

The Stylemetry of Cliché Density and Character in James Joyce's *Ulysses*

Robert Byrnes

Department of Rhetoric, Writing and Communications
University of Winnipeg, MB R3B 2E9 – Canada

Abstract

I've demonstrated, in an article that appeared in JADT 2010, that the "Eumaeus" episode of James Joyce's *Ulysses* is almost entirely composed of clichés, with the Odysseus character, Leopold Bloom, producing nearly 2600 of them, as my database shows. But what about the other main characters, like Bloom's wife Molly, in her famous going-to-sleep monologue after midnight, when she reviews her amorous life and recent infidelity? What of the young university graduate Stephen Daedalus, lyric and aesthetic in sensibility? What about Gerty McDowell, the Nausikaa figure in the novel, the young woman who reads too many romantic novels and has acquired her phraseology therein? This textometric and stylometric study establishes cliché densities for the stream-of-consciousness of each of these main characters, including Bloom, with startling results—Bloom and Gerty MacDowell speak from large repertoires of cliché, inviting us to read them naturalistically, which is to say comically and ironically, in this comic novel. Molly uses them far less, and emerges as a realistic character, one of substance and individuated voice compared to Bloom and Gerty. Stephen alone uses virtually none whatsoever. His language is almost entirely original, often unprecedented, and in accord with his project to escape the paralyzing culture, linguistic and otherwise, of Dublin. This quantitative approach lets us see into the heart of Joyce's art, and into one of the sources for its extraordinary appeal.

Keywords: stylometry, cliché, character, James Joyce, *Ulysses*

1. Introduction

In James Joyce's *Ulysses*, as I reported in a 2010 JADT article (Byrnes: 289-295), there's a 72-page episode ("Eumaeus") that's written almost entirely in clichés—an average of 36 per page, with one page as high as 68. Joyce was portraying the Odysseus figure, Leopold Bloom, as an uneducated man of commonplace information and limited vocabulary who speaks entirely within a repertoire of cliché. This was in accord with his ironic status as a counterpoint to Homer's Greek hero. Bloom was a diminished modern city-dweller of modest intellect, pacifist tendencies, and perverse sexual proclivities. His linguistic platitudinousness was part of his reduction, his diminishment in the ironic design. Odysseus had always been adept verbally, full of tricks and tropes and oratorical success. He spoke from the moral and rhetorical center of his heroic culture. But Bloom is an ironic hero, an antihero. He speaks from the rhetorically diminished center of modern Dublin culture in 1904, rendering its verbal and intellectual degeneration in banalities, recycled platitudes and fossilized locutions.

Of course Joyce critics had understood there were more clichés in “Eumaeus” than a normal speaker would spontaneously produce, and that Bloom’s stature was lessened accordingly, but no one had gotten around to actually counting Bloom’s clichés. The boldest estimate had been that there were about 700 altogether in the entire episode (Raleigh, 1981)—about 10 per page in the Penguin edition of *Ulysses*. I’d decided to do a quantitative study, not simply to produce a more or less definitive number, but to see if a pure quantitative investigation could itself lend us critical insight into Joyce’s intentions. It turned out that there are at least 2,561 formula phrases in “Eumaeus,” nearly 2000 more than the highest critical estimate. And the quantitative difference was a qualitative one as well. We Joyce critics had not only underestimated the number of clichés, we’d underestimated Joyce’s intentions. “Eumaeus” was in fact a bravura stylistic exercise in cliché, to begin with—and an X-ray, a cameo portrait of Bloom in his characterological essence: *homo commonplace*, the antitype and anti-hero to Homer’s rhetorically gifted Odysseus. I’d proceeded this far in the first essay.

But what about the other main characters in *Ulysses*, especially those with long stream-of-consciousness monologues, like Stephen Daedalus, the Telemachus figure in the Homeric parallels, and therefore Bloom’s figurative son? Stephen and Bloom move toward each other throughout the novel, finally meeting, talking, and parting, over the course of the three longest episodes, without any such “atonement” as occurred between Odysseus and Telemachus in the *Odyssey*. What about Gerty McDowell, the fashion-mad, romance-obsessed Nausikaa whose mind we listen to for 30 pages? What about Molly Bloom, Leopold’s wife, whose celebrated 62-page monologue ends the book as she reviews her married life with Bloom and contemplates her afternoon infidelity? Are they equally afflicted with cliché? But first a word about method.

2. Method

To identify most clichés, I began with my own native-speaker intuition, relying in part on my experience with Victorian literature. I followed up with Google searches and checked books of English idioms to verify the slightly outdated ones. Most are still current, or recognizable in fiction from 50 or 100 years ago, and they can hardly be mistaken. And I’ve counted clichés in direct speech (dialogue), stream-of-consciousness, and free indirect discourse. Joyce uses this last technique as much as he does stream-of-consciousness. In free indirect discourse, *Ulysses*’ narrator tells us what characters are thinking, but does so in their own idioms, letting the characters’ voices infuse the narration. For example, “*If I had only received the benefit of a good education, I might have held my own*” becomes “*Had she only received the benefit of a good education, Gerty MacDowell might easily have held her own.*” A narrator speaks, but borrows all his clichés from the character.

For the “Eumaeus” episode (and the 2010 JADT article), I had created a database of all Bloom’s clichés, and was able to generate exact numbers, for example the fact that Bloom averages 36 clichés per page (CPP) over the whole episode. For the additional four characters treated here, I’ve analyzed typical pages from each of the monologues, and report a range of CPPs to compare with the “Eumaeus” 36. For the passages I’ve looked at more closely in what follows, I’ve created an additional measurement, that of Cliché Density: the percentage of words in a selected passage of discourse that are clichés. I count all the words in a section of text, then count all the words that make part of a cliché, and calculate the Cliché Density Index (CDI).

The CDIs are interesting to measure when we study local effects, but the CPPs are the backbone of the study, and I relate them to Joyce's characterization, whether naturalistic or realistic, depending on the character.

3. What's a Cliché?

Technically, it's a combination of words that "co-occur" more often in discourse than is usual for any two randomly selected lexical items. Practically, it's a little phrase that the brain stores as a single unit, as if it were a single word; this reduces retrieval time, and thereby makes it easier for a speaker to maintain fluency, and easier for a listener to comprehend (Wray, 2002). Clichés run along a spectrum from transparent, in which the words add up to the meaning of the whole expression (for example, "highly advisable"), to the opaque, which can't be understood by adding the meanings of the component words (for example, "cooked his goose"). Clichés, or as they are now often called, formula phrases or phrasemes, have latterly spawned a subfield of linguistics called phraseology (Cowie, 1994). Phraseologists have identified a great many phraseme types, in a great many competing taxonomies, including items like collocations, adverbial units, compound units, complex units, sentence stems, routine formulas, sentence idioms and speech acts. No grammatical or semantic commonalities have been discovered among clichés; rather, they share a mnemonic characteristic: they are stored and retrieved holistically, as a single unit, rather than being generated analytically with grammatical rules, words, and morphemes. Of course they *could* be generated and understood analytically, but ordinarily they are not.

4. The "Eumaeus" Baseline

This comparative study needs a baseline, and "Eumaeus" provides one, since it manifests more clichés per page, on average, than any other monologue in the text. It's useful for another reason, in that Bloom's speech is sprinkled with clichés from virtually every subject domain. He doesn't specialize, but promiscuously raises and drops topics of every kind, strewing his remarks with the most miscellaneous of commonplaces. He's a true *homme moyen sensuel*, and an indiscriminate packrat of formula phrases. Here's a sample passage (I've italicized the clichés):

En route, to his taciturn, and, not to put too fine a point on it, not yet perfectly sober companion, Mr Bloom, who at all events, was in complete possession of his faculties, never more so, in fact disgustingly sober, spoke a word of caution re the dangers of nighttown, women of ill fame and swell mobsmen, which, barely permissible once in a while, though not as a habitual practice, was of the nature of a regular deathtrap for young fellows of his age particularly if they had acquired drinking habits under the influence of liquor unless you knew a little jujitsu for every contingency as even a fellow on the broad of his back could administer a nasty kick if you didn't look out. (U 706)

This excerpt, from early in the episode, has a total of 125 words, of which 82 are part of one cliché or another, for a Cliché Density Index (CDI) of 65.6%. But this passage has more clichés in fact than is usual for Bloom. The total number of clichés on the page from which this passage was taken is 61, and Joyce was *lathering on the clichés* here, as Bloom might say, turning the whole episode into a burlesque of naturalistic characterization. For "Eumaeus" as a whole,

Bloom produces an average CPP of 36, and this is the number we'll use for comparisons. Where on this spectrum might other major characters in *Ulysses* find themselves?

5. CDI and CPP for Stephen Daedalus

Let's start with Stephen, the Jesuit-educated Telemachus figure, recently graduated from University College Dublin with a degree in modern languages. If we look just at raw numbers, the difference is breathtaking. For the first 5 pages of "Proteus," for example, Stephen's 20-page monologue, the CPP averages 1.58%, which is to say virtually no clichés whatsoever, not even humble collocations, for a CDI of 0.49%. This is far below the human norm, and Joyce has of course expressly engineered it. All of the phrases in Stephen's stream of consciousness are "original," as we might say, and a great many of them unprecedented. Why? Because Stephen is an artist, a protagonist in the modernist "revolution of the word," in whose work no commonplaces of language can reappear—and so Joyce banishes them from his stream-of-consciousness as well. Stephen's thought is auditorially sharp, a phonetic mouthful,

"The grainy sand had gone from under his feet. His boots trod again a damp crackling mast, razorshells, squeaking pebbles, that on the unnumbered pebbles beats, wood sieved by the shipworm, lost Armada." (*U* 50)

or wildly imagistic,

"He had come nearer the edge of the sea and wet sand slapped his boots. The new air greeted him, harping in wild nerves, wind of wild air of seeds of brightness." (*U* 55)

and always metaphoric,

"A bloated carcass of a dog lay lolled on bladderwrack. Before him the gunwale of a boat, sunk in sand. *Un coche ensablé*, Louis Veuillot called Gautier's prose. These heavy sands are language tide and wind have silted here." (*U* 55)

But there are no clichés anywhere, nothing remotely like one. Stephen thinks like a symbolist aesthete might well do, before writing up his thoughts in purple prose. The modernist revolution of the word purposed to reprimatinate the world by reenergizing language with unexampled locutions. *Locutions reçues* were banned, as a preliminary prophylaxis, and the new literary seer was incited to revision the world with Adamic language.

Clearly, Joyce had a spectrum in mind, with Bloom as one end, bogged down below some abysmal threshold of blinkered imperciption, with a CPP of 36 and CDIs rising above 60%, and Stephen at the other end, virtually transcending the language of common parlance, a Nietzschean *ubermensch* speaking an *ur-sprach* of visionary purity, with virtually no CDI or CPP whatsoever. Existential substance seems spread out along this spectrum as well. Bloom, waterlover, is floundering in tribal formulae, rehearsing his subjection in every utterance, while Stephen, hydrophobe like Icarus, sails above these exhausted idioms, helming toward the sun, oblivious to the Bloomian fate that awaits him, as well, in middle age.

And clearly, Joyce counts clichés. He inflicts them on characters to shadow forth their existential substance. Indeed, his primary mode of characterization is linguistic, which is to say, more clichés, less character. We may note in passing that Bloom loves to immerse himself in water, the material element, symbolically speaking, and that Stephen is a hydrophobe, refusing to

bathe, and declining, Icarus-like, even to swim in Dublin Bay. Clichés are like the watery element: a mind drowns in them, and loses its spiritual or intellectual autonomy. Bloom is the drowned Icarus that Stephen, shunning water and clichés, fears he may become.

If we're correct here, then simply counting clichés may lead us straight to the heart of Joyce's art. The CPP may tell us sooner about Joyce's artistic intentions for a character than careful critical assessment with other approaches. Some characters are merely bearers of inherited platitudes and recycled locutions; Joyce has abased them deliberately, cliché by cliché. We are to see them ironically, or comically, even if they show some rhetorical energy in their streams of consciousness. They are naturalistic characters, naturalism being that novelistic mode, patented by Emile Zola, that portrays humankind as driven and determined by accidents of birth and environment—in *Ulysses*' case, the petit-bourgeois milieu in which Joyce's characters "live and have their being."

6. CDI and CPP for Gerty MacDowell

Gerty McDowell is a likely candidate, as Joyce identifies her with the Homeric Nausikaa, and gives her a 30-page monologue. She beguiles her time inventorying her romantic assets and defects, longing for a proposal, and annotating the season's fashions. Joyce's free indirect discourse takes us directly into her stream of consciousness, the narrator borrowing his clichés from Gerty's own lexicon. We see her describing herself in a language that is both heightened with the phrasemes of popular romance, and entirely clichéd. Gerty has been reading Maria Cummins's *The Lamplighter*, from which her diction and breathless romanticism derive, as Joyce criticism has established, and particularly Cummins's tendency to romantic hyperbole in ceremonious cliché, which Gerty has mastered. Here she describes herself:

There was an *innate refinement*, a languid queenly *hauteur* about Gerty which was *unmistakably evidenced* in her *delicate hands* and *higharched instep*. Had kind *fate but willed* her to be born a *gentlewoman of high degree in her own right* and had she only *received the benefit of a good education* Gerty MacDowell might easily have *held her own beside any lady in the land* and have seen herself *exquisitely gowned with jewels on her brow* and *patrician suitors at her feet vying with one another to pay their devoirs* to her. Mayhap it was this, *the love that might have been*, that lent to her softly-featured face *at whiles* a look, tense with suppressed meaning, that imparted a strange yearning tendency to the beautiful eyes *a charm few could resist* (my italics for the clichés). (U 453)

There are so many clichés here (CDI of 52.27%), they run into strings, like *to be born a gentlewoman, of high degree, in her own right, or suitors at her feet, vying with one another, to pay their devoirs*. Gerty's mastery of Cummins's idiom is impressive, but she's been reading pulp romances too, and knows how to emote tearfully in the received popular phraseology:

And yet and yet! That *strained look* on her face! A *gnawing sorrow* is there *all the time*. Her *very soul is in her eyes* and *she would give worlds to be in the privacy of her own familiar chamber* where, *giving way to tears*, she could *have a good cry* and *relieve her pentup feelings*. (U 456)

Above all she knows how “falling in love” should go, even how a first kiss should be stage-managed:

he who would woo and win Gerty MacDowell must be a man among men. But waiting, always waiting to be asked and it was leap year too and would soon be over. No prince charming is her beau ideal to lay a rare and wondrous love at her feet but rather a manly man with a strong quiet face who had not found his ideal, perhaps his hair slightly flecked with grey, and who would understand, take her in his sheltering arms, strain her to him in all the strength of his deep passionate nature and comfort her with a long long kiss. It would be like heaven. (U 457)

From the passages above, it looks like Gerty’s character has almost dissolved into her own romantic phraseology, but over the course of the monologue her CPP ranges between 19 and 25, averaging 22, much less than Bloom’s average 36. Her stream of consciousness is significantly less stuffed with the formulae of fiction than Bloom’s is with those of the tribe. And her formulae have a redeeming poetics to them, invoking as they do the literary genre of romance, a heightened imaginative response to life, and an aspiration to transcend the determinants of her otherwise abject existence. She may be largely trapped in her popular literary idiom, but she dreams of redemption from mean conditions and emotional constraints. Bloom is entirely trapped in his plain-speaking, graceless tribal language and he’s impressed with himself; Gerty has found a higher idiom, and rues her wretched, unconsummated state. Bloom is entirely imprisoned; Gerty is not, and the CPPs tell us this even before we parse the clichés and identify their valences and provenience.

7. CDI and CPP for Molly Bloom

And now we come to the most famous episode in *Ulysses*, the last, “Penelope,” in which Leopold Bloom’s wife Molly, lying in bed, reviews her amorous life in stream-of-consciousness. This episode is performed all over the English-speaking world every June 16th, the date on which the entire action of *Ulysses* takes place. Famous actresses prove their verbal art by rendering it convincingly, and amateurs do it for pure fun. English graduate students bemuse themselves with the question of whether a man, Joyce, has succeeded in accurately presenting a woman’s mind, and critics debate, yet again, the question of Molly’s personality. Is she a professional singer, a typical petite bourgeoisie, and a neglected wife, on the one hand, or is she an earth mother and a femme fatale, a Greek siren who has pulled Bloom down into the watery element, the mindless organic and reproductive cycles that constrain and determine human lives—and that inspired Zola’s naturalist literary doctrine. In some way, her CDIs and CPPs will answer these questions, at least as much as quantitative data can answer such questions at all.

It’s possible to find a few passages in “Penelope” with a high CDI, but they are misleading, since it’s just as easy to find ones entirely without cliché. Here’s an excerpt with a high CDI, and I offer it simply as a way of showing, in brief compass, the *kind* of clichés Molly uses:

[One] woman is not enough for them *it was all his fault* of course ruining servants then proposing that she could eat at our table on Christmas *if you please O no thank you not in my house* stealing my potatoes and the oysters 2/6 per doz going out to see her aunt *if you please common robbery so it was* but I was sure *he had something on with that one* it takes me to find out *a thing like that* he said you have no proof it was her proof O yes her aunt was very fond

of oysters but *I told her what I thought of her* suggesting me to go out to bealone [sic] with her *I wouldn't lower myself* to spy on them the garters I found in her room the Friday she was out *that was enough for me* (U 873-4)

There are 143 words in this passage, and 51 one of them are clichéd, for a CDI of 35.6%, but as I've said, this isn't representative. More typical is what follows, a selection from near the end of the novel, when Molly remembers how she said “yes” to Bloom's marriage proposal:

I was thinking of so many things he didn't know of Mulvey and Mr Stanhope and Hester and father and old captain Groves and the sailors playing all birds fly and I say stoop and washing up dishes they call it on the pier and the sentry in front of the governors house with the thing round his white helmet *poor devil* half roasted and the Spanish girls laughing in their shawls and their tall combs and the auctions in the morning the Greeks and the jews and the Arabs and *the devil knows* who else from *all the ends of Europe* and Duke street and the fowl market all clucking outside Larby Sharons and the poor donkeys slipping half asleep and the vague fellows in the cloaks asleep in the shade on the steps and the big wheels of the carts of the bulls and the old castle thousands of years old yes (U 932)

With 154 words in total, of which 11 are clichéd, the CDI for this passage is a mere 7.1%. Here, Molly is far from the unselfconscious vessel of Dublin commonplace that her husband is in “Eumaeus.” Only the average CPP will tell us about Molly's burden of cliché overall. I sampled “Penelope” in several places, starting with 9 consecutive pages early on in the episode (pages 873-881) and four pages toward the end (pages 902-903, 919-920). The average CPP is fairly consistent, with just a few pages, early on in the episode, with generally higher ones. The CPPs for pages 873-881 were 23, 16, 7, 9, 7, 9, 22, 12, and 5. The high concentrations of cliché seem to come in spates, at least early on in the episode. For pages 902-903 and 919-920, the scores were 7, 7 and 10, 11. The overall average CPP for Molly Bloom is 11—by far the lowest among the main characters, except for Stephen.

	Average CPP	Maximum CDI in %
Leopold Bloom	36	65.6
Stephen Daedalus	1.58	0.49
Gerty MacDowell	22	52.27
Molly Bloom	11	35.6

In effect, Molly escapes naturalistic characterization altogether, at least in terms of her phraseology. She has a distinct, original voice, readers agree, and we see now why—most of her utterance isn't formulaic, and the little that is, enters her speech as local colour rather than as linguistic constraint. It lends her speech emphasis rather than limiting it to tribal commonplace, as has happened to her husband. In short, while Joyce portrayed her husband as a character linguistically determined by his environment, reduced in his rhetorical resources and human substance to fit his role as anti-hero and *untermensch* in the ironic Odyssean paradigm, Molly by contrast is a character of Joyce's realism, highly individuated in her speech and stylistically opportunistic in her use of received locutions. Unlike a naturalistic novel, for example Zola's *Nana*, *Ulysses* ends not in defeat or despair or death, but in the assertion of free personality.

Like Stephen, Molly lives in a cramping, naturalistic environment, but it doesn't subsume or determine her. Both of them are realistic characters in an otherwise naturalistic novel, and the CPPs bring quantitative support to this critical conclusion.

8. Conclusion

This quantitative approach X-rays the language that so much critical commentary addresses impressionistically, and identifies precisely some of the linguistic correlates of our aesthetic response to *Ulysses*. Bloom has always seemed less free, more ridiculous than his wife Molly. He has always been a kind of naturalistic character, although a "kinder, gentler" one in Joyce's comedy. Gerty, an energetic recycler of romantic cliché, nevertheless uses far fewer than Bloom, and earns the reader's admiration: her romantic desire for a more intensely-lived life is almost Paterian, and not at all pathetic, however comically cute it is at times. Molly, who hardly speaks before "Penelope," emerges as the most powerfully grounded character in the novel, as Stephen is the most aesthetically and linguistically remote from the common life of Dublin. Molly is realistic, not reducible to Earth-Mother, Siren, Femme Fatale, but rather a highly individuated petite bourgeoisie, a professional singer, and a housewife. Stephen is an aesthete, longing for the superhuman, wishing to escape the entangling nets of history he wishes to fly beyond, avoiding even the earthly language of Dublin in which he is immersed, but not abject, as was Bloom. Now, with stylometric help, we see one of the major reasons why: for Joyce, language is character, and *Ulysses*, among other things, is a portrait of character based on an artistically deliberate use of cliché. The burden of cliché Joyce loads upon Bloom and Gerty is what measures, with breathtaking precision, their reality and dignity in our response. Stephen's freedom from cliché, and Molly's largely original voice, is what makes them the realistic characters in this otherwise comic naturalist novel. For *Ulysses*, for Joyce, characterization is linguistic, and only quantitative measures make apparent how deliberate Joyce's characterization has been, how engineered our responses have been—and incidentally, how much of a monument *Ulysses* is to Dublin common parlance in 1904.

References

- Byrnes R. (2010). *Jadt2010: Statistical Analysis of Textual Data*. Milano: Edizioni Universitarie di Lettere Economia Diritto.
- Cowie A. P. (1994). Phraseology. In Asher, R. E. editor, *The Encyclopedia of Language and Linguistics*, 3168-3171. Oxford: Oxford University Press.
- Joyce J. (1922, 1992). *Ulysses*. England: Penguin Books.
- Raleigh J.H. (1981). On the Way Home to Ithaca: The Functions of the 'Eumaeus' Section in *Ulysses*. In Bowen, Z. editor, *Irish Renaissance Annual*. London and Toronto: Associated University Presses.
- Wray A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.