The impact of Oral Pushed Output on Intermediate Students' L2 Oral Production[1]

Impacto de la Producción Oral Inducida en la Producción Oral de Estudiantes de Nivel Intermedio

Kevin López Páez[2]*

Universidad del Tolima, Colombia

Abstract

With the advent of communicative methodologies, the promise to develop both fluency and accuracy was made as a goal for teaching and learning English as an international language. However, it did not happen (Richards, 2008). In an attempt to equalize students' both semantic and syntactic competence, this study investigates the impact of Swain's (1985) oral pushed output hypothesis on EFL intermediate students' L2 oral production under a mixed method approach. The participants were 16 seventh grade EFL students from a private school in Ibagué, Colombia that were randomly assigned to an output and a non-output group. For five weeks, the output group underwent oral pushed output activities while the non-output group was merely exposed to comprehension activities. Quantitative and qualitative instruments to collect the data included pretest and posttest, audiorecordings, stimulated recalls, and interviews. Results revealed that although pushing students to produce meaningful oral output does not promote significant noticing of their linguistic problems in past narrative forms, students can modify more oral output through one-way pushed output activities than two-way activities and equalize their semantic and syntactic competence since they can engage in both processsings. Additionally, students perceived oral pushed output as an affectivity regulator in L2 oral production and as a trigger of exposure to L2 vocabulary, grammar, and pronunciation.

Key words: Pushed Output Hypothesis; Oral Production; Fluency; Accuracy; Semantic Processing; Syntactic Processing.

Resumen

Con la llegada de las metodologías comunicativas, se hizo la promesa de desarrollar la fluidez y la precisión como objetivo para la enseñanza y el aprendizaje del inglés como idioma internacional. Sin embargo, esto no ocurrió (Richards, 2008). En un intento de igualar la competencia semántica de los estudiantes con la sintáctica, este estudio investiga el impacto de la hipótesis de la producción oral inducida de Swain (1985) en la producción oral de estudiantes de inglés como lengua extranjera de nivel intermedio bajo un enfoque de métodos mezclados. Los participantes fueron 16 estudiantes de inglés de grado séptimo de un colegio privado en Ibagué, Colombia que fueron asignados aleatoriamente a un grupo de producción y a uno de no-producción. Durante cinco semanas, el grupo de producción se expuso a actividades orales de producción inducida, mientras que el grupo de no-producción fue expuesto únicamente a actividades de comprensión. Los instrumentos cuantitativos y cualitativos para recolectar la información incluyeron pruebas de entrada y salida, grabaciones de audio, recuerdos provocados y entrevistas. Los resultados revelaron que aunque inducir a los estudiantes a realizar producción oral significativa no promueve una observación significativa de sus problemas lingüísticos en formas narrativas en pasado, los estudiantes pueden modificar más producción oral a través de actividades de producción inducida no correspondidas que de actividades bilaterales e igualar sus competencias semántica y sintáctica al involucrarse en ambos procesamientos. Adicionalmente, los estudiantes percibieron la producción oral inducida como un regulador afectivo en su producción oral y como un detonador de exposición a vocabulario, gramática y pronunciación de la lengua objeto.

Palabras Clave: Hipótesis de la Producción Inducida; Producción Oral en la Segunda Lengua; Fluidez; Precisión; Procesamiento Semántico; Procesamiento Sintáctico.

Resumo

Com a chegada das metodologias comunicativas, fez-se a promessa de desenvolver a fluidez e a precisão como objetivo para o ensino e a aprendizagem de inglês como idioma internacional. Entretanto, isto não ocorreu (Richards, 2008). Este estudo pesquisa o impacto da hipótese da produção oral induzida de Swain (1985) na produção oral de estudantes de inglês como língua estrangeira de nível intermédio sob um enfoque de métodos misturados. Os participantes foram 16 estudantes de inglês de sétima série (oitava série no sistema educativo de 12 anos) de um colégio particular em Ibagué, Colômbia que foram designados aleatoriamente a um grupo de produção e a um de não-produção. Durante cinco semanas, o grupo de produção se expôs a atividades orais de produção induzida, enquanto que o grupo de não-produção foi exposto unicamente a atividades de compreensão. Os instrumentos quantitativos e qualitativos para coletar a informação incluíram provas de entrada e saída, gravações de áudio, lembranças provocadas e entrevistas. Os resultados revelaram que mesmo que induzir os estudantes a realizar produção oral significativa não promove uma observação significativa dos seus problemas linguísticos em formas narrativas em passado, os estudantes podem modificar mais produção oral a través de atividades de produção induzidas não correspondidas que de atividades bilaterais, e igualar suas competências semântica e sintática ao envolver-se em ambos os processamentos. Adicionalmente, os estudantes perceberam a produção oral induzida como um regulador afetivo na sua produção oral e como um detonador de exposição a vocabulário, gramática e pronúncia da língua objeto.

Palavras Chave: Hipótese da Produção Induzida; Produção Oral na Segunda Língua; Fluidez, Precisão; Processamento Semântico; Processamento Sintático.

Introduction

The advent of communicative methodologies for teaching a foreign language in 1970 attempted to improve the teaching practices and develop skills such as listening, reading, writing, and speaking. Methodologies such as CLIL (Content and Language Integrated Learning), CLT (Communicative Language Teaching), and TBLT (Task-Based Language Teaching) promised to develop both fluency and accuracy as goals for teaching and learning English as an international language. However, as Richards (2008) affirms, fluency-based methodologies did not generally manage to help learners develop communicative competence as well as linguistic competence on an equal basis. The issue on what to do about grammar in language learning and teaching was not resolved then. Programs with great deal of input and frequent use of authentic communication determined that learners were developing fluency at the cost of accuracy (Higgs & Clifford, 1982), showing thus good communication skills, but poor grammar proficiency together with high fossilization levels. As Gass (2003) notes, although second and foreign language English learners have been able to comprehend what they listen and read in communicative methodologies, most of them fail to produce the written or oral message they want to convey.

EFL Teaching context in Colombia

In Colombia, public and private schools as well as English institutes have witnessed issues like the above-mentioned. Since the 90s, EFL teachers' attempts to keep learners away from form-focused methodologies, influenced by Krashen's (1985) input hypothesis claims as well as the communicative competence, have led them to implement input-based and fluency-based methodologies (McDougald, 2009) such as CLIL. This methodology aims at having students learn about a subject content through a target language; that is, social studies, biology, chemistry, geometry are taught in English without putting a focus on L2 syntax (Marsh & Langé, 2000). Thus, a great usage of such methodology has drawn teachers' attention more to students' fluency and comprehension (semantic competence) than to their accuracy and production (syntactic competence). Such inequality of competences development, as Higgs and Clifford (1982) state, does not allow students to produce enough L2 output or to be aware of the errors in their interlanguage since instructors do not focus on providing enough feedback on L2 syntax.

The present research study sets out to address the aforementioned issue of competences inequality by investigating the impact of Swain's (1985) pushed output hypothesis and its three functions on 16 Colombian EFL students' L2 oral production employing a mixed method. It is worth noting that, to the best of my knowledge, research examining the effect of oral pushed output has not been conducted within a

Colombian context. I considered it would be interesting not only to analyze whether the hypothesis' claims would be supported or refuted by research in this Colombian setting, but also to gain students' perceptions towards it. This enquiry is different with regard to previous studies described in the literature review section since it involves: L2 oral production, a mixed method approach, the examination of the three output functions, multiple output activities, and insights gaining of students' perceptions towards pushed output. Thereby, this study attempts to answer the following main question: *What is the impact of oral pushed output on students' L2 oral production*? Moreover, four sub-research questions are intended to be answered in this research study: *1. Do output group students outperform significantly students in non-output group in terms of noticing? 2. What type of oral pushed output activity (one-way or two-way) leads to more hypothesis testing episodes? 3. What metalinguistic reflections do learners engage in after performing oral pushed output activities? 4. What are the students' perceptions towards oral pushed output?*

The present study is also undertaken to test the following null hypotheses in order to answer research sub-research question one and two: *H01: Oral Pushed Output participants do not outperform participants in non-output group in terms of noticing. H02: Two-way pushed oral output activities do not lead to more hypothesis testing episodes than one-way oral pushed output activities.*

Literature Review and Theoretical Framework

Research intended to investigate the effect of pushed output on L2 acquisition and linguistic accuracy has been sharply increasing over the past three decades (e.g., Basterrechea, García & Leeser, 2013; Ellis & He, 1999; Garcia-Mayo, 2002; Izumi, 2002; Izumi and Bigelow, 2000; Leeser, 2004; Nobuyoshi & Ellis, 1993; Rezvani, 2011; Shehadeh, 2003; Sitthitikul, 2017; Suzuki and Itagaki, 2007; Uggen, 2012). However, such studies have merely focused on written output and on one of the functions (the noticing function) using quantitative methods; giving little attention to the other two functions (hypothesis-testing and metalinguistic) of the hypothesis. Other studies have yielded results that are inconsistent with Swain's claims (e.g., Izumi and Bigelow, 2000; Rezvani, 2011; Uggen, 2012) and they are attributable to the limited number of output tasks, the short-term treatments, and the learners' foreknowledgement about the task to accomplish. Treatment written output activities in such studies have mostly been text reconstructions. Few studies have investigated the effect of pushed output hypothesis on students' oral production (Mamaghani & Birjandi, 2017; Byrne & Jones, 2014; Sadeghi & Edalati, 2014), yielding findings that are consistent with Swain's claims on grammatical accuracy development, but weak due to short period of implementation of pushed output, lack of interactive tasks, and the absence of the other output functions (hypothesis-testing and metalinguistic).

Pushed Output Hypothesis

The perspective held in second language acquisition (SLA) in the early 1980s involved the notions around Krashen's (1985) comprehensible input hypothesis. As a result, rooted on its premises, L2 teachers began to favor the act of providing learners with extensive input-rich environments and fluency-based instruction arguing that L2 was only acquired effectively, as Krashen (1982, 1985, 1998) established, through interesting, meaningful and relevant input (e.g., reading and listening material learners are provided with).

Nevertheless, Swain (1985, 1995, 1998, 2005) formulated the pushed output hypothesis as a reaction to Krashen's by establishing that comprehensible input alone was not sufficient for L2 acquisition, especially for the development of learners' linguistic competence. While she was conducting research with French immersion students in Canada, she observed that despite students had spent years in immersion programs still had short target-like abilities as well as less grammatical proficiency (syntactic competence) than their native-speaking peers. Swain hypothesized that what learners needed in order to enhance grammatical competence was opportunities to be pushed in L2 production. She claimed that language production, be it written or oral, played a paramount role in the L2 acquisition process. It forces students to move from a purely semantic analysis of the language (as in comprehension) to a syntactic analysis of it (as in production). This movement allows learners to stretch their interlanguage and improve their grammatical competence as long as such an output is meaningful and contextualized. Ever since, many researchers and teachers have increasingly paid attention to output and considered it to have a crucial role in ESL and EFL teaching and learning. The term "pushed" means being obliged to perform beyond ones' normal comfort level (Nation, 2011) and "pushed output" refers to the type of output that "reflects what learners can produce when they are forced to use target language accurately and concisely" (Ellis, 2003, p. 349). According to several scholars, when learners know they have to speak, they are pushed to pay more attention to what is in the input. If they never have to speak, they might be content with always processing the input only for meaning. But if they know that there will be production pressures on them at some point, they may become more active processors of how something is said and not just what is said.

Swain (1993, 1995, 1998, 2005) has further explored more about such a hypothesis and suggested three functions that output serves in the output hypothesis and in SLA provided that it is meaningful and contextualized:

1. The noticing function: Swain and Lapkin (1995) hypothesized that under certain circumstances, output promotes noticing and pushes learners to process language more deeply, with more mental effort than does input. The learners are in control when they produce language output, and can play more active roles in their

90

learning; thence, when attempting to produce the target language (vocally or silently), learners notice a gap between what they want to say and what they can say (they do not know how to say or write precisely the meaning they wish to convey), leading them to identify consciously their linguistic problems and lacks of knowledge, focusing on something they need to discover about their L2. Such a process involves the usage of previous linguistic knowledge in ways it had not been used before and contributes to the enhancement of accuracy.

In examining the noticing function, Izumi (2002) and Basterrechea et al., (2013) investigated whether output prompted learners to notice certain grammatical features on subsequent input using a carefully designed multi-stage text reconstruction and a dictogloss task as treatment for the experimental group, allowing the learners to compare their first production with the input subsequently received. Basterrechea et al., (2013), however, were also interested in the type of reflections the collaborative group engaged in while reconstructing the text in pairs. Results in both studies revealed that such pushed output tasks affected noticing when students became aware of differences or gaps between their interlanguage and target language forms. Additionally, participants' attention was drawn to the events in the text, i.e., content.

2. The hypothesis-testing function: Swain states that language output is an essential way for learners to test hypotheses about the target language. In other words, learners semantically modify or confirm their output (test their hypothetical output) in response to feedback received during fruitful interaction and negotiation of meaning with their peers or teacher; prompting, thus, higher accuracy.

To examine how output can lead to learners' L2 hypothesis testing and output modification, Shehadeh (2003) implemented a picture description task to collect data from sixteen nonnative students in the UK. He audio recorded learners' interactions and examined hypothesis testing episodes. Having them Back-to-back, NNSS (non-native speakers) described the picture to a NSS (native speakers). Results revealed that students tested out their output hypotheses on L2, i.e., experimented which target linguistic form sounded better, appealed for assistance, and requested information every 1.8 minutes.

3. The metalinguistic (reflective) function: Swain maintains that through the production of language, reflection on others' or one's own target language use (L2 output performance) is triggered. Also, she argues that language output enables learners to engage in syntactic as well as semantic processing; that is, to internalize knowledge and raise awareness on forms, rules, function, and meaning of the target language provided that the language production context (task) is communicative, leading thus to learner's deeper understanding of L2 and accuracy development.

In investigating the metalinguistic reflection function, Garcia-Mayo (2002), implementing text reconstruction as a treatment as well, aimed to determine the

effect of output in the metalinguistic function. Learners' interaction was codified and language related episodes (LREs) were identified. She discovered that output tasks promoted learners' attention to form and got engaged in features targeted by the text reconstruction. In the same fashion, Suzuki and Itagaki (2007) examined the type of metatalk learners engaged in after performing writing output-oriented tasks. Intermediate and advanced Japanese learners of English wrote about their thinking processes after having received explicit correct solution feedback. The metalinguistic reflections shed light on the way they had noticed linguistic forms and how they had tested hypothesis, supporting thus Swain's (1995) output hypothesis. Results revealed that learners' metalinguistic reflections enabled them to internalize linguistic competence and allowed them to engage in syntactic processing, essential in SLA.

Methodology

Research design

A mixed method is applied in this study following a true experimental research design with the purpose of interpreting and describing the data obtained from the quantitative and qualitative data collection instruments. According to Creswell (2014), true experimental designs involve the random assignment of participants to each group in the study. Researchers in this design provide a treatment to one (experimental) group and withhold it from another group (control group); then, they analyze how both groups score on the outcome and determine whether it was the treatment and not other factors that influenced that outcome. In this enquiry, the researcher randomly assigned the 16 participants to an experimental (output) group and a control (non-output) group. Table 1 below summarizes the description of the research design in this study and the purpose of both qualitative and quantitative traditions in this study.

Table 1. Summary of Research Design

Research Design	Research Traditions	Research Objectives
Mixed method	Quantitative	To determine whether participants in the oral pushed output group outperform significantly participants in non-output group in terms of noticing.
		To analyze what type of oral pushed output activity leads to more hypothesis testing episodes
	Qualitative	To describe the metalinguistic reflections the learners engage in after performing pushed IGAs
		To unveil students' perceptions towards oral pushed output.

Context

The study was conducted in a private school from Ibagué, Colombia. The school is comprised by 280 students in an only-existing shift. The school implements CLIL methodology (Content and Language Integrated Learning) from pre-school to sixth grade; therefore, the textbooks used in the English class enact this methodology. Then, from seventh grade to eleventh grade, a grammar-focused textbook and methodology is developed. This context was selected for the present study purposes because this school exposes students over six years to a fluency-based methodology, and then they are expected to undergo a grammar-focused textbook and methodology. According to the school teachers, under the implementation of such a methodology in those early grades, accuracy is not demanded from students, being this the likely cause by which they show high fossilization levels.

Participants

The sample of the study consists of 16 seventh grade intermediate level learners (B1) as recommended by their English teacher. Participants' ages range from 11 to 12 years and their strata ranges from three to five. That is, they are medium-class students. The selection of such a population was encouraged by the fact that, similar to what Swain (1985) observed in the Canadian immersion program, these students show low syntactic processing compared to their high semantic processing; that is, their accuracy was shorter than their fluency on L2, as mentioned by their English teachers.

The L2 linguistic form

The target forms emphasized in this study are two of the English past narrative forms: past progressive and simple past tense. By the time output group participants were exposed to a five-week treatment, they had received merely input of these target linguistic forms. That means participants were aware of the target forms receptively but did not have productive mastery of them.

Data collection procedure

Pretest and posttest

Shuttleworth (2013) conceives pretest and posttest design as a pivotal method to compare participant groups and measure the degree of change occurring as a result of treatments or interventions. In this study, the pretest was used to identify students' grammatical errors and writing them down and it was administered in the participants' classroom with the researcher and their teacher's supervision soon after the participants were assigned to the output and non-output groups and four days before starting the treatment. Such pretest involved eight gap-filling short anecdotes and fourteen multiple-choice items comprising the L2 linguistic form and it was aimed at determining both output and non-output participants' noticing of those errors made in the forms.

Audio-record

Button and Lee (1987) explain that the use of tape recordings is a practical strategy for making the data collected through conversation available for extended analysis. These recordings can be implemented in qualitative, quantitative or mixed approach research and must be transcribed subsequently. In this study, the audio-recordings were used in the treatment while participants performed the oral pushed output activities and they were transcribed later in order to analyze, code, classify, and categorize qualitatively the participants' LREs to answer sub-research question three. Furthermore, they were used to identify and calculate quantitatively the participants' hypothesis testing episodes (HTEs) in both one-way and two-way oral pushed output activities to determine which one led to more output modification or HTEs. Finally, such audio-recordings were implemented to gain qualitative insights from participants' perceptions towards oral pushed output through interviews.

Stimulated recall

According to Suzuki and Itagaki (2007), learners are interviewed by the researcher in a stimulated recall; further, using meta-language, they converse about learners' decision-making and thinking processes (e.g., lexical choices, linguistic structure chosen, or activity content) while performing the output task and reflect on how they solved their task. They were undertaken in this study to gain qualitative insights on the metalinguistic reflections the students engaged in while performing the pushed output treatment. It took the form of a semi-structured interview in which learners were asked in Spanish language what they were paying attention to during the performance of the oral pushed output activities as well as what they were thinking of when they stopped in specific moments during the performance.

Interviews

Interviews were used in this study to qualitatively inquire into participants' perceptions on oral pushed output to answer the sub-research question four and were applied to the eight output group participants six days after the posttest was administered. Additionally, they comprised two open-ended questions in Spanish language related to what participants thought about the oral pushed output activities in the treatment as well as how they felt while performing them.

Application and Treatment Materials

The current study was carried out at a private school in Ibagué, Tolima, Colombia during a seven-week period. First, the researcher handed the seventh-grader minor participants the informed consent letter. Two days later, the researcher assigned the participants randomly to output and non-output groups and administered the pretest. During the following five weeks, the researcher asked output group participants one by one or in pairs to come out of the classroom for some minutes to perform three types of information gap activities (2 one-way and 1 one-way) that pushed them to produce oral L2 output. Such types of activities were performed three times during the treatment and their content was changed in the three occasions. First, in the picture-based storytellings, the student was given a set of pictures that composed a story and was asked to tell with their words the story shown in the whole portrayal and sequence of pictures. Second, in the retelling activity, students spent a few minutes reading a passage silently and, then, without looking back at it, started retelling the information to the researcher. Third, in unstructured role-plays, student A was given an assigned task and role and learner B was too. Based on the information given, they simulated a scene. The selection of these three types of activities was motivated by the benefits they have showed in pushed output research, according to Izumi and Izumi (2004), Shehadeh

(1999), and Nation and Newton (2009). Likewise, Nation and Newton (2009) claim that the aim of pushed activities is to make students produce L2 in a way that is beyond their normal comfort level (i.e., pushed way). This is achieved by exposing students to unfamiliar topics and by depriving students from preparation and planning for the production of the target language. These participants were prompted by the researcher after the appearance of an inaccurate use of a past narrative form using the implicit feedback strategy *recast*. Lyster (1998) defines recasts as the interlocutor's repetition of the student's ill-formed utterance while adjusting their intonation to highlight the error (e.g., *They **were walking***). It is worth noting that the three types of activities together with these implicit feedback strategies were piloted by performing them on 6 students who had similar characteristics of the study sample but not included in it, showing reliability to the researcher.

On the other hand, the non-output group was exposed to comprehension activities, i.e., activities that did not push them to speak such as picture sequencing and answering of reading comprehension questions. These participants performed such activities each week during the five weeks of the treatment and their content was different in the five occasions. Once the treatment ended, the researcher interviewed the output group participants and administered the posttest to both groups. Figure 1 summarizes all of the abovementioned stages through the study schedule.
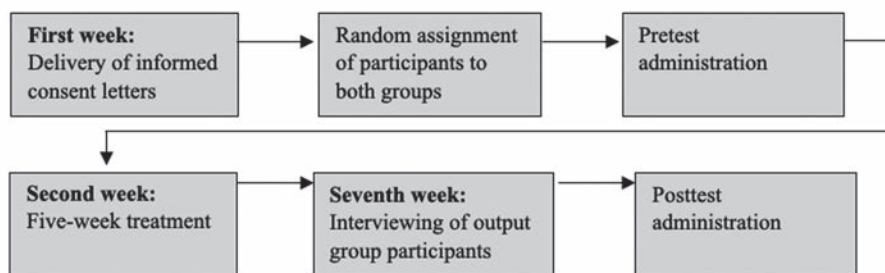

Figure 1. Summary of the schedule of the study.

Data analysis and Interpretation

Research question one

The researcher and the participants' classroom teacher, i.e., second rater, identified, wrote down and compared the grammatical errors made in the target forms in the pretest with those in the posttest to calculate the number of linguistic forms noticed. Both raters agreed on all of the scores they gave to each of the participants as part

of the inter-rater reliability. Quantitative data from such noticing episodes was statistically analyzed using a descriptive statistics approach to determine if pushed output participants did better on noticing than participants in the non-output group. Data distribution was normal. Table 5 portrays the descriptive statistics of the means of output and non-output group on noticing. Table 6 illustrates the independent samples t-test results of means scores difference.

Table 2. Descriptive Statistics of Output and Non-output group on Noticing of past narrative forms

	Group	N	Mean	Std. Deviation	Std. Error Mean
Noticing	Output	8	6,7500	4,39968	1,55552
	Non-output	8	4,8750	1,95941	,69276

Table 2 indicates that mean scores in noticing between output and non-output group were somewhat different. Nevertheless, to determine if the means scores difference was statistically significant or not, as shown in Table 6, an independent samples t-test was run (The mean difference is significant at ,05 level and below).

Table 3. Results of Independent Samples T-Test for Noticing of past narrative forms of Output and Non-output group

t-test for Equality of Means

		t	df	Sig.	Mean Difference	Std Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Noticing	Equal variances assumed	1,101	14	,289	1,87500	1,70281	-1,77716	5,52716
	Equal variances not assumed	1,101	9,672	,297	1,87500	1,70281	-1,93662	5,68662

As it can be seen, although in Table 2 the mean score of the output group (6,75) was higher than the mean score of the non-output group (4,87), Table 3 illustrates that the results of the independent samples t-test did not show any significant differences in the mean scores of noticing between output and non-output group since the sig. level of the test was ,289 (P = ,289), which was greater than the critical value (research confidence interval) ,05 (P > ,05). In other words, pushed output participants did not outperform non-output participants in terms of noticing of their linguistic problems in the targeted linguistic forms. Therefore, the first null hypothesis was not rejected. This result does not provide empirical support to Swain's (1985, 1993, 1995) claim

on the noticing function that output promotes learners' noticing of their linguistic problems. By the same token, this finding does not either coincide with previous studies described in the literature review on the effects of pushed output hypothesis which yielded positive findings for students' noticing of different linguistic forms such as those of Izumi (2002), Basterrechea, García, and Leeser (2013), and Sadeghi and Edalati (2014), and Mamaghani and Birjandi (2017). Nevertheless, the result of this question is in line with and adds up to the negative results trend in studies investigating the effect of output on noticing such as Rezvani (2011), Sitthitikul (2017), Izumi and Bigelow (2000), Song and Suh (2008), and Uggen (2012), noted in the literature review.

The negative result on the noticing function as part of the sub-research question one in this study might be attributable to the way the researcher applied the implicit feedback with students during the performance of the pushed output activities. Recasts may not have drawn students' attention to linguistic problems efficiently. After students produced an ill-formed utterance, the researcher would immediately repeat it in a correct way and students modified their output and kept on performing the activity. However, the researcher did not adjust his intonation or made a considerable pause to highlight the error as Lyster (1998) proposes. This might not have provided students enough time to be sufficiently aware of the linguistic error they just made.

Research question two

Quantitative data from the hypothesis testing episodes (HTEs) identified in the audio-recording transcriptions was also statistically analyzed using the same aforementioned program to determine what type of oral pushed output activity (one-way or two-way) leads to more HTEs. As there were 2 one-way and 1 two-way oral pushed output activities, 1 one-way and 1 two-way activity were considered as variables in the employment of the independent samples t-test for equality purposes in the mean scores analysis of HTEs. Data distribution was normal. Table 4 displays the descriptive statistics of the means scores of output group participants' HTEs in both activities and table 5 displays the independent samples t-test results of means scores difference.

Table 4. Descriptive Statistics of Output group participants' Hypothesis Testing Episodes in One-way and Two-way Output Activities

	Activity	N	Mean	Std. Deviation	Std. Error Mean
HTEs	One-way	3	15,6667	4,50925	2,60342
	Two-way	3	6,6667	2,08167	1,20185

As shown in this table, the mean scores of HTEs in the one-way output activity that involved three picture-based storytellings (15,6) are seemingly greater than the mean scores of HTEs in the two-way output activity that involved three unstructured role plays (6,6). To ensure whether or not the observed difference was statistically significant, an independent samples t-test was performed (see Table 5).

Table 5. Results of Independent Samples T-Test for Output group participants' Hypothesis Testing Episodes in One-way and Two-way Output Activities

t-test for Equality of Means

		t	df	Sig.	Mean Difference	Std Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Noticing	Equal variances assumed	3,139	4	,035	9,00000	2,86744	-1,03871	16,96129
	Equal variances not assumed	3,139	2,815	,056	9,00000	2,86744	-,47344	18,47344

The results indicated that the difference between one-way and two-way output activities, t (4) = 3,139, P = ,035 < ,05, was considered to be statistically significant. Accordingly, it can be claimed that a one-way oral pushed output activity led to more learners' HTEs than a two-way output activity. That is, students modified more output in one-way pushed output activities than in two-way output activities with teachers' assistance through implicit feedback. Thus, the second null hypothesis was rejected. This is consistent with Swain's (1995, 1998) claim on the hypothesis testing function that learners modify output (test their hypothetical output) when producing L2 output and receiving feedback from their peers or teachers. The result in this second question was also congruent with the results of similar studies pointed out in the literature review (Shehadeh, 2003), even though this study was not interested in examining the frequency of occurrence of students' hypothesis testing episodes. Likewise, the beneficial role of one-way output activities that Mackey (2012) and Shedaheh (1999) advocate was confirmed with this finding. They state that this type of activity has the learner do most of the talking as well as provides more opportunities for them to generate and use more pushed output.

Research question three

Qualitative data from the stimulated recalls was carefully analyzed using content analysis in the audiorecording transcriptions of the stimulated recalls before doing

triangulation with the interviews data to ensure validity. This analysis was done in each of the activities so as to identify, classify, and categorize language related episodes (LREs). The resulting LREs categories are: verb tenses, L2 lexis (vocabulary), meaning expression, and events in the activities, i.e. activity content. Each one of these categories were analytically interpreted and grouped, resulting in two emerging themes: (1) content-based metalinguistic reflections and (2) form-based metalinguistic reflections.

Learners' Metalinguistic Reflections Pertaining to Content

Data revealed that while output group participants underwent oral pushed output, students were thinking about the content embedded in the activities, i.e., about the events in the activities: *STUDENT 4*: "[I was thinking] er... about what was happening and making sense out of it". *STUDENT 5*: "in the story itself.... yes... Like to try to tell it well because I didn't know the story". *STUDENT 1*: "In that part, I was thinking of what was going to be before... sorry, after the part when the girls arrived home after the school". *STUDENT 8*: "about the logic of the things that were occurring because it is not like usual like that you see... an alien into the woods... yes"

Students also reported that in the unstructured role plays and retelling activities, they were paying attention to the meaning they wanted to express: *RESEARCHER*: "what were you paying attention to while you were doing the activity?". *STUDENT 7*: "I don't know, to how to keep the conversation going". *STUDENT 6*: "to the topic of the conversation... so I didn't get out of it". *STUDENT 5*: "about what the was going to say in order for me to say something that made sense". *STUDENT 1*: "with respect to what he last mentioned... so I could connect it and keep the same idea"

Learners' Metalinguistic Reflections Pertaining to Form

After asking students what they were thinking about when they paused in certain moment during the activities performance, they said were thinking about the L2 verbs in past tense that were required to tell decently the stories in the picture-based storytellings and the retellings *STUDENT 5*: "I was thinking about how I had to put the verbs". *STUDENT 8*: "well, about things that I didn't know like verbs... or that I didn't know how to tell them in past... for example, I didn't know that of *he found* and things like that". *STUDENT 1*: "I stopped because I didn't remember the past of *see*". *STUDENT 3:* "in my case, well, maybe I stopped to identify... I mean, take the verb and pass it to its past [tense]". *STUDENT 2*: "on the connectors... when to use while and when to join the tenses"

Students would also stop to think about what past tense to use. That is, they were concerned about the target linguistic features selected in this study: *STUDENT 2*: "[I was thinking] of the conjugation... I mean, whether to put the verbs in past continuous or in past simple". *STUDENT 6*: "I stopped because I was thinking of the difference on how to use the past continuous with the other one". *STUDENT 4*: "Sometimes I got stuck in the telling of the story since I had to stop to think about the past"

After being asked about what they were paying attention to and thinking about when stopping during performance, students reported that their attention was drawn to specific L2 vocabulary that was necessary in the performance of the picture-based storytelling and retelling activities as well as in specific contexts of the role plays such as accident and paranormal scenarios: Student 3 Student 7: "to the words.... the key words that are needed. Words that one forgets that doesn't know how to tell them... the words that I asked you for". Student 1: "to some words of vocabulary, for example that of *forest*". *STUDENT 4*: "I didn't remember how to say *letrero* (advertisement, in English)... and the other time was here where I didn't know how to say about the man". *STUDENT 8*: "I don't know the words *robar, esconder, quemar*" (rob, hide, and burn, in English). *STUDENT 5*: "like about the words to say what was happening". *STUDENT 3*: "I don't know the words *chocar* and *afeitar*" (crash and shave, in English). *STUDENT 3* "[I stopped to think about] the vocabulary, because one had to think of what to say". *STUDENT 6*: "I didn't know vocabulary for this case, as it said we had to play a situation about ghosts, I didn't know how to say words related to that". *STUDENT 5*: "that sometimes… sometimes I was kind of recalling or... I kind of like didn't find the specific word".

As observed above, students focused on both content (activities events and expression of meaning) and form (verb tenses and L2 lexis) while performing oral pushed output. In other words, students engaged in both semantic and syntactic processing, a dynamic that starts to shed light on how they can equalize their semantic competence with their syntactic competence.

These findings provide empirical support to Swain's (1998, 2001, 2005) claim on the metalinguistic function of the output hypothesis that output enables learners to engage in both semantic and syntactic processing rather than in purely semantic analysis of the language (comprehension); leading thus to learner's internalization of knowledge, deeper understanding of L2, and accuracy enhancement. In the same fashion, this finding is in line with previous experimental studies noted in the literature review (Byrne & Jones, 2014; Garcia-Mayo, 2002; Suzuki & Itagaki, 2007; Swain & Lapkin, 2002) which claimed that through pushed output tasks, learners internalized linguistic competence and engaged in syntactic processing. In Garcia-Mayo's (2002) results, however, learners paid attention only to form, that is different from this finding which students focused on content (semantic processing needed in comprehension) and form (syntactic processing needed in production) at the same time.

Research question four

Qualitative data from the interviews was carefully analyzed using content analysis in the respective audiorecording transcriptions and triangulated with data from the stimulated recalls. In order to find commonalities among the students' responses, data was coded and categorized. Resulting categories pertained to feelings (relieved nerves, self-confidence, and accompaniment sense) and double purpose of activities (nice and useful, pleasing and subskills pleasing and subskills noticing and exposure trigger). These categories were piled so as to analyze them and interpret them more deeply until broad themes emerged in light of how students perceived oral pushed output. The themes that emerged pertained to (1) oral pushed output as affective regulator in L2 oral production and (2) oral pushed output as a trigger of noticing and exposure to L2 vocabulary, pronunciation, and grammar.

Oral Pushed Output as Affective Regulator in L2 Oral Production

When asked in the interviews how they felt during the performance of the output activities, students said they did not feel comfortable, with lack of self-confidence, and nervous in certain moments, especially in the first weeks of the treatment. However, as treatment passed by, the more students gained familiarity with the researcher and were exposed to a wide repertoire of both one-way and two-way activities, the more they began to overcome such feelings and get controlled even when still being pushed to produce. Further, the fact that students were receiving immediate implicit feedback (researcher's assistance) helped them not to feel incapable of performing well in the activity: *STUDENT 4*: "Uh... I felt awkward because when we began, I wasn't very accustomed". *STUDENT 3*: "I felt comfortable, well... as we did more activities with the teacher, I felt okay". *STUDENT 5*: "I felt good, I understood the dynamic more and more... I just sometimes got stuck in the telling of the stories and in the thinking of the past (tenses)... and for the rest, all good". *STUDENT 6*: "I think this helped me a bit on having more control of... like when one doesn't know someone or something very well and being more relaxed, because at the beginning I was a little bit like that... like stressed, pressed, and scared, but I started to be more relaxed afterwards". *STUDENT 8*: "Perhaps at the beginning I felt a little weird because I didn't know if I was able to say everything well, but later on I didn't feel like that anymore... and well, plus, the teacher gave us a little help whenever we got wrong, and stuff like that".

The unstructured role plays allowed learners to work in pairs to complete the task. Such a fact also made students feel good in the performance of the activities: *STUDENT 7*: "I felt very good... and because we could do the activities accompanied with another classmate"

Oral Pushed Output as a Trigger of Noticing and Exposure to L2 Vocabulary, Pronunciation, and Grammar

When asked what they thought about the oral pushed output activities, students reported that they helped them realize some problems of their interlanguage system as well as be exposed to vocabulary, pronunciation, and grammar: *STUDENT 1*: "I consider them good as this exercise was useful for me to enhance my speaking... and besides, I learned more vocabulary". *STUDENT 3*: "Personally, I could learn things that either I didn't know or I didn't remember how to use them and pronounce them; and plus, I don't know... I mean, they helped me a lot for other English classes because the topics were related". *STUDENT 6*: "I found them nice, and it all also helped me in some things I was bad at in English; also, it helped me improve some words and understand a little better some passages in English"

Students also mentioned that they found the activities appealing and enjoyable and with a learning purpose; besides, they expressed that their content was not isolated from their interests: *STUDENT 4*: "the activities were cool, dynamic and... obviously I took them seriously since it was like a test that told you if you were good or bad at things in English". *STUDENT 5*: "I found them good, well, like the topics were quite related to us". *STUDENT 7*: "I found them nice, like dynamic and not that serious". *STUDENT 8*: "the activities were very cool because... well, at the time we learned English, words, pronunciation; they were also dynamic and... yes... nice".

Overall, students agreed that oral pushed output was appealing and helped them control their feelings of fear and nervousness in L2 production as well as be exposed to several elements of the target language. These results lend support to Swain's (1985, 1993, 1995, 2005) general claim on the pushed output hypothesis that by pushing L2 learners to produce meaningful and contextualized output, they can stretch their interlanguage. The fact that students perceived oral pushed output as a meaningful tool in SLA is also in line with Al-Jamal's (2014) study results which yielded positive students' attitudes towards audio-visual input enhancement in pushed output dictogloss tasks. Al-Jamal measured students' perceptions quantitatively through a survey and observed that all of the students felt that such an activity had improved their learning process, whereas 81% of them felt that doing another similar activity in other courses would be helpful for their language learning. Finally, the fact that students expressed that oral pushed output helped them enhance their vocabulary in L2 is consistent with Ellis and He (1999) and De la Fuente's (2002) studies that revealed that frequently exposing students to output led to development of their L2 vocabulary. In other words, pushed output hypothesis does not only show positive effects on grammatical accuracy, but also on improvement of L2 lexis.

Conclusions and Implications

The results in this enquiry show a positive impact of oral pushed output on students' L2 oral production as well as in SLA process. It can be concluded that students can modify more output with teachers or peers' assistance through one-way oral pushed output activities than through two-way activities. Furthermore, pushing students to produce meaningful, interesting, and contextualized output can contribute to an equality in their semantic and syntactic competence since they can engage in both processings through metalinguistic reflections. Additionally, students perceived oral pushed output as an affectivity regulator that allowed them to get accustomed when producing L2 oral output as well as a trigger of noticing and exposure to L2 vocabulary, grammar, and pronunciation. Nonetheless, these results do not completely support Swain's pushed output hypothesis claims since oral pushed output did not promote significant noticing of output group participants' linguistic problems in L2 past narrative forms.

The present study may have implications for second language teachers and L2 materials developers. This study suggests that EFL and ESL teachers implement oral pushed output to help students gain confidence in L2 oral production and enhance L2 vocabulary and pronunciation. It is also suggested that English instructors apply this type of output to overcome the issue of short syntactic accuracy, especially in students that have long been exposed to fluency-based and input-based methodologies in which their grammatical errors are not often corrected and opportunities to produce L2 oral output are not provided respectively. Additionally, it is recommended that L2 curriculum designers and material developers make changes in the way of constructing curricula, syllabi and English textbooks based on a vast presentation of input. Oral pushed output (tasks) should be included in them so as to provide students more opportunities to produce L2 oral output in the ESL and EFL classroom.

Limitations and Further Research

As many other research studies, this study suffers from limitations that pose the issue of results generalizability and strong conclusions drawing. One limitation is the relatively short five-week period of treatment in this study. Further studies are thus required to carry out long term treatment period. Moreover, these results could be more persuasive if the number of participants were larger to increase generalizability and external validity. The next limitation lays in the lack of rigor in the selection of participants. This study relied on the participants' teacher recommendation to start their random assignment to the groups. It is suggested that future research with similar purposes select participants based on a standard test in order to obtain more accurate results in pretest and posttest processes.

References

Al-Jamal, N. (2014). An Investigation of the Effect of Incorporating Input Enhancement in Dictogloss Tasks to Teach English Grammar on the Development of ESL Female Learners' Grammatical Awareness. *Arab World English Journal.* Retrieved from https://www.awej.org/images/Theseanddissertation/Nadahamadllaljamal/131.pdf.pdf

Basterrechea, M., García, M., & Leeser, M. (2013). Pushed output and noticing in a dictogloss: task implementation in the CLIL classroom. *Portalinguarum, 22,* 7-22. Retrieved from http://www.ugr.es/~portalin/articulos/PL_numero22/1%20%20Maria%20Basterrecha.pdf

Button, G., & Lee, J. (1987). *Talk and social organisation.* Clevedon, United Kingdom: Multilangual Matters Ltd.

Byrne, S., & Jones, C. (2014). Pushed and non-pushed speaking tasks in an EAP context: what are the benefits for linguistic processing and accuracy? *Studies about Languages, 24,* 87- 97.

Creswell, J. (2014). *Research design: qualitative, quantitative, and mixed method approaches* (4th ed.). Thousands Oak, CA: SAGE Publications.

De la Fuente, M. (2002). Negotiation and oral acquisition of L2 vocabulary. *Studies in Second Language Acquisition 24*(1), 81–112

Ellis, R. (2003). *Task-based language learning and teaching.* Oxford, UK: Oxford University Press.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition 21*(2), 285–301

García-Mayo, M. (2002). The effectiveness of two form-focused tasks in advanced EFL pedagogy. *International Journal of Applied Linguistics, 12,* 156-175. doi: 10.1111/1473-4192.t01-1-00029

Greene, J., & Caracelli, V. (1997). Advances in mixed-method evaluation: The challenges and benefits of integrating diverse paradigms. *New Directions for Evaluation, 74.* San Francisco: Jossey-Bass.

Gass, S. M. (2003). Input and interaction. In C. J. Doughty & M. K. Long, M. K. *The handbook of second language acquisition* (pp. 224-254). Massachusetts: Blackwell.

Higgs, T., & Clifford, R. (1982). The push towards communication. In T. Higgs (Ed.), *Curriculum, competence, and the foreign language teacher* (pp. 57–59). Skokie, IL: National Textbook Company.

Izumi, S., (2002). Output, Input Enhancement, and the Noticing Hypothesis. *Studies in Second Language Acquisition, 24*(4), 541–577. doi: 10.1017/S0272263102004023

Izumi, S., & Bigelow, M. (2000). Does Output Promote Noticing and Second Language Acquisition? *Tesol Quarterly, 34*(2), 239-278. http://dx.doi.org/10.2307/3587952

Krashen, S. (1982). *Principles and practice in second language acquisition.* London: Pergamon.

Krashen, S. (1985). *The input hypothesis: issues and implications.* New York: Longman.

Krashen, S. (1998). Comprehensible output. *System, 26*, 175-182. Retrieved from http://www.sdkrashen.com/content/articles/comprehensible_output.pdf

Leeser, M. (2004). Learner Proficiency and Focus on Form during Collaborative Dialogue. *Language Teaching Research, 8*(1), 55–81. doi: 10.1191/1362168804lr134oa

Lyster, R. (1998). Negotiation of Form, Recasts, and Explicit Correction in Relation to Error Types and Learner Repair in Immersion Classrooms. *Language Learning, 48*(2), pp. 183–218. doi: 10.1111/1467-9922.00039

Mamaghani, H., & Birjandi, P. (2017). Oral Pushed Output: The Route to Long-term Grammatical Accuracy. *Journal of Teaching Language Skills, 36*(1), 57-84. doi: 10.22099/jtls.2017.4044

Marsh, D., & Langé, G. (2000). *Using Languages to Learn and Learning to Use Languages.* Retrieved from http://clilcompendium.com/1uk.pdf

McDougald, J. (2009). The state of language and content instruction in Colombia. *Latin American Journal of Content & Language Integrated Learning, 2*(2), 44-48. doi:10.5294/laclil.2009.2.2.15

Nation, I., & Newton, J. (2009). *Teaching ESL/EFL Listening and Speaking.* New York: Routledge.

Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal, 47*(3), 203-210

Rezvani, E. (2011). The effect of output requirement on the acquisition of grammatical collocations by Iranian EFL learners. *Journal of Language Teaching and Research, 2*(3), 674-682.

106

Richards, J. (2008). Growing up with TESOL. *English Teaching Forum, 1*, 2-10.

Sadeghi, A., & Edalati, V. (2014).The impact of pushed output on accuracy and fluency of Iranian EFL learners' speaking. *Iranian Journal of Language Teaching Research*, *2*(2), 51-72. Retrieved from https://files.eric.ed.gov/fulltext/EJ1127288.pdf

Shehadeh, A. (2003). Learner Output, Hypothesis Testing, and Internalizing Linguistic Knowledge. *System, 31*(2), 155–171.

Shuttleworth, M. (2009). *Pretest-Posttest Designs*. Retrieved from https://explorable.com/pretest-posttest-designs

Sitthitikul, P. (2017). The roles of output in second language acquisition: a case study of Thai learners. *Catalyst, 15*(1), 63-76. Retrieved from https://www.apiu.edu/download-individual-articles-pdf/6-the-roles-of-output-in-second-language-acquisition/download

Song, M., & Suh, B. (2008). The effects of output task types on noticing and learning of the English past counterfactual conditional. *System, 36*, 295-312.

Suzuki, W., & Itagaki, N. (2007). Learner metalinguistic reflections following output-oriented and reflective activities. *Language Awareness, 16*(2), 131-46.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass, & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). New York: Newbury House Publishers.

Swain, M. (1993). The output hypothesis: just speaking and writing aren't enough. *The Canadian modern language review, 50*(1), 158-164. Retrieved from https://2010-soph-writing-nccu.wikispaces.com/file/view/05+Swain+(1993)+The+Output+Hypothesis.pdf

Swain, M. (1995). Three functions of output in second language learning. In G. Cook, & B. Seidhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford, UK: Oxford University Press.

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

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-483). London: Routledge.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics, 16*(3), 371-391 http://dx.doi.org/10.1093/applin/16.3.371

Swain, M., & Lapkin, S. (2002) Talking it through: Two French immersion learners' response to reformulation. *International Journal of Educational Research 37*(3&4), 285–304.

Uggen, M.S. (2012). Reinvestigating the noticing function of output. *Language Learning, 62*(2), 505-40.

Author

*Kevin López Páez holds a B.A. in English from Universidad del Tolima and is currently studying an M.A. in English Didactics in the same university. He has been teaching English for over five years. He is currently working as a part-time English teacher at a private English institute. His research interests include Social Justice in ELT, Identity, Multimodality and Foreign Language Didactics.

ORCID: https://orcid.org/0000-0002-6723-2645