CrossMark

An analysis of the evaluation contexts in academic discourse

Xiaoyu Xu[*] and Hilary Nesi

* Correspondence:
xux13@coventry.ac.uk
Department of English and
Language, Coventry University,
Coventry, UK

Abstract

This paper develops a new framework to examine the use of 'evaluation resources' in research discourse. In previous studies of evaluative resources, two distinct, although interwoven, types of context have been identified: 1) the real world where evaluation resources are used to describe situations, and 2) the research world studying the real world, where they are used to engage readers and open up debate. In this paper, we will propose three context types: 1) the research world, where the choice of evaluation must be affected by the writer's argumentative intention, 2) the real world where the choice of evaluation is affected by the writer's argumentative intention, and 3) the real world where the choice of evaluation is not affected by the writer's argumentative intention. This distinction helps us identify more rigorously those evaluative resources that are relevant to arguments. It will also enable us to explore and quantify the way evaluation functions across contexts.

Keywords: Academic writing, Research space, Real world, Research world, Planes of discourse, Appraisal theory

Introduction

In research writing (research articles, proposals and even dissertations and theses), it is considered important to establish a reason for the research that is undertaken. Swales' Create a Research Space (CARS) model for research article introductions (1990, 2004) is widely adopted in the teaching of research writing (see for example Cotos 2014; Swales & Feak 2012), and introduced the notion that writers should establish a territory for their research, and identify a gap or a 'niche' that their research can fill (Swales 1990:141). Swales' model reveals research article introductions to function primarily as a means of engaging and persuading the reader; in order to achieve this end they require extensive use of evaluative resources.

The original CARS model implied that the researcher could only occupy a research space by establishing the centrality of the research topic and/or identifying knowledge gaps relating to the prior research. In a later study, however, Samraj (2002:10) examined Conservation Biology research article introductions in terms of the CARS model and pointed out that the author's own research was sometimes justified through claims about the state of the environment and environmental problems, rather than claims about research centrality and research gaps.

This paper will propose a theoretical framework which acknowledges the possibility of a real world space for the researcher to occupy, alongside the familiar Swalesian

'research space'. The contrast between 'real world' and 'research world' evaluation contexts was briefly mentioned by Samraj (2002) but has been examined more extensively in a number of studies relating to evaluation, such as those of Sinclair (1981), Thetela (1997), Hunston (2000) and Hood (2004). These studies have contributed to our understanding of the way evaluation differs from context to context, and have been widely applied to the analysis of academic texts (see for example, Atai & Falah 2006; Cava 2007; Millán 2012; Shaw 2003; Wu 2008). The studies do not, however, systematically consider the purposes of evaluation within specific moves, in specific genres. This paper aims to build on existing frameworks and the valuable insights provided by Sinclair, Thetela, Hunston and Hood while identifying and addressing some inherent problems with the way they have treated context, bearing in mind the role context plays within the move structure of the research article.

Much anthropological and linguistic research (see, for example, Bakhtin 1981; Hyland 1996; Martin & White 2005; Myers 1989; Swales 1990, 2004) has justified the prevalence of evaluative resources in academic writing on the grounds that academic writing is not solely 'monologic' (i.e. presenting a single interpretation that is not open to discussion), but is also interpersonal and interactive (i.e. offering opportunities for alternative interpretations). Researchers have examined evaluative resources from a variety of perspectives, both in terms of their function (Chafe & Johanna 1986; Hyland 1999; Martin & White 2005; Ochs & Schieffelin 1989; Thompson & Hunston 2000) and in terms of their linguistic form (Biber et al. 1999; Dressen 2003; Hunston 1993, 1994; Tucker 2003; White 2003). The two distinct, although interwoven, types of context that have been identified for these resources in academic writing are the real world being studied, and the research world that studies the real world. These two types are illustrated via the word 'ignore' in example (1) (where it relates to the real world) and example (2) (where it relates to the research world).

(1) Some students **ignored** the photocopies while some others analyzed them either individually or in groups.

(2) Academic literacies research that **ignores** the nature of the texts themselves misses an important source of insights into literacy practices, …

Intuitively, the word 'ignore' in example (1) is much less relevant to the writer's rhetorical purposes and construction of stance and voice than the word 'ignore' in example (2). If a text applies a large number of evaluative resources of the type in example (1), an analysis of this text might wrongly conclude that the writer is being judgmental unless the level of relevance to the writer's rhetorical purposes is taken into account. Therefore, it is not enough simply to count the total number of occurrences of evaluative resources in a text; resources at different levels of relevance to the construction of stance must be counted and analyzed separately. These levels of relevance are highly associated with the context (in the research or the real world), so context is an important consideration when examining the way a particular evaluation resource functions in a text.

The most influential studies of the effect of context on evaluation in academic discourse, those of Sinclair (1981), Thetela (1997), Hunston (2000), and Hood (2004), have taken slightly different approaches to this topic. Moreover, they do not consistently distinguish between the two types of context, and for this reason their approaches are

difficult to apply to new texts. John Sinclair (1981) proposed the existence of an 'autonomous plane' and an 'interactive plane', a notion that was intended to help analysts identify what was being evaluated, and what roles writers and readers were performing. On the autonomous plane Sinclair thought of the writer as an informer who told the reader of the content of the text, while simultaneously on the interactive plane Sinclair thought of the writer as a text-constructor, informing the reader about the structure of the text. Sinclair's planes might possibly be compared to the three Hallidayan metafunctions operating simultaneously within the text, the textual and interpersonal metafunctions managing discourse flow and the interaction between writer and reader, and the ideational metafunction enabling us to construe human experience (Halliday 1994). Building on Sinclair's idea, Hunston (2000) regarded interactive evaluation (the interactive plane) as an aid to the construction, in the whole text, of the argument which aimed to persuade the reader. On the other hand, in Hunston's (2000: 195) view evaluations bestowed on the outside world (the autonomous plane) reflected the ideology of the section of society from which the writer came. Hunston analyzed planes of discourse alongside other perspectives: attribution (language or thoughts presented as deriving from someone other than the writer), averral (the writer's own language or thoughts), and different types of statement (of fact, interpretation, assumption, assessment or recommendation).

In an earlier study, Hunston (1993) had developed the idea that any evaluation that contributes to the achievement of a particular goal will be positive, whilst any evaluation that holds back from that achievement will be negative. Hunston (1993) did not consider context, but her notion of positive and negative evaluation was later used by Thetela (1997) to examine context in academic research articles. Thetela invented the term 'topic-oriented evaluation' (TOE) to describe evaluations related to the real world (as in example (1) above), and the term 'research-oriented evaluation' (ROE) to describe evaluations directly related to the research discourse and its purpose (as in example (2) above). Later Hood (2004) developed Thetela's notion of ROE and TOE, but used slightly different terminology. She replaced the term ROE with 'Field of Research' (FR), defined as intrinsic to the task of introducing and positioning the writer's own research, and replaced the term TOE with 'Field of Domain' (FD), defined as the set of activities that are the focus of the writer's study.

These ideas draw attention to the role of context within evaluation. However, it is difficult to decide which of the four approaches taken by Sinclair, Hunston, Thetela and Hood is the best one to replicate in further studies, as they have all led to slightly contradictory conclusions. For example, Sinclair (1981) claimed that every sentence in a text operates on each plane simultaneously (in Sinclair's terms, examples (1) and (2) would both operate on both planes), while Hood (2004) claimed that sentences do not necessarily operate in both fields (in Hood's terms, example (1) is FD and example (2) is FR). The different studies also employ slightly different analytical tools: Thetela (1997) and Hood (2004) concentrated solely on contexts, but in Hunston's work, planes of discourse, attribution/averral and different types of statement were all analyzed together, making it difficult to isolate findings that are relevant solely to the planes of discourse. Moreover, although Thetela and Hood claimed that their categories were mutually exclusive, some examples in their data seem to fit both ROE/FR (Research-Oriented Evaluation/Field of Research) and TOE/FD (Topic-Oriented Evaluation/Field of Domain) equally well.

This paper aims to establish a more replicable way of distinguishing context types in evaluative discourse. Thus, our objectives are as follows:
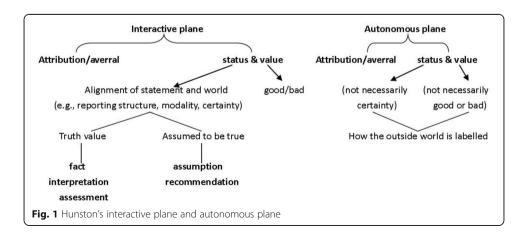
a. To compare and contrast the different types of contexts discussed by Hunston (2000), Thetela (1997) and Hood (2004) in order to identify those aspects of the theory that require further development and clarification.
b. To propose an improved theoretical framework that resolves some of the ambiguities of the previous systems.

A comparison of different approaches to the analysis of context

In this section, we will compare in more detail the way that Hunston (2000), Thetela (1997) and Hood (2004) distinguish between different types of context, with reference to the texts they themselves have analyzed. To begin with, we will look at the distinction between the interactive and autonomous planes, first introduced as a concept by Sinclair (1981) and then developed by Hunston (2000).

Interactive plane vs. autonomous plane

In Sinclair's model (1981) all sentences can function simultaneously on both the interactive plane, on which the writer informs the reader about the structure of the text, and the autonomous plane, on which the writer informs the reader about its content. Hunston (2000: 183) pointed out that the key distinction between the two planes lies in what is being evaluated – "a discourse act in the discourse itself (interactive) or something else (autonomous)". However, Hunston's criteria for identification involve other perspectives such as attribution/averral, and different types of statement (of fact, interpretation, assessment, assumption and recommendation). She suggested that 'attribution' presents the ideas of other people, and attaches different levels of credence to different pieces of information. When a given idea is attributed, the responsibility of the writer decreases and is delegated to the attribute. Hunston also considered the status of the statement, taking into account the variable alignment of 'world', the truth-value of which is evaluated by the writer, and 'statement', assumed for the moment to be true rather than argued. Figure 1 is our attempt to summarize the ideas Hunston expresses in her research, showing the ways in which rhetorical choices are made.

Fig. 1 Hunston's interactive plane and autonomous plane

Hunston (2000: 198) claimed that the advantage of analyzing text at the level of planes is that it "demonstrates the interaction between the world of the text and the world outside it". However, the additional examination of attribution/averral and the status of statements, although important for the study of evaluation, may not always be clear what is relevant solely to planes of discourse and not to the other perspectives, and therefore hinder analysis of the ways the autonomous and interactive planes interact.

In Hunston's system of analysis the two planes function in parallel rather than being mutually exclusive. As Hunston (2000: 183) argues, "every sentence in a text operates on each plane simultaneously, although some sentences draw attention to their status on the interactive plane more explicitly than others". She explained further that "if we take sentence [1.1]….. as a claim … which is evaluated in subsequent sentences, we are seeing sentence [1.1] from the point of view of the writer-as-text-constructor", but "if we take sentence [1.1] as a comment on certain things other than this discourse, including other discourses, we are seeing it from the point of view of writer-as-informer". Thus, in her annotation, every sentence is on the interactive plane and all parts marked as autonomous are also marked as interactive, so a considerable amount of the text functions on both planes. There is no denying that the overlap within Hunston's framework serves a purpose; it allows for the fact that one text (or one sentence, or one word) can simultaneously pertain to both the real and the research world. However, Hunston's approach fails to reveal the interesting complementarity between real and research world resources. This complementarity deserves investigation as it can shed light on the ways evaluation is achieved. For example, writers might intentionally focus on the real or the research world context in order to better establish a niche for their own investigations, and it would be easier to compare these two different evaluative roles if it was impossible for any stretch of text to belong to both.

Research-Oriented Evaluation (ROE) vs. Topic-Oriented evaluations (TOE)

Thetela (1997) provided a simpler and more straightforward distinction between research and real world contexts, but only considered evaluations and evaluated entities where positive or negative attitudes are made explicit (in other words those that are positively or negatively 'inscribed'). She called evaluations directly related to the research discourse and its purpose 'research-oriented evaluations' (ROE), and evaluations related to the real world being investigated 'topic-oriented evaluations' (TOE).

ROE and TOE are mutually exclusive, unlike Hunston's autonomous and interactive planes. They are therefore much more distinct. This is because Thetela took a narrower view of the interactive plane, leaving out what is being investigated and focusing solely on the investigation itself. In Thetela's analysis, the writer therefore functions only as an observer of the research.

The distinction between Hunston's annotation and Thetela's is illustrated in examples (3) and (4), where the same excerpt has been annotated according to the two different systems. In example (3), Hunston identifies Sentence [1] as an averred assessment which is supported by the evidence in sentence [2]. Therefore, both sentences are on the interactive plane although both of them include the real world being investigated. (The interactive plane is in **bold**, the autonomous plane is in *italics*.)

(3) **[1] Right now a new *wave of anti-sect paranoia is sweeping the world*. [2] All ruling bodies, political parties and the media seem unanimous in *their suspicion and hostility towards sects* and any *group of people labelled a 'sect' are automatically viewed with prejudiced eyes.***

However, according to Thetela's system, Sentences [1] and [2] should be annotated as follows, with TOE elements in *italics*. No ROE is found.
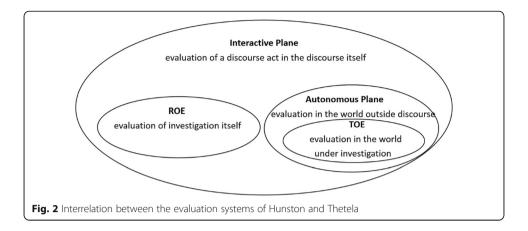
(4) [1] Right now a new wave of anti-sect paranoia is *sweeping the world*. [2] All ruling bodies, political parties and the media seem unanimous in *their suspicion and hostility* towards sects and any group of people labelled a 'sect' are *automatically viewed with prejudiced eyes.*

Thetela's key criterion for distinguishing between TOE and ROE is the idea that with ROE it is the researcher who performs the evaluation, whereas a positive or negative TOE is something that the researcher simply reports, and is "neither good or bad in itself" (Thetela 1997: 105). Positively or negatively evaluated topics are both equally interesting and worth investigation, however.

Broadly speaking, Thetela's ROE is the same concept as Hunston's interactive plane, in that in both cases it marks instances where the writer interacts with the reader. Hunston was interested in the role of evaluation in persuasive discourse generally, however, whereas Thetela was solely concerned with evaluation in research articles and because of this was able to narrow the scope of her analysis. The writer of a research article engages with the research community, and the research community judges the research process when they read the research article. The writer's purpose is therefore to negotiate the perspective "from which the research should be judged" (Thetela 1997: 105). However, although Thetela differed from Hunston in that she paid greater attention to the research process specifically, she retained Hunston's view of the constructive role played by interactive discourse, in that ROE contributes to a global evaluation.
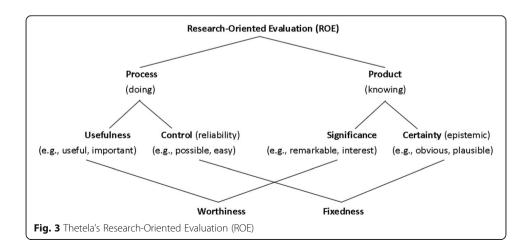
As well as regarding the writer as a research observer, community communicator and discourse constructor (through his or her use of ROE), Thetela also specified the role of the writer as a real world observer (through his or her use of TOE). TOE entities are evaluated and described in research papers, but they neither constitute the research itself nor contribute to the research purpose. The interrelation between Hunston and Thetela's distinctions are represented in our Fig. 2.

Thetela thought of ROE, the research process proper, as consisting of two types, process and product. Her idea of research process related to how the research is done, and how it is reported and interpreted with reference to its usefulness and reliability. She related research product, on the other hand, to the results of the research, generally reported and interpreted with reference to significance and certainty. The evaluation of usefulness and significance, which is called 'worthiness' by Thetela, can be achieved using linguistic items such as 'useful'; 'important', 'remarkable' and 'interest'. The evaluation of certainty and reliability, which is called 'fixedness' by Thetela, can be achieved using linguistic items such as 'possible' and 'obvious'. It is possible to express Thetela's whole system diagrammatically, as in Fig. 3.

Fig. 2 Interrelation between the evaluation systems of Hunston and Thetela

Although there are clear criteria for breaking ROE down in this way, Fig. 3 is incomplete because the categories of process and product do not cover all possible types of ROE. In a research article, for example, a niche (i.e. a gap) usually needs to be established before reporting and interpreting the research process, in order to justify the need for the research (Swales 1990, 2004). Evaluating this gap is not the same as evaluating methods or findings, however, and therefore the evaluation of the gap does not belong in either the process or product categories. This is where Thetela's system runs into difficulty.

Typically, in experimental research a niche is identified either by evaluating the prior research (or lack of prior research) (ROE) or by evaluating the real world being investigated (TOE). However, although Thetela (1997: 105) claimed that when evaluating the real world (TOE) "both good or bad performances are equally interesting topics worth investigating", when a niche is identified in the real world (for example the fact that vocabulary is often **ignored**, as in example (5) below), the evaluation has to be negative in order to justify the research. For this reason it can be said that good or bad performances in the real world are not always equally interesting and worthy of investigation; it is therefore important to distinguish between the two types of real world context – affected and unaffected by the writer's argumentative intention. This problem also arises in other cases, for example when the writer identifies and evaluates the possible

Fig. 3 Thetela's Research-Oriented Evaluation (ROE)

impact of the research (as in example (6)). That is to say, in practice ROE and TOE overlap rather than being mutually exclusive.

(5) Unfortunately, vocabulary is often **ignored** and students are bogged down with a dilemma of guessing words in the EFL teaching context. (Purpose: establishing a research gap)

(6) Although provisional, our model has implications for pedagogy. First of all, the role of L2 vocabulary listening comprehension achievement is **important** information for teachers. (Purpose: identifying possible impact)

The Field of Research (FR) vs. the Field of Domain (FD)

Hood (2004) used the term 'Field of Domain' (FD) rather than Topic Oriented Evaluation (TOE) when referring to the set of activities that are the focus of the writer's study, but her 'Field of Research' (FR) was somewhat different from Thetela's notion of Research Oriented Evaluation (ROE). ROE only considers the process and product of the investigation, but FR constructs a more exhaustive set of research activities, including the identification of research issues (an example of this might be 'establishing a research gap') and the interpretation of findings and outcomes (an example of this might be 'identifying possible impact').

Hood's FR and FD categories, like Thetela's ROE and TOE, can be applied to texts independently of any broader analytical approach. For example, they can be incorporated into Systemic Functional Linguistics, but can equally well be used with other methods of discourse analysis. Hood, like Thetela, is concerned with the categorization of entities and their ascribed evaluations, but, like Hunston, this is not her sole concern. Hood's extension of linguistic boundaries requires more criteria to identify parts of the discourse that Thetela had ignored. For this purpose, Hood identified certain lexical items typical of FR or FD in her sample texts. For example, 'produced', 'found', 'discovered', 'identified', 'achieved', 'findings', 'results', 'evidence', 'answer', etc. were words associated with FR, while 'students', 'performance', 'better', 'worse', etc. were words associated with the FD. Lexis is not always a reliable indicator, however, as lexical meaning can vary according to context. For example, the 'results' of a language exam for students might refer to an FR context or an FD context, depending on whether they were treated as the results of the writer's study or simply as evidence from the real world. Similarly, the FD-associated word 'performance' might refer to prior research performance, in which case it should be categorized as FR.

In order to explain the relationship between clauses constructive of FD and clauses constructive of FR, Hood also proposed a new way of considering projection, drawing both on 'grammatical' projection (Halliday 1994) and 'metaphorical' projection (Christie 1997). In Hood's terms, FD and FR are connected through a range of congruent and non-congruent realizations, where 'sayers' or 'sensers' project 'locutions' or 'ideas' (Halliday 1994). For example, although 'suggest', 'discover' and 'show' were associated by Hood with FR, the locutions or ideas that sayers or sensers suggested, discovered or showed were associated with FD. However, as Hunston (2000) and Thetela (1997) pointed out, the distinction between research and real worlds is made with reference to discourse rather than to grammar, and is thus often open to interpretation. For

example, we argue that in example (7) below (where Hood marked FR in **bold** and FD in *italics*), although 'social actions and interaction' and 'organizations of the conversations' belong to the real world which is being investigated, the rest of the sentence assesses the methodology of the research itself. According to this interpretation, it should be analyzed as in example (8).

(7) *In this sense, all aspects of social actions and interaction can also be examined by looking at the organizations of the conversations* **(Heritage, 1989).**

(8) **In this sense,** *all aspects of social actions and interaction* **can also be examined by looking at** *the organizations of the conversations* **(Heritage, 1989).**
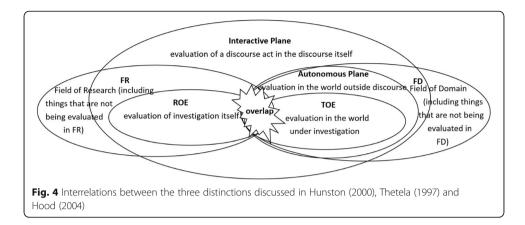
In some cases, Hood's system encounters another difficulty in that the two categories (FR and FD) are not mutually exclusive, just as in Thetela's system. In example (5), for example, the evaluation 'vocabulary is often **ignored**' operates simultaneously in both the FD and the FR. This kind of evaluation is of the activity being investigated (which is a feature of FD), but also serves to justify the necessity of the writer's own research (which is a feature of FR).

Hood herself also noticed that some entities can fit in both fields. She argues that entities such as 'conclusion', 'relationship', 'effects', and 'similarities' are observations of phenomena in the FD, but also present aspects of the FR. In the course of coding, Hood chose between FD and FR by identifying the dominant field in the co-text. In example (9), she considered the words in **bold** to be constructing in the FR, for example, while in example (10), she considered the words in **bold** to be constructing in the FD.

(9) Of the many who have looked at the **relationship between age and performance at universities**, none has as yet produced a definite answer to the question …

(10) …the results have indicated that the **relationship between age and performance** is not a linear one

Although in example (9), Hood considered 'the relationship between age and performance at universities' as FR, her decision does not change the fact that this part of the sentence is related to the real world activities being investigated. This method of identification is not entirely in line with Hood's definitions for the two fields, and thus may be difficult to replicate.

The interrelations between the three systems created by Hunston (2000), Thetela (1997) and Hood (2004) are illustrated in our Fig. 4. The conceptual perspective of the interactive plane is larger than that of ROE and FR in that it accounts for the interactive construction evident in every sentence in the text, while ROE and FR only account for the investigation itself, which is only referenced in some parts of the text. FR is larger than ROE in the figure because FR functions in projecting relations, representing not only every situation that is being evaluated but also all the situations that are not being evaluated, whereas ROE is merely concerned with ascribed evaluations and the entities they evaluate. The autonomous plane and TOE, on the other hand, are presented within the larger category of FD, because FD functions in projecting relations representing not only every situation that is being evaluated but also the situations that are not being evaluated. The linguistic realizations on the autonomous plane also

Fig. 4 Interrelations between the three distinctions discussed in Hunston (2000), Thetela (1997) and Hood (2004)

function on the interactive plane, but this does not necessarily happen the other way round. For this reason, the autonomous plane is placed within the interactive plane in Fig. 4. As noted previously, there are overlaps between ROE and TOE, and between FR and FD, although Thetela and Hood both claimed that their systems were mutually exclusive. This is the biggest problem with their context categorization systems.

The ambiguities of the three systems created by Hunston (2000), Thetela (1997) and Hood (2004) can be summarized as follows: inconsistent conceptual perspectives, non-mutually exclusive categories, and insufficient identification criteria. Building on the prior approaches, these ambiguities will be addressed in the new framework outlined below.

Aspects of the new theoretical framework

In order to build a valid theoretical framework, three issues need to be resolved:

1) the overlap between the two contexts (real world and research world) identified in the prior research needs to be accounted for in some way
2) 'context' needs to be defined with respect to evaluative resources within academic texts
3) criteria need to be set for the identification of each context.

In what follows, these issues are discussed with reference to extracts from research articles in applied linguistics. Research articles have been chosen because of their importance as a channel for the communication of investigative research.

The overlap between the two contexts

Probably the easiest solution to coping with the overlapping resources is to group them into a third 'hybrid' category. However, the problem of the overlap between Hunston's research world and real world contexts cannot be resolved simply by creating a third 'hybrid' category for elements that belong in both contexts, because such a category would result in the fragmentation of propositions (sentences) that in their entirety construe the interaction context. Hunston (2000: 203) presented her sample text as being entirely on the interactive plane, although large parts could also be judged to function on the autonomous plane. In example (11) (the first two sentences from this text), for example, she identified interactivity in the way Sentence [2] provides evidence for

Sentence [1]. However, if we isolate the parts of these sentences which function simultaneously on both planes, and treat them as belonging to a hybrid category, only a few fragments such as 'Right now a new', 'All', 'seem unanimous in' and 'and any' remain on the interactive plane. These are so fragmentary that it is no longer possible to construe any interaction.

In example (11) the interactive plane is in *italics*, the possible hybrid category is in **bold**. No items on the autonomous plane have been identified.

(11) [1] *Right now a new* **wave of anti-sect paranoia is sweeping the world**. [2] *All*
ruling bodies, political parties and the media *seem unanimous in* **their suspicion and hostility towards sects** *and any* **group of people labelled a 'sect' are automatically viewed with prejudiced eyes.**

Unlike Hunston's system, the overlapping resources encountered by Thetela's system and Hood's system can be grouped into a new context. For example, the evaluative items, 'ignore' and 'important', in examples (5) and (6) can be categorized as a hybrid context without causing any new problems. Therefore, three contexts will be considered in the new framework instead of two.

Defining 'context' with respect to evaluative resources

Our intuition about the existence of real world contexts and research world contexts in investigative discourse is supported by the fact that any investigation must include two components: the question of *how* to investigate (in other words the investigative process), and the question of *what* to investigate (in other words the matter being investigated). Thetela (1997) and Hood (2004) associate the question of *how* to investigate with the research world, and the question of *what* to investigate with the real world. This categorization is sometimes problematic, however, because what to investigate might, in fact, be associated with the real world or the research world, given that it is possible to investigate either real world or theoretical matters. Indeed, it is possible to regard theory in a research article as pertaining to both what is being investigated and the method of investigation. Thus in example (12), taken from an introduction to an applied linguistics research article, the 'academic literacies research' relates to theory, and the claim that it 'ignores something' constructs an argument which will be used to justify the selection of a better method of investigation (thereby evaluating investigative processes).

(12) Academic literacies research that **ignores** the nature of the texts themselves misses an important source of insights into literacy practices.

In the context of the research world, any positive or negative evaluation has to be in line with the argument the writer is constructing, for example in order to align or distance readers, and/or to establish a research gap. That is to say, the writer takes full responsibility for choosing a particular evaluation. In example (13), taken from an introduction to another applied linguistics research article, 'well-established' (in **bold**) can be categorized as an evaluation of 'PPP', the research world topic under investigation:

(13) PPP is **well-established** in mainstream ELT methodology but has attracted a lot of criticism.

In examples (14)–(16), also taken from research article introductions, the underlined segments also evaluate research world topics, but in these cases the research world manifests itself in references to prior research (Stone's work in example (14), the current research ('this article') in example (15), and the development of a contrasting train of thought example (16)).

(14) Stone (2003) **found** that, over time, ..., which **suggested** that ....
(15) This article is a step toward **bridging this divide**, **offering insight** into both ....
(16) **However**, we wondered whether .....

Unlike references to the research world, which can relate to the matter under investigation or the process of investigation, references to the real world can only concern the matter under investigation, in the manner illustrated in examples (17) and (18). Example (17), taken from the results section of a research article, is more straightforward in this respect. The evaluation (in this case manifested through the word 'ignore') does not serve to support the writer's concluding argument that the teacher should take actions to help those students who had a negative attitude towards independent learning. The writer is simply observing a real world situation, and had no influence over what the learners did or did not do. A negative evaluation does not change the legitimacy of the investigation, or the relationship between writer and reader.

(17) Some students **ignored** the photocopies while some others analyzed them either individually or in groups.

Every evaluation is of course subjective, and other members of the research community may still query the writer's claim that 'some students **ignored** the photocopies'. However, a researcher's evaluation of a real world situation is more likely to be arguable in terms of whether it is true or false, whereas the researcher's choice of evaluation of the investigation itself is more likely to be arguable in terms of whether it is supported by valid reasoning.

Although it seems that the evaluation of real world situations does not help to forward arguments in research articles, there may be exceptions to this. Example (18) (from a research article introduction) refers to a real world context in the same way as example (17), but it plays a different role – that of 'establishing a niche' for the current research by identifying a problem that the researcher can help to resolve.

(18) Unfortunately, vocabulary is often **ignored** and students are bogged down with a dilemma of guessing words in the EFL teaching context.

A negative evaluation is a typical means of justifying current research; Swales' Move 2 in research article introductions, 'Establishing a Niche', for example, indicates limitations to the existing state of knowledge through the use of adversatives and various forms of negation (Swales 1990:154–5). Similarly, if the writer decides to focus on a real

world phenomenon as a justification for the current research, this phenomenon is likely to be presented in a negative light, so that the writer's research contribution can be presented as a way of improving the situation. The positivity or negativity of the evaluation matters to the argument, regardless of whether the evaluated entity belongs in the real or the research world.

Conversely, evaluations in the conclusions to research articles tend to be positive, as they relate to the writer's contribution to the research field and/or the effect of the research on the real world. In the case of example (19), taken from the concluding section of a research article, the current real world situation (with bold italic) is negatively evaluated, but a predicted future real world situation (in *italics*) is evaluated positively.

(19) This research uncovers the areas of ***persistent challenge*** to EFL learners across different proficiencies, e.g., (correct) use of conjunctions and flexible manipulation of lexical elaboration, which *may merit heed of many* EFL teachers and they can *frame* their future teaching and inquiries to *help* learners *overcome these weaknesses.*

This kind of positive evaluation of a future outcome is more likely to pertain to the real world than the research world. It relates to impact, defined by the UK Economic and Social Research Council (n.d.) as 'the demonstrable contribution that excellent research makes to society and the economy'.

Real world entities do not need to be tangible. Anything that can be evaluated, and which is not a research procedure, theory or framework, can be regarded as falling into the 'real world' category. Thus abstract concepts such as 'relationships' can be considered as pertaining to the real world, and can be given value in order to create a 'research space' for the writer. Swales' examples of Move 1 ('Establishing a Territory') (1990, 2004) include these kind of abstract real world concepts, especially as a means of claiming centrality (Step 1). In example (20), taken from the introduction to a research article, 'the relationship between ideology, context of culture, context of situation and language' is evaluated positively as an 'appealing area' and is a kind of 'topic generalization', Step 2 of Move 1 (Swales 1990: 141).

(20) Systemic functional linguistics offers a number of different models of the **relationship between ideology, context of culture, context of situation and language**, and research on modeling of context continues (Hasan, 2009). Indeed, context of situation as a construct has been criticized by those within and outside SFL for being vague and indeterminate (Bowcher, 2010, p77). Nevertheless, it is an intuitively **appealing area** of research.

Thus, linguistic resources pertaining to the real world represent situations that happen/exist, have happened/existed, are happening/existing, or can happen/exist in the real world. In example (17) the evaluation of the real world entity (learner behavior) did not serve to support the writer's argument, but in other examples the writers take responsibility for their evaluations and use them to create a research space (examples (18)–(19)) or predict impact (example (20)).

Evaluations of the real world can serve the writer's purpose as a means of negotiation with the research community. This is achieved by aligning the writer with the reader or with prior researchers, or by distancing the writer from the reader or the prior research, and by creating a research space, or predicting positive research impact. The work of Thetela and Hood, however, assigns a negotiating function only to the evaluation of entities in the research world, and conversely treats real world evaluations as non-negotiable. Thus, we argue that Thetela and Hood's contextual framework, illustrated in Fig. 5, should be adjusted to take into account the negotiable aspects of real-world evaluation, as in Fig. 6.

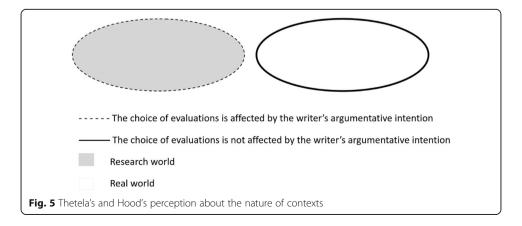Figure 6 identifies not two but three distinct contexts:

1) Context 1 refers to the research world (the investigation process) where the choice of evaluation is affected by the writer's argumentative intention (−real + argumentative intention).
2) Context 2 is the new category, referring to real world situations where the choice of evaluation is affected by the writer's argumentative intention. (+real + argumentative intention)
3) Context 3 refers to real world situations where the choice of evaluation is not affected by the writer's argumentative intention. (+real -argumentative intention)
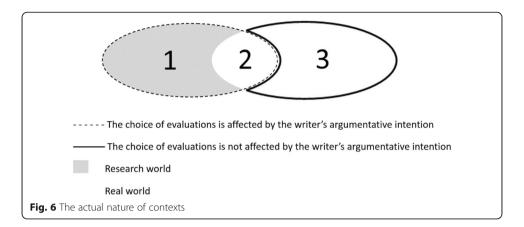

Criteria need to be set for the identification of each context
Before identifying each context, researchers first need to choose an approach they prefer to help with identifying evaluative resources. In this paper, we choose to draw on the Appraisal theory, developed by Martin and White (2005) within Systemic Functional Linguistics. Therefore, in example (21), we identified 'according to' and 'essential' as evaluative items.

(21) According to the framework of test method facets (Bachman, 1990), genre is an essential element in the input test takers receive.

The next step is to identify in which world (research or real world) 'according to' and 'essential' function as evaluations. We have established that what we term the 'research world' is related to the investigative process, whereas what we term the 'real world' is

- - - - - - The choice of evaluations is affected by the writer's argumentative intention

———— The choice of evaluations is not affected by the writer's argumentative intention

Research world

Real world

Fig. 5 Thetela's and Hood's perception about the nature of contexts

Fig. 6 The actual nature of contexts

related to anything being investigated that has happened/is happening/can happen/ can exist in the real world. It is not difficult to distinguish them. In example (21), 'according to' is a reference to the research world and 'essential' is a reference to the real world.

The third step is to identify whether the real world evaluation 'essential' is affected by the writer's argumentative intention (i.e. Context 2: +real + argumentative intention). However, intuition alone is not sufficient for identification; there needs to be some systematic way of testing the function in order to avoid individual variations in interpretation. One way of doing this is to replace any positive evaluation with a negative evaluation, or replace any negative evaluation with a positive evaluation, and seeing what effect this has on the meaning of the text. For example, in example (21), 'genre', as an element in the input test takers receive, is evaluated as 'essential', and if we replace 'essential' with 'useless', a negative evaluation, this will undermine the writer's effort to justify the importance of examining genre. Therefore, 'essential' should be classed as Context 2 (+real + argumentative intention).

Example (22) illustrates how a paragraph is marked up using the criteria given above. References to Context 1 are **in bold**, references to Context 2 are ***bold italics***, references to Context 3 are in *italics*.

(22) This study is also **potentially relevant** to the creation of equivalent versions of tests. This application is ***particularly important*** in view of the trend of allowing students to choose between two or more tests or tasks in one test administration (e.g., Michigan English Language Assessment Battery (MELAB); di Gennaro, 2009; Plakans, 2009), which has received research attention (e.g., Hamp-Lyons & Kroll, 1996; Hamp-Lyons & Mathias, 1994; Peretz & Shoham, 1990; Polio & Glew, 1996). The present study **found** that *different* genres elicited *different* performances from examinees, and their task performances *conflicted* with their perceptions of task difficulty. These findings **might** have effects on test reliability, validity and fairness. Hence, these findings call for attention from test writers who **might** *want* to use ***different*** text types when designing ***equivalent*** summary writing tasks in a test.

Inevitably there will still be some degree of subjectivity in this procedure, because a writer's purpose is always open to interpretation.

Conclusion

This study has analyzed the role of contexts in persuasive academic writing in terms of the function of evaluative resources. We have seen that evaluative resources can create different communicative effects depending on what is being evaluated. Scrutiny and comparison of the most inspiring prior research (Sinclair 1981; Thetela 1997; Hunston 2000; Hood 2004) has led us to the formulation of a new, practical distinction between three categories of context: −real + argumentative intention, +real + argumentative intention, and + real − argumentative intention. Each of these categories is mutually exclusive, and it is therefore easier to identify and compare the evaluative resources within each category. In practice, it enables us to focus on those evaluative resources which are most interesting from a functional perspective – i.e. those that are affected by the writer's argumentative intention, and filter out those evaluative resources which do not serve arguments.

Samraj (2002:15) has added additional options to Moves 1 and 2 of the original CARS model (Swales 1990: 141), so that in Move 1 (Establishing a Territory) claiming centrality can involve the research world and/or the real world, and similarly in Move 2 (Establishing a Niche) indicating a gap can involve either or both of these worlds. The importance of evaluative context has not yet been recognized, however, in materials for teaching and learning academic writing, and does not appear to feature in second language learning syllabuses. Having identified these three categories, a next stage will be to apply this system to the analysis of a considerable quantity of research writing – not only research articles, but also proposals, dissertations, and other types of persuasive writing which concerns research activity. It should then be possible to identify the types of evaluative resource most frequently associated with each of the three categories, and to feed this information into EAP syllabuses and materials for the teaching and learning of research writing.

Abbreviations
CARS: Create a research space; FD: Field of domain; FR: Field of research; ROE: Research-oriented evaluation; TOE: Topic-oriented evaluation

Availability of data and materials
Examples are extracted from the following research articles

Boyno, Mehmet. 2012. Some pedagogical innovations in nurturing autonomous EFL learners. *Energy Education Science and Technology Part B: Social and Educational Studies* Special Issue 2: 197–204.

Elwood, James A. and Jeroen Bode. 2013. Student preferences vis-à-vis teacher feedback in university EFL writing classes in Japan. *System* 42: 333–343.

Gardner, Sheena. 2012. Genres and registers of student report writing: an SFL perspective on texts and practices. *Journal of English for Academic Purposes* 11: 52–63.

Jones, Christian and Ronald Carter. 2013. Teaching spoken discourse markers explicitly: a comparison of III and PPP. *International Journal of English Studies* 14: 37–54.

Li, Jiuliang. 2014. Examining genre effects on test takers' summary writing performance. *Assessing Writing* 22: 75–90.

MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4: 1–16.

Teng, Feng. 2014. Assessing the depth and breadth of vocabulary knowledge with listening comprehension. *PASAA* 48: 29–56.

Vandergrift, Larry and Susan Baker. 2015. Learner variables in second language listening comprehension: an exploratory path analysis. *Language learning* 65: 390–416.

Yang, Wenxing and Ying Sun. 2011. The use of cohesive devices in argumentative writing by Chinese EFL learners at different proficiency levels. *Linguistics and Education* 23: 31–48.

Authors' contributions
XX contributed to the development of this framework, collection of the data, analysis of the data and preparation of the manuscript. HN contributed to the preparation of the manuscript. Both authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

References

Atai, Mahmood Reza, and Simin Falah. 2006. A contrastive genre analysis of result and discussion sections of applied linguistics research articles written by native and non-native English speakers with respect to evaluated entities and ascribed values. In *Proceedings of the 10th Conference of Pan-Pacific Association of Applied Linguistics*, 41–55. Tokyo: PAAL. Edinburgh University, Edinburgh, UK.

Bakhtin, M.M. 1981. *The Dialogic Imagination*. Austin: University of Texas Press.

Biber, Douglas, Johansson Stig, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. New York: Pearson Education.

Cava, Amelia Maria. 2007. *'Good vs. Bad' in Research Articles Abstracts. A corpus-based analysis. Università degli Studi di Napoli Federico II*, PhD thesis.

Chafe, Wallace L., and Nichols Johanna. 1986. *Evidentiality: The linguistic encoding of epistemology*. Norwood: Ablex Publishing Corporation.

Christie, Frances. 1997. Curriculum macrogenres as forms of initiation into a culture. In *Genres and Institutions: Social processes in the workplace and school*, ed. Frances Christie and James Robert Martin, 134–160. London: Cassell.

Cotos, Elena. 2014. *Genre-based automated writing evaluation for L2 research writing: From design to evaluation and enhancement*. Basingstoke, UK: Palgrave Macmillan.

Dressen, Dacia. 2003. Geologists' implicit persuasive strategies and the construction of evaluative evidence. *Journal of English for Academic Purposes* 2: 273–290.

Halliday, M.A.K. 1994. *Introduction to Functional Grammar*. London: Edward Arnold.

Hood, Susan. 2004. *Appraising Research: Taking a stance in academic writing*. Sydney: University of Technology PhD thesis.

Hunston, Susan. 1993. Evaluation and Ideology in Scientific Writing. In *Register Analysis: Theory and practice*, ed. Mohsen Ghadessy, 57–73. London: Pinter.

Hunston, Susan. 1994. Evaluation and organization in a sample of written academic discourse. In *Advances in Written Text Analysis*, ed. Malcolm Coulthard, 191–218. London: Routledge.

Hunston, Susan. 2000. Evaluation and the planes of discourse: Status and value in persuasive texts. In *Evaluation in Text*, ed. Susan Hunston and Geoff Thomson, 176–206. Oxford and New York: Oxford University Press.

Hunston, Susan, and Geoff Thompson. 2000. *Evaluation in text*. Oxford: Oxford University Press.

Hyland, Ken. 1996. Writing without conviction? Hedging in science research articles. *Applied Linguistics* 17: 433–454.

Hyland, Ken. 1999. Disciplinary discourses: Writer stance in research articles. In *Writing: Texts, processes and practices*, ed. C.N. Candlin and Ken Hyland, 99–121. London: Longman.

Martin, J.R., and Peter R.R. White. 2005. *The Language of Evaluation: Appraisal in English*. New York: Palgrave Macmillan.

Millán, Enrique Lafuente. 2012. A contrastive study of generic integrity in the use of attitudinal evaluation in research articles written for different audiences. *Brno Studies in English* 38: 79–96.

Myers, Greg. 1989. The pragmatics of politeness in scientific articles. *Applied Linguistics* 10: 1–35.

Ochs, Elinor, and Bambi Schieffelin. 1989. Language has a heart. *Text* 9: 7–25.

Samraj, Betty. 2002. Introductions in research articles: Variations across disciplines. *English for Specific Purposes* 21: 1–17.

Shaw, Philip. 2003. Evaluation and promotion across languages. *Journal of English for Academic Purposes* 2: 343–357.

Sinclair, John. 1981. Planes of discourse. In *The two-fold voice: essays in honor of Ramesh Mohan*, ed. S.N.A. Rizvi, 70–89. Salzburg: University of Salzburg.

Swales, John Malcolm. 1990. *Genre Analysis*. Cambridge: Cambridge University Press.

Swales, John Malcolm. 2004. *Research genres: explorations and applications*. Cambridge: Cambridge University Press.

Swales, John Malcolm, and Christine Feak. 2012. *Academic Writing for Graduate Students, 3rd Edition*. Anne Arbor: University of Michigan Press.

Thetela, Puleng. 1997. Evaluated entities and parameters of value in academic research articles. *English for Specific Purposes* 16: 101–118.

Tucker, Marc. 2003. Out with the old. *Education Next* 3: 20–24.

UK Economic and Social Research Council. http://www.esrc.ac.uk/research/evaluation-and-impact/what-is-impact/. Accessed 08 July 2016.

White, Peter R. R. 2003. An Introductory Course in Appraisal Analysis. http://www.grammatics.com/appraisal. Accessed 08 July 2016.

Wu, Siew Mei. 2008. Investigating the effectiveness of arguments in undergraduate essays from an evaluation perspective. *Journal of Prospect: An Australian Journal of TESOL* 23: 59–75.