Linguistic analysis of quotations attributed to Daniel Wemp in New Yorker article

Douglas Biber

Northern Arizona University   1731 N. Fox Hill Rd.
Applied Linguistics Program   Flagstaff, AZ 86004
English Department   928-527-1630
Flagstaff, AZ 86011-6032
928-523-6156

1. Background: Corpus-based analyses of conversation and writing

The language of conversation is dramatically different from the language of academic writing. Some of these differences are obvious to all of us, such as contractions and incomplete sentences. However, many other grammatical differences are much more difficult to notice.

Over the last 25 years, a research approach has been developed for the empirical analysis of such grammatical characteristics. Referred to as 'corpus linguistics', the approach is based on the analysis of very large collections of natural texts from thousands of individual speakers and writers. Computer programs aid the analyses, which result in descriptions of the grammatical features that are especially frequent, features that are typical, and features that rarely occur. In addition, by comparing corpora with different kinds of texts, it is possible to contrast the grammatical characteristics that are usually found in conversation to those usually found in academic writing (or any other spoken or written varieties).

The most comprehensive grammatical description of English undertaken from this perspective is the 1,200-page *Longman Grammar of Spoken and Written English* (**LGSWE**; Biber et al. 1999). The research for that project is based on analysis of a very large corpus that represents four major varieties: conversation, fiction writing, newspaper writing, and academic writing. For example, the sub-corpus for conversation includes c. 6.4 million words, produced by thousands of speakers. The sub-corpus for academic writing includes 5.3 million words from 408 different texts. Computational / quantitative analyses of these corpora allow us to make strong generalizations about the grammatical characteristics that are frequent or rare in conversation, contrasted with the features that are frequent/rare in academic writing. The detailed research findings in the LGSWE can be applied to characterize the linguistic style of individual texts, to describe the extent to which the language of that text is typical of conversation or academic writing. (Other corpus-based studies have described the language used by academic professionals in university classroom teaching and advising.)

In the present report, these research findings are applied to evaluate quotations that were supposedly produced in speech by Daniel Wemp, cited in the article 'Vengeance is ours' (by Jared Diamond), published in the *New Yorker* (4/21/08). Section 2 below reports findings from a 3-way comparison: a quantitative analysis of the

grammatical characteristics of Diamond's quotes (i.e., the quotations attributed to Daniel Wemp as his spoken words in the 4/21/08 *New Yorker* article), a quantitative grammatical analysis of Daniel Wemp's actual speech (verbatim transcripts of speech produced by Daniel Wemp collected by Rhonda Roland Shearer), and the research findings from the LGSWE (large-scale corpus analysis of conversation and academic writing).

This linguistic analysis shows that the Diamond quotes (language attributed to Daniel Wemp in the 4/21/08 *New Yorker* article) are atypical of speech. Rather, these claimed quotes contain numerous grammatical constructions that are common in formal academic writing but very rarely used in normal speech. Further, those same grammatical constructions are not used in the verbatim transcripts of actual speech produced by Daniel Wemp (referred to as DW below).

Taken together, the linguistic analyses indicate that it is extremely unlikely that the *New Yorker* quotations are accurate verbatim representations of language that originated in speech. To put it simply, normal people do not talk using the grammatical structures represented in these quotations. However, these quotations do include several grammatical structures found commonly in academic writing, suggesting that the quotations were produced in writing rather than being transcribed from speech.

2. Grammatical characteristics of the quotations in the *New Yorker* article

Certain characteristics of conversation are easy to notice, and so almost any portrayal of speech will include these features. For example, even unskilled novelists are certain to include these stereotypical features in their fictional dialogue. Contractions are probably the most noticeable feature of speech (e.g., *it's, he's, I'm*), and the quotations in the *New Yorker* article are typical of speech (and fictional dialogue) in that they incorporate numerous contractions.

The use of simple coordinators (especially *and*) is also a salient characteristic of conversation, and the Diamond quotes incorporate frequent use of that feature. One major function of *and* is to connect clauses, and this use occurs frequently in both the Diamond quotes and in the actual transcribed speech of DW. For example:

If you die in a fight, you will be considered a hero, <u>and</u> people will remember you for a long time.
I have given all these story <u>and</u> those stories are very true <u>and</u> those names are not fake.

However, many other grammatical characteristics of a variety are less apparent to the casual reader or writer. It can be difficult to read a written transcript of spoken language, because it contains numerous grammatical constructions that we do not normally encounter in a written text. And conversely, readers and writers might not even notice the use of 'written' grammatical constructions in written dialogue, because we expect to find such features in a written text. This is where corpus analysis can be useful: to identify the grammatical features that are actually common or rare in conversation, especially features that would go unnoticed otherwise.

For example, the Diamond quotes frequently employ the coordinator *and* (as well as *but*) in two different ways: 1) to connect clauses, and 2) to connect two adjectives. As noted above, the first use is in fact very common in conversation. However, the second grammatical pattern is rare in actual conversation, although it is common in formal writing. Examples from the Diamond quotes are:

my father was felt to be too <u>old and weak</u>
<u>quick but correct</u> decisions
my <u>tall and handsome</u> uncle

This grammatical pattern is rare in both the actual transcribed speech of DW and in conversation generally. Thus, even in the use of the coordinator *and* (and *but*), the Diamond quotes are more similar to written language than to actual speech.

The noun phrase structures found in the Diamond quotes are especially atypical of normal speech: many of these structures are extremely rare in normal conversation, while they are quite common in academic writing. In normal conversation, a majority of noun phrases are realized as pronouns, and there are numerous pronouns used in both the Diamond quotes and in the actual speech of DW.

However, the Diamond quotes also include numerous examples of noun phrase structures that are extremely rare in conversation and rare (or unattested) in the actual transcribed speech of DW.   One structure of this type is noun phrases that have adjectives as modifiers, referred to as 'attributive adjectives'. These adjectives are very common in the *New Yorker* quotes; for example:

biological father
lower left leg
hot pieces of wood
public battle
unexpected words
experienced fighters
bare hands
constant suffering

Attributive adjectives are 3-4 times more common in formal writing than in conversation. These adjectives occur about 30 times per 1,000 words in the Diamond quotes. This density of adjectives is about 2-3 times as frequent as in normal conversation. (That rate of occurrence is similar to the normal rate in written fiction.) The density of adjectives is also considerably more common in the Diamond quotes than in the transcribed speech of DW.

A second unusual noun phrase structure found in the Diamond quotes is relative clauses that occur with a fronted preposition, as in:

a stone quarry <u>from which</u> the Ombal enemy was throwing stones
a night raid <u>in which</u> we sneak into an enemy village
each battle <u>in which</u> we succeeded in killing an Ombal

These structures are extremely rare in everyday conversation, although they are relatively common in formal writing. There are no examples in the formal statement of DW; one structure in the verbatim interview transcripts might be this structure, but it is difficult to interpret.

A third unusual noun phrase structure in the *New Yorker* quotes is the dense use of prepositional phrases as noun modifiers. Similar to the previous two structures, prepositional phrases as noun modifiers are extremely common in formal writing but rare in normal conversation. There are numerous examples of this structure in the *New Yorker* quotes; for example:

a strong young man <u>in</u> his prime
The original cause <u>of</u> the wars <u>between</u> the Handa and Ombal clans
real enemies <u>of</u> your target
grievances <u>of</u> their own
the mistake <u>of</u> hiring a man who actually does not consider your target to be his own
        enemy
feeling <u>of</u> anger
Both men and women <u>on</u> the other side
our endless cycles <u>of</u> revenge killings

Many of the noun phrases in the *New Yorker* quotes are especially surprising because they contain multiple modifiers (adjectives or nouns as pre-modifiers, and relative clauses or prepositional phrases as post-modifiers). It is very rare in conversation for a noun phrase to have two or more modifiers, while this pattern is relatively common in formal academic writing (with over 20% of noun phrases having multiple modifiers). It thus is highly noteworthy that such structures commonly occur in the *New Yorker* quotes. Many of the examples listed above are of this type:

a <u>stone</u> quarry <u>from which the Ombal enemy was throwing stones</u>
a <u>night</u> raid <u>in which we sneak into an enemy village</u>
a <u>strong young</u> man <u>in his prime</u>
a <u>spear</u> wound <u>on the back of your leg</u>
a <u>single outnumbered</u> enemy
<u>tall and handsome</u> *man*
<u>quick but correct</u> *decisions*
<u>endless</u> *cycles <u>of revenge killings</u>*
The <u>original</u> cause <u>of the wars between the Handa and Ombal clans</u>

Several of these noun phrases have a very complex structure, with multiple levels of embedding; for example:

The original <u>cause</u> [of the wars [between the Handa and Ombal clans]]
The <u>way</u> [that we come to understand things [in life]]
all the <u>stories</u> [that grandfathers tell their grandchildren] [about their relatives [who must be avenged]]

In general, such structures are extremely unusual in speech but common in writing. However, the examples given above are even more unlikely to occur in speech because they often occur in the subject position of a clause. There is a very strong

4

tendency in English conversation (and speech generally) for the subject noun phrase to be short and simple, usually a simple pronoun, as in

Yeah, <u>he</u> went.

Even when an utterance is longer, the subject noun phrase is almost always short, as in:

<u>I</u> *don't think I can go next week.*

More than 70-80% of the subject noun phrases in normal conversation are simple pronouns, and more than 90% of the subjects are a simple noun phrase with no modifiers.

In contrast, what we find in the *New Yorker* quotes is long, complex noun phrases as the grammatical subject, as in:

[<u>The original cause of the wars between the Handa and Ombal clans</u>] was a pig that ruined a garden.

[<u>The way that we come to understand things in life</u>] is by telling stories, like the stories I am telling you now, and like all the stories that grandfathers tell their grandchildren about their relatives who must be avenged.

Structures like these are found in academic writing, although they are not especially common; rather, even in writing, it is more common to use a relatively simple noun phrase as the grammatical subject. However, such structures are virtually unattested in normal speech, and so the Diamond quotes are highly unusual as representations of speech in this regard.

There are other features in the *New Yorker* quotes that are unusual, being much more typical of writing than speech. One of the obvious features is the repeated use of passive voice verbs in the quotes (e.g., *was felt, be considered, be remembered, be forgotten, be avenged*, etc.). Passive voice verbs are generally rare in conversation, but c. 1/3 of the verbs in academic writing are passives. The transcribed statement of DW also includes some passive verbs (several about the article '*being published'*), but not with the same density as the Diamond quotes.

A final noteworthy characteristic is the use of *to*-clauses. Apart from the semi-fixed expression *want to* (and to a lesser extent *would like to*), *to*-clauses are much more common in writing than in speech. However, the Diamond quotes have a very high density of these constructions. What makes this pattern especially noteworthy is the specialized types of *to*-clauses found in the *New Yorker* quotes. In particular, two of the constructions that occur repeatedly in the *New Yorker* quotes are structures that rarely occur in normal speech:

1) Noun + *to*-clause
    the <u>opportunity</u> [to see who really are the best marksmen]
    the necessary <u>experience</u> [to make quick but correct decisions]

2) 'extraposed' *to*-clauses controlled by an adjective
    it's not <u>acceptable</u> [to set fire to the hut]
    it's already <u>extremely dangerous</u> [for us to penetrate enemy territory]
    it will be <u>easy</u> [for the enemy to kill you]

Here again, these structures are relatively common in formal writing, but it is highly unusual to find such structures in normal speech.

3. Summary

In sum, the grammatical characteristics of the *New Yorker* quotes are much more typical of formal writing than of actual conversational speech. In fact, many of these grammatical characteristics are extremely rare in speech. This fact is all the more striking in that there is a whole suite of 'literate' features which appear commonly and pervasively throughout the *New Yorker* quotes.

It would be less noteworthy to find just one or two examples of 'literate' grammatical constructions in speech. Corpus research does not show that such features are impossible in conversation. However, corpus research does show that such features are rare and exceptional in normal conversation.

To indicate just how different the Diamond quotes are from the language of normal conversation, we can compare the rates of occurrence for these specialized grammatical features. All rates are computed for the same basis: a rate per 1-million words of text. For actual conversation, the rates are computed from analysis of a 5-million word corpus (with most specific findings taken from the *Longman Grammar of Spoken and Written English*;LGSWE). The rates for the Diamond quotes are based on the c. 1,500-words of quotes in the *NewYorker* article. The following table compares the rates of occurrence for several of the features discussed in Section 2 above.

Grammatical feature	Rate in actual conversation	Rate in Diamond quotes	Comparison of the two
Attributive adjectives (e.g., *original, biological*)	c. 15,000	c. 30,000	**2 times more frequent in Diamond quotes**
Preposition *of*	c. 12,000	c. 34,000	**3 times more frequent in Diamond quotes**
Noun post-modifier complexes (e.g., *The original <u>cause</u> [of the wars [between the Handa and Ombal clans]])*	c. 500	c. 3,000	**6 times more frequent in Diamond quotes**
Noun phrases with both pre-modifiers and post-modifiers (e.g., *a <u>stone</u> quarry <u>from which the Ombal enemy was throwing stones</u>*)	c. 500	c. 4,000	**8 times more frequent in Diamond quotes**
'extraposed' *to*-clauses controlled by an adjective (e.g., *it's not <u>acceptable</u> [to set fire to the hut]*)	c. 100	c. 2,000	**20 times more frequent in Diamond quotes**

Noun + *to*-clause (e.g., *the opportunity [to see who really are the best marksmen]*)	c. 50	c. 1,200	**25 times more frequent in Diamond quotes**
Adjective *and/but* adjective (e.g., *tall and handsome*)	c. 20	c. 2,000	**100 times more frequent in Diamond quotes**
Preposition + Relative pronoun (e.g., *each battle in which we succeeded in killing an Ombal*)	c. 20	c. 2,000	**100 times more frequent in Diamond quotes**

These comparisons show the magnitude of the discrepancies between the grammatical style of normal conversation contrasted with the grammatical style of the Diamond quotes. To find one of these grammatical features in a normal conversation is noteworthy. To find repeated use of this large constellation of features in actual spoken discourse, some of them occurring c. 100 times more often than in normal conversation, is extremely unlikely. In contrast, these are all features that are typical of academic writing, suggesting that they have their origin in writing rather than actual speech.

Other corpus studies (e.g., the book *University Language*; Biber 2006) have shown that these same features are rare and exceptional in even academic speech, including university lectures. In contrast, what we find in the Diamond quotes is the pervasive use of a suite of grammatical constructions, which are all rare in conversation but common in formal writing. This constellation of grammatical characteristics is also strikingly different from the grammatical style of the verbatim transcripts of speech produced by DW. In sum, the analysis strongly indicates that the Diamond quotes are much more like discourse that was produced in writing, reflecting the typical grammatical features of formal academic prose, rather than verbatim representations of language that was produced in speech.

List of Quotations attributed to Daniel Wemp:

1. " 'Soll did have a son, but he was only six years old at the time of his father's death, much too young to organize the revenge,' Daniel said. 'On the other hand, my father was felt to be too old and weak by then; the avenger should be a strong young man in his prime. So I was the one who became expected to avenge Soll.' "

2. " 'The original cause of the wars between the Handa and Ombal clans was a pig that ruined a garden.' "

3. " 'If you die in a fight, you will be considered a hero, and people will remember you for a long time,' he said. 'But if you die of a disease you will be remembered for only a day or a few weeks, and then you will be forgotten.'… He likened Nipa people to 'light elephants': 'They remember what happened thirty years ago, and their words continue to float in the air. The way that we come to understand things in life is by telling stories, like the stories I am telling you now, and like all the

stories that grandfathers tell their grandchildren about their relatives who must be avenged. We also come to understand things in life by fighting on the battlefield along with our fellow-clansmen and allies.' "

4. " 'I was advancing in battle with my biological father, who was holding a shield to protect me, while I myself held the weapons,' he told me. 'As my father and I went up a hill towards a stone quarry from which the Ombal enemy was throwing stones as well as spears, a stone hit my father on his leg. So I took the shield to protect my father, and I told him to go faster. That was when I was left unprotected, and an Ombal spear struck me on the back of my lower left leg.' ………. 'If, in a fight, you receive a wound on your forehead, then you are considered to have done well, but if you only have a spear wound on the back of your leg, like this one of mine, then you are viewed as not having fought well.' "

5. " 'You have to make sure that the men that you hire as paid killers or allies are real enemies of your target, bearing grievances of their own from years ago,' Daniel said. 'If you make the mistake of hiring a man who actually does not consider your target to be his own enemy, he may seize the chance to kill you, then go to your enemies and claim a reward.' "

6. " 'to provide ladies for the warriors when they were homesick.' "

7. " 'If we had found that a woman married into our clan was squealing, we would have tied her up and burned her with hot wires and hot pieces of wood. That was our plan, but in fact we never found any woman married into our clan who squealed; they all remained loyal to us, not to their blood relatives.' "

8. " 'By killing Isum or arranging for Isum's killing,' Daniel explained, 'I would lose Isum as an uncle, but that would be worth it, because I would gain my revenge.' "

9. " 'Public battles are open not just to experienced fighters but also to new trainees, new allies hired to come and gain confidence, and fun-seekers. In a public battle, the fight-owners have the opportunity to see who really are the best marksmen, with the necessary experience to make quick but correct decisions.'…….. 'That requires nerve, judgment, and presence of mind, to select the right target, and not to panic and shoot the first man who moves into a shootable position,' he said.

'Boys and young men are prone to make such mistakes and hence are excluded from the stealth parties.' "

10. "He said, 'In a night raid in which we sneak into an enemy village and surround the hut of a targeted enemy individual, we can tear down the hut to force the enemy to come out so that we can kill him. But it's not acceptable to set fire to the hut and burn him to death.' I then asked, 'Is it acceptable for six of you surrounding a hut to attack and kill a single outnumbered enemy?' Daniel answered, 'Yes, that's considered fair, because it's already extremely dangerous for us to penetrate enemy territory, where we are greatly outnumbered.' "

11. "He said, 'When you hear that your own brother has been killed in a fight, then you have bad feelings, you feel anger inside yourself, you become aggressive, you cannot think clearly, and you want to tear someone apart with your bare hands.' He went on, 'But, if you fight when that feeling of anger is on top of your mind, you'll expose yourself, and it will be easy for the enemy to kill you.'……. 'Both men and women on the other side sing out unexpected words, which you can hear from far away and which make you feel badly. They'll sing, 'We killed your brother, and he was a coward.' They'll sing war songs to bring up old memories in you: 'I was there on that day of battle, I tried to kill you then, we should have killed you then, you were our target and we missed, but now we won't miss.' Those words make you want to go straight to the attack and to kill the other side, but then you'll end up being killed yourself, because you are not thinking clearly and you're incautious.' "

12. " 'Isum was in the public fight, with his bow and arrow ready for a long-range battle, and he was shooting and dodging arrows in the open. He was concentrating on that public fight, looking at our men far away in the open, and he wasn't prepared for our attack from behind and nearby by one of our hidden parties. It was our group that had gone down along the side of the river that got him. Only one arrow hit Isum, but it was a bamboo arrow, flat and sharp as a knife, and it cut his spinal cord. That's even better than killing him, because he's now still alive today, eleven years later, paralyzed in a wheelchair, and maybe he'll live for another ten years. People will see his constant suffering. Isum may be around for a long time, for people to see his suffering, and to be reminded that this happened to him as proper vengeance for his having killed my uncle Soll.' "

13. " 'I felt that it was a matter of 'kill or else die by suicide.' I was prepared to die myself in that fight. I knew that, if I did die then, I would be considered a hero and would be remembered. If I had personally seen the arrow go into Isum, I would have felt emotional relief then. Unfortunately, I wasn't actually there to see it, but, when I heard that Isum had been paralyzed, I thought, I have everything, I feel as if

I am developing wings, I feel as if I am about to fly off, and I am very happy. After that battle, just as after each battle in which we succeeded in killing an Ombal, we danced and celebrated and slaughtered pigs. When you fight with thinking and finally succeed, you feel good and relieved. The revenge relieves you; now it can be your turn to help someone else get his own revenge.' "

14. "Once he said to me, 'I admit that the New Guinea Highland way to solve the problem posed by a killing isn't good. Our way disturbs our day-to-day life; we won't be comfortable for the rest of our lives; we are always in effect living on the battlefield; and those feelings go on and on in us. The Western way, of letting the government settle disputes by means of the legal system, is a better way. But we could never have arrived at it by ourselves: we were trapped in our endless cycles of revenge killings.' "

15. " 'If I had let the police do it, I wouldn't have felt satisfaction,' he replied. 'I wanted to obtain vengeance myself, even if it were to cost me my own life. I had to ask myself, how could I live through my anger over Soll's death for the rest of my life? The answer was that the best way to deal with my anger was to exact the vengeance myself.' "

16. " 'Now, when I visit an Ombal village to play basketball, and Isum comes to watch the game in his wheelchair, I feel sorry for him,' he said. 'Occasionally, I go over to Isum, shake his hand, and tell him, 'I feel sorry for you.' But people see Isum. They know that he will be suffering all the rest of his life for having killed Soll. People remember that Isum used to be a tall and handsome man, destined to be a future leader. But so was my uncle Soll. By getting Isum paralyzed, I gained appropriate revenge for the killing of my tall and handsome uncle, who had been very good to me, and who would have become a leader.' "