronunciation-based tasks

orthography
Dickens
.ReadFile
orthography
transliterations
transcriptions

MTL
BTL

SR
TTS

ГИК AH Н Z
dIk@nz
d Ik x n z

/dııkınz/
dIkInz

D IH K AH N Z

ディケンズ
Dickens

Диккенс

Нтίкевς
Overview

- $x$ supplemental data for $y$
  - $x \in \{\text{transcription, transliteration}\}$
  - $y \in \{\text{G2P, MTL}\}$

- Rerank outputs from existing system
  - Features similar to base system, but applied to supplemental data
  - $n$-grams, alignment/similarity scores

- Same approach for system combination
  - Use another G2P/MTL system’s outputs as supplemental data
Excellent results
- Up to 8.7% error reduction for system combination
- MTL sees error reduction up to 14% from transliterations and 18% from transcriptions
- G2P sees error reduction up to 43% from transcriptions
Excellent results

- Up to 8.7% error reduction for system combination
- MTL sees error reduction up to 14% from transliterations and 18% from transcriptions
- G2P sees error reduction up to 43% from transcriptions
- But transliterations help G2P for names only
Excellent results *(mostly)*

- Up to 8.7% error reduction for system combination
- MTL sees error reduction up to 14% from transliterations and 18% from transcriptions
- G2P sees error reduction up to 43% from transcriptions
- But transliterations help G2P for names only
Reranking method

- From ACL 2011
- Looks specifically at transliterations as supplemental data for G2P of names
  - Names are hard(er)
  - Transliteration is generally applied to named entities
  - Encodes relevant pronunciation information
- Using supplemental data, rerank $n$-best output list of G2P base system
- Additional findings:
  - Simple similarity-based methods don’t work
  - Multiple languages are helpful
Here, we experiment with:

1. Transcriptions as supplemental data for both G2P and MTL
2. Transcriptions and transliterations simultaneously
3. G2P in general, rather than names only
4. System combination as supplemental data
Here, we experiment with:

1. Transcriptions as supplemental data for both G2P and MTL
2. Transcriptions and transliterations simultaneously
3. G2P in general, rather than names only
4. System combination as supplemental data
Here, we experiment with:

1. Transcriptions as supplemental data for both G2P and MTL
2. Transcriptions and transliterations simultaneously
3. G2P in general, rather than names only
4. System combination as supplemental data
Reranking method

Here, we experiment with:

1. Transcriptions as supplemental data for both G2P and MTL
2. Transcriptions and transliterations simultaneously
3. G2P in general, rather than names only
4. System combination as supplemental data
Here, we experiment with:

1. Transcriptions as supplemental data for both G2P and MTL
2. Transcriptions and transliterations simultaneously
3. G2P in general, rather than names only
4. **System combination** as supplemental data
Related work

- **G2P systems**
  - Neural networks, instance-based learning, ...
  - ..., joint \( n \)-gram models (Sequitur), online discriminative learning (DirecTL+)

- **MTL systems**
  - Similarly many approaches
  - Lately Sequitur and DirecTL+ have performed quite well at NEWS
Using heterogeneous data
- Pivot through a third language for transliteration
- Mostly useful for low-resource environments
- Hard to incorporate more languages
- Linear combination of system scores
input word

Sudan
input word

Sudan → base system
Method

input word
Sudan → base system → n-best outputs
sud@n
sud{n
⁞
sud#n
Method
Method

input word
Sudan

n-best outputs
base system
sud@n
sud{n
⁞
sud#n
supplemental representations
sudAn
S UD D AE N
スーダン
सूडान
Судан
⁞
re-ranker
re-ranked n-best list
sud#n
sUd#n
⁞
sud@n
Method

input  Gershwin

\textit{n-best outputs}  \textipa{dʒɪzmɪn/}  \textipa{gɜːʃwɪn/}  \ldots  \textipa{dʒɜːʃwɪn/}
Method

input

Gershwin

n-best outputs  /dʒɜːʃwɪn/  /ɡɜːʃwɪn/  • • •  /dʒɛɹʃwɪn/

transliterations  गर्शविन  ガーシュウィン  • • •  Гершвин
( /ɡarʃwɪn/ ) ( /ɡaːʃuwin/ ) ( /ɡerʃvin/ )
Method

input  Gershwin

\(n\)-best outputs  /dʒɜːʃwɪn/  /ɡɜːʃwɪn/  \(\ldots\)  /dʒɛɹʃwɪn/

transliterations  गर्शविन  ガーシュウィン  \(\ldots\)  Гершвин

(/ɡarʃvɪn/)  (/gaːɕuwin/)  \(\ldots\)  (/ɡɛʃvɪn/)
Method

**Input**: Gershwin

**n-best outputs**: /dʒɜːʃwɪn/ /ɡɜːʃwɪn/ • • • /dʒɛɹʃwɪn/

**Transliterations**: गश्विन गाशुविन • • • गरश्विन

(/ɡʌrʃʋɪn/) (/ɡaː ɕuwiɴ/) (/ɡerʂvin/)
Method

input

Gershwin

n-best outputs

/gɜːʃwɪn/ /dʒɜːʃwɪn/ /dʒɛɹʃwɪn/

transliterations

गशिवन ガーシュウィン Гершвин

(⁄gʌrʃwɪn/) (⁄ɡaː ɕuwiɴ/) (⁄ɡerʂwɪn/)
Method

\[\text{input} \quad \text{Gershwin} \]

\[n\text{-best outputs} \quad /\text{dʒɜːʃwɪn}/ \quad /\text{ɡɜːʃwɪn}/ \quad \cdots \quad /\text{dʒɛɹʃwɪn}/\]

\[\text{transliterations} \quad \text{गर्शविन} /\text{ɡarʃwɪn}/ \quad \text{ガーシュウィン} /\text{ɡaːɕuwin}/ \quad \cdots \quad \text{Гершвин} /\text{ɡerʃvin}/\]
Method

input

\[ n\text{-best outputs} \]

\[ /dʒɜːʃwɪn/ \quad /dʒɛɹʃwɪn/ \quad \cdots \quad /dʒɜːʃwɪn/ \]

transliterations

\[ गर्शविन \quad गाश्चविन \quad \cdots \quad गेर्शविन \]

\[ (/gəɾʃvɪn/) \quad (/gəʃuɯ̃vɪn/) \quad \cdots \quad (/ɡerʂvɪn/) \]
Method

input

n-best outputs

transliterations

Gershwin

/ðʒɛɹʃwɪn/

/ɡɜːʃwɪn/

/ɡɜːʃwɪn/

/ɡaː ɕuwiɴ/

/ɡʌrʃʋɪn/

/ɡɜːʃwɪn/

/ɡɛɹʃwɪn/

गशिवन

ガーシュウィン

Гершвин

गर्शविन

ガーチュウィン

Гершвин

(/ɡarʃvin/)

(/ɡaːɕuwin/)

(/ɡeɾʃvin/)
Data and base systems

- Transcriptions from Combilex and CELEX
- Transliterations from NEWS 2011
  - Experiment on English-to-Japanese transliteration
- 80/10/10 train/dev/test split
- Sequitur and DirecTL+ as base systems
<table>
<thead>
<tr>
<th>input</th>
<th>candidate outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGee</td>
<td>m@kJi</td>
</tr>
<tr>
<td></td>
<td>m@gi</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>m@CJi</td>
</tr>
</tbody>
</table>
### G2P experiments

**Supplemental transliterations**

<table>
<thead>
<tr>
<th>input</th>
<th>candidate outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGee</td>
<td>m@kJi</td>
</tr>
<tr>
<td></td>
<td>m@gi</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>m@CJi</td>
</tr>
</tbody>
</table>
G2P experiments
Supplemental transliterations

<table>
<thead>
<tr>
<th>input</th>
<th>candidate outputs</th>
<th>supplemental</th>
</tr>
</thead>
<tbody>
<tr>
<td>McGee</td>
<td>m@kJi m@gi ... m@CJi</td>
<td>मगी मギー Макги</td>
</tr>
</tbody>
</table>
G2P experiments: names
Supplemental transliterations

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Sequitur**
  - Base: 70%
  - Reranked: 80%

- **DirecTL+**
  - Base: 80%
  - Reranked: 90%
G2P experiments: full set
Supplemental transliterations

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td>70</td>
<td>80</td>
</tr>
<tr>
<td>DirecTL+</td>
<td>80</td>
<td>90</td>
</tr>
</tbody>
</table>
G2P experiments: core vocab

Supplemental transliterations

Word accuracy (%)

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sequitur

DirecTL+
<table>
<thead>
<tr>
<th>input</th>
<th>candidate outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sudan</td>
<td>sud@n</td>
</tr>
<tr>
<td></td>
<td>sud{n</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>sud#n</td>
</tr>
<tr>
<td>Input</td>
<td>Candidate Outputs</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Sudan</td>
<td>sud@n</td>
</tr>
<tr>
<td></td>
<td>sud{n</td>
</tr>
<tr>
<td></td>
<td>sud#n</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
</tbody>
</table>
G2P experiments
Supplemental transcriptions

(input)
Sudan

candidate outputs
sud@n
sud{n
...
sud#n

(CELEX)

(supplemental)
sudAn
G2P experiments: baselines
Supplemental transcriptions

MERGE

1. Convert Combilex to CELEX
2. Merge with CELEX
3. Train on combined set
G2P experiments: baselines
Supplemental transcriptions

- **MERGE**
  1. Convert Combilex to CELEX
  2. Merge with CELEX
  3. Train on combined set

- **P2P: phoneme-to-phoneme converter**
  1. Intersect Combilex and CELEX
  2. Train a transduction system to convert Combilex to CELEX
  3. If a test word appears in Combilex, grab it from there and convert it to CELEX format
G2P experiments
Supplemental transcriptions: results

![Bar chart showing word accuracy for different models: Sequitur and DirecTL+.](chart.png)
G2P experiments
Supplemental transcriptions: results

Word accuracy (%)

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>MERGE</th>
<th>P2P</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td></td>
<td></td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>DirecTL+</td>
<td></td>
<td>70</td>
<td></td>
<td>90</td>
</tr>
</tbody>
</table>
G2P experiments
Supplemental transcriptions: results

![Graph showing word accuracy for Sequitur and DirecTL+ with different methods: Base, MERGE, P2P, and Reranked.](image)
G2P experiments
Supplemental transcriptions: results

Word accuracy (%)

- Base
- MERGE
- P2P
- Reranked

Sequitur

DirecTL+

0 70 80 90 100 Word accuracy (%)
<table>
<thead>
<tr>
<th>input</th>
<th>candidate outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>John Petrucci</td>
<td>जॉन पेटरूची</td>
</tr>
<tr>
<td></td>
<td>जॉन पटरूक्सी</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>जॉन पटरूक्सी</td>
</tr>
<tr>
<td>input</td>
<td>candidate outputs</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------</td>
</tr>
<tr>
<td>John Petrucci</td>
<td>जॉन पटरूसी</td>
</tr>
<tr>
<td></td>
<td>जॉन पटरूची</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>जॉन पटरूक्सी</td>
</tr>
</tbody>
</table>
### MTL experiments
Supplemental transliterations

<table>
<thead>
<tr>
<th>input</th>
<th>candidate outputs</th>
<th>supplemental</th>
</tr>
</thead>
</table>
| John Petrucci | जॉन पटरूसी  
जॉन पटरूची | जॉन पेट्रुची  
Джон Петруччи |
MTL experiments
Supplemental transliterations

Wikipedia example

- John Petrucci article exists in English & Japanese, but not Hindi
- Want to automatically generate stub article in Hindi
  - Need transliteration of name
- Start from English, use Japanese (etc.) TLs to help generate Hindi TL
MTL experiments
Supplemental transliterations: results

Word accuracy (%)

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>DirecTL+</td>
<td>50</td>
<td></td>
</tr>
</tbody>
</table>
MTL experiments
Supplemental transliterations: results

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td>40</td>
<td>60</td>
</tr>
<tr>
<td>DirecTL+</td>
<td>50</td>
<td>70</td>
</tr>
</tbody>
</table>

Word accuracy (%)
<table>
<thead>
<tr>
<th>input</th>
<th>candidate outputs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sudan</td>
<td>ズーダン</td>
</tr>
<tr>
<td></td>
<td>スーダン</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>スユーダン</td>
</tr>
<tr>
<td>input</td>
<td>candidate outputs</td>
</tr>
<tr>
<td>------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Sudan</td>
<td>スーダン</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>スユーダン</td>
</tr>
<tr>
<td>input</td>
<td>candidate outputs</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Sudan</td>
<td>ズーダン</td>
</tr>
<tr>
<td></td>
<td>スーダン</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>スユーダン</td>
</tr>
</tbody>
</table>
MTL experiments
Supplemental transcriptions: results

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DirecTL+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Word accuracy (%)
MTL experiments
Supplemental transcriptions: results

Word accuracy (%)

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td>60</td>
<td>▲ 70</td>
</tr>
<tr>
<td>DirecTL+</td>
<td>50</td>
<td>▲ 60</td>
</tr>
</tbody>
</table>
Method works across base systems, but magnitude of improvement varies
Sequitur sees higher improvements
1. Lower base score
2. Higher oracle reranker score
3. Reranking features are similar to those used in DirecTL⁺
Analysis

- Feature similarity indicates DirecTL+’s improvement comes from the supplemental representations, not new features
Using transcriptions and transliterations simultaneously doesn’t provide any additional benefit
System combination

cirino
System combination

cirino

DirecTL+ → Cirino

シリーノ
チリーノ
シリノ
チリノ
キリノ
...
チシーリノ
System combination

cirino

DirecTL+ → leranker

シリーノ
チリーノ
シリノ
チリノ
キリノ
...
チシーリノ
System combination

cirino

DirecTL+ → reranker

Sequitur
System combination

![Diagram showing system combination]

- Cirino
- DirecTL+
- Sequitur
- Reranker

Languages:
- Cirino
- DirecTL+
- Sequitur
- Reranker

Languages in cirino:
- Cirino
- Sequitur

Languages in DirecTL+:
- Cirino
- Sequitur

Languages in Sequitur:
- Cirino
- DirecTL+

Languages in Reranker:
- Cirino
- DirecTL+
- Sequitur
System combination
System combination
System combination
Baseline

- Linear combination baseline
  - Merge the base system lists
  - Linearly combine system scores
  - Manually tune linear parameter on training data
System combination

Result

<table>
<thead>
<tr>
<th>System</th>
<th>Word accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td>Base</td>
</tr>
<tr>
<td>DirecTL+</td>
<td>40</td>
</tr>
</tbody>
</table>
System combination

Results

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>LinComb</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DirecTL+</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Word accuracy (%)
System combination
Results

<table>
<thead>
<tr>
<th></th>
<th>Base</th>
<th>LinComb</th>
<th>Reranked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequitur</td>
<td>29 / 31</td>
<td>▲</td>
<td></td>
</tr>
<tr>
<td>DirecTL+</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

- Reranking approach effectively leverages supplemental transcriptions and transliterations for G2P and MTL.
- Improvements across two base systems demonstrates that there is inherently useful information in the supplemental representations.
- Treating another system’s output as supplemental data works, but so does a linear combination.
Future work

- Reranking is post hoc; direct integration might be more effective
- Incorporate supplemental information rather than data
- Other (noisy?) supplemental sources
  - Wikipedia IPA transcriptions
  - Ad hoc approximately-phonetic re-spellings