

Is academic writing becoming more informal?

Ken Hyland & Fang (Kevin) Jiang

There is a general, though largely unexamined, assumption among those interested in such things that writing in many domains has become less formal in recent years (e.g. Adel, 2008; Foster, 2005; Fairclough, 2001) . Observers have noticed a gradual shift away from standard detached and impersonal styles of writing to ones that allow more personal comment, narration and stylistic variation, so that Mair, for example, notes “a trend towards the informal and the colloquial in written communication” (1998: 153) and Leedham (2015) found greater informality in undergraduate essays. This trend towards informality might be seen as part of a contemporary zeitgeist which blurs overt hierarchies and values interpersonal engagement or, alternatively, be regarded as another form of insidious persuasion, what Fairclough (2001: 52) calls “synthetic personalization”. It is possible, however, that the informality which has invaded a large range of written and spoken domains once characterized by formality (journalism, business correspondence, administrative documents, etc.) has also spread to academic writing. Indeed, it has been fashionable among applied linguists in recent years to search for evidence of greater interactivity in academic prose and identify the ways that writers craft an inclusive relationship with their readers (e.g. Hyland, 2004).

In this paper we explore this issue and investigate whether academic writing is becoming less formal and, if so, in what ways and in what disciplines. We first attempt to characterize the notion of ‘informality’ and how it is understood by academic discourse analysts and those who advise authors on academic style. Needless to say, while we know it when we see it, we find ‘informality’ to be a slippery concept, difficult to pin down with a clear definition. It is typically either defined in contrast with formality, or in terms of lists of language features which are thought to comprise ‘informal elements’, such as using imperatives, employing "I" or starting sentences with "but". Focusing on published research articles as the most important genre of academic writing, we then undertake a

comparative study of three corpora composed of papers in four fields drawn from three distinct time periods, examining the frequencies of key ‘informal elements’.

What is informality?

The first place to start is with a definition or characterization of the term, but this is not altogether straightforward. *Informality* is generally hetero-defined, the Latin prefix meaning "not, opposite of, without" (Online Etymology Dictionary, 2010) and demarcated as the absence of formality. It therefore presupposes the existence of formality and a recognized set of practices built on a structure, authority or system. Thus the Cobuild dictionary defines formal speech as “very correct and serious rather than relaxed and friendly” while in pragmatics, formality is associated with ‘negative politeness’ and the use of distancing behaviour to respect the other’s face and their wish not to be imposed on (Brown & Levinson, 1987). More broadly, Heylighen & Dewaele (1999: 1) state that: “A formal style is characterized by detachment, accuracy, rigidity and heaviness; an informal style is more flexible, direct, implicit, and involved, but less informative”. In academic writing, then, formality helps to avoid ambiguity and misinterpretation by minimizing the context-dependence and fuzziness of expressions, while, in contrast, informality rejects stuffy orthodoxy to project a relaxed and approachable persona.

Questions of formality thus relate to tenor, or the grammatical choices that enable speakers to enact their complex and diverse interpersonal relations by selecting language options which project an appropriate persona and a suitable connection with readers (Halliday, 1985). It is, therefore, associated with concepts like *colloquial language* (e.g. Hundt & Mair, 1999) or language used in everyday conversation by ordinary people, and *engagement*, or how writers acknowledge and connect with their readers (Hyland, 2005). We should not, however, jump to the conclusion that the gap between conversation and academic writing is narrowing, despite the recognition that the latter is often strategically interactive (e.g. Hyland, 2004). Academic genres, in fact, appear relatively resistant to penetration by colloquial features (Seone & Loureiro-Porto, 2005) so that Hundt and Mair

(1999: 221), for example, suggest that they reside at the more conservative end of “a cline of openness to innovation ranging from "agile " to "uptight" genres”.

A key reason for this is that in research writing adherence to the conventions of formality suggests impartiality, precision, distance and a faux egalitarianism, allowing authors to construct themselves and their readers as disinterested specialists. The search for truth is couched in objectivity and the features of formality serve to minimize the quirks, foibles and interests of individual authors to suggest an anonymous writer conducting democratic interactions with like-minded peers. This is a context in which status, gender, experience and other social characteristics are subordinated to the accurate and detached presentation of information. The conventions of formality mean that, as far as possible, authors leave their personalities at the door when they sit down to write.

As a result, students are frequently cautioned against informality as it conveys impressions of the author that may be unwelcome in academic writing. Style guide advice for undergraduates often point this out:

In sum, there's a great disadvantage to writing informally in a history or classics class, since it makes you look both casual and rushed, neither of which will help your grade. Conversely, there's a great advantage to writing formally, especially here, since formality forces you into a posture where you appear to create some distance between your own feelings and the cause you're arguing for. That's good in this case, because it puts you in a more objective stance right from the start. Objectivity—or even the mere appearance of being objective—is good in academic writing.

(Guide to writing in history and classics, 2015)

Such advice is also given to research writers, although it is generally expressed more circumspectly:

Academic writers need to be sure that their communications are written in the appropriate style. The style of a particular piece should not only be consistent but also be suitable both in terms of the message being conveyed and the audience. A formal research report written in informal, conversational English may be considered too simplistic, even if the actual ideas and/or data are complex.

(Swales & Feak: 2012: 14)

Informality, then, is generally contrasted to what is seen to constitute formality. An influential early example of such a contrast is Gilbert and Mulkey's (1984) study of how scientists discuss their work in written and spoken modes. They observe, for example, that interviews with academics are typically littered with controversies, speculative insights, intellectual commitments and social biases. This is what the authors refer to as a "contingent repertoire" where actions are not disinterested responses to the natural world but the "judgements of specific individuals acting on the basis of their personal inclinations and particular social positions". In contrast, of course, these features are absent in the same academics' research articles which are governed by a more objective "empiricist repertoire":

Empiricist discourse is organized in a manner which denies its character as an interpretive product and which denies that its author's actions are relevant to its content...it portrays scientists' actions and beliefs as following unproblematically and inescapably from the empirical characteristics of an impersonal natural world.

(Gilbert & Mulkey, 1984: 56)

Differences between the two, however, may be less cut and dried than this as research writing carries obvious traces of speculation, personal biases and group affiliations (e.g. Hyland 2004), but Gilbert & Mulkey's distinction captures how formality helps guide academics' research writing.

While informality may be seen as a deviation from a traditional academic stance of objectivity, it is not just a colloquial style or the opposite of detachment. Informality in academic writing is the expression of a more personal tenor which implies a closer relationship to readers, a willingness to negotiate claims and a positive attitude towards subjectivity. Relevant to this study, it also implies a set of features and rhetorical characteristics which can be identified and counted, allowing us to measure changes in academic argument. Considering informality as a set of distinctive features used to achieve certain rhetorical goals also allows us to see the formal/informal distinction as a continuum, rather than as a language binary. Coffin et al (2003) make this explicit in their definition:

By formality we mean the use of technical, elevated or abstract vocabulary, complex sentence structures and the avoidance of the personal voice. If we think of formality as a cline from the most formal (e.g. legal documents) to the most

informal (e.g. electronic mail between friends), most academic writing falls nearer to the legal documents than the friendly email. (Coffin et al, 2003).

The notion of a continuum is also implicit in Heylighen and Dewaele's (1999) understanding of formality as the avoidance of ambiguity. Informality for them is information which refers to the background knowledge and assumptions which makes communication possible, but complete informal fuzziness "merely signifies that any interpretation is as likely as any other one" (Heylighen & Dewaele, 1999: 9). Any text must carry such assumptions but have at least a minimal formality to be intelligible. In academic writing, then, the basic advantage of formality is that there is less chance of being misinterpreted or persuasion being influenced by the social characteristics of writers. Any increase in informality is therefore likely to reflect changing norms of engagement which accommodate greater flexibility of tenor rather than the abandonment of conventions.

Informality, then, is not simply a reluctance to attend to conventional practice, any more than it is an inappropriately colloquial use of language. Instead it is an attempt to establish a particular kind of relationship with readers, one which makes assumptions about a shared context and seeks to create a collegial familiarity. Such patterns of engagement, of course, display disciplinary membership and so need to ensure that appropriate standards of precision are observed, meaning that academic writers are unlikely to stray too far along the cline towards informality.

Features of informality

Informality in academic writing can be identified by a set of features which have a high chance of co-occurrence, although there is some disagreement about quite what the decisive features are.

One approach to identifying related features is Biber's (1988) multidimensional corpus analyses of spoken and written texts. Using a correlational statistical technique known as factor analysis, he shows how 16 major grammatical categories typically co-occur in five dimensions of variation, with Dimension 1, labelled "involved versus informational production", being closest to what we

understand as informality/formality. Involved genres contain more verb and pronoun forms, adverbs and interjections and have fewer nouns, prepositions and attributive adjectives. Biber (1995: 143) observes that involved features “reflect direct interaction, focus on the immediate circumstance and personal attitudes or feelings, fragmentation or reduction in form, and a less specific, generalized context”. While Biber describes these differences as oral vs literate discourse, they characterize the kind of engagement which is typical among intimates or in encounters where there is no necessity to elaborate a shared context. In this way they help clarify the concept of formality and its expression.

A second approach to formality has been proposed by Heylighen and Dewaele (1999) who see it in the assembly of features which contribute to the precision and context independence of a text. So by listing all words with a deictic function, referring to the communicative context, in one category and features which largely assume knowledge of such a context in the other, they propose a formula to measure formality. Higher frequencies of nouns, adjectives, prepositions and articles are said to reduce the level of background information needed to understand a text and so make a text more formal while more verbs, pronouns and interjections refer more to the immediate context and make it less formal. Thus, as more context is provided, the level of formality increases. Outside of a purely mathematical proof, however, a completely formal text which fails to build a context and a relationship with readers is impossible, but an increase in formal features reduces the background knowledge needed to understand a text.

Formality is thus a result of the need for more context and all texts are situated between the two extremes of complete fuzziness and complete exactitude with the precise degree of formality influenced by the writer’s assessment of the context. So, Heylighen and Dewaele’s characterization is similar to Biber’s in identifying formal texts as more nominal and informal texts as more verbal, with more-or-less rather than all-or-nothing distributions. This is, however, is a very broad approach which equates formality with context independence and makes assumptions about the function of grammatical classes in achieving this. But while it is true that pronouns serve to deixically anchor a statement to a given context, and so may be interpreted as belonging to the ‘informal’ category, we

should be cautious in supposing that nouns are always more formal than verbs. Indeed, by highlighting a writer's stance towards a message, noun complement structures can introduce a more informal and engaging element into texts (Jiang & Hyland, 2015), as here (all examples from our corpus):

(1) Collectivism makes the **mistake** of making the individual fully heteronomous by locating the sovereign source of normativity entirely outside the individual in the community. [Philosophy article]

(2) The procedure offers the **advantage** of obtaining more precise structural estimates. [Marketing article]

Such expressions are among the resources used by academic writers to front load their attitude to the accompanying clause and so can tie utterances very closely to the interactional context.

A third way of identifying features of informality is to turn to the style guides, which reflect something of the concerns and thinking of teachers, students and practitioners. Bennet's (2009) survey of style manuals, for example, found frequent references to 'objectivity' and an insistence "that academic writing is by nature formal and technical", achieved by the extensive use of Latinate vocabulary and impersonal structures. Chang and Swales (1999) researched 40 style manuals published to compile a list of the most frequently mentioned grammatical features recommended to attain an appropriate degree of formality. First person pronouns, WH questions, listing expressions (such as etc., and so on) and contractions appear in their list of 'informal elements' although style manuals authors take different positions with regard to how suitable these might be in research writing. Our own brief survey of the first 25 relevant sites thrown up by a simple *Google* search, including many university language advice sites, reveals similar admonishments to avoid many of these features, although this is perhaps less 'absolutist' than in earlier times.

The uncertainties surrounding informal features remind us that such elements are not monolithic conventions observed in all academic genres, and style guide authors have problems capturing the complexity of the rhetorical choices involved in selecting and employing them. There is also the fact

that formality at the sentence level is likely to be subjectively interpreted by readers, influenced by personal feelings, tastes, or opinions and subject to fashion and changing usage. Nor are these features static and unchanging markers of professional research writing. Using Biber's multidimensional categories, for example, Atkinson(1999) tracked changes in papers published by the Royal Society of London in its journal *The Philosophical Transactions* between 1675 to 1975 finding a relentless growth in 'informational' features until the latter date, a change which Atkinson describes as a move from a less 'author-centred' rhetoric to a highly abstract and 'object-centred one'.

In sum, we have sought to prepare the ground for a study of informality in academic writing. We have sought to define a somewhat slippery term, first in terms of the concept the prefix opposes, formality, and then as linguistic features which establish a close relationship with readers by realizing a relatively personal tenor which allows writers to make assumptions about a shared context. We have also sought to show that this tenor can be identified and measured by the presence of certain features often regarded negatively by style guide authors. We now turn to our corpus and analytical methods to address our guiding question: Has academic writing become less formal and, if so, in what ways and in what disciplines?

Corpus and method

To discover to what extent and in what ways professional academic writing may have grown more informal in recent times, we constructed three corpora to present a snapshot of four disciplines at three periods over the past 50 years: 1965, 1985 and 2015, a period of 20 years between the first dates and then to the present day. The different time spans were chosen to see if any changes were more pronounced in the later or earlier period, although we were concerned with overall changes over the 50 years. We chose applied linguistics, sociology, electrical engineering and biology as representative of both the soft applied fields and the hard sciences, selecting six papers at random from each of the same five journals which had achieved the top ranking in their field according to the

year impact factor in 2015¹. The journals are listed in Appendix 1 and together the corpus comprised 360 papers of 2.2 million words as shown in Table 1:

Table 1: Corpus size and composition

Discipline	1965	1985	2015	Overall
Applied linguistics	110,832	144,859	237,452	493,143
Biology	244,706	263,465	237,998	746,169
Engineering	92,062	97,545,	235,681	425,288
Sociology	149,788	196,232	262,203	608,223
Totals	597,388	604,556	973,334	2,272,823

The corpora were grammatically annotated (part of speech tagged) using *Tree Tagger* then combed for ‘informal elements’ using the *AntConc* concordance software (Anthony, 2011). Some of the features listed in Table 2, such as *first person pronouns* and *unattended anaphoric pronouns*, were retrieved easily through concordancing individual items (e.g. I/we and this/these/that/those/it).

Others, like split infinitives, were searched using a regular expression query based on their syntactic structure². After extracting the items, each one was then checked manually by each author independently to confirm it was a target feature, producing an inter-rater agreement of 93% before resolving disagreements.

We decided to employ the list of informal features identified by Chang and Swales (1999³) as the items here corresponded most closely with our own survey and our experience as language teachers. We made one change as the category of sentence fragments almost never occurs in research writing, where cases are presumably removed by authorial self-editing or the language professionals who are sometimes employed by writers to edit their texts. We substituted this with second person pronouns, which do occur although are generally regarded as taboo by style authors. The widely used *Writing Commons*⁴ website, for example, advises:

¹ Two journals, *TESOL Quarterly* and *Foreign Language Annals* only began in 1967 and so papers were chosen from issues in that year.

² `to_TO\s(\w*_RB\s)+\w*_V\w*` was used to search for split infinitives in the POS tagged corpus.

³ See also Swales and Feak (2004) *Academic Writing for Graduate Students*. Ann Arbor: University of Michigan Press.

⁴ *Writing Commons*, <http://writingcommons.org> is a peer-reviewed resource for academic writers.

Writing from the second person point of view can weaken the effectiveness of the writing in research and argument papers. Using second person can make the work sound as if the writer is giving directions or offering advice to his or her readers, rather than informing or persuading them.

The final list, presented with Chang and Swales' examples, is given in Table 2. These features not only express the concerns of style guide writers and teachers but also correspond to what is generally considered to be 'informal' language (e.g. see Biber, 1988; Chafe, 1982, Nash, 1986), conveying interactiveness, involvement and personal engagement with readers. The list also carries the apprehensions of novice research writers who are often understandably nervous about making use of such stylistic opportunities (see Chang & Swales, 1999) . We also suspected that these features, or at least some of them, might reflect the tension between linguistic prescriptivism and current authorial practice, as we discuss in the next section.

Table 2: List of informal features

1. First person pronouns to refer to the author(s) (*I* and *we*)
e.g., "*I* will approach this issue in a roundabout way.
 2. unattended anaphoric pronouns (*this*, *these*, *that*, *those*, *it*) that can refer to antecedents of varying length
e.g., "*This* is his raw material."
 3. split infinitives -- an infinitive that has an adverb between *to* and the verb stem
e.g., "The president proceeded *to* sharply *admonish* the reporters."
 4. Sentence initial conjunctions or conjunctive adverbs
e.g., "*And* I will blame her if she fails in these ways."
 5. Sentence final preposition
e.g., "A student should not be taught more than he can think *about*."
 6. listing expressions ('and so on', 'etc', 'and so forth' used when ending a list)
e.g., "These semiconductors can be used in robots, CD players, *etc*."
 7. Second person pronouns/determiners to refer to the reader (*you* and *your*)
e.g., "Suppose *you* are sitting at a computer terminal which assigns *you* role R"
 8. contractions
e.g., "Export figures *won't* improve until the economy is stronger."
 9. direct questions
e.g., "*What can be done to lower costs?*"
 10. exclamations
e.g., "*This is not the case!*"
-

Trends in academic style: the demise of formality?

Our results show a fairly complex picture when we seek to measure these ‘informal elements’. Table 1 above shows a substantial growth in the length of published articles over the period, amounting to a 63% increase in the number of words, and we see an almost exactly corresponding growth in the use of informal elements, with a similar increase in raw frequencies. Table 3 presents the changes in informality features in the corpus over this period as occurrences per 10,000 words.

Table 3: Distribution of features of informality over time (per 10,000 words)

Discipline	1965	1985	2015	% change
Applied linguistics	213.8	202.7	191.7	-10.3%
Sociology	205.0	200.5	198.8	- 3%
Electrical engineering	142.4	150.0	155.2	+ 9%
Biology	112.3	121.4	140.2	+ 24.8%
Averages	168.4	168.7	171.5	+ 2%

In broad strokes, there has been a small increase of around 2% in the use of informality features in published academic writing over the past 50 years, when the increase in overall words published is taken into account. This is, however, largely a result of increases in hard science writing rather than in the social sciences, which have become slightly *more* formal. As the table shows, writers in applied linguistics have steadily reduced their use of informal elements since 1965 by about 10.3% and sociology by about 3%. Electrical engineering gradually increased the number of informal features in its main journals over the period, up by 9% , and biology, once again the odd-one-out, rising by a substantial 24.8%.

Table 4 shows that the overall figures are largely influenced by the frequencies of three main features: first person pronouns, unattended reference and sentences beginning with conjunctions and conjunctive adverbs, with second person pronouns/determiners and listing expressions much further behind. The table also shows that informal features are not changing in a single direction nor behaving in uniform ways, either across times or disciplines. Contractions and split infinitives now

seem to be more tolerated by copy editors than in the past and have increased across all fields, as have direct questions, while listing markers show a marked decline overall. Sentences ending with a preposition fell slightly and rose again to a lower level than originally in the soft knowledge fields and barely showed any movement in the sciences. Second person increased greatly in applied linguistics but fell in sociology while first person fell in applied linguistics and rose in the other disciplines. In the next section we discuss these trends in more detail.

Table 4: Changes in use of informality features by discipline (per 10,000 words)

Feature	Applied Ling			Sociology			Electrical			Biology		
	1965	1985	2015	1965	1985	2015	1965	1985	2015	1965	1985	2015
first person	60.9	53.6	52.5	46.5	42.8	68.6	46.9	53.9	68.8	10.9	23.6	34.1
unattended reference	108.9	70.8	71.8	104.9	75.0	66.2	61.9	60.6	58.9	80.7	62.5	44.1
Initial conjunctions	30.6	32.0	37.9	38.3	37.7	49.0	30.1	31.2	38.6	19.3	29.6	40.1
second person	4.9	6.4	10.0	9.3	7.2	4.4	0.0	0.2	0.0	0.0	0.2	0.3
Listing expressions	3.5	2.8	1.9	1.6	1.4	1.0	2.3	1.3	1.0	0.4	0.4	0.1
contractions	1.8	4.8	13.5	2.3	3.3	5.1	0.0	0.1	0.1	0.0	0.1	1.2
preposition ending	1.3	0.3	1.0	1.2	0.5	0.9	0.2	0.3	0.5	0.0	0.0	0.3
exclamation	1.3	1.2	0.8	0.4	0.1	1.1	0.0	0.1	0.0	0.3	0.1	0.0
split infinitives	0.6	0.7	2.1	0.5	0.8	2.5	0.8	2.3	2.4	0.1	1.3	2.3
direct questions	0.0	0.0	0.2	0.1	0.2	0.1	0.1	0.0	0.1	0.5	0.1	0.1

How is formality changing?

As mentioned, first and second person pronouns, unattended reference and sentences beginning with conjunctions/conjunctive adverbs comprise the majority of features, accounting for 97.3% of all informal elements in 1965 and 94.3% in 2015. In this section we discuss each of these features in turn.

First person pronouns, often considered the defining marker of informality, account for much of the increase in informal features with 45% overall growth during the period. This phenomenon has also been noted by Seone (2013) who has traced a decline in the frequency of *be*-passives with respect to

transitive actives between 1905 and 1990 in hard science articles both in British and in American texts. A reduction in the use of the passive in her data suggests an increase in the need for first person pronouns. This drop from around 65% to 53% of cases is more pronounced towards the second half of the century and seems, moreover, to be a more general trend in all types of writing (Mair & Leech, 2006; Leech et al., 2009). In academic genres the increase in first person may be related to the greater participation in published research by authors whose first language is not English and in which first person does not carry the same connotations of personal projection and authority (see Nelson & Castello, 2012). Interestingly, in our data, applied linguistics is the only discipline of the four which has seen a decline in the use of first person, perhaps as a result of the self-consciousness of language-sensitive writers aware of the attention this draws to the individual and the strong claims it makes for agency in research.

Overall, however, the convention of avoiding the first person to convey an impersonal stance, once a hallowed principle for style guide writers and science authors, now seems to be less rigidly adhered to, most dramatically in biology with a massive 213% rise (from 10.9 per 10,000 words in 1965 to 34.1 in 2015). It may seem surprising that it should be a scientific discipline which has been most enthusiastic in undermining what is a cornerstone of positivist objectivity. Empiricism has traditionally stressed the persuasive authority of impersonality to maximize the credibility of the writer and elicit credence from the reader. Lachowicz (1981: 111) for example, argued that impersonality emphasises “objectivity, open-mindedness, and the established factual nature of a given activity”, it functions to underline the “common share of knowledge with the community”, and stresses the collective responsibility of academic endeavour. Eradication of the self is seen as demonstrating a grasp of scholarly persuasion as it allows the research to speak directly to the reader in an unmediated way.

Academic publishing demands that authors display some degree of disciplinarity; that they demonstrate a familiarity with the rhetorical conventions and social understandings of the community, and observe suitable patterns of social and rhetorical interactions (eg Hyland, 2004). Arguments have

to be made in ways that readers find most acceptable and convincing, and research claims framed to project appropriate certainty and maximum plausibility. However, in addition to constructing themselves as plausible members of the discipline, research writers must balance this with vigorous argument for the originality of their claims and the display of an authoritative professional persona. Ivanic (1998) and Hyland (2001) have stressed the importance of *I* in constructing an independent voice and authorial identity in academic writing. Hyland (2012), in fact, refers to choices relating to proximity or positioning. This involves constructing a text which not only persuades but also engages readers and first person pronouns are a powerful means for establishing authority and personal projection. It appears, as we will discuss further below, that in the changing and highly competitive context in which research is now conducted and presented, these considerations are becoming ever more important to writers.

Unattended reference is associated with informality due to its predominance in spoken varieties of discourse and the fact many style guides recommend careful writers to generally avoid it as “Out of control, the unattended *this* points everywhere and nowhere” (Geisler et al, 1995: 153). Swales and Feak (2004) argue that following *this* with an appropriate NP reduces possible comprehension difficulties by readers and helps make the writer appear more professional and authoritative. Strauss (1993) estimates that 40% of uses in speech are unattended and Swales (2005) gives a similar figure of 46% for MICASE dissertation defences, both of which are much higher than in research texts (around 30% of cases). They occur quite frequently in academic prose, however, around 60 times per 10,000 words on average, due to their use in marking immediate text reference (Biber et al. 1999: 349). But we can see from Table 4 that this element has declined dramatically across all disciplines during this period, by 34% (from 108.9 to 71.8 per 10,000 words) in applied linguistics, 37% (from 104.9 to 66.2) in sociology and 45% (from 80.7 to 44.1) in biology.

These anaphoric, or backward-pointing pronouns (*this, these, that, those, and it*) refer to antecedents of varying length, from individual words to entire sections of texts. Gray (2010) argues that they follow the given/new principle of thematic progression and so function to alert readers to the fact that

new information is about to be presented. Wulff et al (2012) suggest the decision to use unattended *this* is related to the accompanying verb, with simple forms such as *this is* and *this means* most common. But what governs the writer’s decision to include an “attended” (Geisler et al. 1985) or “supporting” noun phrase remains uncertain. This is the difference between spelling out the referent (3) or not (4):

(3) This connection between interfaces and optionality in L1 and L2 bilinguals suggests that whenever there is more than one language at stake, there are tangible consequences to the representation of the grammar of speakers. [App ling, 2015]

(4) At the operation bandwidth where S11 is small, transmission coefficient and coupling are very close, while out of this band the coupling level is higher than S21. This demonstrates that for antennas with poor matching S21 is not a good measure of the potential coupling between antennas. [Elec eng, 1965]

For some observers, a singular demonstrative pronoun has a restricted referential capacity and is a poor choice to refer to entire contexts or implied referents (e.g. Hinkel, 2004). In such cases a demonstrative determiner followed by an appropriate noun is needed. Finn (1995), on the other hand, argues that the redundancy of an unnecessary NP slows down the flow of new information to the reader. There is obviously a tradeoff here between snappy economy and ponderous clarity, with the noun phrase reducing the chance of ambiguity. In this example, for instance, it is unclear whether *this* refers to the entire preceding proposition or only to the information in the second clause:

(5) As a result, the announcement that the game will end at some specific point in the future immediately transforms universal defection into the sole subgame-perfect equilibrium (simply by changing expectations). This is true regardless of how far off in the future the end of the game is. [Soc, 1985]

Swales (2005) found that around half of the uses of the singular proximal demonstrative in the Hyland corpus⁵ occur in finite clause subjects, thus confirming the widespread cohesive use of this device so that Swales and Feak (1994: 21) talk of its role in facilitating ‘flow’ or “moving from one statement in

⁵ A corpus of 240 research articles from 8 disciplines

the text to the next”. It does, however, also suggest a more informal or engaging rhetorical stance, assuming that the reader is following the writer’s argument and is looking at things from the same perspective, thus producing an “impression of closeness and solidarity between reader and writer” (Mauranen, 1992: 243). However, by omitting a nominal after a demonstrative, authors may be missing a rhetorical trick. While Swales (2005) shows that the majority of attended NPs are metadiscoursal, referring either to the article itself (*study/ article/paper/research*) or the method (*method, technique, procedure*), many are more interpretive, offering a gloss on the referenced text. As in these examples, a summarizing noun can send a strong attitudinal signal to the reader:

- (6) This defect in the cortical actin cytoskeleton gives rise to ... [Bio, 2015]
- (7) This unpredictability of drinking violence... [Soc, 2015]
- (8) Those striking contradictions opened up the possibility of conflict. [Soc, 1985]
- (9) These dominant discourses can be seen in the ... [App ling, 2015]
- (10) This strong conclusion points to a novel response. [Elec eng, 2015]

More important than appearing ‘informal’ to readers is the need to persuade them of the veracity of a given view and, as Giesler et al. (1985: 151) observe, “writers should make sure that opportunities for useful emphasis or additional characterization have not been lost by leaving a *this* unattended”.

Sentence initial conjunctions and conjunctive adverbs also reflect the spoken language characteristics of fragmentation and real time unplanned production (e.g. Chafe, 1986). But while style guides generally recommended authors to avoid them, some examples have, like first person pronouns, been "legitimized" in English academic writing. This is shown in Table 5.

Table 5: Initial conjunctions and conjunctive adverbs by discipline per 10,000 words (2015).

	Applied Ling	Sociology	Biology	Elec Eng
illicit initial	37.9	49.0	40.1	34.7
initial <i>and</i>	1.2	1.9	0.6	0.0
initial <i>but</i>	1.4	2.8	1.3	1.4
initial <i>so</i>	1.2	1.3	0.0	1.1
initial <i>or</i>	0.2	0.5	0.0	0.0
initial <i>however</i>	8.5	8.0	9.4	4.8
initial <i>also</i>	1.1	0.3	0.4	1.0
initial <i>thus</i>	3.4	4.7	3.9	4.3
initial <i>yet</i>	0.4	2.7	1.2	0.0

initial <i>indeed</i>	1.1	1.2	2.4	0.9
initial <i>again</i>	0.8	0.2	0.1	0.3

Overall these features show an average of 40.4 items per 10,000 words. *But, thus and however*, are particularly common in the corpus with sociologists being the most enthusiastic users and electrical engineers the least. Using a regular expression query we were able to identify a much larger variety of initial conjunctions and conjunctive adverbs than Chang and Swales (1999) original study of informal elements, finding items such as *also, thus, yet* and *again* appearing frequently. Moreover, the list above is not exhaustive and our search revealed a number of other, less common, 'forbidden' initials such *still* and *next*. The following examples illustrate something of the informal stylistic effect of these features

(11) But the paper then contextualises the early studies from the standpoint of my current work with Colleagues on ageing and discourse. [App ling, 2015]

(12) So the assert-the-stronger maxim fails. [Soc, 2015]

(13) Again, whether the shift in mycorrhizal fungal communities is a cause or a symptom of the loss of plant species diversity in grasslands is unresolved. [Bio, 2015]

Diachronically, we found considerable increases in the frequency of these 'informal' initial conjunctions/conjunctive adverbs with a rise of 50% since 1985, due largely to increases in the use of initial *however, so* and *indeed* (see Table 6). Applied linguistics, sociology and electrical engineering had increases of between 24-28% although biology, once again the outlier, recorded a smaller increase of 15.5% (although it has a higher frequency of 9.4 per 10,000 words than the other three disciplines in 2015). This may, speculatively, illustrate something of the distinctive ways that biology pursues and argues problems (Chargaff, 1974; Hyland, 2004) and the developed sense of rhetorical awareness among many of its writers (Halloran, 1984), perhaps extending to issues of 'correctness' in stylistic choices.

Table 6: Sentence initial conjunctions and conjunctive adverbs 1965-2015 (per 10,00 words)

	1965	per 10,000	1985	per 10,000	2015	per 10,000
illicit initial	1661	27.8	2017	28.7	4049	42.9
initial <i>and</i>	87	1.5	91	1.3	95	1.0
initial <i>but</i>	221	3.7	187	2.7	177	1.9
initial <i>so</i>	19	0.3	25	0.4	89	0.9
initial <i>or</i>	7	0.1	13	0.2	18	0.2
initial <i>however</i>	239	4.0	285	4.1	760	8.1
initial <i>also</i>	41	0.7	57	0.8	70	0.7
initial <i>thus</i>	242	4.1	281	4.0	409	4.3
initial <i>yet</i>	46	0.8	74	1.1	108	1.1
initial <i>indeed</i>	23	0.4	54	0.8	137	1.5
initial <i>again</i>	21	0.4	29	0.4	35	0.4

The large increases in initial *however* and declines in *and* and *but* were overwhelmingly accounted for by changes in applied linguistics and sociology. Sociology also showed a large increase in *yet* and biology in the use of *indeed*, although none of these features are highly frequent. Overall, it seems that initial *however* has become a legitimate, indeed commonplace, marker of contrast/adversity across all disciplines and that the need for clear and unambiguous assertion has been filled by the marker *indeed*, especially in applied linguistics and biology.

Second person pronouns is the final feature we want to mention here, and once again, its widespread use in everyday conversation invests it with an informality that academic writers are advised against. Second person *you* and *your* are perhaps the most obvious way of referring to readers and it is 25 times more common in conversation than in academic writing (Biber et al, 1999: 334). This is because it carries certain interactional risks, for by explicitly acknowledging the reader's presence writers are seeking to engage them more closely and create a sense of intimacy, attempting to get them on board and carry them along with an argument. This is not, however, always welcome by interlocutors quickly scanning for bottom line results relevant to their own research.

The second person pronoun is used by authors to refer to readers either as individual or general referents (Biber et al., 1999: 328-331). So, in some cases we see readers spoken to personally as participants in the text:

(14) If we are to disagree, what I say or assert must be what you deny. But this does not commit one to the existence of entities beyond space and time, only to the conceptual truism that if I deny what you assert, and you assert that p, then I deny that p. [Soc, 1965]

(15) You were recently in the market for a new car and, after narrowing your search, chose the Saab 900S over the Volvo 850 and the Pontiac Bonneville. The Saab performs in accordance with your expectations. On the way to dinner one evening, your wife mentions that a car buff co-worker told her that the Volvo 850 is the best designed car he has seen. How do you feel about your Saab now? [Soc, 1965]

In other cases, writers use *you* with a wider semantic reference, referring to people in general, like the indefinite pronoun *one*, rather than with specific discourse participants:

(16) You can see this most clearly with respect to the question of ecology. [Bio, 2015]

(17) The premise upon which Steinberg's reasoning is based is clear. You know you have a good definition of a picture if you can use it to explain what a picture is to a congenitally or early blind person. [Soc, 2015]

Here, *you* carries a more encompassing meaning than rhetorically focusing on an individual reader, seeking instead to engage with an audience by recruiting its members into a world of shared experiences.

The authors analysed all occurrences of second person to determine the referent (agreement of 92%) and, as Table 7 shows, it is general reference which dominates the frequencies and which has seen the largest increase in the past 50 years, growing by 24% despite there being only 3 instances per 10,000

words in 2015. Interestingly, more specific reference, which is more conversational and so might be seen as carrying more personal and informal meanings, has seen no overall increase.

Table 7: Changes in Second person pronoun use by referent (per 10,000 words)

2 nd person reference	1965	1985	2015
general reference	2.5	2.3	3.1
specific reference	0.8	0.7	0.9

Second person remains unknown in electrical engineering and biology texts and has fallen by half in sociology, while doubling in applied linguistics to just one per 10,000 words. Clearly, norms of interpersonal engagement seem to be changing only slowly in the use of this feature, underlining something of a reluctance by writers to involve their readers more directly in their arguments.

Discussion and conclusions

Overall, our findings suggest that the answer to the question posed in our title is ‘it depends’. While academic writing is becoming more informal, this is by small margins and depends on the discipline and features being considered. We have studied only a sample of four disciplines and there is certainly room for further work exploring trends in other fields, especially in the area of how informality manifests itself through lexis, which we have not had space to explore. Even in our corpus, the process of change is not occurring in all disciplines or in all features. Broadly, our selected ‘informal elements’ show a small rise of 2% since 1965 when normed for the increase in published words, although this is mainly due to increases in first person pronouns, unattended reference and sentences beginning with conjunctions and conjunctive adverbs. Moreover, it is the science and engineering disciplines in our sample which have become less formal while trends for the social science fields indicate a reduction in the use of these informal features.

It is also worth pointing out that despite this small increase in informality, the proportion of nouns and adjectives in relation to verbs and pronouns, suggested by Heylighen and Dewaele’s (1999) as markers of relative formality, is actually increasing. Some readers may want to see this as

representing an underlying constant of formality in academic writing, while others might regard it as indicating a necessary precision and informativity in scholarly writing. We would simply caution against strong claims in this regard and emphasize that any trend towards informality is not cut and dried.

Biber and Grey (2016), moreover, observe an inclination during the 20th century towards more compressed, phrasal expressions over-elaborated, clausal expressions in academic writing which, while allowing for faster, more efficient processing by expert readers are more opaque to novices. Although our research suggests that we are witnessing a gradual move towards greater use of some salient informal features in academic writing, it is important not to overestimate either the extent or speed of the change. Genres are supported by powerful interests, not least by academics who have invested years into acquiring their arcane conventions, and so change only slowly. They also change for reasons which more directly relate to the rhetorical purposes of the genre and gradual adjustments to norms of interpersonal persuasion rather than efforts to weaken existing structures in favour of more 'friendly', relaxed' or conversational practices.

The slow increase in the use of informal features in academic writing is, we believe, related to changing rhetorical practices which support greater efforts to involve readers in arguments and secure support for their claims through positioning themselves more explicitly in relation to their ideas and readers. It is, then, part of a wider shift seen in many domains toward more engaged and intimate forms of interaction. Fairclough (2001) has discussed this in relation to the 'synthetic personalization' of texts from the worlds of commerce and media to give the impression of treating mass audiences as individuals. However, this seems to be less to do with overt manipulation than an attempt to embrace a more direct and egalitarian relationship to achieve recognisable persuasive goals. We see here changing textual practices for producing agreement in published academic writing.

The expression of opinions and assessments is nothing new in academic discourse but a ubiquitous feature of all human interaction. Despite its apparently impersonal facade, the evaluation and interpretation of methods, sources, claims and data underpin the construction of knowledge and the

projection of competent scholarly identities in professional academic writing (e.g. Hyland, 2004). Academic texts, however, are only persuasive to the extent they employ the rhetorical conventions that colleagues find convincing and conventions of detachment have often overridden unequivocal interpersonality to emphasize explicitness, objectivity and precision. However, the ways writers seek to claim solidarity with readers, evaluate ideas and acknowledge alternative views seems to be changing to accommodate more obvious interpersonal interactions. Impersonality in the sciences seems to be a growing rhetorical option in this genre and at least some biologists and electrical engineers are taking advantage of this.

One consequence of our results, however, is the potential difficulties this rhetorical change creates for students and novice writers, particularly those writing in a second language. In their study of these features, for example, Chang and Swales (1999) found the majority of their L2 graduate sample were uneasy about the increased flexibility that greater informality might offer. They believed that learning the rules of formal academic English was already a considerable burden and was further complicated by changing styles which allows them to mix formal and informal language effectively. It is true that some studies have shown that L2 English students often make considerable use of some informal features, so that Cobb (2003) and McCrostie (2008) found substantially greater use of first and second person pronouns compared with expert prose and Leedham (2015) found Chinese student writers made substantial use of contractions and informal connectors such as *besides*, *lots* and *what's more* in their essays. Whether these choices simply mimic conversational practices due to limited proficiency is uncertain but it is generally frowned upon by those assessing their writing. So while informality may seem a boon to experienced writers, it may create additional complexities in the relationships the writer is seeking to build with readers and further increases to the compositional burden of novices.

We are aware, of course, that research articles are carefully refined and polished and so a potential drawback of examining a published corpus is the influence exercised by editors, copy editors and reviewers on the style of the eventual finished text. Certainly research articles represent the hard case when considering the spread of informality in written English. But if there is a sea change in such a

rhetorically machined genre, then presumably many community gatekeepers are also subscribing to it, or at least hesitating to require such changes to be corrected. Ultimately, of course, research articles are the sites where academics negotiate and make sense of the issues which preoccupy them as members of particular disciplines. These sites, moreover, are not merely storehouses of arcane, abstract practices, monolithic and forever frozen in time, but responsive to changing contexts and the demands of new conditions. In some fields, these changes are happening before our eyes as scholars make choices to adopt rhetorical options which seek to establish a more engaged and personal connection with their readers.

References

- Adel, A. (2008). Metdiscourse across three varieties of English: American, British and advanced learner English. In U. Connor, E. Nagelhout & W.V. Rozycki (Eds.) *Contrastive Rhetoric: Reaching to Intercultural Rhetoric*. Amsterdam; Benjamins.
- Anthony, L. (2011). *AntConc 3.4.3*. <http://www.laurenceanthony.net/software.html>
- Atkinson, D. (1999). *Scientific discourse in sociohistorical context: The philosophical transactions of the Royal Society of London, 1675-1975*. Mahwah, NJ: Lawrence Erlbaum.
- Bennet, K. (2009). English academic style manuals: a survey. *Journal of English for Academic Purposes*, 8, 43-54.
- Biber, D. (1988) *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.
- Biber, D. (1995) *Dimensions of Register Variation. A Cross-Linguistic Comparison*. Cambridge University Press, Cambridge.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow: Pearson.
- Biber, D., & Gray, B. (2016). *Grammatical complexity in academic English: Linguistic change in writing*. Cambridge: Cambridge University Press.
- Brown, P. & S. Levinson (1978). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.

- Chafe, W. (1986) "Writing in the perspective of speaking." In Cooper, C. and Greenbaum, S. (eds.), *Studying Writing: Linguistic Approaches*. London: Sage Publications.
- Chang, Y-Y & Swales, J. (1999). Informal elements in English academic writing: Threats or opportunities for advanced non-native speakers? In Candlin, C. & Hyland, K. (eds). *Writing: texts, processes and practices*. London: Longman.
- Chargaff, E. 1974. Building the Tower of Babel. *Nature* 248: 776-779
- Cobb, T. 2003. Analyzing later interlanguage with learner corpora: Quebec replications of three European studies. *Canadian Modern Language Review* 59(3): 393–423.
- Coffin, C., Curry, M., Goodman, S., Hewings, A., Lillis, T., & Swann, J. (2003). *Teaching academic writing: A toolkit for higher education*. London: Routledge.
- Fairclough, N. (2001) *Language and Power*. 2nd ed. Essex: Longman.
- Finn, S. (1995). Measuring effective writing: Cloze procedure and anaphoric "this". *Written Communication* 12. 240–266.
- Foster, J. (2005). *Effective writing skills for public relations*. 3rd ed. London: Kogan Page
- Geisler, C., Kaufer, D.S., and Steinberg, E.R. (1985) "The unattended anaphoric 'This': When should writers use it?" *Written Communication*, 2: 129-55.
- Gilbert, G. N. and M. Mulkay. (1984). *Opening Pandora 's box: a sociological analysis of scientific discourse*. Cambridge: CUP.
- Gray, B. (2010). On the use of demonstrative pronouns and determiners as cohesive devices: A focus on sentence-initial this/these in academic prose. *Journal of English for Academic Purposes*, 9, 167-183.
- Guide to writing in history and classics* (2015).
<http://www.usu.edu/markdamen/WritingGuide/00intro.htm>
- Halliday, MAK (1985). *Spoken and written language*. Oxford: Oxford University Press
- Halliday, MAK (1998). Things and relations: regrammaticising experience as technical knowledge. In Martin, J. & Veel, R. (eds.) *Reading science*: London: Routledge. 185-235.
- Halloran, S. (1984). The birth of molecular biology: An essay in the rhetorical criticism of scientific discourse. *Rhetoric Review*, 3 (1), 70-83.

- Heylighen, F. & Dewaele, J-M (1999). *Formality of Language: definition, measurement and behavioral determinants*. Internal Report, Center "Leo Apostel", Free University of Brussels.
- Hundt, M. & Mair, C. (1999). "Agile" and "Uptight" Genres: The Corpus-based Approach to Language Change in Progress. *International Journal of Corpus Linguistics*. 4:2: 221–242
- Hinkel, E. (2004). *Teaching academic ESL writing*. Mahwah: NJ: Earlbaum
- Hyland, K. (2001). Humble servants of the discipline? Self-mention in research articles. *English for Specific Purposes* 20 (3), 207-226.
- Hyland, K. (2004) *Disciplinary Discourses: social interactions in academic writing*. Ann Arbor: MI: University of Michigan Press
- Hyland, K. (2012) *Disciplinary Identities*. Cambridge: Cambridge University Press
- Ivanic', Roz, 1998. *Writing and Identity: The Discoursal Construction of Identity in Academic Writing*. Benjamins, Amsterdam.
- Lachowicz, D. (1981). On the use of the passive voice for objectivity, author responsibility and hedging in EST. *Science of Science*, 2_(6), 105-115.
- Leech, G., Hundt, M, Mair, C & Smith, N. (2009). *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.
- Leedham, M. (2015) *Chinese Students' Writing in English: Implications from a Corpus-driven Study*. Oxford: Routledge.
- Mair, C. 1998. Corpora and the Study of the Major Varieties of English: Issues and Results. In *The Major Varieties of English: Papers from MAVEN 97, Växjö*, ed. by H Lindquist, S Klintborg, M Levin & M Estling, 139–157. Växjö: Edman and Westerlunds Tryckeri.
- Mair, C. & Leech, G. (2006). Current Change in English Syntax. In *The Handbook of English Linguistics*, ed. by B Aarts and A MacMahon, 318–342. Oxford: Blackwell.
- Mauranen, A. (1992). *Cultural differences in academic rhetoric: A contrastive study of Finnish and English writing*. Nordic Research on Text and Discourse. NORDTEXT Symposium (Espoo, Finland, May 10-13, 1990): 237-250.

- McCrostie, J. (2008). Writer visibility in EFL learner academic writing: A corpus-based study. *ICAME Journal*, 32, 97-114.
- Nash, W. (1986) *English Usage: a Guide to First Principles*. London: Routledge & Kegan Paul.
- Nelson, N., & Castello, M. (2012). Academic writing and authorial voice. In: Castelló, M., & Donahue, C. (Eds.) *University Writing: Selves and Texts in Academic Societies*. Emerald, Bingley UK, pp.33-52.
- Online Etymology Dictionary* (2010). Retrieved March 31, 2016 from *Dictionary.com* website <http://www.dictionary.com/browse/in->
- Seone, E. (2013). On the conventionalisation and loss of pragmatic function of the passive in Late Modern English scientific discourse. *Journal of Historical Pragmatics* 14 (1), 70-99
- Seone, E. & Loureiro-Porto, L. (2005). "On the colloquialization of scientific British and American English". *ESP Across Cultures* 2: 106-118
- Strauss, S. (1993) Why 'this' and 'that' are not complete without 'it'. In K. Beals, G. Cooke, D. Kathman, S. Kita, K. McCulloch, D. Testen (eds.), *CLS 29: Proceedings of the 29th Regional Meeting of the Chicago Linguistic Society* (pp. 403-417). Chicago IL: Chicago Linguistic Society.
- Swales, JM (2005). Attended and Unattended "this" in Academic Writing: A Long and Unfinished Story. *ESP Malaysia*. 11: 1-15.
- Swales, J.M. and Feak, C.B. (2012) *Academic Writing for Graduate Students: A Course for Nonnative Speakers of English 3rd ed.* Ann Arbor, MI: University of Michigan Press.
- Wulff, S., Romer, U. & Swales, J. (2012). Attended/unattended *this* in academic student writing: Quantitative and qualitative perspectives. *Corpus Linguistics and Linguistic Theory* 8 (1): 129-157

Appendix 1: Journal list

Applied Linguistics

TESOL Quarterly (1967-)

Language Learning (1948-)

Foreign Language Annals (1967-)

Modern Language Journal (1916-)

College Composition and Communication (1950-)

Sociology

American Journal of Sociology (1895-)

Social problems (1953-)

The British Journal of Sociology (1950-)

American Journal of Economics and Sociology (1941-)

The Sociological Quarterly (1960-)

Biology

The Quarterly Review of Biology (1926-)

Biological Reviews (1923-)

Radiation Research (1954-)

BioScience (1964-)

The Journal of Experimental Biology (1923 -)

Electrical Engineering

Proceedings of the IEEE (1963 -)

Automatica (1963 -)

IEEE Transactions on Automatic Control (1963 -)

IEEE Journal of Solid-State Circuits (1966 -)

IEEE Transactions on Information Theory (1963 -)