# Be Appropriate and Funny: Automatic Entity Morph Encoding

**Boliang Zhang[1], Hongzhao Huang[1], Xiaoman Pan[1], Heng Ji[1], Kevin Knight[2]**
**Zhen Wen[3], Yizhou Sun[4], Jiawei Han[5], Bulent Yener[1]**
[1]Computer Science Department, Rensselaer Polytechnic Institute
[2]Information Sciences Institute, University of Southern California
[3]IBM T. J. Watson Research Center
[4]College of Computer and Information Science, Northeastern University
[5]Computer Science Department, Univerisity of Illinois at Urbana-Champaign
[1]{zhangb8,huangh9,panx2,jih,yener}@rpi.edu, [2]knight@isi.edu
[3]zhenwen@us.ibm.com, [4]yzsun@ccs.neu.edu, [5]hanj@illinois.edu

## Abstract

Internet users are keen on creating different kinds of *morphs* to avoid censorship, express strong sentiment or humor. For example, in Chinese social media, users often use the entity morph "方便面 *(Instant Noodles)*" to refer to "周永康 *(Zhou Yongkang)*" because it shares one character "康 *(Kang)*" with the well-known brand of instant noodles "康师傅 *(Master Kang)*". We developed a wide variety of novel approaches to automatically encode proper and interesting morphs, which can effectively pass decoding tests [1].

## 1 Introduction

One of the most innovative linguistic forms in social media is *Information Morph* (Huang et al., 2013). Morph is a special case of alias to hide the original objects (e.g., sensitive entities and events) for different purposes, including avoiding censorship (Bamman et al., 2012; Chen et al., 2013), expressing strong sentiment, emotion or sarcasm, and making descriptions more vivid. Morphs are widely used in Chinese social media. Here is an example morphs: "由于瓜爹的事情，方便面与天线摊牌. (Because of *Gua Dad*'s issue, *Instant Noodles* faces down with *Antenna*.)", where

- "瓜爹 *(Gua Dad)*" refers to "薄熙来 *(Bo Xilai)*" because it shares one character "瓜 *(Gua)*" with "薄瓜瓜 *(Bo Guagua)*" who is the son of "薄熙来 *(Bo Xilai)*";
- "方便面 *(Instant Noodles)*" refers to "周永康 *(Zhou Yongkang)*" because it shares one character "康 *(kang)*" with the well-known instant noodles brand "康师傅 *(Master Kang)*";

- "天线 *(Antenna)*" refers to "温家宝 *(Wen Jiabao)*" because it shares one character "宝 *(baby)*" with the famous children's television series "天线宝宝 *(Teletubbies)*";

In contrast with covert or subliminal channels studied extensively in cryptography and security, *Morphing* provides confidentiality against a weaker adversary which has to make a real time or near real time decision whether or not to block a morph within a time interval $t$. It will take longer than the duration $t$ for a morph decoder to decide which encoding method is used and exactly how it is used; otherwise adversary can create a codebook and decode the morphs with a simple look up. We note that there are other distinct characteristics of morphs that make them different from cryptographic constructs: (1) Morphing can be considered as a way of using natural language to communicate confidential information without encryption. Most morphs are encoded based on semantic meaning and background knowledge instead of lexical changes, so they are closer to Jargon. (2) There can be multiple morphs for an entity. (3) The Shannon's Maxim "the enemy knows the system" does not always hold. There is no common code-book or secret key between the sender and the receiver of a morph. (4) Social networks play an important role in creating morphs. One main purpose of encoding morphs is to disseminate them widely so they can become part of the new Internet language. Therefore morphs should be interesting, fun, intuitive and easy to remember. (5) Morphs rapidly evolve over time, as some morphs are discovered and blocked by censorship and newly created morphs emerge.

We propose a brand new and challenging research problem - can we automatically encode morphs for any given entity to help users communicate in an appropriate and fun way?

---

[1]The morphing data set is available for research purposes: http://nlp.cs.rpi.edu/data/morphencoding.tar.gz

## 2 Approaches

### 2.1 Motivation from Human Approaches

Let's start from taking a close look at human's intentions and general methods to create morphs from a social cognitive perspective. In Table 1 and Table 2, we summarize 548 randomly selected morphs into different categories. In this paper we automate the first seven human approaches, without investigating the most challenging Method 8, which requires deep mining of rich background and tracking all events involving the entities.

### 2.2 M1: Phonetic Substitution

Given an entity name $e$, we obtain its phonetic transcription $pinyin(e)$. Similarly, for each unique term $t$ extracted from Tsinghua Weibo dataset (Zhang et al., 2013) with one billion tweets from 1.8 million users from 8/28/2012 to 9/29/2012, we obtain $pinyin(t)$. According to the Chinese phonetic transcription articulation manner [2], the pairs $(b, p)$, $(d, t)$, $(g, k)$, $(z, c)$, $(zh, ch)$, $(j, q)$, $(sh, r)$, $(x, h)$, $(l, n)$, $(c, ch)$, $(s, sh)$ and $(z, zh)$ are mutually transformable.

If a part of $pinyin(e)$ and $pinyin(t)$ are identical or their initials are transformable, we substitute the part of $e$ with $t$ to form a new morph. For example, we can substitute the characters of "比尔 盖茨 (Bill Gates) [Bi Er Gai Ci]" with "鼻耳 (Nose and ear) [Bi Er]" and "盖子 (Lid) [Gai Zi]" to form new morph "鼻耳 盖子 (Nose and ear Lid) [Bi Er Gai Zi]". We rank the candidates based on the following two criteria: (1) If the morph includes more negative words (based on a gazetteer including 11,729 negative words derived from HowNet (Dong and Dong, 1999), it's more humorous (Valitutti et al., 2013). (2) If the morph includes rarer terms with low frequency, it is more interesting (Petrovic and Matthews, 2013).

### 2.3 M2: Spelling Decomposition

Chinese characters are ideograms, hieroglyphs and mostly picture-based. It allows us to naturally construct a virtually infinite range of combinations from a finite set of basic units - radicals (Li and Zhou, 2007). Some of these radicals themselves are also characters. For a given entity name $e = c_1...c_n$, if any character $c_k$ can be decomposed into two radicals $c_k^1$ and $c_k^2$ which are both characters or can be converted into characters based on their pictograms (e.g., the radical "艹" can be converted into "草" (grass) ), we create a morph by replacing $c_k$ with $c_k^1 c_k^2$ in $e$. Here we use a character to radical mapping table that includes 191 radicals (59 of them are characters) and 1328 common characters. For example, we create a morph "人呆罗 (Person Dumb Luo)" for "保罗 (Paul)" by decomposing "保 (Pau-)" into "人 (Person)" and "呆 (Dull)". A natural alternative is to composing two chracter radicals in an entity name to form a morph. However, very few Chinese names include two characters with single radicals.

### 2.4 M3: Nickname Generation

We propose a simple method to create morphs by duplicating the last character of an entity's first name. For example, we create a morph "幂幂 (Mimi)" to refer to "杨 幂 (Yang Mi)".

### 2.5 M4: Translation and Transliteration

Given an entity $e$, we search its English translation $EN(e)$ based on 94,015 name translation pairs (Ji et al., 2009). Then, if any name component in $EN(e)$ is a common English word, we search for its Chinese translation based on a 94,966 word translation pairs (Zens and Ney, 2004), and use the Chinese translation to replace the corresponding characters in $e$. For example, we create a morph "拉里 鸟儿 (Larry bird)" for "拉里 伯德 (Larry Bird)" by replacing the last name "伯德 (Bird)" with its Chinese translation "鸟儿 (bird)".

### 2.6 M5: Semantic Interpretation

For each character $c_k$ in the first name of a given entity name $e$, we search its semantic interpretation sentence from the Xinhua Chinese character dictionary including 20,894 entries [3]. If a word in the sentence contains $c_k$, we append the word with the last name of $e$ to form a new morph. Similarly to M1, we prefer positive, negative or rare words. For example, we create a morph "薄 胡来 (Bo Mess)" for "薄熙来 (Bo Xi Lai)" because the semantic interpretation sentence for "来 (Lai)" includes a negative word "胡来 (Mess)".

### 2.7 M6: Historical Figure Mapping

We collect a set of 38 famous historical figures including politicians, emperors, poets, generals, ministers and scholars from a website. For a given entity name $e$, we rank these candidates by applying the resolution approach as described in our previous work (Huang et al., 2013) to measure the similarity between an entity and a historic figure

---

| Category | Frequency Distribution | Examples | | |
|---|---|---|---|---|
| | | Entity | Morph | Comment |
| (1) Avoid censorship | 6.56% | 薄熙来 (Bo Xi-lai) | B书记 (B Secre-tary) | "B" is the first letter of "Bo" and "Secretary" is the entity's title. |
| (2) Express strong sentiment, sarcasm, emotion | 15.77% | 王勇平 (Wang Yongping) | 奇迹哥 (Miracle Brother) | Sarcasm on the entity's public speech: "It's a miracle that the girl survived (from the 2011 train collision)". |
| (3) Be humorous or make descriptions more vivid | 25.91% | 杨幂 (Yang Mi) | 嫩牛五方 (Tender Beef Pentagon) | The entity's face shape looks like the shape of famous KFC food "Tender Beef Pentagon". |
| Mixture | 25.32% | 卡扎菲 (Gaddafi) | 疯鸭上校 (Crazy Duck Colonel) | Sarcasm on Colonel Gaddafi's violence. |
| Others | 23.44% | 蒋介石 (Chiang Kai-shek) | 花生米 (Peanut) | Joseph Stilwell, a US general in China during World War II, called Chiang Kai-shek "花生米 (Peanut)" in his diary because of his stubbornness. |

Table 1: Morph Examples Categorized based on Human Intentions

| No. | Category | Frequency Distribution | Example | | |
|---|---|---|---|---|---|
| | | | Entity | Morph | Comment |
| M1 | Phonetic Substitution | 12.77% | 萨科齐 (Sarkozy) | 傻客气 (Silly Polite) | The entity's phonetic transcript "Sa Ke Qi" is similar to the morph's "Sha Ke Qi". |
| M2 | Spelling Decomposition | 0.73% | 胡锦涛 (Hu Jintao) | 古月 (Old Moon) | The entity's last name is decomposed into the morph "古月 (Old Moon)"? |
| M3 | Nickname Generation | 12.41% | 江泽民 (Jiang Zemin) | 老江 (Old Jiang) | The morph is a conventional name for old people with last name "Jiang". |
| M4 | Translation & Transliteration | 3.28% | 布什 (Bush) | 树丛 (shrub) | The morph is the Chinese translation of "bush". |
| M5 | Semantic Interpretation | 20.26% | 金日成 (Kim Il Sung) | 金太阳 (Kim Sun) | The character "日" in the entity name means "太阳 (Sun)". |
| M6 | Historical Figure Mapping | 3.83% | 薄熙来 (Bo Xilai) | 平西王 (Conquer West King) | The entity shares characteristics and political experiences similar to the morph. |
| M7 | Characteristics Modeling | 20.62% | 金日成 (Kim Il Sung) | 金胖子 (Kim Fat) | "胖子 (Fat)" describes "金日成 (Kim Il Sung)"'s appearance. |
| M8 | Reputation and public perception | 26.09% | 奥巴马 (Obama) | 观海 (Staring at the sea) | Barack Obama received a calligraphy "观海听涛 (Staring at sea and listening to surf)" as a present when he visited China. |
| | | | 马景涛 (Ma Jingtao) | 咆哮教主 (Roar Bishop) | In the films Ma Jingtao starred, he always used exaggerated roaring to express various emotions. |
| | | | 马英九 (Ma Yingjiu) | 马不统 (Ma Secession) | The morph derives from Ma Yingjiu's political position on cross-strait relations. |

Table 2: Morph Examples Categorized based on Human Generation Methods

based on their semantic contexts. For example, this approach generates a morph "太祖 (the First Emperor)" for "毛泽东 (Mao Zedong)" who is the first chairman of P. R. China and "高祖 (the Second Emperor )" for "邓小平 (Deng Xiaoping )" who succeeded Mao.

## 2.8 M7: Characteristics Modeling

Finally, we propose a novel approach to automatically generate an entity's characteristics using Google word2vec model (Mikolov et al., 2013). To make the vocabulary model as general as possible, we use all of the following large corpora that we have access to: Tsinghua Weibo dataset, Chinese Gigaword fifth edition [4] which includes 10 million news documents, TAC-KBP 2009-2013 Source Corpora (McNamee and Dang, 2009; Ji et al., 2010; Ji et al., 2011; Ji and Grishman, 2011)

which include 3 million news and web documents, and DARPA BOLT program's discussion forum corpora with 300k threads. Given an entity $e$, we compute the semantic relationship between $e$ and each word from these corpora. We then rank the words by: (1) cosine similarity, (2) the same criteria as in section 2.6. Finally we append the top ranking word to the entity's last name to obtain a new morph. Using this method, we are able to generate many vivid morphs such as "姚奇才 (Yao Wizard)" for "姚明 (Yao Ming)".

## 3 Experiments

### 3.1 Data

We collected $1,553,347$ tweets from Chinese Sina Weibo from May 1 to June 30, 2013. We extracted 187 human created morphs based on M1-M7 for 55 person entities. Our approach generated 382

new morphs in total.

## 3.2 Human Evaluation

We randomly asked 9 Chinese native speakers who regularly access Chinese social media and are not involved in this work to conduct evaluation independently. We designed the following three criteria based on Table 1:

- Perceivability: Who does this morph refer to? (i) Pretty sure, (ii) Not sure, and (iii) No clues.
- Funniness: How interesting is the morph? (i) Funny, (ii) Somewhat funny, and (iii) Not funny.
- Appropriateness: Does the morph describe the target entity appropriately? (i) Make sense, (ii) Make a little sense, and (iii) Make no sense.

The three choices of each criteria account for 100% (i), 50% (ii) and 0% (iii) satisfaction rate, respectively. If the assessor correctly predicts the target entity with the Perceivability measure, (s)he is asked to continue to answer the Funniness and Appropriateness questions; otherwise the Funniness and Appropriateness scores are 0. The human evaluation results are shown in Table 3. The Fleiss's kappa coefficient among all the human assessors is 0.147 indicating slight agreement.

From Table 3 we can see that overall the system achieves 66% of the human performance with comparable stability as human. In particular, Method 4 based on translation and transliteration generates much more perceivable morphs than human because the system may search in a larger vocabulary. Interestingly, similar encouraging results - system outperforms human - have been observed by previous back-transliteration work (Knight and Graehl, 1998).

It's also interesting to see that human assessors can only comprehend 76% of the human generated morphs because of the following reasons: (1) the morph is newly generated or it does not describe the characteristics of the target entity well; and (2) the target entity itself is not well known to human assessors who do not keep close track of news topics. In fact only 64 human generated morphs and 72 system generated morphs are perceivable by all human assessors.

For Method 2, the human created morphs are assessed as much more and funny than the system generated ones because human creators use this approach only if: (1). the radicals still reflect the meaning of the character (e.g., "愁 (worry)" is decomposed into two radicals "心秋 (heart au-

tumn)" instead of three "禾火心" (grain fire heart) because people tend to feel sad when the leaves fall in the autumn), (2). the morph reflects some characteristics of the entity (e.g., "江泽民 (Jiang Zemin)" has a morph "水工泽民 (Water Engineer Zemin)" because he gave many instructions on water conservancy construction); or (3). The morph becomes very vivid and funny (e.g., the morph "木子月月鸟 (Muji Yue Yue Bird)" for "李鹏" is assessed as very funny because "木子(Muji)" looks like a Japanese name, "月月(Yue Yue)" can also refer to a famous chubby woman, and "鸟人 (bird man)" is a bad word referring to bad people); or (4). The morph expresses strong sentiment or sarcasm; or (5) The morph is the name of another entity (e.g., the morph "古月(Gu Yue)" for "胡锦涛(Hu Jintao)" is also the name of a famous actor who often acts as Mao Zedong). The automatic approach didn't explore these intelligent constraints and thus produced more boring morph. Moreover, sometimes human creators further exploit traditional Chinese characters, generalize or modify the decomposition results.

Table 4 presents some good (with average score above 80%) and bad (with average score below 20%) examples.

| Good Examples | | |
|---|---|---|
| Entity | Morph | Method |
| 本拉登 (Osama bin Laden) | 笨拉灯 (The silly turning off light) | M1 |
| 蒋介石 (Chiang Kaishek) | 草将介石 (Grass General Jie Shi) | M2 |
| 比尔盖茨 (Bill Gates) | 票子盖茨 (Bill Gates) | M4 |
| Bad Examples | | |
| Entity | Morph | Method |
| 科比 (Kobe) | 胳膊 (Arm) | M1 |
| 梅德韦杰夫 (Medvedev) | 梅德育 (Mei Virtue) | M5 |
| 林书豪 (Jeremy Lin) | 老子 (Lao Tze) | M6 |

Table 4: System Generated Morph Examples

To understand whether users would adopt system generated morphs for their social media communication, we also ask the assessors to recite the morphs that they remember after the survey. Among all the morphs that they remember correctly, 20.4% are system generated morphs, which is encouraging.

## 3.3 Automatic Evaluation

Another important goal of morph encoding is to avoid censorship and freely communicate about certain entities. To evaluate how well the new morphs can pass censorship, we simulate the censorship using an automatic morph decoder con-

| | M1 | | M2 | | M3 | | M4 | | M5 | | M6 | | M7 | | **Overall** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Human | System | Human | System | Human | System | Human | System | Human | System | Human | System | Human | System | **Human** | **System** |
| # of morphs | 17 | 124 | 4 | 21 | 10 | 54 | 9 | 28 | 64 | 87 | 9 | 18 | 74 | 50 | **187** | **382** |
| Perceivability | 75 | **76** | 95 | 86 | 94 | 81 | 61 | **71** | 87 | 59 | 66 | 5 | 77 | 34 | **76** | 67 |
| Funniness | 78 | 49 | 92 | 43 | 44 | 41 | 70 | 47 | 70 | 35 | 74 | 28 | 79 | 44 | **76** | 46 |
| Appropriateness | 71 | 51 | 89 | 59 | 81 | 43 | 75 | 49 | 76 | 36 | 78 | 18 | 82 | 38 | **79** | 43 |
| Average | 75 | 59 | 92 | 57 | 73 | 55 | 69 | 56 | 78 | 43 | 73 | 17 | 79 | 39 | **77** | 52 |
| Standard Deviation | 12.29 | 21.81 | 7.32 | 11.89 | 13.2 | 9.2 | 17.13 | 20.3 | 18.83 | 17.54 | 10.01 | 21.23 | 15.18 | 15.99 | **15.99** | **18.14** |

Table 3: Human Evaluation Satisfaction Rate (%)

sisted of a morph candidate identification system based on Support Vector Machines incorporating anomaly analysis and our morph resolution system (Huang et al., 2013). We use each system generated morph to replace its corresponding human-created morphs in Weibo tweets and obtain a new "morphed" data set. The morph decoder is then applied to it. We define *discovery rate* as the percentage of morphs identified by the decoder, and the ranking accuracy $Acc@k$ to evaluate the resolution performance. We conduct this decoding experiment on 247 system generated and 151 human generated *perceivable* morphs with perceivability scores > 70% from human evaluation.

Figure 1 shows that in general the decoder achieves lower discovery rate on system generated morphs than human generated ones, because the identification component in the decoder was trained based on human morph related features. This result is promising because it demonstrates that the system generated morphs contain new and unique characteristics which are unknown to the decoder. In contrast, from Figure 2 we can see that system generated morphs can be more easily resolved into the right target entities than human generated ones which are more implicit.
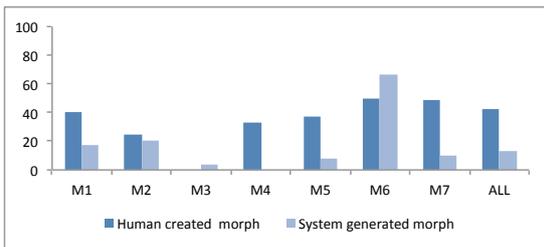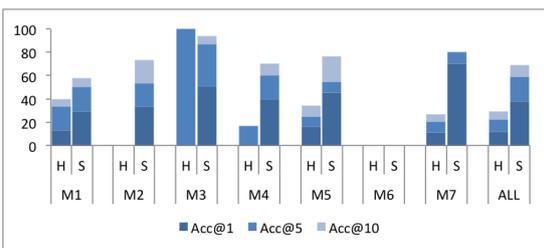


Figure 1: Discovery Rate (%)



Figure 2: Resolution Acc@K Accuracy (%)

## 4 Related Work

Some recent work attempted to map between Chinese formal words and informal words (Xia et al., 2005; Xia and Wong, 2006; Xia et al., 2006; Li and Yarowsky, 2008; Wang et al., 2013; Wang and Kan, 2013). We incorporated the pronunciation, lexical and semantic similarity measurements proposed in these approaches. Some of our basic selection criteria are also similar to the constraints used in previous work on generating humors (Valitutti et al., 2013; Petrovic and Matthews, 2013).

## 5 Conclusions and Future Work

This paper proposed a new problem of encoding entity morphs and developed a wide variety of novel automatic approaches. In the future we will focus on improving the language-independent approaches based on historical figure mapping and culture and reputation modeling. In addition, we plan to extend our approaches to other types of information including sensitive events, satires and metaphors so that we can generate fable stories. We are also interested in tracking morphs over time to study the evolution of Internet language.

# References

David Bamman, Brendan O'Connor, and Noah A. Smith. 2012. Censorship and deletion practices in Chinese social media. *First Monday*, 17(3).

Le Chen, Chi Zhang, and Christo Wilson. 2013. Tweeting under pressure: analyzing trending topics and evolving word choice on sina weibo. In *Proceedings of the first ACM conference on Online social networks*, pages 89–100.

Zhendong Dong and Qiang Dong. 1999. Hownet. In *http://www.keenage.com*.

Hongzhao Huang, Zhen Wen, Dian Yu, Heng Ji, Yizhou Sun, Jiawei Han, and He Li. 2013. Resolving entity morphs in censored data. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL2013)*.

Heng Ji and Ralph Grishman. 2011. Knowledge base population: Successful approaches and challenges. In *Proceedings of the Association for Computational Linguistics (ACL2011)*.

Heng Ji, Ralph Grishman, Dayne Freitag, Matthias Blume, John Wang, Shahram Khadivi, Richard Zens, and Hermann Ney. 2009. Name extraction and translation for distillation. *Handbook of Natural Language Processing and Machine Translation: DARPA Global Autonomous Language Exploitation*.

Heng Ji, Ralph Grishman, Hoa Trang Dang, Kira Griffitt, and Joe Ellis. 2010. Overview of the tac 2010 knowledge base population track. In *Text Analysis Conference (TAC) 2010*.

Heng Ji, Ralph Grishman, and Hoa Trang Dang. 2011. Overview of the tac 2011 knowledge base population track. In *Proc. Text Analysis Conference (TAC) 2011*.

Kevin Knight and Jonathan Graehl. 1998. Machine transliteration. *Computational Linguistics*, 24(4).

Zhifei Li and David Yarowsky. 2008. Mining and modeling relations between formal and informal chinese phrases from web corpora. In *Proceedings of Conference on Empirical Methods in Natural Language Processing (EMNLP2008)*, pages 1031–1040.

Jianyu Li and Jie Zhou. 2007. Chinese character structure analysis based on complex networks. *Physica A: Statistical Mechanics and its Applications*, 380:629–638.

Paul McNamee and Hoa Trang Dang. 2009. Overview of the tac 2009 knowledge base population track. In *Proceedings of Text Analysis Conference (TAC2009)*.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In C.J.C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119.

Sasa Petrovic and David Matthews. 2013. Unsupervised joke generation from big data. In *Proceedings of the Association for Computational Linguistics (ACL2013)*.

Alessandro Valitutti, Hannu Toivonen, Antoine Doucet, and Jukka M. Toivanen. 2013. "let everything turn well in your wife": Generation of adult humor using lexical constraints. In *Proceedings of the Association for Computational Linguistics (ACL2013)*.

Aobo Wang and Min-Yen Kan. 2013. Mining informal language from chinese microtext: Joint word recognition and segmentation. In *Proceedings of the Association for Computational Linguistics (ACL2013)*.

Aobo Wang, Min-Yen Kan, Daniel Andrade, Takashi Onishi, and Kai Ishikawa. 2013. Chinese informal word normalization: an experimental study. In *Proceedings of International Joint Conference on Natural Language Processing (IJCNLP2013)*.

Yunqing Xia and Kam-Fai Wong. 2006. Anomaly detecting within dynamic chinese chat text. In *Proc. Workshop On New Text Wikis And Blogs And Other Dynamic Text Sources*.

Yunqing Xia, Kam-Fai Wong, and Wei Gao. 2005. Nil is not nothing: Recognition of chinese network informal language expressions. In *4th SIGHAN Workshop on Chinese Language Processing at IJCNLP*, volume 5.

Yunqing Xia, Kam-Fai Wong, and Wenjie Li. 2006. A phonetic-based approach to chinese chat text normalization. In *Proceedings of COLING-ACL2006*, pages 993–1000.

Richard Zens and Hermann Ney. 2004. Improvements in phrase-based statistical machine translation. In *Proceedings of HLT-NAACL2004*.

Jing Zhang, Biao Liu, Jie Tang, Ting Chen, and Juanzi Li. 2013. Social influence locality for modeling retweeting behaviors. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence (IJCAI'13)*, pages 2761–2767.