# Lexical Patterns in Venezuelan Spanish Academic Texts*

### Elizabeth Arcay Hands

*Departamento de Idiomas Modernos, Facultad de Ciencias de la Educación.*
*Universidad de Carabobo, Valencia-Venezuela.*
*E-mail: earcay @ thor.uc.edu.ve.*

## Abstract

This paper presents the results of the adaptation and application of M. Hoey's lexical patterns analysis of non-narrative texts to Venezuelan Spanish academic texts. Five academic essays were selected, one written and published by a renowned Venezuelan scholar and 4 written by language teacher-trainees (of Spanish L1 and English L2). The analysis consisted in determining if the properties of the bonds of the nets, created by lexical repetitions among the sentences of the texts, also apply to Spanish texts. The results show that these properties do apply to the examined texts and that they allowed the creation of abridgments and summaries. It is suggested that this analysis carries a great potential for the description of the textual organization of the Spanish language and provides evidence for the genre approach to the study of texts.

**Key words:** lexical patterns, text analysis, academic texts, Spanish.

# Patrones léxicos en textos académicos escritos en español de Venezuela

## Resumen

Esta investigación presenta los resultados de la adaptación y aplicación del análisis de patrones léxicos para textos no narrativos a textos académicos escritos en español de Venezuela. Se seleccionaron cinco ensayos académicos, uno escrito y publicado por un académico Venezolano de reconocida trayectoria y cuatro escritos por futuros profesores lenguas (de español de Venezuela como lengua materna y de inglés como lengua extranjera). El análisis consistió en determinar si las propiedades de las conexiones que se establecen entre oraciones a distancia a través de nueve tipos de repeticiones léxicas (Hoey, 1991) en textos no narrativos escritos en inglés, también ocurren en textos del mismo tipo escritos en español. Los resultados del análisis indican que fue posible comprobar dichas propiedades y que las mismas permitieron la creación de resúmenes de distintos grados de comprensión. Se plantea que el análisis posee un gran potencial para la elaboración de resúmenes utilizando las oraciones originales de los textos y para la descripción de la organización textual del idioma español y que provee evidencia para la utilización del enfoque de género para el estudio de diferentes tipos de texto.

**Palabras clave:** patrones léxicos, análisis de textos, textos académicos, español.

## INTRODUCTION

### Lexis and Texts

One of the possibilities of textual analysis which has been gaining importance and attention among researchers is the one related to lexis, its characteristics and properties. On the one hand, because it plays a major role in written texts comprehension (Perkins, 1992; Hoey, 1991; Grabe, 1991; Vacca, 1981), and production (Hoey, 1991; Willis, 1993; Nattinger y Decarrico, 1992). On the other, because recent findings from corpus linguistics research (Sinclair, 1991) have pointed to features of

language regarding its lexical characteristics as a rich source of information not thought of until now.

Two of the ways of carrying it out are with analysis of topical structure development and of lexical repetitions. The first type consists in determining the sequences of topical development, specially of written texts (Lautamatti, 1987; Cucatto, 1995; Moss, 1995), using the patterns that new and given information form in a text. Enkvist (1987) classifies this type of analysis as based on the "information structure of a text" (p. 29).

The second type of analysis focuses on the repetitions that occur in texts. Youmans (1991) developed what he called a "Vocabulary Management Profile" (VMP). With this instrument, this author analyses topic development in literary texts with a computer program in which 35-word intervals are used to register the appearance of new words. This appearance of new words originates "peaks" and "valleys" in the computer profile which allow a type of text characterization related to the innovation and coherence principles.

Hoey (1991) developed a lexical pattern analysis to study non-narrative texts in which the repetition of lexical items allows to detect organizational patterns in texts. With this analysis cohesion can be studied under a discursive and not only grammatical perspective, as it permits the detection of marginal and central sentences, and with them, forms of global organization in texts. Also, it allows the creation of different types of abridgments and the detection of bonding properties among sentences.

The positive results obtained with this analysis in English texts (Hoey, 1991, 1994) motivated its translation and adaptation to Spanish to determine if these bonding properties among sentences also occur in Spanish texts and if the creation of representative abridgments was also possible. Following, the pilot study carried out with an essay written and published by a renowned Venezuelan scholar (Páez Urdaneta, 1985), and the further testing of the analysis with 4 academic essays, written by language teacher-trainees (of Spanish L1 and EFL) (Arcay, 1996), will be described.

## 1. THE LEXICAL PATTERNS ANALYSIS

This analysis consists in the recording of 9 types of lexical repetitions:

-simple lexical repetition

-complex lexical repetition

-simple mutual paraphrase

-simple partial paraphrase

-antonymous complex paraphrase

-other complex paraphrase

-substitution

-co-reference

-ellipsis

which can occur among the sentences of a text, to determine **links** and **bonds** created by these repetitions among non-adjacent sentences (Hoey, 1991). A **link** is defined by Hoey (1991, p. 266) as "a connection by repetition between items in a text. If one item in one sentence repeats an item in another sentence, then the sentences are said to be linked". A **bond** is defined as "... a connection that exists between a pair of sentences by virtue of there being an above-average number of links relating them. Usually, the requisite number of links is three, and it is never less than three, but sometimes for texts in which there are a great number of repetitions, the threshold may be four links or more".

To record lexical repetitions in the form of links, all the lexical items in each sentence are compared with the lexical items of all the sentences that follow. For example, in a 20-sentence text the lexical items of sentence one are compared with the lexical items of sentences two to twenty. Then, the lexical items of sentence two are compared with the lexical items in sentences three to twenty. This procedure has to be repeated until all sentences have been compared with the ones that follow. In turn, the recordings of these links created by the different types of lexical repetitions allow the construction of a repetition matrix from which the bonds created by an above-average number of lexical repetitions can be identified by the examination of each of the cells. The bonds that result form the analysis of the matrix are then recorded in a table to first determine the number of these bonds that are created with

previous and subsequent sentences, and second to determine the total number of bond per each sentence. Once the frequency of the different number of bonds of each sentence in the text is organized in another table, marginal and central sentences can be identified. **Marginal sentences** are those sentences which do not establish bonds with any other sentence in the text and, as it important to mention, the term is not evaluative. **Central sentences** are those sentences that form an above average number of bonds (Hoey, 1991, p. 265). In turn, this identification of these two types of sentences and of thresholds allows three different abridgment procedures to obtain summaries of various lengths.

Hoey proposes that these bonds have properties which allow a strong and a weak claim to describe the relationships that are created among distanced sentences thus connected in a text, and that this relationships cannot be solely explained by the repetitions themselves because these properties do not emerge in sentences which share lexical items but do not belong to the same text. The consequence of the strong claim is that when a pair of bonded sentences is placed consecutively, they are coherent even though there can be a considerable distance -measured in number of sentences- between them.

## 2. LEXICAL PATTERNS IN SPANISH

### 2.1. Pilot Study

After a successful trial application of the analysis to three academic essays written in Spanish by language assistants: two Venezuelans and one British (Arcay 1995), I decided to try it with a text which could be regarded as a model of its genre to be able to collect more solid evidence (Arcay, 1995/6). The selected text was an academic essay titled ¿Qué lengua enseñar? [*What language to teach?*]which was published as an appendix of a book about the teaching of Venezuelan Spanish as the mother tongue with a communicational approach, and the author was I. Páez Urdaneta, a renowned Venezuelan linguist.

Again, the application confirmed the properties of the bonds and it allowed the creation of 4 summaries. The matrix originated from this 34-sentence text had 561 cells, and 68 of them indicated bonds (an above-average number of lexical repetitions -links- per sentence): 52

with 3 links, 11 with 4, 4 with 5 and 1 with 6. I n turn, these individual bonds originated a total 138 previous and subsequent bonds among the 34 sentences of the text. Five sentences of these 34 were marginal -they did not bond with any other sentence-, and all but one -sentence 22- can be considered metatext because they do not add propositional content to the rest.

> **21** En razón de lo anterior se intuirá que abogamos por una actitud interlectalista.
> [Because of the above, it can become evident that we are in favor of an interlectal attitude]

> **22** Interlectalismo no es panlectalismo, que es disfuncional.
> [Interlectalism is not panlectalism, which is dysfunctional]

> **30** Resulta obvio que es preferible enseñar algo que es propio, antes que enseñar algo que es ajeno, o enseñar primero algo que es ajeno o irreal para luego enseñar lo propio.
> [It is obvious that it is preferable to teach something which belongs to us, before teaching something foreign or to first teach something foreign or unreal to then teach what belongs to us]

> **33** Que de esta integración resulte (o no) una uniformación lingüística, es ya otro asunto.
> [That linguistic uniformity results or does not result from this integration is another matter]

> **34** Todo depende de una voluntad que está más allá del poder de los maestros.
> [It all depends on a will that is far beyond the control of teachers]

Sentence 30 paraphrases previous propositions, sentences 21 and 30 make the author's opinion explicit and sentences 33 and 34 make exoforic references. With regard to sentence 22, the fact that it adds information, but this information is not mentioned again in the rest of the text, considerably lessens its importance with regard to the achievement of coherence in the text and probably explains why it does not bond with any other sentence.

To exemplify the strong claim of the properties of the bonds and the possibility of creating summaries with central sentences, here follows the two sentences which evidenced the highest number of bonds (sentence 05 with 13 bonds and sentence 26 with 11) and which form a summary of the text which represents 5.88% of the total number of sentences. What is even more interesting, these two sentences together answer the question which served as the title to the text ¿ *Qué Lengua Enseñar?* (*What language to teach?* ).

> **05** Un país monolingüe no habla uniformemente la misma lengua; en él se detectan "dialectos" o variaciones regionales, y, aun en sus medios urbanos, se detectan variaciones grupales que suelen denominarse "sociolectos". **26** La enseñanza de la lengua materna consiste, por lo tanto, en facilitar el acceso del individuo a tantas variedades lectales como necesite ( no en atrofiar una(s) mediante otra(s)) y en concientizarlo en la conveniencia de esta actitud plural.

Sentence 5 states that in a monolingual country a language is not spoken uniformly and that dialects or regional variations, together with group variations in urban settings-called "sociolects"- are detected. Sentence 26 states that the teaching of the mother tongue therefore consists in facilitating individual access to as many dialectal varieties as the individual needs, not in substituting one with other, and to make him/her aware of the convenience of this plural attitude. As the reader may very likely agree, these two sentences together completely answer the question that generated a 34-sentence text, thus confirming the property of the analysis that allow the creation of summaries with central sentences. And also, they also confirm the strong claim of the bonding of distanced sentences because there is a distance of 21 sentences between these two.

## 2.2. Lexical Patterns in students' academic essays

After these positive results were obtained, I decided to further try the analysis with texts from a 17-text corpus obtained for a multidimentional Venezuelan Spanish academic texts analysis (Arcay, 1996). The texts were produced by language teacher trainees of EFL and Spanish L1 of the University of Carabobo, Venezuela, and the prompt used to obtain them was *"How to do linguistic research at Carabobo University and why"*. Four of these texts were selected for the

application of the analysis and the criteria used for their selection was
that they had an appropriate length -measured in number of sentences-
for its application. At first, a range of twenty to thirty sentences was
considered an appropriate number to test the properties of the nets and
matrixes because it allowed longer distances between bonded sentences
and the possibility of creating summaries with different degrees of
compression. Regretfully, as only two texts had more than 20 sentences,
I decided to select two more - the ones with the second highest number of
sentences in each group.

## 3. RESULTS

Of the 17 texts , the selected ones written by the EFL language
teacher trainees were identified as EFL01 and EFL02 and the ones
written by the Spanish L1 language teacher trainees as S03 and S04. The
number of sentences of each text was 23 in the EFL01, 13 in the EFL02,
10 in the S03 and 24 in the S04.

### 3.1. Marginal and Central Sentences and Summaries

As it is shown in table 1, the number of marginal sentences differed
in each of the four texts. Also, the number bonds that marked central
sentences differed.

Table 1
Identification of Marginal and Central Sentences per Text

| Text | # of Snte. | Marginal Sentences | Central Sentences |
|------|-----------|--------------------|--------------------|
| E01 | 23 | 05; 12; 15; 16; 17; 21; 22; 23; | Snte. 04 (5 bonds) |
| E02 | 13 | 10 | Snte. 13 (7 bonds) |
| S03 | 10 | - | Snte. 08 (6 bonds) |
| S04 | 24 | 09; 14; 19 | Snte. 04 (10 bonds) |

The difference in number of marginal sentences can be explained
by the logical probability of appearance of more marginal sentences in
longer texts. Additionally, differences in lexical density can add to the
explanation. Text E01 is less dense lexically speaking than text S04, a
characteristic of the text which can be observed in the relatively low
number of bonds (5 in comparison to 10 in text S04 and 7 in text E02)

which constitutes the highest number of bonds of all the sentences in that text. To supply evidence to test the claim of the analysis that marginal sentences are not indispensable for the preservation of coherence, text E01 -the one with the highest number of marginal sentences- is transcribed without them in the appendix section.

With regard to central sentences, table 1 shows that it was possible to detect one central sentence, i.e. one sentence with a higher number of bonds than the rest, in each text. With regard to abridgments, in spite of the short length of two of the texts, it was possible to obtain at least two summaries per text (Arcay, 1996). Here, table 2 only shows the most compressed summaries -one for each text- which were possible to create with the sentences that evidenced the highest number of bonds. Two of the procedures followed to obtain the abridgments and summaries follow Hoey (1991): one consists in the elimination of marginal sentences, and the other in the selection of central sentences through the detection of thresholds in the sequence of the number of bonds per sentence.

However, in some of the texts, the cline that emerges from this sequence of number of bonds in each text did not mark any threshold which could help in the selection of number of bonds per sentence to create summaries of different levels of compression, or to decide how many sentences to take into consideration to create the possible most compressed summary in each text. So, I interpreted this as if those sentences -the ones with the highest number of bonds and therefore the ones which could be deemed to represent the thesis of the text- "needed" to be complemented with the one or ones which followed it in the cline to allow a combination which could satisfactorily answer the prompt that generated the texts.

### Table 2
Percentage and identification of sentences per summary and per text.

| Text | Summary | % | Bonds | Sentences | Total |
|------|---------|------|----------|------------|-------|
| E01  | 1       | 17,4 | 3, 4, 5  | 4; 8; 14; 18 | 4   |
| E02  | 2       | 15,4 | 6 +      | 8; 13      | 2     |
| S03  | 3       | 20   | 5 and 6  | 5; 8       | 2     |
| S04  | 4       | 8,3  | 8 and 10 | 4; 23      | 2     |

The evidence for the "need" of a combination of the two sentences with the highest number of bonds can be found in three of the texts: E01, S03 and S04. In the first one, sentence 04 was the one with the highest number of bonds -5- and sentence 08 the one that followed with 4 bonds. Together, these two sentences can be said to partially answer the prompt:

**Text E01.** Sentences 04 (5 bonds) and 08 (4 bonds)

> **01 04** Pero¿por qué salir de Carabobo cuando aquí también se genera investigación y muchos profesores procuran que los diseños repliquen otros, den nuevas respuestas, apliquen soluciones? **01 08** Hay obligatoriedad y necesidad de estudiar la enseñanza del inglés bajo el calor de Carabobo, en los barrios marginales, y con todo el peso de nuestro devenir socio-económico.

Sentence 04 asks why go outside the state in which the University is placed to do research if there are teachers there trying to give answers to problems, thus answering the "why" in the prompt. On a first look, this may not be considered a direct answer to why it is important or necessary to do research in that University; however, if the fact that it was a student who is answering the prompt is considered, I believe this becomes an appropriate answer because he/she may have interpreted the question as if students had to go somewhere else to do linguistic research because there was not anybody from the University to help them with their research needs. So, if the fact that there is appropriate guidance from the faculty of the University is stated, the reason to do research there is given. Sentence 08 adds to it stating that in that state there is a need and an obligation to study the teaching of the English language within this socioeconomic context.

Additionally, in this text, the summary obtained with the sentences that show the highest number of bonds in each of the paragraphs of this text, rounds up the answer given by sentences 04 and 08 and originates a very reasonable abridgment in four sentences of a 23-sentence text:

**Text E01:** Sentences 04, 08, 14, 18. (Sentences with the highest number of bonds per paragraph)

> **01 04** Pero ¿por qué salir de Carabobo [ *para cumplir con el curriculum de F.A.C.E y para conocer sobre la problemática lingüística que nos hemos planteado estudiar* ] cuando aquí

> también se genera investigación y muchos profesores
> procuran que los diseños repliquen otros, den nuevas
> respuestas, apliquen soluciones? **01 08** Hay obligatoriedad y
> necesidad de estudiar la enseñanza del inglés bajo el calor de
> Carabobo, en los barrios marginales, y con todo el peso de
> nuestro devenir socio-económico. **01 14** Este camino del
> investigar obviamente le da un sentido más profundo y
> propio a nuestro quehacer de docentes o futuros
> profesionales de la lengua inglesa y, en otros, de la española.
> **01 18** Pero la investigación necesita del apoyo documental,
> humano, material.

Sentence 14 states that to do research gives the EFL and Spanish L1
teaching professions a more profound sense of belonging and sentence
18 demands that to do research, material, personal and documentary
support is needed. With these two sentences, the why is complemented
with reasons related to the very same process of research and not only to
the place in which it takes place. The how is answered making
references to the different resources needed to carry it out.

In text S03, a combination of two sentences –sentence 05 with 5
bonds and sentence 08 with 6 bonds– again give a satisfactory answer.
However, as this text was one of the shortest but with the longest
sentences, it was logical to expect that 20% of the sentences, by itself a
relatively high percentage of sentences, could answer the prompt
satisfactorily. The combination of these two sentences can be found in
the appendix section. In text S04, one of the longest, the combination of
the two most connected sentences also gives a complete answer. But in
its case, the situation changes. These two sentences represent 8.3% of the
total number of sentences –24–, which can be said to constitute a very
compressed summary. If a 24-sentence text can be rightfully represented
by 2 of them, then, the analysis and its procedures to summarize can be
deemed successful.

**Text S04.** Sentences 04 (10 bonds) and 23 (8 bonds)

> **04 04** - [*La manera más eficaz para realizar investigaciones*
> *lingüísticas en F.A.C.E. es con* ] Un centro donde las .
> investigaciones a nivel lingüístico sean recopilados, y de esta
> manera, poder obtener material que sea óptimo en el proceso
> de la enseñanza de la lengua. **04 23** [*Puesto que* ] Investigar

en lingüística y formar dicho centro es esencial como guía de estudio y proyección a nuestro futuro.

The translation of these two sentence follows, to include the added metalanguage and in that way to supply a more direct evidence:

**Text S04**. Sentences  04 (10 bonds) and  23 (8 bonds)

**04 04** [*The most efficient way to carry out linguistic research at F.A.C.E (School of Education, University of Carabobo) is with*] A centre where linguistic research    projects are collected and, in this way, to obtain optimal material for language teaching processes. **04 23** | *Given that* ] To do linguistic research and to create that centre is paramount as a study guide and as a projection to our future.

The only text  in which one sentence suffices to answer the prompt is E02. In it, sentence 13 contains all the necessary information as to how and why to do research:

**Text E02**. Sentence 13 (7 bonds)

**02 13** Si contáramos con un lugar o institución y el recurso humano preparado para ayudarnos a investigar en nuestra área del estudio, serían muchos los beneficios que obtendríamos     para     nuestro     mejoramiento     como profesionales y muchos los beneficios que podríamos aportar al campo de enseñanza de lenguas.

Sentence 13 states that if a place or an institution existed, together with a proficient staff in the linguistic area, many benefits would be obtained for the improvement of the future language teachers and many would be the benefits that these professionals could contribute to the field. So, the institution and the staff would be the answer to the how, and the benefits which could be obtained and contributed would be the answer to the why.

A 15,4% summary of this text was created with the combination of two sentences, number 08 and 13, the one described above. The result of this combination is that in sentence 08 the author of the text explicitly states the resources and elements that he/she considers necessary to do research, and in sentence 13 he/she establishes one condition which would allow successful research activities: if the need for these resources

and elements could be met by the creation of a place equivalent to a research centre, then linguistic research can be possible and justified.

**Text E02.** Sentence 08 (6 bonds) and 13 (7 bonds)

> **02 08** Pero todo este proceso [*de aprender inglés y prepararse para enseñarlo* ] no puede ser exitoso si no contamos con estudios recientes del área o por lo menos actualizados o podemos gerenciar investigaciones propias que aporten datos interesantes al estudio de la enseñanza de lenguas en general. **02 13** Si contáramos con un lugar o institución y el recurso humano preparado para ayudarnos a investigar en nuestra área del estudio, serían muchos los beneficios que obtendríamos para nuestro mejoramiento como profesionales y muchos los beneficios que podríamos aportar al campo de enseñanza de lenguas.

Again, the direct translation of the summary is supplied to include the added metalanguage and in that way to supply a more direct evidence:

**Text E02.** Sentence 08 (6 bonds) and 13 (7 bonds)

> **02 08** But all this process [ *of learning English and preparing to teach it* ] can not be successful if we do not have recent -or at least current- studies in the area or if we can not manage our own research projects that, in turn, can supply interesting data to the study of language teaching in general. **02 13** If we could have a place or an institution and a qualified staff to help us do research in our study area, many would be the benefits that we could obtain to improve ourselves as professionals and many the benefits that we could contribute to the language teaching field.

## CONCLUSIONS

As the analysis carried out above with 5 texts show, the characteristics of the summaries obtained give strong evidence for the claim that the lexical patterns analysis of Michael Hoey can also be carried out with satisfactory results in Spanish academic texts. Hoey (1991, p.190) claims that lexically there may be a link with genre because, possibly, bonding occurs to a greater or lesser extent in texts of a certain type or of a certain level of formality. Then, if bonding occurred

in these 5 Venezuelan Spanish academic texts to such extent that the
properties of the nets were confirmed in all of them, and that it was
possible to construct representative summaries of different levels of
compression, evidence is building up about the organizational patterns
of lexis in language and the importance of the genre approach to the
study of texts (Swales, 1990). So far, English and Spanish have not been
considered very similar languages, at least in the way information is
organized in written texts. If further confirmation is obtained from the
analysis of more Spanish texts, the theoretical and practical implications
of the analysis can be more deeply explored. Obviously, more evidence
is needed, but, if confirmation continues to occur judgments about the
nature of lexical organization will probably cease to be hedged.

## NOTES

1. For Hoey, paragraphing is practically idiosincratic (Hoey,
   1994) and therefore not a reliable source of information about
   bonded sentences. However, as this text allowed the identifica-
   tion of one sentence per paragraph with a higher number of
   bonds than the rest, the procedure of selecting them to obtain a
   summary was tried. Eventhough the results were satisfactory,
   for the reasons mentioned above, precaution needs to be obser-
   ved in case the procedure is tried with other texts.

2. The symbols used correspond to Hoey's proposition of rende-
   ring all sentences more neutral, when distanced sentences are
   placed together. To do so, all local non-lexical cohesive features
   must be removed, and conjuncts and cohesive devises should
   be added in some cases to maintain the presentational choices
   of the author of the analysed text. // represent those items pre-
   sent in the original text which establish cohesive relations with
   adjacent sentences. [ ] represent metalanguage which has been
   added to make the relations created among non-adjacent sen-
   tences which are placed consecutively coherent.

## References

ARCAY HANDS, E. 1995. **Patrones lexicales en textos académicos estudiantiles escritos en español de Venezuela** Unpublished raw data.

ARCAY HANDS, E. 1995/6. **Los patrones lexicales de M. Hoey en textos escritos en español.** Manuscript accepted for publication at the **Revista Venezolana de Lingüística Aplicada.**

ARCAY HANDS, E. 1996. **Análisis multidimensional de textos académicos escritos en español de Venezuela.** Unpublished master's thesis, Universidad Simón Bolívar, Caracas (Venezuela).

CUCATTO, A. 1995. La topicalización: entre la sintaxis y la pragmática discursiva. En Resúmenes del **I Coloquio Latinoamericano de Analistas de Discurso,** (pp. 43-44). Universidad Central de Venezuela, Facultad de Humanidades y Educación, Caracas (Venezuela).

ENKVIST, N.E. 1987. Text linguistics for the applier: An orientation. En U. Connor & R. B. Kaplan (Eds.), **Writing across languages: Analysis of L2 text** (pp. 73-86). Reading, Massachusetts: Addison-Wesley.

GRABE, W. 1991. Current developments in second language reading research. **Tesol Quarterly,** 25(3), 375-406.

HOEY, M. 1991. **Patterns of Lexis in Text.** Oxford: Oxford University Press.

HOEY, M. 1994. **Seminario:Discurso, léxico y lectura.** Unpublished data. Caracas: Universidad Central de Venezuela, Facultad de Humanidades y Educación, Dirección de Postgrado, Area de Lingüística.

LAUTAMATTI, L. 1987. Observations on the development of the topic of simplified discourse. En U. Connor & R. B. Kaplan (Eds.), **Writing across languages: Analysis of L2 text** (pp. 87-114). Reading, Massachusetts: Addison-Wesley.

MOSS, M.G. 1995. Análisis del lenguaje de un texto de ciencias sociales y su influencia en los procesos de aprendizaje de un grupo de adolescentes en un colegio de Barranquilla. En **Resúmenes del I Coloquio Latinoamericano de Analistas de Discurso,** (pp. 76-77). Universidad Central de Venezuela, Facultad de Humanidades y Educación, Caracas (Venezuela).

NATTINGER, J. R. & Decarrico, J.S. 1992. **Lexical phrases and language teaching**. Oxford: Oxford University Press.

PÁEZ URDANETA, I. 1985. **La enseñanza de la lengua materna: Hacia un programa comunicacional integral.** IUPC, Caracas (Venezuela).

PERKINS, K. 1992. **The effect of passage topical structure types on ESL reading comprehension difficulty** (pp. 163-172). London: Edward Arnold.

SINCLAIR, J. 1991. **Corpus, Concordance, Collocation.** Oxford: Oxford University Press.

SWALES, J.M. 1990. **Genre analysis: English in academic and research settings.** Cambridge: Cambridge University Press.

VACCA, R. 1981. **Content area reading.** Boston, Massachussets: Little, Brown, & Co.

WILLIS, D. 1993. **The lexical syllabus.** London: Harper Collins.

YOUMANS, G. 1991. A new tool for discourse analysis: The vocabulary-management profile. **Language,** 67(4), 763-789.

## Appendix

*A. Text E01 without marginal sentences:*

**02 01** Entre los roles que desempeñamos en el recinto universitario está el de investigador. **02 02** Este rol va más allá de los galpones de Educación cuando hallamos escaso material de apoyo, falta de recursos y pocos "focos" de investigación con una línea propia. **02 03** Salimos entonces de Carabobo tanto para cumplir con el curriculum de las menciones como para conocer un poco más sobre la problemática que nos hemos planteado estudiar. **02 04** Pero ¿por qué salir de Carabobo cuando aquí también se genera investigación y muchos profesores procuran que los diseños repliquen otros, den nuevas respuestas, apliquen soluciones?

**02 06** En el área de inglés la investigación es capital. **02 07** Como enseñamos esta lengua es producto de otros estudios foráneos que, bajo ninguna circunstancia, parte de nuestra realidad única y diferente. **02 08** Hay obligatoriedad y necesidad de estudiar la enseñanza del inglés bajo el calor de Carabobo, en los barrios marginales, y con todo el peso de

nuestro devenir socio-económico. **02 09** Lo que tomamos prestado de otros contextos nos ayuda, pero no señala soluciones definitivas y a largo plazo. **02 10** Contribuye, y plantea nuevas inquietudes que, en la mayoría de los casos, no trascienden en su problemática.

**02 11** Por ser en este campo de Lengua Segunda o Extranjera donde nos desenvolvemos tal vez pensemos que es imperativa la investigación. **02 13** Si bien necesitamos conocer cómo, por qué y qué caracteriza el proceso de adquisición del inglés, también lo necesitamos saber del español, porque este proceso "natural" nos puede iluminar el camino de la adquisición de la Segunda Lengua .**02 14** Este camino del investigar obviamente le da un sentido más profundo y propio a nuestro quehacer de docentes o futuros profesionales de la lengua inglesa y, en otros, de la española.

**02 18** Pero la investigación necesita del apoyo documental, humano, material. **02 19** Sin ellos es pseudo, incompleta, intento o aproximación.

**02 20** O implica que el investigador viva en el afuera del contexto que le pertenece para llegar a las causas de la problemática.

B. *Summary of Text S03 (20% of the sentences with the highest number of bonds: sentence 05 with 5 bonds and sentence 08 with 6 bonds.)*

**12 05** De hecho de llevarse a cabo la creación de /éste/ [*un* ] Centro [*deInvestigaciones Lingüísticas en F.A.C.E* ] será con la finalidad de recopilar todo trabajo o investigación que se realicen dentro y fuera de la facultad, no sólo a nivel de trabajos de ascensos sino investigaciones provenientes de materias como es Metodología de la Investigación y luego ponerlos a la orden de quien desee continuar trabajos posteriores o investigar aún más sobre el tema.**12 08** [*Para ilustrar la conveniencia de los anterior* ]En semestres anteriores para cumplir con las exigencias de una materia se realizó una investigación en grupo sobre la mala pronunciación del hubieron y no del hubo en el habla Valenciano, la dificultad de la investigación tuvo en no tener un lugar que nos facilitara alguna información de interés, es decir, antecedentes o trabajos previos de los cuales pudiéramos tener alguna Orientación que no fuese la utilización de algunos recursos como entrevistas, encuestas y grabaciones.