Rhetorical Structures of Research Article Abstracts: a Preliminary Proposal of the Auto-featural Model

Jai-Hyoung Park
(Seoul National University)

Park, Jai-Hyoung, 2008. Rhetorical Structures of Research Article Abstracts: a Preliminary Proposal of the Auto-featural Model. SNU Working Papers in English Linguistics and Language 7, 77-106. Writing an abstract of a research article in English is a must among the current academic communities. Unfortunately, for the non-native speakers of English, structuring their abstract is not an easy task. In a number of textbooks on abstract writing, it has been traditionally recommended to write an abstract as a miniature of the full research paper. Recent empirical findings, however, suggest that more than mere summarizing takes place. This paper proposes a new analytic scheme for various rhetorical structures of abstracts in research articles. The proposed model improves earlier models such as IMRD and CARS (Swales 1990, 2004), and allows a descriptively precise analysis, by which we could explore some possible features rather difficult to be explained in the previous research. The proposed model, however, is a preliminary version. Thus, this model has to be refined through a further empirical research. (Seoul National University)

Keywords: CARS model, IMRD model, rhetorical structure, research article (RA) abstract, genre analysis, move, step

1. Introduction

There are a great number of languages in which many local academic research articles (RAs) are written, but when it comes to writing an abstract in an RA, English would be the only choice other than the native languages. According to Santos (1996), the editorial policy of most Brazilian academic journals regulates that an abstract in English should be contained in all papers. In almost every university in Korea, all academic dissertations should include an English version of their abstract. Ventola (1994: 333) argues that the main reason of this strong trend of
writing an English abstract is to circulate one's findings worldwide. For
learners of English, however, it is not always an easy task to write an
RA abstract in English. Even though there would be a lot of causes of
this matter, one of the problems non-native speakers of English often
face is how to organize an abstract in terms of the overall structure.
The studies which investigate the general rhetorical structure of RA
abstracts seem to have reached two prototypical patterns, namely IMRD
and CARS respectively [Swales 1990, 2004]. Still, each of these two models
cannot fully account for a considerable amount of variations because
RA abstracts reflect various discourse environment of their writers.
Reviewing previous research, in this article I propose a preliminary
scheme to analyze the various rhetorical structures of RA abstracts. The
scope of the present paper is to theoretically prove the viability of the
proposed model. If empirical research using this model follows and
supports what this paper argues, we can suggest a variety of choices
of organizational patterns of RA abstracts, which in turn will assist
non-natives in writing an English abstract.

2. Review of the literature

Unlike RA introductions1), there seems to be some disagreement among
scholars on a unified analytic model to schematize general rhetorical
structures of RA abstracts. Some academics traditionally use a
four-partite model for their analysis, while others tripartite one. Still
others employ both of them or even make up a slightly modified version
of them. In course of understanding what rationale lies in accepting a
model or models to analyze macro-structures of RA abstracts, I found
that these choices might be reduced into two rather obvious needs -
either 1) to faithfully summarize the whole research paper or 2) to
maximally interest the readers at the expense of giving up summarizing
every part of the whole research article.

At first glance, this argument looks similar to Day and Gastel’s (2006).
They state that there are two types of abstracts, an indicative (or descriptive)
and an informative abstract. The indicative abstract "is designed to indicate
the subjects dealt with in a paper, much like a table of contents, making it

1) For rhetorical structures of RA introductions, a three-part model, i.e. a CARS model,
is generally admitted to be successful according to Swales (2004).
easy for potential readers to decide whether to read the paper" (Day and Gastel 2006: 53). The other type, the informative abstract is devised to abbreviate the paper. "It can and should briefly state the problem, the method used to study the problem, and the principal data and conclusions" (Day and Gastel 2006: 53). However, the way these two types of abstracts are manifested seems quite unclear. As Day and Gastel call "a table of contents", the indicative abstract is likely to contain summarized content in every chapter of the full paper in order to give the readers a clear and comprehensive enough map to refer to. If this kind of abstract omits some part(s) such as literature review, or the results of the analysis, the readers may not get help in deciding whether to read or not, because obviously there are missing parts. Day and Gastel, however, do not think an indicative abstract can "serve as a substitute for the full paper" (Day and Gastel 2006: 53).

Following Day and Gastel (2006)², Lores (2004) provides much clearer account although confusion about the terms still remains. She defines the function of indicative abstracts as the following.

The function of indicative abstracts is to help readers understand the general nature and scope of the research article; it indicates the subject and the main findings of the paper, but it does not go into a detailed step-by-step account of the process involved (Lores 2004: 281-282).

In this regard, it is apparent that an indicative abstract does not indicate the whole paper but just some important parts of it. Not surprisingly, Lores describes an informative abstract as a miniature or an encapsulation of the full paper. Now, the question is how informative an informative abstract is - perhaps, fully informative. How about an indicative abstract? Is this not informative at all? I also wonder to what extent an indicative abstract indicates. As I intimated above, the terminologies for the two types of abstracts need to be clarified so as to avoid any misunderstanding.

2.1 The IMRD model: a map for the full paper

²) Originally, Lores cited Day (1988), the third edition of the book, but the content discussed here is exactly the same.
It is conventionally believed that an abstract is designed to provide the
readers an exact and summarized information of the whole article (ANSI
1979, Bhatia 1993). Bhatia (1993) argues that an RA abstract consists of
four parts: (a) introducing purpose, (b) describing methodology, (c)
summarizing results, and (d) presenting conclusions. This four-partite
model is basically the same as a model for the rhetorical structure of
an RA abstract as a whole (Bruce 1988, Swales 1990). In fact, one can
notice chapter headings such as "Introduction, Method, Results, and
Discussion" in a majority of scientific articles even though the
nomenclatures may be slightly different for same functions. Intuitively,
this model is called an IMRD (Introduction-Method-Results-Discussion)
model (Swales 1990). Similarly, Graetz (1985) asserts that another
four-part but not very different model (Problem-Method-Results-
Conclusions) is the most common structure for an abstract and
she classifies this type of abstracts as an informative abstract. Hence, there
seems to be quite solid agreement that an abstract can be a substitute
for the full paper among scholars. Nonetheless, even if we took this
an-abstract-as-a-full-paper approach, we would not be so sure whether
typical abstracts really consist of the four parts, namely IMRD until this
scheme is empirically proved.

There has been a great number of research which investigates rhetorical
structures of abstracts with empirical data. Salanger-Meyer (1990) found
52% of 77 abstracts in a corpus from the medical domain followed the
IMRD structure. It is reported that roughly half of the abstracts in the
Kaplan, Cantor, Hagstrom, Lamik-Stlein and Boyd Zimmerman (1994)'s
corpus contained four such identifiable segments. Lores (2004) claims
the IMRD was the most representative in her corpus of 36 RA abstracts
in linguistics and applied linguistics. Lee (2004) analyzes the differences
of rhetorical structures of RA abstracts between Korean and International
journals. He found 8 out of 20 abstracts from International Journals in
his corpus had all four components of the IMRD model. Given these
results, we can see that the IMRD model indeed gives us fairly powerful
criteria in analyzing rhetorical structures of RA abstracts. Yet, the research
cited above could not account for about half of the data in terms of
the IMRD model alone. Perhaps, the whole picture is not that simple.

The structure of an RA as a whole does not always have all the four

3) 22 out of 36 (61.1%) abstracts in her corpus followed the IMRD pattern.
IMRD sections. Some do not include one or two of them, and some contain a section or sections other than the four components of IMRD. Swales (2004) gives us useful hints to tackle this matter. After reviewing Thompson (1999), Bunton (1998), Ridley (2000), and Paltridge (2002), he admits some variations of the traditional IMRD structure of doctoral dissertations. Rather unexpectedly, the traditional IMRD pattern is no longer the standard choice of many doctoral aspirants and their advisers. Instead, following two patterns are convincingly regarded to take up the majority portions of the recent doctoral dissertations.

<table>
<thead>
<tr>
<th>Traditional IMRD</th>
<th>Complex</th>
<th>Topic-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ridley</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>Paltridge</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>Thompson</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Bunton</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Totals</td>
<td>16</td>
<td>44</td>
</tr>
</tbody>
</table>

Table 10. Structures in dissertations: Numbers by category (N=100) from Swales (2004: 110)

Introduction (definitions, justification, aims)
↓
Literature Review (sometimes included in the Introduction)
↓
(General Methods) (optional)
↓
IMRD
↓
IMRD
↓
...
↓
IMRD
↓
Conclusions

4) For more clear understanding, I slightly modified the table in terms of terminologies and the graphical structure.
Figure 1. Complex pattern (after Thomson, 1999) from Swales (2004: 108)

Introduction

(Literature Review)

(Theoretical Framework)

Method

Topic: Analysis-Discussion

Topic: Analysis-Discussion

Topic: Analysis-Discussion

Conclusions

Figure 2. Topic-based pattern (after Buntor, 1998: 114) from Swales (2004: 109)

The data given here are from doctoral dissertations and therefore we cannot safely conclude these varied IMRD patterns could be applied to RAs as well. Still, this could be a partial evidence that there are at least some possibilities that the traditional IMRD model does not work for some RAs either. If an abstract is indeed a faithful miniature of its full RA, the organizational pattern of the RA necessarily influences that of its abstract. In other words, when a macrostructure of an RA does not show the conventional IMRD pattern, it would be hard to expect the strict IMRD structure in its abstract. In my rough guess, that is why the traditional IMRD could not explain the other half of the data in Salanger-Meyer (1990), Kaplan et al. (1994), Lores (2004) and Lee (2004). Thus, there is a strong motivation to understand the macro-structure of various RAs first so that we could figure out how their abstracts are organized. In particular, we need to examine the relation between rhetorical structures of a research paper and its abstract. However, empirically clarifying the co-relationship between the generic
organization of RA abstracts and their main bodies (i.e. the whole RAs themselves) is out of the main interest of this article and hopefully, the future research can cover this issue.\(^5\) Suffice to say in this sub-section, some parts can be repeated and some parts do not follow the established order in an RA, and therefore the abstracts may show these various patterns in their structure.

Hyland (2004)\(^6\) proposes a little modified version of the traditional IMRD model. In his analysis of 800 academic article abstracts, which is a relatively large number compared with the previous studies, he employs a five-part model instead of the four-part IMRD. In table 2 below, I directly repeat the outlined scheme of his.

<table>
<thead>
<tr>
<th>Move</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>Establishes context of the paper and motivates the research or discussion.</td>
</tr>
<tr>
<td>Purpose</td>
<td>Indicates purpose, thesis, or hypothesis, outlines the intention behind the paper.</td>
</tr>
<tr>
<td>Method</td>
<td>Provides information on design, procedures, assumptions, approach, data, etc.</td>
</tr>
<tr>
<td>Product</td>
<td>States main findings or results, the argument, or what was accomplished.</td>
</tr>
<tr>
<td>Conclusion(^7)</td>
<td>Interprets or extends results beyond scope of paper, draws inferences, points to applications or wider implications.</td>
</tr>
</tbody>
</table>

Table 2. A classification of rhetorical moves in article abstracts from Hyland (2004: 67)

It is not difficult to find that the labels, ‘Product’ and ‘Conclusion’ in the table 2 are basically the same as ‘Result’ and ‘Discussion’ in the IMRD model respectively. The Hyland’s scheme, however, has one additional label, ‘Purpose.’ He states that he differentiated the author’s ‘Purpose’ from

---

5) Instead, later in this article, we will discuss the relationship between an abstract and its introduction.

6) This book was republished. The original version had been published in 2000.

7) Hyland regards ‘Conclusion’ as an optional move, but the reason is not stated explicitly. In fact, ‘Product’ and ‘Conclusion’ are sometimes quite confusing to distinguish clearly. I will discuss this matter later in the article.
the 'Introduction,' wherein the usual goal is delivering a tenable context for the research in general while in the 'Purpose' more specific and confined information with the present research alone regarded is given. Even so and probably to one's great surprise, the abstracts containing all five steps in this sequence was extremely rare (less than 5%). Hyland reports approximately half of the abstracts included no method section, about 55% excluded an overt introduction and 78% omitted a conclusion. Then, what happened to the traditional IMRD? Do we have to reject it? Note what claim Hyland has on this obvious discrepancy between what writers practically do and what text books tell them to do.

Despite the admonishments of some researchers (e.g. Salager-Meyer, 1990), writers obviously chose to represent their work in ways that fail to conform to a universal 'ideal' of information structuring. Clearly, then, more than summarizing is taking place (Hyland 2004: 68).

Although Hyland’s data shows a very different tendency from the previously referred studies, the perceptual saliency of the IMRD should not be disregarded. Whether the academics who strongly recommend the conventional IMRD structure in writing an abstract have empirically proved that the model is indeed the standard or not, their intuition or personal experience tells us much. As I mentioned earlier, the answer seems to underlie what an individual writer wants to achieve through his or her abstract - (a) giving potential readers a map to facilitate their decision of where and whether to read or (b) enticing those who have read the abstract to step in the detailed body of the article. To achieve the latter, as Hyland noted above, "more than summarizing" needs to be drawn.

2.2 The CARS model: a promotion of the further details

From the previous sub-section, a naturally entailed question is what others a writer needs to consider than mere summarizing when writing an abstract that sells well. Suppose a certain field of a research, where a large number of academics compete with one another to achieve their academic goals and consequently the field itself is changing rapidly and in very complex ways, accommodating various needs. In addition, the readership is a key indicator of success in the field. Due to the space
constraint on an abstract in academic journals, however, summarizing all four moves or parts in one's RA would take inefficiently much space because readers' minds are usually reluctant to read a lengthy passage with every detailed information in it when they are not yet sure whether to read the rest of the article. This reminds us of advertisements on TV; a good advertisement not only leaves the audience strong impression but also is brief enough. Psychologically speaking, the longer an advertisement is, the less attention it will get. Likewise, a writer of an abstract in a relatively competitive academic field should make choices of which to include or not and simultaneously which to elaborate further or just mention briefly to meet their readers' expectations. In this sub-section, we will review several previous studies on what needs to be done in abstract writing for a researcher to survive and gain the readership in a very competitive field of research, mainly concerning the Create-A-Research-Space (CARS) model by Swales (1990, 2004).

In contrast with the IMRD, Swales (1990) describes the CARS model as consisting of three moves or sections. In each move, there are some steps or submoves - some obligatory, the others optional. Note that the model is originally intended to be applied to analyzing an introduction section in RAs, not an RA abstract. Let us see the outlined scheme on the next page first. The rather belligerent terms for each move in the CARS model are designed to reflect what Swales (2004: 226) calls, "ecological competition for research space in a tightly contested territory." As CARS stands for, CARS-type RA introductions create a research space via re-establishing the importance of the research field itself (Move 1); the demand to locate the present research inquiries with regard to that importance (Move 2); and the requirement to explicitly how to overcome and fill this gap (Move 3) at the presence of the discourse community (Swales 1990: 141-142). So far, the CARS model has been adopted in a considerable number of studies with at least modest positive responses (Anthony 1999, Arvey and Tankó 2004, Chu 1996, Gross, Harmon and Reedy 2002, Lewin, Fine and Young 2001, Nwogu 1990, Samraj 2002, 2005, 2008). Corpus-based research employing this model as their analytical tool for the rhetorical structure of abstracts is also not rare. However, I will review these studies on abstracts later in this section. Instead, it would be appropriate to compare the CARS model with its counterpart, the IMRD model first in order to see how so-called "more than summarizing" takes place. Functions of the moves in both models
are summarized in the following table to clarify their similarities and/or differences.

<table>
<thead>
<tr>
<th>Move 1</th>
<th>Establishing a territory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>Claiming centrality and/or</td>
</tr>
<tr>
<td>Step 2</td>
<td>Making topic generalization(s) and/or</td>
</tr>
<tr>
<td>Step 3</td>
<td>Reviewing items of previous research</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Move 2</th>
<th>Establishing a niche</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1A</td>
<td>Counter-claiming or</td>
</tr>
<tr>
<td>Step 1B</td>
<td>Indicating a gap or</td>
</tr>
<tr>
<td>Step 1C</td>
<td>Question-raising or</td>
</tr>
<tr>
<td>Step 1D</td>
<td>Continuing a tradition</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Move 3</th>
<th>Occupying the niche</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1A</td>
<td>Outlining purposes or</td>
</tr>
<tr>
<td>Step 1B</td>
<td>Announcing present research</td>
</tr>
<tr>
<td>Step 2</td>
<td>Announcing principal findings</td>
</tr>
<tr>
<td>Step 3</td>
<td>Indicating RA structure</td>
</tr>
</tbody>
</table>

Figure 3. A CARS model for article introductions (Swales 1990: 141)
<table>
<thead>
<tr>
<th>Move 1</th>
<th>Introduction</th>
<th>Establishing a territory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>* outlines the author's purpose or goals of the research, or the problems about to be tackled.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* cites previous studies.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* claims interest and/or significance of the research area. (step 1)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* states (i) knowledge or practice or (ii) phenomena, and emphasizes the uniqueness of the topic. (step 2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* recapitulates items of previous research. (step 3)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Move 2</th>
<th>Method(s)</th>
<th>Establishing a niche</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>* indicate the methodology used in the research. (data, materials, research design, etc.)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* counter-claims. (step A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* states that previous work suffers from limitation. (step B)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* raises question or problem. (step C)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* provides weaker challenge of the previous research. (step D)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Move 3</th>
<th>Result(s)</th>
<th>Occupying the niche</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>* summarize the general findings of the research.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* indicates the main purpose of the present study. (step 1A)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* describes the main features of the study. (step 1B)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* summarizes the findings. (step 2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* indicates the structure or the content of the remainder of the RA.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Move 4</th>
<th>Discussion (Conclusion)</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>* interprets the results</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* suggests implications</td>
<td></td>
</tr>
<tr>
<td></td>
<td>* applies the findings to wid-</td>
<td></td>
</tr>
</tbody>
</table>
er fields of study.

<table>
<thead>
<tr>
<th>A typical</th>
<th>an RA itself</th>
<th>introduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>structure</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Comparison of IMRD and CARS model and their functions.

In a brief look at the table 3, first one can notice Move 1 in CARS is more general than that in IMRD in terms of topic specificity. Both of them provide something introductory, but in CARS Move 1 does not directly provide author's purpose of the present research. In fact, this function (or step) appears later in Move 3, where the main concern is on the present study not the research field in broader scope. In a somewhat bold perspective, Move 1 in CARS seems to favor elaborating background information while avoiding to project the thesis statement of the research to be conducted earlier. This relatively peripheral background information in Move 1 of CARS is likely to function as an "interest-getter", because without sufficient understanding of a research field one does not usually pursue further. This argument could be supported more strongly especially when the research field under investigation lacks "well-defined sets of problems" and/or well-defined direction of inquiry (Hyland 2004: 71-72). Certainly, among disciplines which put more value on what was done in a specific research than what has been known in the field, such as hard sciences, a purposive move in earlier position may be more plausible. Secondly, Move 2 in CARS may be included in the introduction of the IMRD pattern. This again might be an evidence of elaborated and thus possibly longer introductory moves in CARS. It seems that the CARS model put great emphasis on justifying the research to be dealt with comparison to the IMRD. Finally and perhaps most noteworthy, the final move in CARS strikingly resembles the four moves in IMRD even though in CARS the conclusive or implicational function in the move 4 in IMRD seems omitted. This final move in CARS can be considered a summary of the whole paper in order. However, the strong tendency of optionality of steps in the move 3 of CARS suggests that the model offers greater flexibility to
writers when summarizing than the IMRD model does. Logically speaking, once a reader of an abstract is very interested in the subject that the abstract conveys, we can expect the reader to search for and study the parts such as methods, results, or discussion in the body of the article in detail even if the abstract itself does not contain any of those parts. Now, we seem to be in a position to give some partial answer to the question asked at the beginning of this section. What could or should an individual do when he or she writes a promotional abstract other than faithful summarizing of the whole research paper? A writer could (a) elaborate the introductory or contextualizing moves in his or her abstract, which would justify the research conducted and/or (b) choose not to summarize all four moves in the IMRD model, but to provide only relatively important moves instead so as to achieve the optimal brevity. Now, we know the IMRD style is not the only choice, and the CARS model may accomplish another purpose of the writer, namely promoting the RA itself. Then, it is time to review how successful the CARS model has been in analyzing RA abstracts with empirical data.

In Lorés (2004), she examined rhetorical structures of a corpus of 36 abstracts. She states 30.5% (i.e. 11) of the abstracts in her corpus belonged to the CARS structure. Interestingly, she provides 3 more instances of CARS-type-like abstracts, but she classifies these abstracts as "combinatory structure" (Lorés 2004: 286). She argues that the combinatory structure is a mixture of both IMRD and CARS structures. According to her, it consists of 3 moves (or sections), the first two moves of which are the same as the CARS model but the last move corresponds to the IMRD structure. In Lorés' account, the third move has the following 3 steps (Lorés 2004: 285).

<table>
<thead>
<tr>
<th>Function</th>
<th>Relevant Moves in IMRD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>the general purpose of the paper</td>
</tr>
<tr>
<td>Step 2</td>
<td>the methodology</td>
</tr>
<tr>
<td>Step 3</td>
<td>the findings</td>
</tr>
</tbody>
</table>

Table 4. The steps in move 3 of Lorés' combinatory structure. (Lorés 2004: 285)

Even though another function, the methodology, is added, the substructure of the third move is basically the same as that of the original
CARS. Therefore, it would not be such a big challenge that we have to posit another model. Furthermore, Swales’ recent revised model of CARS does have an optional step of the third move, which is titled as “summarizing methods” (Swales 2004: 232). Hence, I would say the proportion of the CARS-type abstracts in Loris (2004) could be raised up to about 40% - i.e. 14 out of 36 abstracts. It appears that the CARS model manages to account for the other half unexplainable solely by the traditional IMRD model. As you noticed from the above, however, there have been some minor challenges resulting in modifying the 1990 CARS model into the 2004 version.

Since the challenges against the 1990 CARS model are well introduced and discussed in Swales’ (2004: 227-232) book and they were just minor, it would be more proper to show the product (i.e. the 2004 version of CARS) first and then focus on what has been changed from the previous CARS model without specific descriptions about those individual challenges. In Swales (2004), the revised CARS model is proposed separately, divided into two figures. Yet, there seems no apparent reason to do so, and it could be more inspiring to get the whole picture at once.

![Diagram of CARS model]

Figure 4. Move structures in a revised CARS model (Swales 2004: 230, 232)
The first thing to note is that the revised CARS model enables citations to occur throughout, while literature review statements of the earlier version were restricted to the first move, "establishing a territory." See notes written next to each of the move titles - citations required (Move 1), and citations possible (Move 2, 3). We can find that the requirement for citations in Move 1 still remains. Thus, the new model has not been modified dramatically, but accommodates more possible variations. Next minor change is the incorporation of step 1 and 2 in Move 1 of the 1990 CARS. This is due to difficulties in operationalizing the differentiation of claiming centrality and making topic generalizations. But, this is not the end of the story. The revised CARS also allows topic specification or clarification with the title of the only step in Move 1, that is "topic generalizations of increasing specificity," which captures the general direction of logical flow in the introduction part (general to specific). The third to emphasize is the convergence of the three functionally similar steps in Move 2. Swales (2004: 230) accepts Chu's (1996) criticism that although counterclaiming (step 1A) and question-raising (step 1C) are the far rarer options than indicating a gap (step 1B), the distinction among them is not always so clear. The step 1D of Move 2 underwent a change too because Swales thinks that the nomenclature, continuing a tradition, was quite vague in meaning (Swales 2004: 229-230). Now, we need to discuss the "cyclicity" between Move 1 and Move 2. The "cyclicity" in the new CARS model means that Move 1 and 2 can be made a set and repeated multiple times within an introduction as the figure 3 shows. Actually, the "cyclicity" is not a new finding. In Swales (1990), he does provide some examples of the "cyclicity" in relatively longer introductions. Next, convinced by Samraj (2002), Swales adds an optional step to Move 2. According to Swales, presenting positive justification generally follows a gap indication, which justifies the value or significance of the present research. In my opinion, however, this justificational function or step might be found throughout the introduction. As you can see in Move 3 of the 2004 CARS model, the step 6 appears to work very similarly to presenting positive justification, and Swales himself admits it (Swales 2004: 231). Furthermore, claiming centrality in Move 1 of the former CARS model indicates that the justificational statements may indeed be present in all three moves. While both the first and the second moves have been developed into a simpler scheme avoiding overly rigorous applications,
the revised Move 3 has headed for a slightly different direction. The original step 1A and 1B in Move 3 have been merged under the label, announcing present research descriptively and/or purposely, but the other steps have been diversified. Steps 2 through 7 are all optional, but again they have been ordered just like the structure of the full body of the RAs. However, it seems quite confusing to see how the second step, presenting RQs or hypotheses, can be distinguished from the purposive announcement in step 1, given that research questions are about the same as research purposes. The final feature I would like to mention is, in Swales' terms, "roadmapping" function of the step 7. Referring to Kanoksalapatham (2000), Swales argues that when a research paper is not arranged in IMRD-like sections, this outlining remark becomes close to compulsory (Swales 2004: 232). But, I do not think this option is usual in RA abstracts as well, because of the space constraint.

After reviewing both CARS models, I found out that there are two general tendencies. First, some steps do not have a specific place to occur. For instance, steps such as citations and positive justification may happen throughout the Introduction. Second, some steps seem to be dependent on another step. As Swales states, a gap indication was immediately followed by positive justification in Samraj's (2002) corpus (Swales 2004: 280). Step 6 of Move 3 can usually appear after step 6, announcing principal outcomes. Definitions clarifications also come after more general statements (i.e. Move 1 and/or 3). Cyclicity is activated between Move 1 and Move 2, so they seem to be dependent on each other. In sum, there seem to be some possibilities and even needs that the retention and the dependency of some steps should be explained more systematically.

2.3 The continuum of various RA abstracts

In the previous sub-section, I showed the IMRD and the CARS are similar in some ways. Though it might be a somewhat over-generalized or hasty conclusion, the CARS model is a variation of the conventional IMRD with an addition of the elaborated introductory parts (i.e. Move 1 and 2 of the CARS). What I am trying to argue in the article, then, is that many different organizational patterns of RA abstracts can be placed in the continuum rather than the traditional dichotomous categories of the previous research. Some abstracts may strongly resemble the
macro-structure of their full article, (the IMRD pattern) Others may be
significantly related to the introduction of their RA (the CARS pattern). 
Still, we may find many lying between these two extremes as the figure
4 models.8) 

<table>
<thead>
<tr>
<th>CARS-like structure</th>
<th>IMRD-like structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. similar to the introduction</td>
<td>similar to the full paper</td>
</tr>
<tr>
<td>b. promotional purposes</td>
<td>faithful-map purposes</td>
</tr>
</tbody>
</table>

Figure 5. The continuum of various RA article structures

In Samraj (2005), she used both the 1990 version of CARS model and
the traditional IMRD model in order to find out what structural
relationship exists between introductions and abstracts in RAs. She did
not conduct one-to-one comparisons on an introduction and its abstract,
which is one critical drawback of her research. Adding up the instances
of a certain move in abstracts and introductions in her corpus respectively
and comparing them does not show a direct relation, though it may
reveal general tendency of an individual genre. To put it another way,
multiple comparisons of an abstract and an introduction from the same
sources (i.e. RAs) would be more proper in that purpose.9) Anyway,
her research findings indicate that 1) in both of the disciplines (i.e.
Wildlife Behavior and Conservation Biology) in her research, abstracts
generally follow what the IMRD model regulates, and 2) while
Conservation Biology abstracts show moves often found in introductions,
Wildlife Behavior abstracts do not. Hence, we can now question whether
the status of IMRD and that of CARS are really dichotomous as
traditionally assumed, given that there could be some disciplines where
abstracts contain moves from both IMRD and CARS models. In this
paper, I incorporate IMRD and CARS into a new model, which I hope
can account for what previous studies have overlooked.

8) I assume those just partially explicable by either IMRD or CARS alone lie somewhere
in the middle.

9) Again, the one-to-one comparisons of introductions and abstracts from the same sources
are out of the scope of this paper.
3. The auto-featural model

As you saw in the review of the literature, both CARS and IMRD are not complete enough in precisely describing what variations an abstract may have as its rhetorical structure although both models did provide very clear and helpful advice on organizing the majority of typical abstracts. In order to give an individual writer full choices of various structures, here I propose an analytic scheme, which does not have any established ordering. Instead, the model allows co-occurrences of steps, which were strictly ordered in CARS and IMRD models, but whose order is in fact hard to determine; after all, the steps could be in the same sentence, clause, phrase, or even word. Besides, examining the co-occurrence patterns will help us understand what steps obligatorily co-occur, what steps are possible to co-occur, and what steps are not likely to co-occur. Even if there are indeed some strict orders of steps, the current model can capture them. On the contrary to former models, the current model, "the auto-featural model," is not pre-determined but post-determined. It is a kind of blank sheet, on which one can draw whatever there exists. In this regards, it might not be appropriate to name this "a model." It would be the result(s) of this scheme that we can call a model or models.

3.1 Four types of RA abstracts

As a starting point of this model, a writer should decide how many parts or moves an abstract he or she analyzes can be divided into.

<table>
<thead>
<tr>
<th>2-part type</th>
<th>3-part type</th>
<th>4-part type</th>
<th>5-part type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move 1</td>
<td>Move 1</td>
<td>Move 1</td>
<td>Move 1</td>
</tr>
<tr>
<td>Move 2</td>
<td>Move 2</td>
<td>Move 2</td>
<td>Move 2</td>
</tr>
<tr>
<td></td>
<td>Move 3</td>
<td>Move 3</td>
<td>Move 3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Move 4</td>
<td>Move 4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Move 5</td>
</tr>
</tbody>
</table>

Figure 6. The cognitive partitions of RA abstracts
The reason why I propose the four types in figure 4 is i) abstracts consisting of only one part are very unlikely, ii) the 3-part and 4-part types were already very well attested in CARIS and IMRD model respectively, iii) the 5-part type seems to exist according to some studies (Hyland 2004, Fho 2008, Santos 1996) though not prevailing, and iv) a type with more than 5 moves is cognitively too complex and furthermore it might cut an abstract into unnecessarily tiny pieces, arousing various disagreements among raters. Note that I did not title any of moves in the figure so that a rater should decide the number of moves in an abstract mainly based on his or her cognitive intuition rather than on any biases of functional or lexical markers. This scheme is also advantageous in that it can account for cases, where some expected moves are omitted. For instance, an abstract consisting of introduction, methods, and results without the discussion move will be allotted to the 3-part type instead of the 4-part type. In fact, we do not know whether we could entitle any move in our abstract introduction, methods, results, or even a new categorial name yet. Only after the partition decision, we could judge what functional feature(s) each move in an abstract has with closer and deeper analyses.

3.2 Possible features in a move

The separation of moves only based on personal intuition has to be further supported by examining whether the functional feature(s) of a move separated from the other move(s) are uniformly clustered in accord with those of other abstracts. Since the judgement of functional feature(s) without any well-established criteria can cause huge chaos, we need to draw some criteria from the previous studies reviewed in this article to find out what feature(s) a certain move can have. The following table is a list of functional features or steps within moves or sections in the abstracts previously studied. Features looking basically similar or difficult to distinguish in their functions have been combined into a single unified feature.

---

10) Certainly, there might some factors influencing rater's decision making. One of advantages of the present model, however, is that those factors will be able to be figured out after some analytic processes when there is a clear agreement among raters. When there is not, the factors would be too subjective to operationalize.
<table>
<thead>
<tr>
<th>Title</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 Contextualization</td>
<td>Provides context of the present research, claiming interest or importance of the topic and/or the field of the research.1)</td>
</tr>
<tr>
<td>F2 Gap indication</td>
<td>Indicates a gap in the research field, counterclaims, and/or raises questions against what has been previously established or claimed.</td>
</tr>
<tr>
<td>F3 Extending previous research</td>
<td>Gives a weak challenge to former research while introducing their research as corresponding to the current consensus in the field, and/or adds the present new findings to what has been known so far.</td>
</tr>
<tr>
<td>F4 Positive justification</td>
<td>Explicitly justifies how significant or interesting the present research is.</td>
</tr>
<tr>
<td>F5 Citations</td>
<td>Cite what has been studied or discussed on the topic either explicitly or implicitly.</td>
</tr>
<tr>
<td>F6 Description of the present research</td>
<td>Describes the main features of the research in question.</td>
</tr>
<tr>
<td>F7 Purposive statement</td>
<td>Presents the purpose of the research.</td>
</tr>
<tr>
<td>F8 Presenting RQs or hypothesis</td>
<td>Outlines the research hypothesis or questions.</td>
</tr>
<tr>
<td>F9 Definitional clarification (specification)</td>
<td>Further specifies the relatively general information the adjacent feature introduced.12)</td>
</tr>
<tr>
<td>F10 Summarizing methods</td>
<td>Provides information on subjects, design, procedures, and/or data used in the present research.</td>
</tr>
<tr>
<td>F11 Summarizing findings</td>
<td>Reports the principal findings of the present research.13)</td>
</tr>
<tr>
<td>F12</td>
<td>Interpretation of the findings</td>
</tr>
<tr>
<td>-----</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>F13</td>
<td>Giving recommendations</td>
</tr>
</tbody>
</table>

Table 5. A list of features used in the current analysis

The features listed on table 5 are assigned to every sentence in an abstract. It has been argued that a move and its submoves or steps in genre analysis is a functional unit rather than a formal unit and thus their domain of linguistic realization can be fairly flexible—i.e. several sentences, a single sentence, a clause, or a phrase (Swales 2004: 228-229, Pho 2008: 235). To account for their relative order more precisely, however, we need to set a reliable and fixed domain. At a clausal or phrasal level, for instance, stylistic or syntactical choices alone can affect the order of moves or steps within that specific clause or phrase. In the auto-features model, the domain of a feature is strictly a sentence. What advantages of this decision has over the present flexible approach will be given in detail in the following sections.

3.3 The autonomy of features

11) This function is different from F7 (Positive justification) in that it (F1) justifies the research to be presented via giving background knowledge. If there is any explicit linguistic element validating the current study such as important, valuable, interesting, etc, the sentence will be tagged with F7 regardless of whether F1 is present or not.

12) Unlike others, this feature must be adjacent to another thematically coherent feature in the previous or next sentence, clause, or phrase, because the notion of “specificty” is not absolute, but relative and dependent. In other words, F8 cannot occur in the absence of relatively general statement on the same theme. See examples with this function in Dressen & Swales (2000) and Samraj (2002).

13) This function accompanies with hard evidence, which is objective, and reportive rather than subjective and interpretive. If there are both a report of non-interpretive outcomes and interpretational comments on it, the sentence containing them will have both F11 and F12.

14) In listing, operationalizing, and/or refining of these features, I referred mainly Swales (1999, 2004), Satoke (1995), and Hyland (2004) in decreasing order of significance.
Contrary to the previous models, the analytical features used in the present model are basically autonomous. In both CARS models, the linear order among moves or steps is fixed and a move has its unique step(s) in terms of the hierarchy. In my opinion, those models are too restrictive because, as we already reviewed, some moves or steps can be repeated, omitted, or reversed in their linear order in the CARS models and some steps can be found in every move not restricted to a specific move. Furthermore, some steps seem to co-occur, making us indecisive about what function a given linguistic domain, say a sentence, has. In Santos (1996), we can find an example of this.

(1) This study examines the responses of 60 Spanish, Chinese, and German L2 learners to English sentences with empty pronominal categories (BCs) (Santos 1996: 491).

One can easily notice in the example that "purpose statement (P7)" and "summarizing methods (P10)" occur within the same sentence boundary. Also, it is hard to tell which feature precedes which, and even if one were able to decide the order, there would be no apparent beneficial effect to have done so. In the example above, the order seems not so significant in constructing the overall rhetorical structure of the abstract. Santos (1996) calls this "move embedding," and indicates that this phenomenon is quite common in all five moves of the abstracts in his corpus. Hence, endowing each feature in the table 5 with the autonomous nature would help us tackle a number of problems which the previous models have raised. The idea of autonomous features is originally derived from Goldsmith's (1976) "auto-segmental phonology." The key concept of the autonomy is quite simple; every feature can be combined with another feature still preserving its individual function and each feature is basically independent from each other though some features can be associated, provided that there is a strong tendency of co-occurrence. While the previous models are strictly linear in describing various rhetorical structures of abstracts, the present auto-featural model is multi-dimensional allowing the dependency, recursivity, linearity, optionality and obligatory of features at the same time. The auto-featural approach gives us more descriptive power, specifically when the

19) The move embedding was also found in Pho (2008).
rhetorical structure of an abstract is seemingly very complex and unaccountable in the traditional frameworks of the move analysis.

3.4 The feature dependency and recursivity

The feature, *definitional clarification* (F9), is present only when there is a more general description adjacent. It is because in definition giving more specific information always presupposes the presence of less specific information. Therefore, we can claim that F9 is dependent on a more general statement which has a different function. Likewise, if there is a feature appearing only after or before some other feature with a strong tendency, we need to check whether those features are correlated and figure out the directionality of the dependency. In the auto-featural model, this dependency of features can be shown via some quantitative methods. If some statistical outcomes, say the correlation coefficient, reveal there is the statistically significant correlation of two or more features, it would be a good indicator of the need for a closer qualitative analysis of what makes them co-occur.

As Swales (2004) notes, some steps (in my model, features) can repeatedly appear in different moves or even in the same move. Thus, the current framework allows repetition of features. As stated in section 3.2, the domain of the realization of a feature in the proposed model is a single sentence. Since a multiply recurring feature in the same sentence boundary is intuitively very unlikely, the confinement of the domain of a feature needs not be reduced into the smaller linguistic unit. The problem is, however, if a certain feature extends to several sentences in an abstract, the present model may unnecessarily proliferate the functionally and thematically same feature over the sentences. In this case, the domain may not be a single sentence but several sentences.

To deal with this problem, I adopt a post-hoc modification of the domain. Having allotted a specific feature to several sentences, we need to examine whether the feature of each sentence is thematically coherent. If the feature in adjoining hypothetical sentences A and B is functionally the same but dealing with thematically different subject, this feature, in effect, has to be regarded as the repetition. In other words, the single sentential domain still works. In the opposite case, we should extend the domain to the sentences which are both thematically coherent and functionally same.
3.5 The strict linearity between/among features

Through the post-hoc modification introduced in the previous sub-section, the auto-figural model can somehow preserve the claim that a single sentence is the domain of the feature realization. This time, I would like to introduce "the strict linearity." As I have instructed, at a phrase-internal level it is difficult to identify the linear order of features. In addition, it may also be quite tricky to articulate the relative precedence between/among clauses within a sentence. Even though in a few cases we do notice a clear linearity of clauses within a sentence boundary, it would be better to avoid such a risk. In case that there are indeed many instances where some features appear in specific order, we can conduct a deeper level analysis to see the inter-clausal order of those features afterwards. So, the present framework allows only the inter-sentential linearity. In other words, features can be ordered only if they appear in different sentences, which is called "the strict linearity."

An example of the application of the present model will be provided in the appendix A.

4. Discussion: suggestions for further research

In the following sub-sections, I suggest what further research and analysis have to be conducted in order to support my proposal in this article. What this paper argues is just a preliminary proposal. Thus, depending on the empirical results of subsequent analyses, the proposal is subject to be modified.

4.1 Quantitative analysis

As I mentioned, the model is just a tool for analysis and in fact the product of the analysis using the current model is what we really want to know - the generally acceptable variations of rhetorical structures of an RA abstract. The goal will not be achieved by analyzing just a few number of papers. To draw a convincing generalization, we need to

16 I do not necessarily dismiss other possibilities altogether. The term, "strict linearity," implies there can be "weaker linearity."

construct large enough corpora. Actually, there have been many studies dealing with large corpora, but none of them have used a satisfactory statistical methodology. Most of the studies simply provide proportional information, but how often some instances happened in a sample may not be exactly the same or at least similar in another sample. To expect the behavior of the population, some inferential statistics as well as descriptive statistics are needed. In sum, next procedures are like the following.

A corpus: abstract 1 abstract 2 ... abstract n-1 abstract n

Analyzing using the auto-featural model

Coding the outcomes

Doing statistical analyses

Making conclusion based on statistical data
A. Very frequently occurring patterns
B. Not usual but still possible patterns
C. Very unusual so unrecommended patterns

Figure 5. The future procedures of statistical analysis

4.2 Interview and survey

After attaining results based on the statistical analysis, we need to examine what made the authors construct their abstract in that pattern. The best way to do this is probably asking each author why they did so. However, the number of previous studies do include personal interview data is not large, and for the abstract analysis the only case I know at the time of writing is Hyland (2004). He quotes comments from his several professional informants, and the comments clearly support what Hyland is trying to argue. Yet, even with some interview data we cannot obtain what we want. The interview data is a good starting point, but it just provides a partial evidence. Therefore, what I hope to do in the later research is first to interview some of the authors
who wrote an abstract included in my future corpus. Next, I will make a questionnaire based on the interview data, and then carry out a survey, this time including many more interviewees. Again, through some statistical analyses, we can get more reliable generalizations.

4.3 Broader applications of the auto-featural model

The auto-featural model is based on several previous models. Among them, the 2004 CARS model was the most influential. The CARS models have been used for RA introductions. Thus, it is not hard to assume that the auto-featural model can be employed for genre analysis on RA introductions. Furthermore, if relevant features or steps are identified, we can apply the model proposed in this article to other genres as well - for example, RA conclusions, RA acknowledgements, etc. The model can be adopted in analyzing other sub-genres of abstracts such as master or doctoral dissertation abstracts, abstracts in conference proceedings, and abstracts of review articles. To give pedagogical implications, the findings of those applied analyses above can be compared with those of L2 learners of English. Given that most of textbooks for academic writing are overly prescriptive, possibly neglecting what is really happening in academic discourses, the rather descriptive nature of the outcomes through the auto-featural model will provide a wider variety of choices to the learners, suggesting what is the most frequent pattern and what is not in executing specific purposes of an individual author.

4.4 Concluding remarks

We have not seen the results from any of these suggested in this section yet. Therefore, it is premature to judge how successful the proposed model would be. Maybe, the results will not so different from those of previous research. Maybe, some other features will be identified or some features will be totally merged with other features in the course of analyses. Since this article mainly focuses on theoretical issues, the incompleteness of the article must be filled in the future studies with empirical evidences.
References


Jae-Hyong Park
jehp@paran.com
Appendix: The application of the auto-featural model

<table>
<thead>
<tr>
<th>This study examined whether a younger starting age is advantageous in a situation of minimal exposure to an instructed foreign language (≤4 hours classroom contact per week).</th>
<th>F6</th>
<th>F7</th>
<th>F10-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous theoretical and empirical studies indicated there should be no advantage for an earlier start.</td>
<td>F5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese college students who started studying English between ages three and twelve (n=61) were examined on a phonetic discrimination (/l/-/w/) and grammaticality judgement task (CJT).</td>
<td>F10-2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After controlling for language aptitude and amount of input, statistical correlations were found between starting age and scores on the CJT ( r = -.38 ) but not the phonemic task ( r = .03 ).</td>
<td>F10-2</td>
<td>F11-1</td>
<td></td>
</tr>
<tr>
<td>These earlier starters were also compared to peers who began study in junior high at age twelve or thirteen (n=139) on the same measures.</td>
<td>F10-3</td>
<td>M2</td>
<td></td>
</tr>
<tr>
<td>The earlier starters were found to score statistically higher on the phonemic but not morphosyntactic measure, and this remained true in an ANCOVA analysis where total amount of hours of study input were controlled for.</td>
<td>F11-2</td>
<td>F10-3</td>
<td></td>
</tr>
<tr>
<td>A robust ANCOVA testing for differences at different levels of input found interesting interactions between group affiliation and amount of input.</td>
<td>F10-4</td>
<td>F12-1</td>
<td></td>
</tr>
<tr>
<td>Language attitudes were also tested.</td>
<td>F10-5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The evidence shows there can be perceivable age effects for linguistic measures even in a situation of minimal exposure to a foreign language, but these may not emerge until a substantial amount of input has been gained.</td>
<td>F12-2</td>
<td>M3</td>
<td></td>
</tr>
</tbody>
</table>

* Functions with same tags are related thematically.