In Search of the Use-Mention Distinction and Its Impact on Language Processing Tasks

Shomir Wilson

Department of Computer Science, University of Maryland at College Park
College Park, Maryland 20742
shomir@umd.edu

Abstract. The use-mention distinction is a crucial aspect of natural language which allows us to communicate information about language itself. However, the distinction remains underexamined in computational linguistics, with deleterious effects on common tasks. One reason for this deficiency is a lack of appropriate resources to study the distinction. This paper presents the creation of a corpus of instances of mentioned language gathered from Wikipedia, using a set of "mention words" and cues in text formatting. The corpus demonstrates that recurring patterns do exist among instances of mentioned language, which suggests the potential for automatic identification of the phenomenon in the future.

Keywords: metalanguage, corpus, parsing

1 Introduction

In order to understand the language that we speak, we sometimes must refer to the language itself. Language users are able to do this through an awareness of the use-mention distinction, which separates language produced to illustrate properties of itself from language produced for other reasons. This can be demonstrated in a pair of simple sentences:

The cat is on the mat. **(1)**

The word cat refers to a feline animal. **(2)**

A reader easily understands that *cat* in the first sentence refers to an animal entity in a real or hypothetical world, while the same word in the second sentence refers to the word *cat* itself. The use-mention distinction is well-known and has a history of theoretical examination [1-4], but its actual patterns of appearance in natural language have received little study. This lack of attention has a deleterious impact on common tasks in computational linguistics and natural language processing. Part-of-speech taggers assume by their nature that words are used—which is quite reasonable, since the vast majority of language production *is* use—but this means serious issues arise when language is mentioned. While mentioned words (such as “cat” in (2) above) ostensibly serve as nouns or noun phrases, words that rarely (or never) appear as nouns otherwise are subject to mention as well. For similar reasons, mentioned language also interferes with word sense disambiguation; senses imply language use but have little relevance when a word appears simply “as a word”. Mention also has the potential to interfere with sentiment analysis as well; one can discuss another’s disapproval of something, for instance, without actually disapproving of anything.

The historical lack of attention to the use-mention distinction might suggest that it is peripheral to the study of language, but this is far from the truth. Evidence suggests that human communication frequently employs the use-mention distinction, and we would be severely handicapped without it [5, 6]. In both written and spoken contexts, the mention of letters, sounds, words, phrases, or entire sentences is essential for many language activities, including the introduction of new words, attribution of statements, explanation of meaning, and assignment of names [7] . The distinction is also closely related to the appearance-reality distinction in cognitive science [8].

Moreover, detecting the distinction is a nontrivial task. While cues like italic text are sometimes used to indicate mentioned language, such cues are often inconsistently applied (if at all, in informal contexts) and are “overloaded” with other uses as well. Cues such as pauses and gestures exist for mentioned language in spoken conversation, but these are easily lost in transcription.

Given the common reasons for employing the use-mention distinction, a text whose purpose is to introduce to the reader a broad swath of concepts would seem a good place to begin studying the phenomenon. Wikipedia is such a compendium, and several other aspects make it particularly attractive for study. Some of these are its collaborative nature, its stylistic cues (such as italics) to highlight mentioned language, and its size and article variety. Preliminary studies [9] have validated these observations, but they have left open the question of whether instances of mentioned language can be gathered from Wikipedia accurately and in large numbers.

This paper presents the results of a project to identify instances of mentioned language from English Wikipedia articles using lexical and stylistic cues. First, the use-mention distinction is introduced in greater detail, with some examples to illustrate the variety of the phenomenon. A corpus of mentioned language, named the ML corpus for brevity, is then presented, accompanied by a discussion of the lexical and stylistic cues that were used to gather candidates for inclusion. Although multiple human annotators were available for only a subset of the corpus, their frequency of agreement is a mild indication of reliability and consistency. It is believed that this corpus will be sufficient to bootstrap the first efforts for automatic detection of mentioned language.

2 The Use-Mention Distinction

Although the reader is likely to be familiar with the use-mention distinction, the topic merits further explanation to establish what precisely is being studied. Since the vast majority of language is produced for use rather than mention, this paper will focus on occurrences of mentioned language. Linguistic entities that can be mentioned include letters, sounds, words, names, phrases, and entire sentences.

2.1 An Informal Rubric

In spite of the ubiquity of the phrase *use-mention distinction*¸ it is difficult to find any previous efforts to identify when mention does (and does not) occur. For that purpose, this paper will propose an informal rubric based on substitution. It may be applied, with caveats described below, to determine whether a linguistic entity is mentioned by the sentence in which it occurs.

Rubric: Suppose X is linguistic entity in a sentence. X is an instance of mentioned language if the meaning of the sentence does not change when X is replaced by "that [item]*", where* [item] *is "letter", "sound", "word", "name", "phrase", "sentence", etc., and the replacement phrase is understood to refer to X.*

For example, consider the sentence

Fancy automobiles are called luxury cars. (**3**)

where the phrase "luxury cars" is under consideration. Choosing "that phrase" as a replacement, the sentence becomes

Fancy automobiles are called that phrase. (**4**)

where "that phrase" is understood to refer to "luxury cars". While there might be pragmatic consequences to this change (for instance, a context where a language user wants to avoid producing the phrase "luxury cars"), the reader can verify that the meaning of the sentence is essentially unchanged.

This rubric requires some adjustment when the sentence already explicitly refers to *X* as a word, phrase, or other appropriate entity, such as in (2) above. In such cases it may be appropriate to omit the linguistic entity under consideration without substituting, such as this alteration to (2):

The word refers to a feline animal. (**5**)

where "The word" is understood to refer to "cat". Instances of mixed quotation, which straddle both use and mention [10], also may require some charitable adjustments. This is especially apparent when explicit cues of mention are present. An example of this is

Jane said the cat "is on the mat". **6**) where the reader should understand[1]

Jane said the cat "is on the mat", in that exact phrase. **(7)**

While other permutations exist that challenge the letter of its rubric, this paper will posit that its spirit is sufficiently sound.

2.2 Categories of Mentioned Language

A previous study by Wilson [9] gathered a small 171-sentence corpus of mentioned language from Wikipedia, in order to demonstrate its fertility as a source and to determine some categories for the phenomenon. The study used a set of eight categories of mentioned language inspired by previous theoretical work [7, 10], with the acknowledgement that others may exist. The categories are reproduced below to illustrate the diversity of forms of mentioned language, with examples from the corpus for each. The mentioned entity of interest in each sentence appears between asterisks, and longer sentences have been shortened with ellipses.

Proper name: In 2005, Ashley Page created another short piece on Scottish Ballet, a strikingly modern piece called *The Pump Room*, set to pulsating music by Aphex Twin.
Translation or transliteration: The Latin title translates as *a method for finding curved lines enjoying properties of maximum or minimum, or solution of isoperimetric problems in the broadest accepted sense*.
Attributed language[2]: *It is still fresh in my memory that I read a chess book of Karpov by chance in 1985 which I liked very much*, the 21-year-old said.
Words or phrases as themselves: *Submerged forest* is a term used to describe the remains of trees (especially tree stumps) which have been submerged by marine transgression…
Symbols: He also introduced the modern notation for the trigonometric functions, the letter *e* for the base of the natural logarithm…
Phonetics or sounds: The call of this species is a high pitched *ke-ke-ke*…
Abbreviations: …Moskovskiy gosudarstvennyy universitet putej soobshcheniya, often abbreviated *MIIT* for "Moscow Institute of Transport Engineers"…

---

[1] Depending on context, instead this might be a use of *scare quotes* to imply that the cat is not actually on the mat. We will assume that was not the intent, although it has been argued [4] that scare quotes are a form of mention.

[2] In discussions with other researchers, the author has noted some controversy regarding the inclusion of *attributed language*. While it lacks the "pure mention" quality of some of the other categories, it is discussed as mention (albeit in a "mixed" form) in the cited literature, and the authors would argue that it satisfies the rubric set in 2.1. In the absence of a strong justification for excluding attributed language, this study takes an inclusive approach to it.

Proper names were the most common category found by this previous study, which followed the intuition that many Wikipedia articles describe entities identified by proper names. Translation or transliteration, attributed language, and words or phrases as themselves were the next most common, with fewer instances of the remaining categories.

3 Corpus Creation

As explained in the Introduction, a great deal of theoretical study exists on the use-mention distinction, but little (if any) previous research has been concerned with how language users actually exhibit the distinction. The ML Corpus, whose creation is described in this section, is the first substantial attempt to rectify this gap in research. Although it has some limitations, the value of such a resource is not expected to be diminished by them.

3.1 Rationale for Choosing Wikipedia

Wikipedia is a particularly suitable source for collecting instances of mentioned language. Listed here are four factors that led to its selection for this project:

1) *Wikipedia introduces a wide variety of concepts to the reader*. At the time of writing this paper, Wikipedia contains approximately 3.3 million articles. These articles are written informatively, generally without assuming that the reader is already familiar with the topics they discuss. New names and words are frequently introduced, often explicitly, in a manner that invokes mention.
2) *Stylistic cues that are sometimes used to delimit mentioned language are present in article text*. Wikipedia contributors often (though not always) use quote marks, italic text, or bold text to "highlight" where language is mentioned. This convention is stated in Wikipedia's own style manual, though it is unclear whether most contributors read it there or follow it out of habit.
3) *Wikipedia is collaboratively written*. Its text reflects the language habits of a large sample of English writers. It is unclear how much variation exists between writers on how to mention language, so this large sample is desirable.
4) *Wikipedia is freely available*. Language-learning materials (particularly textbooks) were also considered, but legal issues and electronic availability were deemed obstacles. Moreover, the markup code for Wikipedia articles is easy to access and interpret. This allows for the automatic extraction of the stylistic cues mentioned above.

Naturally, choosing Wikipedia for this project introduced some limitations as well. Since articles are not consistently edited, some mentioned language was not captured by stylistic cues. Such cues are also used by Wikipedia contributors for other purposes, such as emphasis, algebraic symbols, and implicit "non-mention" introduction of words. The particular style of writing in Wikipedia differs from other styles where analysis of mentioned language could be valuable, such as spoken language or pedagogical language. Future research will aim to overcome some of these limitations.

3.2 Candidate Collection and Annotation

The previous study described in Section 2.2 observed that instances of mentioned language are relatively sparse in Wikipedia article text, occurring on average less often than once per article. Since hand annotation was a necessary step in creating the ML corpus, some heuristics were used to gather a rich set of mentioned language candidates.

Articles were randomly selected from English Wikipedia's most current article revisions, and heuristic filtering began at this level. Disambiguation pages were excluded from further examination, since they tend to be repetitive in structure and wording. Inside of articles, text from tables and common end sections (i.e., "Sources", "References", "See also", and "External links") also was excluded, since text from those sources was frequently observed to be non-sentential. The remaining article text was then segmented into sentences using NLTK's [11] Punkt sentence tokenizer. Those sentences that contained stylistic cues (bold text, italic text, or text between double quote marks) were retained, and all others were discarded. Applying this procedure to 3,831 articles produced a set of 22,071 sentences, which in turn contained 28,050 instances of text highlighted by stylistic cues[3].

Initial examinations of these remaining sentences suggested that mentioned language occurred in fewer than one in ten of them, and an additional heuristic was applied before manual annotation commenced. Using observations from the previous 171-sentence corpus, sets of "mention-significant" nouns and verbs were gathered. The appearance of a word from these sets near highlighted text signaled that the highlighted text was likely to be mentioned language. The procedure to gather these words was informal and manual, and a few potential mention-significant words (notably the verb *be*) were rejected because their great frequency reduced their significance as indicators. The eleven selected nouns and twelve selected verbs are listed below. The reader may note that most of the nouns refer to linguistic entities, while most of the verbs can serve as relational predicates or refer to speech acts:

Mention nouns: letter, meaning, name, phrase, pronunciation, sentence, sound, symbol, term, title, word
Mention verbs: ask, call, hear, mean, name, pronounce, refer, say, tell, title, translate, write

Words in the sentences were part-of-speech tagged and stemmed, again using tools from NLTK. The sentences were then filtered for those in which a mention word occurred (respecting the part of speech of its set) in the three-word phrase preceding text highlighted by a stylistic cue. This resulted in a set of 898 sentences, which in turn contained 1,164 instances of highlighted text. This set of instances was named the *ML-0* set.

---

[3] henceforth referred to as "highlighted text", for simplicity

Manual annotation of mentioned language then commenced. To eliminate possible biases, all three stylistic cues were substituted with pairs of asterisks (delimiting the beginning and end of highlighted text) prior to inspection. A human reader who was well-acquainted with the detection of mentioned language considered each instance in the ML-0 set and decided if it qualified by reading the sentence that contained it and applying the rubric from Section 2.1. 1,082 instances were deemed to be mentioned language, and this set was named the *ML-1* set, which also serves as the ML Corpus. This figure suggests that the heuristics leading to the creation of the ML-0 set have approximately 93% precision for retrieving mentioned language, though their recall has not yet been measured.

3.3 Reliability and Consistency

Another limitation of the ML corpus is the lack of participation from multiple readers. To explore the possible impact of this, two additional human readers worked separately (from each other and from the primary reader) to annotate a 30-instance subset of the ML-0 set. These readers were also well-acquainted with the detection of mentioned language. Half of the 30 instances were selected from those annotated by the primary reader as mentioned language, and half were selected from those annotated as not. With that condition, the instances were randomly chosen from the ML-0 set, shuffled, and then distributed to the additional readers.

All three readers produced the same annotation for 25 of the instances, and on each of the remaining five, the additional readers differed with each other. (Since the annotation scheme was binary, this meant that one additional reader agreed with the primary reader and one disagreed). The kappa statistic was 0.779. These results were taken as a mild indication of reliability and consistency of the annotations in the ML corpus, while a second effort is currently underway to provide multiple annotations for all instances.

4 Results and Discussion

This section will present some notable findings distilled from the ML-0 and ML-1 sets. Particular attention was given to the precision of the heuristics used to create the ML-0 set. The combination of heuristics performed better (at 93% precision overall) than had been expected, with some standout performances from specific mention words and stylistic cues.

Below, Table 1 shows the frequency of mention words in the three-word phrases preceding each instance (an instance being a string of highlighted text) in the ML-0 set. Mention words were only counted if they appeared as their set-appropriate parts of speech. In the tables in this section, the precision shown is the percentage of those instances deemed by the primary human reader to be mentioned language and thus placed in the ML-1 set.

Table 1. Frequencies of mention nouns (n) and verbs (v) in the three words preceding each instance in the ML-0 set, with their precisions for retrieving mentioned language.

Mention word	Frequency	Precision (%)
call (v)	349	98.6
name (n)	153	98
name (v)	89	94.4
say (v)	86	94.2
term (n)	79	98.7
title (n)	72	84.7
title (v)	64	96.9
word (n)	55	100
write (v)	52	50
mean (v)	39	100
refer (v)	35	85.7
meaning (n)	20	100
translate (v)	20	20
phrase (n)	18	100
symbol (n)	10	80
pronounce (v)	8	100
tell (v)	7	71.4
letter (n)	6	33.3
pronunciation (n)	4	100
ask (v)	4	75
sentence (n)	3	33.3
hear (v)	3	0
sound (n)	1	0

As shown, the verb *call* and the noun *name* stood out as the most common of the mention words, with all others forming a relatively smooth tail of descending frequency. These top two are intuitive, the informative and descriptive purposes of Wikipedia articles. Both words also had substantially above-average precision. *Word* (n), *meaning* (n), *phrase* (n), *pronounce* (n), and *pronunciation* (v) all had perfect precision, though they appeared less frequently. However, following the multiple-reader experiment in Section 3.3, it was discovered that *meaning* instances were particularly difficult to classify, generating some debate among the participants. Finally, an observant reader may note that the frequencies in Table 1 sum to 1,177 instead of 1,164 (the size of the ML-0 set). This is because 13 instances had more than one mention word in the preceding three-word phrase. All 13 of these instances were annotated as mentioned language.

Although stylistic cues were hidden from the readers while they annotated instances, data on the cues was retained. Table 2 below breaks down their frequencies and precisions.

Table 2. Frequencies of stylistic cues in the ML-0 set and their precisions for retrieving mentioned language.

Stylistic cue	Frequency	Precision (%)
double quote	601	96.7
italic	427	86.4
bold	136	97.1

Double quote marks had the highest frequency, and the reason was first assumed to be frequent quotation (in the sense of speech reporting, for example) in Wikipedia. However, as Table 5 will show, that was probably not the case. Italics had by far the lowest precision. 23 of the 58 non-mention italic instances had *write* (v) as a preceding mention word, which conjures a common construction (as in "Dickens wrote *Great Expectations*…") that does not involve mentioned language. Bold had both the highest precision and lowest frequency. It is worth noting that Wikipedia articles, by convention, contain the article subject in bold text in the first sentence.

Prior to analysis, it was hypothesized that the proximity of a mention word to highlighted text increases its likelihood of being mentioned language. Table 3 shows this hypothesis to be true, albeit in the limited three-word window that was examined. Also shown are overall frequencies and precision percentages (weighted by frequencies) for nouns and verbs.

Table 3. Frequencies of mention nouns and verbs in the three words preceding highlighted text (e.g., word position 1 is the word just before the highlighted text), with their precisions for retrieving mentioned language.

Noun/Verb	Frequency		Precision (%)	
position	Noun	Verb	Noun	Verb
1	281	458	98.6	97.2
2	89	179	91.0	85.5
3	51	119	76.5	84.0
overall	421	756	94.3	92.4

There appears to be a strong correlation between proximity and precision, though *proximity* in this data does not account for the grammatical structure of corpus sentences, which will deserve examination in future research. A mention verb directly preceding highlighted text was by far the most common combination. Overall, mention nouns had a slightly greater precision than mention verbs.

Finally, Table 4 shows the most common mention word-stylistic cue combinations in the ML-1 set.

Table 4. The ten most frequent word and stylistic cue combinations in the ML-1 set, with their percentages of the total (1082) instances. Out of 69 possible different word-cue combinations, 59 were observed.

Word	Cue	Frequency	% of total
call (v)	d. quote	151	14.0
call (v)	italic	133	12.1
say (v)	d. quote	74	6.8
name (n)	italic	60	5.5
name (n)	d. quote	56	5.2
call (v)	bold	53	4.9
term (n)	d. quote	45	4.2
name (v)	d. quote	39	3.6
title (v)	italic	36	3.3
title (n)	italic	32	3.0

The prevalence of *call* (v) is once again apparent, as it appears in the two most common combinations. Double quote marks with *say* is the third most common combination, which matches earlier intuitions on quotation, but the same stylistic cue appears frequently with *call* (v), *name* (n), *term* (n), and *name* (v) as well. Bold makes only one appearance in the top ten, in combination with the previously mentioned *call* (v). These ten combinations account for only 17% of the combinations observed but 62.6% of all instances in the ML set.

Overall, it is believed that these results validate the heuristics that were used to collect candidate instances. They also seem to confirm that Wikipedia is a fertile source of mentioned language, as the instances exhibit a variety of different constructions. Given the size of Wikipedia and the current methods for collecting candidates, future expansion of the ML corpus will be possible and productive.

5 Related Work

Wikipedia's emerging utility as a corpus is well-documented in the literature. A few related uses of Wikipedia include named entity recognition [12], syntactic parsing [13], and lexical semantics [14] among many others. Ytrestøl et al. [15] previously noted the relationship between stylistic cues in Wikipedia and the use-mention distinction, though this observation was incidental to their focus on the automatic extraction of sub-domains of articles. The use of stylistic cues described in this paper appears to be unique.

As mentioned in the introduction, little previous study exists of the use-mention distinction as it appears in linguistic corpora. Notably, Anderson et al. [16] gathered by hand a corpus of metalanguage in human dialogue using a subset of the British National Corpus. Their annotation scheme applied to sentence-level utterances, and focused on metalanguage as used to maintain grounding and recover from perturbations (e.g., misunderstandings and interruptions). Mentioned language generally—perhaps always—requires metalanguage to frame it, and it is likely that many instances of the phenomenon were gathered for such a corpus. However, the annotation scheme was not designed to address mentioned language either as a distinct category (it fit partially into several) or as a phenomenon in the structure of a sentence. This difference in focus, as well as the difference in language context, made the findings of the Anderson corpus and the present one substantially different.

6 Future Work

The ML corpus is significant as the first of its kind, but it has some limitations that will require further work. The heuristics used to identify candidates have high precision, but their recall has not yet been measured. However, it is anticipated that the ML corpus is large and varied enough to provide a basis for "bootstrapping" the detection of mentioned language outside of the heuristics presently used. This will be done through a more detailed examination of syntactic and lexical patterns in the sentences contained in the ML corpus. WordNet [17] is expected to be a useful resource for expanding the sets of mention nouns and verbs currently used. To expand the data on inter-annotator agreement, Amazon's Mechanical Turk service is being considered as a source for additional participants. Although such participants would not be experts, the service was previously tested by Snow et al. [18] with similar annotation tasks and was found to be fairly accurate. Preliminary tests of Mechanical Turk for a use-mention annotation task have yielded positive results.

An eventual goal of this research is the automatic detection of mentioned language in a variety of contexts outside of Wikipedia. Although some retraining of a Wikipedia-based classifier is likely to be necessary, it is hypothesized that a core set of metalinguistic cues are shared across different language contexts. Additionally, previous research has shown that statistically-trained English parsers tend to make egregious errors when faced with even simple forms of mentioned language [9], and rectifying such errors is a further goal. Both goals are motivated by the considerable importance of the phenomenon in the human use of language.

7 Conclusion

This study has created a corpus of instances of mentioned language, using the particularly suitable properties of Wikipedia article text. The ML corpus is a unique resource that should provide a springboard for future research on the use-mention distinction and its relevance to a variety of language processing tasks. Results discussed in this paper show that patterns exist in mentioned language which can be utilized to expand the corpus and to apply machine learning techniques to it. This will eventually benefit both our understanding of the use-mention distinction and our ability to build language systems that recognize and exploit it.

Acknowledgments. Several conversations with Donald Perlis and Scott Fults were valuable in formulating this study. This research was supported in part by the US Office of Naval Research and the Air Force Office of Scientific Research.

References

1. García-Carpintero, M.: The deferred ostension theory of quotation. Noûs. 38, 674–692 (2004).
2. Cappelen, H., Lepore, E.: Varieties of quotation. Mind. 106, 429 -450 (1997).
3. Davidson, D.: Quotation. Theory and Decision. 11, 27-40 (1979).
4. Saka, P.: Quotational Constructions. Belgian Journal of Linguistics. 17, (2003).
5. Anderson, M.L., Okamoto, Y.A., Josyula, D., Perlis, D.: The Use-Mention Distinction and its Importance to HCI. In Proceedings of the Sixth Workshop on the Semantics and Pragmatics of Dialog. 21--28 (2002).
6. Perlis, D., Purang, K., Andersen, C.: Conversational adequacy: mistakes are the essence. International Journal of Human-Computer Studies. 48, 553-575 (1998).
7. Saka, P.: Quotation and the use-mention distinction. Mind. 107, 113 -135 (1998).
8. Miller, M.J.: A view of one's past and other aspects of reasoned change in belief, http://portal.acm.org/citation.cfm?id=164828, (1993).
9. Wilson, S.: Distinguishing Use and Mention in Natural Language. Proceedings of the NAACL HLT 2010 Student Research Workshop. pp. 29–33Association for Computational Linguistics, Los Angeles, CA (2010).
10. Maier, E.: Mixed quotation: between use and mention. Presented at the Logic and Engineering of Natural Language Semantics Workshop, Miyazaki (2007).
11. Garrette, D., Klein, E.: An extensible toolkit for computational semantics. Proceedings of the Eighth International Conference on Computational Semantics. pp. 116-127Association for Computational Linguistics, Tilburg (2009).
12. Balasuriya, D., Ringland, N., Nothman, J., Murphy, T., Curran, J.R.: Named entity recognition in Wikipedia. Proceedings of the 2009 Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources. pp. 10-18Association for Computational Linguistics, Suntec, Singapore (2009).
13. Honnibal, M., Nothman, J., Curran, J.R.: Evaluating a statistical CCG parser on Wikipedia. Proceedings of the 2009 Workshop on The People's Web Meets NLP: Collaboratively Constructed Semantic Resources. pp. 38-41Association for Computational Linguistics, Suntec, Singapore (2009).
14. Zesch, T., Müller, C., Gurevych, I.: Extracting Lexical Semantic Knowledge from Wikipedia and Wiktionary. Proceedings of LREC '08. European Language Resources Association (ELRA), Marrakech, Morocco (2008).
15. Ytrestøl, G., Flickinger, D., Oepen, S.: Extracting and annotating Wikipedia sub-domains. Proceedings of the Seventh International Workshop on Treebanks and Linguistic Theories. (2009).
16. Anderson, M.L., Fister, A., Lee, B., Wang, D.: On the frequency and types of meta-language in conversation: a preliminary report. 14th Annual Conference of the Society for Text and Discourse. (2004).
17. Fellbaum, C.: WordNet: an electronic lexical database. MIT Press, Cambridge (1998).
18. Snow, R., O'Connor, B., Jurafsky, D., Ng, A.Y.: Cheap and fast--but is it good?: evaluating non-expert annotations for natural language tasks. Proceedings of the Conference on Empirical Methods in Natural Language Processing. pp. 254-263Association for Computational Linguistics, Honolulu, Hawaii (2008).