The impact of Focus on Form on the acquisition of passive voice by advanced learners[1]: a quantitative study

Introduction

This paper presents the statistical results obtained from a quantitative study into the effects of 10 hours of instruction on two groups (one experimental and one control) of 52 university learners working with similar materials designed to practice, through differential tasks, the passive voice at an advanced level.

1. Theoretical Background

In the last two decades of second and foreign language acquisition research the study of different types of instruction and its impact has gained more and more importance. Within these a significant portion has been devoted to the effects on learning of Focus on Form (Fonf hereafter), proposed by Long and Robinson (1998). Authors like Long (1991; 2000), Doughty and Williams (1998) and Doughty (2001) claim and have sought to prove that, when formal instruction directs learners' attention to certain grammatical forms presented in context, their learning becomes more effective.

Although the literature does not seem to agree as to what exactly is implied by 'attention', many definitions have been put forward in the field of second language acquisition. Schmidt (2001) in Dörnyei (2009) refers to it as the cognitive mechanism that controls access to consciousness. Dörnyei borrows Baars's (1997b:303) theatre metaphor, and describes attention as the spotlight, whose relevance is crucial since "only events in the bright spot on stage are strictly conscious". Besides, through our attention, we can selectively concentrate on one issue while ignoring other things. In an environment of multiple stimuli, we can direct the spotlight of our attention at specific things. There are even times when the spotlight can even slip out of our hands and we can catch ourselves "mind-wondering". Ortega (2009) highlights the intrinsic link between attention and cognition in that the former "heightens the activation level of input in working memory, allowing it to remain there for longer through rehearsal and thus making it available for further processing and for entering long-term memory" (p. 93).

Within the framework of Fonf, the role that learner attention plays at the moment of processing the grammatical input during "otherwise meaning-oriented

[1] This research corresponds to part of the study carried out by the research team J017, directed by Mgtr. Adriana Álvarez and subsidized by the Secretaría de Investigación, Universidad Nacional del Comahue.

communication" (Izumi and Bigelow, 2000:242) is essential to implicitly establish a link between the form itself and its meaning in that context of use, which will eventually affect students' production as "they learn form, meaning and function of language in an integrated manner" (Izumi and Bigelow, op. cit.) in a context where, according to Robinson (2003), "unguided incidental learning is slow and inefficient (Long 1996), or just not possible for learnability reasons (White 1991).

In Doughty and Williams' words,

> **"...focus on form entails a prerequisite engagement in meaning before attention to linguistic features can be expected to be effective. [...]**The learner's attention is drawn precisely to a linguistic feature as necessitated by a communicative demand. [...] Most important, it should be kept in mind that the fundamental assumption of focus-on-form instruction is that meaning and use must already be evident to the learner at the time that attention is drawn to the linguistic apparatus needed to get the meaning across" (Doughty and Williams, 1998:3)

Even the most fervent proponents of focus on form believe that research on this topic does not necessarily result in the confirmation of hypotheses concerning increased acquisition of the L2. Take for instance what Long and Robinson state,

> "effects for instruction of any kind may be, and almost always are, gradual and cumulative rather than instantaneous and categorical, and they draw on memory for noticed features at subsequent points in development as learners process linguistic material at higher levels. Consequently, **effects of instructions attributable to noticing may not be immediate and may result from the delayed interaction of materials noticed and available for recall (rather than simply detected), with developmental processes occurring along other dimensions of language ability than those specifically targeted by the instructional treatment**." (Long and Robinson, 1998:40)

In an attempt to take into account this plausibly long-term and cumulative effect of instruction, the design of this study has included a Posttest to be administered almost half a year after the treatment phases are over (Posttest 3), taking into account these same authors' advice that "Ideally, future studies should allow for **longer periods of exposure than** has often been the case to date" (op.cit.41; emphasis added).

However, Robinson himself wonders on how to control extraneous variables that can affect the research conditions when doing classroom-based research, as is the case in question,

> "Given the nature of such research [conducted under classroom conditions], it is unclear what learning processes result form instruction and what the significant interactions of learning conditions, as distinct from instructional treatment and the formal characteristics of language are."(Robinson, 1996:51)

Moreover, there are other authors who discredit focus-on-form research and its findings so far. This is the case of Krashen who states that,

> "Lightbown (1991) is the only other study I know of that finds a long-term effect for focusing on form. [...] Lightbown's data also bears on the question of whether form focus is necessary. Although the other classes she studied were not as accurate in the *there is* form, Lightbown notes that in these classes, **students who were the most accurate were those who were in general the most advanced in English**." (Krashen, 1993:723) (emphasis added)

Truscott is particularly critical of all the research carried out with respect to focus on form, and is sceptical of the findings in this field. He claims that,

> "of the studies that reported beneficial effects for form-focused instruction, very many did not use any follow-up testing. Many others did such testing, but after a delay of five weeks or less (often much less). None of these studies provides evidence for form-focused instruction:" (Truscott, 1998: 119)

The effectiveness of this type of instruction is also doubted in Sheen (2002, 2003), where the author criticizes the value that has been attached to the benefits research claims that Fonf brings about.

In a statistical meta-analysis carried out by Norris and Ortega (2000), they investigated the literature published between 1980 and 1998 to determine overall effectiveness of different forms of L2 instruction, among which they researched Fonf with input enhancement and input flood. Their results, also questioned in Sheen (2002, 2003), point to an important effect of Fonf especially when delayed post-tests were included: "At delayed post-testing (in studies where this was carried out), instructed groups both maintained a modest advantage in gains over control/comparison groups and were more homogeneous" (Doughty, 2003: 268).

In the research design that follows, many of the issues raised in this section have been taken into account.

2. The Study

2.1.The Form In Focus

The chosen form to be studied is the passive voice. For the selection of candidates for effective focus on form, Harley (1993), in Williams and Evans (1998), refers to forms that

1. Differ in nonobvious ways from the learners' first language

2. Are not salient because they are irregular or infrequent in the input
3. Are not important for successful communication
4. Are likely to be misinterpreted or misanalysed by learners (op.cit. p. 140).

The passive voice seems to be an instance of 4., where learners misinterpret or misanalyse especially the formal features that are necessary to accurately build this form. Phenomena like the unaccusative passive (**The accident was happened*), wrong participles, incorrect verb forms and wrong prepositions (*for* instead of *by*) have been recorded among EFL learners of different linguistic backgrounds (Oshita, 2001), which suggests that these difficulties are inherent in the acquisition of the passive voice in EFL adult learners.

Doughty and Williams (1998) mention a number of factors underpinning Fonf design: the nature of the rule, its frequency in the input and the relation it bears to the L1, and also the status it has in learners' interlanguage (IL). With respect to its frequency, the passive voice is widely used and appears very often in most of the authentic texts in the foreign language learners are exposed to – especially since they are from the printed press-, and it is very similar in use to learners' L1, Spanish, although in their native language, it may not be as frequent as in English. The components that make up the passive form in Spanish present some differences when compared to the English form. A source of trouble could very well reside in the translation of the English verb *to be* (the auxiliary participating in the passive), which in Spanish can be expressed by two forms, *ser* and *estar*. Only the former can be used as auxiliary verb in the passive form in their L1.

In these learners' interlanguage, the form in focus appears as 'errorred' as evinced in the following IL productions by English IV students,

a) …controversies related to how they are bring up…

b) Another important thing to be mention…

c) Aspects which are taking into account…

d) Parents should be cared about how they raise their children.

The mistakes transcribed above and results from the research carried out in a similar context by Alvarez et al. (2009) show that advanced EFL learners in a university context are ready to notice holes in their production of this particular form (Williams and Evans, 1998; Lightbown, 1998; Doughty and Williams, 1998). They have attempted to use it, and almost do so successfully but for these mistakes. Lightbown (1998) suggests that,

> "…recent classroom-based research has yielded evidence that learners not only benefit from, but may sometimes require, focus on form to overcome incorrect or incomplete knowledge of specific target language features (e.g. Lightbown & Pienemann, 1993; Lyster, 1994a; Swain, 1991b; L. White, Spada, Lightbown and Ranta, 1991)" (p. 179)

All the arguments presented above support the choice of the passive as a form worthy of study through a Focus-on-Form approach.

2.2. Research questions and research hypotheses

Various pieces of SLA research in the literature of Focus on Form (Robinson, 1995 and Truscott, 1998) have attempted to show how different types of input form-focusing tasks result in better acquisition of the foreign language. Some other researchers, such as Izumi (1999), Izumi, Bigelow, Fujiwara and Fearnow (1999) and Izumi (2000), have tried to demonstrate that a combination of output, that is, having learners produce the target form, and Fonf treatment on it can lead to improved production of the L2.

This study has combined features of these two branches of focus-on-form research, for it explores whether attention to formal features in the input, and the need to use these formal features communicatively can help learners move forward in their IL. The following research question guided the investigation:

- **Do input-noticing, combined with output activities, result in improved production of the target form?**

The noticing activities mentioned in the question above refer to input manipulation (input flood and input enhancement), whereas output activities consist of tasks in which learners are encouraged to produce the target language in obligatory contexts. According to Swain (1993, 1995),

> in producing the target language (vocally or subvocally) learners may notice a gap between what they *want* to say and what they *can* say, leading them to recognize what they do not know, or know only partially. In other words, under some circumstances, the activity of producing the target language may prompt second language learners to consciously recognize some of their linguistic problems; it may bring to their attention something they need to discover about their L2. (pp. 125-126)

What is proposed in the Noticing Hypothesis by Schmidt (1994) is that learners' production will facilitate later noticing of formal features in the input they receive. This noticing will help learners move forward in their interlanguage.

Thus, the following hypotheses have been generated:

Hypothesis 1. After being exposed to the treatment and production in Phase 1, the experimental group will not immediately show greater gains in accuracy of their use of the target form than the control group.

Hypothesis 2. After being exposed to the treatment and production in Phase 2, the experimental group will show greater gains in accuracy in the use of the target form than the control group in posttest 2, one month after PrT.

Hypothesis 3 There will be no difference in gains in accuracy of use of the target form between the experimental group and the control group in post test 3, 2 months after the pretest.

Hypothesis 1 and 2 stem from results obtained in several studies of form-focused instruction (White, 1998) in which it was found that, for there to be a significant difference in gains between initial baseline and post-treatment, a considerable amount of time needed to have elapsed. In White's words,

> Findings from previous SLA research carried out in instructional contexts indicate that follow-up posttests sometimes portray a different picture for the effects of instruction than do immediate posttests. (…) although it is difficult to specify the amount of time that should elapse between the immediate and delayed posttests, a minumum of a month would seem to be both reasonable and practical, given the constraints of school-based research. (White, 1998: 92)

Hypothesis 3 has its roots mainly in the literature of studies carried out to assess the effect of focus-on-form instruction on learners' acquisition of the target language (Izumi and Bigelow, 2000).

2.3. Participants

The English IV course at Universidad Nacional del Comahue is traditionally taught to two simultaneous modules due to the large number of students taking it. One of these was haphazardly labelled as the experimental group (EG) and the other one as the control group (CG). A total of thirty-nine teachers- and translators-to-be participated in this study, of which nine were men and thirty, women. Their ages ranged between 20 and 27. EG was made up of twenty-one students while CG consisted of eighteen students[2]. All participants' level of proficiency in the foreign language was rated by their teachers as advanced.

[2] The number of participants in each group differs due to the fact that throughout the four weeks that the experience lasted learners dropped out of the course, or simply did not complete all the tasks assigned.

2.4. Experimental design

This study focused on the passive voice due to the fact that, even though the participants were advanced learners, their written assignments evinced an important number of problems in the use of this structure. Phenomena like the unaccusative passive (**The accident was happened*), wrong participles, incorrect verb forms and wrong prepositions (*for* instead of *by*) have not only been found in our learners' essays but have also been recorded among EFL learners of different linguistic backgrounds (Oshita, 2001), which suggests that these difficulties are inherent in the acquisition of the passive voice in EFL adult learners.

The research plan for both groups (EG and CG) is detailed in Table 1 below:

	Activities and data-collection	
Day 1	Pretest (1 hour)	
Day 2-7	**Phase I** (5 hours)	• Pre-reading tasks • Reading tasks • Dictogloss 1
Day 8	Posttest 1 (1 hour)	
Day 9-15	**Phase II** (4 hours)	• Post-reading Tasks 3.1 Debate, 3.2 Whole-class discussion and 3.3 Essay Writing • Dictogloss 2
Day 22	Posttest 2 (1 hour)	
Day 60	Posttest 3 (1 hour)	

Table 1. Treatment plan for both EG and CG

The study was divided into two Phases, after which a series of posttests was administered to explore the relative gains obtained by the different groups at the different stages in the study. Both phases were organised into the different stages in a reading lesson as proposed by Wallace (1992) and consisted of communicative activities thematically related to the main text in the unit "Argentina's President in Crisis Talks".

2.4.1. Treatment

The treatment in the study was proactive in the sense that the form to which learners' attention was to be drawn was chosen beforehand, motivated by learners' errors because, as Doughty and Williams state,

> Learner error or difficulty comprises the basis of most proactive remedial approaches to instruction (R. Ellis, 1993) and has high face validity, in that teachers are closely connected to instructional decisions. (Doughty and Williams, 1998: 212)

Materials in this experiment were designed with this focus in mind. Focus on form was incidental, with no explicit rule-teaching, and instruction attempted to be unobtrusive.

Most tasks planned in the materials required pair or group work, which encouraged communication between learners, i.e. the type of interaction that triggers negotiation of meaning. Following Long (1996) in the Interaction Hypothesis, "conversation is not only a medium of practice, but also the means by which learning takes place" (Gass, 2003:234).

While getting learners to use the target language to put across their ideas and their answers to the questions in these materials, one of the main conditions for Focus on Form is achieved in the sense that their attention is drawn to form "during otherwise meaning-oriented communication so that they can learn the form, meaning, and function of language in an integrated manner" (Izumi and Bigelow, 2000:242).

We favoured an approach where no rule-seeking was involved on the part of learners, given that they have been taught, and have practised the passive voice in depth in their previous Language and Syntax courses. Instead, learners were implicitly encouraged to notice the passive voice. As a result, the teaching material was designed to get learners to perceive the interaction of both form and meaning in naturally occurring instances of the passive voice in authentic texts from the press on the Argentinean economic crisis of 2001.

The differential treatment, evident in the materials for EG and CG, is shown in Table 2:

	CONTROL GROUP	EXPERIMENTAL GROUP
	----	Input flood
Reading texts in materials	----	Typographical input enhancement
Text-reconstruction	Dictogloss 1* Dictogloss 2*	Dictogloss 1 Dictogloss 2

Table 2: Differential Treatment for EG and CG

* The grammar points to be highlighted in these grammar dictations were the Present and Past Perfect Tenses.

2.4.2.Techniques and differential materials

Fonf resorts to some input-manipulating techniques like *input flood* and *input enhancement* to focalize learners' attention in an implicit way. The former, utilised by Williams and Evans (1998) and White (1998) in their research, refers to exposing the learner to a massive amount of discourse (be it written or spoken) where the selected form appears frequently providing as many opportunities as

possible for learners to encounter the form in focus in the input. In Doughty and Williams' words,

> "…the principle is simply that the more opportunities there are in the input for learners to notice a linguistic feature, the more likely they are to do so". (Doughty and Williams, 1998: 236)

In this particular study, input flooding has been provided through the number of instances present in the main text, together with those in the True-or-False statements, in the dictoglosses and in the quotations for discussion in Post-reading task 3.1. Contrarily, the Control Group is only exposed to the instances present in the main text of this lesson.

As regards the second technique in this experiment, Izumi and Bigelow (2000) define input enhancement as "the external manipulation of the input to encourage better intake of various features of that input" (p.242), which implies visually highlighting that form in the texts by means of colours, underlining, larger font size, bolding, and italics among other techniques. This manipulation is meant to "increase the likelihood that they will be noticed" (Doughty and Williams, 1998: 236) and eventually can create the conditions that will help to advance learners' knowledge of the target language by stimulating internal mechanisms. In this study input has been made salient through larger font size and bolding in every instance of EG materials.

EG worked with materials selected to provide input flooded with multiple instances of the passive voice, each of which was typographically modified (bold type). CG worked with similar materials and tasks, though their texts were not typographically enhanced. Each set of materials centred around three news articles. One article was the basis for a reading comprehension lesson. The other two were selected for text-reconstructions or grammar-dictation tasks (*dictogloss* in Wajnryb, 1990). CG's grammar-dictation tasks were designed around two short texts including a great number of instances of the present and past perfect tenses. This guaranteed that the exposure these students received did not bias their production in favour of the correct use of the passive voice. All three texts in EG materials were flooded with occurrences of the passive voice.

2.4.3. Activities

In **Phase I**, the pre-reading and reading tasks were completed, together with Dictogloss 1.

To activate background knowledge on Argentina's economic crisis in 2001, a set of general questions was posed for learners to discuss in groups. After a first quick reading students were asked to insert three subheadings in the corresponding

place in the text. The following activity explored learners' general understanding of the text through five questions. This was accompanied by a True-or-false exercise where learners were encouraged to rephrase false statements to make them true. A final task was administered to foster deeper understanding through paraphrasing certain expressions from the text. To end this first phase students were asked to do Dictogloss 1, an output task devised by Wajnryb (1990) which, according to Swain,

> … encourages students to reflect on their own output. With this procedure a short, dense text is read to the learners at normal speed; while it is being read, students jot down familiar words and phrases; then the learners work together in small groups to reconstruct the text from their shared resources; the final versions are then analysed and compared. The initial text, either an authentic or a constructed one, is intended to provide practice in the use of particular grammatical constructions. (Swain, 1998:70)

By means of this task, learners are expected to be "forced" to use the target form, and notice a hole in their output, that is, a deficiency in their IL as regards the use and construction of the target form studied. Besides, Doughty and Williams state that "collaborative interaction during dictogloss engages (learners) in the kind of syntactic processing that Swain claims is necessary in order to go beyond communicative effectiveness to a fully accurate L2." (Doughty and Williams, 1998: 239). For this task, both groups (EG and CG) worked with different sections of "Argentina's Cash Crisis Deepens"[3], focussing on the differential use of tenses and active and passive voice in the text. Learners in CG were not presented with passive voice forms, while those in EG were exposed to paragraphs containing this structure, coherently arranged in the shape of a new text.

Phase II of the treatment continued with the post-reading activities, and later, with Dictogloss 2 "Argentine Presidential Poll Next March" following the same procedure as for Dictogloss 1. After the second instance of a grammar dictation task, learners were engaged in a debate of their interpretations of a series of controversial quotes from the British press:

a) *"While stressing that "we need the support of IMF, World Bank and Inter-American Development Bank", Mr Duhalde also insisted that "our crisis is **home-grown**, made in Argentina, by Argentines. Our solutions will also be **home-grown** - **made** in Argentina, by Argentines," he said." (The Guardian, April 22, 2002)*

b) *""The government's brought in the restrictions to prevent a run on the banks. Overnight it created mismatch [on the banks' balance sheets]. The*

[3] Uki Goni in Buenos Aires, Guardian, Thursday July 4, 2002, Guardian Unlimited (c) Guardian Newspapers Limited 2002

*corralito **was introduced** by legislation. It is not something the banks have decided. But public opinion is that the banks are stealing depositors' money," said one banker." (The Guardian, May 8, 2002)*

c) *"Even if Argentines do attempt to mend the economy, and fences with the banking sector, the fall-out will still be considerable. "The anger of the banks' customers is doing terrible damage," said one banker. "Even if the situations **are resolved**... the impact on customer loyalty of these restrictions is going to be very considerable. It is going to take a long time to recover." Assuming the economy recovers too." (The Guardian, May 8, 2002)*

d) *"Argentina's long-term economic progress will depend on generating a new ethos of competitiveness built on research and development and greater investments in education. But between here and there lies and abyss that must **be crossed**." (Jeffrey Sachs, Time, July 23, 2001)*

Learners were expected to work in groups to decode the meaning and intention of each quotation and then venture their opinions on what the authors claimed. As a follow-up, learners then worked in plenary exchanging views on their interpretations and reactions to the quotations discussed. Learners were then presented with a set of questions for them to explore issues regarding their experiences during that and other similar crises and their identity as Argentineans. This discussion was also meant to brainstorm ideas to be included or developed in the subsequent writing task where, in a three-hour lesson, participants wrote an essay of 300-350 words on one of the four topics suggested.

2.4.4. Data collection

The research design tapped into learners' knowledge of the passive voice both in written and oral production and in guided and in less controlled contexts, as well as in de-contextualised discrete-point tests.

Learners' written production was measured through a) the reconstructed versions of two texts (Dictogloss 1 and 2); b) two picture-cued paragraphs (pre- and post- intervention); and c) an essay.

Learners' oral production, on the other hand, was assessed in two consensus tasks (Swain and Lapkin, 2001, in Mackey and Gass, 2005), the discussions of which were recorded in digital format (MP3); and in a debate of four controversial quotations from the British press.

For reasons of length we will only present the data collections measures used in the pre and posttests, namely the multiple choice tests. Learners signed an authorization allowing all the data obtained through these means to be used for research purposes.

2.4.4.1. Multiple-choice tests

Multiple choice tests are experimental, discrete-point measures "employed under controlled conditions with less communicatively driven and decontextualised constraints" (Chaudron, 2003:784). They have typically been used "to assess whether learners have knowledge of particular, structural elements of the language" (Read, 2000:3). In this particular case, to gather data on learners' structural knowledge of the passive voice, four multiple-choice tests were designed and administered at different times during the study:

1) a pre-test (PrT) previous to the intervention,
2) a posttest 1 (PT1) at the end of Phase I (see design above),
3) a posttest 2 (PT2), one week after the experience was over, and
4) a posttest 3 (PT3) two months after PrT.

Each multiple choice test consisted of twenty isolated gapped sentences for which three options were provided (A, B and C). Ten of these items were created to study the passive voice while the other ten functioned as distractors containing other tenses whose components could participate in the formation of both perfect and continuous tenses. These forms were included in the same test due to the fact that they might appear as structurally similar to learners even if they are quite different in terms of the meanings each form can express. Because of this apparent resemblance, even at an advanced stage in their acquisition process, EFL learners can be misled into choosing an incorrect option.

The ten target-form items were distributed haphazardly across the tests (as were the options provided for each of the items). Learners were not told what the tests measured. Each of the prompts was a reduced sentence taken from different articles published in the British press on the issue of Argentina's economic crisis in 2001, which contained either of the forms tested (passive voice or perfect tenses). These sentences were modified where necessary, as will be explained in the following section.

2.4.5. Piloting

The pretest and the three post tests were piloted with thirteen learners of a similar level to that of the participants in the study. The tests were administered at exactly the same intervals as those planned for in the original design.

After the piloting stage and the analysis of the data gathered, all those items across the pretest and posttests that had systematically obtained numerous unexpected answers were modified and retested for validity reasons so as to ensure each of the test items was clear and measured passive and verb tense forms. This

type of modification was implemented on a total of ten sentences. Some examples are presented below.

In PrT, for item 11, half the learners tested chose C, when in fact the correct option should have been A. This result could have originated in insufficient knowledge of the phrasal verb *to be brought in* or in inaccurate chunking, probably due to the same reason.

(a) 11. Cavallo _____ again brought in to head the economy team.
A. was B. had C. has

After piloting, the *to-infinitive* purpose clause in the original sentence was expanded to contain the connector *in order to*, and thus aid learners in the chunking of the sentence:

(b) 11. Cavallo _____again brought in in order to head the economy team.

In the case of item 14, learners were systematically drawn to option B, probably out of ignorance of the pattern of the verb *to default*.

(c)14. Other countries _____ on their lending to the International Monetary Fund and the World Bank.
A. have defaulted B. are defaulted C. have defaulting

The vocabulary item appearing in the original sentence (*defaulted*) was replaced by a more frequent verb (*collapsed*).

Similar procedures were followed with each of the other eight sentences.

3. Results

3.1. Quantitative data analysis of the multiple-choice tests

The analysis was carried out through an entirely randomized factorial design with the following factors: METHOD of instruction (Fonf, non-Fonf), ITEMTYPE (target and distractor) and TIME of data collection (0, 1, 2 and 3, corresponding to Pretest, Posttest1, Posttest 2 and Posttest 3 respectively).

In Table 3, significant effects and interactions (0,05>p>0,01) between ITEMTYPE and TIME have been highlighted in light grey, and very significant effects among METHOD, ITEMTYPE and TIME (of p value lower than 0,01) appear in a darker shade of grey.

F.V.	SC	gl	CM	F	p-valor
Modelo	10,04	15	0,67	3,58	<0,0001
METHOD	2,00	1	2,00	10,72	0,0012
ITEMTYPE	0,17	1	0,17	0,92	0,3383
TIME	1,59	3	0,53	2,83	0,0386
METHOD*ITEMTYPE	0,09	1	0,09	0,49	0,4848
METHOD*TIME	0,85	3	0,28	1,52	0,2102
ITEMTYPE*TIME	3,67	3	1,22	6,56	0,0003
METHOD*ITEMTYPE*TIME	1,54	3	0,51	2,75	0,0428
Error	57,16	306	0,19		
Total	67,20	321			

Table 3.: Variance analysis (SC type III)

To explore the presence of double and triple interactions among variables, multiple comparisons of mean scores were established. A significant interaction among METHOD*ITEMTYPE*TIME can be observed and a very significant effect can be found between ITEMTYPE*TIME. A new analysis was carried out per time.

F.V.	SC	gl	CM	F	p-valor
Modelo	6,61	3	2,20	6,19	0,0007
ITEMTYPE	2,98	1	2,98	8,39	0,0047
METHOD	3,15	1	3,15	8,85	0,0037
ITEMTYPE*METHOD	0,63	1	0,63	1,77	0,1864
Error	34,85	98	0,36		
Total	41,46	101			

Table 4.: Variance analysis (SC type III)

METHOD	Medias	n	E.E.		
CG	9,33	48	0,09	A	
EG	9,69	54	0,08		B

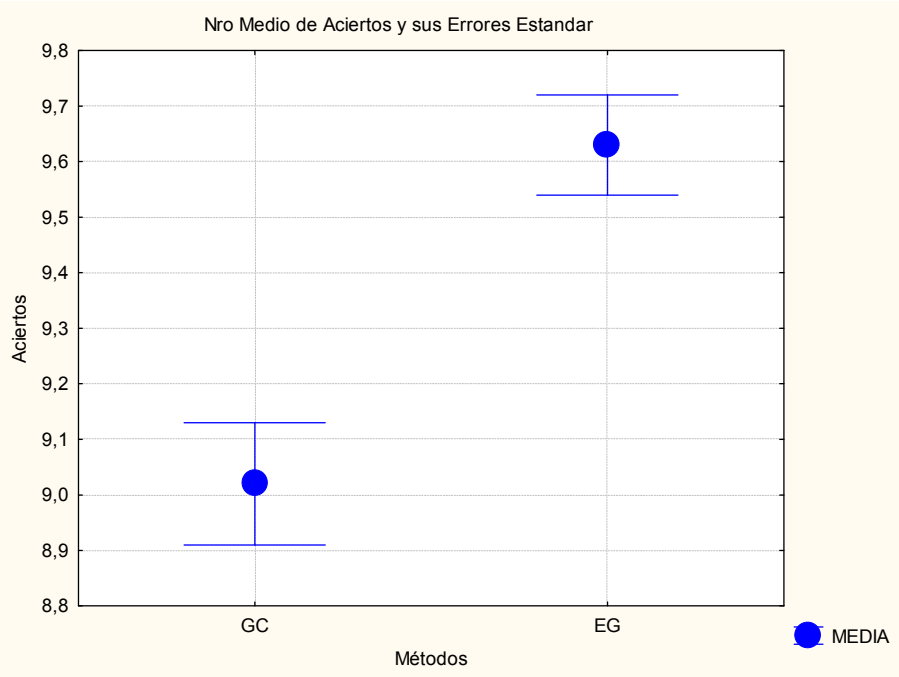
Table 5.: Duncan Test:; Alfa=0,05; *Error: 0,3556 gl: 98*
(Different letters indicate significant differences (p<= 0,05))

Tables 4 and 5 show that there is a significant effect of method of instruction between CG and EG. Performance of CG and EG at time 0 is different from the outset. It is, therefore, necessary to include the differential response as a co-variable in the analysis. The results of this analysis are presented in Table 6:

METHOD	Medias	n	E.E.		
CG	9,10	97	0,09	A	
EG	9,60	122	0,08		B

Table 6.: Duncan Test; Alfa=0,05; *Error: 0,5926 gl: 188*
(Different letters indicate significant differences (p<= 0,05))

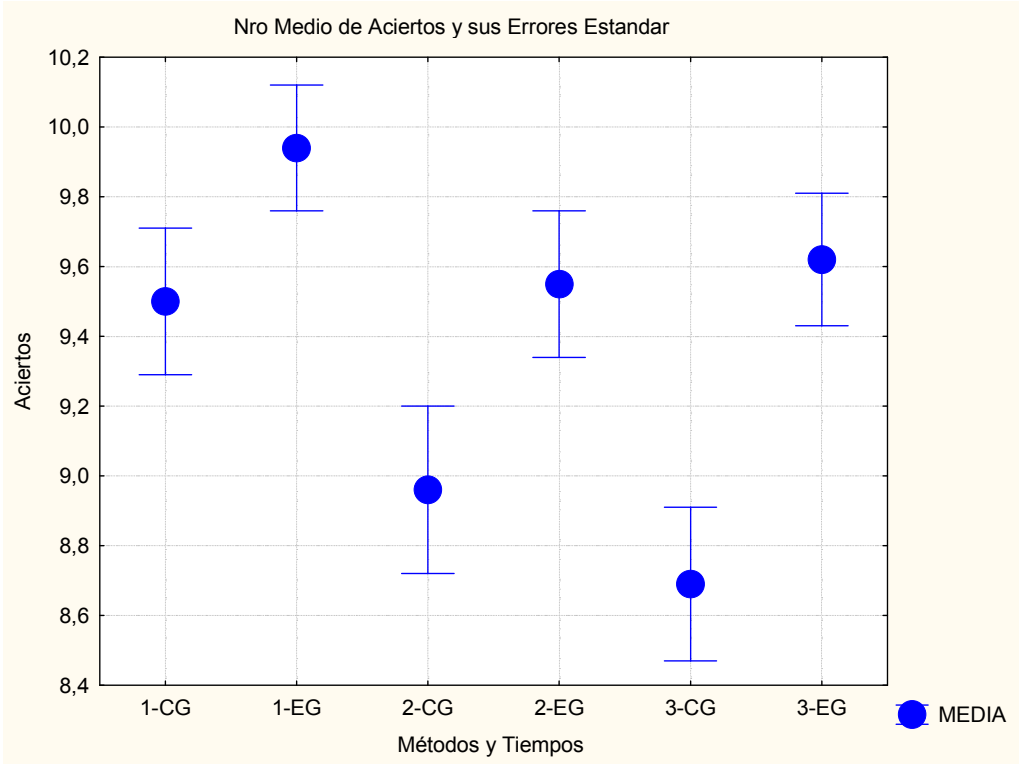
Although the differential response at Time 0 is not significant, this is maintained in the model. As evinced in Graph 1, there is a very significant effect of method of instruction, where EG outperforms CG, irrespective of time.



Graph 1. Box plot of mean scores per METHOD

From Graph 1, it becomes obvious that the intervention had a very significant impact on EG's test performance overall. Even though EG was better than CG at the outset of the experiment, the gains obtained by this group have surpassed this initial advantage.

In Graph 2 below, each group's mean scores have been represented across time. It can be seen that the difference between EG and CG present in Graph 1 remains constant at all times in Graph 2. CG's performance evinces a continuous drop in scores, indicating that time caused attrition in the use of the target form. CG obtained its highest score at Time 1, immediately after working with the same amount of input as EG, even if unenhanced. After this initial result, CG scores decrease at Time 2 and also at Time 3. In the second phase of the treatment, once the input becomes differential for EG and CG, CG's performance begins to fall.



Graph 2

EG's overall scores are significantly higher at all times than those of CG. Although there is a drop in scores for posttest 2 at the end of Phase 1, a slight improvement is observed at Time 3, which indicates that EG constantly and significantly differs in scores from CG.

4. Discussion

In the following section we will analyse the results in terms of the hypotheses formulated.

Hypothesis 1. After being exposed to the treatment and production in Phase 1, the experimental group will not immediately show greater gains in accuracy of their use of the target form than the control group.

The results obtained and presented in Graph 2 show that, in fact, there was a very significant difference in gains between EG and CG as a result of the intervention. From an early stage in the experiment, EG performed better than CG. Even when the initial difference (at Time 0) in favour of EG was cancelled out through statistical procedures, EG continued to perform better. Hypothesis 1 is therefore refuted, for there is statistical evidence of significantly higher scores for EG at Time 1.

Hypothesis 2. After being exposed to the treatment and production in Phase 2, the experimental group will show greater gains in accuracy in the use of the target form than the control group in Posttest 2, one month after PrT.

In Graph 2, it can be observed that, for Posttest 2, there were still significant gains for EG over those obtained by CG, so Hypothesis 2 is confirmed. An impact of the intervention can be shown to continue to exist over time for EG.

Hypothesis 3. There will be no difference in gains in accuracy of use of the target form between the experimental group and the control group in post test 3, 2 months after the pretest.

Based on the findings of research in the literature, Hypothesis 3 predicted that there would be no improvement in accuracy as regards the use of the target form between the experimental group and the control group in Posttest 3, 2 months after the pretest, because the literature offers proof of a dwindling effect for experimental groups over long periods of time (longer than 2 weeks post intervention). However, gains for the EG at Time 3 in our study were higher than those obtained at Time 2. More relevantly, the distance in mean scores between EG and CG for this same testing instance was even larger than that between these two groups for Posttest 2. This evidence disproves Hypothesis 3.

Some researchers argue that such a difference could be due to what some studies refer to as the "test effect" (Baghaei and Amrahi, 2011; Roediger and Marsh, 2005), according to which learners get used to the test type and perform better only through resulting familiarity with the type of test they are given repetitively. In favour of a real difference between posttest 2 and posttest 3, however, we can argue that the third posttest took place much later and the distance between this test and its predecessor was much longer in time, than between each pair of testing instances during all the intervention. The time span between Pt2 and Pt3 was close to a whole month, whereas all the other pairs were carried out no more than fifteen days apart.

A significant effect can still be perceived late in the data collection, perhaps because the combination of input manipulation and output tasks, which mainly took place in Phase II, prompted participants to be readier to notice the form in focus even after the intervention was over (Doughty, 2003 on Norris and Ortega, 2000). This might have enabled the restructuring process in their interlanguage to continue almost unconsciously. Input manipulation on its own does not necessarily lead to improved performance but probably having to produce language through output tasks has contributed to EG's differential results.

Participants in this group were encouraged by the method of instruction to notice the form in focus in its contexts and were readier to restructure their interlanguage. There is substantial evidence of this restructuring process in the clearly higher test scores for EG.

5. Conclusion

The sections above have presented the design and results of a study carried out in an EFL university context with two groups of English IV students. The aim of this study was to assess the effects of Focus on Form on these learners' interlanguage. As stated in the results and discussion, two of the hypotheses were not borne out by the data obtained through successive multiple choice tests and Hypothesis 2, which predicted a decline in EG's performance at PT2, was confirmed. EG performed consistently better at the multiple choice tests, independently of testing time. The variable Method of Instruction appears to account for the significant differences between these two groups. The intervention has proved to have been beneficial for EG learners, who obtained higher scores even two months after PrT.

The experiment combined input manipulation (input enhancement and input flood) with output tasks. However, the experimental plan was not designed to allow researchers to define which of the techniques employed (input flood, input enhancement, or text-reconstruction) was responsible for the increase in EG learners' accuracy in the use of the target form; nor does it allow us to establish whether these results have been brought about by the combination of all these techniques.

The main objective, increasing learners' accuracy of the use of this form, has been achieved. Participants in EG were positively influenced by the intervention and their gains were not subject to attrition, as most studies in the literature seem to suggest. In fact, Doughty reviews research where "the delayed post-test interval is typically quite short (four weeks on average)" and states that "it might be expected that the effects of instruction demonstrated would not remain after a longer period of time, either because control subjects have caught up (…) or because the particular type of instruction favoured in this set of studies leads to the type of knowledge that is easily forgotten." (Doughty, 2003: 269) The data in this study run counter to this phenomenon.

Nonetheless, a word of caution should be issued: the results obtained in this study do not lend themselves to generalisations as regards other populations since the number of participants in this research was not representative. Similar

experiments need to be carried out with larger student groups and delayed post-tests, if possible, should take place three or even four months post-intervention.

However, given that the intervention sought to address a specific problem in the context we teach, the constant, significant difference between EG and CG indicates that the research has been beneficial in solving the problems encountered in this particular situation and context.

References

Alvarez, A., M. Zinkgraf, M.F. Casares, M.T. Araya and A. Monteserín (2009). Acquisition of passives and unaccusatives in an EFL context: a study of elicited production in Spanish-speaking students. Proceedings of Quintas Jornadas "Un espacio de Encuentro: el estado del Conocimiento en la ESI". General Roca, R. Negro.

Baars, B. (1997). *In the theater of consciousness: The workspace of the mind*. New York: Oxford University Press.

Baghaei, P. and N. Amrahi. (2011). The effects of the number of options on the psychometric characteristics of multiple choice items. *Psychological Test and Assessment Modelling, Volume 53*, (2), 192-211.

Batstone, R. 1994. *Grammar.* Oxford: OUP.

Castro, A.E., M. Zinkgraf, M. A. García Álvarez y S. Rodríguez. (2012). Cómo se corrigen los estudiantes de inglés entre sí: paráfrasis y reconstrucción colectiva del mensaje en un estudio de caso. Presentado en XIII Congreso de la Sociedad Argentina de Lingüística, San Luis, Argentina.

Chaudron, C. (2003). Data collection in SLA research. In C. Doughty y M.H. Long (eds.) *The Handbook of Second Language Acquisition* (pp.762-828). Oxford: Blackwell Publishing.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (ed.), *Cognition and Second Language Instruction* (pp. 256-310). Cambridge: Cambridge University Press.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In In C. Doughty y M.H. Long (eds.) *The Handbook of Second Language Acquisition* (pp.762-828). Oxford: Blackwell Publishing.

Doughty, C. and Williams, J. (1998a) Issues and terminology. In C. Doughty y J. Williams (eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 1-11). Cambridge: Cambridge University Press.

Doughty, C. and Williams, J. (1998b). Pedagogical choices in focus on form. In C. Doughty y J. Williams (eds.), *Focus on Form in Classroom Second Language Acquisition* (pp. 197-261). Cambridge: Cambridge University Press.

Ellis, R. 1993. The Structural Syllabus and Second language acquisition. *TESOL Quarterly, 27*(1), 91-113

Gass, S. (2003). Input and Interaction. In C. Doughty y M.H. Long (eds.) *The Handbook of Second Language Acquisition* (pp.224-255). Oxford: Blackwell Publishing.

Hall, N. and J. Shepheard. 1991. *The Anti-grammar Grammar Book*, Essex: Longman

Harley, B. (1993). Instructional Strategies and SLA in Early French Immersion. *Studies in Second Language Acquisition 15*(2): 245-259.

Izumi, S. (1999). A text reconstruction task that promotes noticing. Paper presented at the American Association for Applied Linguistics conference, Stamford, CT.

Izumi, S. y M. Bigelow. (2000). Does output promote noticing and second language acquisition? *Tesol Quarterly Vol. 34*, N°2, pp. 239-278.

Izumi, S., M. Bigelow, M. Fujiwara and S. Fearnow (1999). Testing the output hypothesis: effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition, 21,* 421-452.

Krashen, S. (1993). 'The effect of formal grammar teaching: Still peripheral'. *TESOL Quarterly, 27(4),* 722-725.

Lightbown, P.M. (1998). The importance of timing in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge: Cambridge University Press.

Lightbown, P. and M. Pienemann. (1993). 'Input, instruction and feedback in second language acquisition'. *Second Language Research, 7(2),* ii-iv.

Lightbown, P. & Spada, N. (1993). *How languages are learned*. Oxford: Oxford University Press.

Long, M. H. (1991). Instructed interlanguage development. In L. Beebe (ed.), Issues in Second Language Acquisition: Multiple perspectives (pp.115-41). Rowley, MA: Newbury House.

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie and T. Bhatia (eds.), *Handbook of Second Language Acquisition*. San Diego: Academic Press, 413-468.

Long, M. H. (2000). Focus on form in task-based language teaching. In R.L. Lambert y E. Shohamy (eds.), Language Policy and Pedagogy (pp.179-92). Amsterdam and Philadelphia: John Benjamins.

Long, M. H. and Robinson, P. (1998). Focus on Form: theory, research and practice. In C. Doughty and J. Williams (eds.), Focus on Form in Classroom Second Language Acquisition (pp.15-41). Cambridge: Cambridge University Press.

Lyster, R. (1994). The effect of functional-analytic teaching on aspects of French immersion students' sociolinguistic competence. *Applied Linguistics*, 15: 263–287.

Mackey, A. and S. Gass (2005). *Second Language Research*. New York: Routledge.

Norris, J. M. and L. Ortega (2000). Effectiveness of L2 instruction: a research synthesis and quantitative meta-analysis. *Language Learning*, 50, 417-528.

Ortega, L. (2009). *Understanding Second Language Acquisition*. London: Hodder Education.

Oshita, H. (2001). The unaccusative trap in second language acquisition. In *SSLA,* 23 (pp. 279-301). Cambridge, Cambridge University Press.

Read, J. (2000). *Assessing Vocabulary*. New York: Cambridge University Press.

Robinson, P. (1995). Attention, memory and the "noticing" hypothesis. *Language Learning, 45*, 283-331.

Robinson, P. (1996). 'Learning simple and complex second language rules under implicit, incidental, rule-search and instructed conditions'. *Studies in Second Language Acquisition, 18(1),* 27-68.

Robinson, P. (2003). Attention and memory during SLA. In C. Doughty y M.H. Long (eds.) *The Handbook of Second Language Acquisition* (pp.631-678). Oxford: Blackwell Publishing.

Roediger, H.L. and E.J. Marsh. (2005).The Positive and Negative Consequences of Multiple-Choice Testing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31 N5, p1155-1159.

Schmidt, N. (1994). Deconstructing consciousness in search of useful definition for applied linguistics. *AILA Review*, *11,* 11-26.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Sheen, R. (2002). "Focus on form" and "Focus on forms." *ELT Journal, 56*(3), 303-305.

Sheen, R. (2003). Focus on form: A myth in the making? *ELT Journal*, *57*(3), 225-33.

Swain, M. (1991). French immersion and its offshoots: Getting two for one. In B. Freed (Ed.), *Foreign language acquisition: Research and the classroom* (pp. 91-103). Lexington, MA: Heath

Swain, M. (1993). The output hypothesis: just speaking and writing aren't enough. *The Canadian Modern Language Review*, *50*, 158-164.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook and B. SEidlhofer (Eds.), *Principles and practice in applied linguistics: Studies in honour of H. G. Widdowson* (pp.125-144). Oxford: Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). Cambridge: Cambridge University Press.

Swain, M. and Lapkin, S. 2001, Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan and M. Swain (Eds.), *Researching pedagogic tasks: Second language learning and testing* (pp.99-118). Harlow, UK: Longman.

Truscott, J. (1998). Noticing in second language acquisition: A critical review. *SLA Research* 14, 103-135.

Wajnryb, R. (1990). *Grammar Dictation*. Oxford: Oxford University Press.

Wallace, C. (1992). *Reading.* Oxford: OUP.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research, 7*, 133-161.

White, J. (1998). 'Getting the learners' attention', in Doughty and Williams (eds.) *Focus on Form in Classroom Second Language Acquisition* (85-113) Cambridge: Cambridge University Press.

White, L., N.Spada, P. Lightbown, & L. Ranta. (1991). Input enhancement and L2 question formation. *Applied Linguistics 12,* 416-432*.*

Williams, J., & J. Evans. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition*. Cambridge: Cambridge University Press (pp.139-155).