

# A corpus-based semantic feature analysis of [*down* the NP]<sup>1</sup>

Siaw-Fong Chung  
National Chengchi University / Taiwan

**Abstract** – Using data from the *Wall Street Journal Corpus*, this study investigates the connection between the preposition *down* and the NPs following it in the [*down* the NP] construction. The article demonstrates that the NPs are characterized with features such as ‘concreteness’, ‘continuity’, and ‘verticality’, and that a majority of the NPs following *down* share a certain degree of semantic similarity. The results of this study support the idea that a lexicon in a fixed frame cannot always exert influence on nearby words.

**Keywords** – corpus, preposition, *down*, NP

## 1. INTRODUCTION

Spatial prepositions, also called projective prepositions in Herskovits (1986), have been widely studied (Lakoff 1987; Boers 1996; Grabowski and Miller 2000; Tyler and Evans 2003; van der Zee et al. 2009; Lindstromberg 2010; Vorwerg and Weiß 2010; Geld and Krevelj 2011). In many of these studies, researchers investigated how people comprehend the dimensional concepts encoded in prepositions, for example, the examination of *up* in Geld and Krevelj (2011), *over* in Lakoff (1987), and *near* and *far* in van der Zee et al. (2009), which dealt with the polysemous nature of prepositions (ie. different senses realized by one single prepositional form) and image schemas. In addition to the senses and image schemas for prepositions, researchers such as Grabowski and Miller (2000) channeled their foci on how people perceive and produce the dimensional prepositions *in front of* and *behind*. Later, Vorwerg and Weiß (2010) proposed other possible factors influencing one’s cognitive processing of prepositions and proved that co-appearing verbs do come into play. In addition to these physical and geometrical investigations of spatial concepts, the non-geometrical aspects of prepositions were also examined in Boers (1996) and Lindstromberg (2010), two studies which demonstrated that prepositions can denote temporal readings stemming from different metaphors. In addition to the analysis of prepositions in one particular language, there have also been cross-language examinations of spatial terms (eg. Grabowski and Weiß 1996; Grabowski and Miller 2000; Vorwerg and Weiß 2010).

Some studies conducted research on spatial terms from the perspective of contrasting pairs. For example, Tyler and Evans (2003) and Lindstromberg (2010) analysed *down* in contrast with *up*, and claimed that these two prepositions do not form a simple oppositional pair (Tyler and Evans 2003: 142). In line with other studies regarding prepositions, the geometrical senses and image schemas resulting from the spatial meaning of *down* were also revealed. Moreover, the research illustrated that the paths denoted by *down* involve different metaphors and that the downward sense could be extended for temporal meanings.

Most of these findings from previous studies on the dimensional preposition *down* can be tested on corpus instances. For instance, *down* can be used to convey non-vertical paths, while it prototypically refers to vertical paths, as

---

<sup>1</sup> This research was supported by the research funding of 101-2410-H-004-176-MY2 by the Ministry of Science and Technology, Taiwan. Chao-lin Liu’s help was highly appreciated in extracting data from the BLLIP. My gratitude also goes to Min-Chien Lee for her assistance with this research. Min-Chien extended the topic of this research by using a different set of data as part of her dissertation, but expressed no interest in participating in the revision of this article.

Lindstromberg (2010) has claimed; in addition, the concept of being *down* can encode an end-point focus of a path as shown by Tyler and Evans (2003: 145). However, searching through a dictionary, it seems that an investigation of the meaning of *down* alone ('from a higher to a lower part of (something)' or 'along the course or path of (something)') (*Merriam-Webster's Learner's Dictionary*) cannot explain why *down* collocates with certain NPs. In other words, if we examine the structure [*down* the NP], the ideas of the path emitted by *down* do not perfectly abide by the dictionary definitions (ie. objects going down being vertically positioned and without an end-point focus). *Down* is certainly more complex than what has been emphasized by previous scholars and dictionaries.

In this article, the structure [*down* the NP] is analyzed in order to comprehend the complexity of the preposition *down* and to investigate whether its co-occurring NPs conform to what the dictionary definitions and some previous studies have described, or whether, on the contrary, the dictionaries and the specialized literature have left something behind. The study examines whether the NPs following *down* can be systematically categorized according to a certain number of shared characteristics (eg. 'continuity' and 'verticality'). Through such an examination of the connection between *down* and the NPs following it, it is hypothesized that the NPs will share certain characteristics. In addition, it is believed that it is the neighboring words in the [*down* the NP] structure that influence each other to coerce certain semantic properties.

## 2. METHODOLOGY

The data taken from the *BLLIP 1987–89 Wall Street Journal Corpus Release 1* (Charniak et al. 2000) were used in this research. The BLLIP is a corpus containing Penn Treebank annotation. The following description was taken from the LDC website:<sup>2</sup>

The PTB project selected 2,499 stories from a three-year WSJ collection of 98,732 stories for syntactic annotation. These 2,499 stories have been distributed in both Treebank-2 (LDC1999T42) and Treebank-3 (LDC1999T42) releases of PTB. Treebank-2 includes the raw text for each story.

To extract the data, a computer program was designed to automatically retrieve all instances of prepositional phrases that appear within a verb phrase, as in example (1) below. The sentence contains the Treebank annotation that includes the whole verb phrase *presented the lawyers down the hall*, with an annotation of parts of speech and layers of syntactic structure. In this search, *down the hall* was analyzed as a chunk.

(1) (VP (VBD presented) (NP#1012 (DT the) (NNS lawyers)) (PP-DIR (IN down) (NP (DT the) (NN hall))))

For the analysis, the following semantic features were considered in this study. Table 1 presents these features, their values, the labels we will use hereafter in this article to refer to each of them, definitions, and examples:

Semantic Features	Values of the Features	Name Used Hereafter	Definitions	Examples
Concreteness	[+concreteness]	Concrete	Refers to observable people or things.	<i>... runs her fingers down the neck of a Polar bottle ...</i>
	[-concreteness]	Abstract	Refers to abstract ideas or non-observable qualities.	<i>... been serving the cause of free expression down the years ...</i>
Continuity	[+continuity]	Continual	Refers to continuous, long, and lasting things.	<i>... paddling down the Hudson River ...</i>
	[-continuity]	End-pointed	Refers to things that might have length but have an end-point in the end.	<i>... runs his finger down a list ...</i>
Verticality	[+verticality]	Vertical	Refers to things that are positioned in an up-and-down vertical direction.	<i>... send shivers down the spine ...</i>
	[-verticality]	Horizontal	Refers to things that are positioned in a side-to-side horizontal direction.	<i>... move his offices from Seagram's headquarters here to Claridge's new offices down the street ...</i>

Table 1. Semantic features of the NPs in the [*down* the NP] construction

For each semantic feature there are two values: plus [+] and minus [-]. To correspond to each feature with a value, a name was assigned to each feature (see the third column in Table 1), which will be used hereafter in this article when referring to instances bearing the feature of a specific value. For the first feature, 'concreteness', NPs were categorized as either [+concreteness] or [-concreteness]. The former refers to concrete NPs such as *a way, a route, a leg, a neck, a list*, etc., while the latter includes such abstract NPs as *a year, income, performance*, and so forth. Under the feature 'continuity', the data were grouped as continual or continual with an end-point. For example, *the chart* and *the curve*

<sup>2</sup> <https://catalog.ldc.upenn.edu/LDC2000T43>.

tend to have an end-point at the end of a list so they were grouped as [-continuity], while *the gulf*, *the beach*, and *the street* were grouped as having a [+continuity] value since they have a limitless length. The NPs encoding a ‘verticality’ feature were either the [+verticality] value, such as *the stair* and *the neck*, or the [-verticality] value, with instances like *the coast* and *the sidewalk*, which are positioned in a horizontal direction.

### 3. DATA ANALYSIS

A total of 358 noun phrases for the 356 instances of [down the NP] construction were extracted from the corpus.<sup>3</sup> The data were first grouped into two classes according to the first feature (‘concreteness’). They were then categorized on the basis of the criterion of whether they were observable objects or abstract ideas (see Table 2).

NPs	Number of instances	Percentage (hits/358)	Examples
Concrete	271	75.70%	<i>street, beach, hall, list, throat, staircase</i>
Abstract	85	23.74%	<i>year, summer, level, performance</i>
Erroneous	2	0.56%	<i>payment of purchase, type</i>
Total	358	100.00%	

Table 2. NPs with the ‘concreteness’ feature

For all the instances, *down* collocated mostly with concrete NPs (around 76 percent), as in *some old folks are chasing Congressmen down the street to protest*, *drink at the liquor store down the hall*, and *cram a bunch of securities down investors’ throats*. Around 24 percent of the data co-occurred with abstract NPs, as in *been turned down the past two years*, *push decision-making down a couple of levels to speed up the process*, and *dragged down Xerox’s overall performance*. Two instances were categorized as ‘erroneous’, since the computer program mistakenly retrieved the wrong tree structure for [down the NP]. For example, *types in pressing the right index finger down types a J* is a verb, not an NP.

To examine whether the NPs involved in the structures at issue shared a certain degree of similarity, they were further grouped according to the ‘continuity’ and ‘verticality’ features. The results presented in Table 3 below show concrete NPs with the ‘continuity’ feature, including the number of instances, percentages, and examples of continual NPs and NPs with end-points:

The meaning of [down the NP]	Number of instances	Percentage (hits/271)	Examples
Continual	240	88.56%	<i>road, route, river, coast, gulf, beach, track, tube</i>
End-pointed	31	11.44%	<i>draft, chart, list, cliff face, hill, door, cellar, gas tank</i>
Total	271	100.00%	

Table 3. Concrete NPs with the ‘continuity’ feature

Of these 271 instances, nearly 89 percent of the NPs in the [down the NP] construction were continual, while only 12 percent had both a limitless length and end-points, as the examples in (2) and (3) below demonstrate:

- (2)
- a. Now I drive the tractor in drunken-looking swerves *down the dirt road*.
  - b. IBM says the Micro Channel zips data through a PC like a race car *down a super highway*.
  - c. Mr. Vachon says the government used to pour the liquor *down the drain* – until it realized that the practice violated pollution laws.
  - d. Institutional investors look for sound business arguments now before selling managements *down the river*.
  - e. At the least, Piedmont would greatly strengthen USAir’s presence in the Southeast and extend its operations *down the East Coast*.
- (3)
- a. The White House wanted Mr. Bork depicted as open-minded and respectful of precedent, and as a result toned *down the Justice Department’s draft*.
  - b. As the personal computer becomes standard office fare, the question of how to handle the 8% to 10% of users who seem to fixate on the costly machines has dogged managers up and *down the organizational flow-charts*.
  - c. Surprising things happened: A non-swimmer deathly afraid of water made the rapids trip; an executive afraid of heights made the journey *down the cliff face*.
  - d. Mr. Hornsby boasted that one of his graduates, surprised while snoozing by a would-be thief in the saloon bar of a Lewisham pub, lunged at the intruder, “tripped him and threw him *down the cellar*”.

In (2a) to (2e), the NPs (ie. *the dirt road*, *a super highway*, *the drain*, *the river*, and *the East Coast*) have a limitless length without end-points; thus, the meanings of the [down the NP] construction in these instances were continual. On the other hand, *the Justice Department’s draft*, *the organizational flow-charts*, *the cliff face*, and *the cellar* in (3a) to

<sup>3</sup> If an NP structure had two nouns connected by means of a conjunction (eg. *the Gulf and the East Coast* and *a full pfennig and about one-half yen*), the two nouns were counted separately and analyzed in this study (eg. *Gulf* and *East Coast*).

(3d) are NPs with end-points. Consequently, *down the Justice Department's draft*, *down the organizational flow-charts*, *down the cliff face*, and *down the cellar* can be thought of as conveying an idea of having an end-point.

The further analysis of the 240 instances of the concrete and continual NPs from Table 3 revealed that they shared a similarity in terms of the 'verticality' feature (see Table 4).

The meaning of [down the NP]	Number of instances	Percentage (hits/240)	Examples
Vertical	47	19.58%	<i>tube, stair, ladder, neck, throat, spine</i>
Horizontal	191	79.58%	<i>road, driveway, aisle, hall, gulf, beach, coast</i>
NA	2	0.84%	<i>black hole</i>
Total	240	100.00%	

Table 4. Concrete, continual NPs with the 'verticality' feature

Table 4 shows that around 80 percent of the concrete and continual NPs were horizontally positioned, while only 20 percent were vertical NPs, as shown in (4) and (5) below:

- (4)
- a. Shortly after 8 p.m., two sets of bright headlights begin making their way *down the long driveway of Levi Segopolo's farm*.
  - b. As a bulldozer pulled the roller *down the crowded beach*, it molded the sand – along with banana peels and cigarette butts – into an ankle-high cityscape.
  - c. She would have had to buy the alcohol part of her drink at the liquor store *down the hall*.
- (5)
- a. As a young attorney, he once punched a complaining client *down the courthouse stairs*.
  - b. During the eight days the peregrines remained in the hack-box before release, my partner and I both rappelled down to it, dropped quails *down the feeding tube* in the top of the box, then observed the falcons' highly individual behavior, movements, quirks and habits through peepholes in the box.
  - c. History never repeats itself exactly, but enough sometimes to send shivers *down the spine*.

In these examples, *the long driveway of Levi Segopolo's farm*, *the crowded beach*, and *the hall* were categorized as continual NPs positioned in a side-to-side horizontal direction. On the contrary, *the courthouse stairs*, *the feeding tube*, and *the spine* were positioned in an up-and-down vertical direction. The meanings of [down the NP] for the former cases are horizontal, while those for the latter ones are vertical. There were two instances in which the 'verticality' feature did not apply. In *down a black hole*, for example, we could not identify whether the black hole was vertically or horizontally positioned.

As shown in Table 2, for all the 271 instances of concrete NPs, there were 31 NPs that had an end-point. These concrete NPs with end-points were analyzed according to the 'verticality' feature, as shown in Table 5 below:

The meaning of [down the NP]	Number of instances	Percentage (hits/31)	Examples
Vertical	13	41.94%	<i>door, cliff face, chart, list, curve, crack, dirt</i>
Horizontal	0	0.00%	
NA	18	58.06%	<i>The Grinder, 31-24-40-8, line</i>
Total	31	100.00%	

Table 5. Concrete, end-pointed NPs with the 'verticality' feature

Of these 31 instances, 42 percent of the data represented vertical NPs. The rest of the NPs were categorized as 'NA' because the feature of 'verticality' did not apply, although the NPs were considered as having an end-point. These NPs referred to people or numbers (eg. *the Grinder* and *31-24-40-8*); otherwise, the whole construction of [down the NP] was idiomatic, so a discussion of the NPs alone was meaningless (eg. *down the line*, which means 'in the future'). Some of these 13 instances with vertically positioned NPs are listed in (6) below:

- (6)
- a. On a sunny afternoon three months ago, half a dozen pistol-packing U.S. Customs Service agents battered *down the door of Ramon Cernuda's luxury condo* overlooking Biscayne Bay.
  - b. All large firms in this industry want to make products like DRAMs and Eproms because producing these high-volume chips helps drive a firm *down a very steep "experience curve"*.
  - c. Stuff that might get lost *down cracks* ...

*The door of Ramon Cernuda's luxury condo* in (6a), *a very steep "experience curve"* in (6b), and *crack* in (6c) have shared features: unlike the coast or the street, which are long and continual (ie. have a limitless length), they have a certain length, but also possess end-points, and they are all vertically positioned. These instances exhibit a continual meaning but there is an end-point in the end.

To summarize briefly, from Tables 4 and 5, a total of 191 instances had horizontally placed NPs, which accounted for over 50 percent of the data (ie.  $191/358 = 53\%$ ). This suggests that the 'verticality' semantic feature described in the definition of *down*, 'from a higher to a lower part of (something)', did not fully capture the sense of the [down the NP] structure.

In accordance with the data of concrete NPs, those with abstract NPs were analyzed according to the ‘continuity’ feature. However, unlike the concrete NPs discussed above, the ‘verticality’ feature does not apply to the analysis of abstract NPs. The number of instances, percentages, and examples of abstract NPs with the ‘continuity’ feature are presented in Table 6 below:

The meaning of [ <i>down</i> the NP]	Number of instances	Percentage (hits/85)	Examples
Continuity	0	0.00%	
End-point	85	100.00%	<i>totem pole, ranks, years, summer, value, 4 1/2 points, speech</i>
Total	85	100.00%	

Table 6. Abstract NPs with the ‘continuity’ feature

All the 85 instances with abstract NPs were considered as having a certain length and an end-point. Examples of these abstract NPs are given in (7) below:

- (7)
- a. But he has been turned *down the past two years*.
  - b. Mr. Baker, who disagreed with this approach, made a last-minute fight to tone *down the speech*.
  - c. And there’s at least a chance Texaco might have to be liquidated, which would put stockholders way *down the totem pole*.
  - d. In recent years an increasing number of companies have learned that pushing decision-making *down the ranks* is a management luxury they can’t always afford, especially in a competitive environment.

For example, *the past two years* in (7a) and *the speech* in (7b) have a certain length in time that will end in the future, while *the totem pole* in (7c) and *the ranks* in (7d) can be considered as having a non-observable length that will end in terms of the level of the totem pole or ranks. The preposition *down* found with these NPs is end-pointed as well.

Tables 3 and 6 show that 240 out of the 358 relevant instances bear the continuity trait. That is, nearly 90 percent of the NPs that collocate with *down* do not emphasize an end-point. One of the dictionary definitions of *down* (‘along the course or path of (something)’) seems to have captured this feature of the continuity meaning.

#### 4. DISCUSSION

The results presented in this article demonstrate that when NPs are concrete observable objects and abstract ideas, 67 percent of the [*down* the NP] sequences (240/358) conveyed a continual meaning (eg. *down the road*), while nearly 33 percent ((31+85)/358) presented a meaning that something was long or lasting in space or time but would end at some point (eg. *down the gas tank* and *down the night*). Therefore, it seems that there are two continual senses for [*down* the NP] depending on whether or not there is an end-point. We refer to these meanings as (a) move ‘along’ the NPs (ie. PATH without an end) and (b) move ‘toward’ the NPs (ie. PATH with a final end). Examples of the former type include *chasing Congressmen down the street to protest*, *paddling down the Hudson River*, and *poured more than 80 pots of spaghetti sauce down the drain*, while examples of the latter type include *running his finger down a hand-written list*, *made the journey down the cliff face*, and *force the stock down late last year*. The meanings and the number of instances with the percentages of the [*down* the NP] construction are presented in Figure 1 below:

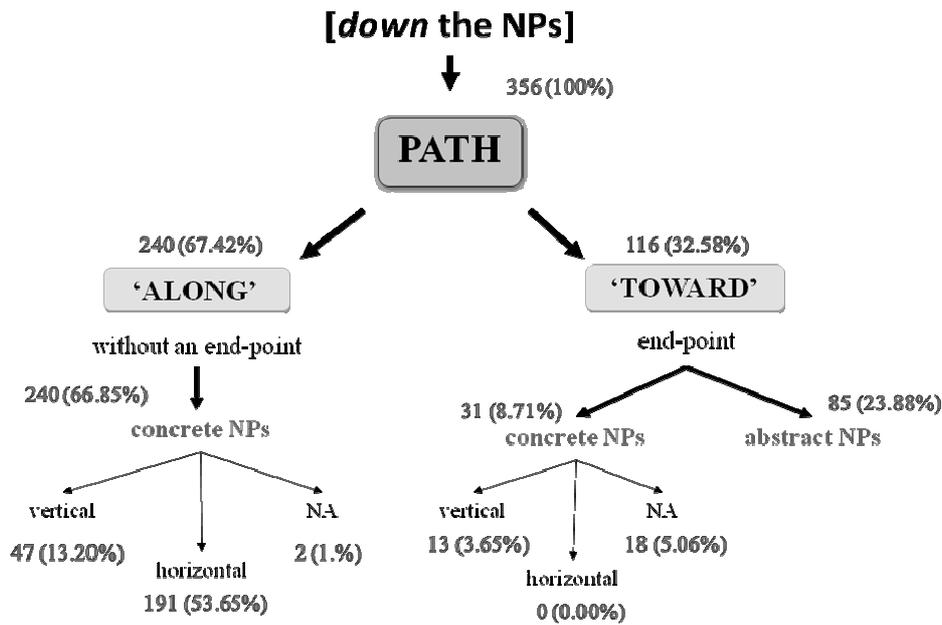


Figure 1. The meanings of the [down the NP] construction<sup>4</sup>

The first type of PATH meaning is to move 'along' the NPs (ie. PATH without an end). Examples belonging to this type accounted for 67.42 percent of the data, and 191 out of 240 instances denoted a horizontal direction, for example, *walked a pig down a Dallas street* and *a bulldozer pulled the roller down the crowded beach*. In these cases, the PATH meaning carried by the [down the NP] construction did not seem to end with a specific final point, so in this case the continuing PATH was emphasized. The actions of walking and pulling continued or lasted as long as the objects *a pig* and *a bulldozer* moved horizontally along the NPs (ie. *a Dallas street* and *the crowded beach*).

The second type of PATH for the [down the NP] construction is to move 'toward' the NPs (ie. PATH with a final end), which accounted for 32.58 percent of the data. This type of PATH meaning conveyed a continual motion that would end at some point. For instance, the action of throwing *him down the cellar* ended when the intruder was inside the cellar; the action of *pushing his siblings down a mound of dirt* ended when his siblings touched the ground; and the abstract idea of the action *selling T-shirts decorated with 10 buttons down the front last year* ended at the last moment of the year.

As shown in Figure 1, 67.42 percent of the data containing the [down the NP] construction denoted a continual meaning, while 32.58 percent of the examples carried an end-pointed meaning. Moreover, only 16.85 percent of the NPs conveyed a meaning related to the vertical direction. However, the dictionary senses of the preposition *down* (ie. 'along the course or path of (something)' and 'from a higher to a lower part of (something)') emphasize continuity and verticality meanings. Therefore, our analysis of the characteristics of the relevant NPs reveals that it is the neighboring words, the NPs in the [down the NP] construction, that coerce certain semantic properties that *down* does not have by itself. Thus, it can be concluded that, given the complexity of prepositions, as shown here with the construction [down the NP], a single examination of a dictionary does not always display their semantic meanings in a fixed frame.

#### REFERENCES

- Boers, Frank. 1996. *Spatial prepositions and metaphor: a cognitive semantic journey along the up-down and the front-back dimensions*. Tübingen: Gunter Narr.
- Charniak, Eugene, Don Blaheta, Niyu Ge, Keith Hall, John Hale and Mark Johnson. 2000. *BLLIP 1987-89 WSJ Corpus Release 1 LDC2000T43*. DVD. Philadelphia: Linguistic Data Consortium.
- Geld, Renata and Stela Letica Krevelj. 2011. Centrality of space in the strategic construal of *up* in English particle verbs. In Mario Brdar, Marija Omazić, Višnja Pavičić Takač, Tanja Gradečak-Erdeljić and Gabrijele Buljan eds. *Space and time in language*. Frankfurt and New York: Peter Lang, 145–166.

<sup>4</sup> Two instances classified in the 'erroneous' group have not been included in this figure.

- Grabowski, Joachim and George A. Miller. 2000. Factors affecting the use of dimensional prepositions in German and American English: Object orientation, social context, and prepositional pattern. *Journal of Psycholinguistic Research* 29/5: 517–553.
- Grabowski, Joachim and Petra Weiß. 1996. The prepositional inventory of languages: a factor that affects comprehension of spatial prepositions. *Language Sciences* 18: 19–35.
- Herskovits, Annette. 1986. *Language and spatial cognition: an interdisciplinary study of the prepositions in English*. Cambridge: Cambridge University Press.
- Hunston, Susan and Gill Francis. 2000. *Pattern grammar: a corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Kreitzer, Anatol. 1997. Multiple levels of schematization: a study in the conceptualization of space. *Cognitive Linguistics* 8/4: 291–325.
- Lakoff, George. 1987. *Women, fire, and dangerous things*. Chicago: The University of Chicago Press.
- Lindstromberg, Seth. 2010. *English prepositions explained*. Amsterdam: John Benjamins.
- Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, Mass.: The MIT Press.
- Tyler, Andrea and Vyvyan Evans. 2003. *The semantics of English prepositions: spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press.
- van der Zee, Emile, Karen Adams and Jussi Niemi. 2009. The influence of geometrical and non-geometrical features on the use of the lexical concepts NEAR and FAR in English and Finnish. *Spatial Cognition & Computation* 9/4: 305–317.
- Vorwerg, Constanze and Petra Weiß. 2010. Verb semantics affects the interpretation of spatial prepositions. *Spatial Cognition & Computation* 10/4: 247–291.