SVETIMŲJŲ KALBŲ STUDIJOS / STUDIES OF FOREIGN LANGUAGES

Pushed and Non-pushed Speaking Tasks in an EAP Context: What Are the Benefits for Linguistic Processing and Accuracy?

Shelley Byrne, **Christian Jones**

crossref http://dx.doi.org/10.5755/j01.sal.0.24.6366

Abstract. This article reports on a mixed methods study investigating the effectiveness of pushed and non-pushed speaking tasks in a UK university setting with upper-intermediate students. Specifically, the study addressed a) if a pushed speaking task produced more language related episodes (LREs) than a non-pushed speaking task, b) the differences in the types of LREs produced by each task and c) whether a pushed speaking task resulted in more accurate usage of past narrative forms. Results showed that the pushed storytelling task produced significantly more LREs than the non-pushed task and it also identified that the most common LRE type for both pushed and non-pushed learners related to some form of output correction. The pushed group achieved greater accuracy gains from pretest and posttest scores but these gain scores were not found to be statistically significant. The study concludes that creating a push during spoken output activities can increase the occurrence of opportunities for linguistic processing, and subsequently interlanguage development, to occur.

Keywords: *Comprehensible Output Hypothesis, task effectiveness, language related episodes, second language acquisition.*

Introduction

Within the field of second language acquisition (SLA), the role of language production and its contribution to interlanguage development has faced much debate (Selinker, 1972). The notion that target language input was solely responsible for acquisition (Krashen, 1982, 1989) appeared to underplay the significance of output in interlanguage development, perceiving it to be only an outcome or result of successful language acquisition. However, following research within an immersion programme setting, Swain (1985) became one of the first advocates for the role of output in the enhancement of SLA. By giving learners opportunities to write or speak in contexts which demanded attention to both form and meaning (given the term of creating a "push" in learner output (Swain, 1985, p. 249)), Swain proposed that comprehensible output could supplement interlanguage development and thus formulated the Comprehensible Output Hypothesis (COH).

The hypothesis and the proposed acquisition-enhancing functions of output have generated much interest (e.g. Izumi, 2000, 2002; Izumi, Bigelow, Fujiwara & Fearnow, 1999; Soleimani, 2008; Swain & Lapkin, 1995). It has been claimed that pushed output activates noticing (when students become aware of differences or gaps between their interlanguage and target language norms), hypothesis forming and testing (in which learners exploit and push the boundaries of their interlanguages during attempts to convey meaning), metalinguistic awareness (which may involve one or more students working towards solving a linguistic problem) and syntactic processing (in which learners acquire linguistic features responsible for creating meaning as opposed to its comprehension). A review of literature shows that many COH studies to date have taken place in an immersion context and have often used written output tasks to explore the hypothesis. In addition, they have not tended to take into account how learners themselves view the production of LREs during pushed and non-pushed output tasks. This study seeks to address these issues by answering the following research questions:

1) Does a pushed speaking task result in more language related episodes (LREs) than a non-pushed task for adult upper intermediate learners at an HE institution in the UK?
2) In what ways do pushed and non-pushed learners differ in the type of LREs they display?
3) Does a pushed speaking task result in more accurate performance when comparing pre- and posttest results for the past simple, past continuous, past perfect simple and past perfect continuous?

Literature Review

The COH is founded on the belief that language production itself can lead to an extension or deeper understanding of a target language in a learner's mind and, ultimately, can result in a greater level of acquisition. Attempts to produce language which accurately and efficiently convey meaning are believed to initiate cognitive processes which may

assist in the development of learners' interlanguages (Swain, 1985; Swain & Lapkin, 1995). Although previous theories had all but dismissed the role of output in second language acquisition (SLA) (see Krashen, 1982, 1989) or had alluded to it as a communication tool useful for receiving feedback and more input (see Long, 1983), Swain was amongst the first researchers to highlight its direct acquisitional effects (Gass, 1997). Whilst she did not dispute the importance of input in SLA, she identified acquisition enhancing opportunities that input alone cannot provide, namely "noticing, hypothesis forming and testing, metalinguistic function and syntactic processing" (Muranoi, 2007, p. 56).

One of the principal studies which informed the COH was conducted in 1985 within a Canadian French immersion class setting. With aims to promote positive attitudes towards the target language and to sustain academic and linguistic achievement, immersion programmes often combine content and second language (L2) subjects. From the start of kindergarten, students are exposed to instruction in the L2 and are encouraged to communicate using the target without "undue" attention to grammatical and structural errors (Ranta & Lyster, 2007, p. 141). Whilst such a setting would satisfy Krashen's "optimal input" criteria (1982, p. 138) and would seemingly be ideal for successful L2 acquisition, Swain noted that learners were noticeably inferior and "off target" in their speech and writing when compared to native speaker peers despite relatively equal capabilities in reading and listening (Swain & Lapkin, 1995, p. 372).

A possible explanation for the differing productive capabilities could have been the setting itself. At the time of the study, the immersion setting was viewed as predominantly teacher-centred and non-facilitative of extended responses. For instance, Allen, Swain, Harley & Cummins (1990)'s study calculated that fewer than fifteen per cent of utterances were longer than a clause in length. Furthermore, the setting did not always necessitate learners to be accurate in their output since students were barely given corrective feedback following grammatical errors (only nineteen per cent of grammatical errors received feedback) and any feedback offered was often "confusing and unsystematic" (Allen et al., 1990, p. 67). Consequently, an environment was created in which there was "little social or cognitive pressure to produce language that reflects more appropriately or precisely their intended meaning" (Swain, 1985, p. 249). Also, although much input is provided in immersion programmes and although input's role in SLA cannot be discredited, it alone is insufficient for acquisition (Gass, 1997); input allows students to employ semantic strategies when extracting and decoding meaning but it does not always involve a need to attend to syntax when *creating* meaning. These notions led Swain to conclude that for students to progress along the interlanguage continuum, they must be encouraged to produce plentiful, extended language responses which adequately and accurately convey meaning. By 'pushing' students "to be more comprehensible than they already are" and by creating a need for greater accuracy and appropriacy, learner interlanguage can be stretched (Swain, 1985, p. 249). This, in turn, could enhance acquisitional opportunities whilst possibly increasing the occurrence of periods when acquisition could take place: i.e. during modified output, noticing, hypothesis formation and testing, metalinguistic awareness and syntactic processing (Swain, 1985, p. 249).

Other, more recent research into the COH has followed various channels. Many studies have focussed on the benefits or shortcomings of individual functions of output: noticing (Izumi, 2000, 2002; Izumi et al., 1999; Pica, 1988; Schmidt, 1990; Schmidt & Frota, 1986; Soleimani, 2008; Swain & Lapkin, 1995), hypothesis forming and testing (Swain, 1998; Shehadeh, 2003), metalinguistic function and collaborative dialogue (Del Pilar Garcia Mayo, 2002; Kim, 2008; Kowal & Swain, 1994; Leeser, 2004; Swain & Lapkin, 2001). Other studies have also investigated the effect of different types of feedback on modified output for example Lyster (1998) and Sheen (2008). Perhaps of more significance to this study are the investigations which have examined the effects of pushing students in their output (e.g. Linnell, 1995; McDonough, 2005; McDonough & Mackey, 2006; Nobuyoshi & Ellis, 1993; Pica, Holliday, Lewis & Morgenthaler, 1989; Shehadeh, 1999; Van den Branden, 1997). Whilst many of these studies have been criticised for focussing more on the occurrence of acquisition opportunities, rather than evidence of acquisition (Shehadeh, 2002), they have still been influential and have identified gaps still to be explored.

For instance, Pica et al. (1989)'s study into learner reactions to various native speaker signals of non-comprehension sought to test hypotheses regarding comprehensible output opportunities and feedback. They discovered that native speaker signals of a lack of understanding, regardless of feedback type, had a significant effect upon non-native responses. They also found that clarification requests produced more modified output than "model utterances" requiring confirmation from the learner since it was the learners' responsibility to resolve the misunderstanding (Pica et al., 1989, p. 83). Nobuyoshi and Ellis (1993)'s exploratory study shared similarities with Pica et al. (1989) in that pushed and non-pushed tasks were compared. The key difference was that instead of focussing on feedback types and the occurrence of modified output, its aim was to ascertain whether either task resulted in differing rates of past tense linguistic accuracy. Although exploratory in nature, the study found that two of the three experimental group learners displayed delayed past tense accuracy gains whilst the control group showed no "overall gain in accuracy" (Nobuyoshi & Ellis, 1993, p. 208). This provides some limited empirical support for Swain's COH claims.

Another study which aimed to discover the effects of modified output on interlanguage development was Linnell's 1995 research into the effects of negotiation on syntacticization. Negotiation is believed to comprise "the provision of corrective feedback that encourages self-repair involving accuracy and precision not merely comprehensibility" (Lyster & Ranta, 1997, p. 42) so, by proving to have an effect on syntax, Linnell was able to demonstrate a possible link between interlanguage and the developmental processes involved in creating a 'push'. The

study consisted of low-intermediate adult learners in a university and aimed to discover the modified output effects on syntacticization, different types of negotiation on syntax and the effect of negotiation over time. Using a pretest, posttest and delayed posttest design to focus upon past tense use, groups were divided into those who received negotiation via clarification requests, negotiation via confirmation requests, interaction with no negotiation and a group with no interaction. The results found that a fifth of syntacticized responses were produced in modifications, although L2 responses were not always accurate. Furthermore, learners who received clarification requests modified their output on more occasions than those who received confirmation checks. Finally, students who modified their syntax during negotiation appeared to syntacticize over time. This study was important in that the occurrence of modified output was directly linked to the effects it may have on syntax development and furthermore in long-term retention of syntactical knowledge.

Van den Branden (1997) conducted a study into the effects of negotiation types with 16 young learners of Dutch. Whereas the aforementioned research had aimed to identify occurrence, effects on linguistic accuracy and effects on syntacticization, Van den Branden's study focused on how task dynamics (in terms of people and feedback) affected the type and impact of modified output. The task involved a verbal picture description communication task relating to a murder mystery context. Interestingly, the push was provided by not allowing participants to see the person with whom they were conversing and negotiation types were divided amongst two groups: some non-native students would partake in peer interaction in which they were partnered with "native speaker friends of theirs" whereas the other dyads communicated with a teacher (Van den Branden, 1997, p. 602). It was discovered that negotiation of output was predominantly focussed on meaning with no deliberate attention to form being identified. Furthermore, output modifications were found once again to be influenced by the feedback type supplied to them (as in Pica et al.'s 1989 study) but were not determined by the person who provided it. Additional analysis of pre and posttest data revealed that there was a delayed effect upon output production; this was attributed to the feedback students received and the amounts of modified output they produced (Van den Branden, 1997). The study suggested that modified output was affected mostly by the way feedback was provided and not by the people involved in the communication. Also, since many modifications concentrated on meaning, the results could bring into question Linnell's (1995) findings, which identified a link between modified output and syntax development.

More recently, Sheen (2008) presents an extension to the debate surrounding modified output, the form of negative feedback and their effects on SLA. Following previous investigations into the effects of recasts on SLA, Sheen (2008) identified that not only did findings appear contradictory, but also that earlier studies consisted of several variables, each able to provide adequate explanations for the contrasting results. More specifically, she declared that research into the effects of individual learner factors upon SLA, output and feedback seemed neglected; in particular, the effect of language anxiety. In an attempt to discover the impact of language anxiety on English article use following recast provision and whether that anxiety influences the amounts of modified output, 45 English learners were divided into four groups (high anxiety with recasts, low anxiety with recasts, high anxiety without recasts and low anxiety without recasts) and participated in a pretest, treatment, immediate posttest, delayed posttest design. Results showed that the low anxiety with recasts group "outperformed" its high anxiety counterpart and the low anxiety control group in article use gains (Sheen, 2008, p. 835). Although no significant difference was found in article use between the two high anxiety groups, a noteworthy finding revealed that the students with low language anxiety who received recasts produced more modified output. The conclusions that recasts appeared effective for learners suffering low anxiety and that they also generated more modified output could substantiate Swain's COH assertions as well as pose important theoretical and pedagogical questions regarding pushing learners with elevated anxiety levels during communicative activities in the L2.

The studies reviewed here do not represent every aspect of COH research but rather reflect this study's objectives. However, the research described does display noteworthy findings suggesting that pushing students in their language production can produce more modified output, may promote gains in accuracy and may reinforce awareness of meaning-form relationships in a second language. However, since Nobuyoshi and Ellis's (1993) study was exploratory and Linnell (1995) and Van den Branden (1997)'s findings appeared to be contradictory, further research is still needed.

Research Gaps

In terms of previous research, this study will be different in three key areas: 1) the setting for the research, 2) the focus for the research and 3) the nature of the research.

Much COH research has taken place in immersion programmes in (e.g., Swain, 1985, 1998; Swain & Lapkin, 1995) and American ESL programmes (e.g., Sheen, 2008; Shehadeh, 1999), there has been little research in regard to comprehensible output within a UK EAP context. Therefore, we felt it would be interesting to see how the hypothesis' claims would be supported or refuted by research within this setting. Additionally, numerous studies have focussed on output in writing (e.g., Cumming, 1990; Donald & Lapkin, 2001; Hanaoka, 2007), and their conclusions cannot be truly applied to spoken output. It is important to study spoken output because in our experience many students place a high premium on enhancing this skill. This is especially true within a UK university setting because students need to improve speaking skills to undertake daily tasks both in and around the academy but it is unlikely that their output is pushed outside of English language classes. Finally, the nature of this study is rather different because it employed mixed-methods design. Quantitative methods were used to examine performance of the chosen linguistic structure and

occurrence of LREs, while a qualitative dimension was incorporated into the findings of the study to interpret the types of LREs produced. This combination was chosen to explore COH from both learning and teaching perspectives.

Methodology

Participants

The sample for the study consisted of 21 Upper Intermediate B2 level learners, as defined by the Common European Framework of References for Language (CEFR) (Council of Europe, 2001), from an English Elective Programme (EEP) at a UK university. Learners on this programme take general English classes as in-sessional support, while undertaking undergraduate or postgraduate courses. Student level was measured through standardised placement test scores. The sample consisted of Chinese L1 learners from one intact class on the EPP, randomly assigned to the pushed and non-pushed treatment groups. There were 11 male learners and 10 female learners; ages ranged from 18 to 34. At the time of the study, learners had been in the UK for a period of six months. The relatively small sample size was chosen due to the desire to include qualitative post-treatment data from each student in the study and in-depth analysis of participants' LREs. It is comparable to the sample sizes used in some studies in this area which have undertaken a similar analysis of a particular feature of the COH (e.g. Sheen, 2008).

Study Design

Firstly, students were tested one week prior to the treatment in their ability to use the selected past forms in a pretest. Selecting a suitable linguistic structure as the focus of this study was crucial if the results yielded were to be meaningful. The structure needed to be one which the students were aware of receptively but did not have productive mastery of. It was felt that the past forms used in narratives would be suitable because Chinese learners of English particularly display errors in this area due to the differences between the two language systems (Jung, 2001).

Following the pretests, both groups then took part in a storytelling task, where they were asked to tell a short story based on a set of pictures. Pushed students were instructed to tell the story using their own words but they were also informed that at times the teacher may interrupt them or ask a question. Teacher intervention was used upon hearing an inaccurate use of a past narrative form (past simple, past continuous, past perfect simple or past perfect continuous) and took the form of error repetition. Repetition can be defined simply as when an interlocutor "repeats the student's ill-formed utterance, adjusting intonation to highlight the error" (Lyster 1998, p. 189) and requires the students to assess their own language. As no explicit clues are given the student must search their own linguistic resources to modify their output. The control group undertook the same storytelling exercise but received no feedback from the teacher.

All students were video recorded during the storytelling task to facilitate post task analysis of LREs and also to provide learners with an appropriate memory aid during a stimulated recall activity. Stimulated recall was selected as a technique at this stage since

> *a subject may be enabled to relive an original situation with great vividness and accuracy if he is presented with a large number of cues or stimuli which occurred during the original situation (Bloom, 1954, p. 25).*

The qualitative data obtained was to uncover the types of LREs produced and why they were produced.

Finally, a posttest was performed on each student to see if their accuracy with the chosen linguistic structures had improved. The posttest was conducted a week after the treatment had taken place.

Data Analysis

In order to answer research questions one and two, video data were analysed to identify, classify and calculate LRE occurrence. Qualitative data was also examined by coding, labelling and grouping themes which emerged so that findings for research question two could be enhanced (Cresswell & Plano Clark, 2007). A Language Related Episode is defined as

> *any part of a dialogue where the students talk about the language they are producing, question their language use, or correct themselves or others (Swain & Lapkin, 1998, p. 326).*

Once results were compiled, statistical significance was calculated for LRE frequency. In order to analyse data for research question three, pre and posttest differences in accuracy were also tested for statistical significance.

Results and Discussion

Research Question One:

Does a pushed speaking task result in more language related episodes (LREs) than a non-pushed task for adult upper intermediate learners at an HE institution in the UK?

In total, pushed students produced 87 LREs, whereas non-pushed students produced only 24 LREs (78.38% and 21.62% of the total 111 LREs, respectively). An independent samples t-test was used to obtain means for this LRE data and to check for statistical significance. Table 1 provides descriptive statistics for LRE numbers for pushed and non-pushed students. The mean number of LREs for pushed students was calculated to be approximately 8 per student but the mean for non-pushed students remained at 2.4 per student.

The results of the t-test were then analysed to see if the findings were statistically significant. The t-test information is presented in Table 2.

The result for this t-test shows that the results are highly significant. We can therefore conclude that, in this study, pushing students during speaking tasks *does* result in higher numbers of LREs than not pushing them. However, since students in the non-pushed group did not receive any teacher initiated LREs it was clear that the previous findings could be criticised if we were to suggest that students were always going to present more LREs if

teacher initiated episodes were to be included. Therefore, another independent samples t-test was conducted, which included only LRE data from learner initiated LREs. This attempted to establish whether a link could be found between pushing students and LRE numbers *excluding* those resulting from direct teacher intervention.

Table 1. Descriptive statistics for LREs

		Levene's test for equality of variances		T-test for equality of means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean difference	Std. Error difference	95% CI of the difference	
									Lower	Upper
LRE number	Equal variances assumed	4.06	.058	5.03	19	.000	5.51	1.09	3.21	7.80
	Equal variances not assumed	--	--	5.19	14.64	.000	5.51	1.06	3.24	7.78

Table 2. Results of LRE t-test

	Group	*N*	*M*	*SD*	*SEM*
LRE number	Pushed	11	7.90	3.14	.95
	Non-pushed	10	2.40	1.50	.48

Table 3. Descriptive statistics for learner-initiated LREs

	Group	*N*	*M*	*SD*	*SEM*
LRE number	Pushed	11	4.18	2.27	.68
	Non-pushed	10	2.40	1.50	.47

Table 4. Results of learner-initiated LRE t-test

		Levene's test for equality of variances		T-test for equality of means						
		F	Sig.	T	df	Sig. (2-tailed)	Mean difference	Std. error difference	95% CI of the difference	
									Lower	Upper
LRE number	Equal variances assumed	1.64	.22	2.09	19	.05	1.78	.85	.00	3.56
	Equal variances not assumed			2.14	17.46	.04	1.78	.83	.02	3.53

In this case pushed students accounted for 46 learner initiated LREs whereas non-pushed students accounted for approximately half of this total at 24 LREs, representing 65.7% and 34.3%, respectively, of the total 70 LREs. Tables 3 and 4 show the descriptive statistics and t-test results for the second set of data.

Whilst mean LRE occurrence for pushed students stood at just over 4 episodes per student, the mean for non-pushed students was approaching half this figure at 2.4 episodes per student. The results of the independent samples t-test also reveal that once again, pushed tasks produced significantly better LRE numbers (p = .05) and we can therefore suggest that that pushing students does have a direct effect upon the number of LREs which are solely initiated by the learner.

Research Question Two

In what ways do pushed and non-pushed learners differ in the type and the success of LREs they display?

After establishing the link between pushed students and increased LRE totals in the story-telling task, data was then analysed according to LRE type. This would help to establish whether pushed and non-pushed students differed in the linguistic features to which they attended during the speaking task and also if they varied according to how successful they were. In order to answer the second research question, the LREs were categorised and then totalled to see how they were distributed across the various classifications, which were as follows:

LRE Categories

A. Learner initiated questioning of meaning of a linguistic term
B. Teacher initiated questioning of meaning of a linguistic term
C. Learner initiated questioning of the correctness of the spelling/pronunciation of a word
D. Teacher initiated questioning of the correctness of the spelling/pronunciation of a word

E. Learner initiated questioning of the correctness of a grammatical form
F. Teacher initiated questioning of the correctness of a grammatical form
G. Learner initiated correction of their own or another's usage of a word, form or structure
H. Teacher initiated correction of their own or another's usage of a word, form or structure

The table below shows this data.

Table 5. LRE distribution across categories

	LRE Type																
Group	A	%	B	%	C	%	D	%	E	%	F	%	G	%	H	%	Total LREs
Non-pushed	9	37.5	0	0.00	0	0.00	0	0.00	0	0.00	0	0.00	15	62.5	0	0.00	24
Pushed	8	9.20	10	11.49	1	1.15	0	0	3	3.45	1	1.15	34	39.08	30	34.48	87

It is clear that all students, regardless of group, produced the majority of their LREs due to some form of correction (LRE codes G and H). Although the two groups appeared rather similar in this factor, further investigation of the data revealed additional findings. First of all, totals for pushed students were greater than those displayed by non-pushed students. The LRE total for codes G and H stood at 64 for all pushed students (constituting approximately 74% of their LRE total) whereas non-pushed students had only a total of only 15 LREs (comprising 62.5% of their LRE total). This shows that pushed students appeared to concentrate much more frequently on correcting or modifying their output than non-pushed students, as we might expect.

To make comparisons more balanced, the data were also examined to determine the differences for pushed and non-pushed students regarding learner initiated correction (LRE code G) only. Although it is evident that pushed students' LREs would still be affected by the presence and intervention of the interlocutor, this was one way in which self-monitoring of learner output for each group could be compared. This particular LRE code constituted 62.5% of non-pushed students' LREs but pushed students only had a total percentage of 39.1%. This clear difference shows that whilst pushed students produced greater numbers of LREs and higher amounts of correction, they did not appear to monitor their own speech internally as much as the non-pushed students had appeared to have done. To find possible explanations for this, stimulated recall data was examined to find what students were thinking when the interlocutor was silent. This revealed interesting attitudes. Of the nine occasions silence was discussed by learners in the non-pushed group, 33% said they believed the silence represented the interlocutor giving them an opportunity to continue speaking, as shown in this example:

> T: So at the moment in the video, I haven't said anything. What were you thinking?
> S: You are wait for I think how to say
> T: So I'm waiting for you?
> S: Yeah
> (Student 5: Non-pushed)

Conversely, 36% of pushed students declared that silence was a signal that their output was correct and error free as seen in the following extracts from pushed students' stimulated recall data:

> T: Were you thinking anything here because I wasn't speaking?
> S: Yes, I thought I suppose you would stop me
> T: OK, and because I wasn't speaking, what were you thinking?
> S: Maybe I'm right
> (Student 6: Pushed)

This difference in opinion could provide a possible explanation for why learner initiated correction was less frequent (in terms of total LRE percentage) for pushed students than for non-pushed students. Due to the feedback pushed students received regarding their use of the past narrative forms, they may have assumed that a lack of feedback signified that their language production was correct.

From this we can gather that whilst frequencies for LRE code G were much more numerous for pushed learners, it appears that there is no positive effect between pushing students in their output and a higher rate of correctly resolved LREs. This means that whilst Swain's claims regarding more frequent processing of language and the production of modified language might be supported, one cannot assume that the modified language is always correct and nor is one type of task more effective than the other in this factor.

Language Related Episodes Pertaining to Meaning

The second most repeated type of LRE presented itself in learner attention to meaning. On inspection of the data, there appears to be a stark comparison between the percentages with which pushed and non-pushed students contemplated meaningful aspects. Table 6 presents a clearer representation of the distribution for meaning-related LREs (LRE A = learner initiated questioning of meaning of a linguistic item, B = teacher initiated questioning of meaning of a linguistic item) for pushed and non-pushed students.

We can see from these results that although the total frequency for pushed students' meaning-based LREs was exactly double that of non-pushed students' (18 and 9 respectively), the percentages revealed that in respect to their total LRE numbers, the total for pushed students only represented a fifth of all LREs whereas the number approximated more than a third of total non-pushed LREs (20.69% for pushed students and 37.5% for non-pushed

students). This shows that whilst only pushed students may have received more feedback regarding their output, they did not attend to features of meaning as much as non-pushed students. However, when submitted to a paired samples t-test, the differences were not found to be statistically significant. The result (p = .058) was slightly above the threshold of statistical significance despite a seemingly large effect size (r = 0.94). Therefore, we cannot assume that either task was more effective in terms of stimulating meaning-based attention to output.

Table 6. Distribution of meaning-related LREs according to task type

Group	A	%	B	%	A + B	%	Total LREs
All non-pushed students	9	37.5	0	0	9	37.5	24
All pushed students	8	9.19	10	11.49	18	20.69	87

In terms of *learner initiated* focus on meaning (LRE code A), we can see that frequencies are relatively equal for both groups. However, in terms of percentages of total LRE numbers, LRE code A represented around a tenth of total LREs for pushed students but symbolised 37.5% for non-pushed students. The results of a paired samples t-test showed this to be highly statistically significant (p = .02). This shows that students who receive no push in their output are more likely to attend to features of meaning using their own linguistic monitoring than pushed students.

After discovering differences between correction and meaning-based features of output for pushed and non-pushed learners, it seemed relevant to investigate what features students attended to during corrective LREs. The previous section suggested that non-pushed students concentrated on meaning-based features more frequently than pushed learners so LRE codes G (learner initiated correction) and H (teacher initiated correction) were examined to see how output was modified and to see if this claim would be substantiated further. Each LRE was then coded according to aspects of form, meaning or both.

The following two tables display the results of this analysis according to group (pushed students were also assessed for LRE H):

Table 7. All non-pushed corrective LRE types

Group	G LREs	Meaning	%	Form	%	Aspects of both	%
Non-pushed	15	11	73.33	3	20	1	6.67

Table 8. All pushed corrective LRE types

Group	G LRE total	H LRE total	Meaning			Form			Aspects of Both		
			G	H	Total G&H	G	H	Total G&H	G	H	Total G&H
Pushed	34	30	8 23.53%	4 13.33%	12 18.75%	24 70.59%	24 80.00%	48 75%	2 5.88%	2 6.67%	4 6.25%

These figures present a clear comparison of the linguistic aspects that non-pushed and pushed students attended to. Whereas non-pushed students attended to meaning for approximately three out of every four LREs, meaning only accounted for approximately every one of five corrective LREs for pushed learners. Conversely, form was the focus for non-pushed students for only 20% of LREs but pushed students looked at aspects of form for 75% of their LREs. In terms of correction without feedback from the teacher (LRE G), non-pushed students again were inclined to amend form for 20% of their LREs but pushed students adjusted their form in 71% of occasions. This undoubtedly confirms that pushed students attend to form more frequently than to meaning-based items. It would therefore appear to support Swain's claims that pushing students can raise their awareness of the importance of form when conveying meanings in a target language. Her claim that creating a push may enhance syntactic processing and also acquisition would also seem to be substantiated by the findings, in conjunction with Schmidt's (1994) views on consciousness and cognition, since many students also appeared aware of any changes they made when producing modified output.

Stimulated recall data regarding learner thoughts during periods of self-correction (LRE G) supports this claim. Although attitudes regarding learner initiated correction did vary, there appeared to be a consensus that correction took place following increased attention to accuracy or due to an innate feeling in the students' minds that output was flawed. These views represented 7 out of 12 responses (58%) for pushed and non-pushed learners regarding self-correction and included comments such as these, where a learner explains that repetition of the same error and information from the title allowed them to realise they could correct their own past tense error:

> (Regarding "He can't go" changed to "He couldn't go went")
>
> T: Ok there I didn't say anything but what were you thinking?
> S: I think I…I'm paying more attention about the tense. I can change it by myself

T: You changed it by yourself ok. How did you know you needed to change it?
S: Before that I always make the same mistake and the pictures shows me the title is last week
(Student 7: Pushed)

These views were just some of those that were typical of those referring to what instigated correction. Whilst some learners implied the use of an instinctive feeling for linguistic correctness, advocating the use of monitoring within Levelt's Model of Speech Production (1989) and indeed Krashen's notion that acquired language is not reliant or dependent upon direct language instruction, other learners did mention that much of their correction was the result of greater attention to form. Although this is by no means exclusive to pushed students, it did represent the views of 37.5% of pushed students' responses regarding correction whereas only one non-pushed student expressed a similar opinion.

Research Question Three

Does a pushed speaking task result in greater accuracy gains pre and posttest results for the past simple, past continuous, past perfect simple and past perfect continuous?

After analysing the effect of pushed and non-pushed speaking tasks on LRE numbers, LRE type and attention to form and meaning, it is important to assess their effect upon the past narrative tense performance of the students. In order to do so, data from the pretest and posttest were examined to observe the effect of the treatment on the accurate use of the target forms. These results are displayed in the table below.

Table 9. Average pre- and posttest scores for all students

Tense	All students			Pushed			Non-pushed		
	Pre-test %	Post-test %	% Difference	Pre-test %	Post-test %	% Difference	Pre-test %	Post-test %	% Difference
Total score	52.07	65.00	12.93	51.10	66.02	14.92	53.26	63.76	10.50
Past simple	84.29	81.85	-2.44	83.77	79.80	-3.97	84.92	84.36	-0.56
Past cont	26.88	60.00	33.12	25.00	67.27	42.27	29.17	51.11	21.94
Past perfect S	19.17	30.17	11.00	18.18	38.96	20.78	20.37	20.63	0.26
Past perfect cont	0.00	1.67	1.67	0.00	3.03	3.03	0.00	0.00	0.00

From the table above, we can see that total percentage scores for both groups of students did increase. The scores of non-pushed learners increased by 10.50% whilst those of pushed students increased by approximately 15%. While this would suggest that all students improved in the use of the target past narrative form use following the storytelling task, the results show that this is not the case. Interestingly, for all students, both pushed and non-pushed, there was a detrimental effect upon past simple tense performance. Between the pretest and posttest, the mean percentage for all students, regardless of group, decreased by 2.44%, with pushed students displaying a mean percentage difference of -3.97% and non-pushed students -.056%. However, examining all the test data and learners' use of the past continuous and past perfect simple, would suggest that pushing students in their language production can have beneficial effects. Whereas non-pushed students achieved a notable mean percentage increase of 21.94%, pushed students accomplished an increase of 42.27% between the pre and posttest for the past continuous Furthermore, in terms of the past perfect simple, non-pushed students remained at a similar level (demonstrating an increase of only 0.26%) whereas pushed students revealed a considerable improvement of 20.78% between pretest and posttest results. Whilst it could be suggested that not pushing students can result in some degree of improvement, as an increase of nearly 22% for non-pushed students in the past continuous shows, it is clear that it is an effect which was not repeated for the other forms. Unfortunately, in terms of the past perfect continuous, all students performed badly and bar one exception showed no improvement and students consistently overlooked this tense. Perhaps this could be an indication that Ellis's (1994, p. 284) view that comprehensible output may not "result in the acquisition of *new* linguistic features" may be right.

These results reveal that whilst some of the percentage increases and decreases appear substantial when pushed and non-pushed learners were compared, none of the differences were statistically significant. The figure closest to statistical significance related to the use of the past continuous tense by the pushed group (p = .073) but this data still demonstrates that results of this study do not indicate that the pushed storytelling task resulted in greater accuracy which was significant in individual past narrative forms. Whilst this result is disappointing, the higher scores of the pushed group, combined with a significantly higher number of LREs suggest that the process of noticing and acquiring more accurate use of these forms has at the least been started by the use of a pushed task.

Conclusion

This study aimed to explore how pushed and non-pushed speaking tasks can vary in effectiveness within a setting with upper intermediate students of English. Pushing students in their spoken output was found to have a significant positive effect upon the number of LREs they produced. Non-pushed students were found to have produced only a quarter of the total number of LREs (for both groups) when teacher initiated and learner initiated episodes were combined. This shows that delivering a 'push' could provide the impetus required for students to assess their output more frequently. Furthermore, analysis of learner-initiated LREs was undertaken to discover if pushing students had a direct effect upon internal monitoring of output by learners. There was also a significant positive effect upon the number of learner-

initiated LREs when students were pushed in their spoken output. This showed that pushed output tasks can increase instances of linguistic processing by learners even when the interlocutor offers no direct feedback. Stimulated recall data also clearly shows that LREs produced by students in the pushed group made them pay more attention to form perhaps prompting them to notice the gaps in their own output**.** The findings of this study may therefore be seen as an expansion of those made by Nobuyoshi and Ellis's (1993) exploratory study into task types. Whilst their data provided a basis for suggesting pushed tasks encouraged more linguistic processing, this study has shown on a larger scale that pushed spoken output tasks can be directly linked to elevated totals of linguistic episodes.

The third research question intended to ascertain whether either type of task resulted in greater gains in past narrative form accuracy. Initial analysis of pre and posttest data revealed that both groups of students improved in their total test score percentages: non-pushed students improved by 10.5% and pushed students improved by 14.9%. This was not found to be statistically significant so no claim can be made that pushing learners in their output led to higher gains in accuracy. Past simple percentages were found to fall for both groups with data for the past continuous demonstrating improvement for both groups, although pushed students had superior percentage gains.

These findings could have several implications for the teaching of second languages and also for further research. One implication would be the application of this knowledge into EAP teaching practice. The findings suggest that pushed speaking tasks can benefit learners in this context because of the functions comprehensible output supplies. Although much research in EAP settings has placed an emphasis on reading and writing, Cornbleet argues that speaking "… is no less important for university life than formal academic work" (Cornbleet, 2000, p. 32). Regular pushed output tasks, targeted towards areas of speech which most learners will need could be used to supplement EAP classes in more traditional areas such as essay writing. It seems sensible to use pushed communicative tasks to achieve this because learners are required to use English to communicate on a daily basis but are unlikely to be offered much in the way of structured feedback outside of class time.

This study could also prompt further research into pushed and non-pushed speaking tasks. Clearly, there are limitations because of the sample size used. Although we have sought to justify this in terms of the study's scope, one suggestion for further research would be to expand the design so that the overall sample size can be increased. For instance, a larger sample within the course (the context for this study) may contain between 80 to 100 students (26-33% of the whole programme) which would constitute approximately 4 to 5 individual classes on that course. The increased sample and greater variation provided by investigating numerous classes could increase the relevance of findings to the larger population. Further studies focussing on different linguistic features would also expand existing research (e.g. Linnell, 1995; McDonough, 2005; McDonough & Mackey, 2006; Nobuyoshi & Ellis, 1993). Furthermore, the study could also be conducted over an extended time-frame to see whether related performance gains or losses are maintained over longer periods.

References

1. Allen, P., Swain, M., Harley, B. & Cummins, J., 1990. Aspects of Classroom Treatment. Towards a More Comprehensive View of Second Language Acquisition. In: B. Harley, P. Allen, J. Cummins & M. Swain, eds. *The Development of Second Language Proficiency.* Cambridge: Cambridge University Press, pp. 57–81. doi: http://dx.doi.org/10.1017/CBO9781139524568.007
2. Bloom, B. S., 1954. The Thought Process of Students in Discussion. In: S. J. French, ed. *Accent on Teaching: Experiments in General Education.* New York: Harper & Row, pp. 23–46.
3. Cornbleet, S., 2000. Illocutionary Force and the EAP Context. In: J. Cutting, ed. *The Grammar of Spoken English and EAP Teaching.* Sunderland: University of Sunderland Press, pp. 21–38.
4. Council of Europe, 2001. *The Common European Framework of Reference for Languages: Learning, Teaching, Assessment.* Available at: http://www.coe.int/t/dg4/linguistic/Source/Framework_en.pdf [accessed December 2013].
5. Cresswell, J. W. and Plano Clark, V. L., 2007. *Designing and Conducting Mixed Methods Research.* Thousand Oaks, California: Sage.
6. Cumming, A., 1990. Metalinguistic and Ideational Thinking in Second Language Composing. *Written Communication,* No. 7, pp. 482–511. http://dx.doi.org/10.1177/0741088390007004003
7. Del Pilar García Mayo, M., 2002. The Effectiveness of Two Form-focused Tasks in aAdvanced EFL Pedagogy. *International Journal of Applied Linguistics,* No. 12 (2), pp. 156–175. doi: 10.1111/1473-4192.t01-1-00029
8. Donald, S. & Lapkin, S., 2001. Exploring the Role of Noticing in a Three-stage Second Language Writing Task. *Journal of Second Language Writing*, No. 10 (4), pp. 277–303. doi: 10.1016/S1060-3743(01)00046-7
9. Ellis, R., 1994. *The Study of Second Language Acquisition.* Oxford: Oxford University Press.
10. Gass, S. M., 1997. *Input, Interaction and the Second Language Learner.* Mahwah: Lawrence Erlbaum Associates Inc.
11. Hanaoka, O., 2007. Output, Noticing, and Learning: An Investigation into the Role of Spontaneous Attention to Form in a Four-stage Writing Task. *Language Teaching Research*, No. 11 (4), pp. 459–479. doi: 10.1177/1362168807080963
12. Izumi, S., 2000. Does Output Promote Noticing and Second Language Acquisition? *TESOL Quarterly*, No. 34 (2), pp. 239–278. doi: 10.2307/3587952
13. Izumi, S., 2002. Output, Input Enhancement, and the Noticing Hypothesis. *Studies in Second Language Acquisition*, No. 24 (04), pp. 541–577. doi: 10.1017/S0272263102004023
14. Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S., 1999. Testing the Output Hypothesis. *Studies in Second Language Acquisition*, no. 21 (3), pp. 421–452.
15. Jung, C., 2001. Chinese Speakers. In: M. Swan & B. Smith, eds. *Learner English: A teacher's Guide to Interference and Other Problems.* Cambridge: Cambridge University Press, pp. 310-325.
16. Kim, Y., 2008. The Contribution of Collaborative and Individual Tasks to the Acquisition of L2 Vocabulary. *Modern Language Journal*, No. 92, pp. 114–130. doi: 10.1111/j.1540-4781.2008.00690.x
17. Kowal, M., & Swain, M., 1994. Using Collaborative Language Production Tasks to Promote Students' Language Awareness. *Language Awareness*, No. 3 (2), pp. 73–93. doi: 10.1080/09658416.1994.9959845
18. Krashen, S., 1982. *Principles and Practice in Second Language Acquisition.* London: Pergamon.
19. Krashen, S., 1989. We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *The Modern Language*

Journal, no. 73 (4), pp. 440–464.
doi: 10.1111/j.1540-4781.1989.tb05325.x

20. Leeser, M. J., 2004. Learner Proficiency and Focus on Form during Collaborative Dialogue. *Language Teaching Research*, No. 8 (1), pp. 55–81.
doi: 10.1191/1362168804lr134oa
21. Levelt, W. J. M., 1989. *Speaking: From Intention to Articulation.* Cambridge, MA: MIT Press.
22. Lewis, M., 1997. *Implementing the Lexical Approach.* Hove, England: Language Teaching Publications.
23. Linnell, J., 1995. Can Negotiation Provide a Context for Learning Syntax in a Second Language? *Working Papers in Educational Linguistics*, No. 11, pp. 83–103.
24. Long, M. H., 1983. Linguistic and Conversational Adjustments to Non-native Speakers. *Studies in Second Language Acquisition*, No. 5 (2), pp. 177–193.
doi: 10.1017/S0272263100004848.
25. Lyster, R., & Ranta, L., 1997. Corrective Feedback and Learner Uptake. *Studies in Second Language Acquisition*, No. 19 (1), pp. 37–66.
doi:10.1017/S0272263197001034
26. Lyster, R., 1998. Negotiation of Form, Recasts, and Explicit Correction in Relation to Error Types and Learner Repair in Immersion Classrooms. Language Learning, No. 48 (2), pp. 183–218.
doi: 10.1111/1467-9922.00039
27. McDonough, K., & Mackey, A., 2006. Responses to Recasts: Repetitions, Primed Production, and Linguistic Development. *Language Learning*, No. 56 (4), pp. 693–720.
doi: 10.1111/j.1467-9922.2006.00393.x
28. McDonough, K., 2005. Identifying the Impact of Negative Feedback and Learners' Responses on ESL Question Development. *Studies in Second Language Acquisition*, No. 27 (01), pp. 79–103.
doi: 10.1017/S0272263105050047
29. Muranoi, H., 2007. Output Practice in the L2 Classroom. In: R. M. DeKeyser, ed. *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology*. New York: Cambridge University Press, pp. 51–84.
30. Nobuyoshi, J. and Ellis, R., 1993. Focused Communication Tasks and Second Language Acquisition. *ELT Journal*, No. 47, pp. 203–210.
doi: 10.1093/elt/47.3.203
31. Pica, T., 1988. Interactive Adjustments as an Outcome of NS-NNS Negotiated Interaction. *Language Learning*, No. 38, pp. 45–73.
http://dx.doi.org/10.1111/j.1467-1770.1988.tb00401.x
32. Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L., 1989. Comprehensible Output as an Outcome of Linguistic Demands on the Learner. *Studies in Second Language Acquisition*, No. 11 (01), pp. 63–90.
doi: 10.1017/S027226310000783X
33. Ranta, L. & Lyster, R., 2007. A Cognitive Approach to Improving Immersion Students' Oral Language Abilities: The Awareness-Practice-Feedback Sequence. In: R. M. DeKeyser, ed. *Practice in a Second Language: Perspectives from Applied Linguistics and Cognitive Psychology.* Cambridge: Cambridge University Press, pp. 141–160.
34. Schmidt, R. W., 1990. The Role of Consciousness in Second Language Learning. *Applied Linguistics,* No. 11(2) pp. 129–158
35. Schmidt, R. W., 1994. Deconstructing Consciousness in Search for Useful Definitions for Applied Linguistics. *AILA Review,* No. 11, pp. 11–26.
36. Schmidt, R. & Frota, S., 1986. Developing Basic Conversational Ability in a Second Language. A Case Study of an Adult Learner of Portuguese. In: R. Day, ed. *Talking to Learn: Conversation in Second Language Acquisition*. Rowley, MA: Newbury House, pp. 237–326.
37. Selinker, L., 1972. Interlanguage. *IRAL-International Review of Applied Linguistics in Language Teaching*, No. 10 (1-4), pp. 209–232.
38. Sheen, Y. 2008. Recasts, Language Anxiety, Modified Output, and L2 Learning. *Language Learning*, No. 58 (4), pp. 835–874.
39. Shehadeh, A., 1999. Insights into Learner Output. *The English Teaching Forum,* No. 37, (4), pp. 2–5.
40. Shehadeh, A., 2002. Comprehensible Output, From Occurrence to Acquisition: An Agenda for Acquisitional Research. *Language Learning*, No. 52 (3), pp. 597–647.
41. Shehadeh, A., 2003. Learner Output, Hypothesis Testing, and Internalizing Linguistic Knowledge. *System*, No. 31 (2), pp. 155–171.
42. Soleimani, H., 2008. The Noticing Function of Output in Acquisition of Rhetorical Structure of Contrast Paragraphs of Iranian EFL University Students. *Linguistik Online*, No. 34 (2). Available at: http://www.linguistik-online.de/34_08/soleimani.html [accessed December 2013].
43. Swain, M. & S. Lapkin, 1995. Problems in Output and the Cognitive Processes They Generate: A Step Towards Second Language Learning. *Applied Linguistics*, No. 16, pp. 371–391.
http://dx.doi.org/10.1093/applin/16.3.371
44. Swain, M., & Lapkin, S., 1998. Interaction and Second Language Learning: Two Adolescent French Immersion Students Working Together. *Modern Language Journal*, No. 82, pp. 320–337.
http://dx.doi.org/10.1111/j.1540-4781.1998.tb01209.x
45. Swain, M., & Lapkin, S., 2001. Focus on Form through Collaborative Dialogue: Exploring Task Effects. In: M. Bygate, P. Skehan, & M. Swain, eds. *Researching Pedagogic Tasks: Second Language Learning, Teaching and Testing.* London, UK: Longman, pp. 99–118.
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.* Rowley, MA: Newbury House, pp. 235–253.
47. Swain, M., 1998. Focus on Form through Conscious Reflection. In: C. Doughty & J. Williams, eds. *Focus on Form in Classroom Second Language Acquisition.* Cambridge: Cambridge University Press, pp. 64–81.
48. Van den Branden, K., 1997. Effects of Negotiation on Language Learners' Output. *Language Learning*, No. 47, pp. 589–636.
http://dx.doi.org/10.1111/0023-8333.00023

Shelley Byrne, Christian Jones

Stimuliuojamos ir nestimuliuojamos kalbos užduotys EAP (*English for Academic Purposes*) kontekste: kaip jos pasitarnauja lingvistiniam apdorojimui ir kalbos tikslumui?

Santrauka

Šiame straipsnyje aptariamas mišrius metodus naudojantis tyrimas, nagrinėjantis stimuliuojamų ir nestimuliuojamų kalbėjimo užduočių efektyvumą JK universitete su „upper intermediate“ lygio studentais. Tyrime analizuojama a) ar stimuliuojama (užduota) kalbėjimo užduotis gali sukurti didesnį skaičių kalbinių epizodų (LREs), negu nestimuliuojama užduotis, b) skirtumai tarp LREs tipų, kuriuos nulemia kiekviena užduotis ir c) ar stimuliuojama kalbinė užduotis sąlygoja tikslesnį naratyvą būtojo laiko formomis. Gauti rezultatai parodė, kad stimuliuojamas siužeto atpasakojimas leido sukurti daugiau kalbinių epizodų, negu nestimuliuojamas. Be to, galima teigti, kad dažniausiai panaudoti tiek stimuliuojamų, tiek ir nestimuliuojamų kalbos epizodų tipai studentų turėjo būti taisomi. Stimuliuotoji grupė pasiekė didesnio tikslumo, tačiau statistiškai jis buvo gana nežymus. Darytina išvada, kad stimuliavimas kalbinėje veikloje gali padidinti lingvistinio apdorojimo galimybių skaičių, dėl to atsiranda ir vystosi „interlanguage“, t. y. šiek tiek iškreipta antrosios kalbos forma, atsirandanti dėl klaidingo gramatinių formų ar tarimo vartojimo.

Straipsnis įteiktas 2014 01
Parengtas spaudai 2014 06

About the authors

Shelley Byrne, Associate Lecturer and PhD candidate, University of Central Lancashire, UK.
Academic interests: second language acquisition, spoken language production, corpus-informed teaching and testing.
Address: University of Central Lancashire, Fylde Road, Preston, Lancashire, PR1 2HE, UK.
E-mail: sbyrne@uclan.ac.uk

Christian Jones, Senior Lecturer in TESOL, University of Central Lancashire, UK.
Academic interests: spoken discourse, corpus-informed language teaching, lexis and the pedagogical treatment of spoken grammar.
Address: University of Central Lancashire, Fylde Road, Preston, Lancashire, PR1 2HE, UK.
E-mail: Cjones3@uclan.ac.uk