Detecting Location-Indicating Phrases in User Utterances for Chat-Oriented Dialogue Systems

Hiromi Narimatsu, Hiroaki Sugiyama, Masahiro Mizukami
NTT Communication Science Laboratories
{narimatsu.hiromi, sugiyama.hiroaki, mizukami.masahiro}@lab.ntt.co.jp

Abstract

This paper establishes a method that detects words or phrases that indicate location in Japanese spoken language for a chat-oriented dialogue system. Although conventional methods for detecting words or phrases focus on named entities (NE)s, humans frequently use non-NE words to signify locations. For example, we can say "I went to that famous tower in Paris" instead of "I went to the Eiffel Tower" if we forget its proper name. Since conventional NE recognizers extract only Paris as a location from the utterance, they cannot correctly understand because the phrase "that famous tower in Paris" denotes the location in this utterance. Such insufficient understanding may allow a system to ask "Where did you go in Paris?" next, and easily result in dialogue breakdown.

To correctly understand location phrases, we focused on conditional random field (CRF)-based model as a representative method for NE extraction. Since there is no chat corpus that such location-indicating phrases are annotated, we firstly created a corpus by annotating location-indicating phrases to actual human-human chat-oriented dialogues. Then, we evaluated with the corpus how the model work. The evaluation shows that human utterances include various location phrases except for NEs. It also shows that a CRF-based model trained a new annotated corpus detects the target phrases with high accuracy.

1 Introduction

Recently, chat-oriented dialogue systems have been attracting attention for social and entailment aspects [Bickmore and Picard, 2005; Ritter *et al.*, 2011; Higashinaka *et al.*, 2014; Otsuka *et al.*, 2017]. In chatting situation, there is a significant problem that systems precisely understand users' utterances. Although the systems need to grasp the meaning of words or phrases in utterances [Higashinaka *et al.*, 2015], it is difficult because the domain in not limited in chats.

In this study, we focused on the understanding of location phrases. Locations are frequently used as background of a dialogue, which should be shared between talkers. In addition, location phrases are important in a slot filling-based conversational agents [Han *et al.*, 2013]. An example system is that uses 5W1H (who, what, when, where, why, how) slots for filling by conversation. The target words or phrases are extracted from user utterances. Since the targets of *when* and *where* slots particularly appear in the beginning of dialogue, the system needs to detect whether they are included in the utterance.

User : I played tennis at a park.
System: (understand only *play tennis*.)
System: Oh you played tennis, where did you play?
User : Hmm, I played at a park close to my home...

Figure 1: Example dialogue that system undetected location phrase and caused a dialogue breakdown.

User : I went to the capital of France last week.
System: (understand only *France* as location.)
System: Oh great, where did you go in France?
User : Hmm, I visited the capital...

Figure 2: Example dialogue that the system misunderstood location phrase and caused a dialogue breakdown.

For the purpose of detecting location in sentences and documents, previous work has been adopted named entity (NE) recognition. However, we human often use and understand location words or phrases except for NEs in chatting situation. We describe two cases using Figure 1 and Figure 2.

First case is that we human use and understand a common word as location. In the example shown in Figure 1, *a park* represents location but it is not a named entity. If the system takes 5W1H information extraction strategies, it is important to detect it as location. However, NE recognizers usually undetect it as location, and it leads a dialogue breakdown.

The second case is that humans use various words to tell a location. For instance, the following two utterances "I went to Paris" and "I went to the capital of France," have identical meaning. However, conventional NE recognizers correctly extract Paris as the location in the first utterance, but they only extract France as the location while whole the phrase "the capital in France" is the correct location phrase in the second. Such insufficient detection also results in a dialogue breakdown, as shown in Figure 2.

The simplest way to detect these phrases as location is that developing a location phrase list as a dictionary and matching the target phrase against the list, but it is possible to lead misdetection such as *park* in "Can I park my car?" for the first case. Moreover, location phrases inlcude not only words but also phrases like "the capitable in France," and "the electricity shop near XX station" as shown in the second case. Therefore, simply adding these location phrases to a list is not effective.

To overcome the difficulties, we conduct this research as follows. First, we newly annotated such location-indicating phrases to human-human chat-oriented dialogues because there is no such corpus available. Then, we evaluated the location phrase detection accuracy using the chat corpus. We focused on CRF-based model that is a representative method for NE extraction, and compared three models; one is trained only NEs, another is trained the chat corpus, and the other is combined the above two models. The evaluation results show that human represents location with various phrases except for NEs, and training the chat corpus with CRF- based model is effective for detecting them.

2 Related Work

For the purpose of grasping the meaning of words or phrases, there are two types of related work. The first type is a named entity task initiated by the Defense Advanced Research Projects Agency (DARPA) [DARPA, 1995] at the Sixth Message Understanding Conference (MUC-6). It identified seven types of NEs: person, organization, location, and numeric expressions such as date, time, and money. Sekine *et al.* proposed an extended named entity [Sekine *et al.*, 2002]. There are many NE recognition approaches [Sekine *et al.*, 1998], and the scheme using conditional random fields (CRF) [Lafferty *et al.*, 2001] has been the primary one [Nadeau and Sekine, 2007]. The characteristics using CRF is that it can estimate the sequence probability dealing with relations between n-th prior and posterior words and their features, i.e., part-of-speech (POS) tags and character types. For this task, approaches using Bi-directional LSTM, RNN have also been proposed [Chiu and Nichols, 2015; Lample *et al.*, 2016; Wang *et al.*, 2017]. They obtain higher performance than CRF-based methods, but they need a certain amount of training datasets to obtain stable results. Although these approaches detect NEs with high accuracy, the target NE locations are different from location phrases in chats.

The second type is an information extraction task for task-oriented dialogue systems [Lee *et al.*, 2010; Eric *et al.*, 2017; Bordes *et al.*, 2017]. Basically, this is a slot-filling task, which assumes that the target words or phrases that fill the slots are predefined. For example, in a restaurant reservation task, slots are prepared for date, location, and the number of people, and they are filled through a dialogue by checking words in user's utterances against words and phrases list that are predefined. Although this approach is effective if the words and phrases list is prepared in advance, they are unsuitable for chatting situation such that target words or phrases cannot be predefined.

3 Location Phrase Dataset

To examine what kinds of words or phrases except for NE are used as locations, we analyze human utterances in chats. Since there is no available chat data with location phrase annotations, we create a corpus by annotating location words or phrases in human-human chat-oriented dialogues.

3.1 Location phrase annotation

We use chat dialogues collected by human-human text-based chats, and annotated location words or phrases to them. The dialogue data are collected by the previous study [Meguro *et al.*, 2009] and the dialogues are conducted without limiting the topic or contents. We use 600 dialogues and 24,888 utterances in the dataset. Each dialogue consists of about 40 utterances.

Then, we extract location-indicating phrases by manual annotations. To define the instructions for the annotation, we examined 10 chat dialogues including about 400 utterances and extracted the features of location phrases. These are example location phrases:

Example 1
I went to the capital of France yesterday.
I ate at a ramen shop near my office.

In the examples, the underlined phrases *the capital of France* and *a ramen shop near my office*, were the target locations of the utterance. Although *France* and *ramen shop* are also location words or phrases, they are partial phrases of the target locations. Therefore, we assumed a whole phrase that indicates a location is extracted as a single location.

Then, we determined the instructions as follows:

1. Annotate a sequence of words (including modifiers) as a single location, such as *the capital in France* instead of *France*.
2. Annotate words or phrases that can identify a location, such as the *area around the tower* and *the place where I ate ramen*.
3. Regard words or phrases that evoke "location" even if only slightly, as annotation target. (This definition helps to avoid overlooking any words.)
4. Clarify the ambiguity of the annotation, by attaching one of the labels shown in Table 1. (It helps to omit superfluities that may be occurred by the third instruction.)

We assumed that most of the location phrases can be intuitively understood as location, but it is possible that human cannot decide whether the phrase is location, and where the phrase is segmented. Therefore, we decided the ambiguity labels as shown in Table 1. These labels help to precisely measure the system performance by removing phrases which human cannot simply decide.

To decide the number of annotators, we firstly verified the annotation agreement using the first 30 dialogues. We employed two annotators and gave them the above instructions and the entire sequential dialogues.

Table 1: Ambiguity label.

Label	Criteria
$L1$	The words/phrase that annotated without any hesitation.
$L2$	The words/phrase that annotated without a certain about segmentation.
$L3$	The words/phrase that the annotator annotated but had no confidence that it is a location.
$L4$	Applies both labels 2 and 3.

Table 2 shows the annotation agreement results. We calculated agreement score v by

$$v = \frac{\text{Number of phrases detected by both annotators}}{\text{Number of phrases detected by the reference annotator}}.$$

The score using all the detected phrases is shown as *all* and that using only label $L1$ is shown as $L1$. The agreement scores using $L1$ data exceeded 0.89 in both evaluations. Since the 0.89 score is high enough to use the data of a single annotator, one annotator worked on the remaining 570 dialogues in accodance with the above instructions.

Table 2: Annotation agreement.

Reference	Detector	v (*all*)	v ($L1$)
Annotator 1	Annotator 2	0.87	0.89
Annotator 2	Annotator 1	0.83	1.0

3.2 Dataset analysis

We analyzed the annotated data by counting the number of ambiguity labels. The total number of location words or phrases annotated by this work was 4,202. Table 3 shows the number and the ratio of the ambiguity labels annotated to these phrases. The $L1$ results show that 70% of the location phrases were annotated without any ambiguity. The $L2$ results show that 25% were annotated with segmentation ambiguity. The other labels were much less than $L1$ and $L2$.

Table 3: Number of phrases with ambiguity labels.

Label	$L1$	$L2$	$L3$	$L4$	all
Number	2914	1025	216	47	4202
(Ratio)	(0.69)	(0.24)	(0.05)	(0.01)	

Then we analyze the feature of sentences in each ambiguity labels by taking some representative examples. Figure 3, Figure 4, and Figure 5 are the example three sentences assigned into each label $L1$, $L2$, and $L3$ respectively. For label $L1$ that human understand the words or phrases as locations without any ambiguity, there were many location phrases except for NEs such as general nouns and the phrases including modifiers. For label $L2$ that human uncertainly annotated the words in regard to the segmentation place, there were words used to ambiguate the locations for example *around* and *about*. For label $L3$ that human annotated the words with less confidence, there were words that it is difficult to identify the unique location, and words included in other phrases that represent other entities except for location.

1	[JP] 電車の中で隣の人とおしゃべりしました。 [EN] I talked with the person next to me in the train.
2	[JP] 暇なときはよく電気屋にいきます。 [EN] I often go to electricity shops in my free time.
3	[JP] 水が美味しいところに行きたい。 [EN] I want to go a place where the water is delicious.

Figure 3: Representative examples assigned into $L1$.

1	[JP] 国内を三地域ほど旅をしました。 [EN] I travelled about three areas in Japan.
2	[JP] X というお店にいきました。 [EN] I went to the shop named X.
3	[JP] 京都の辺りは暖かいです。 [EN] Area around Kyoto is hot.

Figure 4: Representative examples assigned into $L2$.

1	[JP] 私は実家暮らしです。 [EN] I am living at home.
2	[JP] ファミレスよりファーストフードにいきます。 [EN] I often go to fast food restaurants than family restaurants.
3	[JP] イタリア料理を良く作ります。 [EN] I usually made Italian food.

Figure 5: Representative examples assigned into $L3$.

From the results, we focused on detecting location phrases assigned $L1$ because it is not a big difference that understanding only *Kyoto* as location and *area around Kyoto* as location phrases. In addition, the location phrases assigned into $L3$ are different from others because they are some parts of other entities. Since such phrases are understood as other entities, we assumed that it is not necessary to detect them as location. Furthermore, although the location phrases assigned into $L3$ include phrases that cannot identify the location as *fast food restaurants*, human does not always understand them as location. Therefore, we use the location phrases assigned $L1$ as evaluation target.

4 Location Phrase Detection using Annotated Dataset

To detect target location phrases except for NE, we develop a new model using the dataset that is newly annotated in Section 3. We used CRF [Lafferty *et al.*, 2001] to detect location phrases by training word sequences with their features and tags. Since the performance of CRF-based approach is stable and it can work with less datasets than neural network based methods, we take CRF-based approach.

We use grammatical and superficial features: the original words, the POS tags for each word estimated *a priori*, and five character types: hiragana, katakana, kanji, mark, and tag.

Table 4: Features and LOC-tag that is an estimation target where the input sentence is [JP]: 昨日、エッフェル塔に登ったよ。/ [EN]: I went to the Eiffel Tower yesterday. The underlined words represent location.

Word	Char type	POS	LOC-tag
<S>	tag	bos	O
昨日 (yesterday)	kanji	noun	O
、 (,)	mark	noun	O
エッフェル (Eiffel)	katakana	noun	B-LOC
塔 (Tower)	kanji	noun	I-LOC
に (to)	hiragana	pp	O
登っ (go)	kanji	verb	O
た (-ed (past))	hiragana	verb	O
よ (expression)	hiragana	sep	O
。 (.)	mark	sep	O
<S>	tag	eos	O

Table 4 shows the example features where the input sentence is "[JP] 昨日、エッフェル塔に登ったよ。([EN] I went to the Eiffel Tower yesterday.) The underlined words represent location." First, the sentence is split into words using a Japanese morphological analyzer, JTAG [Fuchi and Takagi, 1998], and POS tags were estimated simultaneously. Char type represents the character type that is determined by its unicode symbols. The LOC-tags are labeled using BIO-tags that B-LOC is attached to the first word of location phrase, I-LOC is attached to its intermediate words, and O is attached to the other words that are not location words or phrases. BIO-tags are the estimation targets. Here, the i-th word is represented as x_i. To train and estimate the tag of i-th word x_i, we used the features of $x_{i-2}, \cdots, x_{i+2}$.

5 Evaluation

We evaluated the performance of the location phrase detection using the new model described in Section 4 comparing with conventional models trained only NEs.

5.1 Experimental setup

We compared the following three models:

- **NE** CRF trains the NE location tags annotated to 1995 Mainichi newspapers.
- **Dial** CRF trains the location tags newly annotated to our text dialogue data.
- **NE+Dial** CRF trains both **NE** and **Dial** dataset.

For **NE** evaluation, we only used B-LOC, I-LOC, and O location tags instead of all NE-tags in this experiment. All the 24,888 annotated utterances were used as test data for the evaluation. For **Dial** evaluation, we calculated the evaluation scores by 5-fold cross-validation. For **NE+Dial** evaluation, we combined both of the above dataset and trained them using CRF. We evaluated the detection performance using precision, recall, and f-measure, which is the harmonic mean of the precision and recall. If the detected phrase partially matched the annotated one, it was counted as incorrect because extracting partially matched phrase such as *Paris* in "that famous tower in Paris" easily leads dialogue breakdown.

Table 5: Location phrase detection performance.

Label	model	Precision	Recall	f-Measure
all	**NE**	0.58	0.22	0.32
	Dial	0.91	0.70	0.79
	NE+Dial	0.87	0.74	**0.80**
$L1$	**NE**	0.66	0.03	0.07
	Dial	0.89	0.67	0.76
	NE+Dial	0.91	0.84	**0.87**

1	[JP] 海も山もあるのでいろいろできました。 [EN] I can do many things because there is a sea and mountain.
2	[JP] 私は実家暮らしです。 [EN] I am living at home.
3	[JP] 私の近所の図書館にも子供がたくさんいます。 [EN] There are many children at the library in my neighborhood.

Figure 6: Example of location phrases that **Dial** successfully detected. **Dial** detected underlined words and phrases as locations, but **NE** did not detect any locations.

5.2 Results

Table 5 shows the results. Score *all* represents the detection performance for the annotated location phrases in all the utterances. Score $L1$ represents the performance using only the utterances that are annotated $L1$ ambiguity labels. The results of recall scores using all labels indicate that only 22% of location phrases in human-human chat dialogue are NEs, and **Dial** can detect non-NE location phrases by training the suitable dataset. Then, the results of precision scores show that the correctness of detected phrases using **Dial** are improved 0.33 points over **NE**. Therefore, the overall score f-Measure is improved 0.47 points. The results of label $L1$ remarkably indicate that human use various phrases except for NEs as location in chatting situation. Finally, combined models **NE+Dial** reached 0.80 for all, and 0.87 for $L1$ label.

To demonstrate the effectiveness of training the newly annotated data, we analyzed the detected location phrases and compared the results of the two models; **NE** and **Dial**. Figure 6 shows the example phrases of **Dial** successfully detected utterances and **NE** undetected utterances. The underlined words or phrases represent the location phrases. Although humans understand sea, mountain, and home as locations, these terms are undetected by **NE** because they are not location NEs. However, these words were correctly detected as locations by training the chat corpus annotated in this study.

Figure 7 shows example phrases of **Dial** undetected and **NE** successfully detected utterances. The underlining is represented as well in Figure 6. The words *Florence*, *Palma*, and *Bologna* are named locations. Famous place names are of course included in the data of **NE**. However, **Dial** includes some famous place names only in the annotated dialogue data. Therefore, combining the training data of **NE** and **Dial** is effectively improved the detection performance. However, some named locations that are not so famous cannot be detected by both **NE** and **Dial**. Therefore, adding some named

1	[JP] フィレンツェのステーキはオススメです。 [EN] The steak in Florence is my recommendation.
2	[JP] パルマやボローニャは本当においしいものがたくさんある。 [EN] There are many delicious foods in Palma and Bologna.

Figure 7: Example of location phrases that **NE** successfully detected utterances. **NE** detected underlined words and phrases, but **Dial** did not detect any locations.

locations may be necessary in case that further higher accuracy is required.

From these results, **Dial** extracts location words and phrases that are not named entities, and a group of phrases such as *the library in my neighborhood* by traning features of words and words' sequence. Since the detected phrases from **NE** and **Dial** are different each other, the combined model **NE+Dial** is effective for detecting them. The results also show that CRF trained NE with small dialogue dataset is effective for detecting location phrase in chat-oriented dialogues.

6 Conclusion

We addressed the importance of understanding location phrases in chatting situations. To verify the performance of conventional CRF models of NE extraction for phrases that indicate locations in chatting situation, we created a new corpus of annotated location phrases in a textualized human-human chat-oriented dialogue. Our evaluation using the corpus shows that the conventional NE recognizer is insufficient for understanding location phrases in chatting situation, but the conventional method CRF is effective for detecting location-indicating phrases in chats by training the target words and phrases that are newly annotated in this studies.

In future work, we will further annotate an essential location phrase in phrases assigned to $L2$, $L3$, and $L4$ ambiguity labels, and evaluate the performance in detail. Then, we will implement the detection function in 5W1H based chat-oriented dialogue systems, and evaluate the effectiveness. Some dialogue examples using this location-phrase detection are described in Section A. Finally, we will extend this work to other targets of 5W1H except for locations.

A Appendix

We show some dialogue examples using the location phrases detection. In the case of Figure 8, the system conducts dialogue by choosing one sentence from many options. Although the similarity score between the user utterance and the option sentences is high, the system can filter the options with different locations.

In the case of Figure 9, the system correctly understands mountain as location and asks "which mountain" to identify the location in detail. Actually, *the mountain near by Mt. Fuji* easily makes a system misunderstand only *Mt. Fuji* as location. Therefore, showing the correct understanding to users may look smarter than ever.

In case of Figure 10, the system rephrases the location phrase to a correct NE. Detecting location phrase that is not NE may be used for identifying the location and rephrasing it as smart agents. These rephrasing may makes us feel the intelligence of the system.

References

[Bickmore and Picard, 2005] Timothy W Bickmore and Rosalind W Picard. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 12(2):293–327, 2005.

[Bordes *et al.*, 2017] Antoine Bordes, Y-Lan Boureau, and Jason Weston. Learning end-to-end goal-oriented dialog. *Proc. of the 5th International Conference on Learning Representations (ICLR)*, 4 2017.

[Chiu and Nichols, 2015] Jason PC Chiu and Eric Nichols. Named entity recognition with bidirectional lstm-cnns. *arXiv preprint arXiv:1511.08308*, 2015.

[DARPA, 1995] DARPA. Proc. of the sixth message understanding conference. *Morgan Kaufmann Publishers, Columbia, MD, USA*, 1995.

[Eric *et al.*, 2017] Mihail Eric, Lakshmi Krishnan, Francois Charette, and Christopher D. Manning. Key-value retrieval networks for task-oriented dialogue. *Proc. of the 18th Annual SIGdial Meeting on Discourse and Dialogue (SIGDIAL)*, pages 37–49, 8 2017.

[Fuchi and Takagi, 1998] Takeshi Fuchi and Shinichiro Takagi. Japanese morphological analyzer using word co-occurrence. *Proc. of the 36th Annual Meeting of the Association for Computational Linguistics (COLING)*, pages 409–413, 1998.

[Han *et al.*, 2013] Sangdo Han, Kyusong Lee, Donghyeon Lee, and Gary Geunbae Lee. Counseling dialog system with 5w1h extraction. In *Proceedings of the SIGDIAL 2013 Conference*, pages 349–353, 2013.

[Higashinaka *et al.*, 2014] Ryuichiro Higashinaka, Kenji Imamura, Toyomi Meguro, Chiaki Miyazaki, Nozomi Kobayashi, Hiroaki Sugiyama, Toru Hirano, Toshiro Makino, and Yoshihiro Matsuo. Towards an open-domain conversational system fully based on natural language processing. *Proc. of the 25th International Conference on Computational Linguistics (COLING)*, 2:928–939, 2014.

[Higashinaka *et al.*, 2015] Ryuichiro Higashinaka, Kotaro Funakoshi, Masahiro Araki, Hiroshi Tsukahara, Yuka Kobayashi, and Masahiro Mizukami. Towards taxonomy of errors in chat-oriented dialogue systems. In *Proceedings of the 16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pages 87–95, 2015.

[Lafferty *et al.*, 2001] John Lafferty, Andrew McCallum, and Fernando C.N. Pereira. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. *Proc. of the Eighteenth International Conference on Machine Learning (ICML)*, pages 282–289, 2001.

System:	お昼はどこで食べましたか？(Where did you eat lunch today?)
User:	東京行きの新幹線で食べました。 (I ate lunch in the bullet train bound for Tokyo.)
System:	*Thinking as follows... These options are similar to user's utterance, but they should be filtered because the location is not the bullet train. Option1: Lunch at Tokyo is great! Option2: There are many restaurants in the Tokyo station.*
System:	新幹線、ってことは駅弁ですか？(Bullet train! So, you mean you ate Ekiben (box lunch)??)
User:	そうそう。京都駅で駅弁買ったんだよ。(That's right! I bought the box lunch at Kyoto station.)

Figure 8: Example dialogue using the location phrase detection.

System:	夏休みにどこへ行きましたか。(Where did you go last summer?)
User:	山に行きました。(I went to a mountain.)
System:	*Thinking as follows... Understand a mountain as location, but that is not named entity.*
System:	山行ったんですね。どちらの山に行かれたんですか？(Oh, sounds good! Which mountain did you go?)
User:	富士山の近くの山に登りました。(I climbed the mountain near by Mt. Fuji.)
System:	*Thinking as ... Understand the mountain near by Mt. Fuji as location.*
System:	富士山じゃないんですね。 (Oh, that's not Mt. Fuji.)

Figure 9: Example dialogue using the location phrase detection.

User:	先週末、パリにあるタワーに登りました。(I went to the tower in Paris last weekend.)
System:	*Thinking as follows... Understand the tower in Paris as location and identify it as Eiffel Tower.*
System:	あっ、エッフェル塔ですね。どうでしたか。(Oh, that's the Eiffel Tower, isn't it? How was that?)
User:	そうそう。すごく眺めが良かった。(Yes! The view was great!)

Figure 10: Example dialogue using the location phrase detection.

[Lample *et al.*, 2016] Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, Kazuya Kawakami, and Chris Dyer. Neural architectures for named entity recognition. *arXiv preprint arXiv:1603.01360*, 2016.

[Lee *et al.*, 2010] Cheongjae Lee, Sangkeun Jung, Kyungduk Kim, Donghyeon Lee, and Gary Geunbae Lee. Recent approaches to dialog management for spoken dialog systems. *Journal of Computing Science and Engineering*, 4(1):1–22, 3 2010.

[Meguro *et al.*, 2009] Toyomi Meguro, Ryuichiro Higashinaka, Kohji Dohsaka, Yasuhiro Minami, and Hideki Isozaki. Analysis of listening-oriented dialogue for building listening agents. In *Proceedings of the 10th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 124–127. Association for Computational Linguistics, 2009.

[Nadeau and Sekine, 2007] David Nadeau and Satoshi Sekine. A survey of named entity recognition and classification. *Named Entities: Recognition, classification and use*, pages 3–26, 2007.

[Otsuka *et al.*, 2017] Atsushi Otsuka, Toru Hirano, Chiaki Miyazaki, Ryuichiro Higashinaka, Toshiro Makino, and Yoshihiro Matsuo. Utterance selection using discourse relation filter for chat-oriented dialogue systems. In *Dialogues with Social Robots*, pages 355–365. Springer, 2017.

[Ritter *et al.*, 2011] Alan Ritter, Colin Cherry, and William B Dolan. Data-driven response generation in social media. In *Proceedings of the conference on empirical methods in natural language processing*, pages 583–593. Association for Computational Linguistics, 2011.

[Sekine *et al.*, 1998] Satoshi Sekine, Ralph Grishman, and Hiroyuki Shinnou. A decision tree method for finding and classifying names in japanese texts. *Proc. of the 6th Workshop on Very Large Corpora*, 1998.

[Sekine *et al.*, 2002] Satoshi Sekine, Kiyoshi Sudo, and Chikashi Nobata. Extended named entity hierarchy. *Proc. of the 3rd International Conference on Language Resources and Evaluation (LREC)*, pages 52–57, 5 2002.

[Wang *et al.*, 2017] Chunqi Wang, Wei Chen, and Bo Xu. Named entity recognition with gated convolutional neural networks. In *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*, pages 110–121. Springer, 2017.