Coordinated Pronoun Variation in American English

Kristin Van Engen, Northwestern University

English pronouns in coordinated structures have long captured the attention of prescriptive grammarians and linguists alike, largely because these little words exhibit variation with respect to both case and word order (e.g. John and I, John and me, me and John). Prescriptive grammars typically dictate that coordinated pronouns should bear the same case marking as non-coordinated pronouns (Blount and Northup 1931; Greenbaum 1989; Perrin 1972; Wales 1996; Weiner 1983; Williams 1999). Further, it has been argued that politeness influences word order variation in coordinated pronoun structures as it is considered polite for first person pronouns to be placed at the end of coordinated noun phrases (NPs) (e.g. Quirk et al. 1985, cited in Wales 1996). However, English speakers regularly flout the prescriptive rules. This paper will show that this coordinated pronoun variation differs across modality (speech versus writing). It will also show, crucially, that linguistic factors that affect word ordering in other domains—specifically constituent weight and information structure—play a significant role in the usage of coordinated pronoun forms that defy prescriptive rules.

(1a) exemplifies the prescriptively correct form for a coordinated first-person, singular (1sg) pronoun in subject position: the pronoun I is nominative and in second position. However, as indicated by (1b-d), this is not the only possible form. (1b-c) show the accusative pronoun me in subject position, and (1c-d) show the nominative 1sg pronoun I in first position.

(1) Subject position:
a) *John and I* had a good time at the party.

b) *John and me* had a good time at the party.

c) *Me and John* had a good time at the party.

d) %*I and John* had a good time at the party.

All of the above forms are attested in English, although (1d) is notably rare by most accounts of speech data. Parker et al. (1988) claim that “As a rule, coordinate structures of the form *X and I* are acceptable … while those of the form *I and X* are not.” Angermeyer and Singler (2003, 178) also mark this form as ungrammatical in their analysis of object position coordination, but they do note that this form occurs in *subject position* in speech (rarely) and in writing. Aside from an example of *I and X* in the writing of William Labov, however, they provide limited evidence for these claims. The %-marking is used above to indicate that a form is attested in language use, though rarely, and that speakers vary as to whether or not they accept it.

As shown in (2)-(3), both nominative and accusative coordinated pronouns also occur in object positions.

(2) Object position:

a) The cow kicked *John and I*.

b) The cow kicked *John and me*.

c) The cow kicked *me and John*.

d) %The cow kicked *I and John*.

(3) Object of a preposition:

a) The gift is for *John and I*.

b) This gift is for *John and me*.

c) The gift is for *me and John*.
The gift is for I and John.

The form in (2a) and (3a), in which the nominative 1sg pronoun is used in object position (and in second position) is often referred to as the “hypercorrect” form (e.g. Boyland 2001). This interpretation stems from the assumption that people are frequently corrected for using coordinated me in subject position, so that they not only learn to use I in subject position, but they overextend I into the object position.

Crucially, the pronoun case variation described here does not occur in non-coordinated structures:

(4) I/*me had a good time at the party.
(5) The cow kicked *I/me.
(6) The gift is for *I/me.

According to Webster’s Dictionary of English Usage (1994, 778), pronoun case variation in coordinated NPs has been around for hundreds of years, and it’s been raising the ire of grammarians for hundreds of years as well: prescriptivist complaints against uses of the nominative pronoun in object position (e.g. told my lord and I; between you and I) have been raised at least as far back as the 1760s (Leonard 1962, cited in Angermeyer and Singler 2003). The endurance of both the variation and the prescriptive injunctions against it make this particular grammatical structure an interesting case for studying linguistic variation in general, as well as for investigating the effects of prescriptivism on language variation and change. And indeed, linguists of various stripes have taken up the phenomenon as a topic of study. A variety of accounts have been proposed, some emphasizing the role of syntactic properties of the construction (e.g. optionality or underspecification in the syntax) (Johannessen 1998; Parker et al. 1988; Shorrocks 1992), some taking syntax and other factors such as phonological properties
into account in constraint-based analyses (Quinn 2005; Sadock 2005) and others emphasizing the importance of language-external factors (e.g. prescriptive rules) in the realization of pronoun case (Emonds 1986; Parrot 2007; Sobin 1997). Recently, variationist and sociolinguistic studies have taken up variation in pronoun coordination in order to examine the nature of the variation itself along with the social and linguistic factors that correlate with the usage of particular variants (Angermeyer and Singler 2003; Boyland 2001; Grano 2006).

These recent studies have provided important empirical information about rates of variant usage, speaker preferences, and some of the social and linguistic factors that influence coordinated pronoun variation. However, several factors remain to be examined. In this paper, I provide usage data from the second release of the American National Corpus (ANC) (Reppen et al. 2005), particularly focusing on 1sg pronouns in subject position. I present evidence that linguistic modality (speech versus writing), the constituent weight of the conjuncts, and the information structural status of the NP conjoined with the pronoun are all contributors to the distribution of the variants. I show that these factors must be considered in a comprehensive account of pronoun variation in coordination.

In this paper, I will adopt the following definitions:

*Coordinated Noun Phrase:* a noun phrase in which two (or more) pronouns or noun phrases are combined by a coordinating conjunction such as *and* to form a single noun phrase (as in examples 1-3).

*Conjunct:* one of the two (or more) nouns in a coordinated NP. Since this paper focuses on pronouns in coordinated NPs, “conjunct” will typically be used in reference to the *other* NP in the coordinated NP.

*Conjunction:* the word(s) that join the two conjuncts (e.g. *and*, *or*). Note that only
coordinated noun phrases conjoined with and will be included in this analysis.

**Coordinated Pronoun:** a pronoun that is a conjunct in a coordinated NP.

**Previous Findings**

Angermeyer and Singler (2003) conducted a study of 1sg pronouns in object position, terming the variants the “Vernacular” (*me and X*), the “Polite” (*X and I*), and the “Standard” (*X and me*). They showed that social factors such as age and education level were predictors of the usage of the different variants—the Vernacular was used by the youngest and least educated speakers in their corpus, the Polite was used by those intermediate in age and education, and the Standard was used by the oldest, most educated speakers. Angermeyer and Singler also identified a number of linguistic factors that affected the usage of different variants. These included characteristics of the other conjunct such as grammatical weight, animacy, plurality, and familiarity of the referent to the speaker. They also found that aspects of the constituent that followed the object-position coordinated NP affected which variant was used: where the coordinated NP acted as the underlying subject of a gerund, bare verb, or infinitive, a coordinated nominative pronoun was more likely to be used. Likewise, when the object-position coordinated NP was followed by a verb phrase but was *not* the underlying subject of that VP, the nominative was more likely to be used (e.g. *The priest who married my wife and I told me*). In addition to identifying these factors, Angermeyer and Singler conclude, based on their evidence, that all three of the object-position variants (*me and X, X and I, X and me*) currently exist in stable variation.

Angermeyer and Singler (2003) (also Boyland 2001) focused on 1sg variation in object position only. Recent work by Grano (2006) has examined a greater range of data, considering case, person, and word order. Grano’s study included speech data from the Fisher English
Training Speech corpus along with acceptability judgments elicited via magnitude estimation. Grano found that 1sg coordinated pronouns are more likely to be nominative than are 3sg pronouns in both subject and object positions. In addition, the magnitude estimation study showed that, even when speakers conform to prescriptive standards to a high degree, they impose two additional constraints when coordinating a pronoun and a full NP: the “X and I” constraint—that 1sg nominative is limited to second position; and the “s/he and X” constraint—that 3sg nominative is limited to first position.

The Present Study

The research discussed in the previous section showed that social and linguistic factors contribute to variation in coordinated pronouns. The study presented in this paper extends this line of inquiry to 1sg subject position coordinated pronouns. It also examines linguistic factors that are likely contributors to the variation but that have been overlooked in previous work: linguistic modality, grammatical weight of the conjuncts, and information structure. To do this, I examined the four variants illustrated in (7):

(7) Subject position, coordinated 1sg pronouns

a. X and me: cuz her and me are warring
b. X and I: my ex-boyfriend and I got into this debate like at least once a month
c. me and X: but now me and my mom are a lot closer
d. I and X: and then I and someone else intervene and stop him

I investigated coordinated first person pronouns in subject position for several reasons. First, while first person pronouns have been the focus of previous studies (Boyland 2001; Angermeyer and Singler 2003), these studies analyzed object position, leaving subject position relatively understudied. In addition, the issue of hypercorrection (the extension of the nominative X and I
form into object position due to its strong enforcement in subject position) is avoided in analyzing subject position itself. Finally, Grano’s (2006, 62) study indicated that, in speech, \textit{X and I} is prevalent in subject position, that \textit{me and X} is also in use in this position, and that \textit{I and X} and \textit{X and me} are rarely used (if at all). This variability, and the near-categorical lack of \textit{I and X} and \textit{X and me}, make the set of variants in (7) a good locus for examining the relative importance of modality in coordinated pronoun variation, as well as for examining linguistic and non-linguistic factors that may correlate with the usage of the infrequent variants.

**Factors Under Investigation**

**Modality.** Prescriptive grammars call for the usage of the same pronoun cases in coordinated structures that would be used in non-coordinated structures. In written contexts, as opposed to in speech, usage manuals are more likely to be consulted and/or professional editors who are aware of the standards may be employed to enforce them. Furthermore, speakers/writers who know the standard forms are more likely to put them to use in writing, even if not always in speaking. Presumably, then, standard forms (\textit{X and I} for subject position coordination) are used at a higher rate in written English as compared to spoken English. If this is so, then it is possible, as suggested by Angermeyer and Singler (2003, 200), that the use of standard forms in writing may contribute to their continued use in spoken English. Previous studies on variation in coordinated pronouns have not included naturally occurring written data. This study will provide a comparison between speech and writing in order to determine the role of linguistic modality in coordinated pronoun variation.

**Weight.** The gradient acceptability and usage of coordinated pronoun variants may also be affected by factors that affect word order in other contexts. One such factor is constituent weight, which affects the ordering of grammatical constituents cross-linguistically. In general, longer
and/or more syntactically complex NPs tend to be placed after comparatively lighter noun phrases when syntactic structures allow variable ordering. Quirk et al. (1972) termed this trend “end-weight,” and described it as “the tendency to reserve the final position for more complex parts of a clause or sentence” (quoted in Wasow 1997). Recent studies on end-weight effects in English (Wasow 1997, 2002; Wasow and Arnold 2003) show that this tendency holds across a range of structures. In the dative alternation (e.g. *Kim handed a toy to the baby* vs. *Kim handed the baby a toy*), for example, there is a tendency for speakers to use the alternant that allows for later placement of the heavier constituent.

Since end-weight is a factor in other structures that allow for word order variation in English, it is also a likely contributor to word order variation with coordinated pronouns. Angermeyer and Singler (2003) found that conjuncts with four or more syllables (relatively heavy conjuncts) greatly favored the Vernacular *me and X* variant in object position—the one variant they considered in which the other conjunct could occur after the coordinated pronoun. The patterns of end-weight that occur elsewhere in English, combined with specific evidence for its effect in coordinated pronouns in object position, motivates the following prediction: subject-position variants in which the pronoun comes first (*me and X, I and X*) are more likely to occur with heavy NP conjuncts than with light NP conjuncts.

*Information structure.* A factor that has been given relatively little explicit attention in discussions of coordinated pronoun variation is information structure, which refers to the organization of information in a discourse, and hence to the ordering of constituents² (Prince 1981; Ward and Birner 2004; Birner and Ward 1998). Information structure has been shown to affect constituent ordering in other contexts such that new information tends to follow old (or given) information.
Although constituent weight and information structure have both been cited as factors influencing word order, Wasow and Arnold (2003, 128) note that the literatures on weight and information structure are largely disjoint, even though the two factors are highly correlated and perhaps inseparable. They point out that references to old information can often be quite short, since the information is already known to listeners. In particular, pronouns are single words in English, and also typically point to information already in the discourse, so that length and newness are likely to be highly correlated—both factors that would favor a constituent’s placement later than comparatively short/old constituents in structures that allow word order variation.

If the old-before-new patterning holds true for pronoun coordination, then it is predicted that structures where the pronoun comes first (me and X; I and X) will be more likely to be used when the other conjunct represents new information.

**Corpus Study**

The factors of modality, weight, and information structure were examined by studying case and word order variation in coordinated 1sg pronouns in subject position in a corpus of written and spoken American English.

**Methods.** The data for the study came from the American National Corpus Second Release, which contains more than 22 million words of written and spoken American English (Reppen et al. 2005). This corpus was selected because it contains both spoken and written data and because the data is very recent (primarily from the 1990s and 2000s), allowing for assessment of contemporary English usage.

First person singular pronouns were collected by using GREP to identify all occurrences of the following strings: *I and, and I, me and, and me*. The output of the search was then manually
filtered in order to include only the tokens in which the pronoun was indeed a conjunct in a coordinated noun phrase. (This eliminated, for example, sentences like *I got angry and I slammed the door.*) All possible coordinated pronoun tokens from the corpus were extracted, yielding a total of 1722 tokens. The data were further filtered to include only tokens in which a non-coordinated pronoun would be either consistently nominative or consistently accusative (see Grano 2006, 34-35 for a similar method). This procedure eliminated tokens in which the coordinated pronoun was isolated (comprising an entire utterance and/or in a syntactic fragment), in a predicative complement, right- or left-dislocated, or in apposition. Tokens in which the coordinated pronouns were used metalinguistically, such as (8) were also omitted.

(8) …lowering of the standards of civility and yielding, typically, constructions like *Me and her* went to the park (as contrasted with “*Her and me* went to the park,” I suppose).

(Verbatim: journal on language)

Tokens were then categorized according to grammatical role (subject, object, or prepositional object) and the subject tokens (1150 total) were subjected to further analysis. Divided by modality, the spoken data constituted approximately 34.4 percent of the subject-position tokens, while the written data constituted approximately 58.4 percent. The remaining tokens (7.2 percent of the data) were tokens of quoted speech in written texts. These were omitted from subsequent analysis due to the complications of characterizing them as either spoken or written.

Examples of each token type, taken from the corpus, are shown in (9)-(12):

(9) *X and me*

a. i wanted to ask you if you'd be willing to go in on a scheme with me to um cuz *her and me* are warring (Spoken academic discourse: Micase)

b. Thanks for the link, Sean, I answered a question 3 *brilliant people and me* were
unable to figure out a couple days ago. "Singapore: country or city?!" Turns out, it's both. (Online chat: Buffistas.org)

(10) X and I

a. Some friends and I have collected $23 to get your campaign started.³ (Slate magazine)

b. … my husband and I haven't done a whole lot of camping we but we bought a van last year … (Telephone speech: Switchboard)

(11) me and X

a. … I went and laid down with him and me and him fell asleep.

   (Face-to-face speech: Charlotte narratives)

b. In Jax, Fla, you could take PE for six weeks of half-days in the summer and not have to take it during the school year. Me and all my jr. hi geek friends did that and spent the day like this: Walk 12 times around track (required for everyone). Play chess, ping-pong … (Online chat: Buffistas.org)

(12) I and X

a. I never presented a paper at a conference, but I and several of my friends (not all in the same department) found our master's level graduate work less challenging than the courses we'd taken as juniors and seniors at Brown. (Online chat: Buffistas.org)

   b. I and my co-workers happen to be Ebonics speakers. (Slate magazine)

Results and discussion. The combined data (spoken and written) from the American National Corpus (Table 1) shows that the prescriptively standard X and I variant is overwhelmingly dominant (94.38% of the tokens).
Table 1 approximately here.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>X and me</td>
<td>5 (.47%)</td>
</tr>
<tr>
<td>X and I</td>
<td>1008 (94.38%)</td>
</tr>
<tr>
<td>me and X</td>
<td>36 (3.37%)</td>
</tr>
<tr>
<td>I and X</td>
<td>19 (1.78%)</td>
</tr>
<tr>
<td>Total</td>
<td>1068</td>
</tr>
</tbody>
</table>

Table 1. Coordinated 1sg pronouns in subject position. All tokens from the American National Corpus (after filtering), collapsed across modality. For each variant, the number of tokens is given, along with its proportion of total subject-position 1sg coordinated pronoun usage.

This finding is in accordance with previous findings, e.g. Grano (2006, 62), and with the hypothesis that the high frequency of *X and I* in its standard position may contribute to its extension to object position (Boylan 2001; Grano 2006). In addition to the standard *X and I* variant, *me and X* was observed in 3.37% of the tokens and *I and X* in 1.78% (note that these *I and X* tokens came nearly exclusively from written data, as will be discussed below). The rarest variant in subject position was *X and me* (.47%).

Each of the subject position tokens was coded by source text, modality, and by the weight of the other NP conjunct (measured in number of words). Finally, each token was coded for the information structure status of the conjunct—specifically for whether or not the NP conjoined with the pronoun represented discourse-new information. Conjunct NPs were coded as discourse-old if the entity denoted by the conjunct was explicitly evoked in the preceding 100 lines of text.
**Modality.** Pearson’s chi-squared test shows that the two modalities differ significantly in the distribution of coordinated pronoun variants (p < .001), though this must be taken as suggestive given the small number of tokens in some cells. This finding is supported by Fisher’s Exact test (p < .001), which handles smaller numbers better and also indicates that usage is significantly different across modalities.

Table 2 approximately here.

<table>
<thead>
<tr>
<th>Variant</th>
<th>Spoken tokens</th>
<th>Written tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>X and me</td>
<td>2 (.50%)</td>
<td>3 (.45%)</td>
</tr>
<tr>
<td>X and I</td>
<td>367 (92.21%)</td>
<td>641 (95.67%)</td>
</tr>
<tr>
<td>me and X</td>
<td>28 (7.04%)</td>
<td>8 (1.19%)</td>
</tr>
<tr>
<td>I and X</td>
<td>1 (.25%)</td>
<td>18 (2.69%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>398</td>
<td>670</td>
</tr>
</tbody>
</table>

Table 2. Coordinated 1sg pronouns in subject position. All tokens from the American National Corpus (after filtering), divided by modality. In each modality, the number of tokens and proportion of subject-position 1sg coordinated pronoun usage is given for each variant.

In general, the data from the ANC shows that the prescriptively standard *X and I* variant is used with by far the greatest frequency in both speech and writing. Usage rates by modality reveal that the rate of *X and I* usage is slightly higher in written English (95.67% compared to 92.21%). Another difference is observed in the rate of usage of *me and X*, which is higher in speech than in writing (7.04% versus 1.19%), whereas the usage of *I and X* is higher in writing versus speaking (2.69% compared to .25%). The appearance of *I and X* in writing may reflect the authors’/editors’ desire to maintain the standard case (nominative) in a context where other
factors such as conjunct weight or information structure favor pronoun-first ordering. For example, in (13), old information comes before new, and the light conjunct comes before the heavy one:

(13) *I and many of my friends* think that Mrs. Dole has impeccable credentials and would have been a very knowledgeable president.

(Slate Magazine)

In summary, the prescriptive standard is used more consistently in writing than in speech, the most common non-standard variant for speech is *me and X*, and the most common non-standard variant in writing is *I and X*.

*Weight.* With respect to weight, the data show that the pronoun-first variants (*me and X, I and X*) do occur with longer conjuncts on average than the variants in which the pronoun is last—the average number of words in X was 2.50 for *me and X* and 3.58 for *I and X*, whereas for *X and me* and *X and I* it was 1.80 and 1.72 words respectively.

Figure 1 approximately here.

![Figure 1. Average weight, measured in number of words, of the NP conjuncts (e.g. X in X and I) for each of the four subject-position coordinated 1sg pronoun variants.](image)
A t-test comparing the length of the conjuncts for *and X* variants versus *X and* variants shows the difference in conjunct weight based on its position in the coordinated NP to be significant (p < .001). This pattern suggests that end-weight plays an important role in subject position coordinated pronoun variation such that the pronoun-first variants are more likely to be used when the conjunct is relatively heavy.

**Information structure: discourse status of the other conjunct.** It was predicted that variation in coordinated pronoun usage would also reflect the general (and related) tendency of new information to follow old information, so that the usage of *me and X* and *I and X* variants would be correlated with second conjuncts that represent new information as shown in (13) and (14):

(13)  
If you were rebelling against the status quo in the '50s and '60s, no one was more uncool than Frank Sinatra. He was everything *I and my friends* rejected—the phony suave with fake self-control, the finger-popping slick. (Slate magazine)

(14)  
But yeah, when Hugo hit, I was at work that day, and I worked third shift at Coca-Cola. And, there was only, *me and one other girl* were the only two people in the whole building. (Face-to-face speech: Charlotte narratives)

The results show that this is indeed the case: the (non-first singular pronoun) conjuncts in these variants represented discourse-new information 72% (for *me and X*) and 84% (for *I and X*) of the time, as compared to 40% for *X and me* and 54% for *X and I*. Pearson’s chi-squared test shows the difference between all *X and* tokens and all *and X* tokens to be significant (p = .003).

Figure 2 approximately here.
Figure 2. Proportion of NP conjuncts (e.g. X in X and I) that represent information new to the discourse for each of the four subject-position coordinated 1sg pronoun variants.

The old-before-new pattern would also predict that X and I and X and me, in which the second position necessarily contains old information, would be relatively unlikely to convey new information in X. This is borne out in the case of X and me, which occurred with new information in less than half of the tokens, showing a “preference” for old information. X and I, however, occurred with new information more than half the time (54%). It may be that in this case, the old > new tendency, which influences the variation in general, is trumped by the “X and I” rule.

**Conclusion**

This study of variation in coordinated 1sg pronouns in subject position contributes to the research on coordinated pronouns in two ways. First, it provides data on coordinated pronouns in subject position, which serves as a complement to previous variationist work on object position (Angermeyer and Singler 2003; Boyland 2001). It shows that variation exists in subject position, although the prescriptively standard form is more prevalent than it is in object position. Second,
and most importantly, this paper has considered several underexplored factors that play a role in the distribution of the coordinated pronoun variants. The data on modality support the intuition that standard forms such as \( X \) and \( I \) (for subject position) are used with higher frequency in writing than in speaking, although the difference between the modalities in this data is quite small. The analysis of conjunct weight and information structure show that the gradient acceptability of coordinated pronoun variants and the usage of non-standard variants may result from more general linguistic processes involving constituent weight and/or the structuring of information in discourse.

Acknowledgements. I would like gratefully acknowledge the insights and advice of Brady Clark and Gregory Ward, along with audiences at Northwestern University, the Linguistic Society of America, and GURT 2007.
Notes

1. As noted by Grano (2006, 17) and my own review of prescriptive literature, this politeness ordering rule is often not explicit in usage manuals and grammars, but is merely implicit in their examples and exercises.

2. Grano (2006, 41) does briefly note both weight and information structure as factors that could possibly independently motivate the ordering constraint he proposes for coordinated pronouns.

3. Note that this sentence is ambiguous between distributive and collective readings. No attempt was made in this study to systematically correlate particular readings with particular coordinated pronoun variants. This is an important consideration for future research on coordinated pronoun variation.

4. The rarity of \( X \text{ and me} \) compared to \( me \text{ and } X \) is notable since \( X \text{ and me} \) does at least conform to the prescriptive “politeness” pattern, whereas \( me \text{ and } X \), which is more frequent, does not conform to prescriptive case or ordering. Quinn’s (2005) survey also showed that \( me \) tends to appear in first-position. Her analysis takes phonological considerations into account, arguing that \( me \) (along with \( he, she, we, they, \) and \( who \)) are less phonologically complex than their counterparts (\( I, him, her, us, them, whom \)) and that this phonological factor partially accounts for ordering preferences.

5. Wasow (1997) found that length (number of words), and complexity measures (number of nodes, number of phrasal nodes) were so highly correlated that one measure could not be isolated as the single best measure for constituent weight.

6. It was ensured that a minimum of 10 clauses preceding the coordinated NP were examined for each token. This 10-clause measure was extended to 100 lines of text (always significantly
more than 10 clauses) because 10 clauses did not always provide enough context to reveal that an entity had been explicitly evoked in the discourse.
References


Grano, Thomas. 2006. "me and her" meets "he and I": Case, person, and linear ordering in English coordinated pronouns. Honors Thesis, Linguistics Department, Stanford University.


Parrot, Jeffrey. 2007. Distributed morphological mechanisms of Labovian variation in morphosyntax. PhD, Linguistics, Georgetown University, Washington, DC.


