

Move two: *establishing a niche*

Wasima Shehzad

University College Yanbu (Kingdom of Saudi Arabia)

wasima.shehzad@yahoo.com

Abstract

The significant purpose of the author in the Introduction of a research article is to convince the reader about the importance of the work to be presented. To achieve this end, a convincing “niche” needs to be built by evaluating, rejecting or indicating gaps in previous related work. The purpose of “establishing a niche” is to emphasize the current research project presented by the author. The present paper investigates how Computer scientists use this obligatory step of “Create a Research Space” (CARS) model (Swales & Feak, 1994 & 2004) to highlight their own research work. This paper not only compares the results with other similar studies but also presents an in-depth analysis of various types of gap statements used in Computer Science research article Introductions. The issue of cyclicity of this step and the linguistic indicators used for the establishment of “niche” (the gap statements) are both discussed.

Key words: corpus-based study, CARS, niche, Computer Science, gaps.

Resumen

Sub-sección dos: “creando el nicho”

El autor de un artículo de investigación persigue, como finalidad principal en la introducción de dicho artículo, convencer al lector sobre la importancia de su trabajo. Para alcanzar su objetivo, habrá de “crear un nicho”, con lo que evaluará, rechazará o señalará las carencias encontradas en trabajos anteriores y relacionados con el que se presenta. La elaboración o construcción de un “nicho” tiene por objeto resaltar la investigación que en ese mismo momento está exponiendo el autor. En el presente artículo se investiga cómo los especialistas del campo de la informática realizan esta tarea obligatoria ajustándose al modelo de “crear un espacio de investigación” (Modelo CARS) propuesto por Swales & Feak (1994 & 2004) y así hacer notorio su trabajo de

investigación. En este artículo se comparan los resultados con otros estudios análogos y, además, se analizan de forma pormenorizada los distintos tipos de enunciados que denuncian carencias (*gap statements*) y que son característicos de las introducciones de los artículos de investigación en informática. Asimismo se delibera sobre el asunto de la “ciclicidad” presente en este paso del modelo y los indicadores lingüísticos de carencias que con más frecuencia se utilizan para la construcción del “nicho”.

Palabras clave: estudios de corpus, modelo CARS, nicho, informática, carencias.

1. Introduction

The writing of the introduction section of a research article has been studied by Swales (1981, 1983, 1984 & 1990), Cooper (1985), Swales & Najjar (1987), Hughes (1989), Fredrickson & Swales (1994), Lindeberg (1994), Posteguillo (1995 & 1999), Anthony (1999 & 2002), Hyland (1999a, 1999b & 2000), Kuo (1999), Lewin et al. (2001), and Shehzad (2006 & 2007), among others. Out of these, the works of Cooper (1985), Posteguillo (1996 & 1999), Posteguillo et al. (1998), Anthony (1999 & 2002), and Shehzad (2006 & 2007) have focused on the discipline of Computer Science (CS) research articles. However, Cooper (1985) and Anthony (1999) have used a relatively smaller corpus (12 research articles) and the focus of Posteguillo (1995, 1996 & 1999); Posteguillo et al. (1998) and Anthony (2001 & 2002) has been on the overall schematic structure of research articles, titles and abstracts, whereas Shehzad’s (2006) focus is on Move One and Shehzad (2007) concentrates on “indicating the structure” part of the introduction. Furthermore, none of these scholars has investigated in detail Move Two: “Establishing a niche” (indicating a gap in the previous research or extending previous knowledge in some way) of the “Create A Research Space” (CARS) model when writing the introduction of a research article/paper, which is often thought to be obligatory in major research articles (Swales & Feak, 1994 & 2004), in detail.

In the CARS model, Move 2 is the key move as it is the hinge that connects Move 1 to Move 3 –i.e., “what has been done” to “what the present research is about”. It functions as a “mini-critique” and often consists of not more than a sentence. By indicating a knowledge gap, the writer builds up a “demand” for the current contribution. “Essentially the gap represents an unresolved question that the current contribution seeks to solve” (Lindeberg, 2004: 89). The present study aims to bring forward different

ways through which this move is realized in CS research articles for a comprehensive understanding of Computer scientists' rhetorical style rather.

2. Literature review

The research article is essentially a “product” in its finished form which is the outcome of a complex process (Swales, 2004), introduction holding a fundamental position in this process.

An introduction of a research article is a “crafted rhetorical artefact” and a “manifestation of rhetorical maneuver” (Swales, 1990: 155). This maneuvering also involves building up a convincing “niche” (Swales, 1990: 142) through “elaborated criticisms or denials of previous knowledge claims” (Lindeberg, 1994: 138) –i.e. statements of gaps as indicated through Move 2 of CARS model (Swales & Feak, 1994 & 2004). Lindeberg (2004: 40) uses the term “foil” (a thin layer of bright metal placed under a displayed gem or piece of jewelry to lend it brilliance) for gap statements which “emphatically foreground and enhance the current contribution by contrast”. Foil is also “A very thin layer of polished metal placed under a gem, especially an inferior or imitation gem, to give it more color or sparkle” (*The World Book Dictionary*, 2001: 828) thus emphasizing its quality of setting off contrast through which a “demand” for the current contribution is built up.

“Niche” is also a marketing term indicating the process of finding small but potentially profitable market segments and designing custom-made products to cater for these specialized markets. However, the “niche” should be large enough to be profitable but ignored by the major industry players. Swales (1990) uses this term in its ecological sense describing the relational position of a species or population in an ecosystem. The “niche” (Swales, 1990: 142) thus established includes how a population responds to the abundance of its resources and enemies. Relating, both, the ecological and marketing connotations to writing the introduction of a research paper, the “niche” provides a firm “background against which the magnitude or relevance of the current contribution stands out clearly” (Lindeberg, 2004: 89).

Another objective of displaying writers' knowledge and understanding of the state of the art of their discipline is also achieved through gap statements (Bazerman, 1988). Unfortunately, the available literature about CS is limited to the analysis of overall structure of the CS research article (Posteguillo, 1995 & 1999) and the generic evaluation of the introduction of Software

Engineering article (Anthony, 1999). Hence, in the present paper I deal with the pertinent question of how Computer scientists make use of gap statements to build up a convincing argument in the “tightly contested” genre of research article.

3. Methodology

In the previous works of this kind either researchers have limited their studies to a close analysis of a small number of texts such as Tarone et al.’s (1981 & 1998) detailed description of two astrophysics articles, or a general analysis of a larger corpus (Dudley-Evans’s (1984) corpus of 156 articles, or even a comparative analysis of the Hyland’s (2000) large corpus of 240 research articles from soft and hard sciences). These approaches have their limitations (see Anthony, 2002). Therefore a balanced approach was adopted for the study in focus. A corpus called, Shehzad Computer Science Corpus (SCS Corpus) (Shehzad, 2005) of 56 introductions of CS research articles from five different journals published by the Institute of Electrical and Electronics Engineers (IEEE) Computer Society was created (see Table 1 for detail). The initial intention though was to have 60 texts but the December issues of four journals had not been published by the time the corpus was created. The data was analysed with the help of computer-based tools such as WordSmith (Scott, 1997).

The remainder of the paper first presents a general description of Move 2 in CS research articles introductions followed by taxonomy of the different types of Move 2 identified in the SCS Corpus. The next section deals with the linguistic indicators of gaps followed by a discussion about the cyclical nature of this move.

4. Results and discussion

The first step of the rhetorical analysis was to identify the realization of Moves which showed 94.64% occurrence of Move 2, thus Swales & Feak’s (2004) assertion of its being obligatory. This percentage is higher than Anthony’s (1999) report (91.7%) which was much higher than Posteguillo’s (1995) results (57%) thus implying a growing trend in its usage. The reason could be the fact that because of increased research being conducted in this discipline, Computer scientists feel a competitive pressure more than ever and need to create a research space for their work.

More than the overall number of occurrences, the patterns of Move 2 were found to be interesting as they were not only complex but also complicated. Before going into the details, a Journal-wise summary of Move 2 is given in Table 1.

Name of journal	Abbreviation for journal	Move 2 occurrence	Percentage
Transactions on Computers	ToC	10	90.90%
Transactions on Pattern and Machine Analysis	PAMI	9	81.81%
Transactions on Software Engineering	SE	11	100%
Transactions on Parallel and Distributed Systems	PADS	11	100%
Transactions on Knowledge and Data Engineering	KDE	12	100%

Table 1. Journal-wise occurrences of Move 2 in the SCS Corpus.

In comparison to a hundred percent presence of Move 2 in three journals 81.81% occurrence in PAMI looks slightly lower, but still it is high enough to generalize that Move 2 is an important part of the research papers in this field and Computer scientists use it quite frequently. In one extreme case, the introduction opened with Move 2.

Example (1)

SOME recent investigations [1], [2] revealed that some of the major problems today in developing software systems are not technical but managerial. (INTR 25)

4.1 Taxonomy of Move 2

The writers show their knowledge of “the state of the art” of their field by stating a problem, a set of conflicting evidence or a question. On the basis of extensive elaborations of such gap statements made in the introductions of CS, a taxonomy of Move 2 types has been prepared, defined and discussed as follows.

(a) *Simple/Short*

About 20% of the occurrences could be called “simple gap statements” because of their easy-to-follow and straightforward manner. As far as brevity is concerned, “establishing a niche” is supposed to be a brief transitory period like the one between touching the water and starting kicking while swimming. Swales & Feak (2004) thus recommend it to be fairly short. A few instances of brief Move 2s were found in the SCS Corpus.

Type	Meaning	Occurrence	Percentage
Short/Simple	brief, easy to identify	15	26%
Lengthy/Numbered	extended to many sentences and paragraphs	14	25%
Embedded	intertwined with Move One, step b. i.e., "reviewing items of previous research"	13	23%
Multiple Act	based on "acts" each realizing a sub-purpose	6	10%
Reported	already identified by the researchers/scholars of the target community	3	5%
Research Questions/Rationale	purpose/research question/rationale statement of the present research serving as the gap statement	3	5%
Extension	continuation/augmentation of the previous research	2	3%
Contrastive	presentation of a comparable idea/algorithm/technique/method to the existing works	2	3%

Table 2. Types of gap statements in CS RAs introductions.

Example (2)

Previous attempts to compare items from different ontologies have been based on an integrated ontology that was derived manually or semi automatically. This work aims at ... (INTR 51)

(b) *Lengthy/Numbered*

Contrary to Swales & Feak's (2004) belief, that Move 2s are quite short, often consisting of no more than a sentence, most of the Move 2s in CS introductions are fairly lengthy. The few one sentence examples look more like exceptions. An example of the Lengthy gap is given here.

Example (3)

These data sets are far too large to fit in main memory and are typically stored in secondary storage devices. Linear scans are the only cost-effective access method; random access is prohibitively expensive. Some data sets, such as router packet statistics, meteorological data, and sensor network data, are transient and need not be realized on disk; the data must be processed as they are produced and discarded in favor of summaries whenever possible. As the size of such data sets far exceeds the amount of space (main memory) available to an algorithm, it is not possible for a data stream algorithm to "remember" too much of the data scanned in the past. This scarcity of space necessitates the design of a novel kind of algorithm that Each scan of a large set on a slow device is expensive, and so the criteria In the case of "transient" streams, i.e., ... only one scan is possible. Furthermore, unlike

an online algorithm, a data stream algorithm may not be required to take an irrevocable action after the arrival of each point; it may be allowed to take action after a Clustering, a useful and ubiquitous tool in data analysis, is, in broad strokes, the problem of finding a partition of a data set so that, under some definition of “similarity,” similar items are in the same part of the partition and different items are in different parts. ... (INTR 54)

Move 2 took about 45% of the space of this introduction. Another interesting type of Move 2 is that one in which the authors use the discourse markers to logically explain the reasons why there is a need for the research space they intend to occupy.

Example (4)

Several characteristics of computers that execute such applications make them unsuitable for storing sensitive information. First, the devices may be poor and Second, can be easily stolen or Third, when data size become large, storage management is expensive and prone to errors. (INTR 45).

Even in one case, the gap statement does not remain just a brief statement as suggested by Swales & Feak (1994 & 2004). It rather becomes a report covering over two pages and then mingling with Move 1, b. towards the end. Unlike casual reporting it begins in an organized manner and has been divided into four points that are numbered.

Example (5)

However, most of the existing database models are designed under the assumptions that the data/information stored is precise and queries are crisp. In fact many of these assumptions are not valid for many of the next generation information systems since they may involve complex information with uncertainty. In general, data/information in databases may be uncertain for the following reasons:

- A decision in
- Integrating data from ...
- Information in some traditional ...
- In natural languages, ... (INTR 58)

(c) *Embedded*

Twenty two percent of Move 2s in CS research articles introductions were embedded in Move 1, b. “reviewing items of previous research”. In most instances it is difficult to demark Move 2 from Move 1, b because of their

overlapping natures. Consider the example (6) in which a gap has been supported by the relevant literature:

Example (6)

DNS-dispatcher-based systems can easily scale from locally to geographically distributed Web-server systems ... [23]. However, dispatching requests through the DNS has three problems that prevent ... These issues have been addressed through ...[8], or multiple tiers of ... [15]. This last approach has three main drawbacks. To avoid a system bottleneck at the A-DNS, the traffic ... requires ... Moreover, if any client request... [21]. Finally, the TTL period ... is considered too low. (INTR 40)

(d) *Multiple Act*

“Communicative purposes are realized in moves, which in turn, are composed of one or more act (each realizing a sub purpose)” (Lewin et al. 2001: 44). Up to four to five acts in the realization of gap statements in the SCS Corpus have been noted. This shows that the purpose of “establishing a niche” may also have sub purposes depending upon the kind of research one is involved with.

Example (7)

In high resolution palmprint images ... many features that are similar to singular points and minutiae points in a fingerprint can be obtained; however, these features cannot be observed in low-resolution images ... Nevertheless, we can extract principal lines and wrinkles from In fact, line features play an important role in palmprint identification [3], [4]. Unfortunately, it is difficult to obtain a high recognition rate using only principal lines because of ... (INTR 21)

Example (8)

Despite the advantages of a NOW, the potential computation power of these parallel architectures has not been exploited in depth. One reason is the difficulty in developing ...Traditionally a programmer must However, this approach is time consuming and impractical for ..., particularly since... (INTR 38).

(e) *Reported*

In some cases, gaps are not identified by the authors, they are rather reported as having been generally known or found by other researchers. So, instead of

pointing out the shortcomings, the writer “merely reports them second-hand, relying on previous authorities, presumably a less face-threatening act” (Lindeberg, 2004: 92). This is done in their effort of not to “tread on the toes of other citizens of the academic universe” (Lewin et al., 2001: 43). Computer scientists also use this device sometimes.

Example (9)

Reference [10] shows that none of the existing architectures meets the universal access requirement. (INTR 29)

In such a statement, the author does not create a gap himself; he rather uses the weaknesses and shortcomings found by the previous research, to serve as a background to present his own work.

Example (10)

Previous research [21], [23] suggests that cross-site communication and coordination issues cause a substantial loss of development speed. (Move 2)
In this paper, we investigate relationships among delay, communication, coordination and ... (INTR 30)

Considering it a promotional strategy, Lindeberg (2004: 134) further states that “[t]he promotional effect of referring to outside sources as having pointed out the gap is that the present writers’ claim of a gap is supported by other researchers working in the same problem area and this presumably makes the criticism less confrontational”.

(f) *Research Question(s) / Rationale*

In this type of Move 2, a gap is established by posing research question(s) based on the weaknesses of the previous research.

Example (11)

Qualitative studies [21], [18] have shown how individuals are disrupted by cross-site coordination challenges. But question remains about the cumulative effects, for example, how distance affects the speed with which software engineering tasks are accomplished, and how distance is related to other important variables that influence speed, such as the size of a task, or the number of people involved. In addition to being important research questions, these are critical pragmatic issues as ... (INTR 30)

In the following example, we notice the rationale of the need of “interagent communication and synchronization in multiagent applications” which serves as a step forward to work on a new research area rather than bridging an already existing gap.

Example (12)

In this paper, we address the issue of interagent communication and synchronization.... (Move 3 a.) Interagent communication and synchronization in multiagent applications is needed due to two reasons. First, the application may be inherently distributed in nature.... Second, the performance of the application can be significantly improved by employing interagent communication and synchronization techniques.... (Move 2) We propose (INTR 39)

(g) *Extension*

The gap created may not always be for a new research work, it could be the extension of the author’s previous work.

Example (13)

... this paper extends the logic language presented in previous work of ours [5], [6], [7] with a second order logic syntax ... (INTR 57)

(h) *Contrastive*

Here Move 2 is realized in the form of a contrast with other work(s). This is mostly done when an improved version of the already existing design or technique is presented. In one introduction, the second occurrence of Move 2 was seen as a contrastive statement to the preceding announcement of Move 3, step a. Lexical indicators such as “unlike” are used to make this contrast.

Example (14)

Unlike existing algorithms, when making decisions about reducing or ordering a test suite, our algorithms consider the complexities of MC/DC. (INTR 27)

It may be noted here that whereas this taxonomy of gap statements offers a myriad of choices to Computer scientists for the realization of Move 2, it may make this genre tradition complex and lead to comprehension difficulties for the audience.

4.2 Linguistic indicators of gaps

The following linguistic indicators used for the establishment of a gap in the previous research (Swales & Feak 1994 & 2004, and personal discussions with Prof. Swales) were studied in the present Corpus. For the purpose of discussion these have been classified into three categories.

- (a) *Contrastive statements*: however; while; but; although; nevertheless; as opposed to; rather than; with a few exceptions.
- (b) *Quantifiers and quasi-negatives*: limited; few; little.
- (c) *Negatives*: none of; not been; no [work/research/data/study].

(a) *Contrastive Statements*

Sometimes to avoid negative or quasi-negative comments, the use of a contrastive statement such as “however” “while” “rather than” “as opposed to” “although” “considerable”, “with few exceptions”, “tended to”, etc., can be useful. These are used to soften the negative impact of gap statements. Thereby, concordance search was conducted and studied to determine their role in the establishment of gap statements and has been listed in the descending order of gap statements occurred in the SCS Corpus.

Table 3 illustrates significantly high usage of “however” for the enactment of gaps as compared to the usage of “while” and “but” for such executions. The percentage of the occurrence of “nevertheless” is the same as that of “while”; however, having one gap statement in the Corpus, it could easily be neglected.

Conjunctions	Hits	Gap Statements	Percentages
However	58	36	62%
While	41	14	34%
But	46	11	23.91%
Although	21	11	52.38%
Nevertheless	3	1	33.33%
As opposed to	4	2	50%
Rather than	14	1	7.14%
With a few exceptions	1	1	100%
TOTAL	188	77	40.95%

Table 3. Contrastive statements used for Move 2.

(a.1) However

“However” with 36 occurrences seems to be the obvious indicator to establish a gap, but this word alone does not guarantee a clear indication of

the gap. It is the phrase following “however” that contains and conveys the meaning, for example:

Example (15)

However, software visualizations are often too simplistic and lack visual clues for the viewer to correctly interpret them. (INTR 33)

Example (16)

However, none of this work contains a complete and a systematic investigation of different probabilistic modeling techniques for query approximation. (INTR 59)

Example (17)

However, dispatching requests through the DNS has three problems that prevent load balancing among the Web nodes: (INTR 40)

Interestingly, four instances of “however” represented Step a. “stating the nature of research” and two were used in the context of “Principal findings”.

Example (18)

We will show below, however, that a straightforward remedy is to base the similarity measure on a suitable path metric. (INTR 23) (stating the nature of research)

Example (19)

However, our algorithm took more time to find its better answers. (INTR 54) (Principal finding)

In one exceptional case it was used to state the “limitation” of the work undertaken.

Example (20)

... of packets dropped for networks with soft linklevel flow control that handle congestion by dropping packets, however, this issue is not explored in this work. (INTR 44)

(a.2) *While*

The next common linguistic item that played a part in the establishment of the knowledge gap was, “while”. Out of the 41 entries, 14 were used to convey gaps.

Example (21)

While most existing works on MANET take flooding as a straightforward and direct solution, we show in [15] that a blind flooding may result in excessive redundancy, contention, and collision. (INTR 5)

Apart from the gap statements, an interesting juxtaposition was seen in these entries. Two each were used to write the “limitations” and “value” of the current work.

Example (22)

While an in-depth discussion of all possible applications is not possible, we next briefly discuss work that relates to the applications we use to illustrate our optimization approach. (INTR 23)

Example (23)

These studies show that our testsuite reduction techniques can be effective in reducing test suites while providing acceptable performance. (INTR 27)
(Value)

(a.3) *But*

“But” usually signals two types of contrast: (i) “denial-of-expectation” (what is expected after a reading of the first conjunct turns out not to be true from a reading of the second); and (ii) the “marker of semantic contrast” (Celce-Murcia & Larsen-Freeman, 1999). Some examples from the eleven instances of “but” used for making gap statements are given below:

Example (24)

Qualitative studies (e.g., [21], [18]) have shown how individuals are disrupted by cross-site coordination challenges. But question remain about the cumulative effects, for example, how distance affects the speed ... and how ... (INTR 30)

Limited use of “but” to augment the “value” of their own work was also noticed.

Example (25)

After a thorough and detailed study ... we have discovered centralized and distributed schemes that not only achieve good results, but also guarantee stable performance. (INTR 40)

(a.4) *Although*

With eleven examples, “although” seems to be a popular choice after “however” and “while” for the expression of knowledge gap when the Computer scientists want to hedge the weaknesses identified in the already used algorithms, methods developed or design mechanisms, etc.

Example (26)

Although a large part of the behavior of grouping algorithms is determined by the behavior of the grouping measure, in this paper, we will focus on the considerations for the grouping framework. (INTR 16)

The use of “although” has not been restricted to contrastive statements. An example each of Move One step a. and step b. and Move Three step c. and step d. were also found.

Example (27)

Although mobile ad hoc networks have been mainly used in military applications, they are being increasingly used for civilian applications such as virtual class rooms, wireless local area networks, and law enforcement. (INTR 37) (Move 1 step a)

Example (28)

The definition of ... was first characterized ... by Henzinger et al. [38], although the works of Munro and Patterson [64] and of Flajolet and Martin [23] predate this definition. (INTR 54) (Move 1 step b)

Example (29)

... we have found out that, although in some cases a distributed redirection algorithm may also achieve slightly worse performance than some centralized alternatives it has three major advantages that make its use preferable ... (INTR 40) (Move 3 step c)

Example (30)

... our heuristic approach performs better in terms of all the above parameters, although the percentage improvement is relatively low. (INTR 1)
(Move 3 step d)

(a.5) *Nevertheless*

Out of the three, one clear-cut case of the use of “nevertheless” in gap statements was:

Example (31)

Nevertheless, limited work has been reported on palmprint identification and verification, despite the importance of palmprint features. (INTR 21)

(a.6) *As opposed to*

This contrastive statement included two examples of gap and one of “Value”.

Example (32)

E-textiles impose specific challenges as opposed to other applications in the general area of networked systems. (INTR 8)

(a.7) *Rather than*

Surprisingly, the entries of “rather than” were over shadowed with Move 3, having four examples of Step a., one of Step b. and two of Step d. The only clear example of gap statement is given here.

Example (33)

On the negative side, requiring explicit identification of states requires much more consistency from stakeholders when constructing scenarios and forces them to reason about their system in terms of state “rather than” sequences of actions. (INTR 26)

(a.8) *With a few exceptions*

Writers often hedge their full commitment by referring to deficiencies in the research model, theory or method which may compromise the accuracy of their results, thus allowing “the prudent researcher to present results having anticipated criticism of the premises or methods by which they were

achieved” (Hyland, 1998: 143). So the writers use phrases like “with a few exceptions”, rather than “nothing is known about” or “we have no knowledge about”. As Hyland (1988: 145) points out “Again, this is a way of limiting the damage that might result from a bald proposition; the hedge functioning to contextualize the claims made by suggesting parameters within which they might be accepted as true”. An example from the Corpus is given here.

Example (34)

In fact, with a few exceptions, the creators of Web pages create links to other pages, usually with an idea in mind that the linked pages are (INTR 55)

(b) *Quantifiers and quasi-negatives*

The “most common way to indicate a gap is to use a ‘quasi-negative’ subject”. (Swales & Feak, 2004: 258). Quantifiers, such as the following, could be helpful in pointing out the shortcomings of the previous research and thus preparing ground to accommodate the current work. “Few” and “little”, although not morphologically negative, are negative in meaning and in syntactic behavior. The concordance results were studied for gap statements and are depicted in Table 4.

Quantifiers	Hits	Gap Statements	Percentages
Limited	28	17	60.71%
Few	16	4	25%
Little	9	6	66.66%
TOTAL	44	27	61.36%

Table 4. Gap statements and quantifiers/quasi-negatives.

“Little” with a 66.66% occurrence seems to be at the top, but “limited” having lesser percentage comparatively (with 17 occurrences) appears to be a more popular choice for the exercise of gap statements.

(b.1) *Limited*

One way of establishing gap is to show that other investigators’ work has been limited in certain ways or to certain areas and thus intended to be encompassed by the present research. The word “limited” registered 28 hits. Despite having some open cases such as: “How do we search for a queried palmprint in a given database and obtain a response within a limited time?”,

17 clear examples were found that exhibited the limitations of the earlier researchers. These included limitations in the designs, methods adopted, efficiency and accuracy, and limited work/research etc.

Example (35)

The disadvantages of the repeated application of the BOM test... are test time inefficiency and “limited” fault coverage for This paper presents an improved... (INTR 10)

(b.2) *Few*

From the concordance list having 16 entries for “few”, only four could be categorized as gaps.

Example (36)

There are only a “few” studies to incorporate methods for handling impreciseness in such an integrated environment. (INTR 58)

(b.3) *Little*

The next quantifier looked for was “little” as according to Swales & Feak (2004), the chances of using a full negative like “no studies” are that somebody can find an exception to such a strong statement. Six instances out of the nine entries for “little” had interesting nouns (work, research, information, detection effect, empirical information/evidence) accompanied with them to report insufficient job done by the other discourse community dwellers.

Example (37)

... there is so far very “little” empirical evidence that communication of defect information among inspectors yields significant gains (INTR 29)

(c) *Negatives*

The third category of the linguistic indicators studied in the Corpus was negatives. “The chief use of negatives is directed at a proposition already in the discourse” and are used more to respond than to initiate, Celce-Murcia & Larsen-Freeman (1999: 196) explain. They further elaborate that one might assume that negative would receive prominent stress whereas it is reduced to mitigate the disagreement, presumably in the interest of maintaining social harmony. Negatives are generally known to be used for pinpointing

something which is either faulty or has not been paid sufficient attention to. So, one would assume the use of negatives as a natural choice to pave way for one's own work. Contrary to this speculation, the tendency of the usage of negatives is lower here, when compared to the usage of conjunctions and quantifiers/quasi-negatives for gap statements (see Table 5).

Negatives	Hits	Gap statements	Percentages
None of	6	5	83%
Not been	3	3	100%
No work/research...	1	-	0%
TOTAL	10	8	80%

Table 5. Gap statements and negatives.

(c.1) *None of*

In the six entries of *none of*, one looked like an expression of the “value” of the present work. In the rest of the examples, efforts to establish gaps can easily be noticed.

Example (38)

... “none of” the above-mentioned approaches can guarantee to find the global minimum ... (INTR 23)

Referents of “None of” used for gaps were: “the above-mentioned approaches”, “this work”, “the published papers”, “the approaches above”, “group communication services”, or “existing architectures”.

(c.2) *Not been*

The three entries of “not been” were clear cut cases of gap statements: “has not been exploited in depth”, “has not been pursued in the literature until recently”, “has not been studied in the past”. This negative was accompanied with verbs such as “exploited”, “pursued”, “studied”, etc., used as present perfect tense in the passive voice.

(c.3) *No work/research/data/study*

Contrary to the previous two negatives, the only instance of “no work” could be taken as a “value” statement.

Example (39)

To the best of our knowledge, no work has been published on the ... (INTR 4)

4.3 Overall indicators of gaps

The overall situation of the linguistic indicators found in Move 2 (i.e., gap statements used by Computer scientists) is depicted in Table 6.

Linguistic indicator	Hits	Occurrence as gap	Percentage	The most dominant
Conjunctions	148	62	41.89%	However
Quantifier	44	27	61.36%	Limited
Contrastive statements	48	15	31.25%	Although
Negatives	10	8	80%	None of

Table 6. Linguistic indicators of Move 2.

Out of the total hits for negatives, 80% serve the purpose of establishing gaps in the introductions, which is followed by 61.36% presence of Quantifiers. On the other hand, the number of occurrences corresponding to conjunctions in gap statements is the highest but lower than the percentages for negatives and quantifiers. This is because of the usage of Conjunctions in other moves too as we saw earlier on.

(a) *Negative Verbs and Adjectives*

Swales & Feak (2004) suggest twelve verbs and seven adjectives for students to tell if they are strongly negative, neutral or slightly negative. The concordance search showed that Computer scientists do not seem to be interested in this list except for the rare use of “concentrated on”.

Example (40)

hence, people have concentrated on designing on heuristic algorithms for designing ... for ... and analyzing their performance. (INTR 37)

According to Swales & Feak (2004), five verbs (i.e., “misinterpreted”, “failed to consider”, “disregarded”, “ignored” and “overlooked”) are strong negatives and thus should be avoided. Out of these five, the use of one verb (such as “misinterpreted”) has been found only once in the CS research articles introductions. But interestingly, neither of the other verbs considered less negative were used by the Computer scientists. Similarly, three strongly negative adjectives (“misguided”, “unconvincing” and “unsatisfactory”) were not found in the Corpus, neither the rest of the adjectives except one rare occurrence of “incomplete” as shown in example (41).

Example (41)

Parnas [4] assessed that most legacy systems suffer from several typical problems, including ... and outdated, incomplete, or missing documentation. (INTR 33)

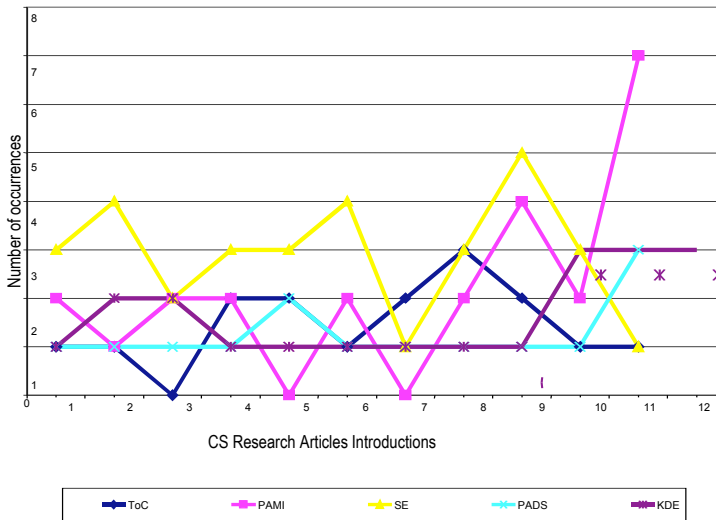
Hence, a word of caution in this regard for the teachers of CS rhetoric while using general guide books for academic writing widely available in the market.

(b) *Cyclicity*

A fairly typical feature of gap statements in longer introductions is that they tend to occur at several different places, “being reintroduced as another aspect of the current research question is elaborated” (Lindeberg 2004: 91). Swales (1990: 158) acknowledges that “Niche-establishment does not necessarily occur only at the end of a literature review, but may follow reviews of individual items, so that cycles of Move 1, step 3 and Move 2 recur” (see also Shehzad (2006) on cyclicity for “establishing a research territory move”).

In Posteguillo’s (1995: 67) study, 75% introductions showed a cyclical pattern for Move 2 as he informed that “[i]t is normal to find this move repeated a series of instances throughout the same introduction, usually alternating with steps in move 1”. Although the occurrence of Move 2s in the present corpus ranged from one to seven, on an average it remained 1.84. This gives us the percentage of 51.78% which is much lower than Posteguillo’s figure. The lower percentage could be attributed to the fact that many Move 2s that were embedded in Move 1, b. were excluded in this calculation. The comparative cyclicity pattern of Move 2 (journal-wise gap statements) is presented here.

Although a straight line of one time occurrence of Move 2s in many introductions is obvious from this graph, yet, the graph clearly reflects variations in the number of recurrences of Move 2 in different journals and different introductions. Seven recurrences in PAMI is the highest number but most Move 2 recurrences fall between one and three giving an average of 1.84 per introduction.



Graph.1. Comparative recurrences (journal-wise) of Move 2.

Example (42)

- Move 2: An inadequate optimization criterion will not solve the application problem, no matter how easy it is to compute the optimum. Conversely, sophisticated criteria that ...are useless in practice as well. For this reason, optimization approaches are attractive ...
- Move 3.a: In this paper, we introduce a novel optimization technique..
- Move 2: Numerous problems in computer vision, including partitioning and grouping, lead to combinatorial optimization problems of this type.
- Move 3.a: In contrast to related work, no specific assumptions are made with respect to As a consequence, our approach covers...
- Move 2: The absence of any specific assumptions about ... as well as ... listed above motivated this investigation.
- Move 1, b: ... approaches to efficiently compute good minimizers have a long history in literature...
- Move 2: Apart from ..., none of the above-mentioned approaches can guarantee to find ... and, in general, this goal is elusive due to the combinatorial complexity of these minimization problems.

- Move 3.b: Consequently, the important question concerning the ...
- Move 2: To the best of our knowledge, none of the approaches above (apart from simulated annealing) seems to be immune against ... and, hence, does not meet these criteria.
- Move 3.a: Our approach belongs to ...
- Move 2: Second, under certain conditions, bounds can be derived with respect to the quality of the suboptimal solution. At present, these bounds are not tight with respect to the much better performance measured in practice.
- Move 2: ...these authors were able to show that ... cannot be worse than ... to the unknown global optimum. Besides ... this fact has motivated our work. (INTR 23)

Another determining factor in the cyclicity of Move 2 is its physical placement in the introduction. The total number of words of each introduction was divided into four parts. Table 7 shows these linguistic indicators divided into four quarters of the introduction, in terms of their physical placement which reflects that most of the Move 2s occurred in the first three quarters of the introductions.

Indicator	Occurrence	1st quarter	2nd quarter	3rd quarter	4th quarter
However	36	9	9	17	1
Limited	17	6	5	6	-
While	14	3	8	3	-
Although	11	2	2	7	-
But	11	4	2	4	1
Little	6	2	1	3	-
None of	5	-	2	3	-
Few	4	-	2	2	-
Not been	3	1	1	-	1
As opposed to	2	-	-	2	-
Rather than	1	-	1	-	-
With a few exceptions	1	1	-	-	-
Nevertheless	1	1	-	-	-
No work	-	-	-	-	-
Total	112	29	33	47	3
Percentage		25.89%	29.46%	41.96%	2.69%

Table 7. Physical placement of Move 2 indicators.

Although the highest percentage is found in the third quarter of the introductions, a fairly large percentage of gap indicators is present in the first and second quarters also –hence, strengthening the proposition of the cyclicity of Move 2 in CS research articles introductions. Since this is a preliminary study, rather than resolving the issue of cyclicity, it raises questions for further research such as: why does this cyclicity occur? Why are

journals such as PAMI more prone towards this characteristic? Does the length of the article and conceptual profile of the paper affect the pattern of Move 2?

5. Conclusion

Gap statements function as a contrasting background to strengthen the significance of the present contribution. They serve as a stage with a brief announcement of the new drama to be presented, thus raising the expectations mingled with the expected applause of the audience. An important decision to be made is how much prominence can be or should be afforded for this step as we saw that gap statements in CS research articles are not instances of brief announcements only, neither are these fixed in their physical or rhetorical nature. They can be found at different places in the texts. These can vary from simple to complex, from creating a rationale (that serves as a gap) to the extension of the previous work, and from contrastive statements to reported gap statements. Conjunctions, quantifiers, negatives and contrasts are typically used as the signal of the onset of this move.

With rapid developments in CS research, Computer scientists have to find a strong “niche” to increase the chances of their audience’s acceptability and the target community’s recognition of their research work which is essential for scientific progress. The editors and reviewers’ increasing pressures for reliable, quality controlled scholarly scientific information with their high quality standards for assessing research for publication in academic journals may be the compelling factor for the authors to use innovative strategies to create a “gap”, a “niche” so that they can provide the cognitive bridge between the known and unknown, old and new, faulty and improved, obsolete and latest, ineffective and efficient, slow and rapid, delayed and subsequent etc., successfully. However, individual disciplinary practices when writing a research article may vary according to their communicative purposes as the choice of the rhetorical strategies and linguistic indicators used, and the notion of cyclicity in the CS research article introductions reflect. This selection is deliberate or content/audience specific needs to be researched further.

Acknowledgements

I am grateful to Professor John Swales for his kind guidance, to Santiago

Posteguillo for sharing scholarly thoughts, and also to the anonymous reviewers for their invaluable insights.

(Revised paper received July 2007)

References

- Anthony, L. (1999). "Writing research article introductions in software engineering. How accurate is a standard model?" *IEEE transactions on Professional Communication* 42: 38-46.
- Anthony, L. (2001). "Characteristic features of research article titles in computer science". *IEEE Transactions on Professional Communication*, 44: 187-194.
- Anthony, L. (2002). *A Machine Learning System for the Automatic Identification of Text Structure and Application to Research Article Abstracts in Computer Science*. PhD Dissertation. Birmingham: The University of Birmingham.
- Bazerman, C. (1988). *Shaping Written Knowledge: The Genre and Activity of the Experimental Article in Science*. Madison, WI: The University of Wisconsin Press.
- Celce-Murcia, M. & D. Larsen-Freeman (1999). *The Grammar Book*, 2nd ed. Boston: Heinle and Heinle.
- Cooper, C. (1985). *Aspects of Article Introductions in IEEE Publications*. MSc unpublished dissertation. Birmingham: The University of Aston in Birmingham.
- Dudley-Evans, T. (1984). "A preliminary investigation of the writing of dissertation titles" in G. James (ed.), *The ESP Classroom Methodology, Materials, Expectations*. 40-46. Exeter: University of Exeter.
- Fredrickson, K.M. & J. Swales (1994). "Competition and discourse community: introductions from Nysvenska studier" in B.L. Gunnarsson, P. Linell & B. Nordberg (eds.), *Text and Talk in Professional Context*, 9-22. Uppsala: ASLA.
- Hughes, G. (1989). *Article Introductions in Computer Journals*. Unpublished MA dissertation. Birmingham: The University of Birmingham.
- Hyland, K. (1998). "Boosting, hedging and negotiation of academic knowledge". *Text* 18: 349-382.
- Hyland, K. (1999a). *Writing: Texts, Processes and Practices*. London: Longman.
- Hyland, K. (1999b). "Disciplinary discourses: writer stance in research articles" in C. Candlin & K. Hyland (eds.) *Writing: Texts, Processes and Practices*, 99-121. London: Longman.
- Hyland, K. (2000). *Disciplinary Discourses: Social Interactions in Academic Writing*. Harlow, Essex: Longman.
- Kuo, C.H. (1999). "The use of personal pronouns: role relationships in scientific journal articles". *English for Specific Purposes* 18: 121-138.
- Lewin, A., J. Fine & L. Young (2001). *Expository Discourse: A Genre-based Approach to Social Science Research Texts*. New York: Continuum.
- Lindeberg, A.C. (1994). "An exploratory study of knowledge claims in article introductions in three disciplines: Finance, Management, and Marketing" in M. Majapuro & T. Nikko (eds.) *Talous ja kieli II (Business and Language II)*, 647-655. Helsinki: Publications of the Helsinki School of Economics and Business Administration.
- Lindeberg, A.C. (2004). *Promotion and Politeness: Conflicting Scholarly Rhetoric in Three Disciplines*. Vasa: Abo Akademi University Press.
- Posteguillo, S. (1995). *Genre Analysis in English for Computer Science*. Unpublished PhD dissertation. Valencia: Universitat de València.
- Posteguillo, S. (1996). "A genre-based approach to the teaching of reading and writing abstracts in Computer Science" in J. Piqué, J.V. Andreu-Besó & D.J. Viera (eds.), *English in Specific Settings*, 47-58. Valencia: Nau Llibres.
- Posteguillo, S. (1999). "The schematic structure of computer science research articles". *English for Specific Purposes* 18: 139-160.
- Posteguillo, S., I. Fortanet, J.C. Palmer & J.F. Coll (1998). "Linguistic analysis of research article titles: disciplinary variations" in I. Vázquez Orta & I. Gullén Galve (eds.), *Perspectivas pragmáticas en lingüística aplicada*, 443-458. Zaragoza: Universidad de Zaragoza.
- Scott, M. (1997). *WordSmith Manual*. Oxford: Oxford University Press.
- Shehzad, W (2005). *Corpus-based Genre Analysis: Computer Science Research Article In-*

- roductions. PhD Dissertation, Islamabad, National University of Modern Languages, Pakistan.
- Shehzad, W. (2006). "Computer scientists' approach to 'establishing a research territory'". *Selected Papers from the Fifteenth International Symposium on English Teaching*, 127-141. Taipei: Crane Publishing Co and ETA-ROC.
- Shehzad, W. (2007). "How to end an introduction in a Computer Science article" in E. Fitzpatrick (ed.), *Corpus Linguistics Beyond the Word: Corpus Research from Phrase to Discourse*, 243-255. New York: Rodopi.
- Swales, J.M. (1981). *Aspects of Article Introductions*. Aston ESP Research Report No. 1, Language Studies Unit. Birmingham: The University of Aston in Birmingham.
- Swales, J.M. (1983). "Developing materials for writing scholarly introductions" in R.R. Jordan (ed.), *Case Studies in ELT*, 180-200. London: Collins.
- Swales, J.M. (1984). "Research into the structure of introductions to journal articles and its application to the teaching of academic writing" in R. Williams, J. Swales & J. Kirkman (eds.), *Common Ground: Shared Interests in ESP and Communication Studies*, 77-86. Oxford: Pergamon.
- Swales, J.M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge: Cambridge University Press.
- Swales, J.M. (2004). "Then and now: a reconsideration of the first corpus of scientific English". *Ibérica* 8: 1-14.
- Swales, J.M. & C. Feak (1994). *Academic Writing for Graduate Students: A Course for Nonnative Speakers of English*. Ann Arbor: The University of Michigan Press.
- Swales, J.M. & C. Feak (2004). *Academic Writing for Graduate Students. Essential Tasks and Skills*, 2nd ed. Ann Arbor: The University of Michigan Press.
- Swales, J.M. & H. Najjar (1987). "The writing of research article introductions". *Written Communication* 4: 175-192.
- Tarone, E. S. Dwyer, S. Gillette & V. Icke (1981). "The use of the passive in two astrophysics journal papers". *English for Specific Purposes* 1: 123-140.
- Tarone, E. S. Dwyer, S. Gillette & V. Icke (1998). "On the use of the passive and active voice in astrophysics journal papers: with extensions to other languages and other fields". *English for Specific Purposes* 17: 113-132.
- _____ (2001). *The World Book Dictionary*. Chicago: The World Book.

Dr. Wasima Shehzad currently teaches linguistics at the University College Yanbu, KSA and has taught technical communication/writing at National University of Sciences and Technology, Pakistan. Her research interests include corpus-based genre analysis, academic/technical writing and teacher education. Her recent publications include *Teaching Communication Skills* (2006), Bookman Publishers.

