Marina Bondi*

What came to be called: evaluative *what* and authorial voice in the discourse of history

Abstract: While acknowledging the importance of specific lexis in profiling discourse communities, studies on academic discourse have paid growing attention to general lexis and its disciplinary specificity. This is particularly true of recent approaches to phraseology in English for Academic Purposes (EAP), including both statistically significant clusters and grammar patterns and semantic sequences. The current paper explores the notion of semantic sequences through a case study of sequences involving relative *what*. Based on a corpus of academic journal articles in the field of history (2.5 million words), the analysis highlights co-occurrence of *what* with a range of signals referring to a shift in time perspectives or attribution, and points to the "redefining" function of *what*, introducing a re-formulation of what has just been said or proposing an interpretation on the basis of the cotext. A tentative classification of the sequences is provided, building on previous studies on a local grammar of evaluation. The sequences are shown to highlight the argumentative voice of historians interacting with their sources and their discourse community, by showing awareness of different interpretations of people and events in history.

Keywords: academic discourse, historical discourse, evaluation, *what*-nominal clauses, semantic sequences, authorial identity

1 Introduction

Research into specialized discourses has often paid particular attention to specific lexis, and to how this instantiates the meanings and values associated with specific communities or institutions. More recently, however, studies on academic discourse have shown growing interest in general lexis and its disciplinary specificity, for example in studies on metadiscourse and reflexivity (Dahl 2004; Hyland 2005; Holmes and Nesi 2009; Bondi 2010) or evaluative

---

*Corresponding author: Marina Bondi,** Department of Studies on Language and Culture, University of Modena and Reggio Emilia, Largo S. Eufemia 19, 41121 Modena, Italy, E-mail: Marina.bondi@unimore.it

language use (Del Lungo Camiciotti and Tognini Bonelli 2004; Charles 2006; Groom 2005, 2009; Hyland and Diani 2009). If specific lexis can be seen to identify the ontologies on which specific discourses are based, i. e. the categories, entities and relations that constitute the object of study of a discipline, general language often points at the epistemology of its discourse community, i. e. its fundamental values and views on how truth is produced, established or justified (cf. Groom 2009: 123). Groom (2010), for example, provides an illustration of how patterns with *of* can be studied to illustrate peculiarities of historical discourse by looking at the most frequent semantic sequences involved: PROCESS + *of* + OBJECT (*the building of a new church*), PROPERTY + *of* + PHENOMENON (*the essential values of academic life*), CONCEPTUALISATION + *of* + PHENOMENON (*the institution of Mothers' Day*).

The purpose of this study is to explore the notion of semantic sequence (Hunston 2008) and its potential in identifying distinctive features of specialized discourses, with reference to the academic discourse of historians. Starting from the observation that *what* is particularly frequent in history, the analysis focuses on the use of nominal relative *what* in combination with signals referring to a shift in time perspectives or attribution. The aim is to show that the high frequency of *what* in history is related to the interpretative nature of its discourse. Nominal relative *what* appears to typically identify discourse referents through the process that characterizes them – *what took place in the South*, *what Britain did not control*. Many of these *wh*-clauses, however, identify and classify discourse referents (people and events) through either their explicit evaluative interpretation (*what is perhaps the earliest popular account of success*) or through a denomination that evokes their interpretation (*what came to be called "Vichy France"*). By doing this, they highlight a pattern of a "local grammar of evaluation" (Hunston and Sinclair 2000) that seems to characterize historical discourse and the "interpreter identity" of historians as writers, their need to interpret events by expressing evaluation (Coffin 2006).

The rest of the paper is organized as follows. The next section introduces the background to the study of semantic sequences with *what*. Section 3 presents the corpus used and the methodology adopted for the study of *what* in academic history. Section 4 offers an overview of quantitative data on *what* and its combinations with other words, whereas Section 5 focuses on the semantic sequences involving *what*, as an extension of previous studies on a local grammar of evaluation (Hunston and Sinclair 2000). Co-textual analysis looks both at the lexico-semantic patterns and at the pragmatic functions involved. Frequencies and patterns are then interpreted – in Section 6 – in the light of factors characterizing academic discourse and specific features of writer identity in historical discourse (Coffin 2006), before conclusions are drawn in Section 7.

2 Theoretical background: Phraseology, semantic sequences and evaluation

Attention to words in association has been greatly influenced by Sinclair's work, starting from his notion of the idiom principle – viewing "semi-preconstructed phrases that constitute single choices" as the core unit of linguistic meaning and structure (Sinclair 1991: 110) – to the idea of the "extended unit of meaning" (i. e. "an element of meaning which is the function of the item in its cotext and context" [Sinclair 2004: 121]) and that of a general phraseological tendency of language (the tendency for words to be co-selected by speakers to achieve meanings) (Cheng et al. 2008).

The study of disciplinary language and discourse has profited considerably from recent corpus-informed approaches to the phraseology of academic discourse. These include studies paying attention to automatic derivation of statistically significant continuous sequences of word forms (clusters or lexical bundles) (Biber 2004; Biber et al. 2004; Hyland 2008; Bondi 2009b; Pecorari 2009), as well as work on discontinuous sequences, such as lexical frames, i. e. words which form a "frame" surrounding a variable slot (Gray and Biber 2013) or concgrams, i. e. "the permutations of constituency variation and positional variation generated by the association of two or more words" (Cheng et al. 2006: 414).

Studies on academic discourse often start with a recurrent framework and build up a genre- or discourse-specific phraseology by examining the framework's common collocates (Gledhill 2000; Luzón Marco 2000; Groom 2005, 2010; Charles 2006, 2012; Bondi 2010; Vincent 2013). This means taking phraseology in its most inclusive sense, including work on collocation, but also grammar patterns (Hunston and Francis 1999) and semantic sequences (Hunston 2008, 2011).

Semantic sequences can be defined as sequences of meaning elements that have no single realization: "recurring sequences of words and phrases that may be very diverse in form [...], sequences of meaning elements, rather than as formal sequences" (Hunston 2008: 271). Strictly speaking, these do not pertain to the level of language analysis, but to the analysis of discourse: "they do not tell us what language is like [...]. Rather, they tell us what things are often said" (Hunston 2011: 170). Phraseology is studied here at the intersection of form, meaning and function, thus combining lexico-grammatical and lexico-semantic perspectives and integrating them in the study of discourse.

Focusing on *what* inevitably requires considering the ample literature on its use and grammar patterns, particularly in pseudo-cleft sentences. Pseudo-cleft

constructions are usually explained in terms of information focus and textual emphasis. In reference grammars they are often dealt with as devices "giving prominence" to an element by means of a "division of the sentence into two clauses, each with its own verb" (Quirk et al. 1985: 1383), thus making explicit the division between given and new parts of the communication. They are one of the many constructions that help make clauses "fit in within the context, thereby building a coherent text that conveys emphasis and related stylistic effects" (Biber et al. 1999: 896). Cleft sentences have been paid great attention in a textual perspective as tools for presenting and highlighting new information (cf. Declerck 1984; Collins 1991).

More recently, interest has shifted to the evaluative and interpersonal functions of the cleft construction. Pseudo-cleft constructions have been shown to have important semantic and pragmatic associations characterizing them as a prominent pattern in a "local grammar" of evaluation (Hunston and Sinclair 2000), e. g. *what's interesting is the tone of the statement*. Hunston and Sinclair's fifth pattern (2000: 89–90) involves a link verb, an "Evaluative Category" (*interesting*) and a "Thing Evaluated" (*the tone of the statement*). The *wh*-clause typically introduces what they would call attitudinal or affective evaluation expressing the writer's opinion regarding entities – *what is unusual, what is largely overlooked, what is individual, subjective* etc. – rather than modal evaluation, i. e. different degrees of likelihood, obviousness and relevance (Hunston 2000). They also notice that evaluation is often attributed in patterns involving explicit mention of an "Evaluator", as illustrated in Table 1.

Table 1: Pseudo-clefts. Fifth pattern (iii) in Hunston and Sinclair (2000: 90).

Hinge	Evaluator		Evaluative category		Thing evaluated
What	Noun group	Verb group	Adjective group	Link verb	Clause of noun group
What	*I*	*find*	*so amazing*	*is*	*that my dad is a very strict Hindu*

Herriman (2008) extends the perspective, highlighting that within the framework of Appraisal (Martin 2000; White 2003; Martin and White 2005), clefts are "heteroglossic rhetorical strategies used by speakers and writers to negotiate an authorial position" (Herriman 2008: 144). In White's (2003) terms, these would be heteroglossic utterances characterized by markers of Engagement, "a cover-all term for resources of intersubjective positioning" including "modality,

polarity, evidentiality, hedging, intensification, attribution, concession and consequentiality" (White 2003: 260). The basic distinction would be between bare assertions, presented as "facts", acknowledging no alternative position (also called "monoglossic" or "undialogized") and evaluative propositions, variously engaging with alternative positions (White 2003: 265), also called heteroglossic or dialogistic, in line with Bakhtin's (1986) widely influential notion of dialogism and heteroglossia.

Herriman's work is also related to the preference of cleft constructions for specific semantic classes of verbs: in her analysis of *wh*-clefts in the Freiburg update of the Lancaster-Oslo-Bergen Corpus (F-LOB), she notices that they seem to show a preference for mental and relational processes (Herriman 2004: 451), contrasting her data with Matthiessen's (1999) analysis, where the most frequent category was that of material processes. Work by Deroey (2012: 112) also takes up the issue and provides a picture of *wh*-clefts in academic lectures, where material processes are the most frequent but mental, verbal and relational processes are also well represented.

The present study takes an interest in these developments, but looks at other forms of relative nominal clause frequently found in academic history with an evaluative function. Drawing on Cheng et al.'s (2006: 423–424) observation on the association of *call* and *what*, seemingly "in relation to a speaker reformulating what he/she has just said (*what I would rather call*) or to introduce something based on what has just been said (*what we call*), the analysis proposed here is meant to explore this association at the level of discourse, looking at the possibility of establishing recurrent sequences of elements that might be semantically related. Focusing on semantic sequences with *what*, this paper aims at exploring how historians use these patterns to position themselves and their readers.

3 Materials and methods

The study is based on a corpus of 306 academic journal articles (2.5 million words) published in a given time period (1999–2000) in a set of international journals chosen to cover a wide range of sub-disciplines in the field of history. The principle of selection of the articles for this history corpus (HC) was purely temporal: all the articles published in all the issues of each journal were collected over a year. The journals are listed below with the acronym used to identify the source of examples and the number of articles published by each journal in brackets: AHR – *American Historical Review* (30), AQ – *American*

Quarterly (32), GaH – *Gender & History* (51), HoEI – *History of European Ideas* (29), HR – *Historical Research* (39), JoIH – *Journal of Interdisciplinary History* (19), JoMH – *Journal of Medieval History* (42), JoSH – *Journal of Social History* (14), LHR – *Labour History Review* (17), SiH – *Studies in History* (33).

Other corpora used for comparison were: the academic component of the Corpus of Contemporary American English (COCA-A) (Davies 2008), consisting of 103.4 million words, and a corpus of international journal articles in the field of business and economics (BEC), compiled following the same principles as the history corpus (HC) and comprising of about 5 million words.[1] The comparison was meant to highlight the peculiarity of *what* in the corpus of historical articles in relation to academic discourse in general, as well as to journal articles in a major area in the social sciences. The choice of business and economics was also meant to highlight the interpretative nature of history against more empirical and formal disciplines (Kyburg 1990: 16) within the general area of the humanities and the social sciences.

The methodology adopted combined a corpus and a discourse perspective. A preliminary analysis of frequency data was carried out on wordlists and keywords, as calculated by *WordSmith* 6 (Scott 2012). A word is identified as a keyword if it occurs with unusual frequency in the text or corpus under investigation as compared to its frequency in a reference corpus. Keywords were calculated (using log-likelihood) against the COCA-A and the BEC. The preliminary analysis also considered other phraseological patterns around *what*, always using *Wordsmith Tools*, namely clusters, lexical frames and WSconcgrams. These were studied to see what types of verbs *what* was mostly associated with.

Having established some of the closest associations of *what* in our corpus, the main study consisted in a co-textual analysis of concordances. This meant looking at syntactic patterns, as well as at the lexico-semantic patterns and at the pragmatic functions involved. Concordance analysis was carried out on a random sample of 200 concordances of *what* from the HC, the COCA-A and the comparable BEC. The analysis began by describing our sample concordance in

---

1 The business component comprises 370 journal articles published in the following journals: *Administrative Science Quarterly* (ASQ), *Journal of World Business* (JWB), *Academy of Management Journal* (AMJ), *Marketing Science* (MS), *Journal of Marketing Research* (JMR), *Business & Society Review* (B&SR), *Business Strategy Review* (BSR). The economics component comprises 436 journal articles published in the following journals: *European Economic Review* (EER), *European Journal of Political Economy* (EJPE), *International Journal of Industrial Organization* (IJIO), *International Review of Economics and Finance* (IREF), *Journal of Corporate Finance* (JCF), *Journal of Development Economics* (JDE), *Journal of Socio-Economics* (JSE), *The North American Journal of Economics and Finance* (NAJEF).

simple structural terms. Clause types were classified following Biber et al. (1999) into dependent and independent clauses. Independent clauses introduced by *what* could be either exclamative or interrogative. Dependent clauses could be either dependent interrogative clauses (presenting an indirect question) or nominal relative clauses (typically paraphrasable by a general head noun modified by the *Wh*-clause functioning as a relative clause) (Biber et al. 1999: 683).

Following this, the nominal relative clauses were submitted to a semantically oriented analysis. Concordances were studied first to reconstruct semantic sequences and then to explore the potential of semantic sequences in identifying distinctive features of discourse. The focus lied on the possibility of identifying Evaluative Categories, considering both inscribed evaluation and evoked evaluation (Martin and White 2005: 61). This meant considering both explicit use of evaluative lexis and a range of forms that provoke or invite evaluation, either for their ideational content or for the choice of specific denominations. It should be noticed, in fact, that in historical discourse many expressions such as nouns referring to historical periodization (i. e. chrononyms such as *the Terror*) and events (i. e. eventonyms such as *the First World War*) actually constitute summary representations of historical debate (as for example in the shift from *The Great War* to the *First World War*) often involving largely shared evaluative meanings.

Focus on these elements allowed us to reconstruct recurrent sequences of semantic elements referring to how historical characters, events or ideas can be or have been variously interpreted and denominated. The sequences were identified as "polyphonic". A quantitative assessment of these sequences was made with reference to the COCA-A and the BEC, by establishing their frequency out of the total number of nominal relative clauses. To contextualize the figures, data are also provided as to other heteroglossic clauses (typically but not exclusively modalized) or bare assertions.

The sequences were also related to the choice of verbs in the *wh*-clause, classified according to the general classes of processes proposed in systemic functional analyses of transitivity: mental, relational, verbal, material, existential and behavioural (Halliday 1985: 101–130).

Sequences were further specified in relation to what kind of heteroglossia (potential diversity of point of view) they involved, usually with reference to authorial voice in implicit or explicit relation to the point of view of other historians or historical characters. A major distinction was drawn between diversity of "perspective", i. e. spatio-temporal and cognitive point of view, and "position", i. e. the attitudinal, argumentative and ideological point of view adopted by the writer (Bondi 2009a). Polyphonic sequences were classified accordingly into "position polyphony" and "perspective polyphony".

Both sequences were studied in the context of the stretches of discourse that follow them, to check whether the text actually developed the potential diversity of points of view highlighted in the *what*-clause. Other language resources involved in marking this diversity through attribution or temporal shifts were also listed.

The qualitative interpretation of the sequences involved the notion of authorial voice or identity, keeping in mind that consistent patterns in the use of evaluative resources appear to "construct particular authorial identities or personas" (Martin and White 2005: 161). Following Tardy (2012: 34), I take the notion of voice to include the individual aspects of authorial voice (its uniqueness, authenticity and authoritativeness), its social aspects (repertoires of self-representation and authorial presence), and its dialogic features (the interaction between the individual and the social dimensions, including writer-reader interaction). In this light, adapting Coffin's work on the voice of historians as Recorders and Appraisers (e. g. Coffin 2002: Figure 3; Coffin 2006: 152–155), three different roles for historians can then be identified: Recounters (in the textual narrative), Interpreters (in the interpretation of narrated events) and Academic Arguers (in the dialogic argumentation of the interpretation put forward) (Bondi 2007: 68–69). The distinction between these latter roles is more one of degree than of kind: writers present their arguments in ways that are sensitive to a reader in both cases, but overtly highlighting either internal consistency or dialogic positioning.

4 Preliminary analysis: Focus on frequency data

The preliminary analysis of quantitative data aimed at establishing the discipline-specificity of *what*. Table 2 offers an overview of the frequency of *what* in the HC, in the BEC and in the COCA-A. The word-form *what* stands out as significantly frequent in history when compared to all of them, as shown in Table 2.

Table 2: Frequency of what.

	HC		BEC		COCA-A	
	Frequency	Normalized frequency (per thousand words)	Frequency	Normalized frequency (per thousand words)	Frequency	Normalized frequency (per thousand words)
What	3,232	1.3	4,158	0.8	147,052	0.1

Table 3: Top twenty clusters with *what* (3- and 4-word clusters) in HC.

3-word clusters	Frequency	4-word clusters	Frequency
what it is	36	the question of what	14
what has been	33	what we might call	14
of what was	31	what might be called	13
of what is	30	what it meant to	11
what might be	30	what it means to	10
to what extent	25	it means to be	8
of what it	25	what he calls the	8
what is the	23	what we mean by	8
what he called	20	thing what it is	7
what can be	20	we do not know	7
what had been	19	what is at stake	7
what we might19	19	it meant to be	7
this is what	18	what they saw as	7
of what they	17	of what has been	7
what he calls	17	what came to be	6
what it was	16	what appears to be	6
what i have	16	what it is not	6
in what is	16	in what might be	6
we might call	15	to find out what	6

Table 3 illustrates the top twenty 3-word and 4-word clusters, fixed strings of word forms, used in association with *what* in the HC. The most frequent associations of *what* include relational linking verbs (typically forms of *be*), modals (*might*) and verbs referring to verbal processes (*call*, *mean*) or mental processes (*know*, *see as*).

A similar search in the BEC would produce a slightly different range of verbs, including linking verbs (*be*) and verbal processes (*call*, *refer*), but also material processes (*happen*, *follow*, *do*).

The centrality of the co-selection *what–call* noticed by Cheng et al. (2006) is confirmed in my data, and its special role in history is also foregrounded: there are 274 occurrences of the concgram *what–call* in the HC against 190 in the BEC (with a frequency of 1.1 versus 0.4 per ten thousand words, respectively). Table 4 illustrates the frequency of different combinations of *what* with forms of the verb *call* in the HC (with a span of 5 words to right and 5 words to the left), showing that the *–ed* form of the verb is clearly dominant.

As only 10 of the concgrams (out of 274) involve the verb in first position, it appears that the vast majority of the occurrences of *what* introduce a clause with a form of the lemma *call*. Table 5 illustrates what other full lexical verbs are

Table 4: Concgrams involving *what* and verb forms of *call* (HC).

	Frequency of co-occurrence	**Percentage**
What – called	151	55.1
What – calls	66	24.1
What – call	52	18.9
What – calling	5	1.8
TOTAL	274	100

Table 5: Top 10 full lexical verb collocates in R1-R4 position (HC).

R1	**R2**	**R3**	**R4**
makes	*called*	*called*	*take*
follows	*calls*	*call*	*used*
happened	*saw*	*mean*	*saying*
happens	*meant*	*calls*	*know*
made	*means*	*said*	*perform*
constitutes	*considered*	*described*	*expect*
appears	*perceived*	*termed*	*work*
became	*happened*	*considered*	*think*
seems	*said*	*written*	*mean*
distinguishes	*happening*	*seen*	*lead*

found as collocates of *what* in the first four positions on the right, showing that *call* is found in all positions, but also that many other verbs representing mental or verbal processes are found in the clauses following *what*, especially in second and third position.

A preliminary classification of the types of clauses introduced by *what* was carried out on a sample of 200 concordances. The data in Table 6 shows that the nominal relative clause is largely dominant in the corpus, as well as in other academic corpora, for comparison.

Table 6: Types of clauses.

	HC	**BEC**	**COCA-A**
Exclamative	1 (0.5 %)	1 (0.5 %)	0
Interrogative	21 (10.5 %)	41 (20.5 %)	40 (20 %)
Dependent interrogative	29 (14.5 %)	30 (15 %)	42 (21 %)
Nominal relative	149 (74.5 %)	128 (64 %)	118 (59 %)
TOTAL	200 (100 %)	200 (100 %)	200 (100 %)

The data also shows that history does not make extensive use of *wh*-interrogatives, but is certainly prominent in the use of nominal relatives.

5 Concordance analysis: Semantic sequences

Concordance analysis focused on semantic sequences in nominal relative clauses. The discourse function of *what*-clauses was confirmed to be that of identifying discourse referents through the process that characterizes them. Some of these clauses could simply be explained in ideational terms, as in (1), where the nominal clause identifies referents by characteristics that are presented as "fact":

(1)   These musicians and dancers took from what *they found* around them. (AQ 1999)

Rather than factual we could say – following Martin and White (2005: 99) – that they are "monoglossic" bare assertions, "not overtly referencing other voices or recognising alternative positions".

Many other clauses involve resources that Martin and White (2005) would classify as "heteroglossic", such as modality or negation: *what was not/what they ought to be/what was clearly a market economy*. Focus on collocates confirmed that *called* is the most frequent full lexical item (followed by *calls*, *know*, *call*, *mean*, *said* etc.), but modals (*would, might, could, should*) proved to be very high in the list. The analysis also revealed a tendency of *what* to co-occur with inscribed or evoked evaluative meanings in general, including both evaluations of Status (modal evaluation) and of Value (attitudinal stance) (Hunston 2000).

More importantly, the cotext was often noticed to signal diversity in perspective (*the first step of what would become a full retreat after 1842*) or position (*what Feather calls a "cycle of development"*). The *wh*-clause identifies a referent through a specific interpretation or denomination, and at the same time signals the possibility of a different cognitive or evaluative point of view. The nominal clause creates a double focus of information, where two noun groups could have simply been listed, as in (2/2a*).

(2)   Herder placed das Volk and what *he deemed to be* traditional texts at the centre of emergent German nationality. (AQ 1999)

(2a)* Herder placed das Volk and traditional texts at the centre of emergent German nationality.

Semantically, this is equivalent to introducing an Evaluator in a sequence also involving Evaluative Category and Thing Evaluated. Pragmatically, however, introducing a third person Evaluator allows the writer to avoid expressing alignment with the Evaluation presented. The writer chooses to isolate the Evaluative Category (*traditional texts*) in a separate clause and present it as the result of Herder's interpretation, rather than his/her own category. The Evaluative Category is preceded by a "Prefacing" element that isolates it as the object or complement of a verb expressing verbal or mental processes: *what was referred to as "voluntary resettlement", what counts as a "theory"*.

In most cases, the writer uses this prefacing pattern to disassociate from the Evaluative Category, either through attribution (3) or through modal hedging (4).

(3)   He begins by describing what he calls the "traditional ideal of induction". (SIH 2000)

(4)   These include a long verse history of the Exodus, containing what may be the longest identifying acrostic in the history of Hebrew poetry. (JOMH 2000)

These patterns – while not strictly speaking pseudo-cleft sentences involving a *wh*-clause, a form of the verb *be* and a focused element (noun phrase, infinitive clause or finite nominal clause (Biber et al. 1999: 959) – share many of the textual and pragmatic functions of a cleft by giving emphasis to the Evaluative Category. One of the clauses (mostly the subordinate) carries the evaluative weight of the sentence and qualifies the interpreting activity of the writer, thus foregrounding the possibility of different interpretations.

These evaluative uses of *what* can thus be seen as one of the many resources of Engagement, acknowledging diversity and negotiability of positions (White 2003; Martin and White 2005). Many of the examples we have analyzed can be called "polyphonic", in that they introduce an Evaluative Category (whether inscribed in or evoked by the lexical elements chosen) and signal the writer's attitude to (agreement or disagreement with) the values inscribed or evoked by the Evaluative Category. Rather than just a double information focus (as in clefts), we have here a dissociation between the degree of adhesion of the writer to different elements in the sentence.

Paraphrasing Hyland and Tse's (2005) analysis of "evaluative *that*", we can talk of an "evaluative *what*", signalling patterns of polyphonic interaction of the writer's voice with different interpretations of historical fact. The evaluative element is not isolated as superordinate (as in Hyland and Tse's *we believe that*) but rather shifted to a subordinate clause, where it is prefaced by signals of

attribution or of spatio-temporal or cognitive shift. Separating the two elements of the sentence signals that they have different status: this can mean that one section is more factual and the other more explicitly interpretative, or that they are both evaluative, but represent different positions, as illustrated in (5).

(5)   What John naïvely reads as Viney's "home instinct" (170), one might see as a much more wilful and complicated decision. (AQ 1999)

The structure highlights that the referent identified can be interpreted in different ways, thus clearly foregrounding the voice of the historian as Interpreter and Arguer.

The quantitative relevance of these evaluative polyphonic uses is illustrated in Table 7, which also shows the proportion of other heteroglossic patterns and bare (monogloss) nominal clauses in our sample concordances.

Table 7: Subtypes of nominal relative.

	HC	BEC	COCA-A
Polyphonic	83/149 (55.7 %)	37/128 (28.9 %)	41/118 (34.7 %)
Bare	33/149 (22.1 %)	49/128 (38.3 %)	53/118 (44.9 %)
Otherwise heteroglossic	33/149 (22.1 %)	42/128 (32.8 %)	24/118 (20.3 %)
TOTAL	149 (100 %)	128 (100 %)	118 (100 %)

The data suggests that historians are particularly interested in signalling explicit interpretations and evaluative denominations (inscribed and evoked evaluation).

The concordances of polyphonic sequences were also studied to classify the processes involved. Table 8 shows the relative importance of verbal and mental processes, with relational processes lagging behind.

Table 8: Verb types in polyphonic nominal clauses (HC).

Process types	Occurrences	Percentage	Examples
Mental processes	31	37.3	*conceive, consider, deem, envision, interpret, mean, neglect, see, understand, view*
Verbal processes	34	41.0	*account for as, call, condemn as, explain, pass as, refer to, term,*
Relational processes	18	21.7	*be, become, constitute*
TOTAL	83	100	

6 Evaluative *what* in polyphonic utterances: Perspective and position shifts

Closer study of polyphonic uses concentrated on the prefacing elements and on the following cotext. The prefacing elements typically introduce a shift in perspective or position, i. e. temporal or attitudinal point of view, pointing at a potential divergence of point of view, to be further explored in the following cotext. Examples (6) and (7) illustrate different types of prefacing and show that the following text develops the diversity and negotiability of the highlighted denominations.

(6)   A distinction must be made between art criticism and what *is more properly regarded as* art journalism. By differentiating one from the other, I do not intend to suggest anything about their respective merits, but merely to establish the important fact that they constituted very different vehicles for conveying information about African American art and artists. (AQ 1999)

(7)   He might even have had his southern home built in what *would have been identified in the 1890s as* the colonial style, depending how literally we take Julius's comments about the younger Murchison's intentions at the end of the story. (AQ 1999)

The examples were classified into two categories, according to whether they were attributable to shifts in perspective or in position, i. e. signalling different perceptions deriving from:
a.   spatio-temporal perspective: *what was once the "bottom" of the town*, *what had initially amounted to individual symbolic behaviors*, *what eventually became the capstone story to "the conjure woman"*, *what was even then an already established finding*;
b.   argumentative (theoretical, ideological, moral) positions: *what she saw as an aggression*, *what Ruth Little-Stokes called "the west Indian house"*, *what is more properly regarded as art criticism*, *what might be expected in the familiar phrase "town and gown"*, *what he deemed to be traditional texts*, *what is perhaps the earliest popular account of success*.

Perspective is characterized by the spatio-temporal coordinates of discourse, as reflected for example in deictic elements (or other elements of the construction of time or space in discourse). Cognitive point of view is strictly

related: the spatio-temporal conditions provide different access to knowledge to writer and historical characters. Looking more closely at an example may help clarify:

(8)   Carnap's new view on analyticity utilised the so-called Ramsey-sentence, which was first proposed by Frank Ramsey in his paper "Theories". In fact, Carnap re-invented *what came to be called* the "Ramsey-sentence approach", where all theoretical terms that feature in a theory are replaced by variables, bound by existential quantifiers. He called it "the existentialised form of a theory", and first presented it at a conference at Los Angeles in 1955. (SIH 2000)

What came to be called is used to signal the typical historical awareness of a double time perspective, introducing a denomination that belongs to later interpretations and not to the historical time in focus (1955): in narratological terms the text flashes forward to the "time of discourse" before going back to the "time of the story" (Chatman 1978). The text that follows elaborates on the changes in perspective predicted by the prefacing element.

Perspective shifts can be signalled by:

a.   shifts in the temporal axis (present/past) of the text, for example by instances of "future in the past" or perfect tenses;

b.   time adverbials: *once*, *initially*, *eventually*, *in the 1890s*, *heretofore*, *now*, *even then*, *already* etc.;

c.   place adverbials marking a shift in space: *what transpired in Nanjin*

d.   relational processes (verbs of state or verbs of change): *be*, *come to*, *become* etc.;

e.   glosses: *what may be the conception of the next century*.

The sequence typically involves a range of verbs of state and verbs of change instantiating Relational processes, as well as Evaluative Categories and signals of perspective shift, as shown in Table 9.

Table 9: Sequences of perspective polyphony.

What	Relational process (+ spatio-temporal shift)	Evaluative category: evoked or inscribed (+ spatio-temporal shift)
What	*would eventually be known as*	*"comfort"*
What	*became*	*one of the country's largest banks*
What	*may be*	*the conception of the next century*

Position refers to attitudinal (emotional and ideological) point of view. It is reflected in attitudinal language and in the cultural frameworks that constitute our argumentative "common places" (*topoi*), i. e. widely accepted values inscribed in a linguistic representation of reality and determining the argumentative value of an utterance (Ducrot 1995: 85). The category is exemplified in (9), where *what may be termed* signals momentary disassociation from the negative evaluation inscribed in the word *paradox* and predicts an analysis of the paradox itself.

(9)   A resulting interpretative shift in the understanding of Augustinian theory brought an awareness of *what may be termed* the Augustinian paradox. Augustine's semiotics, presented in Book 2 of *De Doctrina Christiana*, are a confusing tangle of claims and doubts. (AHR 2000)

The cotext of the *wh*-clause reveals that the predictive function of the item is once again fulfilled: the text that follows the example introduces a detailed account of the paradox, leading to a statement of the final position of the writer on the issue.

Signals of position are varied and typically include markers of attribution and epistemic modality. Here are some examples:

a.   verbs referring to mental processes (expressing opinion): *see*, *regard as*, *deem*, *identify as*, *read as*, *figure as*, *criticize as*, etc.;

b.   verbs referring to verbal processes (verbs of saying): *call*, *describe*, etc.;

c.   other signals of attribution: *according to*, *after*, etc.;

d.   markers of agreement, disagreement or contrast, e. g. adverbials signalling contrasting evaluations: *actually*, *misleadingly*, etc.;

e.   epistemic modal markers: *may*, *could*, *might*, etc.

The semantic sequence created in such contexts involves markers of attribution, verbs referring to verbal or mental processes and signals of alignment/disalignment, as well as evaluative categories, as illustrated in Table 10.

Table 10: Sequences of position polyphony.

What	Source (attribution)+ verbal / mental process+ (dis) agreement	Evaluative category: evoked or inscribed
What	*we call*	*the comforts of life*
What	*Patricia Nelson Limerick has provocatively, if misleadingly, categorized as*	*an "unbroken past"*
What	*I like to call, after Stuart Hall,*	*"the popular"*

When the respective role of perspective and position is considered, position turns out to be largely dominant (accounting for 60/83 occurrences) and focuses especially on different perceptions of historical characters and evaluative implications of different chronomyms, i. e. denominations of specific periods of time. Perspective has a lower incidence (accounting for 23/83 occurrences), and yet it is much more noticeable than in other disciplines (e. g. 3/37 in the BEC and 6/41 in the COCA-A), related as it is to the characteristic features of narrative discourse and to the diachronic dimension of the object of historical inquiry.

Both types of sequences contribute to constructing an authorial voice and become signals of the authoritative persona of the writer. The historian appears here as both Interpreter of events and Arguer of positions. Perspective shifts contribute to the writer's identity as Interpreter and "knowledgeable narrator", emphasising awareness of knowledge that was not available at the time (and place) of the historical narration or previous interpretations. The expertise of the historian is presented in terms of knowledge and capacity to relate different spatio-temporal perspectives. Position shifts contribute to the writer's identity as Arguer and expert member of a discourse community. They foreground the writer's awareness of values and positions that may be inscribed in or evoked by specific expressions in the textual narrative, or that may be part of the reader's own background. The expertise of the historian is presented in terms of judgement and the capacity to identify positions, find an argumentative space for one's own claims and negotiate this with the reader.

From the point of view of textual development, these marked references to shifts in position or perspective typically act as elements predicting an elaboration on the category/ies introduced. They thus become clear markers of the onset of a specific discussion of the issue for the attentive reader.

7 Conclusions: *what* as a "re-defining" relative

Starting from existing literature on evaluation, the study has focused on the discourse of academic history and looked at evaluative patterns that were only marginally dealt with in Hunston and Sinclair (2000), although reflecting similar concerns and functions. Extending the association between *call* and *what* noticed by Cheng et al. (2006) to a global consideration of semantic sequences (Hunston 2008) in nominal *what*-clauses, we have found that our history corpus makes significant use of patterns that involve evaluative meanings and may be called polyphonic for their signalling awareness of a plurality of views.

In specific, we have noticed that *what* often associates with other resources of intersubjective positioning including modality, evidentiality, hedging or

attribution. These may signal a shift in perspective (spatio-temporal point of view) and position (argumentational point of view). Both perspective and position sequences typically signal disassociation from the Evaluative Category, either in the form of an explicit contrast with other possible interpretations or by pointing at the need to clarify terms better. In all forms, however, they contribute to giving evaluation greater discourse prominence, thus preluding to an elaboration on the issue in the cotext that follows.

These evaluative uses of *what* are typically associated with functions of reformulating what has just been said or proposing an interpretation on the basis of what is said in the cotext. Nominal relative *what* may be looked at here as a "Re-defining relative", thus adapting the traditional distinction between Defining and Non-defining relatives. The "Re-defining" function of the relative refers to its use in contexts that place emphasis on the negotiability of representations, typically questioning denominations and the values attached to them.

The incidence of these uses in our corpus of historical journal articles can be explained in terms of the role of interpretation in history and the authorial voice of historians (Coffin 2006; Bondi 2007). The voice of historians is always clearly recording events, but also interpreting them and arguing for an interpretation in the context of a debate. The writer's academic identity will variously highlight these roles and variously foreground the writer's knowledge of facts or the writer's (and reader's) judgement on different positions. The interpreting voice of historians is foregrounded in the double temporal awareness of narrative texts: showing knowledge that is only available to the historian-writer, and not to historical characters, is one way of constructing the writer's persona in terms of expertise, knowledge and access to knowledge. The argumentative voice of historians is highlighted by "re-defining" relatives that create dialogue with the extended discourse community and with the sources. By showing awareness of different interpretations of people and events in history, writers identify an argumentative space for their own discourse, one where they can take position, acknowledge other positions and redirect the argument in the direction that best corresponds to their aims.

More generally, the study was meant to explore ways of integrating lexico-grammatical and lexico-semantic perspectives on discourse. Lexical frequencies, lexical combinations, lexico-semantic patterns and semantic sequences were shown to offer complementary views on discourse. The results suggest that the integration of lexical and semantico-pragmatic analyses can illuminate features of discourse organization and discursive identity, confirming that "small words" (Hunston 2008) can contribute greatly to profiling registers and discourses, when studied in the context of the semantic and pragmatic sequences they are involved in.

The pedagogical implications of the study also deserve attention, with a view to evaluative uses of *what* and semantic sequences in general. The specific sequences studied can be taken to represent important aspects of authorial academic voice in general, thus adding to the increasing interest in developing an appropriate voice in learners (e. g. Hyland and Sancho Guinda 2012). We are dealing here with language resources that help writers and readers identify diversity of perspectives or positions to be elaborated upon in the following text, and therefore with resources that may contribute to developing an appropriate disciplinary voice in student writers and efficient predictive reading in student readers. The presence of these sequences in other disciplinary fields could be explored further, possibly as a sign that interpretative or qualitative elements are at play, as in qualitative research articles or in article sections dealing with placing authorial position in context.

More generally, the perspective offered by semantic sequences in describing discourses might be instrumental in a pedagogical representation of the meanings often expressed in a particular discourse and the range of resources available to student writers. Semantic sequences as such draw our attention to the fact that learning to read and write is learning to combine meanings and language resources in ways that allow learners to understand more efficiently the development of textual interaction and to use more effectively resources that characterize academic discourse.

References

Bakhtin, Michail. 1986. *Speech genres and other late essays*. Austin, TX: University of Texas Press.

Biber, Douglas. 2004. Lexical bundles in academic speech and writing. In Barbara Lewandowska-Tomaszczyk (ed.), *Practical applications in language and computers*, 165–178. Frankfurt am Mein: Peter Lang.

Biber, Douglas, Susan Conrad & Viviana Cortés. 2004. *If you look at*: Lexical bundles in university teaching and textbooks. *Applied Linguistics* 25(3). 371–405.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. London: Longman.

Bondi, Marina. 2007. Authority and expert voices in the discourse of history. In Kjersti Flottum (ed.), *Language and discipline perspectives on academic discourse*, 66–88. Newcastle: Cambridge Scholars Publishing.

Bondi, Marina. 2009a. Perspective and position in museum websites. In Sara Radighieri & Paul Tucker (eds.), *Point of view: Description and evaluation across discourses*, 113–128. Roma: Officina.

Bondi, Marina. 2009b. In the wake of the terror: Phraseological tools of time setting in the narrative of history. In Maggie Charles, Diane Pecorari & S. Hunston (eds.), *Academic writing. At the interface of corpus and discourse*, 73–90. London: Continuum.

Bondi, Marina. 2010. Metadiscursive practices in introductions. Phraseology and semantic sequences across genres. *Nordic Journal of English Studies* 9(2). 99–123.

Charles, Maggie. 2006. Phraseological patterns in reporting clauses used in citation: A corpus-based study of theses in two disciplines. *English for Specific Purposes* 25(3). 310–331.

Charles, Maggie. 2012. "Proper vocabulary and juicy collocations": EAP students evaluate do-it-yourself corpus-building. *English for Specific Purposes* 31(2): 93–102.

Chatman, Seymour. 1978. *Story and discourse. Narrative structure in fiction and film*. Ithaca, NY: Cornell University Press.

Cheng, Winnie, Chris Greaves, John McH Sinclair & Martin Warren. 2008. Uncovering the extent of the phraseological tendency: Towards a systematic analysis of concgrams, *Applied Linguistics* 30(2). 236–252.

Cheng, Winnie, Chris Greaves & Martin Warren. 2006. From n-gram to skipgram to conc-gram. *International Journal of Corpus Linguistics* 11(4). 411–433.

Coffin, Caroline. 2002. The voices of history: Theorising the interpersonal semantics of historical discourses. *Text* 22(4): 503–528.

Coffin, Caroline. 2006. *Historical discourse*. London: Continuum.

Collins, Peter. 1991. *Cleft and pseudo-cleft constructions in English*. London: Routledge.

Dahl, Trine. 2004. Textual metadiscourse in research articles: A marker of national culture or of academic culture?. *Journal of Pragmatics* 36. 1807–1825.

Davies, Mark. 2008. *The corpus of contemporary American English: 520 million words, 1990-present*. http://corpus.byu.edu/coca/.

Declerck, Reneé. 1984. The pragmatics of *it*-clefts and *wh*-clefts. *Lingua* 64. 251–289.

Del Lungo Camiciotti, Gabriella & Elena Tognini Bonelli (eds.). 2004. *Academic discourse: New insights into evaluation*. Bern: Peter Lang.

Deroey, Katrien. 2012. *What they highlight is…*: The discourse function of basic wh-clefts in lectures. *Journal of English for Academic Purposes* 11. 112–124.

Ducrot, Oswald. 1995. Topoï et formes topiques. In Jean-Claude Anscombre (ed.), *Théorie des topoï*, 85–99. Paris: Kimé.

Gledhill, Chris. 2000. *Collocations in science writing*. Tübingen: Gunter Narr Verlag.

Gray, Bethany & Douglas Biber. 2013. Lexical frames in academic prose and conversation. *International Journal of Corpus Linguistics* 18(1). 109–135.

Groom, Nicholas. 2005. Pattern and meaning across genres and disciplines: An exploratory study. *Journal of English for Academic Purposes* 4(3). 257–277.

Groom, Nicholas. 2009. Phraseology and epistemology in academic book reviews: A corpus-driven analysis of two humanities disciplines. In Ken Hyland & Giuliana Diani (eds.), *Academic evaluation. Review genres in university settings*, 122–139. Houndmills: Palgrave.

Groom, Nicholas. 2010. Closed-class keywords and corpus-driven discourse analysis. In Marina Bondi & Michael Scott (eds.), *Keyness in text*, 59–78. Amsterdam: Benjamins.

Halliday, Michael. 1985. *An introduction to functional grammar*. London: Arnold.

Herriman, Jennifer. 2004. Identifying relations: The semantic functions of *wh*-clefts in English. *Text* 24(4). 447–469.

Herriman, Jennifer. 2008. The interpersonal function of clefts in English and Swedish. *Languages in Contrast* 8(2). 143–160.

Holmes, Jasper & Nesi, Hilary. 2009. Verbal and mental processes in academic disciplines. In Maggie Charles, Diane Pecorari & Susan Hunston (eds.), *Academic writing: At the interface of corpus and discourse*, 58–72. London: Continuum.

Hunston, Susan. 2000. Evaluation and the planes of discourse: Status and value in persuasive texts. In Susan Hunston & Geoffrey Thompson (eds.), *Evaluation in text*, 176–207. Oxford: Oxford University Press.

Hunston, Susan. 2008. Starting with the small words: Patterns, lexis and semantic sequences. *International Journal of Corpus Linguistics* 13(3). 271–295.

Hunston, Susan. 2011. *Corpus approaches to evaluation: Phraseology and evaluative language*. London: Routledge.

Hunston, Susan & Gill Francis. 1999. *Pattern grammar*. Amsterdam & Philadelphia: John Benjamins.

Hunston, Susan & John Sinclair. 2000. A local grammar of evaluation. In Susan Hunston & Geoffrey Thompson (eds.), *Evaluation in text*, 74–101. Oxford: Oxford University Press.

Hyland, Ken. 2005. *Metadiscourse*. London: Continuum.

Hyland, Ken. 2008. Academic clusters: Text patterning in published and postgraduate writing. *International Journal of Applied Linguistics* 18(1). 41–62.

Hyland, Ken & Giuliana Diani (eds.). 2009. *Academic evaluation. Review genres in university settings*. Houndmills: Palgrave Macmillan.

Hyland, Ken & Carmen Sancho Guinda (eds.). 2012. *Stance and voice in written academic genres*. London: Palgrave.

Hyland, Ken & Polly Tse. 2005. Hooking the reader: A corpus study of evaluative that in abstracts. *English for Specific Purposes* 24(2). 123–139.

Kyburg, Henry. 1990. *Science and reason*. New York: Oxford University Press.

Luzón Marco, Maria José. 2000. Collocational frameworks in medical research papers: A genre-based study. *English for Specific Purposes* 19. 63–86.

Martin, James R. 2000. Beyond exchange: Appraisal systems in English. In Susan Hunston & Geoffrey Thompson (eds.), *Evaluation in text*, 142–175. Oxford: Oxford University Press.

Martin, James R. & Peter White. 2005. *The language of evaluation. Appraisal in English*. Houndmills: Palgrave.

Matthiessen, Christian. 1999. The system of transitivity: An exploratory study of text-based profile. *Functions of Language* 6(1). 1–51.

Pecorari, Diane. 2009. Formulaic language in biology: A topic-specific investigation. In Maggie Charles, Diane Pecorari & Susan Hunston (eds.), *Academic writing. At the interface of corpus and discourse*, 91–104. London: Continuum.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive Grammar of the English language*. Harlow: Longman.

Scott, Michael. 2012. *Wordsmith tools*. Version 6. Stroud: Lexical Analysis Software.

Sinclair, John McH. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press

Sinclair, John McH. 2004. *Trust the text*. London: Routledge.

Tardy, Christine. 2012. Current conceptions of voice. In Ken Hyland & Carmen Sancho Guinda (eds.), *Stance and voice in written academic genres*, 34–48. London: Palgrave.

Vincent, Benet. 2013. Investigating academic phraseology through combinations of very frequent words: A methodological exploration. *Journal of English for Academic Purposes* 12. 44–56.

White, Peter. 2003. Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *Text* 23(2). 259–284.

Bionote

Marina Bondi

Marina Bondi is Professor of English Language and Translation at the University of Modena and Reggio Emilia, Italy, where she is currently Director of an inter-university research centre – *CLAVIER* (*Corpus an Language Variation in English Research*). The main line of her research has been on genre analysis, English for Academic Purposes and corpus linguistics, with a focus on the argumentative features of academic discourse, metadiscourse, evaluation and phraseology.