"NEWSPEAK" IN GEORGE ORWELL'S
NINETEEN EIGHTY-FOUR

William Steven Dodd
University of Exeter

Bernard Crick, his biographer, writes[1] of George Orwell that two adjectival constructions have been based on his name "Orwell-like" and "Orwellian". The first of these, he says, conveys the idea of a love for England and its language; the second, a very different image. This second sense he directly attributes to the novel "Nineteen Eighty-Four"[2]. What is meant may be judged by the reaction of the American National Council of Teachers of English, who in 1972 set up a Committee on Public Doublespeak to combat the use of officialese, and in 1974[3] debated the creation of an "Orwellian Award" for the coiners of the most grandiose and jargonesque expressions to hide simple meanings. George Orwell wrote a number of articles attacking officialese and sloppy use of language, in particular, "Propaganda and Demotic Speech"[4] and "Politics and the English Language"[5], but also many other briefer items.[6] Nonetheless, his name has come to be associated with a distorted form of language of the very sort he criticised. This must be seen as a result exclusively of his last great novel, "Nineteen Eighty-Four", with its invented language "Newspeak" and other elements of linguistic alienation. This article is an attempt to investigate these features by means of a stylometric analysis. Apparently the only major work using a similar approach that concerns this author is the general survey of Orwell's essays prepared by Ringbom[7] —though this was compiled by manual methods, rather than computer aided like the present article.

The book[8] consists of four sections, amounting to just under 104,000 words, defined rather crudely, but in the fashion normally accepted in automatic treatment of text, as a group of letters coming between two punctuation marks or spaces. It is divided into four sections of varying length. The first section, entitled "Part One", of eight chapters, is 34,329 words long. The second, "Part Two", runs to 40,530 words and is nine chapters in lenght. "Part Three", the third section, is composed of six chapters and amounts to 24,328 words. The fourth and final section is headed "Appendix" and subtitled "The Principles of Newspeak". It is not divided into chapters, is only 4,374 words in extent and, as its title suggests, is not intrinsically linked with the main body of the novel, occurring, as it does, after the words "THE END", which appear as the last line of "Part Three".

Literary statistics frequently make use of the concepts of "type", as a unique combination of letters, of which one or more examples may be found in a work studied, and of "token", as each of these occurrences as words on a page. The ratio

between types and tokens for each of the sections of the work, expressed to one decimal place, is 6.7, 7.3, and 6.8, with the appendix having a ratio of 3.6. The type-token ratio for the whole text has not been calculated; if it were, it would undoubtedly be rather higher even than the 7.3 of "Part two", because of the overlap of types between sections.

At first, the impression given by the book is one of considerable linguistic alienation, so that it would initially appear that the amount of Newspeak in the text is very great. In fact, over the whole text the total number of Newspeak words is 425, so that the ratio of Newspeak tokens to all tokens is roughly 1 to 244. The number of Newspeak types is only 89. If the Appendix, which, as its title suggests, concerns itself exclusively with Newspeak, is left out of consideration, the total of Newspeak types is reduced to 40, represented by 281 tokens. This means that in the novel as a novel there is only one Newspeak word in every 353 roughly speaking: barely one and a half Newspeak words per page averaged over the 234 pages of text of the Penguin edition[9]. Even this gives a false view of the frequency with which Newspeak words are met, because all but a few are bunched together in several short passages rich in such vocabulary: the official memoranda listing the day's tasks received by Winston Smith, the protagonist, in chapter 4 of Part One (pp. 34, 39); the conversation between Smith and his friend Symes, an expert on Newspeak, in chapter 5 of the same part (pp. 44-47); the memorandum dictated by the party official O'Brien in the eighth chapter of Part Two (p. 137). Further, the text of the book of political science, "The Theory and Practice of Oligarchal Collectivism", by the fictional Emmanuel Goldstein, that Smith reads in the final chapter of Part Two (pp. 150-161, 162-173), is noticeably richer in Newspeak than the text of the novel, but even in it there are only two pages with more than three Newspeak words per page: pp. 169 and 171, with fewer than thirty between them both. Of the words outside these "hot spots", several occur in footnotes or parenthetical comments such as: "... to wear an improper expression on your face... was itself a punishable offence. There was even a word for it in Newspeak: *facecrime*, it was called. (p. 53) or "... artificial insemination (*artsem* it was called in Newspeak)." (p. 56).

Apart from this, of the forty types occurring in the main body of the novel, several could be coalesced by lemmatization: *telescreen* and *telescreens*, *unperson* and *unpersons*, *speakwrite* and *speakwrites*, with possibly in addition the participle *speakwritten*. The last word is not accurate Newspeak, if the prescriptions of the Appendix are observed —it should really be *speakwrited*. If this lemmatization of singulars and plurals is carried out, the total number of types having more than a dozen occurrences would be reduced to only five. Many of the types are represented by a single token alone, as can be seen from the tabular presentation below.

Number of tokens per type	Number of types such	(Types)
1	21	
2	8	
3	3	
6	2	
13	1	speakwrite
15	1	telescreens
26	1	Ingsoc
30	1	doublethink
46	1	Newspeak
93	1	telescreen

It seems clear that the impression of linguistic alienation that strikes the reader is not due in any extensive way to the concentration of Newspeak words in the text, at least in that part of the book, the main novel, where Newspeak is not under detailed discussion as a topic.

It is highly instructive to compare this paucity of occurrence of Newspeak in "Nineteen Eighty-Four" with the use of another imaginary language in another futuristic novel, Anthony Burgess's "A Clockwork Orange". This book[10] is barely half the length of "Nineteen Eighty-Four", yet the glossary of the invented teenagerspeak "Nadsat" (a name based on the Russian numeral ending corresponding to "-teen") at the end of the book includes in excess of 200 entries, with even this not by any means a complete list of all the non-standard English in use in the work. A single page taken at random (p. 88 of the Penguin edition) shows thirty types in use, represented by forty-three tokens: *chellovech, creeching, crunk, dratsing* (twice), *fillying, gromky, gulliver* (meaning "head"; thrice), *heighth* (twice), *horrorshow* (an adjective meaning "good"), *krovvy* (twice), *litso* (thrice), *malchiks* (twice), *malenky, minootas, nadsat, platties, poogly, ptitsa, raz, rooker* (meaning "hand"), *rot* (meaning "mouth"), *skorry* (twice), *smeck, smecking* (thrice), *starry* (meaning "old", especially of people), *tolchock, tolchocking, ultra-violent* (used as a noun), *vaysay, viddied* (thrice). This count is no more than average for Burgess's novel, but there is not one single page of "Nineteen Eighty-Four" with such a concentration of abnormal words, even in the Appendix.

Orwell clearly avoids any extensive use of the articial language which he invented, perhaps unsurprisingly, as it incorporates many of the features of English that he did not like. For instance, there is the inexorable destruction of words in the ever more reduced vocabularies of successive versions of Newspeak, as explained by the expert Symes: "Don't you see that the whole aim of Newspeak is to narrow the range of thought? In the end we shall make thoughtcrime literally impossible, because there will be no words in which to express it".[11] This is diametrically opossed to Orwell's own proposal[12] for the creation of a group of "several thousands of people" to devote their time and energy to the creation of new words,

in order to allow the expression of "things an ordinary man has to keep locked up because there are no words to express them". On a more specific point, Newspeak's extensive use of the negative prefix unparallels Orwell's sentence ridiculing of the "not un-" construction: "A not unblack dog was chasing a not unsmall rabbit across a not ungreen field."[13]

How, then, *does* Orwell achieve the strong feeling of linguistic alienation which is undoubtedly present? There are several linguistic strands that lie behind it, and Newspeak is probably less significant than at least two of them.

One strand which is marginally less important in numerical terms than Newspeak is the employment of metric measurements, dollar-cent currency and the twenty-four hour clock. The clock system is used several dozen times, starting at the very first sentence, where "... the clocks were striking thirteen." The measurements and currency between them occur in over seventy places, but are more repetitious than the Newspeak, since, if plurals and singulars are lemmatized together, they involve only nine types, none of which has only one token associated, while most of the Newspeak types are "oncers", represented by single tokens, as previously mentioned. It is true that this strand is rapidly losing its power to alienate, and the present generation will be the last to feel its effects naturally, because pounds-shillings-pence currency has now been abolished, being replaced, what is more, by dollar-cent in several of the countries where it once was used. Measurements in metres, litres, and grammes are no longer unfamiliar, while yard, pint, and pound, even in the United States, are destined to become the oddities that their metric versions were in 1949 when "Nineteen Eighty-Four" was published[14]. Apparently, the New York copy-editors initially produced proofs in which all the metric measurements were converted back to imperial[15], so strange did they seem in that period.

Another strand of an importance equal to or greater than that of the Newspeak words is that of names. The three imaginary super-powers of the book are *Eastasia, Eurasia,* and *Oceania,* and these names plus forms derived from them occur more than 150 times in the novel, though absent from the Appendix. The name *Airstrip One* for Britain adds a further four occurrences. The frequent repetition of the first three names is far more striking than the scattered Newspeak "oncers", and is to be compared with the Newspeak terms *Newspeak* and *doublethink,* the second and third most frequently used of all.

The most weighty source of linguistic alienation, however, must lie in the group of unusual collocations of words, and unusually frequent use of words. In the first of these two categories would come phrases such as *floating fortress, Junior Anti-Sex League, Three-Year Plan, memory hole, rocket bomb, Youth League,* and the constant use of the title *Comrade* with people's names. *Hate Week* and *Two Minutes' Hate* between them occur some three dozen times in the first two parts of the book. *Victory,* in collocations such as *Victory Mansions, Victory Square, Victory cigarettes*[16], *Victory gin,* has roughly the same frequency. In the second category, an obvius example is the word *Party,* which alone or in the groupings *Inner Party* and *Outer Party,* refers to the ruling clique in Oceania, and has a total of 376

occurrences in the main text, all but 52 in the first two parts. This is more than the total of Newspeak words in the novel proper. Likewise, *prole*, a word not invented specially for "Nineteen Eighty-Four", but uncommon elsewhere[17], totals 55 appearances, which are restricted to the first two parts of the book, 35 in Part One, and 20 in Part Two.

This tendency for elements producing linguistic alienation to become less common as the book proceeds into its final part is shared by the Newspeak words and almost all other factors mentioned here: names, measurements, clock. The overall distribution of occurrences would yield a time series graph with a J curve (so called because it looks rather like a sans serif capital "J" tilted ninety degrees to the right), if plotted in terms of words on the Y scale and progression through the book on the X scale: a relatively high proportion at the start, with a downwards trend, and a low proportion at the end.

The best-known collocation of all, which, along with the lexical items *Newspeak, doublethink, unperson, prolefeed*, and the title of the book itself, has passed into current English, is *Big Brother*, the name of Oceania's semi-mythical dictator. Together with its abbreviation *B.-B.*, it can be instanced from some eighty places in the main text, and once in the Appendix. Ironically, his description in the book, "... a man of about forty-five, with a heavy black moustache and ruggedly handsome features", fits Orwell himself at the time of his writing the novel. This is not in accord with the view of Bernard Crick[18] that it is a description of Joseph Stalin, or that of William Steinhoff[19] that it is the actor who played Stalin in a widely distributed 1943 film[20]. However, Stalin was nearly seventy at the time the book was written, and over sixty when he first came to be portrayed in Allied wartime propaganda, though it must be admitted that particularly favourable photographs of him were always used.

While discussing the question of the origin of elements in "Nineteen Eighty-Four", it may be worth considering what exactly Orwell is trying to satirise in Newspeak. Bernard Crick considers[21] that it is a parody of C.K. Ogden's Basic English, though he also mentions Esperanto and journalists' "cablese", pointing out that "Newspeak" and "news-speak" are almost identical in sound, and quoting an extract from Eugene Lyons' "Assignment in Utopia", a book reviewed by Orwell, which reads very like the official memoranda sent to Winston Smith. It is true that an aunt of Orwell's, Nellie Limouzin, lived with a Breton, Eugène Adam who was a fanatical Esperantist, and Orwell's time in Paris is likely to have brought him into considerable contact with them, as also with the proprietors of the Hampstead bookshop where he later worked for a while[22]. He seems to have been put off any initial sympathy by the extreme position held by such people, who had an ideology rather than a language. Basic, on the other hand, he does not seem to have disapproved of in the same way, for he corresponded with C.K. Ogden, did not discourage his friend William Empson's support of it, wrote approvingly of it on several occasions[23] and saw it as "a sort of corrective to the oratory of statesmen and publicists", in which "you cannot make a meaningless statement without its being apparent that it is meaningless"[24]. This in itself is enough to refute Bernard

Crick's idea, let alone the narrower view of the origin of Newspeak held by Howard Fink[25]. W. F. Bolton suggests several other strands as being incorporated in the language, including the concept of thought control by a Dictionary Bureau, probably from H. G. Wells' "The Shape of Things to Come"; the reductionist nature of the language maybe from Zamiatin's "We"; and the concept of absolute vacuity of official writing, which could arguably be seen as mentioned in Aldous Huxley's "Brave New World"[26].

Finally, what of the technical skill of Orwell as a linguist in the production of "Newspeak"? Bolton says[27] "Orwell knew no more about language... than the average Briton of his time and class might have known, and perhaps a trifle less" and points out that "even the celebrated 'Appendix' on Newspeak in *Nineteen Eighty-Four*, though it uses some technical terms of linguistic descriptions, uses none that the intended non-specialist reader would fail to grasp". This is probably fair comment, and, indeed, Bolton's own work shows[28] that, although Orwell claims that all the Newspeak words of the B vocabulary are compounds, many are blends, like *Ficdep* or *Minilove*. Much of the reduction of vocabulary of Newspeak is also shown to rely on the process of conversion, whereby a word is re-used with a different grammatical category from its norm, such as "boot", used as a verb "to boot", for "to kick". However, Orwell was not setting out in any serious way to produce an artificial language, and what is most important is the extremely skilled fashion in which he satirises a number of undesirable features of language while still avoiding anything but minimal use of the features he abhors and attacks in a number of his publications over a long period.

Notes

1. Bernard Crick: "George Orwell: A life" (Penguin Books, Harmondsworth, 1982 edition). The mentions of the two forms is on page 580.
2. *Op. cit.*, p. 23.
3. As reported in "Newsweek" of 25 March, 1974.
4. First published in "Persuasion", Summer 1944. Reprinted in "The Collected Essays, Journalism and Letters of George Orwell" edited by Sonia Orwell and Ian Angus (Secker and Warburg, London, 1968) as item III-35. (Hereafter, this set of books will be referred to as "The Collected Essays &c.").
5. First published in "Horizon", April 1946. Reprinted in "The Collected Essays &c." as item IV-38.

6. For example, the section on "The English Language" in "The English People" which was commissioned by Collins for their "Britain in Pictures" series in 1944, but actually published in 1947, and included in "The Collected Essays & c.") as item III—1; the "As I Please" column from "Tribune" of 17 March 1944 ("The Collected Essays & c." III-27); part of the "As I Please" column from "Tribune" of 18 August 1944 ("The Collected Essays &c." III—62); the Editorial in "Polemic" no. 3, May 1946 ("The Collected Essays &c." IV-45).

7. Hakan Ringbom "George Orwell as Essayist: A Stylistic Study". Acta Academiae Aboensis, Series A, Humaniora, 44.2, 1973. This looks at uses of certain series of synonyms, at punctuation and sentence length, and at abnormally frequent use of some words, but does not touch on "Newspeak" or "Nineteen Eighty-Four".

8. "Nineteen Eighty-Four" was first published by Secker and Warburg in the summer of 1949, having been completed in December of the previous year, and was Orwells's last novel.

9. The Penguin edition of the book was first published in 1954 and has subsequently been reprinted many times. All page references are to this edition.

10. Anthony Burgess "A Clockwork Orange" (Penguin Books, Harmondsworth, 1972), was originally published by Heinemann in 1962.

11. This comes from the conversation on p. 45 of the Penguin edition of "Nineteen Eighty-Four".

12. This proposal is contained in an essay entitled "New Words" that was published for the first time as item II-1 of "The Collected Essays &c.".

13. This critical comment is in "Politics and the English Language", item IV-38 of "The Collected Essays &c.", as an author's footnote on p. 138.

14. Orwell himself wrote "Another thing I am against in advance... is ... scrapping of our present system of weights and measures" in his "As I Please" column of "Tribune" 14 March 1947 (item IV-78 of "The Collected Essays &c.").

15. Bernard Crick mentions this in his annotation number 8 to his edition: George Orwell, "Nineteen Eighty-Four, with a Critical Introduction and Annotations by Bernard Crick" (Clarendon Press, Oxford, 1984), and in his biography, *op. cit.,* pp. 554-555).

16. Bernard Crick mentions that there actually were Victory Cigarettes made in Lebanon and issued to Eighth Army troops in North Africa during the war, but adds that Orwell almost certainly invented the term without any knowledge of this make, in annotation number 13 to his edition.

17. The Oxford English Dictionary Supplement finds its first citation of this form in a George Bernard Shaw letter of 1887. The word was also used by Jack London in his "The Iron Heel" of 1908 and by James Joyce, both of whom were authors Orwell had read and written about.

18. Bernard Crick in his annotation number 3 to his edition says: "This is the 'Good Old Uncle Joe' face of Stalin".

19. Quoted by Bernard Crick in his annotation number 3 to this edition. The film was "Mission to Moscow", and Steinhoff's comments on it are in his "George Orwell and the Origins of 1984" (University of Michigan, Ann Arbor, 1975) p. 186.

20. Orwell himself certainly had knowledge of the film, for he mentions details of it in the Editorial to "Polemic" previously cited. (Item IV-45 of "The Collected Essays &c.).

21. In his annotation number 9 to his edition.

22. See Crick's biography, pp. 189-90, and comments in W.F. Bolton "The Language of 1984: Orwell's English and Ours" (Basil Blackwell/André Deutsch, Oxford, 1984) on pp. 20-21.

23. For example, as part of his "As I Please" column for "Tribune" 28 January 1944 (Item III-17 of "The collected Essays &c.").
24. "As I Please" column, "Tribune" 18 August 1944. Item III-62 of "The Collected Essays &c.", p. 210.
25. Howard Fink "Newspeak: The Epitome of Parody Techniques in *Nineteen Eighty-Four*" in "Critical Survey", 5 (1971), pp. 155-163.
26. In "The Language of 1984", p. 152, Bolton puts forward these possible sources.
27. In "The Language of 1984", p. 21.
28. In "The Language of 1984", p. 60.