Antonymy and gradability: A corpus-based approach on English gradable antonyms*

Hye-Kyung Lee
(Ajou University)

Lee, Hye-Kyung. 2013. **Antonymy and gradability: A corpus-based approach on English gradable antonyms.** *Linguistic Research* 30(2), 335-354. This study examines the nature of antonymy relations of six gradable pairs (*old/young, high/low, thick/thin, wide/narrow, big/small* and *tall/short*) in terms of their frequencies and collocation in COCA (the Corpus of the Contemporary American English created by Mark Davies). The frequencies of the chosen twelve adjectives are counted and then the mean frequencies of the unmarked members and the marked members are calculated. The results reveal that the unmarked members of the pairs are used more frequently than their marked counterparts, which backs up the prediction that the unmarked adjectives are employed more frequently than their marked counterparts in general. The analysis of the collocation of the adjectives shows that the chosen pairs differ in terms of the extent to which they share common collocates. Four pairs (*old/young, high/low, thick/thin* and *wide/narrow*) have common collocates, which implicates that the antonyms can play in the same semantic fields. On the other hand, two pairs (*big/small* and *tall/short*) do not share any collocates. Hence in terms of the common collocates, some pairs are claimed to better qualify as antonyms than the others. The reason the latter pairs do not have common collocates is claimed to be the presence of other salient antonyms of one member. **(Ajou University)**

Keywords antonyms, gradability, collocation, phraseology, COCA

1. Introduction

The antonymy relation is one of the topics which have received extensive and intensive attention in the literature. It has also been proved that antonymy plays a pivotal role in organization of mental lexicon and discourse (e.g., Cruse, 1986; Jones, 2002; Lyons, 1977; Murphy, 2008). Accordingly, the research on antonymy varies

ranging from theoretical (e.g., Cruse, 1986; Lyons, 1977) to experimental approaches (e.g., Paradis & Willners, 2006). Among the various types of antonymy relations, contraries or gradable antonyms account for the vast majority of the antonymous pairs. Also it is widely known that most gradable antonyms show the marked and unmarked distinction.

This paper is an attempt to investigate the nature of the gradable antonyms from the perspective of corpus linguistics. The reason a corpus is used for this study is that corpora can provide authentic resources where the ordinary speakers' use and perception of gradable antonyms can be witnessed. By doing so, the paper tries to reveal any discrepancies between researchers' intuition on and ordinary speakers' use of the phenomenon under discussion.

Specifically, this study verifies the salience of the so-called unmarked member of gradable pairs by examining the frequencies of six representative gradable pairs in a well-established corpus, which is www.byu.edu/coca.[1] The six gradable pairs are *short/tall, narrow/wide, low/high, thin/thick, young/old,* and *small/big.* It also examines the extent to which these representative English gradable pairs are entitled to be antonyms by searching the words that frequently co-occur with the antonyms, i.e., collocates. The possible reasons for the differing extent are provided.

2. Background

2.1 Antonyms and gradability

Antonymy has been proved as crucial in the mental organization of words and discourse (e.g., Cruse, 1986; Jones, 2002; Murphy, 2008). The scope and pervasiveness of antonymy are much greater than other meaning relations such as synonymy and hyponymy (Jones, 2002). For instance, antonyms are acquired much earlier than other sense relations and furthermore acquired as pairs rather than separately (e.g., Kagan, 1984). Antonymy is also dubbed paradoxical because

---

[1] COCA stands for the Corpus of Contemporary American English created by Mark Davies of Brigham Young University. "The corpus contains more than 450 million words of text and is equally divided among spoken, fiction, popular magazines, newspapers, and academic texts. It includes 20 million words each year from 1990-2012 and the corpus is also updated regularly." (http://corpus.byu.edu/coca/)

antonymous words are both close and distant; antonyms share nearly all the same semantic features except for one feature.

In the semantics literature, antonyms have been classified according to their logical properties. Cruse (1986) and Lyons (1977), for instance, grouped antonymy relations into such major categories as contradiction, contrariety, and converseness.[2] Based on these categories, later semantic research (e.g., Jaszcsolt, 2002; Saeed, 2009) suggests various types of antonyms, such as complementary antonyms (contradictions), gradable antonyms (contraries), relational antonyms (converses), directional antonyms (reverses), and taxonomic sisters. Each type is illustrated in the examples in (1).

(1) a. dead/alive   (complementary)
    b. old/young   (gradable)
    c. employer/employee   (relational)
    d. up/down   (directional)
    e. red/blue   (taxonomic sisters)

Of main interest in the literature are still the first and the second categories. As for complementary antonyms, if one member is negated, the state described by other member holds. That is, they presuppose a definite 'either-or' mode (e.g., Paradis, 2001). On the other hand, the negation of one gradable member does not necessarily guarantee the truth of other member, opening the possibility of being other states than the one described by the antonyms. This is because the gradable antonyms are based on a range of scale.[3] Note that one member of a gradable pair is the unmarked expression employed in the so-called measure expression, such as (2), and in a *how*-question, such as (3).

(2) Mr. Jesperson is 75 years *old.*   (Quirk et al, 1985: 470)
(3) How *old* is your daughter?   (Quirk et al, 1985: 471)

---

[2] These scholars actually used the term antonym in a more restricted sense in such a way that antonyms only refer to gradable antonyms. The more general umbrella term opposites was employed.

[3] For these reasons, the first relation is called bounded and the second is called unbounded. (Paradis and Willners, 2006)

This classification of antonymy is mainly based on the logical properties of the lexical items as well as on the researchers' introspection. However, it is missing in these semantics-oriented studies how the language users actually perceive and employ the antonymy relations. In a bid to fill this gap, attempts have been made to investigate antonyms from the perspective of their functions in discourse. Jones (2002), for instance, proposed several functions of antonyms through English written corpus. The functions proposed by Jones (2002) are roughly related to specific English contrastive constructions. The two major categories are Ancillary and Coordinated antonyms, exemplified in (4) and (5) below.

> (4) I **love** *to cook* but I **hate** *doing the dishes*- so I'd have a dishwasher or
>   a family of gypsies to do the washing up.　(Jones 2002: 43)
> (5) The government will encourage everyone, **rich** and **poor**, to rely for
>   their retirement mainly on money ...　　　(Jones 2002: 61)

In (4), the pair *love* and *hate* have the function of generating the second antonymous pair, *to cook* and *doing the dishes*. That is why the pair, *love* and *hate,* are dubbed Ancillary antonyms. The example (5) contains the pair *rich* and *poor* coordinated by the conjunction *and*. Here the contrast between the two antonyms is neutralized, and thus the sentence signals inclusiveness of everyone in terms of the triggered scale of wealth. In addition to these two major categories, minor categories are also proposed such as Comparative, Distinguished, Transitional, Negated, Extreme, Idiomatic and Residual antonyms. Most categories are matched with specific English constructions. These categories are later attested in different modalities/registers in English (e.g., Jones, 2006; Murphy and Jones, 2008) and in a different language (i.e., Swedish) (Murphy et al, 2009). Noteworthy is that Ancillary antonyms and Coordinated antonyms comprise approximately 70 percent of the antonyms both in English and Swedish (Jones, 2002). Hence from the perspective of discourse, antonyms have a strong tendency to serve the function of signaling another contrast or to neutralize the contrast inherent in word meanings. These studies are meaningful in that the use of antonyms is investigated in the longer context rather than in isolation.

The categories suggested in the semantics literature are also tested against ordinary speakers in experimental settings. For instance, Paradis and Willners (2006) proved in their experimental study that language users do not always equate the

negation of one complementary antonym with the affirmation of its counterpart, contrary to the prediction of semantics-oriented studies. In Paradis and Willners' (2006) research, several complementary pairs can be classified differently in terms of the interpretation of their negations. For instance, in the case of the pair *dead* and *alive*, the logical property of complementarity holds, i.e., *dead* and *alive* are judged to be the same as *not alive* and *not dead*, respectively.[4] In contrast, in the case of the pair *wrong/right*, *wrong* was judged to the same as *not right*, while *right* was perceived to be different from *not wrong*. The other two pairs, *bound/free* and *empty/full*, also show different behavioral patterns from these two. Paradis and Willners (2006: 1075) successfully demonstrate that "alternative readings of lexical items are formed in context through construals that operate on their content and configuration structure", while ascribing the asymmetry to gradability (scaling potential in their term) of some complementary antonyms.[5] Nevertheless, as the authors acknowledge, their research uses a very limited number of bounded antonyms and thus more extended set is needed to draw a robust conclusion concerning the gradability or scalar potential of complementary antonyms. Presumably, as natural languages can function differently from logic, complementary antonyms have the potential to be used as gradable antonyms.

2.2 Collocation and phraseology

Collocation is one of the most pivotal topics in corpus linguistics. According to Biber et al (1998: 24), collocation refers to "the patterned way that words group

---

[4] In an experimental study about the complementary antonyms, *open* and *closed*, by Kaup et al (1986), a similar finding was reported, i.e., the subjects perceived a door that is not open to be the same as to be closed.

[5] Indeed, gradability of complementary antonyms was noted by theoretical semantists.

But with many terms, a proviso 'in all conceivable in any circumstances' seems necessary before the inferences which establish complementarity can be accepted at valid. This is perhaps true, for example, of dead and alive; could one not say of ghosts, or better still, vampires that they existed in a state which was neither death nor life? (Cruse, 1986: 200)

There are occasions when we will grade a pair of normally ungradable antonyms, because we do reject their interpretation as contradictories. 'Male' and 'female' are obvious example. ... we may well recognize that certain people cannot be satisfactorily classified in terms of this yes/no opposition of 'male' and 'female'. (Lyons, 1977: 278-279)

together". Collocation can provide a means to investigate the senses of words or the sense relations in an effective way, which conventional definitional theories of senses cannot supply. For instance, the senses of the seemingly synonymous words, *big, large* and *great* become differentiated by means of their collocates.

Table 1. Ten most common right collocates of *big, large* and *great* in Longman–Lancaster Corpus (Biber et al, 1998: 46)

Academic prose (2.7 million words)					
big		*large*		*great*	
right collocate	freq. per million	right collocate	freq. per million	right collocate	freq. per million
enough	2.2	number	48.3	deal	44.6
traders	1.1	numbers	31.3	importance	12.5
		and	29.4	number	8.9
		scale	28.0	majority	8.1
		enough	15.9	variety	7.0
		proportion	11.8	extent	7.0
		amount	10.7	part	4.1
		quantities	10.3	care	3.3
		part	10.0	advantage	2.6
		extent	8.9	detail	2.6
				interest	2.6

The collocational patterns in Table 1 reveal that the three synonymous words are employed distinctively. That is, *big* is used in association with physical size, while *large* is used to refer to amounts or quantities. On the other hand, *great* is used in a wider range of senses, even though its main sense is to express amounts.

Collocational patterns can be presented by means of phraseological units or lexical bundles. The uses of these two terms vary in the literature; they are used synonymously (e.g., Cheng & Leung, 2012) or differently (Ädel & Erman, 2012). According to Ädel & Erman (2012), phraseological units coincide with grammatical units such as noun phrases or verb phrases, whereas lexical bundles refer to multi-word sequences that recur frequently across different texts (Biber, 2010). In the synonymous context, on the other hand, they both refer to "the recurrent contiguous words that constitute a phrase or a pattern of use in texts or a corpus" (Cheng & Leung, 2012: 618). Cheng & Leung (2012), for instance, examine the phraseological

variations of five two-word lexical bundles, which are *fair/value, management/risk(s), interest/rate(s), energy/saving(*s)*, and *energy/use* in HKEC and HKFSC.[6] They claim that the investigation of the phraseologies contributes to the better understanding of language use and meaning. In any case, phraseological units or lexical bundles have been proved to be very crucial in language learning and hence drawn intensive research attention both in corpus linguistics and language teaching (e.g., Biber, et al, 2004; Cheng and Leung, 2012; Hyland, 2008; Lieu, 2012 and references therein). In this study, the two terms, phraseological units and lexical bundles, will not be differentiated, mainly because most of lexical bundles found in this study coincide with phraseological units.

3. The study

3.1 Aims of the study

This study is an attempt to investigate the nature of the antonymy relations of several representative gradable adjectives. For that investigation, the study has two main aims. The first one is to verify the unmarkedness of the so-called unmarked members of gradable pairs by showing their prominent occurrences in the corpus. It is because the unmarked members are predicted to be used more frequently than their marked counterparts. At the same time, by observing the collocational patterns, it examines the extent to which the gradable antonyms actually serve as antonyms.

3.2 Materials and procedures

In order to test the aims of this paper, an investigation of a well-established corpus was conducted, which is www.byu.edu/coca. The use of a group of gradable antonyms was examined in the corpus. For the analysis, six pairs of gradable antonyms were chosen based on the data used in Paradis and Willners (2006) and Quirk *et al* (1985). Paradis and Willners used the Swedish gradable adjective pairs presented in Table 2.[7]

---

6 HKEC and HKFSC are abbreviations of the Hong Kong Engineering Corpus and the Hong Kong Financial Services Corpus respectively.

Table 2. Gradable adjectives and their antonyms
in Paradis and Willners' (2006)

adjectives having little of the designated property (X)	adjectives having much of the designated property (Y)
short	tall
light	dark
narrow	wide
low	high
light	heavy
thin	thick
ugly	beautiful
cold	hot
poor	rich
sad	happy

Among the ten pairs, the pairs which are in the marked/unmarked distinction were chosen for the analysis: *short/tall, narrow/wide, low/high* and *thin/thick*. In addition to the four adjectival pairs, two more representative gradable adjective antonyms were chosen from the ones suggested by Quirk et al (1985). They are *young* and *old*, and *small* and *big*. The pairs used in the current study are presented in Table 3.

Table 3. Gradable adjectives and their antonyms in the current study

adjectives having little of the designated property (marked: X)	adjectives having much of the designated property (unmarked: Y)
short	tall
narrow	wide
low	high
thin	thick
young	old
small	big

Then the raw occurrences of these adjectives were counted and the means of X's and Y's occurrences were calculated. Next, the words collocated with the adjectives were extracted from the corpus site. Specifically, the collocates are extracted in order of frequency with a span of four words right and left.[8] Ten most frequent collocates

---

7 Paradis and Willners' (2006) research was conducted on Swedish adjectives. Their English translations are presented in Table 2.

of the relevant expressions were dealt with in the current study.

3.3 Results of analysis

The occurrences of the adjectives are shown in Table 4. The mean occurrences of *Y* and *X* members are 132,588 and 88,960 respectively.

Table 4. Occurrences of gradable pairs *Y* and *X*[9]

Y	Counts	Mean	*X*	Counts	Mean
big	288,075		small	181,035	
high	209,747		low	74,248	
old	207,448	132,588	young	160,898	88,960
wide	42,179		narrow	19,313	
tall	24,082		short	73,017	
thick	23,999		thin	25,250	

The mean occurrence of *Y* outnumbers that of *X*. As for *Y* members, the frequency of *thick* is the lowest (N=23,999), whereas that of *big* is the highest (N=288,075). As for X members, the adjective *narrow* occurs least frequently (N=19,313), whereas the adjective *small* appears most frequently (N=181,035). Out of six sets, two sets show deviations from the distributional pattern presented in Table 4: *short/tall* and *thin/thick*. The counts of *Y* and *X* are reversed.

The ten most frequent collocates of each adjective is presented in the Appendix.

4. Discussion

4.1 Frequency of gradable pairs

Overall the occurrence of the antonym having much of the property is greater than that of its *X* counterpart, as shown in Table 4. Hence the prediction that the unmarked antonyms are used much more frequently than their marked counterparts is borne out. The exceptions to this tendency are the two pairs, *short/tall* and *thin/thick,*

---

[8]   The extracted collates have a Mutual Information score of 3.0 or higher.

[9]   The pairs are listed in the decreasing order of the frequencies of *Y*s.

i.e., the marked members occur more frequently than their unmarked counterparts. The deviance of the first pair, *short* and *tall*, can be attributed to the fact that the adjective *short* can be used as an antonym of another common adjective, *long*. The occurrence of the adjective *long* in the corpus is 277,641. If we add the occurrences of *tall* and *long*, it outnumbers the occurrence of *short*. As for the second pair, even though the occurrence of *thin* is greater than that of *thick*, the difference between them is pretty small. Also, as with the *short/tall* pair, the adjective *thin* is a candidate for an antonym of the adjectives *fat* or *chubby*. That must have affected the greater frequency of *thin* than that of *thick*.

4.2 Collocates of gradable pairs

The function of gradable pairs is examined with reference to their collocates. The main function of adjectives is to modify or predicate an entity. Hence, it is predicted that adjectives mostly co-occur with nouns. The extracted collocates of the adjectives are listed in Table 5.

Table 5. Top ten collocates of *Y*s and *X*s

Y	Collocates of Y only	Common collocates	Collocates of X only	X
big	deal, difference, picture, ten, bang, brother, screen, mistake, fan, bucks		town(s), business, bowl, relatively amount(s), size, percentage, pieces	*small*
old	years, friend, testament, plain, dominion, twelve, fashioned	man/men, lady/ladies	woman, children, girl(s), adults, boy	*young*
high	school(s), **low**, junior, pressure, blood	level(s), rate(s), heat	**high**, vision, voice, **relatively**, prices,	*low*
thick	black, inch(es), smoke, glasses, dark	air, layer, hair, slices,	tall, line, cut, lips, strips, ice	*thin*

tall	feet, man, inches, thin, woman, six, grass, trees, stood, building		term, hair, story/stories, cut, period, break, stopped, distance, fall	*short*
wide	variety, open, receiver, feet, web, inches, mouth, array	eyes, range	street(s), road, focus, path, gap, straight, strip	*narrow*

Among the six pairs, four pairs share common collocates, as shown in Table 5: *old/young, high/how, thick/thin* and *wide/narrow* pairs. The two pairs do not share any collocates at all: *big/small* and *tall/short* pairs. Hence as far as the common collocates are concerned, the first four pairs more qualify as antonyms than the latter two do. Let us first examine the pairs which have common collocates.

The pair *old* and *young* shares two lemmas, *man* and *lady*, which reveals that the two adjectives are employed to describe persons' age. The adjective *old* co-occurring with *years* is used to refer to a person's age, forming a lexical bundle with a number, *number years old*. Among the numbers of age, *twelve* occurs most frequently with the phrase *years old*. Except for the number twelve, other numbers are also included in the top 100 recurrent collocates of *old*. Though why these age groups are more prominently mentioned should be dealt with in another study, the frequent use of *twelve* can be inferred from some of the data. It is probably because at the age of around twelve, a child starts a secondary education and thus is regarded as turning into an adolescent, as shown in concordance lines in (6).

(6) a. "By the time I was **twelve** years **old**, I had gone to thirteen public schools," she says.
   b. I should have known. When I was **twelve** years **old**, I entered the 7th-grade science fair.

The other collocates, *plain, fashioned, testament* and *dominion*, are due to semi-idiomatic lexical bundles or compounds, *plain old, old fashioned, Old Testament* and *Old Dominion*. On the other hand, the collocates of *young* show that the adjective *young* is used to describe the youngness of a person.

The pair, *high* and *low*, share the words *level(s), rate(s),* and *heat*. The word *low*

is used to refer to lowness of such qualities as voice, prices and vision. On the other hand, the two collocates of *high*, *school* and *junior*, are used to refer to educational institutions or school grades, as shown in concordance lines (7).

(7) a. total number of teachers (not specified as to whether they were middle school or **high school)**
    b. all adopted from China and Sophie is the eldest and she's a **junior** in **high** school.
    c. if you want to call it a joke, since I was in junior high school. It hasn't improved with age.

The remaining two collocates of *high*, *blood* and *pressure*, come from the compound, *high blood pressure*. Noteworthy is that among the examined six pairs, only these two adjectives have their antonyms in the top ten collocates as highlighted in Table 5. This fact accords with Stubbs' (2006: 23) observation that "semantically contrasting words co-occur in text with a frequency much higher than chance."[10]

The pair, *thick* and *thin*, share the collocates, *air, layer, hair,* and *slices*, which mostly refer to physical entities. The word *thick* is found to co-occur with nouns such as *glasses* and *smoke*. It also tends to co-occur with the measurement expression *inch(es)*. Interesting is that *thick* co-occurs with other adjectives, *black* and *dark*. *Thick* in these cases is used either to modify (as in (8a) and (8b)) or to be juxtaposed with the following adjective (as in (8c) and (8d)).

(8) a. Rita had **thick dark** red hair which glinted in the sunlight that streamed through her big picture.
    b. The second man wore **thick dark** blond hair pulled straight back in a band.
    c. Stocky and bowlegged, with cropped **black** hair and a **thick** mustache
    d. The honey was **thick** and **dark**, with a strong country taste.

---

10 The other pairs also have their corresponding counterparts as recurring collocates, albeit not in the top ten. For instance, the adjective *wide* has *narrow* as its 34th frequent collocate and *thin* has *thick* as its 20th frequent collocate. The two pairs, *big/small* and *old/young* do not have their counterparts in the top 100 frequent collocates.

On the other hand, *thin* is associated with concrete subjects such as *line, lips* and *strips*. *Thin* also makes a lexical bundle with the verb *cut* as in (9), especially in cooking recipes.

(9) a. Slice tenderloin crosswise into 1/4-inch-thick pieces, then **cut** each into **thin** strips.

b. Arrange onions and tomatoes in alternating slices on 4 serving plates. **Cut** steak into **thin** slices.

That is, *cut* + something + *into* + *thin* + *slices/strips* appears to form a lexical bundle. Noteworthy is that *thin* is strongly associated with another adjective *tall*. Presumably, in describing persons' appearances, people tend to associate thinness with tallness.

The pair, *narrow* and *wide*, share the two words, *eyes* and *range*. The first collocate *eyes* pertains to a physical/concrete entity, while the second collocate *range* refers to an abstract notion. Except for the common words, the collocates of *narrow* show that *narrow* is commonly used to refer to physical sizes. When *narrow* co-occurs with another adjective *straight*, they are juxtaposed to describe an object such as *a path, a road* or *a street*, which is also one of the most frequent collocates of *narrow*. On the other hand, the adjective *wide* is used to describe a variety of notions, from the range or measurement (e.g., *variety, array,* and *feet*) to physical size (e.g., *feet, mouth* and *open*). The concordance lines of *wide* occurring with *feet* in (10) show the different meanings of *feet*.

(10) a. brothers, perhaps in their late 20s, both slight, muscled, with **wide feet**.

b. of wagon box or platform made of four or five planks, about three **feet wide** and four **feet** long.

The two collocates *receiver* and *web* are due to the compounds *wide receiver* and *World Wide Web* respectively.

The two pairs, *big/small* and *tall/short*, do not have any common collocates. Let me first turn to the *big/small* pair. The examination of the top ten collocates of *big* reveals that *big* is collocated with another word to form compounds (such as *Big*

Bang and *Big Ten*) or lexical bundles (such as *big difference* and *big mistake*).[11] The use of *big* in the latter cases must have to do with the intensifying function of *big*. As noted by Quirk et al (1985: 430), the adjective *big* can be employed as an intensifying adjective as in the example *a big baby* (very babyish) or *a big fool* (very foolish). The examination of the top 100 collocates of *big* reveals that most lexical bundles contain the intensifying *big*. This intensifying function of *big* must contribute to the high frequency of *big* and the semi-fossilized expressions, which again differentiated the collocates of *big* from those of *small*.

Neither do the two adjectives *big* and *small* share common collocates, nor does the adjective *big* have any of the top ten collocates of *small* in its top 100 collocates. This leads us to ask the question about the antonymy relationship between the two adjectives. Instead, the adjective *large* shares two collocates with *small*, which are *bowl* and *amount*. Hence, *large* can better qualify as an antonym of *small* as far as collocation is concerned. On the other hand, the adjective *small* is used to describe the size of objects such as a company, a business, a bowl or pieces. It is also used to refer to an amount or a portion of something. Two of the ten collocates clearly carry this meaning: *amount(s)* and *percentage*, as shown in (11).

(11) a. But doctor after doctor we talked to said there is such a **small amount** of mercury.
      b. It's a very **small percentage** of the nearly 5,000 visas we give a year.

Noticeable is that the adverb *relatively* is collocated with *small*. The same adverb also appears in the top 10 collocates list of *low*.

The pair *tall/short* does not share common collocates either. Even the top hundred collocates of *tall* does not contain any of top ten collocates of *short*. The adjective *tall* is used to describe the physical size of such entities as a person, grass, trees and buildings. Also it forms a lexical bundle with the measure expression *feet*, which is normally accompanied by a specific number, as shown in (12).

(12) a. By the time I was seventeen, I was two inches over six **feet tall**

---

11 All the ten collocations are actually registered as lexical entries in an online dictionary.

and was beginning to look freakish.

b. Was he a man - he was five **feet** eleven inches **tall** - or was he a boy?

Interestingly, number *six* is most frequently collocated with the lexical bundle *feet tall*. Why *six feet tall* is more prominent than the others should also be examined in terms of cultural aspects concerning persons' height in the language community. The adjective *short* is used usually refer to either concrete objects (*hair* or *story*) or a certain period of time or space (*period, term, break* or *distance*). It is also employed mainly in (semi)-idiomatic expressions such as *cut something short*, *stop short (of)* or *fall short (of)*, as shown in the concordance lines (13).

(13)   a. Poverty **cut short** her education. After graduating from high school, she worked for four years.

b. The pen, already poised above the document, **stopped short** of the paper.

c. But the supply is expected to **fall short** of the demand ...

As mentioned above, the top ten collocates of both adjectives do not overlap at all. Instead, the other adjective *long* shares three collocates with *short*, which are *hair*, *term* and *period*. Hence, the adjective *long* can be better regarded as an antonym of *short* in terms of collocation.

Among the pairs examined in the current study, two pairs do not share collocates: the *big/small* pair and the *tall/short* pair. Instead, the other *Y* adjectives, *large* and *long*, can better qualify as the antonyms of the respective *X* adjectives, as far as collocation is concerned. Except for these two pairs, the other pairs have overlapping collocates, which means the antonyms are used to modify or predicate the same types of entities. At the same time, the antonyms are found to be employed in different semantic fields. These two pairs, *big/small* and *tall/short*, also belong to the category in which the *Y* member is not included in the top hundred collocates of the *X* member (cf. Stubbs, 2006). The three pairs *high/low*, *thick/thin,* and *wide/narrow* have their respective antonyms within their top 100 collocates, as discussed in footnote 10 above. These two facts taken into account, it could be concluded that the last three pairs, *high/low*, *thick/thin,* and *wide/narrow*, can better

qualify as antonyms than the remaining ones. In the case of the other three pairs, one member of the pair has another prominent alternative counterpart. That must account for the less overlapping between the two members. For instance, the *X* adjectives *small* and *short* have other salient antonyms, *large* and *long*. In the case of the *old/young* pair, the *Y* adjective *old* has another salient antonym *new*.

5. Conclusions

This study examined the nature of antonymy relations of six gradable pairs in terms of their frequencies and collocation in COCA. For the analysis, the frequencies of twelve adjectives were counted. At the same time, the mean frequencies of the unmarked members and the marked members were counted. The counted numbers revealed that the unmarked members of the pairs were used more frequently than their marked counterparts, which supports the prediction that the unmarked adjectives are employed more frequently than their marked counterparts in general.

The analysis of the collocation of the adjectives showed that the chosen pairs differ in terms of the extent to which they share common collocates. Four pairs have common collocates, which implicates that the antonyms can play in the same semantic fields. The pairs *old/young, high/low, thick/thin* and *wide/narrow* belong to this category. On the other hand, two pairs, *big/small* and *tall/short,* do not share any collocates. Hence in terms of the common collocates, the first four pairs were claimed to better qualify as antonyms than the latter two. As discussed in section 4, the reason the latter two pairs do not have common collocates was claimed to be the presence of other salient antonyms. In the case of the *big/small* pair, the adjectives such as *large* and *great* can also be regarded as antonyms of *small*. Likewise, in the case of the *tall//short* pair, *short* can be used as an antonym of another common adjective *long*.

References

Ädel, Annelie, and Britt Erman. 2012. Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for*

Specific Purposes 31: 81-92.

Biber, Douglas. 2010. Corpus-based and corpus-driven analyses of language variation and use. In B. Heine, and B. Narrog (eds.), *The Oxford Handbook of Linguistic Analysis*, 159-191. Oxford: Oxford University Press.

Biber, Douglas, Susan Conrad, and Randi Reppen. 1998. *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Biber, Douglas, Susan Conrad, and Viviana Cortes. 2004. *If you look at. . .*: Lexical bundles in university teaching and textbook. *Applied Linguistics* 25: 371-405.

Cheng, Winnie, and Maggie Leung. 2012. Exploring phraseological variations by concgramming. *Linguistic Research* 29(3): 617-638.

Cruse, Alan. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.

Hyland, Ken. 2008. As can be seen: Lexical bundles and disciplinary variation. *English for Specific Purposes* 27(1): 4-21.

Jaszcsolt, Katarzina. M. 2002. *Semantics and Pragmatics: Meaning in Language and Discourse*. London: Longman.

Jones, Steven. 2002. *Antonymy: a Corpus-based Perspective*. London: Routledge.

Jones, Steven. 2006. The discourse function of antonymy in Spoken English. *Text and Talk* 26: 191-216.

Kagan, Jerome. 1984. *The Nature of the Child*. New York: Basic Books.

Kaup, Barbara, Jana Ludtke, and Rolf Zwaan. 2006. Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics* 38: 1033-1050.

Liu, Dilin. 2012. The most frequently-used multi-word constructions in academic written English: A multi-corpus study. *English for Specific Purposes* 31: 25-35.

Lyons, John. 1977. *Semantics*. Cambridge: Cambridge University Press.

McEnery, Tony, and Andrew Wilson. 2001. *Corpus Linguistics*. 2nd edition. Edinburgh: Edinburgh University Press.

Murphy, Lynne. 2008. *Semantic Relations and the Lexicon: Antonyms, Synonyms and other Semantic Paradigms*. Cambridge: Cambridge University Press.

Murphy, Lynne, and Jones Steven. 2008. Antonymy in childhood. *First Language* 28(4): 403-430.

Murphy, Lynne, Carita Paradis, Carikube Willners, and Steven Jones. 2009. Discourse function of antonymy: A cross-linguistic investigation of Swedish and English. *Journal of Pragmatics* 41: 2159-2184.

Paradis, Carita. 2001. Adjectives and boundedness. *Cognitive Linguistics* 12(1): 47-65.

Paradis, Carita, and Caroline Willners. 2006. Antonymy and negation- The boundedness hypothesis. *Journal of Pragmatics* 38: 1051-1080.

Saeed, John. 2008. *Semantics*. 3rd edition. Chichester: Wiley-Blackwell.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A*

Comprehensive Grammar of the English Language. London: Longman.

Sinclair, John. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Stubbs, Michael. 2006. Corpus analysis: the state of the art and three types of unanswered questions. In G. Thomson, and S. Hunston (eds.), *System and Corpus: Exploring Connections*, 15-36. London: Equinox.

Appendix

Figure 1: Collocates of *short*

		CONTEXT	FREQ	ALL	%	MI
1	☐	TERM	2216	38634	5.74	5.51
2	☐	HAIR	1729	79398	2.18	4.11
3	☐	STORY	1720	138611	1.24	3.30
4	☐	STORIES	1605	52855	3.04	4.59
5	☐	CUT	1411	103499	1.36	3.44
6	☐	PERIOD	1394	61527	2.27	4.17
7	☐	BREAK	1039	73063	1.42	3.50
8	☐	STOPPED	861	46584	1.85	3.88
9	☐	DISTANCE	852	32423	2.63	4.38
10	☐	FALL	809	67794	1.19	3.25

Figure 2: Collocates of *tall*

		CONTEXT	FREQ	ALL	%	MI
1	☐	FEET	2615	92315	2.83	6.09
2	☐	MAN	1738	306672	0.56	3.76
3	☐	INCHES	762	25395	3.00	6.18
4	☐	THIN	626	25396	2.46	5.89
5	☐	WOMAN	608	164165	0.37	3.16
6	☐	SIX	606	104504	0.58	3.80
7	☐	GRASS	574	18475	3.11	6.23
8	☐	TREES	557	40849	1.36	5.04
9	☐	STOOD	388	52966	0.73	4.14
10	☐	BUILDINGS	371	23418	1.58	5.25

Figure 3: Collocates of *narrow*

		CONTEXT	FREQ	ALL	%	MI
1	☐	STREETS	630	30626	2.06	5.95
2	☐	ROAD	424	79214	0.54	4.01
3	☐	EYES	404	147008	0.27	3.05
4	☐	STREET	314	104013	0.30	3.18
5	☐	RANGE	279	50822	0.55	4.04
6	☐	FOCUS	271	57821	0.47	3.82
7	☐	PATH	266	27029	0.98	4.89
8	☐	GAP	226	15018	1.50	5.50
9	☐	STRAIGHT	205	41438	0.49	3.89
10	☐	STRIP	196	10938	1.79	5.75

Figure 4: Collocates of *wide*

		CONTEXT	FREQ	ALL	%	MI
1	☐	RANGE	4856	50822	9.55	7.04
2	☐	EYES	2946	147008	2.00	4.79
3	☐	VARIETY	2808	32717	8.58	6.88
4	☐	OPEN	2416	131266	1.84	4.66
5	☐	RECEIVER	1791	7701	23.26	8.32
6	☐	FEET	1495	92315	1.62	4.48
7	☐	WEB	1352	38619	3.50	5.59
8	☐	INCHES	805	25395	3.17	5.45
9	☐	MOUTH	656	42448	1.55	4.41
10	☐	ARRAY	618	8553	7.23	6.64

Figure 5: Collocates of *low*

		CONTEXT	FREQ	ALL	%	MI
1	☐	HIGH	3955	210567	1.88	3.88
2	☐	RATES	2032	43560	4.66	5.19
3	☐	HEAT	1964	52462	3.74	4.87
4	☐	LEVELS	1964	53309	3.68	4.85
5	☐	VISION	1791	32925	5.44	5.41
6	☐	VOICE	1556	92631	1.68	3.72
7	☐	RELATIVELY	1513	29820	5.07	5.31
8	☐	PRICES	1410	39611	3.56	4.80
9	☐	LOW	1309	75051	1.74	3.77
10	☐	RATE	1120	58839	1.90	3.90

Figure 6: Collocates of *high*

		CONTEXT	FREQ	ALL	%	MI
1	☐	SCHOOL	52942	278048	19.04	5.72
2	☐	LEVELS	4443	53309	8.33	4.53
3	☐	SCHOOLS	4199	78414	5.35	3.89
4	☐	LOW	3816	75051	5.08	3.81
5	☐	LEVEL	3498	96310	3.63	3.33
6	☐	JUNIOR	3330	17053	19.53	5.76
7	☐	RATES	3001	43560	6.89	4.25
8	☐	PRESSURE	2639	50329	5.24	3.86
9	☐	HEAT	2294	52462	4.37	3.60
10	☐	BLOOD	2243	64837	3.46	3.26

Figure 7: Collocates of *thin*

		CONTEXT	FREQ	ALL	%	MI
1	☐	AIR	961	120296	0.80	4.19
2	☐	LAYER	664	11193	5.93	7.08
3	☐	TALL	637	24084	2.64	5.92
4	☐	LINE	616	118127	0.52	3.58
5	☐	SLICES	506	7596	6.66	7.25
6	☐	CUT	497	103499	0.48	3.46
7	☐	HAIR	494	79398	0.62	3.83
8	☐	LIPS	410	21104	1.94	5.47
9	☐	STRIPS	368	5270	6.98	7.32
10	☐	ICE	339	37048	0.92	4.39

Figure 8: Collocates of *thick*

		CONTEXT	FREQ	ALL	%	MI
1	☐	HAIR	1306	79398	1.64	5.31
2	☐	BLACK	801	177907	0.45	3.44
3	☐	INCH	556	10639	5.23	6.97
4	☐	AIR	496	120296	0.41	3.31
5	☐	INCHES	354	25395	1.39	5.07
6	☐	SMOKE	353	22546	1.57	5.24
7	☐	GLASSES	352	13146	2.68	6.01
8	☐	DARK	347	68270	0.51	3.61
9	☐	LAYER	321	11193	2.87	6.11
10	☐	SLICES	288	7596	3.79	6.51

Figure 9: Collocates of *small*

		CONTEXT	FREQ	ALL	%	MI
1	☐	TOWN	4686	81713	5.73	4.20
2	☐	BUSINESSES	3138	23585	13.31	5.41
3	☐	RELATIVELY	2658	29820	8.91	4.84
4	☐	BOWL	2096	33544	6.25	4.32
5	☐	AMOUNT	1884	46715	4.03	3.69
6	☐	SIZE	1810	51689	3.50	3.49
7	☐	TOWNS	1547	11550	13.39	5.42
8	☐	PERCENTAGE	1116	20672	5.40	4.11
9	☐	AMOUNTS	1028	12442	8.26	4.73
10	☐	PIECES	990	32356	3.06	3.29

Figure 10: Collocates of *big*

		CONTEXT	FREQ	ALL	%	MI
1	☐	DEAL	5580	87850	6.35	4.15
2	☐	DIFFERENCE	2300	52713	4.36	3.61
3	☐	PICTURE	1812	52081	3.48	3.28
4	☐	TEN	1627	46743	3.48	3.28
5	☐	BANG	1584	5163	30.68	6.42
6	☐	BROTHER	1524	49371	3.09	3.11
7	☐	SCREEN	1085	31436	3.45	3.27
8	☐	MISTAKE	1053	20028	5.26	3.87
9	☐	FAN	1052	15689	6.71	4.23
10	☐	BUCKS	988	8195	12.06	5.07

Figure 11: Collocates of *young*

		CONTEXT	FREQ	ALL	%	MI
1	☐	MAN	14334	306672	4.67	4.08
2	☐	MEN	8355	177900	4.70	4.08
3	☐	WOMAN	8323	164165	5.07	4.19
4	☐	CHILDREN	6400	253844	2.52	3.19
5	☐	WOMEN	6337	237921	2.66	3.26
6	☐	GIRL	2926	74847	3.91	3.82
7	☐	ADULTS	2495	26765	9.32	5.07
8	☐	GIRLS	2076	54635	3.80	3.78
9	☐	LADY	1831	31143	5.88	4.41
10	☐	BOY	1698	73233	2.32	3.06

Figure 12: Collocates of *old*

		CONTEXT	FREQ	ALL	%	MI
1	☐	YEARS	25116	528247	4.75	3.73
2	☐	MAN	13393	306672	4.37	3.61
3	☐	FRIEND	2832	72754	3.89	3.45
4	☐	LADY	1964	31143	6.31	4.14
5	☐	TESTAMENT	1265	4299	29.43	6.36
6	☐	LADIES	563	13728	4.10	3.52
7	☐	PLAIN	498	14077	3.54	3.31
8	☐	DOMINION	425	1448	29.35	6.36
9	☐	TWELVE	381	11142	3.42	3.26
10	☐	FASHIONED	342	2226	15.36	5.43

Hye-Kyung Lee
Department of English
Ajou University
206 World cup-ro, Yeongtong-gu, Suwon-si
Gyeonggi-do 443-749, Korea
E-mail: hklee@ajou.ac.kr