Distinguishing Use and Mention in Natural Language

Shomir Wilson
Computer Science
University of Maryland
College Park, MD 20742, USA
`shomir@umd.edu`

Abstract

When humans communicate via natural language, they frequently make use of metalanguage to clarify what they mean and promote a felicitous exchange of ideas. One key aspect of metalanguage is the mention of words and phrases, as distinguished from their use. This paper presents ongoing work on identifying and categorizing instances of language-mention, with the goal of building a system capable of automatic recognition of the phenomenon. A definition of language-mention and a corpus of instances gathered from Wikipedia are discussed, and the future direction of the project is described.

1 Introduction

> *Costello: Well then who's on first?*
> *Abbott: Yes.*
> *Costello: I mean the fellow's name.*
> *Abbott: Who.*
> *Costello: The guy on first.*
> *Abbott: Who.*

In Abbott and Costello's famous routine "Who's on First?", Costello asks Abbott for the names of the players on a baseball team. In the above excerpt[1], Abbott tries unsuccessfully to explain that the name of the first baseman is *Who*, but Costello interprets this as another question instead of a response to his own. If Abbott had been more explicit and less terse (by responding with "The fellow's name is the word 'Who'", for instance), he would have avoided the ambiguity in his answers and might have succeeded in conveying to Costello the name of the first baseman. Instead, this misunderstanding is repeated throughout the routine with comic results, as the two become increasingly agitated by their breakdown in communication.

As Abbott and Costello unwittingly demonstrated, we sometimes must refer to the language that we speak and write in order to understand language itself. *Metalanguage* is our facility for doing this, and its interpretation often implicitly relies on the *use-mention distinction*—that is, simply, the distinction between using elements of language and mentioning them. In both written and spoken communication, the mention of letters, sounds, words, phrases, or entire sentences (termed *language-mention* in this paper for brevity) is essential for indicating titles, explaining meaning, introducing new words, attributing exact utterances to others, and other common functions of metalanguage (Saka 2005). There is evidence that human conversation makes frequent use of the use-mention distinction, and that we would be severely handicapped without it (Perlis et al., 1998). Moreover, this distinction has close ties to the appearance-reality distinction in cognitive science (Miller 1993).

It is surprising, then, that the use-mention distinction has thus far received little attention in computational linguistics. The need for greater study is unmistakable, as human audiences gener-

---

[1] Quoted from `http://www.phoenix5.org/humor/WhoOnFirst.html`.

ally navigate through this linguistic phenomenon with a natural ease that computers do not possess. The complexity behind this natural ease is apparent in our ability to understand simple sentences such as "Sky is spelled S K Y" and "The letters S, K, and Y make the word sky", which mean essentially the same thing but are structured and worded very differently. To gain the benefits of understanding the use-mention distinction, natural language processing systems must detect the subtle cues that signal this phenomenon.

This paper presents some preliminary results from a project on characterizing and identifying instances of language-mention in the English language. The use-mention distinction is first explained in greater detail, and a working definition is proposed for the phenomenon of language-mention. A corpus of instances of language-mention from Wikipedia is then introduced, with analysis of the categories in which the phenomenon appears to occur. The hypothesis of this continuing project is that lexical and syntactic cues will be sufficient to automatically identify at least a large subset of instances of mentioned language.

2 The Use-Mention Distinction

The use-mention distinction, as previously stated, is the distinction between using linguistic entities (such as letters, sounds, words, phrases, or entire sentences) and mentioning them. Since this explanation is slightly opaque at best and possibly circular, some examples and a proposal for a definition are appropriate. Consider the following sentences:

(1) The cat is on the mat.
(2) The word "cat" is spelled with three letters.

In (1), the reader's attention to meaning does not focus on the words themselves, but instead upon the presumed cat on the mat. In (2), the reader understands that the word *cat*—a string of three letters, as opposed to any particular cat or an abstract idea of a cat—is in the focus of the sentence. Quotation marks around *cat* in (2) are a convention to further reinforce that the word is being mentioned, and in some contexts (such as this sentence) italics may serve the same purpose.

The other linguistic entities listed above can also be mentioned, and the reader may easily conjure appropriate examples. Of particular note is *quotation*, a form of language-mention in which lan-

guage from another source is reproduced as part of a statement, as in (3) below:

(3) Eric said, "We should meet for lunch."

In (3), the phrase between quote marks is mentioned as what Eric has said. However, the reader is likely to treat the quoted text in (3) as a string with semantic depth, indicating that the *use* half of the use-mention distinction is present as well. Examples such as this illustrate that use and mention are not mutually exclusive (Maier 2007).

If writers always and consistently used cues such as quotation marks and italics, and if speakers followed a convention for delimiting mentioned utterances[2], recognizing language-mention would be an easier task. However, it frequently falls upon the intuition of the audience to determine when, where, and how it occurs (Anderson et al. 2002). Sentences (2) and (3) above, if typed less formally (sans quotation marks) or transcribed from speech, would still be easily understood by a human reader. Moreover, cues such as italics and quotation marks are also used for other purposes, such as distancing ("scare quotes") and emphasis, meaning that they are uncertain indicators of language-mention. It seems that subtler cues are responsible for our ability to distinguish use and mention.

In spite of the ubiquity of the phrase *use-mention distinction*, it is difficult to find an explicit definition for either the distinction itself or its two halves. The effort here will be to define language-mention, since this will aid in identifying where and how it occurs. What follows is a working definition, in the sense that it is a "rough draft"; suggestions for improvement are invited. For the moment, it restricts the scope of this work to *sentential* language-mention, where the mentioned linguistic entity is referred to inside of the same sentence that it occurs. (An example of a sentence that fails this additional requirement is: "Disregard the last thing I said.") This restriction is necessary to reduce the complexity of the identification and labeling problems, and it will be assumed for the rest of the paper.

Definition: For T a token or a set of tokens in a sentence, if T refers to a property of the token T or the type of T, then T is an instance of language-mention.

---

[2] One might observe that spoken language sometimes contains nonverbal cues for language-mention. While worthy of study, these cues fall beyond the scope of this paper, which will focus on written or transcribed language.

Here, a *token* can be any one of the linguistic entities listed at the beginning of this section—letters, sounds, words, phrases, or entire sentences. A *property* might be its spelling, pronunciation, original source (in the case of quotation), meaning (for a variety of interpretations of that term), or another aspect for which language is shown or demonstrated[3]. The *type* of T is relevant in some instances of language-mention (such as in (2)) and the *token* itself is relevant in others (including unusual cases such as "*The* is the first word in this sentence").

3 A Language-Mention Corpus

The second task of this project has been to create a corpus of sentences that contain instances of language-mention. The corpus will be valuable to move beyond laboratory examples and to begin mining for patterns in syntax and vocabulary that predict the occurrence of the phenomenon.

Wikipedia was chosen as a source of text for several reasons. Its text is freely available and covers a wide variety of subjects. Articles are written to be informative, which suggests that new names and terms are introduced frequently—a common function of language-mention. Contributors tend to highlight language-mention with italicization, bold text, or quotation marks. (This convention is mentioned in the Wikipedia Manual of Style, though it is unclear whether most contributors read it there or simply follow it out of habit.) While language-mention can certainly occur outside of those stylistic cues, the decision was made to concentrate on sentences that contained them, since this greatly accelerated the annotation process.

The annotation effort focused on the markup text of 1000 randomly chosen articles from English Wikipedia. Except for delimiters for bold and italic text, most of the markup was removed, and the remaining text was segmented into sentences using NLTK's implementation of the Punkt sentence tokenizer (Kiss and Strunk, 2006). The sentences then were filtered for only those that contained bold text, italic text, or text between single or double quotation marks, yielding a set of 1339 sentences that contained one or more of them.

Hand annotation required approximately three person-hours, with that time heavily skewed toward approximately the first third of the sentences,

as the set of categories for language-mention was also developed during this labeling process. Categories were formed with an informal "diagnostic test" of substitution of the category's theme (e.g., "this proper name", "this translation", "this symbol", "this quotation") in the place of the candidate token or tokens. Only text highlighted by one of the cues mentioned above was considered for labeling. Although only one researcher participated in the annotation, at the time of writing this paper an effort was in progress to build a much larger corpus using multiple annotators via Amazon's Mechanical Turk service. This service has shown promise in other natural language annotation tasks (Snow et al., 2008).

Out of the 1339 sentences inspected by hand, 171 contained at least one instance of language-mention. Many of those sentences contained several instances. Table 1 below lists the categories observed and the frequencies of each one, and Table 2 provides examples from each category.

Language-Mention Category	Frequency
Proper name (PN)	119
Translation or Transliteration (TR)	61
Attributed Language (AT)	47
Words/Phrases as Themselves (WD)	46
Symbols/Nonliteral Marks (SY)	8
Phonetic/Sound (PH)	2
Spelling (SP)	2
Abbreviation (AB)	1

Table 1: Frequencies of the different categories of language-mention found in the corpus.

Cat.	Example
PN	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.
TR	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".
AT	"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.
WD	"'Submerged forest'" is a term used to describe the remains of trees (especially tree

---

[3] These properties are based upon the ostentions of language in Paul Saka's treatment of the use-mention distinction (1998).

	stumps) which have been submerged by marine transgression, i.e. sea level rise.
SY	He also introduced the modern notation for the trigonometric functions, the letter "<u>e</u>" for the base of the natural logarithm (now also known as Euler's number) …
PH	The call of this species is a high pitched "<u>ke-ke-ke</u>" like American Kestrel.
SP	'''James Breckenridge Speed''' (middle name sometimes spelled '''<u>Breckinridge</u>''') (1844-1912) was a successful businessman in Louisville, Kentucky and an important philanthropist.
AB	… "Moskovskiy gosudarstvennyy universitet putej soobshcheniya", often abbreviated '''<u>MIIT</u>''' for '''"Moscow Institute of Transport Engineers'''" …

Table 2: Examples from the corpus of each category of language-mention. Triple quote marks indicate bold text in the original markup. The longer sentences for SY and AB have been truncated. The relevant instance of language-mention in each example appears underlined.

As shown, proper names were by far the most common category, with almost twice as many instances as the next most frequent category. This follows intuition, since Wikipedia articles often describe entities identified by proper names. In contrast, there were just a few instances of pronunciation (phonetic/sound) or spelling. Either the pre-filtering of sentences eliminated many instances of these before human annotation could find them, or Wikipedia is not a fertile source for them.

Of particular note are the 46 instances of words or phrases as themselves, since these are examples of language being either introduced or clarified for the reader. While there exists a body of work on named entity recognition (Nadeau and Sekine, 2007), very little exists on identifying when words serve a very similar function, essentially as rigid designators for their types. One of the future goals of this project will be to fill that gap.

4   Related Work

A similar corpus-building project was undertaken by Anderson, et. al (2004) to study the occurrence of metalanguage in human dialogue. In addition to the difference in focus (metalanguage broadly versus language-mention in particular), their project concentrated on the classification of utterances from casual speech, as opposed to the structure of well-formed sentences. The automatic recognition of language-mention will require a specific focus on the phenomenon, since it differs from other forms of metalanguage in its unusual syntactic structure (as shown in the next section).

In applications, the use-mention distinction has also received some treatment within dialog management and commonsense reasoning, as implemented in the ALFRED system (Josyula et al., 2003). However, its ability to recognize language-mention is limited to the task of learning new words from a limited set of sentence structures. The ongoing project described in this paper instead has the goal of recognizing and eventually interpreting language-mention in a wide variety of natural cases.

5   Future Work

The next step in this project will be to enlarge the language-mention corpus, using more data from Wikipedia and other promising sources. Language learning materials have also been considered for this purpose, as they necessarily contain a high frequency of metalanguage. The presence of stylistic cues in the text will be useful but perhaps not essential, as it is anticipated that bootstrapping the annotation process will become possible once enough indicators in sentence structure and vocabulary have been identified. This identification will be accomplished through a combination of eyeballing of patterns in parse trees and automated searching through the corpus using a tool such as Tregex (Levy and Andrew, 2006).

One eventual goal of this project is to detect language-mention without the presence of stylistic cues, with the intent of correcting egregious errors common in syntactic parsing of the phenomenon. Statistically-trained parsers have achieved great levels of accuracy at the macro level of examining large quantities of text, but this comes at a cost. Such accuracy tends not to include the phenomenon of language-mention, which often has an unusual structure. Consider the following two sentences paired with the resulting output from the Stanford Parser (Klein and Manning 2003):

(4a) Car is spelled c a r

(4b) (ROOT (S (NP (NNP Car)) (VP (VBZ is) (VP (VBN spelled) (S (NP (SYM c)) (NP (DT a) (NN r)))))))

(5a) The pronunciation of potato is pough tayh toe

(5b) (ROOT (S (NP (NP (DT The) (NN pronunciation)) (PP (IN of) (NP (NN potato)))) (VP (VBZ is) (NP (JJ pough) (NN tayh) (NN toe)))))

Both of these sentences are easily interpretable by a human audience, but the parser garbles their structure where language-mention occurs. Such unusual structure and vocabulary are likely not to lend well to the methods used to train such a parser. Because of this, the feasibility of a "hybrid" system is being investigated, which would combine an existing high-performance parser with a rule-based system to modify and correct its output where appropriate.

Preliminary work on a language-mention parser has shown the feasibility of this hybrid approach. A trial system has been built that uses parse trees produced by the Stanford Parser as input to five rules that detect common syntactic patterns indicating the phenomenon occurs in a sentence. In (4a), for instance, the presence of the verb *spell* and the sequence of two or more single-letter words indicates that the sequence is likely an instance of language-mention and falls into the category of spelling. Although language-mention exhibits substantial variety in its forms (and certainly will not be conquered by the five rules in the trial system), this approach should be able to take advantage of additional patterns mined from the corpus of the phenomenon currently being created. It is hypothesized that such a parser, using lexical and syntactic cues, will be sufficient to identify and categorize a large percentage of instances of language-mention in the absence of any stylistic cues.

References

Anderson, Michael L., Andrew Fister, Bryant Lee, and Danny Wang. 2004. On the frequency and types of meta-language in conversation: a preliminary report. Paper presented at the 14th Annual Conference of the Society for Text and Discourse.

Anderson, Michael L., Yoshi Okamoto, Darsana Josyula, and Don Perlis. 2002. The use-mention distinction and its importance to HCI. In *Proceedings of the Sixth Workshop on the Semantics and Pragmatics of Dialog*.

Josyula, Darsana, Mike Anderson, and Don Perlis. 2003. Towards domain-independent, task-oriented, conversational adequacy. In *Proceedings of IJCAI-2003 Intelligent Systems Demonstrations*.

Kiss, Tibor and Jan Strunk. 2006. Unsupervised multilingual sentence boundary detection. *Computational Linguistics,* 32(4): 485-525.

Klein, Dan and Christopher Manning. 2003. Accurate Unlexicalized Parsing. In *Proceedings of the 41st Meeting of the Association for Computational Linguistics*.

Levy, Roger and Galen Andrew. 2006. Tregex and Tsurgeon: tools for querying and manipulating tree data structures. In *Proceedings of the 8th International Conference on Knowledge-Based Intelligent Information and Engineering Systems*.

Maier, Emar. 2007. Mixed quotation: between use and mention. In *Proceedings of LENLS2007*, Miyazaki, Japan.

Miller, Michael. 1993. A view of one's past and other aspects of reasoned change in belief. Ph.D. thesis, University of Maryland, College Park, Maryland.

Nadeau, David and Satoshi Sekine. 2007. A survey of named entity recognition and classification. *Linguisticae Investigationes*, 30(1):3–26.

Perlis, Don, Khemdut Purang, and Carl Andersen. 1998. Conversational adequacy: mistakes are the essence. *International Journal of Human-Computer Studies,* 48:553-575.

Saka, Paul. 1998. Quotation and the use-mention distinction. *Mind,* 107(425):113–135.

Saka, Paul. 2005. Quotational constructions. *Belgian Journal of Linguistics,* 17(1):187–212.

Snow, Rion, Brendan O'Connor, Daniel Jurafsky, and Andrew Y. Ng. Cheap and fast—but is it good? Evaluating non-expert annotations for natural language tasks. 2008. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Honolulu, Hawaii.