Stance adverbials in research writing

Matthew Peacock

Department of English, City University of Hong Kong (China)

enmatt@cityu.edu.hk

Abstract

This paper describes an analysis of eight categories of stance adverbials, for example "definitely" and "usually", in a corpus of 600 research articles (RAs) across 12 disciplines, six science and six non-science. Stance adverbials may play an important role in the key RA functions of putting forward claims and propositions. However, there has been very little previous research in the area. A new list of stance adverbials was developed and frequency, function and disciplinary variation were examined using WordSmith Tools. Stance adverbials in two categories, Limitation and Doubt and Certainty, were much more prevalent than hitherto suspected. Numerous statistically significant disciplinary differences, for example between the sciences and non-sciences, were also found. A closer examination of science RAs was undertaken. Authors were found to develop claims in a different way, putting greater weight on methods and procedures, while non-science authors tended more towards discursive argument. The techniques of semantic preference, the creation of meaning through multiple occurrences of collocates (Hunston, 2007), were also adopted to further examine function. Conclusions are that stance adverbials play an important role in the construction of stance in RAs, though this differs by discipline. Finally, semantic preference techniques may be a valuable method for corpus-based research.

Keywords: stance adverbials, corpus analysis, genre analysis, research articles.

Resumen

Expresiones adverbiales epistémicas en artículos de investigación

En este trabajo se analizan ocho categorías de expresiones adverbiales epistémicas, por ejemplo "definitely" y "usually", en un corpus de 600 artículos de investigación pertenecientes a 12 disciplinas, seis de ciencias y seis no de

ciencias. Las expresiones adverbiales epistémicas pueden desempeñar un importante papel en funciones clave en los artículos de investigación tales como presentar argumentos y proposiciones. Sin embargo, este tema ha sido escasamente investigado. Tras obtenerse una nueva lista de expresiones adverbiales epistémicas, se examinó su frecuencia y variación según la disciplina utilizando WordSmith Tools. Las expresiones adverbiales epistémicas pertenecientes a dos categorías, Limitación y Duda/Certeza, resultaron ser mucho más frecuentes de lo que se había sospechado. Se encontraron también numerosas diferencias estadísticamente significativas dependiendo de la disciplina, por ejemplo entre los artículos de ciencias y los de disciplinas no científicas. Se llevó a cabo un análisis más detallado de artículos científicos, descubriéndose que sus autores presentaban sus argumentos de forma diferente, poniendo más énfasis en los métodos y procedimientos, mientras que los autores en disciplinas no relacionadas con las ciencias tendían más hacia la argumentación discursiva. También se adoptaron técnicas de preferencia semántica, de creación de significación por medio de múltiples ocurrencias de colocaciones (Hunston, 2007) como método adicional de análisis funcional. Se concluye que las expresiones adverbiales desempeñan un importante papel en la construcción de una posición epistémica en los artículos de investigación, aunque existen diferencias según la disciplina. Además, las técnicas de preferencia semántica pueden resultar útiles en proyectos de investigación basados en el análisis de corpus.

Palabras clave: expresiones adverbiales epistémicas, análisis de corpus, análisis de género, artículos de investigación.

Introduction

Stance adverbials, for example "clearly" and "generally", may play an important role in expressing and constructing epistemic stance. The aims of this study were to develop a more comprehensive list of stance adverbials and to investigate this role in research articles (RAs), along with their form, frequency, function, distribution, and discipline variation in a corpus of 600 research articles across 12 disciplines, six science and six non-science: Biology, Business, Chemistry, Computer Science, Economics, Environmental Science, Language and Linguistics, Law, Neuroscience, Physics and Materials Science, Psychology, and Public and Social Administration. The RA was chosen for this research because of its importance for the dissemination of knowledge. Williams (2002: 45) says discourse communities develop systems for communication through the use of patterns and these, rather than individual words, are among their essential

attributes. RAs have been called a vital channel for substantiating findings and disciplines (Hyland, 1996: 252), and the preferred method for communication among discourse communities (Williams, 1998: 153). Their language defines these communities.

Stance is how "writers present themselves and convey their judgments, opinions and commitments" (Hyland, 2005: 176). Sancho-Guinda and Hyland (2012: 1) add that it is how writers "appropriately engage with readers". The language used to achieve this function has had different names historically: Halliday (1993: 107) refers to "modality" such as "I think that" and "I doubt that" and also to "attitude", Thompson and Hunston (2000: 1) to "evaluation" which expresses positive or negative views and frequently supports claims, and McGrath and Kuteeva (2012: 162-163) to "evaluative language" which communicates attitudes regarding the reliability and impact of results. Stance adverbials are items which express stance. Silver (2003: 372) notes that they function to construct knowledge claims and a "writer's professional persona". In this paper Biber's (2006: 99) definition is used: items which express attitude or assessment towards a proposition. Epistemic stance is defined as the expression of commitment to the truth of a proposition presented by the writer (Hyland, 1999: 101). It is represented by "linguistic mechanisms used by speakers and writers to convey their…personal feelings, attitudes, value judgments, or assessments" (Biber, 2006: 97-98). The following extract from a Law RA in the corpus shows an example: "Violence by young people *generally* does not involve the use of knives". "Generally" here expresses the author's attitude or stance towards the proposition – in this case, an assessment of a limitation of the proposition. Suggestions, claims, and propositions are an important part of the RA: and in RAs the construction and expression of epistemic stance is part of the important function of claiming, confirming, and expressing membership of and position in the discourse community of peers, academics, and other researchers, and therefore in constructing identity.

Biber et al. (1999: 762, 875-881; also see Conrad & Biber, 2000) indicate that adverbials fall into three different classes, all of which are important cohesive devices: (i) circumstance (e.g. "nowadays"), which add circumstantial information about propositions in clauses; (ii) Stance; (iii) Linking (e.g. "however", "therefore"), which make the relationship between two units of discourse clear. They note (pages 854-857) that stance adverbials fall into three major categories – epistemic, attitude, and style:

A. Epistemic. These "convey one of the following six major areas of meaning":

1. Doubt and Certainty (e.g. "perhaps", "definitely"): judgments of certainty, or level of probability.

2. Actuality and Reality (e.g. "actually"): the status of a proposition as real life fact.

3. Source of Knowledge/Allude to Evidence (e.g. "according to"): the source of information reported in a proposition.

4. Limitation (e.g. "generally"): the limitation of a proposition.

5. Viewpoint or Perspective (e.g. "in our view"): the viewpoint or perspective from which a proposition is true.

6. Imprecision (e.g. "kind of"): these mark a proposition as being imprecise.

B. Attitude (e.g. "fortunately"): these convey an evaluation or value judgment towards a proposition.

C. Style (e.g. "frankly"): these comment on the style or manner of conveying propositions.

It is noted that there is potential functional overlap between "Doubt" in category 1 and category 4, "Limitation": for example, "perhaps" could function either to express probability or to limit a proposition. Biber et al. (1999) did not discuss this.

A careful literature search found only two previous empirical studies on stance adverbials. First, Biber et al. (1999: 859-860) examined usage in four registers in the *Longman Spoken and Written English* (LSWE) corpus: conversation, news articles, fiction, and academic prose (book extracts plus RAs, 2.6 million words each). Academic prose consisted of seventy-five book extracts, mostly technical trade books, from thirteen different disciplines, and RAs from fifteen different disciplines. Stance adverbials were most common in conversation, followed by academic prose, then fiction, and news. They do not give exact figures, but their bar charts show a few frequency (pmw/per million words) results for major categories and individual items. They say (pages 860, 868) that epistemic markers (all six areas combined) were "surprisingly" common, resulting from the importance of showing the doubt or certainty of propositions and

constructing epistemic stance. (Conrad & Biber, 2000: 65 add to this that academic authors "pay considerable attention to certainty, actuality, and imprecision"). Second, Biber (2006) presents corpus results for stance adverbials in two corpora, textbooks across five disciplines (760,600 words) and "written course management" (course packs and course management, 159,600 words), though he does not give exact figures. Table 1 shows all these findings:

Category	Biber et al. (1999): LSWE Academic Prose	Biber (2006)		Biber et al. (1999)	
		Textbooks	Course Management		
Doubt and Certainty		1950	1300	probably	200
				maybe	> 50
				perhaps	300
				of course	200
				certainly	100
				definitely	> 50
				really	100
Actuality and Reality	3600	--	--	actually	100
				in fact	100
Source of Knowledge/ Allude to Evidence		--	--	according to	100
Limitation		--	--	generally	200
Viewpoint or Perspective		--	--	--	--
Imprecision		--	--	like	> 50
				sort of	> 50
				kind of	> 50
Attitude	350	150	150	--	--
Style	100	700	350	--	--

Table 1. Previous empirical findings: frequency pmw.

Stance adverbials may play an important role in epistemic stance and in the very important RA functions of putting forward suggestions, claims, and propositions, and claiming, confirming, and expressing membership of and position in discourse communities. They would therefore be valuable persuasive devices and an important part of research writing including RAs, and worth investigating further. Yet very little previous research seems to have investigated discipline variation in the area. Also, if stance adverbials are important, they must also be acquired by aspiring NNS research writers. Bhatia (2000: 147) says a strong justification for genre research is that it informs the teaching of research writing.

Research method

The aims of this study were to extend previous research on the form, frequency, function, and distribution of stance adverbials in RAs across twelve disciplines, and to develop a more comprehensive list of stance adverbials.

Research aims and research questions

The aims of this research were, within the corpus, to:

1. Build up a list of stance adverbials in the three target categories.

2. Investigate the frequency of all stance adverbials in the target categories.

3. Investigate disciplinary variation.

4. Investigate function.

The following questions are directly addressed:

1. How frequently do RA authors use stance adverbials across a range of disciplines? Are there any interdisciplinary differences?

2. What stance adverbials do RA authors use across a range of disciplines? Are there any interdisciplinary differences?

3. How do stance adverbials function across a range of disciplines?

The RA corpus

The corpus included 600 RAs published from 2000-2008, 50 from each discipline – see Table 2 below. These disciplines were chosen because they represent a variety of subjects and also have large numbers of research writers, mostly NNS, around the world. Disciplines were classified as science or non-science by asking experts to discuss the classification. Only Economics and Psychology caused any controversy, although the experts did classify both as non-sciences. Leading journals were chosen from each discipline (see Appendix 1). Visits were made to the pertinent departments and two sources from each asked to name key journals in their area:

Discipline	Number of RAs	Total Word Length
Biology	50	286,440
Business	50	329,599
Chemistry	50	182,472
Computer Science	50	359,003
Economics	50	364,710
Environmental Science	50	249,874
Language and Linguistics	50	320,847
Law	50	372,748
Neuroscience	50	303,098
Physics and Materials Science	50	226,253
Psychology	50	381,709
Public and Social Administration	50	306,624
ALL TWELVE DISCIPLINES	600	3,683,377

Table 2. Lengths of disciplinary corpora.

RAs were randomly chosen from each journal by numbering them and drawing numbers from a box. No distinction was made between native- and non-native writers. Only empirical data-driven RAs with an Introduction-Method-Results-Discussion format were chosen, following Hyland (1998: 97), who notes that this is an important genre. Discussions and RAs by writers previously chosen were not used. It is suggested that the disciplinary corpora are adequately representative because of their size and because of the use of discipline informants to select journals.

Investigating the corpus

Analysis was done in these steps, using the Concord and Contexts functions of WordSmith Tools 4.0 (Scott, 2004). Explanation of certain steps follows:

1.

(a) To build up a preliminary list of stance adverbials, we turned first to the standard work on the topic, Biber et al. (1999: 853-875), who list 78.

(b) A further 80 stance adverbials were then identified from other grammars, thesauruses, and the RAs themselves, for a total of 158. This is more extensive than previously published lists.

(c) Biber et al. (1999: 857-858) call two of their items ambiguous, "really" and "in fact". "Really" functions to construct epistemic stance only with the meaning "in reality" or "in truth", for example "*Really* you've noticed the difference?" "Really" is not a stance adverbial when it functions as an intensifier, for example

"It's a *really* wonderful day". "In fact" is a linking adverbial, not a stance adverbial, when it connects a proposition to a preceding sentence, e.g. "I was out for hours yesterday. *In fact* I was very busy".

2. A preliminary examination of the corpus was conducted to check the function of all 158 stance adverbials, to see whether they always function as stance adverbials or not. This was done by individually checking a large number of occurrences of the adverbials in each discipline corpus and evaluating function by reading the relevant sentence and surrounding sentences. 118 of the 158 appear in the corpus. After a careful examination of the functions of all 118, it was found that seven, in addition to "really" and "in fact", do not always function as stance adverbials:

(i) "Clearly": Functions as a Doubt and Certainty stance adverbial only with the meaning "obviously", not with the meaning "apparent" or "easily perceived", for example "as *clearly* seen in Table 2".

(ii) "About": Functions as an Imprecision stance adverbial only with the meaning "approximately", not "on the topic of".

(iii) "In short": Only functions as a stance adverbial in sentence initial position.

(iv) "Absolutely": Functions as a Doubt and Certainty stance adverbial only with the meaning "definitely".

(v) "Kind of" and "sort of": Function as an Imprecision stance adverbials only with the meaning "approximately", not "type".

(vi) "Indeed": Functions as a Doubt and Certainty stance adverbial only with the meaning "without a doubt". With the meaning "in reality", it functions as an Actuality and Reality stance adverbial.

3. The frequency of all stance adverbials was checked, along with disciplinary variation. All cases of the nine items above which do not always function as stance adverbials were excluded from the count. This required manual examination of every occurrence of each using the Concord function.

4. The function of every occurrence of all stance adverbials was individually checked by reading the relevant sentence and surrounding sentences.

5. Statistical significance was set at p<.05 and was tested with the log-likelihood calculator.

6. For this research, high-frequency is defined as 40 pmw or higher, following Biber, Conrad and Cortes (2004: 376): "we take a conservative approach…[a] frequency cut-off of 40 times per million words to be included in the analysis".

Regarding steps (2) and (3), the corpus was split into discipline corpora as required to examine discipline variation. Individual manual examination of the function of all occurrences is crucial.

Two evaluators were involved in step (4): this writer and a local university lecturer. The second coder independently evaluated the function of every occurrence in order to measure inter-rater agreement. This writer reassessed the function of every occurrence after one month in order to measure intra-rater agreement. Inter-rater agreement was 97%, rising to 100% after discussions. Intra-rater agreement was 99%. Both were calculated by measuring correlations between the results using Statistical Package for the Social Sciences.

Results

The results for categories and individual items can be seen below in Tables 3 to 7. Totals include all stance adverbials, though only items with a frequency of 20 pmw or higher are shown. Asterisks mark statistically significant differences – bold significantly higher, italics significantly lower:

Function	All 12 disciplines	Business	Economics	Language and Linguistics	Law	Psychology	Public and Social Administration
Doubt and Certainty	825	707*	698*	1417*	1202*	817	926
Actuality and Reality	414	397	407	595*	757*	565*	452
Source of Knowledge/ Allude to Evidence	375	295*	352	478*	260*	425	298*
Limitation	1476	1627*	1186*	2094*	1717*	1209*	1622*
Viewpoint or Perspective	6	4	10	6	0*	20*	0
Imprecision	208	190	236	293*	276*	154*	307*
Attitude	141	96*	157	182*	204*	163	181
Style	47	24*	116*	41	102*	72	36
ALL STANCE ADVERBIALS	3493	3339	3163*	5106*	4518*	3424	3821*

Function	All 12 disciplines	Biology	Chemistry	Computer Science	Environmental Science	Neuro-science	Physics and Materials Science
Doubt and Certainty	825	601*	573*	643*	602*	699*	663*
Actuality and Reality	414	312*	137*	308*	113*	379	291*
Source of Knowledge/ Allude to Evidence	375	378	451	367	276*	424	486*
Limitation	1476	1381	858*	1458	1705*	1197*	1265*
Viewpoint or Perspective	6	0	0	3	0	25*	0
Imprecision	208	142*	97*	262*	169	119*	174
Attitude	141	71*	94*	192*	113	115	96*
Style	47	24	11*	24*	39	45	15*
ALL STANCE ADVERBIALS	3493	2909*	2221*	3258*	3016*	3003*	2989*

Table 3. Functional categories – discipline differences. Frequency pmw.

Item	All 12 disciplines	Business	Economics	Language and Linguistics	Law	Psychology	Public and Social Administration
clearly	154	144	147	201*	193	167	154
perhaps	114	126	92	228*	174*	127	163*
probably	111	64*	51*	180*	145	69*	99
of course	65	36*	89	141*	169*	59	45
possibly	65	46	82	90*	46	52	68
certainly	47	50	17*	80*	99*	69	59
obviously	43	28	31	82*	43	33	45
most likely	42	22*	27	16*	67*	65*	18*
presumably	38	18*	14*	59*	62*	69*	18
I / we think	33	36	14*	113*	21	20	145*

	All 12 disciplines	Biology	Chemistry	Computer Science	Environmental Science	Neuro-science	Physics and Materials Science
clearly	154	123	144	115	90*	128	195
perhaps	114	85	29*	66*	56*	111	15*
probably	111	128	119	112	163*	107	116
of course	65	9*	11*	59	39	16*	41
possibly	65	52	47	45	79	62	105*
certainly	47	19*	18*	31	34	58	9*
obviously	43	47	36	35	51	16*	55
most likely	42	47	65	66*	34	37	47
presumably	38	43	61	17*	0*	70*	20
I / we think	33	0*	0*	7*	0*	0*	0*

Table 4. Individual Doubt and Certainty Epistemic Stance Adverbials – Discipline Differences Frequency pmw.

Item	All 12 disciplines	Business	Economics	Language and Linguistics	Law	Psychology	Public and Social Administration
Actuality and Reality:							
indeed	163	120*	202	207*	255*	323*	104*
actually	124	114	106	201*	268*	118	158
in fact	70	82	55	86	140*	91	50
really	28	34	21	68*	30	13	90*
Source of Knowledge/Allude to Evidence:							
according to	325	277*	328	388*	201*	363	236*
apparently	42	22*	14*	70*	59	56	63
Viewpoint or Perspective:							
in our / my view	4	4	10	6	0	20*	0
Imprecision:							
about	168	174	137	260*	236*	131	239*
roughly	37	12*	92*	29	35	16*	59

Item	All 12 disciplines	Biology	Chemistry	Computer Science	Environmental Science	Neuro-science	Physics and Materials Science
Actuality and Reality:							
indeed	163	170	50*	101*	28*	169	137
actually	124	76*	25*	122	45*	115	58*
in fact	70	47	50	52	23*	53	64
really	28	0*	0*	21	17	12	6*
Source of Knowledge/Allude to Evidence:							
according to	325	307	418*	343	270	370	410*
apparently	42	61	25	24	0*	53	49
Viewpoint or Perspective:							
in our / my view	4	0	0	3	0	0	0
Imprecision:							
about	168	104*	72*	199	158	91*	137
roughly	37	38	25	59	11*	29	38

Table 5. Other Individual Epistemic Stance Adverbials – Discipline Differences. Frequency pmw.

Item	All 12 disciplines	Business	Economics	Language and Linguistics	Law	Psychology	Public and Social Administration
generally	222	186	178	258	453*	163*	194
typically	146	210*	171	178	142	134	86*
usually	146	158	89*	228*	121	95*	172
in general	125	138	113	137	169*	147	99
primarily	112	158*	51*	166*	94	121	185*
mainly	108	132	65*	141*	38*	82	117
frequently	105	124	58*	238*	150*	88	99
largely	92	56*	85	94	145*	72	126
commonly	75	58	96	96	83	59	108
essentially	58	70	34*	72	43	59	50
mostly	56	84*	68	107*	32*	33*	68
normally	47	46	58	55	40	20*	59
rarely	40	24	24	72*	70*	29	81*
predominantly	31	36	10*	45	19	33	54*
in most cases	22	24	14	14	21	29	41
basically	20	36*	14	49*	19	3*	18

Item	All 12 disciplines	Biology	Chemistry	Computer Science	Environmental Science	Neuro-science	Physics and Materials Science
generally	222	203	141*	227	293*	136*	206
typically	146	109	72*	182	62*	148	169
usually	146	104	119	136	236*	132	148
in general	125	66*	87	182*	141	37*	140
primarily	112	95	50*	63*	225*	82	41*
mainly	108	137	126	70*	191*	74	137
frequently	105	52*	32*	154*	68	66*	26*
largely	92	90	50*	87	169	115	47*
commonly	75	61	32*	94	45	91	64
essentially	58	58	54	56	62	29*	73
mostly	56	61	22*	21*	56	45	32
normally	47	175*	18*	14*	23	58	20*
rarely	40	24	7*	45	51	25	15*
predominantly	31	19	4*	14*	0*	53	70*
in most cases	22	33	18	17	23	12	26
basically	20	14	7	14	17	4*	20

Table 6. Individual Limitation Epistemic Stance Adverbials – Discipline Differences. Frequency pmw.

Item	All 12 disciplines	Business	Economics	Language and Linguistics	Law	Psychology	Public and Social Administration
Attitude:							
unfortunately	40	24	27	47	97*	36	45
surprisingly	28	22	48	37	32	**59***	14
Style:							
strictly	28	16	109*	23	64*	20	5*

Item	All 12 disciplines	Biology	Chemistry	Computer Science	Environmental Science	Neuro-science	Physics and Materials Science
Attitude:							
unfortunately	40	14*	22	56	39	21	38
surprisingly	28	14	14	28	11	33	6*
Style:							
strictly	28	19	7*	10*	28	16	9*

Table 7. Attitude and Style Individual Stance Adverbials – Discipline Differences.

Frequency – semantic categories

Table 3 shows these results. One category, Limitation, makes up 42% of all stance adverbials; two categories together, Doubt and Certainty and Limitation, make up 66%; and these two categories plus Actuality and Reality and Source of Knowledge/Allude to Evidence make up 89%. The other four categories (Viewpoint or Perspective, Imprecision, Attitude, and Style) combined make up only 11%. These proportions are consistent across disciplines, with some minor differences. Table 3 reveals considerable disciplinary variation, and large numbers of statistically significant differences. One clear and broad difference is between the non-sciences and sciences: the latter show significantly lower frequencies. However, this varies considerably by category. The six sciences are 30% lower over all categories, but 50% lower for Doubt and Certainty, and only 20% lower for Limitation. Among the distinctive areas of individual discipline variation are: (1) the very high frequencies in Language and Linguistics. All four major categories are significantly higher. (2) Law, where three major categories, Doubt and Certainty, Actuality and Reality, and Limitation, are significantly higher. The second, Actuality and Reality, was much higher than in any other discipline. (3) Chemistry shows even lower frequencies than the other sciences in two major categories, Doubt and Certainty and Limitation.

Frequency – individual forms

Tables 4 to 7 show these by category, in frequency order, with the most common first. Four notable findings across all disciplines were observed: (1) authors used a wide range of forms. 118 of the list of 158 appear in the corpus, as noted above, of which 38 appear in these tables. (2) The range of forms is much greater in two categories, Doubt and Certainty and Limitation, than in other categories. (3) Science authors used a narrower range of forms than non-science authors in just one category, Doubt and Certainty. They used an equally wide range in the other seven categories. (4) The limited number of just 20 forms make up the bulk of occurrences in five categories. The top eight Doubt and Certainty forms make up 78% of all occurrences in that category. "Indeed" and "actually" are dominant in Actuality and Reality at 69%, "according to" in Source of Knowledge/Allude to Evidence at 87%, and "about" in Imprecision at 81%. Finally, the first eight forms make up 72% of usage in Limitation. Two prominent areas of individual discipline variation are (1) the significantly higher frequencies across a wide range of forms in Language and Linguistics

in two categories, Doubt and Certainty and Actuality and Reality. This is also noticeable, but to a lesser degree, in one other category, Limitation. (2) Law shows very high usage of three forms in Actuality and Reality, and significantly higher frequencies, though across a narrower range of forms, in Doubt and Certainty and in Limitation.

Function

Individual manual checking of the function of every occurrence of all stance adverbials confirmed that they all functioned to construct epistemic stance, and all in line with Biber et al.'s (1999) categories. The only exceptions were the nine items noted above. Function will be explored further, along with examples from the corpus, in the next section.

Discussion and Conclusions

Semantic categories

Over all twelve disciplines, a striking finding is revealed in the proportional breakdown by semantic categories as percentages of the whole: Limitation makes up no less that 42% of the total in all eight categories, and Doubt and Certainty 24%. The other six categories combined make up only 34%, and all these proportions are remarkably consistent across disciplines, with only minor differences. Limitation, and to a lesser extent Doubt and Certainty, are much more prevalent and therefore presumably much more important to RA authors than hitherto suspected: the two previous empirical studies (Biber et al., 1999; Biber, 2006) do not discuss the topic or present category percentages. And the higher frequency of Limitation makes it appear particularly important. It is therefore suggested that the functions expressed in Doubt and Certainty and in Limitation, commitment to the truth of judgments of certainty, level of probability, and the limitation of propositions, are of particular value to RA authors. It seems that these two categories play a very important role in the important function of claiming, confirming, and expressing membership of and position in relevant discourse communities (although the potential functional overlap between these two categories means that these results must be handled with care). It is also suggested that the fact that these two categories each contain a very much larger number and variety of linguistic forms than the other six categories lends support to this conclusion.

Stance adverbials in the corpus appear to be less common overall than in the two previous empirical studies, though this is perhaps not surprising as they had different corpora: academic prose, book extracts plus RAs, textbooks, course packs, and course management.

Individual forms

Two findings seem particularly noteworthy. First, the range of forms employed by authors is wide: 118 appear in the corpus. It is also noted that science authors used an equally wide range in seven out of eight categories, the only exception being Doubt and Certainty. Second, just 20 forms make up a very large percentage of forms. This research has thus revealed the top 20 forms apparently preferred by authors, and the prevailing terminology used to express the target functions across twelve disciplines.

Disciplinary variation

Regarding the broad science/non-science difference in semantic category frequency, the sciences using significantly fewer stance adverbials overall than the non-sciences, this varied by category: 30% lower for all categories, 50% lower in Doubt and Certainty. Hyland (2008: 549-555) proposes that different disciplines value different kinds of arguments and also vary in what their readers already know and how they might be persuaded. He says the result is that physicists do not write like philosophers or applied linguists, and theorizes that disciplines range along a cline with hard knowledge sciences and softer humanities at opposite ends. His hypothetical cline describes sciences as empirical, objective, quantitative, showing linear and cumulative growth, utilizing experimental methods, not relying on rhetoric, and putting greater weight on methods, procedures and equipment; and humanities as explicitly interpretive, qualitative, utilizing discursive argument and more fluid discourses, and putting greater weight on strength of argument to present claims.

A closer examination of science RAs was then undertaken to try to understand some of the reasons for this much lower rate of occurrence. It was observed that authors tended to present and to develop claims in a different way, using less argument. They described their research justifications, methods, results and conclusions in a much more narrative and descriptive style: they seemed merely to describe the steps they took, and their findings, one by one, and let readers work out their claims. Presumably

the aim is to show the order of events or rather, this order is sufficient for readers, who perhaps do not need to be explicitly told the connections between facts, arguments, and claims. These authors used far fewer "linguistic mechanisms…to convey personal feelings, attitudes, value judgments, or assessments" (2008: 549-555).

Hyland's proposals were a helpful starting point for analysis of the present corpus: the twelve disciplines did range along a cline with sciences and non-sciences at different ends. However, while Biology, Chemistry, Environmental Science, Neuroscience, and Physics and Materials Science authors did show tendencies to rely less on rhetoric and to put greater weight on methods, procedures and equipment. It was also found that Economics (to a large extent), and Language and Linguistics and Psychology (to some extent), were empirical, objective, and quantitative, and put a lot of emphasis on methods and procedures. However, Business, Language and Linguistics, Law, Psychology, and Public and Social Administration were found to tend more towards interpretive and discursive argument, and to place greater weight on argument to present claims, than did the five sciences. Science authors, though, certainly employ the Doubt and Certainty functions of judgments of certainty and level of probability at a comparatively lower level. Finally, the fact that the sciences were only 20% lower in Limitation seems to be because only three sciences, Chemistry, Neuroscience, and Physics and Materials Science, were significantly lower in this category, while Environmental Science was higher. Some typical and illustrative examples from the corpus will now be shown. Examples (1)-(2) are from the sciences (more will be given later in this section):

(1) Periplasmic expression in E. coli as opposed to expression in the cytoplasm is preferred for proteins which are secreted in their native host and need a more oxidising environment for disulphide bond formation (Biology)

(2) Styrene and MMA formed a helical copolymer in conventional free radical vinyl copolymerization with captodatively substituted chiral acrylate, (–)-menthyl 2-acetamidoacrylate, near T_c. It is noted that styrene tends to undertake an alternating copolymerization with the chiral acrylate (Physics and Materials Science)

Examples (3)-(8) are from the non-sciences:

(3) They *typically* do not bring the same shared values, thought patterns, and actions to the situation (Business)

(4) In this case, environmental information is *clearly* required to re-establish orientation (Psychology)

(5) *Perhaps* children simply failed to correctly categorize the sounds in our novel words (Psychology)

(6) Experiment 6 investigates if this is *indeed* the case (Psychology, Actuality and Reality)

(7) The preceding variables capture *primarily* supply-side credit issues (Public and Social Administration)

(8) Scores above 215 are *generally* considered clinically significant (Public and Social Administration)

Closer examination of the corpus was then undertaken to investigate the striking individual discipline differences seen in Tables 3 to 7, which are not easy to explain. Hyland's (2008) hypotheses also had some value here, as it was found that Language and Linguistics and Law authors do not write like (for example) Biology or Chemistry authors.

1. Language and Linguistics. Further analysis shows that authors use 46% more stance adverbials overall than other disciplines, with all four major categories being significantly higher. Doubt and Certainty is 72% higher, and authors tend to rely more heavily on four terms, "clearly", "perhaps", "probably", and "of course", to express this function. Limitation is 42% higher, and authors rely more heavily on four terms, "usually", "primarily", "mainly", and "frequently", in this category. Presumably it is correspondingly more important and necessary in Language and Linguistics to express judgments of certainty and the level of probability of propositions, to signal the limitations of propositions, and to put greater weight on the strength of argument in these areas.

2. Law. Three out of four major categories, Doubt and Certainty, Actuality and Reality, and Limitation, were significantly higher. The second, Actuality and Reality, was higher than any other discipline. Law authors rely more heavily on two terms, "perhaps" and "of course", to express Doubt and Certainty; on three Actuality and Reality terms, "indeed", "actually", and "in fact", to express the status of propositions as real life fact; and on four terms, "generally", "in general", "frequently", and "largely", to express Limitation. Seemingly it is correspondingly more important and necessary in Law than in most other disciplines to utilise stance adverbials to express judgments of certainty/the level of probability towards propositions, the

status of propositions as real life fact, and assessments of the limitations of propositions concerning discipline-specific topics.

3. Chemistry. This discipline shows particularly low frequencies in two major categories, Doubt and Certainty and Limitation: lower than the other five sciences, or any other discipline. A closer examination of Chemistry RAs was then carried out to try to identify the reasons for this. After careful searches revealed that authors do not appear to cover Doubt and Certainty and Limitation in ways aside from the use of stance adverbials, it was concluded that these authors present and develop arguments in a different way, with less reference to these functions than even the other five sciences. Authors appear to rely almost exclusively on describing their research materials and equipment and findings. Apparently this is sufficient for Chemistry readers, who may not need to be openly told the connections between propositions and Doubt, Certainty, or Limitation. Randomly selected and representative examples from Language and Linguistics, Law, and Chemistry follow:

(9) It is *perhaps* not so surprising that the deficits are restricted in this way (Language and Linguistics)

(10) Dialogue annotation is not *usually* time-aligned (Language and Linguistics)

(11) A different consideration *frequently* overrides the notion of gender (Language and Linguistics)

(12) Robbers, *of course*, want to maximize their net gains (Law)

(13) The final decision is often *actually* made by jail administrators (Law)

(14) *In fact*, most victims apparently do not have guns (Law)

(15) Parolees *generally* suspend their identity while in prison (Law)

(16) The treatments are *largely* unavailable in developing countries (Law)

(17) In this study, it has been demonstrated that the array biosensor can be employed for the detection of E. coli O157:H7 in a variety of matrices and in the presence of high levels of extraneous bacteria. The assay was completed in less than 30 min with minimal sample preparation. The limit of detection without sample concentration or enrichment is 5×10^3 cells mL^{-1} in buffer (Chemistry)

(18) The slides were then incubated in a solution of 2% MTS in toluene. After 1 h, the slides were rinsed with toluene and dried with nitrogen.

The slides were then exposed to 2 mM GMBS in ethanol for 30 min. The slides were again rinsed with water, placed in 30 $\square$g mL$^{\square 1}$ NeutrAvidin in PBS (Chemistry)

Function

The next step was to look more closely at function. While the primary function of all stance adverbials in the corpus does fall into one or another of Biber et al's. (1999) categories, it was decided to adopt the philosophy and techniques of semantic preference to further explore the meaning and the function of stance adverbials. Understanding of the terms "semantic prosody/preference" has been evolving recently. Earlier work defined semantic prosody as the assessment of negative/positive meanings, but this approach has received criticism (e.g. Hunston, 2007) as these are hard to identify. Semantic preference is the creation of meaning through multiple occurrences of collocates, manifested only in context (e.g. Hunston, 2007; Bednarek, 2008).

The first step in this further analysis was to use the Patterns, Collocates, and Cluster functions of Concord in WordSmith Tools to isolate the clusters (groups of words which always appear in the same order, Mahlberg, 2007) and collocates associated with the top 20 stance adverbials across all twelve disciplines. Table 8 below shows the results. Selected representative examples extracted from the corpus follow below. The numbers in brackets in the "Stance adverbial" column refer to examples extracted from the corpus, which follow Table 8:

Category	Stance Adverbial	Clusters	Collocates
Doubt and Certainty	clearly	clearly show/s that, clearly indicate/d that, would clearly be, is clearly rejected	~ demonstrates, ~ related, ~ defined, ~ significant, ~ due, ~ importantly, ~ surprisingly/surprising, ~ because
	perhaps	perhaps the most, perhaps due to, perhaps most/more importantly, perhaps not surprising	
	probably (19)	is/was probably due to, probably the most, would probably be, is most probably	~ require, ~ less, ~ related, ~ lower
	possibly	possibly due to, could possibly be, possibly a result, and possibly also	could ~, ~ result, ~ inconsistent, ~ due
	certainly	it is certainly [not], there is certainly, and certainly more, certainly more complex	~ support, would ~, almost ~, most ~
	most likely	is/are/was/were most likely, most likely due, most likely explanation, will most likely	~ explanation, ~ due, will ~, which ~
Actuality and Reality	actually (20)	is/was actually a, than they actually, they actually did, may actually be	~ made/make, ~ present, ~ provided, ~ occurred
Source of Knowledge/ Allude to Evidence	according to	according to the following, according to the literature, according to the manufacturer's, according to this view	calculated ~, performed ~, prepared ~, vary/ies ~
Imprecision	about	about # of the, about # percent of, about half of the, about # of all	~ half, ~ # percent, ~ two, ~ a third
Limitation	generally	generally assumed that, it is generally accepted, generally considered to, generally associated with	~ speaking, ~ defined, ~ thought, ~ believed
	typically	have typically been, typically associated with, is typically the, typically used for	~ assumed, ~ have, ~ carried, ~ based
	usually	which is/are usually, it is usually, and are usually, usually associated with	~ have, ~ less, ~ involves, ~ assumed
	primarily	primarily due to, primarily concerned with, focused primarily on, primarily focused on	~ focused, ~ responsible, ~ due, ~ based
	mainly (21)	is/are mainly due, mainly composed of, mainly because of, mainly responsible for	~ through, ~ caused, ~ determined, depends ~
	frequently	the most frequently, more/less frequently than, frequently used in, frequently encountered in	~ cited, ~ observed, ~ occurring, ~ mentioned
	largely (22)	largely due to, largely based on, has largely been, have been largely	~ ignored, ~ unknown, ~ remain/ed, depend/s ~

Table 7. Most Frequent Clusters and Collocates of Common Stance Adverbials, in Order of Frequency - All Disciplines.

(19) It would *probably* require a legislative amendment to the statute to effect this change (Law)

(20) As readers are not *actually* present during the research activity, they must be provided with information (Computer Science)

(21) Venture capital is a type of business financing provided *mainly* through the acquisition of a stake in small and medium sized firms (Business)

(22) Religiosity appears to have been *largely* ignored in conceptual and empirical work (Law)

No clusters or collocates were found for "indeed", "of course", "obviously", or "in general", nor any useful results for individual disciplines because of relatively low item occurrence. Table 8 reveals conventional stance adverbial patterns in RAs across twelve disciplines. It is suggested that these clusters and collocates represent the patterns which are accepted as standard ways for authors to present and discuss their research, making them standard terminology. This also makes them an important part of the meaning and the function of these common stance adverbials. It is also suggested that these patterns are a useful finding, for a number of reasons. Gledhill (2000) points out that collocations are fundamental units in texts, that they validate the existence of discourse communities, and that they are subconscious efforts to conform to discipline norms. They may also be more quickly recognized than individual words (Cantos & Sanchez, 2001) and reduce processing effort for readers (Jones & Haywood, 2004). Schmitt and Carter (2004) say that collocations are stored and processed as unitary wholes, and Schmitt, Grandage and Adolphs (2004: 127) that writers use the same clusters repeatedly because they are "prepackaged in the memory". Mahlberg (2003) says that meaning develops across word clusters and not through single words, and Durrant (2009) that learners need to acquire high-frequency collocations. Morley and Partington (2009) propose that members of discourse communities share very large numbers of collocations, and project community membership through them. Channell (2000) suggests that theories of meaning built through semantic preference research are useful for language teaching. She adds that meaning is hidden from introspection and observation until we have a large number of instances of a word, derived through the observation of naturally occurring corpus data.

Further research

More complete lists of stance adverbials can be developed, perhaps combining lists used in previous research, forms found in dictionaries and a thesaurus, and forms found inside RAs. The latter method may be especially useful. Other questions to research are: How frequent are stance adverbials in other disciplines, and how are they used? How and when are they acquired by research writers? How do Chemistry authors achieve certain functions? What are the implications of the potential functional overlap between Doubt and Limitation?

This study has revealed some conventional forms in RAs across twelve disciplines. Analysis of the corpus leads to the suggestion that stance adverbials play an important role in the construction of epistemic stance, a key part of research writing, in RAs. Authors employ them to express attitudes, value judgments, and assessments towards their suggestions, claims, and propositions, and thereby accomplish the essential functions of claiming and confirming membership of their discourse community, and constructing identity. Additionally, sciences and non-sciences, and certain disciplines, achieve this in significantly different ways, confirming the need to consider discipline variation when researching their use, and adding to knowledge of ESP. It is also proposed that Doubt and Certainty, and Limitation, stance adverbials in particular are more important in RAs than previously thought, and that semantic preference has added valuable information to the understanding of the meanings and functions of stance adverbials. Finally, it is hoped this study helps us better understand scientific expression and the RA.

References

Bednarek, M. (2008). "Semantic preference and semantic prosody re-examined". *Corpus Linguistics and Linguistic Theory* 4: 119-139.

Bhatia, V.K. (2000). "Genres in conflict" in A. Trosborg (ed.), *Analysing Professional Genres,* 147-161. Amsterdam: John Benjamins.

Biber, D. (2006). "Stance in spoken and written university registers". *Journal of English for Academic Purposes* 5: 97-116.

Biber, D., S. Johansson, G. Leech, S. Conrad & E. Finegan (1999). *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Biber, D., S. Conrad & V. Cortes. (2004). "*If you look at*…: Lexical bundles in university teaching and textbooks". *Applied Linguistics* 25: 371-405.

Cantos, P. & A. Sanchez. (2001). "Lexical

constellations: what collocates fail to tell". *International Journal of Corpus Linguistics* 6: 199-228.

Channell, J. (2000). "Corpus-based analysis of evaluative lexis" in S. Hunston & G. Thompson (eds.), 38-55.

Conrad, S. & D. Biber (2000). "Adverbial marking of stance in speech and writing" in S. Hunston & G. Thompson (eds.), 56-73.

Durrant, P. (2009). "Investigating the viability of a collocation list for students of English for academic purposes". *English for Specific Purposes* 28: 157-169.

Gledhill, C.J. (2000). *Collocations in Science Writing*. Tubingen: Gunter Narr Verlag.

Halliday, M.A.K. (1993). "Towards a language-based theory of learning". *Linguistics and Education* 5: 93-116.

Hunston, S. (2007). "Semantic prosody revisited". *International Journal of Corpus Linguistics* 12, 249-268.

Hunston, S. & G. Thompson (eds.) (2000). *Evaluation in Text: Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press.

Hyland, K. (1996). "Talking to the academy: Forms of hedging in science research articles". *Written Communication* 13, 251-281.

Hyland, K. (1998). *Hedging in Scientific Research Articles*. Amsterdam: John Benjamins.

Hyland, K. (1999). "Disciplinary discourses: Writer stance in research articles" in C. Candlin & K. Hyland (eds.), *Writing: Texts, Processes, and Practices,* 99-121. London: Longman.

Hyland, K. (2005). "Stance and engagement: A model of interaction in academic discourse". *Discourse Studies* 7: 173-192.

Hyland, K. (2008). "Genre and academic writing in the disciplines". *Language Teaching* 41: 543-562.

Jones, M. & S. Haywood (2004). "Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context" in N. Schmitt (ed.), 269-300.

McGrath, L. & M. Kuteeva (2012). "Stance and engagement in pure mathematics research articles: Linking discourse features to disciplinary practices". *English for Specific Purposes* 31: 161-173.

Mahlberg, M. (2003). "The textlinguistic dimension of corpus linguistics: the support function of English general nouns and its theoretical implications". *International Journal of Corpus Linguistics* 8: 97-108.

Mahlberg, M. (2007). "Clusters, key clusters, and local textual functions in Dickens". *Corpora* 2: 1-31.

Morley, J. & A. Partington (2009). "A few frequently asked questions about semantic – or evaluative – prosody". *International Journal of Corpus Linguistics* 14: 139-158.

Sancho-Guinda, C. & K. Hyland (2012). "Introduction: A context-sensitive approach to stance and voice" in K. Hyland & C. Sancho-Guinda (eds.), *Stance and Voice in Written Academic Genres,* 1-11. Basingstoke, UK: Palgrave Macmillan.

Schmitt, N. (ed.) (2004). *Formulaic Sequences: Aquisition, Processing, and Use*. Amsterdam: John Benjamins.

Schmitt, N. & R. Carter (2004). "Formulaic sequences in action" in N. Schmitt (ed.), 1-22.

Schmitt, N., S. Grandage & S. Adolphs (2004). "Are corpus-derived recurrent clusters psycholinguistically valid?" in N. Schmitt (ed.), 127-151.

Scott, M. (2004). *WordSmith Tools* Version 4. Oxford: Oxford University Press.

Silver, M. (2003). "The stance of stance: A critical look at ways stance is expressed and modeled in academic discourse". *Journal of English for Academic Purposes* 2: 359-374.

Thompson, G. & S. Hunston (2000). "Evaluation: An introduction" in S. Hunston & G. Thompson (eds.), 1-27.

Williams, G. C. (1998). "Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles". *International Journal of Corpus Linguistics* 3: 151-171.

Williams, G. (2002). "In search of representativity in specialised corpora: Categorisation through collocation". *International Journal of Corpus Linguistics* 7: 43-64.

Matthew Peacock teaches in the Department of English at the City University of Hong Kong. His research interests include English for Specific Purposes, corpus analysis, research writing, genre analysis, and TEFL methodology. He co-edited (with John Flowerdew) a collection from Cambridge University Press, *Research Perspectives on English for Academic Purposes*.

Appendix 1

Journals in the corpus

Biology
Applied Soil Ecology
Biochimica et Biophysica Acta
Biomass and Bioenergy
Chemistry and Biology
Current Biology
Journal of Biotechnology

Business
Industrial Marketing Management
International Business Review
International Journal of Information Management
International Journal of Project Management
International Journal of Research in Marketing
Journal of Business Venturing
Journal of International Management
Journal of Operations Management

Chemistry
Analytica Chimica Acta
Analytical Biochemistry
Corrosion Science
Inorganica Chimica Acta
International Journal of Inorganic Materials
Journal of Chemical Thermodynamics
Journal of Organometallic Chemistry
Journal of Solid State Chemistry

Computer Science
Computers in Human Behavior
Computer Speech and Language
Information and Software Technology
International Journal of Human-Computer Studies

Economics
Economic Modelling
Journal of Economic Behavior and Organization
Journal of Economics and Business
Journal of Financial Economics

Environmental Science
Applied Energy
Atmospheric Environment
Biomass and Bioenergy
Ecological Modelling
Environmental Pollution
Global Environmental Change

Language and Linguistics
English for Specific Purposes
Journal of English for Academic Purposes
Journal of Neurolinguistics
Journal of Second Language Writing
Language and Communication
Language Sciences

Speech Communication
System

Law
California Law Review
Canadian Journal of Criminology
International Review of Law and Economics
Journal of Criminal Justice

Neuroscience
Cognition
Brain and Cognition
Neuropsychologia
Neuroscience

Physics and Material Science
Acta Materialia
Biomaterials
Chemical Physics
Corrosion Science
International Journal of Fatigue
Journal of Luminescence
Journal of the Mechanics and Physics of Solids
Physica C: Superconductivity
Polymer

Psychology
Acta Psychologica
Cognitive Psychology
Journal of Anxiety Disorders
Journal of Research in Personality

Public and Social Administration
Child Abuse & Neglect
Evaluation and Program Planning
Habitat International
International Journal of Public Sector Management
Social Science & Medicine
World Development