

# Optimized Chinese Pronunciation Prediction by Component-Based Statistical Machine Translation

Shunle Zhu\*

## Abstract

To eliminate ambiguities in the existing methods to simplify Chinese pronunciation learning, we propose a model that can predict the pronunciation of Chinese characters automatically. The proposed model relies on a statistical machine translation (SMT) framework. In particular, we consider the components of Chinese characters as the basic unit and consider the pronunciation prediction as a machine translation procedure (the component sequence as a source sentence, the pronunciation, pinyin, as a target sentence). In addition to traditional features such as the bidirectional word translation and the n-gram language model, we also implement a component similarity feature to overcome some typos during practical use. We incorporate these features into a log-linear model. The experimental results show that our approach significantly outperforms other baseline models.

## Keywords

Chinese Pronunciation Prediction, Component, Features, Statistical Machine Translation (SMT)

## 1. Introduction

Chinese characters are widely used when writing Chinese and other Asian languages. In standard Chinese, they are called Hanzi (simplified Chinese: 汉字). Except for Chinese use in China, they have been adapted to write various other languages such as Japanese and Vietnamese. Modern Chinese have several homophones. Therefore, the same spoken syllable may be represented by more than one character, based on the meaning. Additionally, cognates in several varieties of Chinese are generally written with the same Chinese character, and they typically have similar meanings but often quite different pronunciations [1]. Therefore, it is challenging for a beginner to learn the pronunciation of a Chinese character, especially many uncommon words, such as “饕,” “霏,” “犇,” and “淼,” not to mention sharing the same spoken syllable with others; for example, “阿(a)姨,” “东阿(e),” “西藏(zang),” “储藏(cang),” puzzling many beginners.

If you are a native English speaker and know nothing about other languages written in the Latin alphabet, you can still read some words [2]. The reason is that English and these languages share most characters. Moreover, the same Latin character usually has a similar pronunciation (Table 1).

Empirically, someone may think that if there are similar components in two words (characters), they

※ This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Manuscript received August 31, 2018; first revision November 14, 2018; second revision March 12, 2019; third revision May 9, 2019; accepted May 29, 2019.

Corresponding Author: Shunle Zhu ([zhushunle@126.com](mailto:zhushunle@126.com))

\* Donghai Science and Technology College, Zhejiang Ocean University, Zhoushan, Zhejiang, China ([zhushunle@126.com](mailto:zhushunle@126.com))

should have similar pronunciation. Unfortunately, this is not always true in Chinese (Table 2). The pronunciation of Chinese characters relies heavily on other components and their surroundings.

Accordingly, there are two main challenges in Chinese pronunciation prediction:

- (1) Several polyphones exist in China [3]: Polyphones in this paper refer to Chinese characters with more than one pronunciation. In Chinese, there are more than 1,000 polyphones. For example, “和” can be pronounced as “he1,” “he2,” “hu2,” “huo2,” and “huo4”; “还” can be pronounced as “huan2” and “hai2.”
- (2) Several character components are distorted or changed in form to fit into a block with other components: Therefore, the actual shape of the component when used in a character depends on its placement relative the other elements in the character (its context information). For example, the Chinese characters “池,” “驰,” “弛,” “地,” “他,” and “她,” they all take “也” as their phonetic component; however, none of them take the pronunciation of “也(ye3).” This was caused by the simplification in the characters “她” and “他,” as their phonetic parts were originally “它(ta1).”

**Table 1.** Examples of English and French pronunciation

English		French	
Word	Pronunciation	Word	Pronunciation
President	[ <sup>1</sup> prezɪdənt]	président	[prezidɑ̃]
Restaurant	[ <sup>1</sup> restrɒnt]	restaurant	[rɛstɔrɑ̃]
Piano	[pi <sup>1</sup> ænəʊ]	Piano	[pjano]
Government	[ <sup>1</sup> gʌvənmənt]	gouvernement	[guvɛrnɑ̃]
Expert	[ <sup>1</sup> ekspɜ:t]	Expert	[eksper, -t]
Flower	[ <sup>1</sup> flaʊə(r)]	Fleur	[flœr]

**Table 2.** Examples of similar components with different pronunciation

Component	Pronun1	Pronun2	Pronun3	Pronun4
云	天坛(tan)	运气(yun)	昙花(tan)	动(dong)作
月	朋(peng)友	钥(yao)匙	刖(yue)足	衣服(fu)
木	森林(lin)	花朵(duo)	杀(sha)害	杂(za)技
日	元旦(dan)	南昌(chang)	水晶(jing)	黄昏(hun)

There are relatively few studies on this topic. In this regard, in [4], a method based on the Bishun (order of strokes) of Chinese characters to predict pronunciation was proposed. An important disadvantage of this approach is that any prior knowledge is not considered (for example, whether the learner knows the basic pronunciation of some Chinese characters, which can be very helpful in pronunciation prediction). More than 80% of Chinese characters are composed of a semantic (meaning) component and a phonetic (sound) component. Phonetic components are elements in a Chinese character providing clues on a character’s pronunciation. They can be used to deduce the intonation of an unknown Chinese character. Learning the pronunciation of common sound components is crucial. The main contribution of this paper is that we attempt to predict the pronunciation of Chinese characters based on components. We considered the Chinese pronunciation prediction as a machine translation problem [5-7]. The proposed approach relies on an SMT framework; the source part is a sequence of components, and the target part is a pinyin sequence. In addition to traditional features such as the bidirectional word translation and the

n-gram language model, we also include a component similarity feature to overcome some errors existing in the component-based approach. We combine these features with a log-linear model. The experimental results show that the proposed model significantly improves the performance compared to several baseline models in all tasks.

## 2. Background

First, we briefly introduce the Chinese character component, which is the most important topic in this paper. Then, we use components to represent Chinese characters. We also present some previous studies related to this study.

### 2.1 Component of Chinese

English words are organized by letters, and Chinese words are often classified according to their components. There are approximately 214 components in Chinese. A Chinese component is a graphical component of Chinese characters. These components are often semantic indicators such as a phonetic component or even an artificially extracted portion of the character. Examples of Chinese components are listed in Table 3.

The functions of the components are: (1) To indicate the meaning of a Chinese character such as “is made of metal” (“铜,” “铁,” “银”), “is one kind of bird” (“鹊,” “鹞,” “鹅”), “is for female” (“妈,” “嫁,” “妇”), and (2) To search Chinese characters in a dictionary.

**Table 3.** Examples of components in Chinese characters

Character	Semantic component	Phonetic component
河	氵	可
坡	土	皮
灿	火	山
蚁	虫	义
猴	犛	侯
框	木	匡

### 2.2 Previous Studies

In the field of Chinese pronunciation prediction, a system for foreigners to speak Chinese was proposed in [8]. This system uses phonetic spelling in the foreigner’s own orthography to represent the input text. In [9], a generative model was proposed based on existing dialect pronunciation data and medieval rime books to discover phonological patterns in multiple dialects. Moreover, a Bishu-based Chinese pronunciation prediction model was proposed in [10].

In [10] and [11], the pronunciation of Japanese was predicted based on a phrasal statistical machine translation (STM) model, which combined word and character-based pronunciation from a dictionary within an SMT framework to handle OOV words.

This study is closely related to [11] and [4], which are also based on an SMT framework. However, the

proposed method has two significant differences: (1) our model is component-based, and (2) our study is Chinese learner-oriented; and we mainly studied Chinese.

### 3. Proposed Method

First, we introduce in this section the details of the proposed pronunciation prediction model. After that, we describe some additional features that can further optimize the proposed model performance. Finally, we describe how the model is trained.

#### 3.1 Definition

In this paper, we consider the Chinese pronunciation prediction as an SMT problem. We consider the component as the basic unit in the SMT. Given a source sentence (component sequence)  $x = \{x_1, x_2, x_3, \dots, x_l\}$  ( $x_i$  is a component of the current Chinese character) and a target sentence (pinyin sequence)  $y = \{y_1, y_2, y_3, \dots, y_m\}$ , we can reformulate the translation probability to predict the pronunciation of a Chinese word  $w$  given a component sequence  $x$  according to Bayes's rule as follows:

$$p(y|x) = \arg \max_{y'} p(x|y)p(y) \quad (1)$$

where  $p(y)$  is the language model and  $p(x|y)$  is the translation probability.

#### 3.2 Method

To optimize the proposed method performance, we proposed a log-linear framework [12] to integrate several effective features. In [13], a log-linear framework was introduced into an SMT to integrate eight features, such as the bidirectional translation probabilities  $p(f|e)$  and  $p(e|f)$ , bidirectional lexical weights  $p_{lex}(f|e)$  and  $p_{lex}(e|f)$ , language model, reordering model, word penalty, and phrase penalty to improve the performance of phrase-based translation models.

Unlike previous studies, we do not require all features defined in standard phrase-based SMT. Instead, we select some features, such as bidirectional word translation, language model, and extend features such as component similarity features to adapt our Chinese pronunciation prediction task.

##### 3.2.1 Feature definition

We use the following three features in the proposed approach:

###### **Bidirectional word translation feature (BWT)**

In our model, we estimate the word translation probabilities between the target candidates and the corresponding source components from both directions as follows:

$$\text{Target (pinyin)} \rightarrow \text{Source (component)}: H_{wtp1} = \sum_{j=1}^J \sum_{i=1}^i \alpha_{ji} \log(p(y_i|x_i)) \quad (2)$$

$$\text{Source (component)} \rightarrow \text{Target (pinyin)}: H_{wtp2} = \sum_{j=1}^J \sum_{i=1}^i \alpha_{ji} \log(p(x_i|y_j)) \quad (3)$$

Here,  $p(y|x)$  and  $p(x|y)$  represent the word translation probabilities, estimated from the word-aligned bilingual corpus, where the word alignment model is trained based on the open source tool GIZA++ [14]. The results are combined with the “*grow-diag-final-and*” method. Therefore,  $p(y|x)$  and  $p(x|y)$  can be computed as follows:

$$p(x|y) = \frac{N(x,y)}{\sum_{x'} N(x',y)} \quad (4)$$

$$p(y|x) = \frac{N(y,x)}{\sum_{y'} N(y',x)} \quad (5)$$

where  $N(x, y)$  is the number of times component  $x$  and pinyin  $y$  co-occurrence.

### N-gram language model feature (NLM)

N-gram is a contiguous sequence of  $n$  items from a given text. An  $n$ -gram [14] models natural language sequences using the statistical properties of  $n$ -grams. Practically, an  $n$ -gram model predicts  $y_i$  based on  $y_{i-(n-1)}, \dots, y_{i-1}$ . This can be indicated as probability terms as follows:

$$H_{lm} = \sum_{j=1}^J \log(p(y|y_{i-(n-1)}, \dots, y_{i-1})) \quad (6)$$

When used in language modeling, independent assumptions are made so that each item word replies to its last  $n - 1$  words.

The language model is trained on a monolingual corpus and used in the proposed pronunciation prediction model to ensure the fluency of the output pinyin sequences. Due to LM, the language model feature allows us to use a large-scale monolingual corpus of the target language (pinyin sequence).

### Component similarity feature (CS)

Our study is user-oriented; thus, we propose string similarity and hybrid language model features to overcome some input errors existing in Chinese learners.

Usually, we may make some Chinese character writing mistakes, especially for people using Chinese as a second language. We considered a component similarity feature based on the Levenshtein distance algorithm [15].

$$H_{lev} = \min \begin{cases} D(i-1, j) + del[x(i)] \\ D(i, j-1) + ins[x'(j)] \\ D(i-1, j-1) + sub[x(i), x'(j)] \end{cases} \quad (7)$$

In the proposed model, we define the component similarity feature as the Levenshtein distance between two components. We compute the minimum number of single-character editing (insertions, deletions, or substitutions) required to change one component into another.

## 3.3 Model

In this study, we propose a component-based Chinese pronunciation prediction model, which relies on an SMT framework. To further optimize the performance of the proposed model, we suggest several features. The log-linear model used in our proposed method can be formulated as follows:

$$p(E|F) = \frac{\exp(\sum_{i=1}^m \lambda_i H_i(F, E))}{\sum_{E'} \exp(\sum_{i=1}^m \lambda_i H_i(F, E'))} \quad (8)$$

where  $H_i(F, E)$  is a feature function defined in this section, and  $\lambda_i$  is the weight of  $H_i(F, E)$ . We extend the decoder to perform pronunciation under the log-linear framework. The weights of the log-linear models are tuned using the minimum error rate training (MERT) algorithm [16].

## 4. Experiments

This section presents the experiments conducted to evaluate the proposed model performance for Chinese pronunciation prediction. Moreover, we compare our approach with several existing methods.

### 4.1 Data and Settings

To evaluate the potential of the proposed approach, we conducted experiments on the Bilingual Learning corpus, which we collect from several websites. The corpus contains 160K sentences in total. We selected these datasets from the Sougou Lab’s corpus (<http://www.sougou.com/labs/>); thus, these datasets are publicly accessible. We only need to transform these corpora into a specific format with some open-source tools. We first annotated these sentences with their corresponding pinyin sequences based on an open-source toolkit pinyin4j (<http://pinyin4j.sourceforge.net/>). Then, we divide these “bilingual” sentences into three datasets: training set (150K sentences/22.5M characters), develop set (2K sentences/0.3M characters), and test set (8K sentences/1.2M characters). We converse all Chinese characters into components, according to the rules extracted from Baidu Chinese (百度汉语) (<http://hanyu.baidu.com>). The average length of a component sequence in these data sets was 8. Additionally, to evaluate the domain adaption of our model, we also provide test sets on five different domains, such as Government Document (8K sentences/1.3M characters), Tourism (8K sentences/1.1M characters), Named Entities (8K sentences/0.9M characters), Weixin (8K sentences/0.8M characters), and Dialog (8K sentences/0.7M characters). We classify these as two groups, one is -ph, and the other is +ph. The groups are similar except that we annotate phonetic tones on the +ph group. We train the proposed pronunciation prediction model using Moses [17]. Specifically, we train a 5-gram language model (Table 4). To achieve the best performance, we select the n-gram equal to 5 according to a group of experiments. We use the same data sets as in pronunciation prediction experiments. As shown in Table 4, we can observe that when we set n=5, our model achieves the best performance both in -ph and +ph. on the whole pinyin part of the corpus using the SRILM [18] with the modified Kneser-Ney Smooth algorithm. We use the MERT [16] to optimize the feature weights on the developed set. Finally, we evaluate the performance of the models, considering the accuracy and rate of out-of-vocabulary (OOV).

In the experiments, we compared the proposed method with two baseline models. The first was the Bishun based Chinese Pronunciation Prediction Model (BcPPM) [4]. The BcPPM model also considers pronunciation prediction as an SMT problem. The model uses Bishun as the basic unit; features such as the global language and local language models were integrated into it.

Another baseline model used in our experiments was KyTea, an open-source Japanese word segmentation and pronunciation prediction tool. The model achieves state-of-the-art performance on the task of Japanese pronunciation prediction. We extend this tool to adapt it to our task.

**Table 4.** Evaluation on different n-grams

	+ph		-ph	
	Accuracy (%)	Out-of-vocabulary (%)	Accuracy (%)	Out-of-vocabulary (%)
n=1	78.04	6.9	79.76	6.4
n=2	79.10	6.2	80.88	4.1
n=3	82.94	2.9	83.56	3.1
n=4	84.03	2.3	84.90	2.2
n=5	85.20	1.7	85.24	1.6
n=6	83.95	3.2	84.52	2.1
n=7	80.25	4.2	82.77	3.8
n=8	79.23	5.9	80.59	5.3
n=9	78.50	6.7	79.72	6.5
n=10	78.14	6.8	79.01	6.6

## 4.2 Results and Analysis

To evaluate the model performance completely, we conducted several groups of experiments and compared the proposed method with other existing models.

Tables 5 and 6 list the experimental results for different models. In Table 5, the proposed model outperforms all baseline models. BcPPM cannot achieve better results than the proposed method. The reason is that our approach selects the component as the basic unit, which is more meaningful than Bishun. Moreover, we use a log-linear model to combine features, while the BcPP only uses several features empirically. KyTea performs poorly, which implies that using a unigram model does not work properly on our task. Table 6 also lists the performance of different models. In contrast to Table 5, Table 6 focuses on pronunciation prediction of polyphones. The proposed approach achieves the best results on this task, indicating that the features used in our model have a powerful disambiguate ability.

**Table 5.** Evaluation on different models (general)

Model	-ph		+ph	
	Accuracy (%)	Out-of-vocabulary (%)	Accuracy (%)	Out-of-vocabulary (%)
BcPPM [4]	80.50	4.5	80.24	4.7
KyTea [11]	82.18	2.0	82.09	2.2
The proposed model	85.24	1.6	85.20	1.7

**Table 6.** Evaluation on different models (polyphones)

Model	-ph	+ph
	Accuracy (%)	Accuracy (%)
BcPPM [4]	79.40	79.18
KyTea [11]	80.23	80.04
The proposed model	89.45	89.20

Table 7 lists the evaluation results on different features. We found that the proposed model performs poorly if only the BWT feature is used. The reason is that BWT relies only on the translation probability of words, which cannot reflect the fluency of the output sequence. The features combining BWT and LM

together outperform the BWT only feature. An important reason is that language model information is imported, and the LM features are especially useful when a Chinese character has more than one pronunciation (polyphone). The features (BWT+LM+CS) used in the proposed approach achieve the best results, which included not only the language model information but also the component similarity feature (CS feature). The CS is substantially powerful when some typos or errors are present in Chinese input characters.

Table 8 lists the experimental results on different test sets (domains) using the proposed method. We can conclude that our model performs much better on formal corpora (such as news reports, government documents, and commentaries used in tourism) than informal situations (such as Weixin and Dialog corpus). This may be due to two reasons: (1) the model was trained on a formal corpus and (2) the formal corpus was compared with informal texts that may include some typos or new words, which may never appear in the train set (OOV rate: Weixin 3.2% and Dialog 2.4%). The results on Named Entities (NE) outperform other domains significantly; the most important reason is that the NE corpus is relatively formal and close to the domain of our train set.

**Table 7.** Evaluation on several features

Model	-ph		+ph	
	Accuracy (%)	Out-of-vocabulary (%)	Accuracy (%)	Out-of-vocabulary (%)
BWT	60.36	9.3	60.16	9.5
BWT+LM	76.27	5.7	76.07	5.9
BWT+LM+CS	85.24	1.6	85.01	1.9

**Table 8.** Evaluation on different domains

Model	-ph		+ph	
	Accuracy (%)	Out-of-vocabulary (%)	Accuracy (%)	Out-of-vocabulary (%)
News (default)	85.24	1.6	85.06	1.8
Government document	86.57	0.9	86.41	1.1
Tourism	85.30	0.6	85.12	0.9
Named Entities	87.18	0.2	86.08	0.5
Weixin	84.80	3.2	82.96	3.6
Dialogue	83.62	2.4	82.01	2.6

From the above experimental results, we find that models considering phonetic tones (+ph) (there are four phonetic tones in Chinese, such as high-level tone, rising tone, falling-rising tone, and falling tone) perform poorly than those that did not consider this situation. The reason is that models trained based on phonetic tones annotated corpus may suffer from data sparseness during model training, which weakens the performance and increases the OOV rate.

## 5. Related Work

This study is mainly inspired by topics like pronunciation prediction and STM. In the field of Chinese pronunciation prediction, in [13], a system for foreigners to speak a language they do not know it was



proposed. This system uses phonetic spelling in the foreigner's own orthography to represent the input text. In [14], the authors presented a generative model based on existing dialect pronunciation data and medieval rime books to discover phonological patterns in multiple dialects. Moreover, a Bishun based Chinese pronunciation prediction model was proposed in [17].

The pronunciation of Japanese was predicted in [15] and [16] based on a phrasal STM model, which combined word and character-based pronunciation from a dictionary within an SMT framework to handle OOV words.

This study is closely related to [15] and [16], which are also based on the SMT framework. However, the proposed method has some significant differences: (1) our model is component-based, (2) our study is Chinese learner-oriented; and we mainly studied Chinese.

## 6. Conclusion

To simplify Chinese pronunciation learning for beginners, especially for foreigners, we proposed a component-based Chinese pronunciation prediction approach, which was based on an SMT framework that considered a component sequence as the source sentence and a pinyin sequence as the target sentence. In addition to traditional features such as the bidirectional word translation and the n-gram language model, we also included a component similarity feature to overcome some errors that exist in components. These features were combined using a log-linear model. Several groups of experiments were conducted to evaluate the performance of our proposed approach. The experimental results showed that the proposed model significantly outperforms several existing methods.

In future work, we plan to improve further the performance of our Chinese pronunciation prediction model from two aspects. (1) Investigating the unsupervised representation of components in Chinese to alleviate data sparseness during model training. (2) Analyzing errors existing in Chinese pronunciation prediction and developing new techniques to optimize the performance of the pronunciation prediction model.

## Acknowledgement

This work was supported by the Natural Science Foundation of Zhejiang (No. LY16F020014), Scientific Research Funding Project of Zhejiang Education Department (No. Y201840288), and Youth Natural Science Foundation of Zhejiang (No. LQ16A010003).

## References

- [1] S. K. Hsieh, "Hanzi, concept and computation: a preliminary survey of Chinese Characters as a Knowledge Resource in NLP," PhD dissertation, Universitat Tübingen, Tübingen, Germany, 2006.
- [2] R. J. Byrd and E. Tzoukermann, "Adapting an English morphological analyzer for French," in *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, Buffalo, NY, 1988, pp. 1-6.
- [3] F. L. Huang, S. Y. Ke, and Q. W. Fan, "Predicting effectively the pronunciation of Chinese polyphones by extracting the lexical information," in *Advances in Computer and Information Sciences and Engineering*. Dordrecht, Germany: Springer, 2008, pp. 159-165.

- [4] C. Mi, Y. Yang, X. Zhou, L. Wang, X. Li, and T. Jiang, "Exploiting Bishun to predict the pronunciation of Chinese," *Computación y Sistemas*, vol. 20, no. 3, pp. 541-549, 2016.
- [5] P. Koehn, F. J. Och, and D. Marcu, "Statistical phrase-based translation," in *Proceedings of Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics (HLT-NAACL)*, Edmonton, Canada, 2003.
- [6] R. Zens, F. J. Och, and H. Ney, "Phrase-based statistical machine translation," in *KI 2002: Advances in Artificial Intelligence*. Heidelberg, Germany: Springer, 2002, pp. 18-32
- [7] F. J. Och and H. Ney, "The alignment template approach to statistical machine translation," *Computational Linguistics*, vol. 30, no. 4, pp. 417-449, 2004.
- [8] X. Shi, K. Knight, and H. Ji, "How to speak a language without knowing it," in *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, Baltimore, MD, 2014, pp. 278-282.
- [9] C. C. Lin and R. T. H. & Tsai, "A generative data augmentation model for enhancing Chinese dialect pronunciation prediction," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 4, pp. 1109-1117, 2012.
- [10] J. Hatori and H. Suzuki, "Predicting word pronunciation in Japanese," in *Computational Linguistics and Intelligent Text Processing*. Heidelberg, Germany: Springer, 2011 pp. 477-492.
- [11] J. Hatori and H. Suzuki, "Japanese pronunciation prediction as phrasal statistical machine translation," in *Proceedings of 5th International Joint Conference on Natural Language Processing*, Chiang Mai, Thailand, 2011, pp. 120-128.
- [12] R. Christensen, *Log-Linear Models and Logistic Regression*. New York, NY: Springer, 2006.
- [13] F. J. Och and H. Ney, "Discriminative training and maximum entropy models for statistical machine translation," in *Proceedings of the 40th Annual meeting of the Association for Computational Linguistics*, Philadelphia, PA, 2002, pp. 295-302.
- [14] P. F. Brown, V. J. Della Pietra, P. V. Desouza, J. C. Lai, and R. L. Mercer, "Class-based n-gram models of natural language," *Computational Linguistics*, vol. 18, no. 4, pp. 467-480, 1992.
- [15] W. J. Heeringa, "Measuring dialect pronunciation differences using Levenshtein distance," Ph.D. dissertation, University Library Groningen, The Netherlands, 2004.
- [16] F. J. Och, "Minimum error rate training in statistical machine translation," in *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, Sapporo, Japan, 2003, pp. 160-167.
- [17] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, et al., "Moses: open source toolkit for statistical machine translation," in *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions*, Prague, Czech Republic, 2007, pp. 177-180.
- [18] A. Stolcke, "SRILM-an extensible language modeling toolkit," in *Proceedings of the 7th International Conference on Spoken Language Processing*, Denver, CO, 2002, pp. 901-904.



**Shunle Zhu** <https://orcid.org/0000-0003-1982-925X>

He is an associate professor in Donghai Science and Technology College, Zhejiang Ocean University. He received his master's degree in computer science from the Huazhong University of Science and Technology in 2006. His current research interests include natural language processing, signal processing, and wireless communication.