

# Algorithms for Identifying the Multiple Syntactic Categories and Meanings of the Word *Over*

Yukiko Sasaki Alam

Dept. of Digital Media, Hosei University, Tokyo, Japan

Email: sasaki@hosei.ac.jp

**Abstract**—The word *over*, among others, is associated with a great variety of syntactic categories and meanings. Although *over* has received attention from scholars in different frameworks for various aspects, there is little research, to my knowledge, to take a comprehensive approach to both syntax and semantics of the word *over*. This article proposes algorithms for identifying from the context a case of *over* with one of four syntactic categories and with one of seventeen meanings. The test was carried out manually on five hundred instances of *over* from British National Corpus. The results are promising, with 95 percent of the classification of the instances being correct. This study, while taking an idealistic approach, brings to light methods that may stretch the limits of natural language processing.

**Index Terms**—preposition, particle, category and sense identification, lexicon, syntactic parsing, phrase boundary

## I. INTRODUCTION

In the past the word *over* has been studied in the literature of prepositions, and various aspects of prepositions have been examined: to name a few, the syntax in general ([11], [13]), the semantic network within a cognitive linguistic framework ([4], [15]), the pragmatic aspects ([5]), the deictic properties ([6]), the sense disambiguation ([1]), the labeling of such thematic roles as Place, Path, Temporal, Direction ([12], [17], [20]), the prepositional phrase (PP) attachment ([14]), the phrasal verbs ([9], [10]), and the verb-particle issues ([3], [16]).

The word *over*, however, appears in a variety of syntactic categories other than a preposition. Among 500 uses of *over* automatically and randomly chosen from British National Corpus (BNC) in the current research, 43.27% of the uses of *over* are not used as a preposition. Little research, however, has examined the syntactic aspects of *over*. The current article deals with the syntactic aspects, in particular, how to label the syntactic categories of *over* in terms of syntactic information, and

proposes algorithms for assigning the syntactic categories, enriched for the wholeness with the ability of providing the meanings of the prepositional uses of *over* adopted from my previous study ([1]), which is the result of research inspired by pedagogical studies ([8], [18], [19]) and the in-depth study of the preposition *over* in the cognitive linguistic framework ([14]).

Four syntactic categories of *over* (noun, adjective, adverb and preposition) and six meanings are identified in the present algorithm, together with the eleven meanings of the prepositional uses, which are carried over from [1]. The identification draws on the syntactic and semantic contexts in which *over* occurs. Although benefited much from previous research on prepositions, the current study takes a quite different approach from other such computational works, and hence offers no comparative discussion. The second section discusses the processes of assigning the syntactic categories and the meanings of *over*, in the order of nominal, in-phrase idiosyncratic, adverbial, adjectival, and prepositional uses. The third section evaluates the identification algorithms while discussing what remains to be solved. Some concluding remarks are offered in the fourth section.

## II. IDENTIFICATION PROCESSES

The processes begin with the identification of the syntactic categories of *over* other than the prepositional uses. The algorithm attempts to check for an idiosyncratic use of *over* among the non-prepositional uses. It has three idiosyncratic uses registered for checking: a deictic use as a noun, the use as an attributive adjective, and the use inside an idiomatic phrase. After not finding an idiosyncratic use of *over*, it checks for an adverbial use of *over*, and, if it detects no such use, it looks for the use as a predicative adjective.

When the algorithm finds the use of *over* to be none of a noun, adverb or an adjective, it assumes the use to be a prepositional use, and starts to determine the meaning because the prepositional use of *over* has eleven different meanings. Discussed below are further details of the processes for identifying the syntactic categories of *over* other than a preposition, and the meanings of the prepositional uses.

---

Based on "Algorithm for Identifying the Syntactic and Semantic Categories of Prepositions: Case of *Over*", by Yukiko Sasaki Alam, which appeared in the *Proceedings of the International Conference on Computing: Theory and Applications Platinum Jubilee Conference 2007*, Kolkata, India, March 2007. © 2007 IEEE.

### A Syntactic Categories Other Than A Preposition

According to a survey conducted in the current study, a little less than half of the uses of *over* do not appear as a preposition, and three other syntactic categories have been observed. They are noun, adverb and adjective, and the uses as an adjective are further divided into two types: as a predicative adjective and an attributive adjective. The meanings of *over* in the three categories other than a preposition are 'a number larger than that', Path, Side,<sup>1</sup> 'everywhere', 'finished' and 'more than'. Uses of *over* as a preposition occur in eleven meanings, depending upon the contexts. The identification moves from specific to general cases, and thus the idiosyncratic uses of *over* are first examined ahead of the adverbial and adjectival uses.

#### A.1 Idiosyncratic uses of *over*

The idiosyncratic uses of *over* are uses as a noun as well those inside idiomatic phrases.

##### a) Deictic uses as a noun

The algorithm for identifying the syntactic categories and the meanings (see Appendix A) begins by checking for the deictic uses of *over* such as *over* in:

- (1) [<sub>S</sub> *Entrants* [<sub>VP</sub> *must be* [<sub>NP</sub> *18 or over*]]].

This use is deictic because the meaning of *over* depends on the meaning of the noun preceding the conjunction *or*. In this example, it means 'a number larger than 18'. Syntactically this use of *over* has specific features of (Deictic Context 1) always following a conjunction such as *or* and *and*, and (Deictic Context 2) being followed by a closed noun phrase (NP) boundary. Therefore, to be successful in detecting and interpreting this use, the parser needs to be able to recognize syntactic boundaries such as open and closed NP boundaries. In this instance, it must identify that *over* is located immediately before a closed NP boundary, followed by a closed verb phrase (VP) boundary and a closed sentence (S) boundary.

Relying only on information on the conjunction preceding *over*, the parser will wrongly interpret *over* in such sentences as [<sub>S</sub> *Try* [<sub>PP</sub> *with pasta*] *or* [<sub>PP</sub> *over* [<sub>NP</sub> *poached fish*]]] as a deictic use. Although *over* is preceded by the conjunction *or*, it is not followed by a closed NP boundary, which is a boundary required for identifying this deictic use, as indicated in Deictic Context 2 above. In addition, this example violates another condition (Deictic Context 3) that the noun or the noun phrase preceding the conjunction must refer to quantity.

As the element preceding the conjunction is a numeral in most cases, a heuristic rule using this information might be formulated, but it would be unable to handle such an example as ... *mothers* [<sub>CP</sub> *who were* [<sub>PP</sub> *under age 20*] *or* [<sub>PP</sub> *over age 39*] *at their birth*]<sup>2</sup>. For this deictic use to be properly identified, the parser needs to be able to recognize syntactic boundaries such as phrase, clause and sentence boundaries, which are sometimes marked by a comma, an adverb, a relative pronoun and a

period. The syntactic category of this use should be noun, because it is conjoined with a noun, often a numeral, by the conjunction.

##### b) Attributive adjectives

Now we examine a special use of *over* which is immediately followed by a noun or noun phrase denoting quantity (Attributive adjective Context 1). The use in question is illustrated by the following examples:

- (2) a. [<sub>PP</sub> *During* [<sub>NP</sub> [<sub>NP</sub> *over 30 years*] [<sub>PP</sub> *of adventures*]]] *he met* ...  
 b. [<sub>NP</sub> *Over 60 people*] *were injured*.  
 c. [<sub>NP</sub> *A little over 60 people*] *were injured*.  
 d. [<sub>NP</sub> *Interest* [<sub>PP</sub> *over 10 years*]] *was* ...

*Over* in (2a) precedes a noun that indicates the number of years. *Over* in (2b) precedes a noun denoting the number of people. In addition to the restriction on what should follow *over*, this use of *over* appears at the beginning of the noun phrase (Attributive adjective Context 2). The phrases headed by *over* in (2a) and (2b) are both noun phrases. Although a modifying adverb may interfere between the open noun phrase boundary and *over*, as *A little* in (2c) does, yet it does not change the syntactic category of the phrase, and it would be safe to say that the *over*-phrase is a noun phrase. This second requirement (Attributive adjective Context 2) distinguishes the use as an attributive adjective from a prepositional use given in (2d), where *over* appears not at the beginning of a noun phrase, but at the beginning of a prepositional phrase, and the semantics of *over* differs from the semantics in (2a), (2b) and (2c).

The word *over* in this type modifies the following noun denoting quantity, giving rise to the meaning of excess of that quantity, and therefore it is regarded as an attributive adjective.

##### c) *Over* used inside idiomatic phrases

An idiomatic phrase that contains *over* is the next item to be searched for: among such are *over again*, *all over again*, *over and over*, *over and over again*, *over the moon*, and *over the counter*. The success in the detection of these in-phrase uses depends upon the inventory of such idiomatic phrases. An insufficient inventory would result in a wrong interpretation of *over* by the parser. The syntactic category of *over* inside an idiomatic phrase does not have to be specified, but rather the idiomatic expression as a whole must be assigned a syntactic category. For instance, *all over again*, *over and over*, and *over the counter* could be assigned the category of adverb or adverbial phrase.

The categories of some such uses would be unspecified between several categories, because they depend upon how the uses occur in the sentences. For instance, the category of adverbial phrase would be appropriate for *over the counter* in *You get them from the hospital or your doctor, not over the counter* while the category of adjective or adjectival phrase would be preferable in *over the counter medicines*. At any rate a syntactic category should be assigned to such an idiomatic expression as a whole, not just to *over*.

<sup>1</sup> *Path* and *Side* are the names of thematic roles.

<sup>2</sup> *CP* stands for *complementizer phrase*.

A phrasal verb that contains *over* is then checked for: among such are *cloud over*, *wash over*, *sweep over*, *gloss over*, *peer over*, *pore over*, *go over*, *get over*, *take over*, *hand over*, and *sign over*. Phrasal verbs are considered to be those that have metaphorical senses, as opposed to their counterparts that refer to physical events. In reality, it would not be so straightforward to determine whether *over* is part of a phrasal verb or not. Many phrasal verbs have the correspondents that should not be recognized as phrasal verbs. Observe the following:

- (3) a. *He went over the conversation with the lady again.*  
 b. *I went over the hills to bring her back.*

*Go over* in (3a) (followed by a noun phrase referring to an abstract concept) is a phrasal verb, whereas *go over* in (3b) (followed by a noun phrase representing a concrete entity) is the non-phrasal case, with the verb denoting directional movement and the preposition *over* indicating Path. To identify a phrasal verb from the non-phrasal counterpart, it would require semantic information of the noun phrase that follows *over*. The success in distinction between phrasal and non-phrasal uses rests on the inventory of phrasal verbs as well as provided semantic information that characterizes phrasal verbs. To equip the inventory with sufficient semantic information would not be simple in some instances. For example, in *He went over his notes again before the test*, it is not easy to determine whether the phrase *his notes* refers to an abstract entity (i.e. information) or a physical object. More research in this area is called for. The category of verb should be assigned to a phrasal verb as a whole.

#### A.2 Adverbs

After not finding the idiosyncratic uses of *over* mentioned in the subsection A.1, the algorithm for identification now checks whether *over* is followed by an open noun phrase (NP) boundary (Preposition Context 1), because if it is, it is probably used as a preposition, but otherwise, it is used as an adverb or an adjective. When *over* is followed by an open NP boundary like in [<sub>NP</sub> *the sign* [<sub>PP</sub> *over* [<sub>NP</sub> *the door*]]], it is a preposition followed by an NP, *the door* in this instance.

When *over* is not followed by an NP boundary (Non-preposition Context), the algorithm further checks whether *over* is located in the construction of the Subject of the sentence, followed by a BE verb, followed by *over* or *all over*. If not (Adverb Context 1), the category of adverb is assigned. Examples are:

- (4) a. [<sub>S</sub> *A cuckoo* [<sub>VP</sub> *flew over*]].  
 b. [<sub>S</sub> *He* [<sub>VP</sub> *went over* [<sub>CP</sub> *to pick up a coffee*]].  
 c. [<sub>S</sub> *The phone* [<sub>VP</sub> *is over* [<sub>PP</sub> *in the corner*]].  
 d. [<sub>S</sub> *He* [<sub>VP</sub> *bent over* [<sub>CP</sub> *to pick up the keys*]].  
 e. [<sub>S</sub> [<sub>VP</sub> *Place the cheese filling in the middle of the pastry*] and [<sub>VP</sub> *fold it over*]].  
 f. [<sub>S</sub> *I* [<sub>VP</sub> *am aching all over*]].  
 g. [<sub>S</sub> *I* [<sub>VP</sub> *used to go all over* [<sub>PP</sub> *to the Country Dances*]]].

*Over* in (4a) is immediately followed by a closed verb phrase (VP) boundary, followed by a closed sentence (S) boundary. *Over* in (4b) is followed by an open

complementizer phrase (CP) boundary, while *over* in (4c) by an open prepositional phrase (PP) boundary. All the three instances of *over* are followed by other than an open NP boundary, and it is used as an adverb. *Over* in (4a) and (4b) occurs with verbs denoting directional movement, indicating Path. On the other hand, *over* in (4c) appears with a verb indicating state and a prepositional phrase denoting a stationary place, and thus it indicates Side (the edge of a path), not Path.

*Over* in (4d) and (4e) are also followed by other than an open NP boundary: an open CP boundary and a closed VP boundary, both indicating that the objects are moved so that they are not straight, or flat any longer. The semantics of *over* in (4d) and (4e) is common to a group of verbs similar in meaning to *bend* and *fold*.

As in (4f), when *over* is preceded by *all* and followed by a closed verb phrase (VP) boundary, it, together with *all*, denotes ‘everywhere’. However, it should be noted that in such a case the verb should not indicate directional movement. If it does, as in (4g), *over* denotes Path. *Over* in (4g) is not followed by a closed VP boundary, but by an open PP boundary, and it is used as an adverb.

Unlike the prepositional use of *over*, these instances as an adverb are not followed by an open noun phrase boundary, and at the same time not located in the construction of the Subject + BE + [*all*] *over*, and hence they are identified to be adverbs, semantically indicating Path, Side or ‘everywhere’, depending upon the contexts.

#### A.3 Predicative adjectives

The syntactic category of *over* is adjective, meaning ‘finished’, when not followed by an open NP boundary (Non-preposition Context), which is the same requirement as for an adverbial use above, but located in the construction with a BE-verb as the predicate (BE-verb Context), and without *all* between the BE verb and *over* (Without-all Context). Examples are given below:

- (5) a. *The vacation is over.* [Subj. + BE + *over*]  
 b. *Don't talk about Chris. He is over.*

#### A.4 Instances of all over

Located in the Non-preposition Context and the BE-verb Context, with *all* existing between the BE verb and *over* (With-all Context), and possessing the Subject referring to a physical entity (Physical-Subject Context), the syntactic category of *over* is adverb and the meaning is ‘everywhere’, as in:

- (6) *The police are all over.*  
 [(physical) Subj. + BE + *all over*]

On the other hand, situated in the Non-preposition Context, the BE-verb Context, and the With-all Context, but having the Subject that does not refer to a physical entity (Non-physical-Subject Context), the syntactic category of *over* is adjective, meaning ‘completely finished’, as in:

- (7) *I told him we wouldn't be enemies once the war was all over.* [(abstract) Subj. + BE + *all over*]

To sum up the algorithm for handling uses of *over* in the construction of the Subject + BE + [*all*] + *over*, the following can be stated. Without *all* that precedes *over*,

*over* is an adjective, meaning ‘finished’. With *all* preceding *over*, the category depends upon the semantics of the Subject. When the Subject represents a physical entity, *over* is an adverb, meaning ‘everywhere’ while with the Subject referring to an abstract entity, it is an adjective, meaning ‘completely finished’.

To make the distinction, however, between ‘everywhere’ and ‘completely finished’ is not so straightforward when the Subject is the pronoun *it*, *this* or *that*, as in:

- (8) a. *In a few minutes it was all over.*
- b. *All you have to do is open your eyes and it is all over.*

In (8a) *it* could refer to cat fur, and *all over* means ‘everywhere’. Or *it* could mean *ordeal*, and *all over* means ‘completely finished’. For a proper interpretation, the referent of the pronoun must be correctly identified. The use of *all over* often has the pronoun *it* as the Subject of the sentence. In most cases, however, in a singular pronoun like *it*, *this* or *that*, the meaning is ‘completely finished’. Therefore, with the singular pronoun Subject in this construction, the default meaning could be set to ‘completely finished’, whereas with the plural pronoun, it could be set to ‘everywhere’.

A.5 A summary of non-prepositional uses

We have seen uses of *over* as a noun, an adverb and an adjective. These uses altogether occupy a little less than half of all the uses of *over*. The meanings indicated by these uses and the contexts for identifying the syntactic categories and meanings are summarized in Table 1.

TABLE 1. NON-PREPOSITIONAL USES OF *OVER*

Syntactic categories	Meanings	Contexts
Noun	number larger than that	N (quantity) + <i>or/and</i> + <i>over</i> + closed NP boundary (Ex. <i>4 or over</i> )
Adverb	<i>Path</i>	V (directional movement) + ( <i>all</i> ) <i>over</i> + PP (to-PP, etc.) (Ex. <i>He went over to Florida</i> )
	<i>Side</i>	V (state) + <i>over</i> + PP (Location) (Ex. <i>The phone is over in the corner</i> )
	everywhere	Subject (physical object) + BE + <i>all over</i> (Ex. <i>The Police are all over</i> ) Subject + [ <sub>VP</sub> V (non-directional movement) + <i>all over</i> ] (Ex. <i>I'm aching all over</i> )
Adjective	finished	Subject + BE + <i>over</i> / Subject (abstract entity) + BE + <i>all over</i> (Ex. <i>The harvest season is over/The war is all over</i> )
	more than that quantity	[ <sub>NP</sub> <i>over</i> + N (quantity)] (Ex. <i>Over 60 people came</i> )

B. Prepositional Uses of *Over*

We have seen uses of *over* as a noun, an adverb, part of an idiomatic expression, part of a phrasal verb, a predicative adjective, and an attributive adjective. When

*over* is none of these uses, it is a preposition. Once it is recognized as a preposition, the next task is to interpret the meaning, because the prepositional uses have several meanings.

B.1 Governor and Complement

To determine each meaning of the prepositional uses, two components need to be syntactically identified, the Governor and the Complement of the prepositional phrase, because the two are the elements that help identify the meanings of the prepositional uses. The Governor is a component that governs or calls for the prepositional phrase, and the Complement is a noun phrase that complements the preposition to construct together a prepositional phrase.

In Figure 1 below, in the verb phrase (VP) *fly over the sea*, the verb *fly* is the Governor of the prepositional phrase (PP) *over the sea*, and the noun phrase (NP) *the sea* is the Complement of the PP. In the noun phrase *disputes over authorship*, the noun *disputes* is the Governor of the PP *over authorship*, and the bare NP *authorship* is the Complement of the PP.

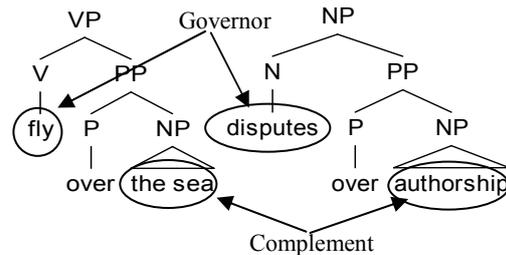


Figure 1. Governor and Complement

There are three types of Governor, although the Complement is always the noun phrase following *over*. What is the Governor depends upon the three grammatical functions played by the prepositional phrase: as an adverbial phrase that modifies a verb, a verb phrase (VP) or a sentence, as a nominal modifier that modifies a noun or a noun phrase, and as the predicate in the BE-verb construction. When the prepositional phrase (PP) is a nominal modifier, the PP is within a noun phrase as in:

- (9) [<sub>NP</sub> *disputes* [<sub>PP</sub> *over authorship*]]

The PP *over authorship* is inside the NP *disputes over authorship*. The Governor in such a case is (Case 1) the preceding noun, *disputes*, modified by the prepositional phrase *over authorship*. Therefore, once the use is recognized as a prepositional use, the first question to be asked by the algorithm is whether the *over*-headed PP is inside a noun phrase (NP) or not, and if it is, it assigns the preceding noun as the Governor, and the noun phrase following as the Complement.

If the prepositional use is not used as a nominal modifier, the algorithm next checks whether the Governor is the Subject of the sentence. To illustrate such a case, an example is given below:

- (10) [<sub>S</sub> [<sub>NP</sub> *The decision*] [<sub>VP</sub> [<sub>V</sub> *is*] [<sub>PP</sub> *over* [<sub>NP</sub> [<sub>CP</sub> *who goes on the cover*]]]]]]]

The tree diagram of the sentence is shown in Figure 2.

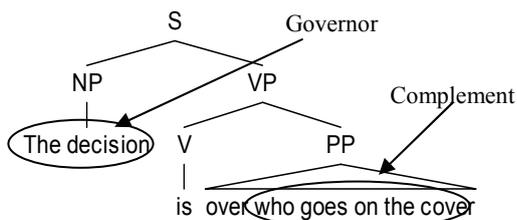


Figure 2. Governor in a sentence with a BE verb

Thus the second question to be asked after the identification of a prepositional use is whether the sentence is in the construction of Subject + BE verb + over-PP. If it is, (Case 2) the Subject is recognized as the Governor, as illustrated in Figure 2.

Finally, when the prepositional phrase is not inside a noun phrase (and thus not a nominal modifier) nor part of the predicate of the sentence in the BE-verb + over-PP construction, it is then an adverbial PP that modifies a verbal element. In the adverbial instance, (Case 3) the preceding verbal element is identified to be the Governor, and the noun phrase following *over* as the Complement. As illustrated in the example [VP fly [PP [P over][NP the sea]]] in Figure 1, the Governor is *fly*, and the Complement is *the sea*.

B.2 Identifying the meanings of the prepositional uses

Once the Governor and the Complement of the prepositional phrase are identified, the algorithm starts to determine the meaning of the prepositional use of *over*, first in terms of the semantics of the Governor and then in terms of the semantics of the Complement. The order of the checking process is from specifics to generics in this case, too. The semantics of the Governors and Complements that helps interpret the meanings of the preposition *over* is derived from Figures 1 and 2 in [1]. See Appendix B for the algorithm for interpreting the meanings of the prepositional uses of *over*.

The following tables are modified versions of the tables listed in [1].

TABLE 2. MEANINGS OF PREPOSITIONAL USES OF OVER (1)

Meanings	Identified by the Semantics of the Governor
<i>over_locus</i> (Ex. <i>Keep the net over the pond.</i> )	physical object or <i>stay/hit event</i>
<i>over_path</i> (Ex. <i>It flew over the sky.</i> )	<i>Movement_event</i>
<i>over_control</i> (Ex. <i>He presided over the sale.</i> )	<i>Control_event</i>
<i>over_prefer</i> (Ex. <i>Transformation was preferred over transaction.</i> )	<i>Prefer_event</i>
<i>over_about/over_because</i> (Ex. <i>Give elderly people choice over residential accommodation.</i> ) (Ex. <i>He quit over the failure.</i> )	<i>Communication_event/ Agree_event/ Psych_Cog_event/State</i> ( <i>over_about</i> when the governing event denotes a durative event while <i>over_because</i> when the Governor indicates a punctual event)

TABLE 3. MEANINGS OF PREPOSITIONAL USES OF OVER (2)

Meanings	Identified by the Semantics of the Complement
<i>over_coffee</i> (Ex. <i>Charles told me over tea.</i> )	an indefinite singular noun meaning meal or drink
<i>over_during</i> (Ex. <i>David has done a great job over the last few weeks.</i> ) (Ex. <i>That car's been parked there over a week.</i> )	an expression of time denoting duration: <i>over_during</i> when the expression is a definite NP while <i>over_duration</i> in an indefinite NP.
<i>over_means</i> (Ex. <i>It was sold over telephone.</i> )	<i>Communication_tool</i>
<i>over_many_parts</i> (Ex. <i>He has asked questions all over the village.</i> )	<i>All</i> precedes an NP meaning place or physical object.

The example sentences used in the original tables are replaced here. The second row of Table 2 shows that the meaning of the prepositional use of *over* is *over\_locus* in the context in which the Governor of the prepositional phrase denotes a physical object or an event of the class *stay/hit\_event*.

III. EVALUATION AND DISCUSSION

Five hundred instances containing *over* were selected automatically from British National Corpus (BNC) to test manually the ability of the algorithms. Ten instances are removed because they are incomplete sentences. The algorithm recognizes cases of *over* in (11a) and (11b) as adverbs:

- (11)a. *We have plenty of time to talk it over.*
- b. *Why don't you think it over for a while, and give me a call.*

It would be controversial whether these expressions should belong in phrasal verbs or not. The meaning of *over* in (11a) and (11b) is 'thoroughly', and *over* could be regarded as an adverb, so that the size of the inventory of phrasal verbs would be reduced. These examples suggest that the issue of phrasal verbs should be examined carefully together with the adverbial uses.

The algorithm identifies adverbial uses of *over* such as (4a) through (4e) successfully, but the semantic study in this area is in need, because the semantics of the adverbial uses varies according to the verbs they modify.

The algorithm fails to interpret the correct meaning of the following:

- (12) [NP Anything [? over this amount]] is...

This example seems to be similar to *women over 40* in terms of the semantics and the syntactic structure. It would be an open question whether *over* in (12) should be treated as an exceptional case or a preposition indicating excess. To identify this type of *over*, the following should be specified: the *over*-phrase should be a postnominal modifier located inside an NP, and the following noun phrase should denote quantity (a numeral in most cases). If we assume this case to be a prepositional use, another interpretation, which is to be

identified by the Complement, should be added to those presented in Table 3 as well as Appendix B.

The current algorithm assigns the meaning of 'place' (*over\_locus*) by default. Five instances obtain the default interpretation, which turns out to be a correct meaning. Among them are:

- (13) a. ... *may use a practice skirt for one selection which can be used over jeans, or another ...*  
 b. *Having applied cement render over the external brickwork, ...*  
 c. *I got up and had another little walk around, finishing up just over the road from ...*

The above verbs *use*, *apply* and *finish up* do not belong in any verb class listed in the algorithm, thus resulting in the assignment of the default meaning.

The algorithm fails to assign a correct interpretation to *over* in 17 instances (3.47%), where the default meaning is not a correct one. Among them are:

- (14) a. *Most of the time spent over it should be given to looking at its pictures.*  
 b. *In the debates over the last twenty years ...*  
 c. *I got up and walked quickly out of the coffee shop, up over the footbridge and on to the other platform ...*  
 d. ... *spending more time talking with journalists than working over hot ovens in the kitchen...*

(14a) is a case in which *over* is used interchangeably with *on*. The lexical entry for *spend* may need information about a possible use of *over* in place of *on*. In (14b) the algorithm wrongly assigns the meaning of 'about' (*over\_about*) because of the Governor *debates*, the semantics of which is checked before that of the Complement. The algorithm fails to assign the correct interpretation to *over* in (14c). It would be possible to interpret it correctly only when the meaning of the entire sentence is understood. The algorithm wrongly assigns the meaning of place (*over\_locus*) to *over* in *working over hot ovens in the kitchen* in (14d).

Table 4 below illustrates a breakdown of the uses of *over* in four hundred and ninety instances. About forty three percent of the uses of *over* appear in syntactic categories other than a preposition. A little over half of all the uses of *over* are used as a preposition, and a little more than a half of the meanings of the prepositional uses are identified by the Governors of the prepositional phrases while a little less than a half by the Complements. The default settings of the algorithms contributed to correct identification by one percent. The algorithms failed to identify about four percent of the uses of *over*.

For a successful implementation of the algorithms, the lexicon needs to have in the lexical entry for *over* a satisfactory list of idiomatic expressions containing *over* such as *over the counter* and *over and over*. It also needs to be furnished with a sufficient list of phrasal verbs containing *over*, together with information that helps distinguish from the non-phrasal counterparts. All the nouns and verbs in the lexicon must be categorized into semantic classes such as classes denoting physical object, abstract entity, quantity and directional movement. Verbs

and deverbal nouns must be provided with such an aspectual feature as  $\pm$ punctual (event).

The parser should be able to parse sentences properly into syntactic units so that the information on syntactic boundaries is available for determining the syntactic categories of *over*. It should be able to recognize if a noun phrase is definite and if it is singular. To identify the Governor, the parser needs to be able to determine if the *over*-prepositional phrase is a nominal modifier or not. If the PP is a nominal modifier, the preceding noun is the Governor while otherwise the predicate verb is the Governor. The task of determining is not trivial because the parser needs to understand semantic relations among the syntactic units. It is much easier to determine the Complement of the prepositional phrase, because the noun phrase that follows *over* is always the Complement. In view of the relative difficulty in determining the Governor, it might be more feasible to find the meaning of a prepositional use of *over* first by the Complement. The possibility remains to be explored.

TABLE 4. BREAKDOWN OF *OVER* IN 490 INSTANCES

Syntactic categories and Interpretations of <i>over</i>	count	%
<b>1. Non-prepositional uses (subtotal):</b>	<b>212</b>	<b>43.27</b>
idiomatic adverbial expressions	23	4.70
phrasal verbs	46	9.39
adverbs	61	12.45
excess 'more/higher/greater'	59	12.04
adjective 'finished'	23	4.69
<b>2. Prepositional uses (subtotal)</b>	<b>256</b>	<b>52.24</b>
<b>2.1 By the Governor (subtotal)</b>	<b>136</b>	<b>27.75</b>
a physical object over a place	14	2.86
<i>over_locus</i> (with <i>place/stay/hit</i> events)	33	6.73
<i>over_path</i> (with movement events)	37	7.55
<i>over_control</i> (with <i>control_events</i> )	15	3.06
<i>over_prefer</i> (with <i>prefer_events</i> )	2	0.41
<i>over_about</i> (with communication/ agree/psych/cognitive_events)	35	7.14
<i>over_because</i> (same as above)	0	0
<b>2.2 By the Complement (subtotal)</b>	<b>120</b>	<b>24.49</b>
<i>over_coffee</i> (with <i>meal/drink</i> nouns)	2	0.41
<i>over_during</i> (time nouns)	64	13.06
<i>over_duration</i> (time nouns)	37	7.55
all over (with place nouns)	15	3.06
<i>over_means</i> (with communication tool nouns)	2	0.41
<b>3. BY DEFAULT</b>	<b>5</b>	<b>1.02</b>
<b>4. FAIL</b>	<b>17</b>	<b>3.47</b>
<b>TOTAL</b>	<b>490</b>	<b>100.00</b>

#### IV. CONCLUSION

This article has proposed algorithms for assigning four syntactic categories (noun, adverb, adjective, and preposition) and seventeen meanings of the word *over*.

The algorithms draw on syntactic boundaries such as noun phrase and verb phrase boundaries as well as semantic classes of nouns and verbs. This suggests that the syntactic and semantic interpretation of multifaceted words like *over* requires the accurate parsing into phrases as well as the semantic classification of nouns and verbs.

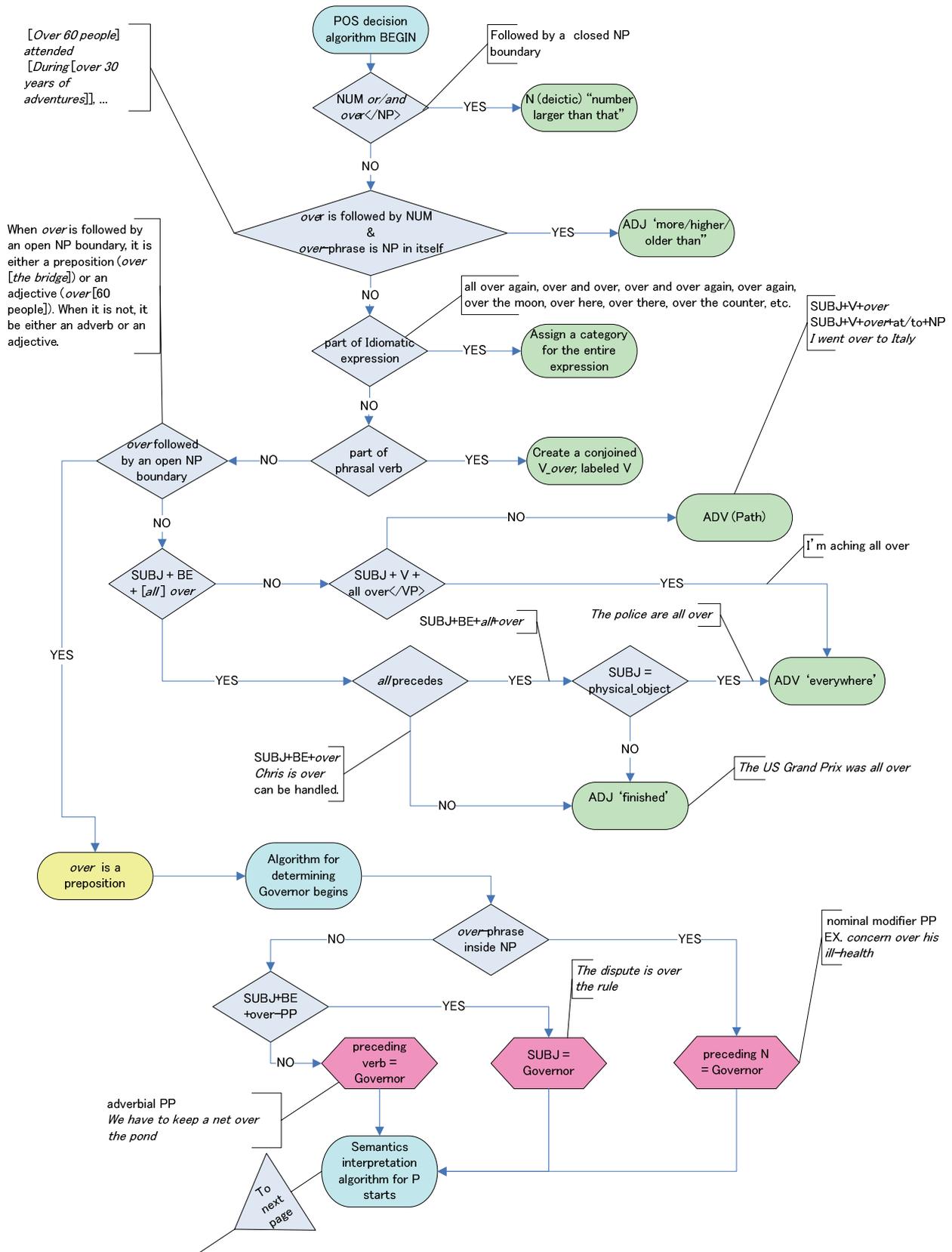
This study also reveals a need for study of the contexts that distinguish between phrasal verbs and the non-phrasal counterparts. The ability of the algorithms was tested manually by using five hundred instances of *over* from British National Corpus (BNC). The results are encouraging, with over 95 percent of the instances being correctly identified. This study, while pointing to an ideal direction, has revealed several aspects of what remains to be solved for a syntactic and semantic treatment of such words as *over* that are used in multifarious syntactic categories and meanings.

## REFERENCES

- [1] Y. Sasaki Alam, "Decision Trees for Sense Disambiguations: Case of *Over*", in *Proc. of the Workshop on Computational Lexical Semantics Held in Cooperation with HLT-NAACL*, 2004, pp.52-59.
- [2] Y. Sasaki Alam, "Algorithm for Identifying the Syntactic and Semantic Categories of Prepositions: Case of *Over*", in *Proceedings of the International Conference on Computing: Theory and Applications Platinum Jubilee Conference 2007*, Kolkata, India, 2007, pp. 434-440.
- [3] Colin Bannard, Timothy Baldwin and Alex Lascarides, "A Statistical Approach to the Semantics of Verb-Particles", in *Proc. of the ACL-2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, Sapporo, Japan, 2003, pp. 65-72.
- [4] Claudia Brugman, *The Story of Over: Polysemy, Semantics and the Structure of the lexicon*, Garland Press, New York, 1988.
- [5] Fauconnier, Gilles, *Mental Spaces*, Cambridge University Press, Cambridge, 1994.
- [6] Clifford Hill, "Up/Down, Front/Back, Left/Right. A Contrastive Study of Hausa and English", In J. Weissenborn and W. Klein, eds., *Here and There Cross Linguistic Studies on Deixis and Demonstration*, 1982, pp. 13-42.
- [7] Beth Levin, *English verb classes and alternations*, The University of Chicago Press, Chicago, 1993.
- [8] Seth Lindstromberg, *English prepositions explained*, John Benjamins, Amsterdam, 1998.
- [9] Wei Li, Xiuhong Zhang, Cheng Niu, Yuankai Jiang and Rohini Srihari, "An Expert Lexicon Approach to Identifying English Phrasal Verbs", in *Proc. of the 41<sup>st</sup> Annual Meeting of the ACL*, 2003, pp. 513-520.
- [10] D. McCarthy, B. Keller, and J. Carroll, "Detecting a Continuum of Compositionality in Phrasal Verbs", in *Proc. of the ACL-SIGLEX Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, Sapporo, Japan, 2003.
- [11] Geoffrey Pullum, and Rodney Huddleston, "Prepositions and Prepositional phrases" In R. Huddleston and G. Pullum, eds., *The Cambridge Grammar of the English Language*, 2002, pp. 597-661.
- [12] Tom O'Hara, and Janyce Wiebe, "Preposition Semantic Classification via Treebank and FrameNet", in *Proc. Conference on Natural Language Learning (CoNLL-03)*, Edmonton, Canada, 2003.
- [13] Gisa Rauh, "On the Grammar of Lexical and Non-lexical Prepositions in English", In Zelinsky-Wibbelt, eds., *The Semantics of Prepositions*, 1993, pp. 99-150.
- [14] Joseph M. Sopena, Agusti LLoberas and Joan L. Moliner, "A Connectionist Approach to PP Attachment for Real World Texts", in *COLING-ACL '98*, 1998, pp. 1233-1237.
- [15] Andrea Tyler and Vyvyan Evans, "Reconsidering Prepositional Polysemy Networks: The Case of *Over*", *Language* 77, 2001, pp. 724-765.
- [16] Aline Villavicencio, "Verb-Particle Constructions and Lexical Resources", in *Proc. of the ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*, Sapporo, 2003, p. 57-64
- [17] Clare Voss. *Interlingua-based machine translation of spatial expressions*, Ph.D. Dissertation, University of Maryland, 2002.
- [18] Gloria Wahlen, *Prepositions illustrated*, The University of Michigan Press, Michigan, 1995.
- [19] Jean Yates, *The ins and outs of prepositions A guidebook for ESL students*, Barron's, New York, 1999.
- [20] Patrick Ye and Timothy Baldwin. "Semantic Role Labelling of Prepositional Phrases". In *Proceedings of the Second International Joint Conference on Natural Language Processing (IJCNLP-05)*, Jeju, South Korea, 2005, pages 779-791.

**Yukiko Sasaki Alam** is currently a full professor at the Department of Digital Media of Hosei University, Tokyo, Japan. Her main research areas include machine translation with current focus on the lexicon and educational software.

APPENDIX A ALGORITHM FOR IDENTIFYING THE SYNTACTIC CATEGORIES AND MEANINGS OF *OVER*



APPENDIX B ALGORITHM FOR IDENTIFYING THE MEANINGS OF THE PREPOSITIONAL USES OF *OVER* (ADOPTED AND MODIFIED FROM THOSE ON FIGURES 1 AND 2 IN [1])

