Non-concord in Existential-*There* Constructions: A Corpus-Based Study

Seo Jeong Park

(Seoul National University)

Park, Seo Jeong. 2014. Non-concord in Existential-*There* Constructions: A Corpus-Based Study. *SNU Working Papers in English Linguistics and Language 12, 111-137.* Non-concord between the notional subject and the verb is an oft-cited characteristic of existential there-constructions. Recognizing some limitations of the previous studies (Meechan & Foley, 1994; Insua & Martinez, 2003), this paper seeks to resolve them by leading a more focused corpus-based research with in-depth analyses of the factors involved: the coordination of NPs, the characteristics of the quantifiers positioned between the verb and the NP, and the medium of expression (i.e. speech vs. writing). The findings confirm the prevalence of non-concord in speech, to a degree that non-concord in existential-*there* is deemed acceptable in speech. It is also suggested that it is not merely the length of the NP but the number information conveyed by the element immediately following the verb that affects one's memory to produce non-concordant sentences. **(Seoul National University)**

Keywords: existential-*there*, non-concord, corpus, memory, speech

1. Introduction
1.1 Existential-*there*

In present-day English there are two kinds of *there*: existential and locative[1] (Breivik, 1981). This study concerns the former, existential *there*-constructions (hereafter 'TC's). An important aspect of existential *there*-sentences is that he subject in the original basic clause serves as the notional subject of the *there*-sentence (Quirk et al., 1985, p. 1403). Consider the following examples for the relationship between the basic clauses and the *there*-constructions.

---

[1] also coined adverbial by others (Quirk et al., 1985, p. 944)

(1)   a. Something must be wrong.
       b. There must be something wrong. (Quirk et al., 1985, p.
       1404)
(2)   a. Plenty of people are getting promotion.
       b. There are plenty of people getting promotion. (Quirk et
       al., 1985, p. 1404).

1.1.1 Concord in existential-*there*

Prescriptive grammar dictates that the number of the verb following existential *there* must agree with that of the notional subject (Cowan, 2008, p.134). This kind of agreement is considered as one type of *notional concord*[2] (Biber et al., 1999, p. 757), as it is not between the grammatical subject (*there*) and the verb as in the usual agreement but between the notional subject and the verb. Therefore, using the plural form of *be* with a plural subject as in *There are five students in the playground* is grammatically correct.

Nevertheless, as pointed out in numerous studies (Biber et al., 1999, p. 944; Quirk et al., 1985, p. 1405; Meechan & Foley, 1994; Insua & Martinez, 2003), a singular form of *be* is frequently followed by plural NPs. In the majority of such instances of non-concordant *there*-constructions[3], the verb is in present tense, and is contacted as *'s* and attached to *there* (Meechan & Foley, 1994; Biber et al., 1999, p. 186). According to prescriptive grammar, it is an oft-cited instance of incorrect usage, as suggested in Cowan (2008): "A commonly heard violation of the agreement rule is the use of *there's* before an NP that has a plural count noun as its head" (p. 134). Biber et al. (1999) explain

---

[2] Another instance of notional concord is when a singular verb is chosen with a plural noun viewed as a single entity, as in "*King prawns cooked in chili salt and pepper was very much better, a simple dish succulently executed*" (Biber et al., 1999, p. 187). In this example, "king prawns" are considered as one dish, rather than individual prawns.
[3] Henceforth termed *NC*s, as opposed to concordant *there*-constructions, abbreviated as *C*s. 'TC' stands for all *there*-constructions.

that the widely attested phenomenon of non-concord between present-tense be and a plural NP is simply due to "the special behaviour of *there's*: because of the contraction, *there's* tends to behave as a single invariable unit for the purposes of speech processing" (p. 186). Other scholars have attempted to offer different explanations, such as the syntax of existential *there*, social classes of speakers, information structure regarding the length of NPs, and so forth to be discussed in Section 2.

2. Literature Review
2.1 Meechan & Foley (1994)

Meechan & Foley (1994) present a variationist analysis of non-concord in TCs and explore a wide range of factors for the attested phenomenon: syntax, verb tense, determiners preceding the NPs, adjacency of the NP and the verb, and speakers' level of education and geographical background.

2.1.1 Data selection

Meechan & Foley (1994)'s study is based on their analysis of the data from two sources. The first is the *African Nova Scotian English Control Group Corpus* of English speakers residing in Nova Scotia who were over the age of 60 (Meechan & Foley, 1994). The second, the main data for analysis, is "the *Linguistics Department Archives of Spoken Language Materials* at the University of Ottawa, which features recordings of the speech of 31 residents of Ottawa, Ontario" who, all but one, were "between the ages of 55 and 95" (Meechan & Foley, 1994).

From these corpora, Meechan & Foley (1994) analyze "[e]xistential constructions with present tense markings".

2.1.2 Results and conclusion

Meechan & Foley (1994) view non-concord for plural nouns in spoken TCs (i.e. *'s* or *is*) as the default agreement, as its proportion accounts for 72% of the whole. The factors behind the prevalence of such non-concord are further investigated. The two tables below summarize the central linguistic and social factors accounting for the phenomenon.

Table 1. "Percentage of concord with plural nouns for linguistic
        factors" in Meechan & Foley (1994) (Meechan & Foley,
        1994, Table 5)

Factor	Plural Nouns % Concord	Distribution	
		n	% of Total
Copula form			
Contracted	8	119	51
Full present	100	16	7
Past	42	100	43
Plural mark			
No mark	21	48	24
-s	32	151	76
Determiner			
Weak	27	182	84
Strong	37	35	16
Other intervening material			
Present	28	72	31
Absent	29	163	69
Small clause			
Present	27	66	28
Absent	29	169	72
Specificity			
Nonspecific	27	164	86
Specific	26	27	14

Table 2. Social factors in existential concord: Percentage of plural concord according to social factors" in Meechan & Foley (1994) (Meechan & Foley, 1994, Table 9)

Factor	% Plural	*n*	% of Total
Sex			
Female	24	105	45
Male	32	130	55
Location			
Rural	18	106	45
Urban	37	129	55
Education			
< 11 years	16	68	35
> 11 years	35	129	65
Topic			
School, church	26	58	22
Other	24	200	78

Meechan & Foley (1994)'s observations of these factors lead to two ultimate conclusions. First, contrary to what minimalist syntacticians claim, LF-raising of the postverbal NP must be optional instead of obligatory in TCs. Second, concord in TCs is learned through formal education and thus urban and more educated speakers tend to avoid non-concord.

Although the research is noted for its bridging between syntax and variation and between society and language, there lies a limitation: the samples are too limited, the total being 670 tokens of TCs and with no subgroup number in Table 1 and 2 exceeding 200. Also, only spoken data were analyzed, with 31 speakers of a limited range in age, hampering the reliability of the conclusion.

2.2 Insua & Martinez (2003)

Insua & Martinez (2003) lead a corpus study on non-concord in TCs in which 2410 samples of TCs from the British National Corpus (BNC)

were analyzed.

2.2.1 Data selection

Insua & Martinez (2003) acquired their selective data of TCs through two steps of sampling processes. The following table illustrates the outcome of each procedure. The values under the column *Total of TCs* signify the data from the first stage of selection, and those under *TCs here analyzed* are the number of the TCs resulting from the second.

Table 1. "Total number and distribution of the TCs found in the corpus under analysis TC samples" in Insua & Martinez (2003) (Insua & Martinez, 2003, Table 1)

	Total of TCs	*TCs here analysed*	
	No.	No.	%
Written English	1071	963	89.91
Spoken English	1619	1447	89.37
Total	2690	2410	89.59

During the course of the first stage, a small subpart of TCs (2690 in total, as seen in Table 3) dated from the years 1989 onwards was picked out of the whole corpus. The sub-corpus is characterized by an equal proportion of each genre (e.g. *business*) of the tokens and also equal proportions of words for both speech (half a million) and writing (also, half a million) (Insua & Martinez, 2003). In this process, a *disproportionally stratified sample* (Butler, 1985, p. 6) was employed, which is a sample that "creates the optimum situation for the comparison of subgroups" and "avoids the possibility that those types of unit with a small overall proportion in the population…are not represented in the sample at all" (Insua & Martinez, 2003).

When the set of TCs were collected, they underwent a second procedure in order for the samples to be "identified, counted and entered in a database", resulting in a smaller total of 2410 samples as shown in Table 3 (Insua & Martinez, 2003). This was a "*multi-step* sampling process" which ruled out TCs containing modal auxiliaries, semi-auxiliaries, and modal idioms and those somehow structurally incomplete (Insua & Martinez, 2003). Therefore, sentences like the following were further excluded from the sub-corpus.

(3)   I say, Ro ro Rod, there's gonna be (yeah) sorry, there's gonna be some people who don't even ha have no pupil comment on at all then? (F7F 1017/1019[4], Insua & Martinez, 2003)

(4)   I think there's a, there's <unclear> (F7C 1622, Insua & Martinez, 2003)

(5)   Er are we all agreed then we'll put down that there's a third </> (F7J 418, Insua & Martinez, 2003)

Sentence (3) contains a semi-auxiliary, *be gonna*, and (4) and (5) are incomplete sentences.

Although Insua & Martinez (2003) explain the sampling procedures, they lack the justification on the quantity and the quality of these samples. Firstly, the reader is at a loss as to how the total number of 2690 was chosen as appropriate to represent the whole in the first place. After going through the first process, the big corpus of all TCs—of which, by the way, the total number of TCs has not been given—was somehow reduced to a sub-corpus of not 5000, 10000, or a million, but of 2690 samples, but no rationale is given for that number. If data reduction were the purpose behind this extraction, narrowing down the scope of the research with specific goals in mind, rather than selecting out an arbitrary number of TCs could have served the purpose better.

---

[4] Citations in this format indicate the *file names* in the BNC.

This is a limitation that I try to resolve in Section 3.

Next, the representative power of the sub-corpus is also in question when the ratio of speech and writing or different genres is taken into account. The BNC consists of about 10 million words of spoken English and 90 millions of written counterparts (Insua & Martinez, 2003), with the latter exceeding the former surpassingly[5]. Nevertheless, the researchers *adjust* the proportions of spoken and written words to be equal in the first sampling, hence the spoken and written TCs of similar numbers in their sub-corpus (1071 vs. 1619 in Table 3, before undergoing the second sampling). Moreover, although the researchers acknowledge that the different genres of texts "are represented in *different* proportions within the BNC" (Insua & Martinez, 2003, emphasis by me), they fine-tune the proportions to be equal. Despite the fact that these two are evident manipulations against simple random sampling which is a reliable way of extracting samples, Insua & Martinez (2003) fail to give sufficient justification of why such artificial adjustments are indispensable.

Therefore, one cannot but question the credibility of the results to be discussed in 2.2.2 which are based on the samples extracted in the above manner.

2.2.2 Results and discussion

Insua & Martinez (2003) conclude that the fundamental cause behind non-concord in TCs is "short-term memory", which is attested to by the factors such as the on-line characteristic of speech communication and the length and complexity of the elements following the verb. In particular, the latter aspect is highlighted. As it is directly linked to the motivation of the present study, I stretch it out in great detail.

According to Insua & Martinez (2003), a TC is considered lengthier

---

[5] This is probably due to the difficulty in collecting oral productions, as keeping record of speech demands more work than writing.

Table 2. "Distribution of the TCs showing concord and non-concord" in Insua & Martinez (2003) (Insua & Martinez, 2003, Table 2)

	TCs showing concord		TCs showing non-concord	
	No.	%	No.	%
Written English	932	96.78	31	3.22
Spoken English	1255	86.74	192	13.26
Total	2187	90.75	223	9.25

and more complex when there is an extension after the NP, a coordination of NPs, or an intervening element between the verb and the NP. If there exists one of the three factors, a TC is more likely to be non-concordant.

2.2.2.1 Presence of an extension

Insua & Martinez (2003) observe that there is a contrast in the frequency of non-concord between minimal TCs and "those containing different types of complements or modifiers after the PVNP[standing for postverbal NP] itself" (henceforth, extended TCs), exemplified in (6).

(6)   Mobiles could be made smarter as there's no exams <u>at the end</u> (F7C 14; Insua & Martinez, 2003, underlined by me).

They conclude that non-concord occurs more frequently in the extended TCs (indicated as TCs with X), based on the following tables.

Table 3. "Length of the total number of TCs considered" in Insua
& Martinez (2003) (Insua & Martinez, 2003, Table 11)

	Minimal TCs		TCs with X		Total	
	No.	%	No.	%	No.	%
Written	122	12.67	841	87.33	963	100.00
Spoken	365	25.22	1082	74.78	1447	100.00
Total	487	20.21	1923	79.79	2410	100.00

Table 4. "Length of the TCs showing non-concord" in Insua &
Martinez (2003) (Insua & Martinez, 2003, Table 12)

	Minimal		TC with X		Total	
	No.	%	No.	%	No.	%
Written	4	12.90	27	87.10	31	100
Spoken	60	31.25	132	68.75	192	100
Total	64	28.70	159	71.30	223	100

Regarding the results, Insua & Martinez (2003) make the first mistake
in mis-analyzing the given statistics. They conclude that since the *raw
number* of NCs with an extension outnumbers minimal NCs as shown
in Table 6 (159 vs. 64), the former show greater non-concord.
Nevertheless, this is a misinterpretation of the data, not taking into
account the original proportion of each subgroup. That is, the
conclusion is logically flawed since the total number of the extended
TCs is already much greater than the minimal TCs, as seen in Table 5
(1923 vs. 487). The percentage of the extended TCs, in total, is 79.79%
and the minimal TCs, 20.21% (Table 5). As for the NCs, the percentage
of the former is 71.30% and the latter, 28.70% (Table 6). Therefore,
considering the *total* number, the authors' claim is contrary to fact.
There also arises a fundamental question on why such a factor as

extension must account for the phenomenon of non-concord in TCs. I argue that there is a missing link between the presence of an extension after the NP and the so-called "short-term memory" that Insua & Martinez (2003) deduce to be the ultimate cause underlying non-concord. In the extended TCs, because the extension occurs *after* the NP, as in (8), there is no memory loss or processing difficulty for the speaker/writer that can account for their misplacing a singular verb instead of a plural; the information about plurality of the NP immediately follows the verb, unlike sentences with modifiers *before* the NP like (7):

(7)   And there's <u>only about three</u> marks for it anyway so (FM4 1023; Insua & Martinez, 2003, underlined by me).

When one or more elements intervene between the verb and the NP as in (7), Insua & Martinez (2003)'s explanation seems satisfactory, but the TCs with an extension is not a valid target for survey.
Taking the two major flaws into consideration, the whole argumentation concerning the TCs with an extension loses its validity.

2.2.2.2 Coordination of NPs and presence of intervening material

Table 5. "TCs with coordinated PVNPs as regards concord and non-concord" in Insua & Martinez (2003) (Insua & Martinez, 2003, Table 13)

	Coordin.+concord		*Coordin.+non-concord*		*Total*	
	No.	%	No.	%	No.	%
Written	16	47.05	18	52.95	34/963	3.53
Spoken	8	25.81	23	74.19	31/1447	2.14
Total	24	36.92	41	63.08	65/2410	2.70

Table 6. "Distribution of TCs with intervening material" in Insua
& Martinez (2003) (Insua & Martinez, 2003, Table 14)

	Non-concord+ intervening material		Concord+ intervening material		Total	
	No.	%	No.	%	No.	%
Written	6	19.35	25	80.65	31	100
Spoken	26	13.54	166	86.46	192	100
Total	32	14.35	191	85.65	223	100

Based on Table 7 and 8, Insua & Martinez (2003) affirm that non-concord occurs more frequently in TCs with coordinated NPs and intervening material. There lie several limitations here as well.

First, the samples for both Table 7 and Table 8 are too limited in number. The total number of TCs with coordination is 65 and that of TCs with intervening material, 223, resulting in an even smaller number of each analyzed subgroup. For instance, the proportion of NCs with intervening material is 14.35% in Table 8, but, taking into account the total number, 223, and the small difference between that proportion and the average proportion of NCs in total, 9.25% in Table 4, TCs with intervening material cannot be generalized so easily to be more non-concordant than the average. Because the number of the bigger set is limited, the analysis coming from each Table lacks sufficient support from the data.

Second, it is dubious what consists of non-concordant instances in coordination. Insua & Martinez (2003) suggest that "singular agreement was found to be overwhelmingly predominant in TCs showing non-concord and co-ordination" but do not analyze exactly what kind of coordination follows the verb. It could be two NPs, three, or more, with all NPs being singular, only one of them singular, or both plural; the research paper chooses not to disclose.

Last but not least, the characteristics of intervening material in the given TCs are also not analyzed. Some elements in the samples could

be adverbs, and in some, quantifiers or other types of modifiers. I take the stance that not only the presence of the intervening elements matters but their different *qualities* affect non-concord in TCs, which is taken into consideration when collecting data in the present study.

Because Insua & Martinez (2003) lack sufficient amount of data for the analysis of the factors of coordination and intervening material and also characterization of the samples analyzed, the points made by Insua & Martinez (2003) are not strongly supported.

3. Research Method and Data

Although I used the same corpus as Insua & Martinez (2003), the *British National Corpus*, I re-collected data from the *BNCweb* (http://bncweb.lancs.ac.uk/), due to the limitations the research contained, as stated in Section 2.2. Of course, reducing the size of the corpus into an appropriate, analyzable subset was necessary. Nevertheless, as pointed out in 2.2.1 and 2.2.2, Insua & Martinez (2003)'s sample lacks representativeness, as the sampling size is somewhat arbitrary and small for a wholesome analysis, and the way of extracting the data is quite manipulative. Instead, I collected relevant data in the concrete manner I stretch out in this section.

First, the pivot of my research is narrowed down to non-concord between present-tense *be* and plural NPs, based on the generalization that this type takes up, beyond comparison, the majority of NCs. The previous works and the descriptive grammar books (Biber et al.,, 1995, p. 186; Quirk, 1985, p. 757) all confirm this generalization. Thus, for a more effective investigation, I executed a focused research and collected samples from the BNC web with plural nouns and a present form of *be—is, are, 's,* and *'re*[6]. This method reduces the whole number of data considerably, yet rendering the results more trustworthy than

---

[6] There was no sample with the contracted form of plural-be, there're.

Insua & Martinez (2003) who selectively pick out a rather unrepresentative sample out of the whole and Meechan & Foley (1994) who solely deal with a small number of oral production samples.

Second, TCs with a plural noun immediately following the verb *be* without intervening elements were collected. Therefore, examples like below were not included in my sub-corpus.

(8)   a. …so there is definitely benefits for being in partnership (KSV 1537).

Searching for TCs containing one or more intervening words between *be* and the semantic subject not only expanded the sub-corpora exceedingly but also caused parsing problems since in sentences of the form [there + be + (1$^+$ X) + plural NP], which serves as the input for the corpus search, the plural NP in that position could either be a subject or in the majority of cases, not. This meant that distinguishing between sentences exemplified as (9a) and (9b) with the same form [there + is + X + X + plural NP] had to be manually done.

(9)   a. We contest that…there <u>is</u> room for <u>hybrids</u> (A0B 440)
        b. …there <u>is</u> the slightest <u>foundations</u> for these remarks… (HRB 344)

In the above example, (9b) is an instance of non-concord while (9a) is not, but the website has no machinery for singling out only the relevant samples. Hence, sentences of this kind were ruled out.

Nevertheless, recognizing the potential importance of the role of elements positioned between *be* and the plural NP as suggested by Insua & Martinez (2003) and Meechan & Foley (1994), I separately collected TC samples with quantifiers of the highest frequency: *a lot of*, *a number of*, *a variety of*, *many*, and *(a) few*.[7] Quantifiers were chosen

---

[7] I allowed intervening segments between, for instance, *a* and *number*, or *of* and the

as the relevant class of intervening material, because these are *determiners*[8] that agree with the number information of their head nouns and thus can affect a language user's choice of the number of the preceding verb, to a greater degree than other elements like an adjective or an adverb can. For instance, quantifiers like *many* and *(a) few* only appear with plural nouns, while an adverb intervening between the verb and the NP like *greatly* can come regardless of the number of the notional subject. In this way, unlike Insua & Martinez (2003) who indiscriminately consider TCs with intervening material *one uniform class* as opposed to those without, I attempted to identify each quantifier and investigate how different *qualities* of quantifiers can affect the proportion of NCs.

Lastly, in order to examine non-concord in TCs with the coordinated notional subject, samples in the form of [there + be + NP1 + and + NP2] were included in the data. Nevertheless, coordinated structures in which the NPs cannot be judged as a valid coordination of separate notional subjects were ruled out, as in (10) and (11).

(10)   There is <u>love</u> **and** <u>love</u>, girl, and his feelings go beyond it (HWE 367).

(11)   Since there is <u>speculation</u> **and** <u>King Richard</u> must surely be

---

plural NP in order to include sentences like *There is/are a growing number of young people on the street* and *There are too many clever kids in the class*. Unlike the aforementioned parsing difficulty, this input rarely had a non-subject as the plural NP. Of course, unaccepTable samples with the above form, such as *There is in many countries a strange phenomenon* were ruled out, though very rare.

[8] Quantifiers belong to one type of determiners, *postdeterminers*, that again fall into four different groups: cardinal numerals (e.g. *two*), ordinal numerals (e.g. *first*), closed-class quantifiers (e.g. *many*), and open-class quantifiers (e.g. *a lot of*) (Quirk et al., 1985, pp. 260-264). The relevant groups in the present study are closed-class and open-class quantifiers: *many, few,* and *a few* pertaining to the former, and *a lot of, a number of, a range of,* and *a variety of* to the latter. While words like *variety, range*, and *number* are singular nouns on their own, when these are used in contexts to specify the quantity of the head noun, they behave as quantifiers. Their role as quantifiers is more apparent when a modifier precedes them (e.g. *a wide range of activities, an infinite variety of books, a great number of people*).

aware of it… (CCD 2416)

In so doing, I attempt to provide with more transparent criteria in judging non-/concord and a deeper analysis of non-concord in TCs with coordinated NPs with taking each conjunct into account, compared to Insua & Martinez (2003) who do not analyze the coordinated NPs any further, as explained in 2.2.2.2.

To sum up, all data from the BNC, not a selected sample as in the previous studies, were collected that met the criteria above. Thus, this paper aims to conduct a focused study with clear-cut criteria of the data, yet employing the largest sub-corpus possible in order to increase the reliability of the results.

4. Results and Discussion

Following is a table summarizing the results of the overall research.

Table 7. Distribution of Cs and NCs

NP type	Concord		Non-concord		Total
	No.	%	No.	%	No.
NP	5654	91.43	530	8.57	6184
a lot of NP	191	61.02	122	38.98	313
a few NP	220	81.78	49	18.22	269
a number of NP	910	94.30	55	5.70	965
a range of NP	33	19.53	136	80.47	169

a variety of NP	59	*50.43*	58	*49.57*	117
many NP	2308	*96.09*	94	*3.91*	2402
few NP	495	*98.61*	7	*1.39*	502
NP1 and NP2	177	*54.97*	145	*45.03*	322
Total	10047	*89.36*	1196	*10.64*	11243

The potential factors affecting non-concord to be investigated are the existence of coordination, the presence and quality of an intervening quantifier, and the distinction between speech and writing.

4.1 Speech vs. writing

As many linguists have pointed out, spoken samples showed greater non-concord than written counterparts did. Table 10 and Table 11 illustrate the distribution of NCs between the spoken and written styles.

Table 8. Distribution of Cs and NCs in spoken and written TCs

	Concord		Non-concord		Total
	No.	**%**	**No.**	**%**	**No.**
Spoken	978	*61.13*	622	*38.88*	1600 (*14.23%*)
Written	9069	*94.05*	574	*5.95*	9643 (*85.77%*)
Total	10047	*89.36*	1196	*10.64*	11243

Table 9: Distribution of Cs and NCs in spoken and written TCs across NP types

NP type	Concord		Non-concord		Total
	Spoken	Written	Spoken	Written	
NP	628 (11.11%)	5026 (88.89%)	332 (62.64%)	198 (37.36%)	6184
a lot of **NP**	57 (29.84%)	134 (70.16%)	92 (75.41%)	30 (24.59%)	313
a few **NP**	20 (9.09%)	200 (90.91%)	40 (81.63%)	9 (18.37%)	269
a number of **NP**	99 (10.88%)	811 (89.12%)	15 (27.27%)	40 (72.73%)	965
a range of **NP**	3 (9.09%)	30 (90.91%)	17 (12.50%)	119 (87.50%)	169
a variety of **NP**	5 (8.47%)	54 (91.53%)	7 (12.07%)	51 (87.93%)	117
many **NP**	128 (5.55%)	2180 (94.45%)	75 (79.79%)	19 (20.21%)	2402
few **NP**	18 (3.64%)	477 (96.36%)	5 (71.43%)	2 (28.57%)	502
NP *and* **NP**	20 (11.30%)	157 (88.70%)	39 (26.90%)	106 (73.10%)	322
Total	978 (9.73%)	9069 (90.27%)	622 (53.21%)	574 (49.10%)	11243

From Table 10, one can notice a much higher percentage of non-concord for the spoken TCs compared to that of the written ones (38.88% to 5.95). In fact, the former is significantly higher than the result from Insua & Martinez (2003), which was only 13.26%. Likewise, from Table 11, it is evident that for the NCs, the spoken samples *outnumber* their written counterparts. This is striking, considering the paucity of spoken constructions in the database, accounting for only 14.23% of the total sub-corpus as seen from Table 10.

4.2 Coordination of NPs

From Table 9, it can be observed that a significant number of TCs with coordinated structures (NP1 and NP2) show non-concord. So far, Insua & Martinez (2003)'s analysis of TCs with coordinated NPs is confirmed. However, as discussed in 2.2.2.2, their research lacks details about what exactly is seen as non-concord and about the number information of each conjunct. The first limitation is resolved by stating the rule for concord in coordinated structures, and the second, by further analyzing the given data.

Firstly, the prescriptive rule for subject-verb agreement in coordination is that when the two coordinated NPs are separate entities, the plural form of a verb must be used. This obtains even when each member in coordination is singular count noun or a mass noun[9] (Quirk et al., 1985, p. 759).

(12)   a. Marks and Spencers is offering a new sale.
         b. John and Mary are going to the shop.

In (12a), the subject is a coordinated NP, but denotes one entity, a brand name,[10] while in (12b), the coordinated subject consists of two entities, and thus plural-*be* is used. Likewise, in TCs, the same norm applies. Thus, the following contrast in grammaticality (in terms of strict grammar) occurs:

(13)   There are/*is a vase and a book on the table.
(14)   There are/*is plates and cups all over the place.

---

[9] For convenience, mass nouns (e.g. *water*) will be included in the notion of singular nouns as they also agree with a singular verb.

[10] The TC counterpart of this sentence would be, '*There is Marks and Spencers offering a new sale*'. Examples like this are guaranteed to be ruled out, because as stated in Section 3, only those with a coordination of *separate notional subjects* were collected.

Next, a table with the number information of each NP in coordinated TCs is provided below.

Table 10. Distribution of Cs and NCs with coordinated NPs

	Concord		Non-concord	
	No.	%	No.	%
sg NP1 + sg NP2	0	*0*	125	*86.20*
pl NP1 + pl NP2	177	*100*	16	*11.03*
sg NP1 + pl NP2	0	*0*	3	*0.21*
pl NP1 + sg NP2	0	*0*	1	*0.07*
Total	177	*54.97*	145	*45.03*

The table shows that in all Cs with coordinated NPs, NP1 and NP2 are both plural, exemplified in the sentences in (15).

(15)   a. Many people believe that there are <u>schoolrooms</u> and <u>football grounds</u> where civilized order is forever on the verge of breaking down. (ECN 8)
b. There are <u>tables</u> and <u>anecdotes</u> to satisfy the enthusiast… (K9V 541)
c. There are <u>advantages</u> and <u>disadvantages</u> in both methods. (HJE 935)

On the other hand, most of the coordinated NCs have singular conjuncts ((16)), a small number having plural forms for both ((17)),

and even less, one singular and the other plural ((18) and (19)). In other words, when singular conjuncts are coordinated, there is, without exception, non-concord.

(16) Well, I love you for a start — and there's <u>Mum</u> and <u>Dad</u>… (FS1 1411)

(17) I'm warm and free and there's <u>lights</u> and <u>people</u>, and nobody's asking me questions. (A74 2723)

(18) There's <u>bacon</u> and <u>eggs</u> in the fridge. (HHA 658)

(19) There's <u>mushrooms</u> and <u>cream</u> in your starter. (EDJ 999)

This finding suggests that when a singular noun immediately follows the verb *be*, a speaker/writer has difficulty in selecting the correct plural form of the verb. This is possibly due to the discrepancy between the number of the nominal element to be uttered right after the verb (singular) and that of the whole coordinated structure (plural). Also, there is a time gap between when the first conjunct appears and when the actual number of the coordinated subject is finally resolved—at the end of the coordinated structure—which complicates the matter of concord for the speaker/writer.

4.3 Presence of an intervening quantifier

As *length* is the most crucial factor determining non-/concord for speakers and writers in Insua & Martinez (2003), as pointed out in 2.2.2, TCs with an intervening element are more prone to be NCs. Nevertheless, our results show that it is not the mere *presence* of intervening material that encourages non-concord, but its *quality*.

In Table 9, not all TCs with intervening material (*a lot of*, *a few*, *a number of*, *a range of*, *a variety of*, *many*, and *few*) uniformly show more frequent non-concord than those without, contrary to Insua & Martinez (2003)'s prediction. In particular, for quantifiers like *many*

and *few*, the percentage of NCs is even substantially lower (3.91% and 1.39%, respectively) than the average.

Accordingly, a distinction can be made between the quantifiers taking the form '*a X* (*of*)' and '*X*', with the result in Table 13.

Table 11. Distribution of TCs across quantifier type

Quantifier	Concord		Non-concord		Total
	No.	%	No.	%	No.
Ø	5654	*91.43*	530	*8.57*	6184
a X (*of*)	1413	*77.09*	420	*22.91*	1833
X	2803	*96.52*	101	*3.48*	2904
Total	9870	*90.38*	1051	*9.62*	10921

In Table 13, *a X* (*of*)-quantifiers, in total, show more variation in concord than when the verb is directly followed by a plural noun (Ø - quantifier) or a single-word quantifier (*X*-quantifier). The sentences in (20) and (21) exemplify Cs and NCs with a *X* (*of*)-quantifier.

(20)   Cs with *a X* (*of*)-quantifier
        a. Administration takes up much of the day, especially if there are <u>a lot of</u> patrols. (A77 1255)
        b. Today there are <u>a few</u> very old men in Lebanon... (A1V 532)
        c. There are <u>a number of</u> cherries which will interest the wildlife-oriented gardener. (A0G 1046)
        d. There are <u>a</u> whole <u>range of</u> features designed to help save

you time and money... (EEJ 580)

e. There are <u>a variety of</u> reasons for this, which may be deeply rooted in our past. (B21 565)

(21)   NCs with *a X* (*of*)-quantifier

a. ...there's <u>a lot of</u> things been said about these benifits... (HDP 331)

b. So obviously, there's <u>a few</u> tough games ahead... (KS7 752)

c. There is <u>a</u> limited <u>number of</u> pitch patterns in any one language... (H0Y 2898)

d. There is <u>a</u> wide <u>range of</u> standards at which you can compete, ranging from weekly club races to national series events… (AT6 1487)

e. There is <u>a variety of</u> approaches in graduate schools... (A1A 1207)

In parallel with these results in Table 13, in Table 9, each quantifier in the form *a X* (*of*) shows non-concord to a far greater extent than the average. An exception to this generalization is the quantifier *a number of*. Perhaps it is due to prescriptive teaching that puts forward *a number of* as a representative quantifying expression that comes before a plural noun and agrees with a plural verb. The majority of the prescriptive grammar books seem to contain the famous distinction between *a number of* NP and *the number of* NP: the former takes a plural verb while the latter, a singular verb.

Other than *a number of*, the effect of *a* is manifest, especially when the striking contrast between the results of *few* and *a few* is taken into account. Although both only occur with plural count nouns as their head, the speaker/writer seems to mistakenly refer to the element immediately following *be* (i.e. *few* and *a*, respectively) for number information. This may be linked to what Quirk et al. (1985) note of the

indefinite article: "[T]he indefinite article derives historically from the unstressed form of *one*, and in present-day English there are still many contexts in which the numeral function is uppermost" (p. 273). Whether this explanation is satisfactory, it is a fact that the co-occurrence requirement for *a* is that a singular noun must follow. Therefore, although the *head* noun, the actual target of the subject-verb agreement, is in its plural form, due to the close proximity of the indefinite article and a singular noun (e.g. *range*) and to the relative distance of the head noun, non-concord is prevalent.

On the other hand, when the *X*-quantifiers (i.e. *many* and *few*) appear, there is the effect of having a plural noun directly following the verb since the number information is readily seen. One could also argue that in the presence of the plural quantifiers *many* and *few*, the plurality of the NP is so accentuated that speakers/writers make a conscious effort to use *are*, because if not, the mismatch between the number of the verb and the NP would be too obvious. The sentences in (22) are examples of Cs with *many* and *few*, which are much more common than examples in (23), as contrasted in Table 13.

(22)   Cs with an *X*-quantifier
       a. It is plain that there are <u>many</u> ways of being and of imagining a someone other. (A05 498)
       b. There are <u>a few</u> details I want to clear up with the other girls. (A0F 896)

(23)   NCs with an *X*-quantifier
       a. Erm, there's <u>many</u> suggestions I've got here... (KM4 164)
       b. ...there's very <u>few</u> sentences in between to explain what's going on. (FA6 1160)

In brief, it is not the mere presence of an intervening element, as argued in Insua & Martinez (2003), but rather its inherent *quality* that accounts

for the frequency of non-concord in the TCs with an intervening element. Depending on whether the number information indicated by the quantifier matches that of the head noun (*X*-quantifier) or not (*a X* (*of*)-quantifier), the intervening element is further classified into two different categories, rather than behaving as *one uniform class* encouraging non-concord, as suggested in Insua & Martinez (2003).

5. Conclusion

Two general conclusions can be drawn from the findings in Section 4. The first is the contrast between speech and writing, and the second concerns the speaker/writer's memory and concord.

Firstly, non-concord in TCs occurs more frequently in speech (38.88%) than writing (5.95%). The proportion of NCs in speech is smaller than that in Meechan & Foley (1994) (72%) and bigger than Insua & Martinez (2003) (13.26%). Although non-concord does not occur in the majority of TCs like in Meechan & Foley (1994) and thus cannot be considered as the default agreement, it still accounts for a significant part of TCs to be considered quite *acceptable* in speech.

Secondly, an element closer to the verb is likely to serve as a reference point for determining the number of the verb, as seen from the results in Section 4.2 and 4.3. Thus, contrasts occur in the following NCs:

(24) a. There is <u>many</u> people. (*X*-quantifier)
　　　b. There is <u>people</u>. (Ø -quantifier)
　　　c. There is <u>a few</u> people. (*a X* (*of*)-quantifier)
　　　d. There is <u>an apple</u> and a banana. (NP1 and NP2)

In (24), the proportion of NCs for each subclass increases from (a) to (d), with that in (a) the smallest and (d), the largest. The factor involved here is memory correlated with the apparent number information of the

immediately following element. In this respect, Insua & Martinez (2003)'s claim about "short-term memory" based on the length and complexity of TCs is only half right at its best. It seems correct at a first glance because the farther the head noun is from the verb, the harder it is to correctly make them agree ((24c)). Nevertheless, the case does not stand in examples like (24a): even though the head noun is intervened by a quantifier, the latter does not encourage non-concord. Hence, Insua & Martinez (2003)'s explanation is only a partial analysis and rather a hasty generalization, since when analyzed in depth, there are differences even among the class of so-called lengthy and complicated TCs.

As illustrated in Section 4.3, the mere presence of an intervening element does not determine the frequency of NCs but rather whether that quantifier overtly expresses the number information of the head noun (*X*-quantifier), or has the surface number information not matching that of the head noun (a *X* (*of*)-quantifier) does. In the latter case, there is a memory loss for the speaker/writer due to the presence of the indefinite article *a* followed by a *singular* noun (e.g. *variety*) that mismatches the number information of the *plural* head noun.

This also explains coherently why, in the coordinated TCs with singular conjuncts, non-concord occurs without exception. Although the number of the whole coordinated subject is plural, the number of the NP immediately following the verb is singular, resulting in a discrepancy. Thus, the speaker/writer is prone to losing track of the right number information for the verb. Without the characterization of the concord rule in coordinated structures and of the conjuncts, a sweeping generalization like that of Insua & Martinez (2003) cannot hold.

References

Biber, D., Johansson, S., Conrad, S., & Finegan, E. (1999). *Longman*

grammar of spoken and written English. London: Pearson.

Breivik, L. E. (1981). On the Interpretation of Existential *There*. *Language*, *57*(1), 1-25.

Butler, C. S. (1985). *Statistics in Linguistics*. Oxford, New York: B. Blackwell.

Cowan, R. (2008). *The teacher's grammar of English: a course book and a referential guide*. New York, NY: Cambridge University Press.

Insua, A. E. & Martinez, I. M. (2003). A corpus-based approach to non-concord in present day English existential *There*-constructions. *English Studies*, *84*(3), 262-283.

Meechan, M. & Foley, M. (1994). On resolving disagreement: Linguistic theory and variation—*There's* bridges. *Language Variation and Change*, *6*, 63-85.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York: Longman.

Seo Jeong Park
sj.park@snu.ac.kr