UNIVERSITAT DE BARCELONA

Task complexity effects on the acquisition of an L2 vowel contrast: A task-based pronunciation teaching study

Ingrid Mora Plaza

M.A. Thesis

Supervised by: Dr. Joan Carles Mora and Dr. Roger Gilabert

Applied Linguistics and Language Acquisition in Multilingual Contexts

M.A. Program

Departament de Llengües i Literatures Modernes i Estudis Anglesos

Universitat de Barcelona

September 2017

Acknowledgements

This MA thesis would not have been possible without the help and support of many people to whom I would like to dedicate the following paragraphs.

First and foremost, I am deeply indebted to my two supervisors, Dr. Joan Carles Mora and Dr. Roger Gilabert, who have guided me from the initial stages of planning this project to the very last details of writing it. They have been incredibly helpful, efficient and dedicated. Moreover, it has been a privilege and a pleasure to benefit from their wide knowledge and expertise in the area of SLA, specifically, L2 speech acquisition and task-based language teaching, respectively. I really appreciate the time you have devoted to me.

Secondly, I am grateful to all the professors of the Applied Linguistics M.A. program who have shared with me their priceless knowledge and experience: Dr. Roger Gilabert, Dr. Joan Carles Mora, Dr. Carme Muñoz, Dr. Maria Luz Celaya, Dr. Elsa Tragant, Dr. Raquel Serrano, Dr. Imma Miralpeix, Dr. Sara Feijoo and Dr. Julia Barón. I cannot express with words how much I have learned in this intense academic year thanks to each of you. Also, my special thanks to Dr. Raquel Serrano, who has been very comprehensible and flexible with my working hours in the department as an internal collaborator, and Dr. Olena Vasylets for giving me very useful recommendations and being so helpful.

Thirdly, I would like to thank all my classmates, most of whom have become very close friends. Without all of them, this masters would not have been so enjoyable. All my love and most sincere gratitude to Kimberli Rodrigues, who has been very supportive and has very kindly accepted to be my second inter-rater. I am also grateful to Vanessa Robinson, Nicole Elliot, Tom Flaherty and David Hetherington, who volunteered to record all the stimuli for my tests. Finally, Daniela Avello, Luke McCarthy, Jacey Cargill, Irina Stan, you will always be in my heart.

Fourthly, I owe my gratitude to my closest friends, who have encouraged me during this process and offered me company, love and laughs. I owe a great deal to Claudia López and Isabel Archs for their help in the design of the task illustrations. You are talented artists! In addition, I must thank Alba Montaño, not only for being my third inter-rater but also for her valuable advice and her wholehearted support. You are an older sister to me.

Fifthly, I am very thankful to the Escola La Mercè (Martorell) for allowing me to collect data from their students. I cannot thank the school director, Montse Font, as well as the teachers, Maria Pujol and Zaira Álvarez enough. Without their help and their great flexibility with schedules, this research would have been impossible.

Last but not least, the most deserved acknowledgement goes to my parents, Frederic and Rosa, and my sister, Diana, who strongly believed in my work and did not hesitate to help me anytime. Finally, thank you Matteo Radice for your words of encouragement, endless patience and unconditional love.

Barcelona, September 2017

Abstract

L2 pronunciation is often neglected in the EFL classroom and, when addressed, it is typically decontextualized from communicative practice. Additionally, limited research has been conducted in SLA on the role of task manipulation for the improvement of L2 pronunciation accuracy during meaning-focused interaction. This study investigates the impact of decision-making tasks, organized in increasing complexity, on the perception and production of English /æ/-/ʌ/ in order to improve learners' pronunciation in foreign language exchanges. L1 Catalan/Spanish young adults (*n*=18) performed four dyadic problem-solving, reasoning-gap tasks over a three-week period. Tasks were always preceded by form-focused pre-tasks that contained lexical items contrasting the target vowels (e.g., *bag-bug, cap-cup*) to be used during task performance. Furthermore, tasks were sequenced on the basis of increasing level of cognitive complexity (+S, -S, -C, +C) in order to progressively enhance the occurrence of pronunciation-based language-related episodes. Perception and production accuracy were pre- and post- tested through identification and ABX discrimination tasks and a delayed-sentence repetition task, respectively. Individual differences in learners' L2 proficiency and attention control were also assessed. In line with the Cognition Hypothesis (Robinson, 2001, 2007, 2011), the results revealed that orienting attention to a phonological contrast during interactive tasks improves its perception and production significantly, and increased task demands along resource-directing variables (i.e. +/- reasoning demands and +/- elements) generate more pronunciation-focused LREs. Finally, auditory selective attention was the main moderator factor in explaining inter-subject variability in the perception and production of the English vowel contrast.

Keywords: pronunciation instruction, L2 vowel contrast, task-based pronunciation teaching (TBPT), focus on phonetic form, task complexity, language-related episodes (LREs), L2 proficiency and attention control.

Resumen

La pronunciación de la L2 no suele ser prioritaria en las clases de inglés como lengua extranjera y, cuando se aborda, queda normalmente descontextualizada de otras prácticas comunicativas. Además, hay escasa investigación en el campo de la adquisición de segundas lenguas sobre cómo afecta la manipulación de actividades a la mejora de la pronunciación de la L2, cuando ésta forma parte del contenido presente en la interacción. Este estudio investiga el efecto de tareas, en las cuales hay que tomar decisiones superando niveles de complejidad, en la percepción y la producción de los fonemas ingleses /æ/-/ʌ/ con tal de conseguir la inteligibilidad de los estudiantes durante la comunicación en la lengua extranjera. Jóvenes catalanes y españoles (*n*=18) llevaron a cabo durante tres semanas cuatro tareas en pareja donde tenían que resolver problemas a través del razonamiento. Las tareas estaban siempre precedidas de otras tareas previas centradas en el lenguaje y que contenían las vocales principales (p. ej. *bag-bug, cap-cup*) las cuales iban a ser utilizadas durante la resolución de éstas. Asimismo, las tareas estaban ordenadas por escala de complejidad (+S, -S, -C, +C) para provocar, progresivamente, episodios relacionados con el lenguaje y basados en la pronunciación. Se evaluó la precisión en la percepción y la producción de estos fonemas antes y después del tratamiento a través de tareas de identificación y discriminación (ABX) así como una tarea de repetición de frases con acción retardada. También se valoró el dominio lingüístico de la L2 y el control de la atención. De acuerdo con la Hipótesis Cognitiva (Robinson, 2001, 2007, 2011), los resultados revelaron que orientar la atención hacia el contraste fonológico durante tareas interactivas mejora significativamente su percepción y producción, y un aumento del foco de atención (p. ej. +/- razonamiento y +/- elementos) genera más episodios relacionados con el lenguaje (*LREs*) y, específicamente, la pronunciación. Finalmente, la atención selectiva auditiva fue el principal factor moderador que marcó diferencias entre individuos en la percepción y la producción del contraste vocálico en inglés.

Palabras clave: enseñanza de la pronunciación, contraste vocálico de la L2, enseñanza de la pronunciación centrada en las tareas (*TBPT*), enfoque en la pronunciación, complejidad de la tarea, episodios relacionados con el lenguaje (*LREs*), dominio lingüístico de la L2 y control de la atención.

Table of contents

List of figures

List of tables

[1] T1: Pre-test
T2: Post-test
T3: Delayed post-test

1. Introduction

Research in second language acquisition has shown that L2 sounds which are non-existent in the learners' L1 -but are similar to the learners' native phonetic inventory- are difficult to perceive and produce (Flege, 1995). As a matter of fact, the learning of L2 phonological representations requires practice through long periods of exposure to the foreign language. Nevertheless, in school contexts, which is the instructional setting of the current study, input is not always present and it is also limited outside the classroom (Muñoz, 2008). Apart from the lack of linguistic experience in the FL environment, L2 pronunciation is conceived as one of the most challenging skills to be taught and learned in the EFL classroom.

According to Murphy & Baker's (2015) historical overview of the teaching of pronunciation, none of the methodologies appeared to be effective enough due to an interplay of factors such as old methods, outdated materials, lack of teacher training, among others. In actual fact, only 30% of ESL programs in Canada offered phonetics and phonology courses, and TESOL programs around the world only taught metalanguage-heavy classes which did not address the practical applications of L2 pronunciation (Isaacs, 2009). After the Cognitive Approach and the Natural Approach in the 60s, and the Silent Way in the 70s, pronunciation was reincarnated in a broadly-constructed communicative approach which encouraged prioritized pronunciation instruction rather than attainment of a native-like accent (Morley, 1991, p.490). Some speech characteristics that affected intelligibility in communication were word boundaries such as linking sounds (i.e. *go in* /gəwɪn/), sound mergers (i.e. *nice shoe* /naɪʃuː/) and composite sounds (i.e. *this year* /ðɪʃɪə/) (Kenworthy, 1987; op cit Mora, 2016, p. 17) as well as phonological contrasts such as /æ/ vs. /ʌ/ (e.g. *cap/cup*) or /iː/ vs. /ɪ/ (e.g. *sheep/ship*), among others[2].

Nowadays, high-variability phonetic training (HVPT) has proved to be effective for the acquisition, retention and generalization of L2 sounds (Cebrian & Carlet, 2014, 2015); however, it lacks a communicative component because it is usually performed individually on drill-like tasks in phonetic laboratories. In contrast, task-based language teaching encourages interaction following an analytic focus on form approach but has rarely dealt with the area of L2 pronunciation. In order to bridge the gap between L2 speech acquisition and TBLT research, the present study advances a pre-/post-test design inspired by HVPT testing method but uses

[2] Intelligibility can be defined as "the extent to which a speaker's message is understood by a listener" considering that understanding entails "recognizing and grasping the meaning of most or all the individual words that the speaker has produced" (Munro and Derwing, 1995; op cit Munro & Derwing, 2015, p.379).

tasks in increasing cognitive complexity as the central element for the acquisition of the phonological contrast. As a result, the present study investigates the effects of task complexity on the manifestation of language-related episodes (LREs) and the repercussion of learner factors on L2 outcomes. Results are interpreted in light of the Cognition Hypothesis (Robinson 2001, 2007, 2011).

2. Literature review

2.1. The role of input, noticing and attention in learning L2 sounds

According to Schmidt (1995), there is no learning without attention so unattended stimuli persist in immediate STM for only a few seconds at best but it cannot be stored in the LTM without sufficient attention[3]. Van Pattern (1994; op cit Schmidt, 1995) argues that learners may detect everything in the input but attention to specific phonological items is necessary in order to encode their information. Nevertheless, selective attention may not be required when items occur with one or more features that match an already existing representation in the L1 (Long, 2015). Attention can be involuntarily attracted to certain stimuli so learners may not need intentional focus in order to learn L2 items; however, "some aspects of the L2 input are so subtle and abstract that they cannot possibly be attended to" (Schmidt, 2001, p.30) with only incidental learning.

Concerning L2 speech learning models, Flege's (1995) Speech Learning Model (SLM) contends that the speaker's phonological system remains malleable over the life span and can be modified depending on the distance between L1 and L2 phonetic categories and the amount of input. The model posits that new categories can be created for L2 sounds if the L2 phonemes are perceived as dissimilar to the L1 phonemes (e.g. Eng. /ʃ/ vs. Sp. /s/). Conversely, if L1 and L2 categories are similar (e.g. Eng. /i:/ vs. Cat. /i:/), L2 sounds are mapped onto L1 categories, thus, learners do not have any problem acquiring them. Finally, Flege (1995) states that learners' production of a specific sound eventually corresponds to the properties represented in its phonetic category representation. Another theory is the Perceptual Assimilation Model (Best & Tyler, 2007) which claims that learners do not create new categories for new sounds but they

[3] STM: Short-term memory
LTM: Long-term memory

assimilate them according to the acoustic cues of their L1. Therefore, learners' ability to discriminate L2 contrasts depends on how sounds are assimilated to L1 representations: (a) two-category assimilation, (b) single-category assimilation and (c) category-goodness assimilation. The current study presents an instance of a single-category (SC) assimilation, where learners tend to perceive two non-native sounds (Eng. /æ/ & /ʌ/) as one native sound (Sp. /a/).

Whereas native speakers have a great awareness of speech as a sequence of sounds and develop a high degree of sensibility to speech contrasts (Piske, 2008), L2 learners struggle to notice foreign phonemes and establish new categories. Guion-Anderson & Pederson (2007) propose that directing learners' attention to phonetic aspects help acquiring difficult contrasts; therefore, pronunciation-based instruction may be beneficial for learners to 'notice the gap' between L1 & L2 phonetic categories and produce more accurate pronunciation.

2.2. Focus on forms, focus on meaning and focus on form

Phonetic training has always been very useful for the acquisition of foreign phonemes, especially in situations where experience with the target language is limited. Cebrian and Carlet (2014) -among others- show that HVPT may be very efficient to train new phonetic categories and generalize them to new contexts (Pereira, 2014); however, it lacks communicative peer interaction, which has been shown to be conductive to L2 learning.

Concerning group-oriented teaching, explicit instruction tends to divorce pronunciation from the rest of the lesson by encouraging learners to attend to particular phonetic forms of the language. Explicit focus on L2 phonetics may help advanced learners reduce their pronunciation errors and increase intelligibility (Sturm, 2013) and comprehensibility (Gordon et al, 2003, Saito, 2011) so "only a relatively time-limited explicit pronunciation component in a primarily communicative classroom [may already lead] to beneficial results in production". Contrary to this synthetic focus on forms -where learners are thought to incorporate ready-made target L2 structures to their interlanguage after instruction-, the analytic focus on meaning approach is more ecologically valid as students learn incidentally from exposure to comprehensible target language samples. Nevertheless, attention to meaning may not be sufficient for learners to notice and internalise the phonological properties of L2 speech (Trofimovich & Gatbonton, 2006).

Beyond pronunciation, many practitioners follow an analytic focus on form approach, which implicitly draws learners' attention to form in the context of meaningful communication. This

approach is motivated by the Interaction Hypothesis (Long, 1991), which claims that interaction is crucial in SLA and the modifications that result from negotiation of meaning increase input and output comprehensibility. Although its primary aim is to promote accuracy, phonetic learning is not instantaneous and learners may first exhibit emergent interlanguage forms that need to be repetitively practiced in content-based contexts in order to be internalised (Saito, 2013). Once learners successfully restructure and develop their phonetic representations, they are ready to transfer the target feature in production to communicatively authentic contexts (Saito, 2015). Saito and Wu (2014) advocate for orienting attention to phonetic form while maintaining the primary focus on meaning, and emphasize the integration of suprasegmental features in formal environments. Taking into consideration that negotiation of form in content-based lessons improves L2 phonological accuracy, tasks are crucial to direct learners' cognitive resources to phonetic forms during real-world activities (Salaberry & López-Ortega, 1998).

2.3. TBLT and TBPT

Task-based language teaching has been defined as an analytic communicative-based approach in which focus on meaning is of primary concern (Long & Robinson, 1998). Although TBLT has been shown to potentially lead to success in the acquisition of L2 grammar, lexis and pragmatics in formal contexts, it is also an issue whether this can be applied to L2 speech. Consequently, this section explores a task-based pronunciation teaching approach.

2.3.1. Definition of 'task'

According to Long (2015), TBLT starts with a task-based needs analysis of a particular group of learners and focuses on the elaboration of 'tasks', which are defined in different ways by different researchers (see table 1). Nevertheless, they all agree that tasks have a clear goal and well-defined outcome which learners need to fulfil. Also, tasks increase the focus on form during communicative activities that bear resemblance to real-world events and involve several cognitive processes that promote L2 development and performance. In Long's (2015) words, tasks are typically associated with focus on form, that is, a reactive use of a wide variety of pedagogical procedures to draw learners' attention to linguistic problems in context, as they arise during communication, thereby increasing the likelihood that attention to code features will be synchronized with the learners' internal syllabus, developmental stage and processing

ability. In other words, tasks induce a focus on language without disturbing the fundamental communicative nature of the task and help learners along the stages of acquisition.

Author	Definition
Long (1985)	A piece of work undertaken for oneself or for others, freely or for some reward. Thus, examples of tasks include painting a fence, dressing a child, filling out a form, buying a pair of shoes (…). In other words, by "task" is meant the hundred and one things people do in everyday life, at work, at play, and in between. Tasks are the things people will tell you they do if you ask them and they are not applied linguists.
Crookes (1986)	A piece of work or an activity, usually with a specified objective, undertaken as part of an educational course, or at work.
Prabhu (1987)	An activity which requires learners to arrive at an outcome from given information through some process of thought, and which allows teachers to control and regulate that process.
Candlin (1987)	One of a set of differentiated, sequenceable, problem posing activities involving learners and teachers in some joint selection from a range of varied cognitive and communicative procedures applied to existing and new knowledge in the collective exploration and pursuance of foreseen or emergent goals within a social milieu.
Bachman & Palmer (1996)	An activity that involves individuals in using language for the purpose of achieving a particular goal or objective in a particular situation.
Skehan (1998)	A task is an activity in which meaning is primary, there is some communication problem to solve, there is some relationship to real-world activities, task completion has some priority, and the assessment of the task is in terms of outcome.
Bygate et al. (2001)	An activity which requires learners to use language, with emphasis on meaning, to attain an objective.
Ellis (2003)	A [task] involves a primary focus on meaning (…), it involves real-world processes of language use (…), it can involve any of the four language skills (…), it engages cognitive processes (…), and it has a clearly defined communicative outcome.
Van den Branden (2006)	A task is an activity in which a person engages in order to attain an objective, and which necessitates the use of language.
Samuda and Bygate (2008)	A holistic activity, which engages language use in order to achieve some non-linguistic outcome while meeting a linguistic challenge, with the overall aim of promoting language learning, through process or product or both.

Table 1. *Definitions of 'task'.*

According to Pica et al. (1993), task conditions can be psycholinguistically classified as [a] interactant relationship (one-way/two-way); [b] interaction requirement (required/optional); [c] goal orientation (convergent/divergent); and [d] outcome options (open/closed). They can also be cognitively divided into [a] information-gap activity; [b] reasoning-gap activity; and [c] opinion-gap activity (Prabhu, 1987).

2.3.2. Manipulation of tasks and LREs

In task-based research and in pedagogical contexts, tasks features can be manipulated in order to generate further focus on form in meaning-driven interactions. This study follows the Cognition Hypothesis (Robinson 1996, 2001, 2011) which emphasizes the flexibility of attentional capacity and claims that greater effort at conceptualization induces learners to stretch and develop their L2 linguistic resources. Furthermore, Robinson (2001; op cit Robinson & Gilabert, 2007, p.162) states that "pedagogic tasks should be designed, and then sequenced for learners on the basis of increases in their cognitive complexity" because they have the potential to lead to more accurate and complex language. Following Robinson's Triadic Componential Framework (Robinson & Gilabert, 2007), tasks may be created by considering *task complexity*, *task condition* and *task difficulty*, establishing a relationship between learners' factors and tasks as well as linguistic performance (see figure 1). In this study, only task complexity is tackled, which is "the result of attentional, memory, and other information processing demands imposed by the structure of the task on the language learner" (Robinson, 2001, p. 29). As for task complexity, Robinson (2011) distinguishes cognitive/conceptual (resource-directing) variables and performative/procedural (resource-dispersing) variables. Robinson (2011), on the basis of the work of Talmy (2000, 2008), agrees that increasing the complexity of resource-directing task characteristics has the potential to direct learners' attentional and memory resources to the way the L2 structures and concepts differ from the L1, with the likelihood of future automatization. Within resource-directing dimensions, tasks can be manipulated by increasing task complexity through ± few elements and ± reasoning demands, which guides resources to specific functional and linguistic requirements. Concerning ± reasoning demands, they can be described as a task component which makes learners reason about certain actions and justify their choices (Robinson, 2007). Turning to ± few elements, Malicka (2014) states that the more elements involved in the task, the more complex it is; however, the number of elements needed to distinguish simple from complex ones is undefined and depends on the operationalization of the construct of cognitive complexity.

Task Complexity (Cognitive factors)	*Task Condition* (Interactive factors)	*Task Difficulty* (Learner factors
(Classification criteria: cognitive demands) (Classification procedure: information-theoretic analyses)	(Classification criteria: interactional demands) (Classification procedure: behavior-descriptive analyses)	(Classification criteria: ability requirements) (Classification procedure: ability assessment analyses)
(a) Resource-directing variables making cognitive/conceptual demands	*(a) Participation variables* making interactional demands	*(a) Ability variables* and task-relevant resource differentials
+/− here and now +/− few elements −/+ spatial reasoning −/+ causal reasoning −/+ intentional reasoning −/+ perspective-taking	+/− open solution +/− one-way flow +/− convergent solution +/− few participants +/− few contributions needed +/− negotiation not needed	h/l working memory h/l reasoning h/l task-switching h/l aptitude h/l field independence h/l mind/intention-reading
(b) Resource-dispersing variables making performative/procedural demands	*(b) Participant variables* making interactant demands	*(b) Affective variables* and task-relevant state-trait differentials
+/− planning time +/− single task +/− task structure +/− few steps +/− independency of steps +/− prior knowledge	+/− same proficiency +/− same gender +/− familiar +/− shared content knowledge +/− equal status and role +/− shared cultural knowledge	h/l opennes to experience h/l control of emotion h/l task motivation h/l processing anxiety h/l willingness to communicate h/l self-efficacy

Figure 1. *The Triadic Componential Framework for task classification (from Robinson & Gilabert, 2007).*

Michel (2011) posits that there may be a confound between reasoning demands and number of elements, as a result, the factor ± few elements in combination with ± reasoning demands influences cognitive task complexity and, thus, affects task-based L2 performance. For example, Gilabert et al. (2009) and Baralt (2014) advance that complex tasks, where learners have to stretch their attention and memory resources, trigger more language-related episodes (LREs), which are defined as "any part of a dialogue where students talk about language they are producing; question their language use, or other- or self-correct their language production" (Swain & Lapkin, 1995; op cit Bygate et al., 2001, p.104). Therefore, increasing task demands along resource directing dimensions is likely to draw attention to how messages are being encoded during performance and, consequently, lead to interlanguage development (Gilabert, 2007). Nonetheless, complex tasks are more prone to inducing LREs when they are not extremely challenging and understanding between interlocutors is sufficient for communication (Révész, 2011). In sum, on the basis of the task-based literature revised here, the prediction is that task complexity -along a higher occurrence of pronunciation-based LREs- may indirectly help learners attain a more accurate L2 performance.

2.3.3. Tasks and L2 pronunciation

Task-based pronunciation teaching (TBPT) emerges when TBLT theories are applied to L2 speech acquisition. Considering the effectiveness of tasks for L2 learning, TBPT presents tasks which generate form-focused episodes that target phonological elements during interaction. In other words, tasks raise awareness of pronunciation elements by making target items essential and enhance the occurrence of pronunciation-focused LREs in conversations (Mora & Levkina, 2017). Despite the limited number of empirical studies on TBPT, some researchers have recently applied already extensively researched TBLT dimensions (i.e. task complexity, task repetition and task modality) to improve L2 pronunciation accuracy. For instance, Solon et al.'s (2017) findings support the Cognition Hypothesis (Robinson, 2001, 2007, 2011) in that the more complex version of the task generates more accurate realizations of L2 vowels; however, pronunciation-based LREs are produced at a higher rate during the simple task than the complex one, even if mean rates are not statistically significant. Jung et al. (2017) investigated the role of task repetition in the development of L2 stress patterns through collaborative priming tasks and found that repetition enhanced L2 speech intelligibility. Finally, in relation to task modality, Loewen and Isbell (2017) reveal that the occurrence of LREs is not statistically significant between face-to-face conversation and computer-mediated audio-only conversation but, on the contrary, Parlak and Ziegler (2017) report that learners in the face-to-face condition benefit more from recasts when they need to identify the correct position of lexical stress in the target words.

2.4. Learner factors affecting L2 speech learning

Studies following a communicative approach have shown that L2 proficiency modulates the results in form-focused instruction. For example, low-level learners may not benefit from corrective feedback as much as high-level learners (Saito & Akiyama, 2017) or may not be at the adequate developmental stage for awareness and repetition (Trofimovich & Gatbonton, 2006). As Mora and Levkina (2017) mention, "lower-level learners may need opportunities for developing their L2 phonology through repetition and awareness development without being under pressure for producing modified output". Nevertheless, Lee et al. (2015) have not found a clear pattern for the effects of proficiency on pronunciation instruction, suggesting that learners at different proficiencies may all benefit from TBPT.

Cognitive abilities require learners to recollect linguistic information under difficult conditions and to have attentional flexibility for different stimulus dimensions (Segalowitz, 1997). Certain phonological domains seem to be related to particular cognitive abilities more than other domains, suggesting that "phonological processing is a complex task, requiring recruitment of various cognitive abilities" (Darcy et al., 2015, p.71). In the field of L2 speech acquisition, working memory, attention control and inhibitory capacity are related to attaining high L2 outcomes. In fact, Darcy et al. (2016) contend that inhibitory control -together with selective attention- is implicated in L2 phonological processing because high inhibitory control helps supress L1 phonological categories and develop more accurate L2-specific categories, thus, explaining inter-subject differences in the perception and production of L2 linguistic representations (Lev-Ari & Peperkamp, 2014). The present study takes individual differences in proficiency and cognitive ability as mediating factors.

3. The current study

3.1. Justification and goals of the study

The aim of this study is to bridge the gap between L2 pronunciation instruction, which is often explicit and minimally communicative, and task-based language teaching, which is based on the incidental negotiation of form and meaning during interaction but has rarely been extended beyond grammar and lexis. The design of this experiment is based on four communicative tasks, embedded in a real-world situation, whose L2 phonological forms are essential for task completion. Specifically, the aim of the present study is to examine the effectiveness of task design on the perception and production of a difficult vowel contrast for EFL learners. The selected phonological contrast is /æ/ - /ʌ/ (e.g. *cat-cut*), two English sounds that are challenging for Catalan/Spanish speakers because they are perceptually assimilated to their single L1 low vowel category /a/ (Flege, 1995; Best & Tyler, 2007; Rallo-Fabra & Romero, 2012). Tasks were manipulated with various levels of cognitive complexity (along ± elements and ± reasoning demands) with the objective of enhancing pronunciation-based LREs that may facilitate the improvement of L2 pronunciation accuracy. Finally, the role of individual differences in L2 proficiency and attention control in explaining inter-subject variation in perception and production gains was explored.

3.2. Research questions

Taking into consideration the objectives discussed above, the following research questions and hypotheses are formulated:

RQ.1. Do learners' perception and production of English /æ/ & /ʌ/ improve after the performance of four decision-making tasks?

RQ.1.1. Are gains retained two weeks after the treatment?

RQ.1.2. Do gains generalize to novel items (non-words) and speakers?

Hypothesis RQ.1. Given that the tasks are designed to direct learners' attention to the specific phonological contrast through the use of task essential language (Loschky & Bley-Vroman, 1993), improvement in perception and production is expected. It may also be the case that gains are greater at the perceptual level than the production level because perception usually precedes production (Flege, 1995) and learners take more time to modify the articulation of L2 vowels. Moreover, learners may show retention of the vowel contrast in the delayed post-test, as in HVPT training (Bradlow et al., 1999), and transfer to novel items (Hazan et al., 2005) and speakers (Flege, 1995), suggesting in-depth learning of the phonetic categories.

RQ.2. Does increasing task complexity have an effect on the occurrence of pronunciation-based language-related episodes?

RQ.2.1. How does it relate to learners' gains in perception and production?

Hypothesis RQ.2. An increase in reasoning demands and number of elements is predicted to produce greater occurrence of LREs in the more complex tasks (Gilabert et al., 2009). No hypotheses are made about the relationship between LREs and gains in perception and production, a relationship unattested in previous research.

RQ.3. Are perception and production gains affected by proficiency, inhibitory control and selective attention?

Hypothesis 3: It is hypothesised that learners with better inhibition and attention control will obtain greater gains in L2 speech production and perception (Darcy et al., 2016). We are unable

to specify the direction of the relationship between gains and L2 proficiency, as mixed results have been found in pronunciation teaching studies (see section 2.5).

4. Methods

4.1. Participants

Thirty-six Catalan-Spanish bilingual speakers from a semi-private secondary school took part in the study. In the experimental group, there were 18 students (9 females) who were 16.4 years old and had been learning English together since the age of 6 at school. Moreover, 61.1% of the class stated that they had received extra-curricular English instruction between 2 and 13 years (*M*= 6.36) for 2.5h/week. One person had studied English in a naturalistic context during 6 summers (2 weeks each). Their self-assessment of English, reported in the questionnaire, indicated a B1-B2 level according to the *Common European Framework of Reference for Languages.* Other foreign languages spoken were French (B1) [5 students], Italian (A2) [1 student] and Portuguese (A1) [1 student].

In the control group, there were eighteen Catalan/Spanish bilinguals (7 males and 11 females) from the same school but different year. Although all of them were born and raised in Catalonia, one of the students was born in Ghana but raised in Catalonia, achieving a high proficiency in both Catalan and Spanish. They were 15.2 years old and they had received English instruction since primary school, at the age of 6. Furthermore, 61.1% of the participants had been formally instructed in language schools between 1 and 9 years (*M*=5.27) for 2.5h/week and their level of English proficiency was a B1. Finally, French (B1) was the only L3 that learners spoke as it is instructed in the school during 4 years. None of the learners in either group had been previously taught any English phonetics but may have engaged in pronunciation activities from textbooks.

4.2 Research design

The experimental and the control groups were tested before and after the one-week treatment period, and two weeks after it (delayed post-test). Testing included perception (vowel identification and discrimination) and production (delayed sentence repetition) tasks. Pre-test was not equal to post-test because the post-test and delayed post-test included novel items and speakers learners had not been exposed to during testing or training. This was done in order to

avoid familiarity to new contexts before the post-test and ensure the reliability of generalization tests. The treatment involved 4 sessions of 15 to 30 minutes each, which were carried out every day (see appendix D). The target items were practiced repetitively during pre-task and tasks. See figure 2 for a summary.

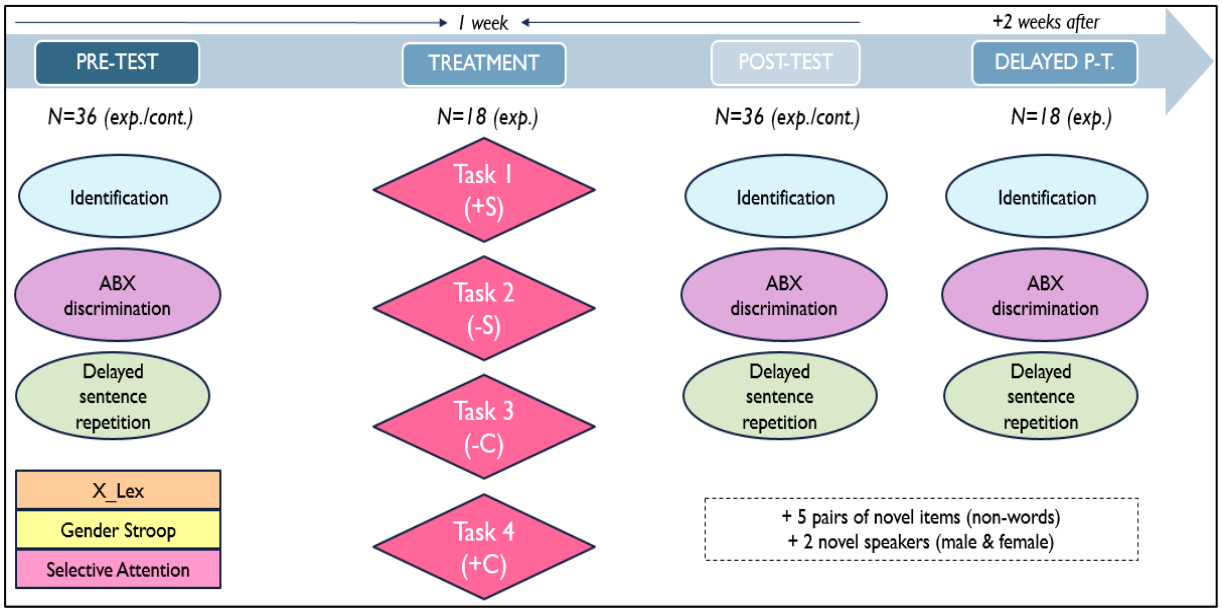


Figure 2. Research design of the experiment.

4.3 Materials

4.3.1 Stimuli

Stimuli were specifically designed and collected for this study. The selected target sounds were the two standard Southern British English vowels /æ / and /ʌ/. This contrast was embedded in 10 pairs of real words, which appeared in the pre-test, post-test and delayed post-test, and 5 pairs of non-words, which only appeared in the post-test and delayed post-test (see appendix C for words and sentences).

4.3.2. Speakers and elicitation procedure

Four British native speakers (2 females) of similar ages produced the stimuli for the perception and production tests and pre-tasks. The words and non-words in the perception tests were elicited in carrier phrases (*I say X, I say X again*). The sentences in the production test (DSR) were elicited from reading lists. Also, two out of the four speakers (male and female) produced

a dialogue that was latter segmented and combined to create the listening comprehension of two of the pre-tasks.

4.3.3. Testing

4.3.3.1. Perception

Concerning the identification test (in *Praat*), it was formed by 80 random trials + 8 trials for practice, which were composed by 10 minimal pairs (20 tokens) per speaker which were repeated twice. The contrasts contained the vowels "a" and "u" (/æ/- /ʌ/) and had one syllable (e.g. *bag-bug*) or two syllables (e.g. *amber-umber*); however, 70% of vowel contrasts were monosyllabic. In the post-test, untrained items (i.e. non-words)[4] were incorporated together with trained items. Stimuli were composed by 10 trained minimal pairs (20 words) which were half of them uttered by untrained voices and half of them by trained voices and 5 untrained minimal pairs (10 non-words).

In the discrimination ABX test (in *DmDx*), 88 test trials + 4 practice trials were presented. Within the 88 trials, 8 trials were released as control items (e.g. *A*: Bin, *B*: Bag, *A*: Bin) to ensure participants' correct performance of the test. The 10 minimal pairs were randomly presented twice in the four orders (ABA, ABB, BAA, BAB). In the post-test, 10 trained minimal pairs were spoken by old and new voices (20 trials) and 10 untrained minimal pairs (non-words) were also spoken by old and new voices (20 trials). All the trials were randomly presented in the four orders and four voices (M-M-F and F-F-M).

4.3.3.2. Production

As for the delayed sentence repetition task (in *DmDx*), learners were exposed to 44 test trials + 2 practice trials. The 44 test trials included 40 sentences with the target vowels (/æ/ & /ʌ/) and 4 sentences with other non-target minimal pairs (/iː/ & /ɪ/). These distractors were used in order to avoid learners focusing too much on the target vowels and get a more natural performance. The sentences were only uttered once by male and female speakers. In the post-test, sentences included 10 minimal pairs with trained and untrained voices (20 sentences) and 5 minimal pairs that contained non-words with trained and untrained voices (20 sentences).

[4] Non-words had the same phonetic context as real words (e.g. *bat/butt* vs. *dat/dutt*); see appendix C.

4.3.3.3. Proficiency

Proficiency was measured with the vocabulary size test X_Lex and attention control with an auditory selective attention test and an auditory gender stroop test. The X_Lex test (Meara & Milton, 2003) has been shown to correlate strongly with proficiency and has been used as a reliable indicator of L2 proficiency in a number of previous empirical studies (Gilabert et al., 2009). Learners were presented with 120 words and they needed to indicate, in a yes/no format, if they knew the word.

4.3.3.4. Attention

The auditory selective attention test (Hummes et al., 2006) presented a call signal (e.g. Charlie) and two sentences with conflicting information so learners had to focus on the sentence that contained the call signal (e.g. *ready Charlie* vs *ready Arrow*) and select the colour (*white, green, red, blue*) and digit (*1-8)* that the sentence expressed (e.g. *ready Charlie go to white, eight now*). Learners responded to 33 trials whose answers corresponded to correct/incorrect digit or colour, and combined (digit + number) with a 0 meaning wrong and a 1 correct. Concerning inhibitory control, in the gender stroop test, learners were exposed to different words in Catalan (i.e. oca, núvol, noia, home, nata, oli) and had to select whether the voice was male or female. RT were registered for the analysis of inhibitory control.

4.3.4. Training

The materials followed a pre-task/task design and they were specially created for this experiment.

4.3.4.1. Pre-tasks

A general pre-task was used to train the meaning of the words that appeared in the tasks. Then, participants practiced their pronunciation through word imitation and sentence imitation practices (in *DmDx*) where feedback was provided. In the word imitation and sentence imitation, 10 minimal pairs were presented (20 tokens) plus 5 minimal pair distractors (10 tokens), containing the vowels /iː/ and /ɪ/ (i.e. *bean/bin, feast/fist, sheep/ship, teen/tin, weep/whip*) with their corresponding images -designed by the researcher and collaborators- (see appendix C).

Also, mini pre-tasks were carried out before each of the tasks to remind students of the meaning and pronunciation of the target words. Before task 1 (+S) and task 4 (+C), there were two listening comprehensions spoken by the two English speakers (1 female) who appeared in the pre-test, whereas task 2 (-S) and task 3 (-C) were preceded by two listening comprehensions narrated by the researcher (see appendix D).

4.3.4.2. Tasks

Concerning the four decision-making tasks, they were two-way, split, close and convergent (Pica et al., 1993) because the two interlocutors had different information and they had to come up with one single solution. Moreover, learners could not solve the task if they did not produce the L2 phonological contrast accurately and so, they were made "task-essential" language (Loschky & Bley-Vroman, 1993). These tasks were designed around a trip to Kenya that students had to plan. In a sequential manner, learners had to decide on what they wanted to see and buy in a natural park in Kenya (Task 1); the objects they wanted to bring to Kenya (Task 2); the organization of a "roleplay" party in Kenya (Task 3) and what they wished to post in the school website (Task 4). All tasks involved two mental operations: information-sharing and decision-making; nevertheless, in order to complexify the tasks in increasing order (+S, -S, -C, +C), more elements and reasoning demands were added in subsequent versions of the tasks (appendix D). In addition, task complexity was independently assessed (Révész, 2011) by ten experienced language teachers, who critically evaluated the tasks according to the degree of difficulty and mental effort. Students themselves also rated task difficulty after each task (appendix E). All the tasks were previously piloted on a similar sample of learners and proved to be adequate for their level of proficiency and difficulty.

4.4. Procedure

Participants filled out an informed consent form (appendix A), a language background questionnaire (appendix B) and the Bilingual Language Profile. All the tests and tasks took place in a small classroom in the school, except for the X_Lex proficiency test, which took place in a computer room. Students came in pairs and did the tests in a counterbalanced order in front of a computer for 30-40 minutes (see table 2).

	STUDENT A	STUDENT B
PRE-TEST	Identification	Delayed-Sentence Repetition
	Delayed-sentence repetition	Identification
	Discrimination	Discrimination
	Auditory Gender Stroop	Auditory Gender Stroop
POST-TEST	Delayed-Sentence Repetition	Identification
	Identification	Delayed-sentence repetition
	Discrimination	Discrimination
	Auditory Selective Attention	Auditory Selective Attention
DELAYED POST-TEST	Identification	Delayed-Sentence Repetition
	Delayed-sentence repetition	Identification
	Discrimination	Discrimination

Table 2. *Schedule of tests for student A and B.*

Perception tests and pre-tasks were administered through headphones (*Sennheiser PC8*). In the identification test, learners had to select the word that they heard by clicking on the corresponding box. In the ABX discrimination test, learners heard three different stimuli and had to say whether X corresponded to A or B. Finally, in the DSR task, learners had to read the sentence on the screen, listen to it, wait for a *beep* sound and repeat it. *DmDx* presented the stimuli and the productions were recorded with a digital *Tascam Dr-40* recorder with an external *Shure SM58* microphone. The tests of proficiency and attention control were administered before the treatment. The X_Lex proficiency test was done in a computer room with 18 computers whereas the attention control tests were carried out in pairs in a small classroom.

After the pre-test and before the post-test, participants in the experimental group did the general pre-task and tasks, whereas the control group did not receive any phonetic training and was only exposed to grammar-based lessons at school that did not include pronunciation. On the following day, learners did a very short pre-task which lasted 3-5 minutes in pairs. Tasks lasted 15-30 minutes, depending on their complexity, and were registered using two *Tascam Dr-40* solid-state recorders. Learners were placed in front of a microphone and facing each other so they could not see each other's piece of information. After each one of the 4 sessions, learners filled in a short survey indicating, on a nine-point scale, the mental effort of the task, attractiveness, performance and perceived time-on-task. Tasks were carried out one day after another during four days and were recorded, transcribed and analysed in terms of LREs.

4.5. Analyses

4.5.1. Analysis of perception, production and learner factors.

Accuracy scores of the identification test as well as response latencies (ms) of the discrimination test were aggregated for the experimental and control groups. RT scores were screened above and below 2.5 standard deviations. In addition, practice items and control items were discarded from the general analysis (see appendix G.2 for detailed information). Since the post-test analyses contained instances of non-words, gains were only calculated for the words which participants were trained on. Therefore, pre-test scores were compared to post-test scores and individual gains were obtained by subtracting pre-test scores to post-test scores. Moreover, post-test scores were subtracted from the delayed-post test scores to analyse retention of the vowel contrast after 2 weeks. Furthermore, the independent variables of voice and vowels -in the ID/DSR test- as well as sequence order (in the DIS test) were checked to confirm that they did not have any effect on perceptual gains (Appendix G.4). Finally, generalization to novel items and speakers was examined by comparing pre-test trained items and voices to post-test untrained items and voices.

Concerning the production task, learners' accuracy scores were obtained by analysing the quality of vowels /æ/ and /ʌ/. Analyses focused on the first formant (height) and second formant (advancement), which were transformed into Bark following this formula: [B_i = 26.81/(1+1960/F_i) - 0.53] (Syrdal & Gopal, 1986) and then, the spectral distance between the two vowels (Euclidean distance) was also calculated[5]. Having discarded the two practices and four distractor trials, data was aggregated in a by-subjects dataset to contrast learners' B1, B2 and Euclidean distance values with native speakers' values at pre-test, post-test and delayed post-test.

As informed in appendix G.3, vowels in words produced by male and female speakers were analysed separately because they were not acoustically comparable due to pitch and vocal track differences and the normalized values obtained with the Bark Distance Metric were difficult to interpret in relation to the research questions. This may be due to the fact that this intrinsic method only works with high front vowels. The values for our four native speakers are reported in appendix G.3 and illustrated in figures 3 and 4. Besides, they are closely related to the measures in Deterding's (1997) study.

[5] Euclidean distance formula: $\sqrt{(B2\ æ - B2\ ʌ)2 + (B1\ æ - B1\ ʌ)2}$.

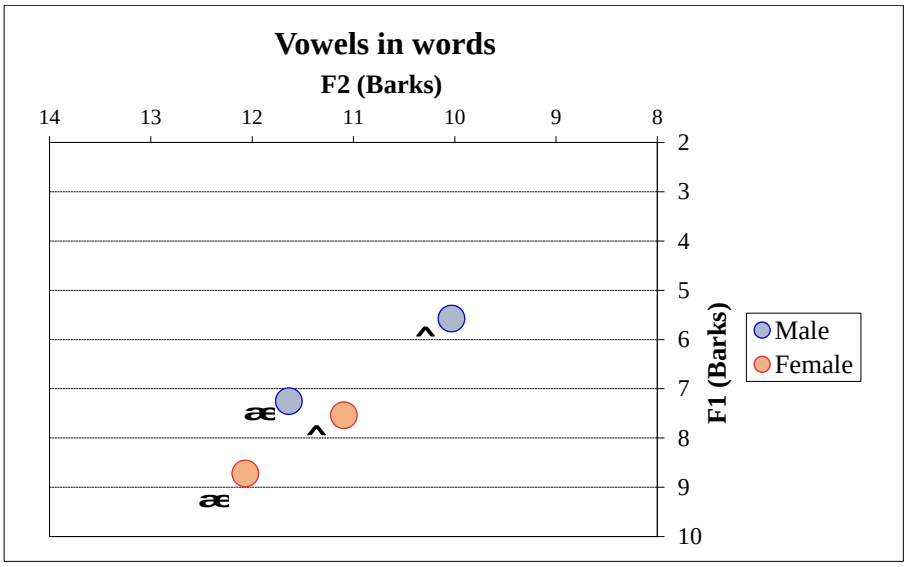


Figure 3. *Vowel plot for male and female NSs values: Eng. /æ/ and /ʌ/ in words.*

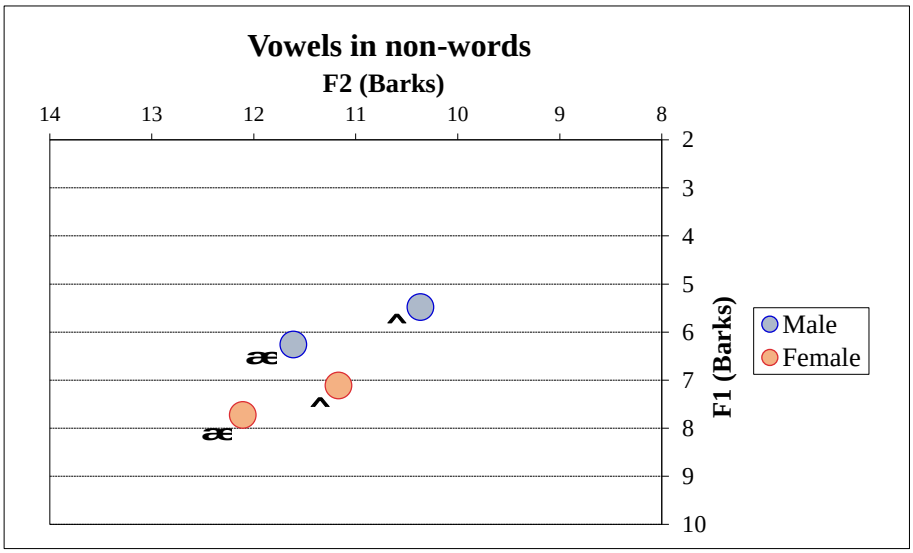


Figure 4. *Vowel plot for male and female NSs values: Eng. /æ/ and /ʌ/ in non-words.*

In addition, learners' mean scores for B1 and B2 were interpreted in the discussion section taking into account their native language counterpart: Sp. /a/. Following Martínez Celdrán and Fernández Planas (2007), the Bark measures for height were 6.51 (males) and 7.81 (females) and, for advancement, were 10.96 (males) and 11.96 (females).

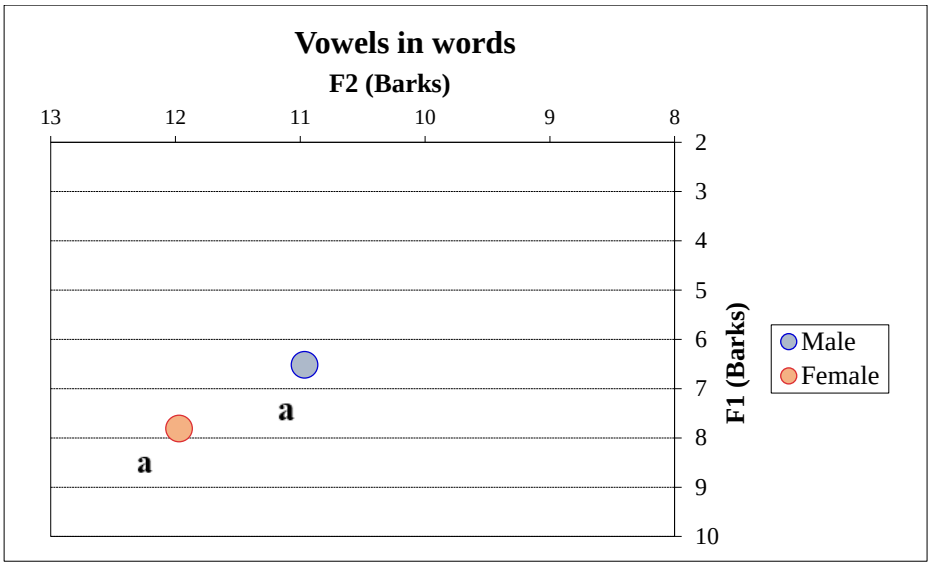


Figure 5. *Vowel plot for male and female learner values: Sp. /a/ in words.*

Finally, proficiency scores, inhibition scores and attention scores were correlated with perception and production gains.

4.5.2. Analysis of LREs

Three raters (*M= 24.7 years old*), experienced English teachers living in Barcelona (Spain) at the time of testing, were instructed on the analysis of LREs. They listened to 100% of the recordings (4 tasks), transcribed the pronunciation focused LREs and classified them into four types: (a) general LREs, (b) recasts, (c) self-repairs, (d) repetitions (see appendix F for LREs' guidelines and definitions). The number of LREs per person was calculated with all types of LREs as well as with only general LREs. In addition, an LRE ratio (LRE/time-on-task) was estimated to interpret the results in relation to Solon et al.'s (2017) study. Finally, in order to correlate gains and amount of LREs, each dyad was assigned the same number of LREs.

5. Results

5.1. Perception and production

For all tasks where control items were included, analyses were performed to ascertain that performance on control items was significantly higher than performance at test items (see appendix G.2 [tables 8.10 and 8.11, figure 8.2], and later discarded for general analyses.

5.1.1. Identification

A one-way ANOVA, with *testing time* as the within-subjects factor, reported that there were statistically significant differences among the three testing times (F (2,16) = 301.306, p<.001, η^2= .974). Bonferroni pairwise comparisons revealed that learners significantly improved from pre-test (M=.463, SD= .092) to post-test (M=.944, SD=.048) and delayed post-test (M=.944, SD=.045) (p<.001) but obtained similar accuracy scores from post-test to delayed post-test (p=1.000). See appendix G.1 (table 8.1, 8.2 and 8.3) and figure 5.1.

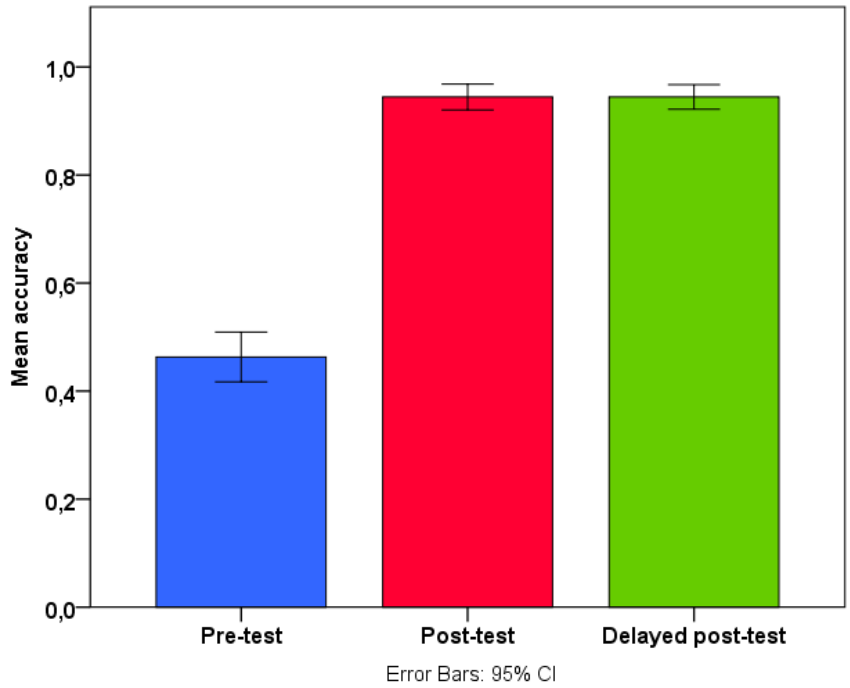


Figure 5.1. *Bar graph for pre-test, post-test and delayed post-test mean correct identification scores (experimental group).*

Concerning the control group, the paired samples t-test revealed that participants did not significantly improve from pre-test (M= .651, SD= .118) to post-test (M= .686, SD= .121), t (17) = -1.804, p= .089, r= .156, see appendix G.1 (table 8.4) and figure 5.2.

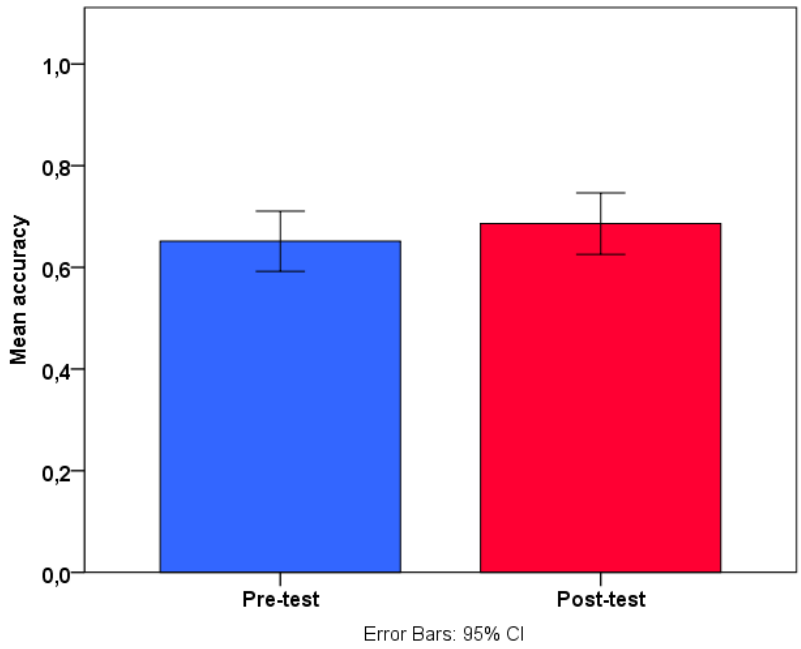


Figure 5.2. *Bar graph for pre-test, post-test and delayed post-test mean correct identification scores (control group).*

In addition, a mixed design ANOVA revealed a significant interaction between *testing time* (pre-test and post-test) and *group* (experimental and control) (F (1, 34) = 221.158, p<.001, η^2= .867) so the main effects of *time* had to be interpreted independently for the experimental group and control group (appendix G.1, figure 8.1). See appendix G.1 (tables 8.7, 8.8 and 8.9) for further analysis that confirmed gains in the experimental group despite differences at the onset.

5.1.2. Discrimination

Boxplots and extreme values table (appendix G.2, tables 8.12 and 8.13) reported that outlier number 30 had to be eliminated from the control group.

Regarding the experimental group's accuracy, a one-way ANOVA reported significant differences across the three testing times (F (2,16) = 23.478, p<.001, η^2= .746); nonetheless, Bonferroni pairwise comparisons detected that participants significantly improved between pre-test (M= .665, SD= .094) and post-test (M= .813, SD= .120) as well as pre-test and delayed post-test (M= .836, SD= .136) (p<.001) but no significant changes were found between post-test and delayed post-test (p =.809). See appendix G.2 (tables 8.14, 8.15 and 8.16). In terms of RT, learners in the experimental group also performed better between pre-test (M= 817.198, SD=194.839) and post-test (M= 707.393, SD= 152.156), and between pre-test and delayed post-

test (*M*= 663.196, *SD*= 148.508), (*F* (2,16) = 10.101, *p*=.001, η^2= .558). According to the pairwise comparisons, changes reached significance between pre-test and post-test (*p*=.025), and between pre-test and delayed post-test (*p*=.001); however, from post-test to delayed post-test, learners became faster at discriminating the target contrasts but the gains did not reach significance, *p*=.169 (see figure 5.3 and appendix G.2: 8.14, 8.17, 8.18).

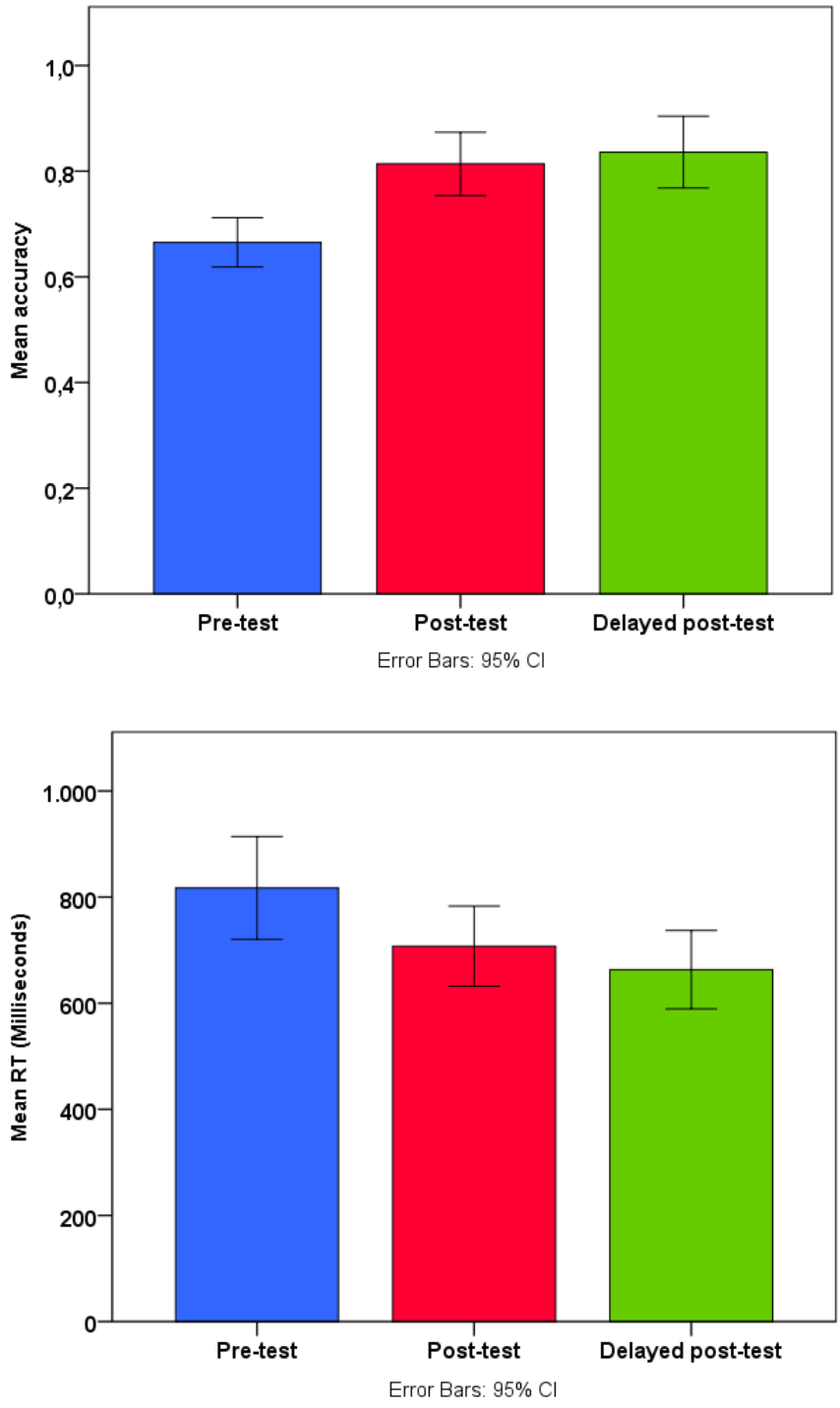


Figure 5.3. *Bar graphs for pre-test, post-test and delayed post-test mean accuracy and RT scores (experimental group).*

As far as the control group is concerned, a paired-samples t-test showed that the control group did not significantly improve from pre-test to post-test, t (16) = -2.095, p=.052, r=.465; nonetheless, they became significantly faster at discriminating vowels in the post-test, t (16) = 4.724, p<.001, r=.763, possibly due to task repetition effects (see appendix G.2: table 8.19).

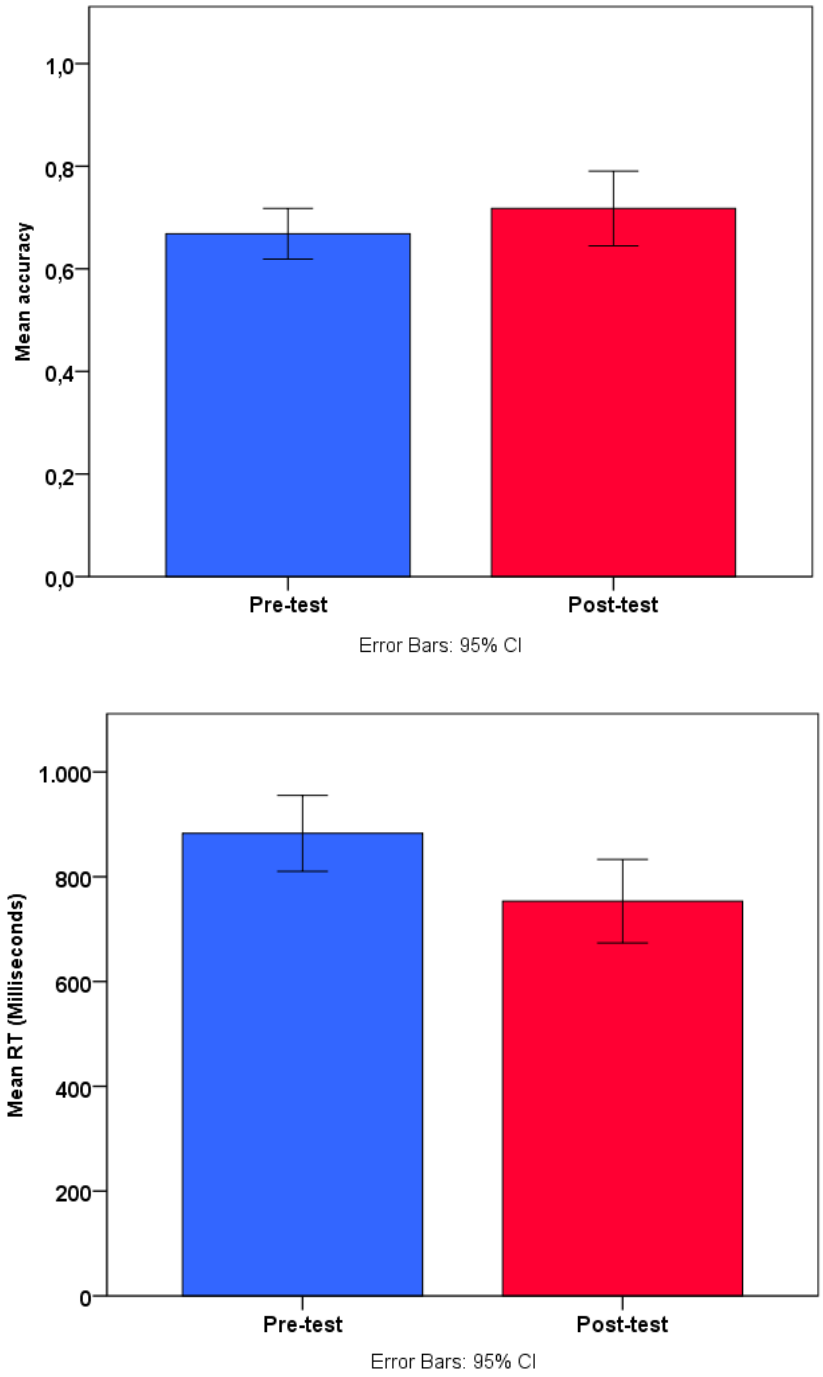


Figure 5.4. *Bar graphs for pre-test, post-test and delayed post-test mean accuracy and RT scores (control group).*

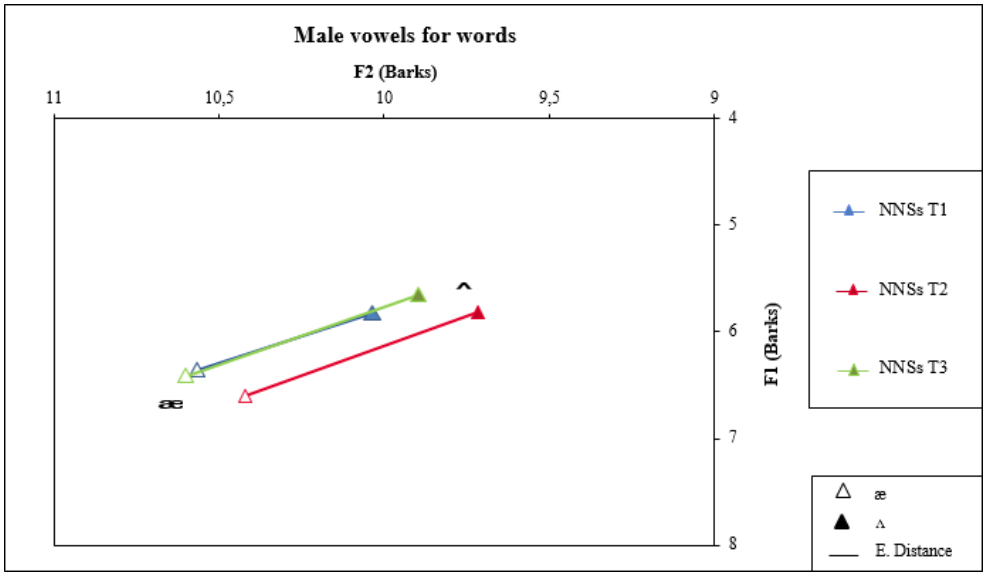
Concerning accuracy, a mixed design ANOVA informed about a significant interaction between *testing time* (pre-test and post-test) and *group* (experimental and control) (F (1,33) =

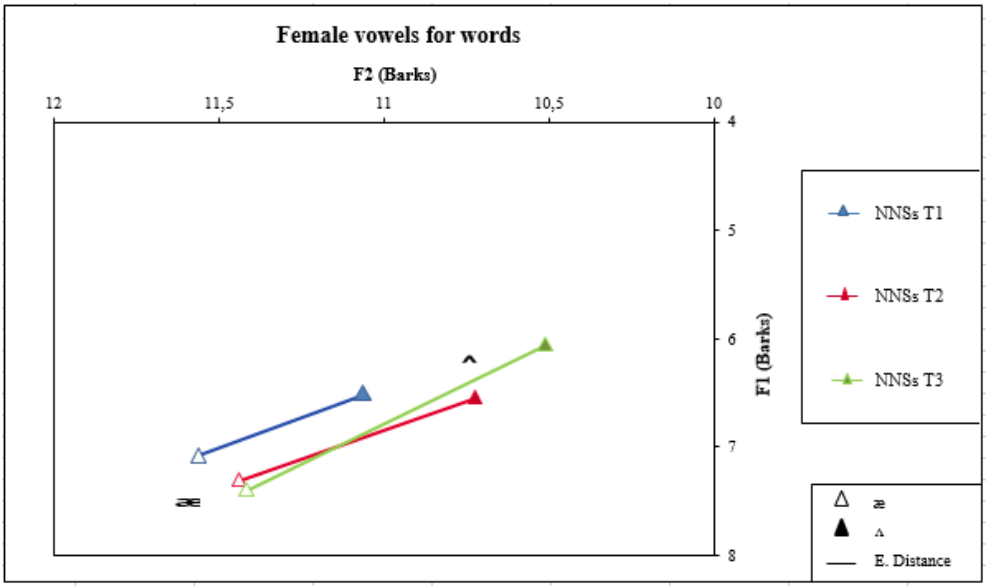
9.410, p=.004, η^2= .222) suggesting that the variation in gains needed to be understood in relation to the two different groups (see appendix G.2. (figure 8.4) and tables 8.20, 8.21 and 8.22 for further analyses). As for RTs, both groups improved statistically from pre-test to post-test (F (1,33) = 26.690, p<.001, η^2= .447). Nevertheless, the between-subjects analysis showed that the experimental and control groups were not statistically different (F (1,33) = 1.253, p= .271, η^2= .037) at pre-test and post-test (see table 8.23 and 8.24, and figure 8.5).

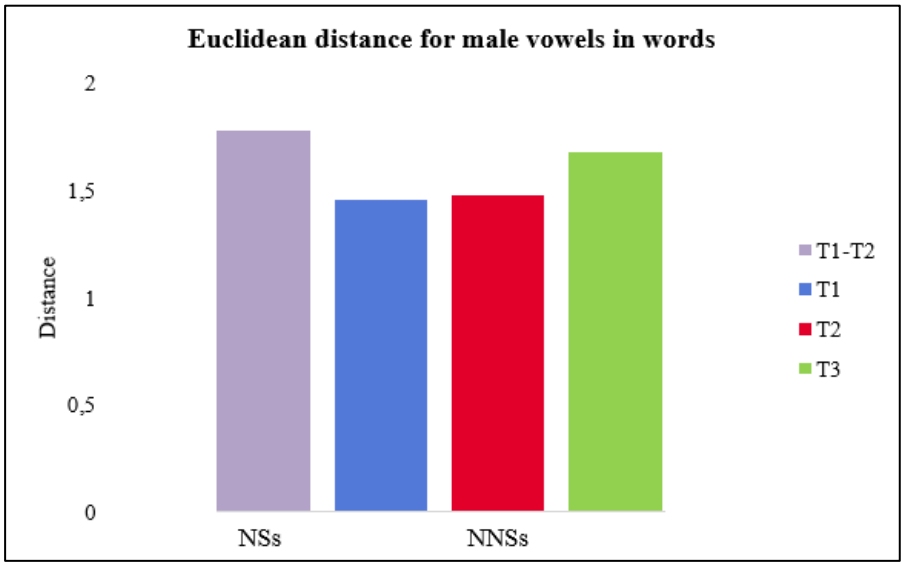
5.1.3. Delayed sentence repetition

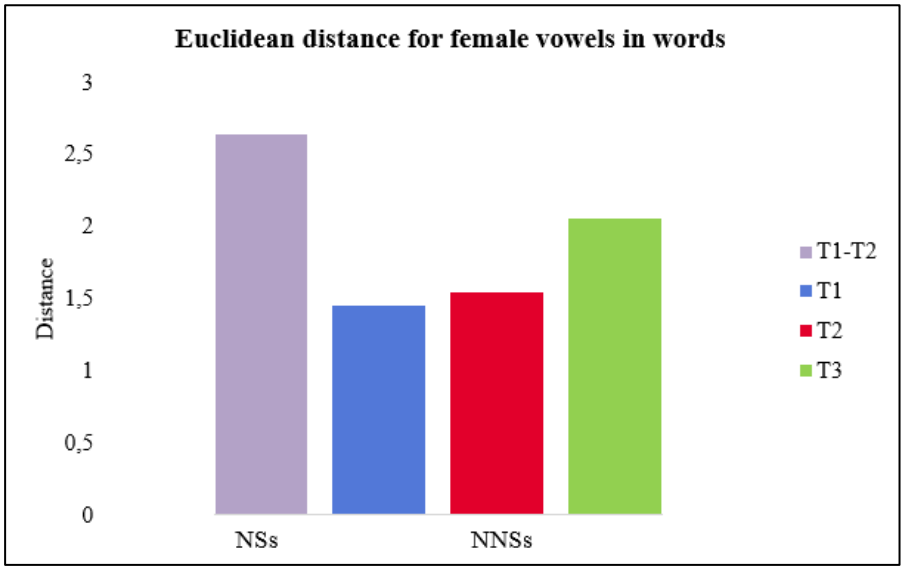
Given that gender differences were found for vowel quality in native and non-native speakers (see appendix G.3), the analysis of production was carried out separately for males and females.

The descriptives in table 8.26 (appendix G.3) show that, overall, learners did not equal the production of native speakers concerning the B1, B2 and Euclidean distance (ED) values; nevertheless, learners in the experimental group changed their productions from pre-test to post-test and delayed post-test. Averages for males and females exhibited an increase in B1 for the /æ/ vowel from pre-test to post-test -approaching L2 values- but male speakers did not behave likewise in the delayed post-test. The tendency of vowel /ʌ/ was to become higher than /æ/ so male and female learners tended to decrease their B1 values from pre-test to delayed post-test. Finally, the ED for male and female speakers of the experimental group became bigger due to the distinction of the two target vowels from pre-test to delayed post-test thanks to the treatment. In the case of the control group, it appears that the Euclidean distance was reduced (see figures 5.5 and 5.6).

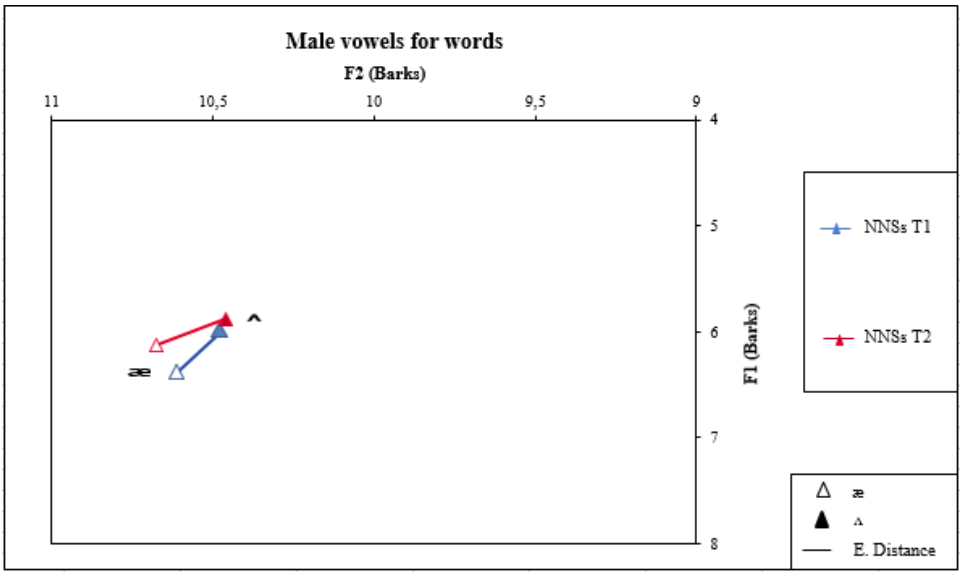


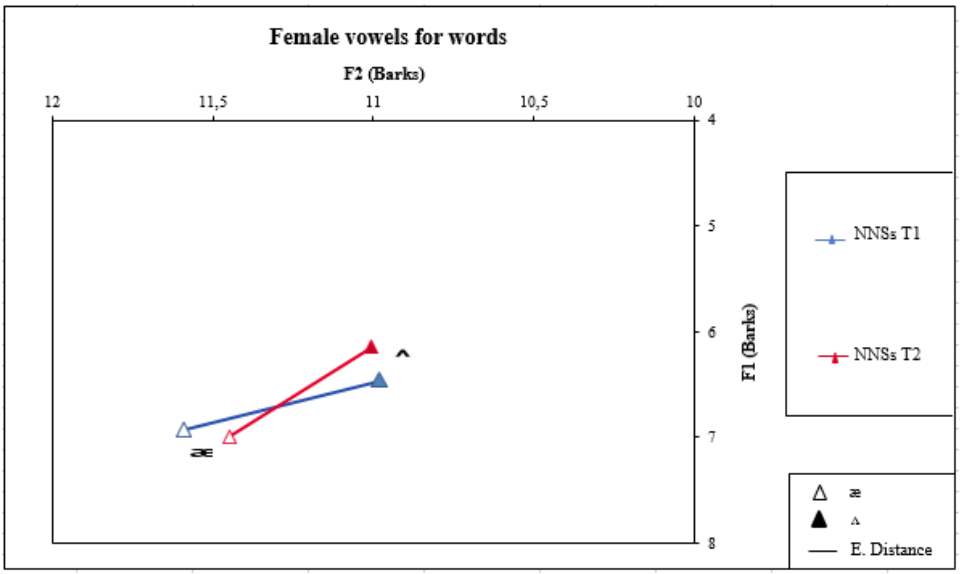


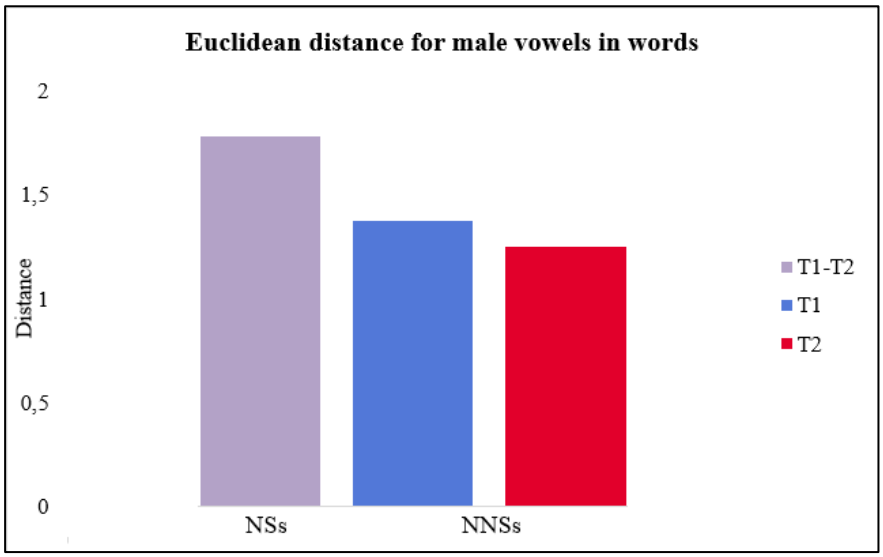


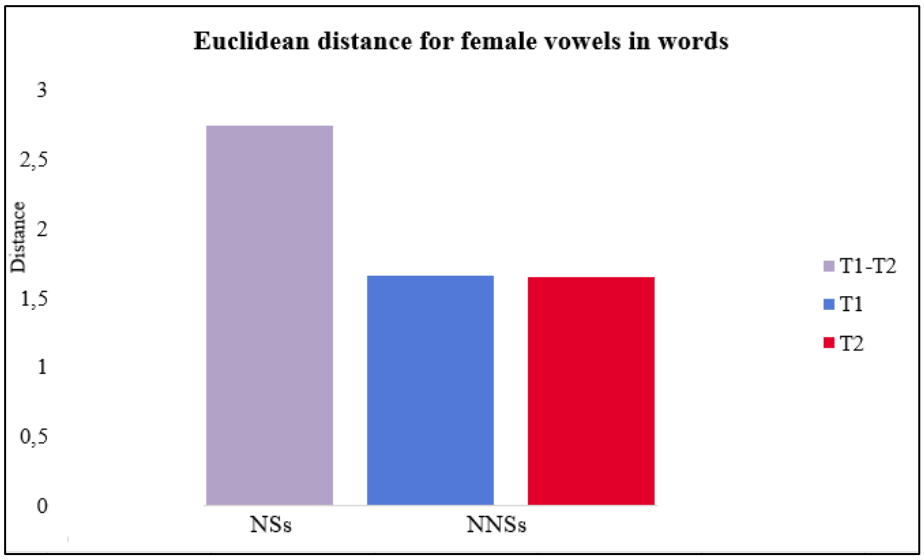


Figures 5.5. *B1 and B2 values for vowels /æ/ & /ʌ/ (pre-test, post-test and delayed post-test) and Euclidean distance for words: experimental group.*


Male vowels for words
F2 (Barks)
11
10,5
10
9,5
9
4
5
6
7
8
F1 (Barks)
æ
ʌ
NNSs T1
NNSs T2
æ
ʌ
E. Distance


Female vowels for words
F2 (Barks)
12
11,5
11
10,5
10
4
5
6
7
8
F1 (Barks)
æ
ʌ
NNSs T1
NNSs T2
æ
ʌ
E. Distance


Euclidean distance for male vowels in words
2
1,5
1
0,5
0
Distance
NSs
NNSs
T1-T2
T1
T2



Figures 5.6. *B1 and B2 values for vowels /æ/ & /ʌ/ (pre-test and post-test) and Euclidean distance for words: control group.*

Given that English speakers perceive and produce the vowels /æ/ and /ʌ/ in a significantly different way (see appendix G.3, tables 8.27 and 8.28), learners' vowels of the experimental and control groups were also analysed for differences in the B1 and B2 values across tests. A one-way ANOVA was run across tests to demonstrate the differences between /æ/ and /ʌ/ for B1, B2, and ED. Concerning males' vowels in the experimental group, they already produced a little difference at pre-test in B1 (F (1,16) = 6.506, p=.021, η^2 =.288) and B2 (F (1,16) = 8.086, p=.012, η^2 =.335); nevertheless, significance improved at post-test for B1 and B2 values (F (1,16) = 14.248, p=.002, η^2 = .470, F (1,16) = 11.664, p=.004, η^2 =.421; respectively) and even more, in the delayed post-test: B1 (F (1,16) = 17.556, p=.001, η^2 =.523) and B2 (F (1,16) = 14.864, p=.001, η^2 =.481). Similarly, females significantly differentiated /æ/ from /ʌ/ at pre-test for B1 and B2 (F (1,16) = 5.776, p=.029, η^2 =.265, F (1,16) = 9.748, p=.007, η^2 =.378; respectively); however, effect sizes increased at post-test (F (1,16) = 8.655, p=.010, η^2 = .351, F (1,16) = 17.289, p=.001, η^2 =.421; respectively) and, also, at the delayed post-test (F (1,16) = 27.805, p<.001, η^2 =.634, F (1,16) = 22.771, p<.001, η^2 =.587; respectively), see table 8.29. In order to verify that learners were acquiring the contrast between the two vowel sounds, a one-way ANOVA was run, where the within-group factor was the *ED* at pre-test, post-test and delayed post-test of the experimental group. The multivariate tests showed that males EDs did not significantly vary across testing times (F (2,7) = 1.977, p=.209, η2 =.361) whereas females EDs increased form pre- to delayed post-test (F (2,7) = 1.977, p=.209, η2 =.361), see table 8.30. This implies that, overall, learners in the experimental group increased the distance between the two vowels in production, even if this did not reach significance. In comparison to the control

group, the repeated-measures ANOVA showed that learners' ED for male speakers (F (1,6) = 2.713, p=.151, η^2 =.311) and female speakers (F (1,10) = .081, p=.782, η^2 =.008) was not significantly different from time 1 (pre-test) to time 2 (post-test), see table 8.30 and 8.31 (appendix G.3).

As shown in the descriptives (table 8.26), learners of the experimental group tended to distinguish the vowels by modifying their height (B1) rather than changing advancement (B2). As a result, we calculated a spectral distance measure of vowel height between mean native speaker height values and those of learners. Concerning the ash vowel (/æ/), male and female learners reduced the distance with respect to NSs from pre-test (M= .820, SD= .424; M= 1.955, SD= .425, respectively) to post-test (M= .682, SD= .523; M= 1.236, SD= .693) but, whereas females kept reducing these B1 values at delayed post-test, males did not (M= .847, SD= .484). The mixed design ANOVA showed a significant interaction between *gender* and *time* (F (2,15) = 11.872, p=.001, η^2 =.613), indicating that the female group approached the native speakers values closer than the male group did (appendix G.3 [table 8.36 and figure 8.6]). Concerning /ʌ/ of the experimental group, the interaction confirmed that differences in time depended on gender (F (2,15) = 11.566, p=.001, η^2 =.607) because female learners approached native speakers' values for /ʌ/ more (see figure 8.7 and tables 8.37, 8.38, 8.39).

Finally, these pre-test values for B1 were subtracted from post-test values in order to obtain a gain measure.

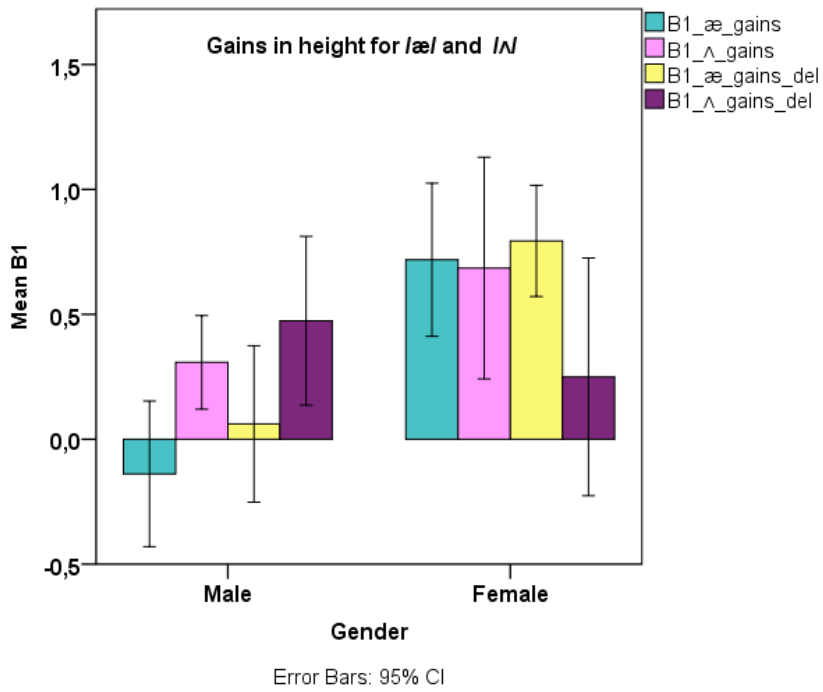


Figure 5.7. *Graph for gains in the experimental group (Pre-test – post-test vs. pre-test – delayed post-test).*

In general, female learners gained more in height than male learners. The male group obtained more gains in /ʌ/ than /æ/ from pre-test to post-test (*M*= .307, *SD*= .244) and delayed post-test (*M*= .474, *SD*= .439). The female group obtained gains in /æ/ and /ʌ/, especially from pre-test to post-test (*M*= .718, *SD*= .398; *M*= .685, *SD*= .577); but their gains in /ʌ/ decreased from post-test to delayed post-test (*M*= .249, *SD*= .618), see figure 5.7.

5.1.4. Generalization to new contexts

Before tackling generalization, the stimuli's voices (male vs. females), vowels (/æ/ vs. /ʌ/) and trial sequences (ABB, ABA, BAA, BAB) were submitted to statistical analysis to reinforce the internal validity of the instrument (appendix G.4).

Concerning the identification test, pre-test accuracy scores for trained items (words) (*M*=.463, *SD*=.092) were compared to post-test accuracy scores for untrained items (non-words) (*M*=.877, *SD*=.059) to see whether there was improvement regardless of the type of item. The paired samples t-test showed that there was a statistically significant difference, *t* (17) = -23.835, *p*<.001, *r*=.985, hence, we can infer that learners are able to generalize their learning to new contexts. See table 8.40 and figure 5.8.

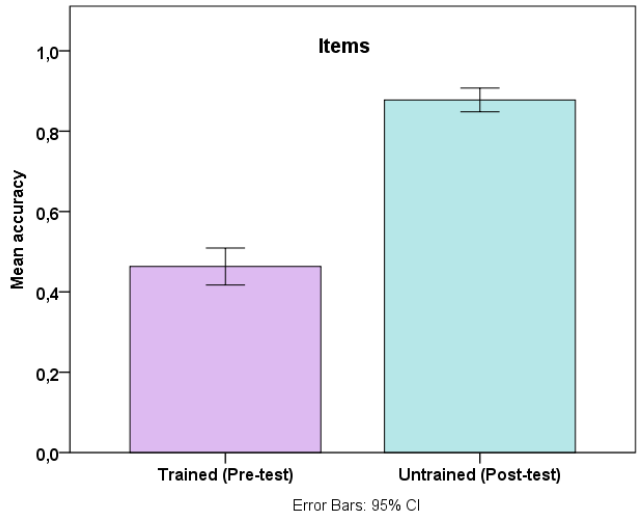


Figure 5.8. *Bar graph for accuracy scores in the identification of trained vs. untrained items at pre-test and post-test.*

Moreover, trained speakers in the pre-test (*M*=.463, *SD*=.092) were compared to untrained speakers in the post-test (*M*=.926, *SD*=.047) to observe whether learners benefitted from the multiplicity of voices from the treatment and could discriminate the vowel contrasts with new English voices. A paired samples t-test indicated statistically significant differences, *t* (17) = -24.414, *p*<.001, *r*=.986, meaning that learners improved regardless of the speaker producing the stimuli. See table 8.41 and figure 5.9.

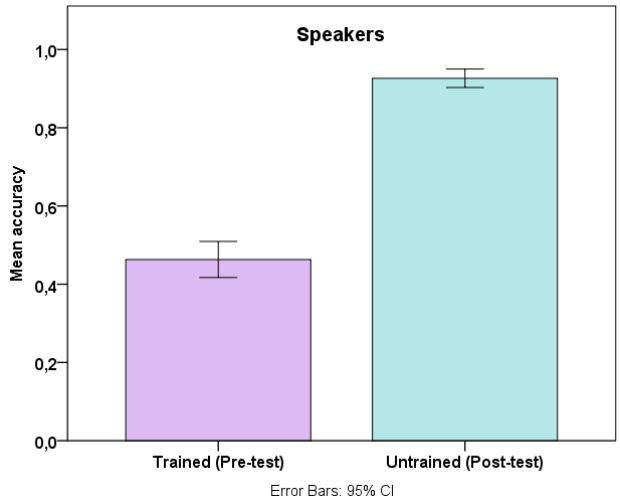


Figure 5.9. *Bar graph for accuracy scores in the identification of trained vs. untrained speakers at pre-test and post-test.*

Concerning the discrimination test, pre-test accuracy scores for trained items (*M*=.665, *SD*=.094) were compared to post-test accuracy scores for untrained items (*M*=.745, *SD*=.123). Response latency mean scores were also contrasted between trained (*M*=817.19, *SD*=194.83) and untrained (*M*=740.90, *SD*=156.43) items. The paired samples t-test reported statistically significant differences for accuracy scores (*t* (17) = -3.465 *p*=.003, *r*=.643). As for reaction time responses, learners were also faster at untrained items in the post-test, *t* (17) = 2.151 *p*=.046, *r*=.462). See table 8.41 and figure 5.10.

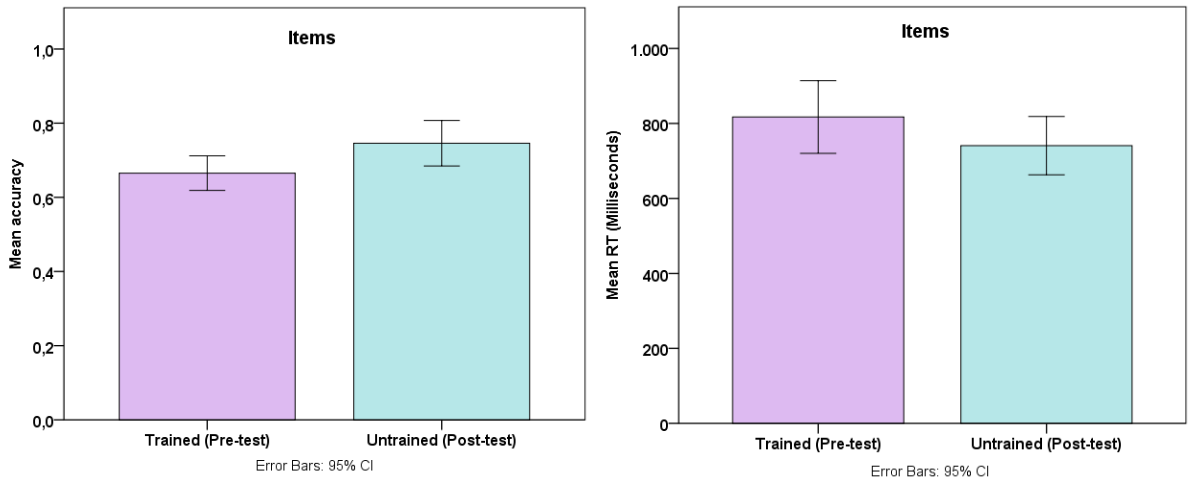


Figure 5.10. *Bar graph for accuracy and RT scores in the discrimination of trained vs. untrained items at pre-test and post-test.*

In addition, a paired samples t-test showed that there were statistically significant differences between the identification of trained speakers (*M*=.665, *SD*=.094) in the pre-test and untrained speakers (*M*=.768, *SD*=.112) in the post-test, *t* (17) = -5.194 *p*<.001, *r*=.783) (see table 8.41 and figure 5.11). As for RT measures, stimuli from trained speakers (*M*=817.19, *SD*=.194.83)

obtained significantly slower RT than untrained speakers (*M*=728.67, *SD*=156.88), *t* (17) = 2.588 *p*=.019, *r*=.531.

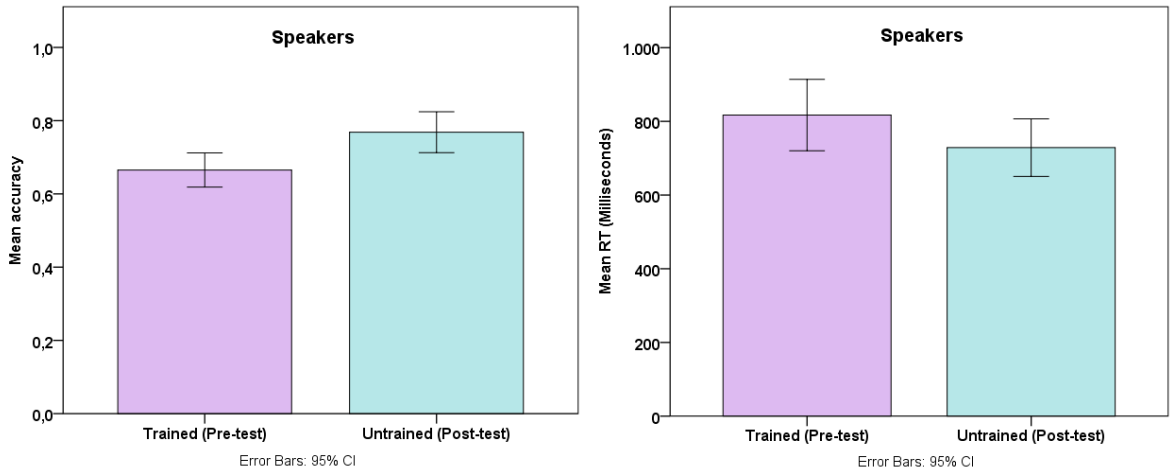
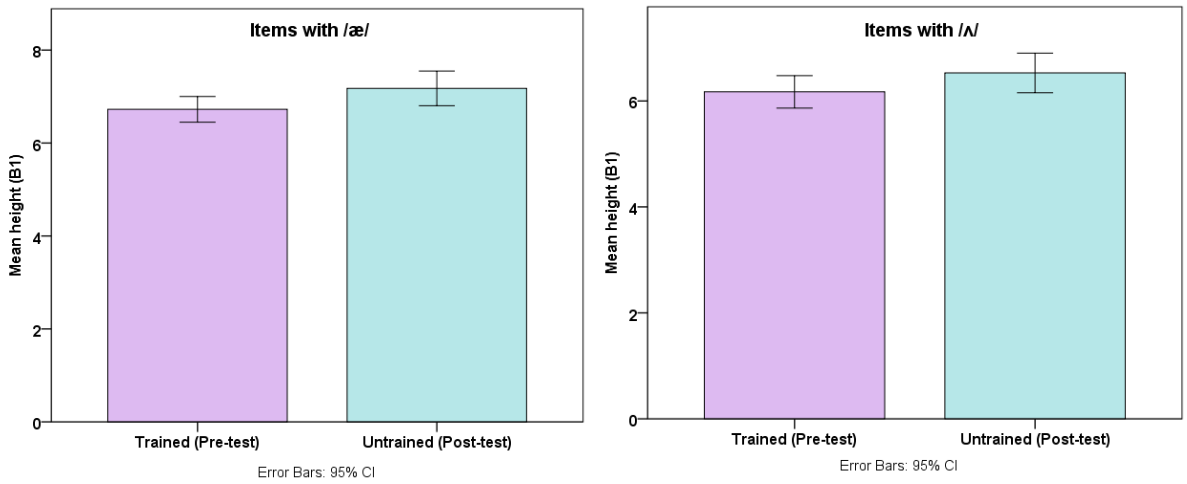


Figure 5.11. *Bar graph for accuracy and RT scores in the discrimination of trained vs. untrained speakers at pre-test and post-test.*

When analysing the independent variables of the production test (i.e. height, advancement and Euclidean distance), three paired-samples t-test determined that, concerning vowel /æ/, learners significantly moved their vowel to a lower position with pre-test untrained (*M*=7.175, *SD*=.749) than post-test trained (M=6.725, SD=.553) items, *t* (17) =-3.505, *p*=.003, *r*=.647, but no significant changes were observed for advancement (*t* (17) =-.377, *p*=.711, *r*=.091). As regards vowel /ʌ/, learners also experienced significant changes in height, generalizing the lowering of the vowels in post-test untrained items (*M*=6.52, *SD*=.754), *t* (17) =-2.376, *p*=.030, *r*=.499; however, they did not move their /ʌ/ vowel towards the back of the oral cavity, *t* (17) =-391, *p*=.701, *r*=.094. Finally, the Euclidean distance confirmed that learners were able to generalize the acquired contrast to non-words in the pre-test by making the distance between the two vowels larger (*t* (17) =-4.742, *p*<.001, *r*=.754). See table 8.46,8.47 and figure 5.12.



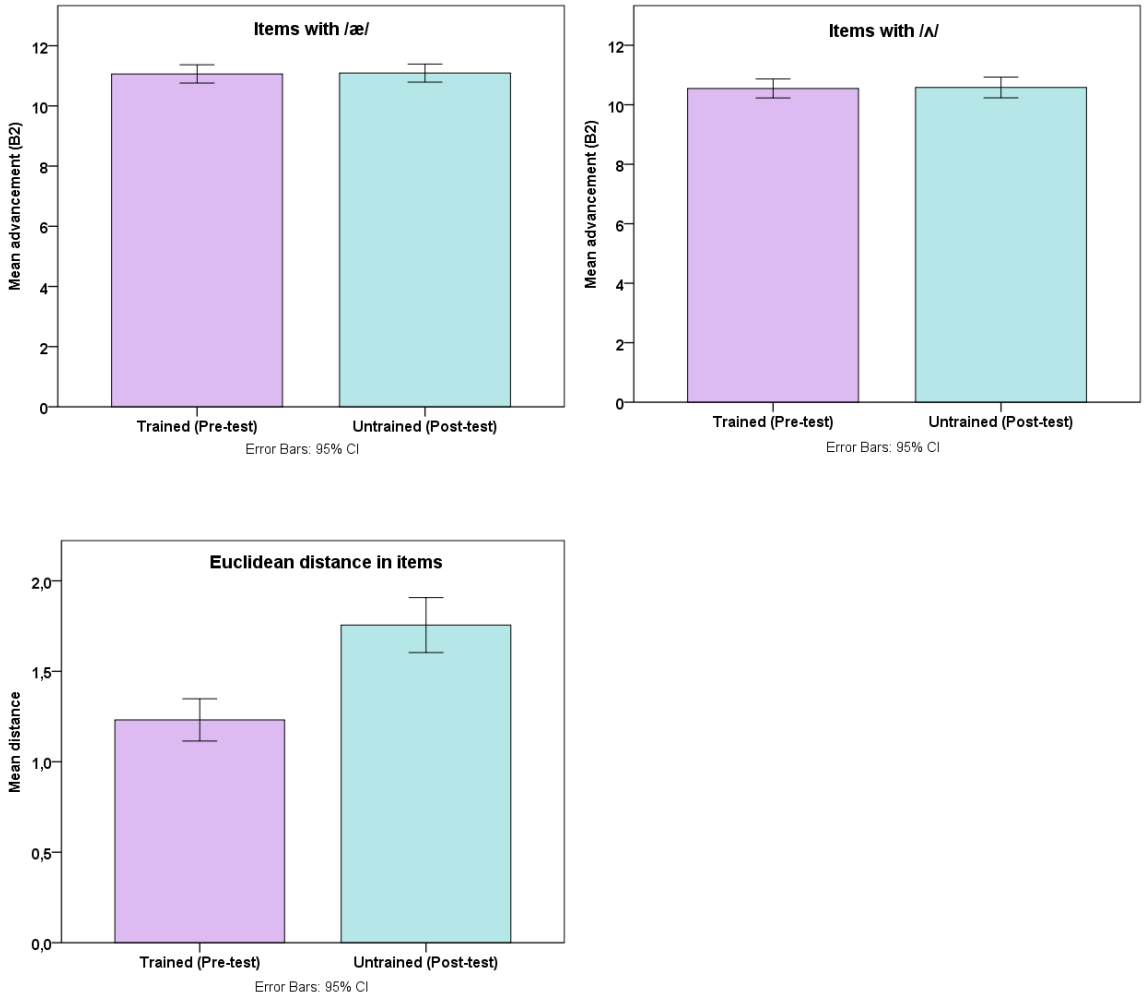


Figure 5.12. *Bar graphs for B1, B2 and E.D. gains in the production of words uttered by trained vs. untrained items at pre-test and post-test.*

Furthermore, a paired samples t-test revealed no differences between trained and untrained voices for vowel /æ/: B1 (*t* (17) =-1.056, *p*=.306, *r*=.248), B2 (*t* (17) =-.450, *p*=.659, *r*=.108), and vowel /ʌ/: B1 (*t* (17) =.728, *p*=.477, *r*=.173), B2 (*t* (17) =.890, *p*=.386, *r*=.210). Finally, ED results (*t* (17) =.739, *p*=.470, *r*=.176) confirmed that, since the differences were very small even to appreciate gains, we could not talk about generalization to new speakers (see figure 5.13).

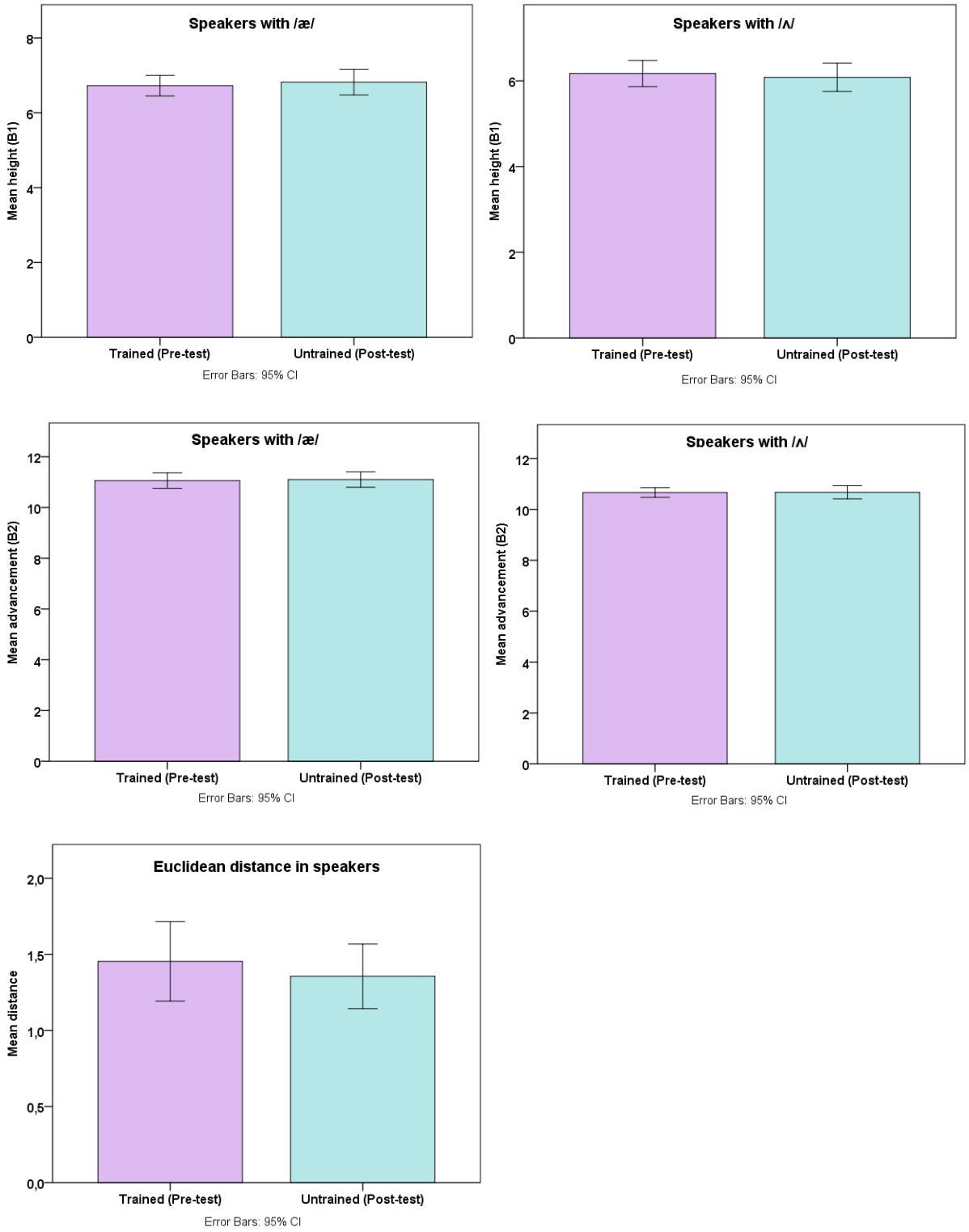


Figure 5.13. *Bar graphs for B1, B2 and ED gains in the production of words uttered by trained vs. untrained speakers at pre-test and post-test.*

Summary of perception and production results

- Learners in the experimental group became significantly more accurate in the identification of the target phonological contrast from pre-test to post-test and retained their gains in the delayed post-test, whereas the control group did not obtain any gains.
- Learners in the experimental group became significantly more accurate and faster at discriminating the target phonological contrast from pre-test to post-test and retained their gains in the delayed post-test. The control group did not become more accurate.
- Learners in the experimental group increased the spectral distance for /æ/ - /ʌ/ from pre-test to post-test and did not lose this knowledge in the delayed post-test. Although none of them reached native-likeness, females approached NSs values more.
- Learners in the experimental group were able to generalize their perception gains to novel items and speakers, and to transfer production gains to novel items mainly.

5.2. Task complexity and LREs

5.2.1. Effects of task complexity

Three paired-samples t-tests showed that learners and teachers perceived task demands in a similar way. Statistical analyses showing significant differences across tasks are reported in appendix E.3. Due to space limitations, learners' answers on different affective factors will not be thoroughly discussed here; however, learners thought that they had spent significantly less time when the task was simple than when it was complex (X_2 (3) =39.705, p<.001). See appendix E.3 for specific analyses.

Given that the inter-rater reliability was 91.6% across tasks (see appendix G.5, figures 9.1 and 9.2), a one-way ANOVA was used to assess the effect of task complexity (±reasoning demands and ±number of elements) on the occurrence of all LREs. The parametric test showed that there was a significantly main effect of LREs across tasks, (F (3,15) = 42.630, p<.001, η^2 =.895) (table 9.2). Bonferroni pairwise comparisons revealed that there were statistically significant differences for LREs between task 1 (M=3.89, SD=2.96), task 2 (M=6.00, SD=2.91) and task 3 (M=10.00, SD=4.82), (p<.05); however, the occurrence of LREs between task 3 and task 4 did not reach significance (M=13.11, SD=4.56), p=.086, see figure 6.1. In addition, see appendix G.5 (figure 9.3, table 9.1) for the same analysis with general LREs, exclusively, which obtained similar results.

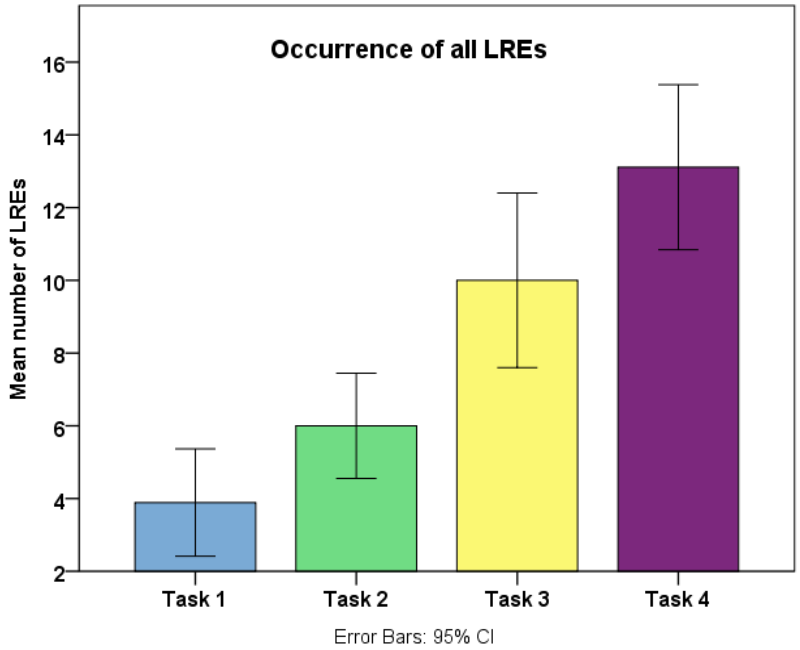


Figure 6.1. *Mean scores for the occurrence of all LREs (general LRE, recasts, self-repairs and repetitions).*

When analysing LREs per minute, the one-way ANOVA confirmed that a significantly lower number of LREs occurred in simple rather than complex tasks, (F (3,15) = 7.747, p=.002, η^2 =.608) (figure 6.2); however, Bonferroni pairwise comparisons revealed that main differences were among task 1 and task 2,3,4, (p<.05) but no statistically significant differences were found between task 2 and 3 or 4 (see appendix G.5 [tables 9.1, 9.5, 9.6 and 9.7). In short, this shows a significant main effect of complexity on the number of LREs so the more complex the task is, the higher the occurrence of LREs.

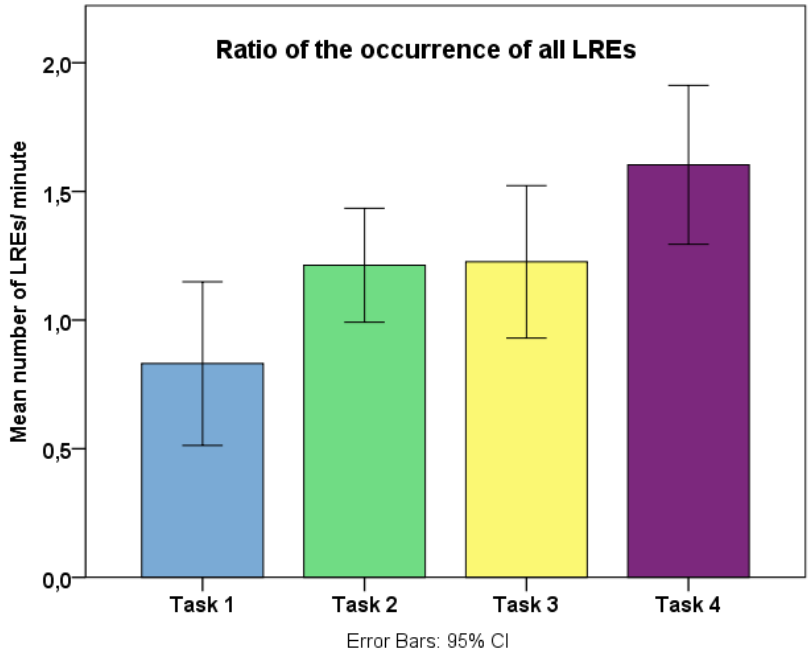


Figure 6.2. *Mean scores for the occurrence of all LREs per minute.*

To sum up, learners significantly engaged in more LREs when tasks were cognitively more complex irrespectively of time-on-task. Finally, the different kinds of LREs were analysed separately in order to appreciate differences among them (within-subjects analysis) and across tasks (between-subjects analysis) but will only be reported in appendix G.5, due to space limitations.

5.2.2. LREs and gains

LREs were also analysed in comparison to learners' accuracy gains in the perception and production of /æ/ and /ʌ/. When selecting all types of language related episodes (GenLRE, LRERC, LRESR-err, LREREP), the *Pearson-r* correlation revealed that the more language related episodes learners produced, the larger the size of gains in Euclidean distance (r=.479, p=.044) learners obtained. However, it was only in the male group that there was a strongly significant correlation (r=.704, p=.034) whereas the female group did not show such strong relationship. No other significant correlations were found (see appendix G.5., table 9.12). We can conclude that the more often they paid attention to the phonological contrast, the better able they were to distinguish between the two contrasting vowels in production.

5.3. Learner factors

5.3.1 Proficiency

No correlations were found between proficiency (X_Lex adjusted scores) and gains in identification, discrimination or production for male and female learners. (See table 10.1).

5.3.2 Attention

Two measures were taken: one of inhibition control (RT) and one of selective attention (accuracy). Before analysing the effect of inhibitory control on learners' perception and production gains, the experimental group showed differences between the target congruent (e.g. female voice = 'noia') and the target incongruent condition (e.g. female voice= 'home'), t (17) = -2.720, p= .015, r=.519. In contrast, learners responded equally fast with congruent fillers (e.g. female voice= 'oca' [feminine noun]) and incongruent fillers (e.g. female voice= 'oli' [masculine noun]), t (16) = 1.416, p=.176, r=.333. (See appendix G.6., table 10.2, figure 10.1). A *Pearson-r* correlation showed that there were no significant correlations (p>.05) between

learners' differences in inhibitory control and gains in production and perception (see table 10.3, appendix G.6).

Concerning the auditory selective attention test, a paired samples t-test informed that there were no significant differences for accuracy in terms of colour or digit responses, t (17) = -.559, p=.584, r=134 (see table 10.4 and figure 10.2 in appendix G.6). After this, a *Pearson-r* negative correlation showed that selective attention predicted more than 75% of gains in the discrimination of the target vowels (r=-.862, p=.003). This indicates that learners who were better at focusing their attention were faster at recognizing the L2 phonological contrast (see figure 7.1 and appendix G.6: table 10.5). In addition, when separating low from high-proficiency learners, it was found that low-proficiency learners' selective attention explained more than 50% of the variance in height gains for /æ/ (r=.726, p=.027), (see table 10.6).

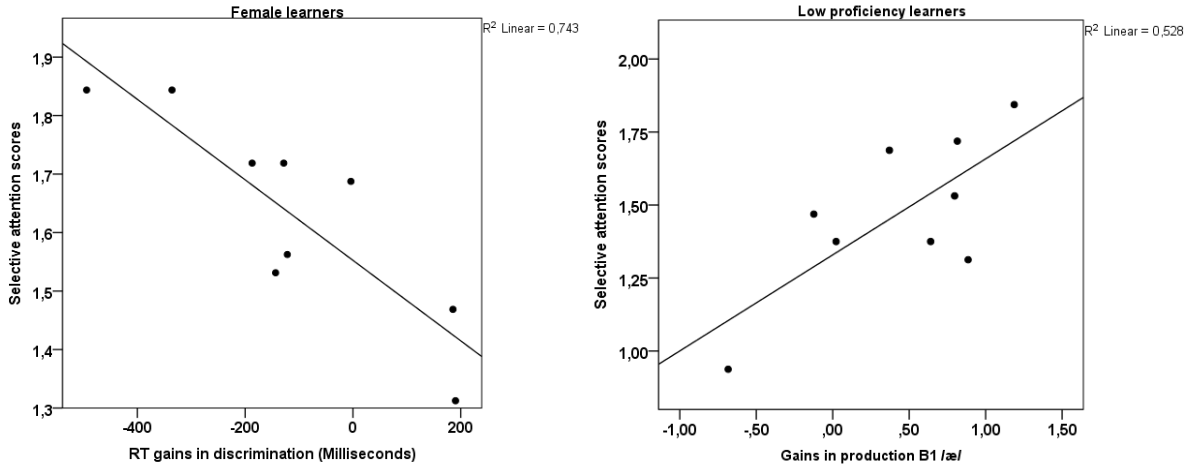


Figure 7.1. *Scatterplot showing the Pearson correlation between selective attention and RT gains in discrimination (female learners) [left side] and Pearson correlation between selective attention and B1 gains in /æ/ (low proficiency learners) [right side].*

Last but not least, different measures of attention control were contrasted. Given that the results were expressed with different values (RT vs. accuracy), a chi square test was used to observe the significant relationship between two nominal variables (low and high inhibitory control vs. low and high selective attention). The test reported that there was no statistically significant association between inhibition control and selective attention neither in the experimental group (males: $\chi(1)$ =.900, p =.343, females: $\chi(1)$ =.225, p =.635) nor in the control group (males: $\chi(1)$ =1.215, p =.270, females: $\chi(1)$ =2.396, p =.122), which may mean that the two tests measured two very different constructs (see appendix G.6, table 10.7).

6. Discussion

6.1. Task effects on perception and production

6.1.1. Perception and production gains

The first research question of this study asked about the effects of task-based instruction on the perception and production of the English vowel contrast /æ/ - /ʌ/. In general, results confirmed previous findings in that pronunciation-focused interaction and negotiation during task performance provides excellent opportunities for L2 pronunciation learning (Saito 2013, 2015).

Concerning identification test outcomes, learners from the experimental group significantly improved from pre-test to post-test -outperforming the control group- as well as retained their gains in L2 speech perception after two weeks of the treatment. These results are in line with HVPT studies, which also found identification gains after focusing on phonetic form (Carlet & Cebrian, 2015); nevertheless, learners may have acquired the contrast in a more naturalistic and motivating way during interaction. Turning our attention to discrimination (ABX) test gains, the experimental group became more accurate and faster at post-test as well as at delayed post-test, indicating initial internalization of the English vowel contrast. Although the control group did not improve their accuracy, response latencies significantly decreased from pre-test to post-test, possibly coming from practice effects. In any case, the control group did not improve their accuracy in vowel discrimination and the experimental group was faster when the three testing times were considered. This positive outcome can only be explained by the effectiveness of tasks along resource-directing variables to push learners to focus on phonetic form during interaction, hence, increasing their sensitivity to the L2 vowel contrast. With respect to production results, the experimental group showed a larger amount of improvement than the control group, which did not learn to distinguish the two central vowels because did not do any task. Despite not reaching native-like vowel quality, learners produced several interesting movements in the vowel space. As Saito (2013, 2015) mentions, learners exhibit several interlanguage forms before reaching L2 speech intelligibility because target phonological features need to be practiced in authentic contexts during a long period of time. However, increasing communicative and cognitive demands forces learners to push production, stretch interlanguage and destabilize fossilized forms (Gilabert, 2007). In general, learners relied on vowel height (F1) to produce the contrast, hence, they lowered their /æ/ and heightened their /ʌ/ in their vowel spaces. When it comes to advancement (F2), it appears that learners had a tendency to retract both vowels towards the back of their oral cavity. Following the SLM and

PAM-L2 models, L2 vowels that are acoustically more distinct from the nearest L1 vowels are perceived more accurately. Initially, learners perceive English /æ/ and /ʌ/ as Spanish/Catalan /a/ in a 100% and 85% percentage assimilation for B1 and B2, as reported by Cebrian et al. (2010). In this study, post-test the target vowels became less fronted and lower in the oral cavity. In addition, learners' Euclidean distance between the two vowels became larger at post-test. When assessing learners' distinction in terms of B1 and B2, both males and females perceived a substantial difference at post-test but the Euclidean distance increased significantly for female learners. As for gains in height, although all learners reduced the distance with respect to NSs between pre- and post-test for vowel /æ/, only female learners showed a significant reduction at delayed post-test. Similarly, male and female learners reduced the distance with respect to NSs for /ʌ/at post-test but only males retained this vowel position at delayed post-test. Individual gains showed that female learners approached the native-like height of /æ/ and /ʌ/ more than male learners, even if their production of /ʌ/ decreased at delayed post-test. In contrast, males performed significantly better only for the /ʌ/ vowel but showed more retention than female speakers at time 3. In sum, the performance of these four tasks, which increased in resource-directing variables and decreased in resource-dispersing variables[6] allowed them to engage in L1 and L2 form-meaning mappings and draw attention to how the phonological contrast was encoded during performance (Gilabert, 2007). Finally, the initial hypothesis concerning higher gains in perception than production was discarded because no significant differences were reported between the perception and production of L2 vowels. This is reasonable because complete intelligibility and native-like levels of production and perception can only be attained after a long treatment.

6.1.2. Generalization to new contexts

The experimental group demonstrated not only improvement and retention at the perceptual level, but also generalization of gains to new speakers and items, suggesting that a focus on phonetic form during task-based performance allowed learners to use their recently acquired knowledge. These findings corroborate previously reported generalization effects coming from the high-variability phonetic training regime (Iverson & Evans, 2009; op cit Pereira, 2014), which can now be applied to more interactive tasks.

[6] Resource-dispersing variables were reduced by (a) increasing the familiarity with the target items during pre-task; (b) giving online planning time before tasks; and (c) repeating the same type of tasks four times.

In the case of production, where gains were not so large, learners did not generalize the trained L2 vowel knowledge to novel speakers but they did so to novel items, where there were statistically significant differences in terms of height (B1). As a matter of fact, post-test oral reports informed that some native speakers had not been easy to understand, which may have been referred to the new voices that were incorporated. Perhaps generalization to absolute new speakers is talker-specific and requires much more practice in second language environments. In any case, the present results suggest that task-based pronunciation teaching is clearly effective for the development, retention and generalization of L2 segmental phonology and, besides, may be comfortably incorporated in EFL lessons.

6.2. Task complexity effects on the occurrence of LREs

The second research question aimed to examine whether the prediction of the Cognition Hypothesis (Robinson, 2001, 2007, 2011) could be extended to L2 pronunciation. According to this theory, greater task complexity encourages greater incidence of form-focused episodes that, consequently, may lead to the development of interlanguage. In line with previous studies on grammar (Baralt, 2013) and pragmatics (Kim & Taguchi, 2015), this study revealed that the employment of four decision-making tasks, which had a clear focus on pronunciation, improved learners' L2 accuracy. Furthermore, in line with Baralt (2013) and Gilabert's (2007) findings, learners' attention towards language was especially directed in complex tasks than simple tasks because they required more precise linguistic resources from the conceptual demands of the task. This may also reflect multiple-resource models (Wickens, 1989; op cit Gilabert, 2007) that suggest that attention may be distributed among different resource dimensions, as opposed to attention models that perceive attention as a single volume that runs out of resources.

This study found that a significantly greater amount of pronunciation-based language-related episodes were generated in the more complex tasks when all kinds of LREs were considered (i.e. general LREs, recasts, self-repairs and repetitions) as well as with only general LREs. Interestingly, the occurrence of LREs per minute was also higher in the complex tasks, which indicates that learners were reflecting on form to a higher extent in complex tasks regardless of time-on-task. This finding goes against Solon et al.'s (2017) study, who found that simpler tasks produced more LREs, albeit not significantly. These researchers argued that whereas grammatical targets have specific forms that can be described with metalinguistic rules, phonetic targets cannot because they are part of a gradient range of production possibilities and require the physical modification of the articulators. Contrary to this statement, here we suggest

that, if tasks have a clear focus on phonetic form and make the target phonological contrast essential during meaningful interaction, learners are able to negotiate the target form explicitly and implicitly (i.e. direct corrections, recasts or repetitions). As in Solon et al.'s (2017) experiment, pronunciation training was not a part of the learners' curriculum and, even if they were not used to verbally reflecting on phonetic form, learners developed strategies to reach intelligibility during conversation. The result was a higher production of LREs regardless of the real time-on-task. All in all, it seems that increasing cognitive complexity through task design geared attention towards form in productive ways (i.e. interlocutors paid attention to form more often in complex tasks).

In addition, the number of LREs was related to the size of the gains in production, especially, for male learners. The fact that the correlation was not significant considering the whole group may be due to the fact that each dyad was assigned the same number of LREs, therefore, individual differences in the generation of LREs could not be accounted for. A possible solution would be dividing learners into generators vs. gainers of LREs; nonetheless, this may be carried out in further TBPT studies. Another reason may be that it was not only the negotiation of form that generated the LREs but it was the overall task design *per se* that triggered many opportunities to focus on the language they were producing. In other words, gains in perception and production may have occurred beyond the presence of LREs because the nature of the task design already induced a focus on phonetic form. For example, Sicola (2009) found that learner-learner dyads modify their productions in the target-like direction if tasks are carefully designed with a clear task-essential language during interaction.

6.3. Learner factor effects on perception and production gains

Out of all individual factors that may have affected learners' performance in this task-based pronunciation teaching study (see Szalkowska-Kim, 2014), this study assessed the effects of proficiency and attention on learners' gains in L2 speech perception and production. Similar to Lee et al.'s (2015) findings, gains in perception and production did not depend on proficiency. Several reasons can account for these results. Firstly, the X_Lex may not have been the appropriate measure of proficiency regarding this experiment, hence, a measure of L2 speech perception/production from an elicited imitation task (Ortega et al., 2002) would have been ideal. Secondly, the use of task-based instruction may have levelled out individual differences not only because it is an analytic approach that respects learners' developmental stages, but also

because tasks involve two minds working and competing at the same time. These findings contrast those of Trofimovich and Gatbonton's (2006), who found significant differences in terms of accuracy for low-proficiency learners.

Besides proficiency, differences in attention control were assessed. According to Moyer (2014), phonological development relies heavily on speech motor control and auditory-perceptual mechanisms so processes which are extremely efficient in the L1 may not be so in the L2. This TBPT study did not find any relationship between inhibitory control and perception gains, perhaps because the inhibitory control test may not have been sensitive enough to capture inter-subject variation. In this study, gains in L2 phonology seemed to be more closely linked to auditory selective attention. In line with Aliaga et al. (2011), who found that phonological short-term memory was linked to L2 vowel discrimination, learners who had a high auditory selective attention were faster at selecting and discriminating the target vowels in the perception test. Moreover, when learners were classified according to proficiency, low-level learners who had a high selective attention exhibited higher gains in the production of /æ/ than high-level learners, which may point at the fact that these learners use attention to a greater extent when it comes to learning the phonology of the second language. Therefore, learners develop more accurate representations of the L2 segments by virtue of their capacity to select the target information in L2 speech. Having analysed these two learner factors, I would call for further research in the area of individual differences and TBPT in order to understand which factors play a significant role for L2 phonological learning in EFL contexts.

7. Conclusion

7.1. General discussion and implications

After decades of investigation on task-based language teaching and L2 speech acquisition, this is one of the first studies that explores whether the benefits of tasks can be extended beyond grammar and lexis (Gurzynski et al., 2017) and whether task manipulation helps enhancing intentional focus on phonetic form. Considering that L2 speech learning requires a considerable use of auditory- and articulatory-based attentional resources, this study dealt with four carefully-designed, real-world tasks which directed learners' attention to phonetic form during meaningful interaction. In this way, learners were able to notice the gap between their peers' productions and their own as well as engage in metalinguistic reflection in the form of output (Robinson, 2011). Moreover, this analytic focus on form method respected learners'

developmental stage and processing ability, while making the target form essential for task completion. In addition, tasks were manipulated along cognitive complexity, which resulted in the higher occurrence of LREs and, hence, improvement in L2 segmental (vowel) accuracy. Nevertheless, internal cognitive processes generated by task design (i.e. instances of noticing, intake…) were not targeted nor captured by our output measures. Despite the general improvement in perception and production of the vowels /æ/ and /ʌ/, results showed large individual differences in the amount of gains obtained. Whereas proficiency did not seem to exert a strong effect on L2 outcomes (Lee et al., 2015), this is the first study to show the role of auditory selective attention in learning an L2 vowel contrast; consequently, further research on TBPT should consider selective attention as a potential moderator of L2 speech gains. Finally, this study has provided solid evidence about the potential benefits of tasks on L2 pronunciation as well as the importance of complex tasks for the acquisition of L2 phonological features. Nevertheless, TBPT is not likely to work efficiently if English teachers do not have a sufficient understanding of pronunciation and develop awareness of the suprasegmental features (e.g. syllabic structures, rhythm, stress and intonation) and segmental features of the second language (e.g. vowel contrasts and voice onset time). Furthermore, teachers need to integrate the aforementioned phonological features in the EFL classroom by paying special attention to pronunciation while taking into account all the other aspects such as grammar, semantic discourse and pragmatics involved in transactions, interaction and communication generally (Taylor, 1991). Finally, according to Burgess and Spencer (2000), pronunciation is best dealt with as the need arises rather than in an extremely predetermined way so it is recommended to carry out needs analyses with questionnaires, listening comprehension/discrimination tests and production samples (Celce-Murcia & Godwin, 1991). In sum, phonetic forms can be processed and learned in motivating task-based lessons where the overriding focus is on meaning if tasks are well-crafted and make the target phonetic forms essential.

7.2. Limitations and further research

This small-scale short-term training study is not exempt of several limitations. Firstly, the experimental and control groups were not comparable in terms of L2 proficiency. To eliminate any confounds, statistical analyses were run to ensure the comparability of the groups. Secondly, according to learners' reports, the delayed-sentence repetition task appeared to be very complicated, hence, perhaps a delayed word repetition task would have allowed learners to focus more on accuracy and slightly less on memorization. Likewise, despite the fact that

vocabulary size test scores were used as a proxy for proficiency, an elicited imitation task (Ortega et al., 2002) might have provided a better measure of oral proficiency. Thirdly, vowels were taught contrastively and not in connection with a subset of difficult vowels. This may be problematic because the phonological features of a language cannot be learned in relation to one single representation as they are surrounded by multiple contrasts. Finally, the analysis of production only focused on one specific acoustic correlate -formant structure of vowels- without taking into account other aspects such as duration, that may definitely be important for other phonological contrasts (e.g. /i:/ and /ɪ/).

After acknowledging those limitations, I would like to encourage further investigation on the role of tasks and task manipulation (i.e. task complexity, task repetition, task modality, etc.) for the attainment of intelligibility in L2 segmental and suprasegmental learning. Although this experiment has taken sequencing into consideration (SSARC[7] model, Robinson 2010), it has not specifically tested the effect of different sequences, which would be interesting to analyse, as some studies have shown an advantage of simple-to-complex sequences (Levkina & Gilabert, 2014) whereas others have not (Baralt, 2014). Moreover, researchers need to invest time on the exploration of learner factors which may be crucial for TBPT research. These include L1 background, age, proficiency, affective and cognitive factors, among others. Finally, it would be interesting to study cross-linguistic influence in the occurrence of LREs; for instance, whether an L1-mixed classroom enhances more awareness of L2 pronunciation errors (i.e. more LREs) given that learners do not share the same phonetic repertoire. Video-based interactive tasks would be the perfect avenue to provide an answer to this research question.

As a conclusion, this study has contributed to paving the way for empirical research on the effectiveness of tasks to promote L2 pronunciation learning. Therefore, it is my hope that specialists on L2 speech acquisition and TBLT join their areas of expertise to carry further research on the field of task-based pronunciation teaching which, without a doubt, will have a huge impact on SLA research and pronunciation instruction.

[7] SAARC stands for *stabilize, simplify, automatize, restructure,* and *complexify.* This model posits that (a) task sequencing should be based on cognitive complexity factors and (b) tasks should increase first in resource-dispersing dimensions and then, in resource-directing dimensions.

References

Aliaga-Garcia, C., Mora, J. C. and Cerviño-Povedano, E. (2011). Phonological short-term memory and L2 speech learning in adulthood. *Poznań Studies in Contemporary Linguistics,* 47.1, 1-14.

Bachman, L. F. & Palmer A. S. (1996). *Language Testing in Practice*. Oxford: Oxford University Press.

Baralt, M. (2013). The impact of cognitive complexity on feedback efficacy during online versus face-to-face interactive tasks. *Studies in Second Language Acquisition,* 35, 689-725.

Baralt, M. (2014). Task complexity and task sequencing in traditional versus online language classes. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 95–122). London, UK: Bloomsbury.

Best, C. & Tyler, M. (2007). Nonnative and second language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. Munro (Eds.), Language Experience in Second Language Speech Learning -In honour of James Emil Flege (pp.13-34). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D.B. & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/ contrast. Studies in Bilingualism Phonology and Second Language Acquisition, 287-308.

Burgess, J., & Spencer, S. (2000). Phonology and pronunciation in integrated language teaching and teacher education. *System,* 28, 191-215.

Bygate, M., Skehan, P. & Swain M. (2001). *Researching pedagogic tasks: second language learning, teaching and testing.* Harlow: Longman.

Candlin, C.N. (1987). "Towards task-based language learning". In Candlin, C.N. and D. Murphy (eds.), *Language Learning Tasks*. Lancaster Practical Papers in English Language Education, vol. 7 (pp. 23-46). London: Prentice-Hall International (UK) Ltd and Lancaster University.

Carlet, A. & Cebrian, J. (2015). Identification vs. discrimination training: Learning effects for trained and untrained sounds. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences.* Glasgow, UK: University of Glasgow. Retrieved from http://www.icphs2015.info/pdfs/proceedings.html.

Cebrian, J., & Carlet, A. (2014). Second-language learners' identification of target-language phonemes: a short-term phonetic training study. *Canadian Modern Language Review, 70,* 474-499.

Cebrian, J., Mora, J.C. & Aliaga García, C. (2010) Assessing crosslinguistic similarity by means of rated discrimination and perceptual assimilation tasks. A K. Dziubalska-Kolaczyk, M. Wrembel and M. Kul (eds), *New Sounds 2010 Proceedings of the 6th International Symposium on the Acquisition of Second Language Speech*, pp. 77-82.

Celce-Murcia, M. & Godwin, J. M. (1991). Teaching pronunciation. In Marianne Celce-Murcia (Ed.), *Teaching English as a Second or Foreign Language* (pp. 136-153). United States of America, USA: Heinle & Heinle Publishers.

Crookes, G. (1986). *Task classification: A cross-disciplinary review* (Tech. Rep. No.4). Honolulu: University of Hawaii at Manoa, Social Science Research Institute, Center for Second Language Classroom Research.

Darcy, I., Mora, J. C. & Daidone, D. (2016). The role of inhibitory control in second language phonological processing. *A Journal of Research in Language Studies*, 1-33.

Darcy, I., Park, H., & Yang, C. L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences, 40,* 63-72.

Deterding, D. (1997). The formants of monophthong vowels in standard southern British English Pronunciation. *Journal of the International Phonetic Association,* 27, 47–55.

Elliot, A. R. (1997). On the teaching and acquisition of pronunciation within a communicative approach. *American Association of Teachers of Spanish and Portuguese,* 80.1, 95-108.

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

Flege, J. E. (1995). Second-language speech learning: theory, findings, and problems. In Strange, W. (ed), *Speech Perception and Linguistic Experience: Issues in Cross-language research.* Timonium, MD: York Press, Pp. 229-273.

Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching, 45,* 215-240.

Gilabert, R., Baron, J., & Llanes, M.A. (2009). Manipulating cognitive complexity across task types and its impact on learners' interaction during task performance. *International Review of Applied Linguistics, 47,* 367-395.

Gordon, J., Darcy, I., & Ewert, D. (2013). Pronunciation teaching and learning: Effects of explicit phonetic instruction in the L2 classroom. In J. Levis & K. LeVelle (Eds.). *Proceedings of*

the 4th Pronunciation in Second Language Learning and Teaching Conference. (pp. 194-206). Ames, IA: Iowa State University.

Guion, S. G., & Pederson, E. (2007). Investigating the role of attention in phonetic learning. In O-S. Bohn & M. J. Munro (Eds.) *Language experience in second language learning* (pp. 57-77). Amsterdam: John Benjamins.

Gurzynski-Weiss, L., Long A.Y. & Solon M. (2017). TBLT and L2 Pronunciation: Do the benefits of Tasks Extend beyond Grammar and Lexis? *Studies in Second Language Acquisition,* 39, 213-224.

Hazan, V., Sennema, A., Iba, M. & Faulkner, A. (2005). Effect of audiovisual perceptual training on the perception and production of consonants by Japanese learners of English. *Speech Communication,* 47.3, 360-378.

Isaacs, T. (2009). Integrating form and meaning in L2 pronunciation. *TESL Canada Journal,* 27.1, 1-12.

Jung, Y., Kim, J. & Murphy, J. (2017). The role of task repetition in learning English stress patterns through auditory priming tasks. *Studies in Second Language Acquisition*, 39, 319-346.

Kim, Y. & Taguchi, N. (2016). Learner–learner interaction during collaborative pragmatic tasks: The role of cognitive and pragmatic task demands. *American council on the teaching of foreign languages*, 49.1, 52-57.

Lee, J. Jang, J. & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A Meta-Analysis. *Applied Linguistics,* 36.3, 345-366.

Lev-Ari, S., & Peperkamp, S. (2014). The influence of inhibitory skill on phonological representations in production and perception. *Journal of Phonetics, 47,* 36-46.

Loewen, S. & Isbell, D. R. (2017). Pronunciation in face-to-face and audio-only synchronous computer-mediated learner interactions. *Studies in Second Language Acquisition*, 39, 225-256.

Long, M. (1985). A role for instruction in second language acquisition: Task-based language teaching. In K. Hylstenstam & M. Pienemann (eds.), *Modelling and assessing second language acquisition* (pp. 77-99). Clevedon: Multilingual Matters.

Long, M. H. (1991). Focus on form: A design feature in language teaching methodology. In K. de Bot, R. Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp. 39-52). Amsterdam: John Benjamins.

Long, M. H. (2015). *Second language acquisition and task-based language teaching*. Malden, MA: Wiley-Blackwell.

Long, M. H., & Robinson, P. (1998). Focus on form: theory, research and practice. In Doughty, C. J., & Williams, J. (eds.), *Focus on form in second language acquisition* (pp. 15-41). Cambridge: Cambridge University Press. Reprinted in Ortega, L. (ed.), *Second language acquisition: Critical concepts in linguistics*. London: Routledge.

Loschky L., Bley-Vroman R. (1993). Grammar and task-based methodology. In Crookes G., Gass S. (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123–67). Clevedon: Multilingual Matters.

Malicka, A. (2014). *The role of task complexity and task sequencing in L2 monologic oral production*. Retrieved from Dipòsit Digital UB (http://diposit.ub.edu/dspace/bitstream/2445/62643/1/Aleksandra_Malicka_THESIS.pdf).

Martínez Celdrán, E. & Fernández Planas, A.M. (2007). *Manual de fonética española*. Barcelona: Ariel.

Meara, P. M., & Milton, J. (2003). *X_Lex*. Swansea, UK: Lognostics.

Michel, M. C. (2011). Cognitive and interactive aspects of task-based performance in Dutch as a second language. Oisterwijk, University of Amsterdam: Boxpress.

Mora-Plaza, I. (2016). The "Cinderella" of foreign language teaching. An insight to the integration of pronunciation in EFL contexts. Retrieved from Dipòsit Digital UB (TFG - Estudis Anglesos), (http://hdl.handle.net/2445/107295).

Mora, J. C. & Levkina, M. (2017). Task-based pronunciation teaching and research: key issues and future directions. *Studies in Second Language Acquisition*, 39, 381-399.

Morin, R. (2007). A neglected aspect of the standards: preparing foreign language Spanish teachers to teach pronunciation. *Foreign Language Annals*, 40.2, 342-360.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly*, 25.3, 483-520.

Moyer, A. (2014). Exceptional outcomes in l2 phonology: the critical factors of learner engagement and self-regulation. *Applied linguistics*, 1-24.

Munro, M. J. & Derwing, T. M. (2015). Intelligibility in research and practice: teaching priorities. In M. Reed and J.M. Levis (Eds.), *The Handbook of English pronunciation* (pp. 377-396). Malden, MA: Wiley Blackwell.

Muñoz, C. (2008). Symmetries and asymmetries of age effects in naturalistic and instructed L2 learning. *Applied Linguistics*, 24(4), 578-596.

Murphy, J. & Baker, A. (2015). History of ESL pronunciation teaching. In M. Reed & J. Levis (eds.), *The handbook of English pronunciation*, (pp. 36-65). Malden, MA: Wiley Blackwell.

Ortega, L., Iwashita, N., Norris, J. M., & Rabie, S. (2002, October). An investigation of elicited imitation tasks in crosslinguistic SLA research. Paper presented at the *Second Language Research Forum*, Toronto, Canada.

Parlak, Ö. & Ziegler, N. (2017). The impact of recasts on the development of primary stress in a synchronous computer-mediated environment. *Studies in Second Language Acquisition*, 39, 257-285.

Pennington, M. C. (1996). *Phonology in English Language Teaching.* London and New York, NY: Longman.

Pereira, Y. I. (2014). *Perception and production of English vowels by Chilean learners of English: Effect of auditory and visual modalities on phonetic training* (Published doctoral dissertation). University College London, London, UK.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communication tasks for second language research and instruction. In G. Crookes & S. Gass (Eds.), *Tasks and second language learning* (pp. 9–34). Clevedon, UK: Multilingual Matters.

Piske, T. (2008). Phonetic awareness, phonetic sensitivity and the second language learner. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of language and education Vol. 6: Knowledge about language* (pp. 155-166). New York, NY: Springer Science.

Prabhu, N.S. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Rallo-Fabra, L. & Romero J. (2012). Native Catalan learners' perception and production of English vowels. *Journal of Phonetics*, 40, 491-508.

Révész, A. (2011). Task complexity, focus on L2 constructions, and individual differences: a classroom-based study. *The Modern Language Journal*, 95, 162-181.

Robinson, P. (2001). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27-57.

Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on speech production, interaction, uptake of premodified input and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching (IRAL), 45,* 195-215.

Robinson, P. (2010). Situating and distributing cognition across task demands: The SSARC model of pedagogic task sequencing. In M. Putz & L. Sicola (Eds.), *Cognitive processing*

in second language acquisition: Inside the learner's mind (pp. 243-268). Amsterdam/Philadelphia, PA: John Benjamins.

Robinson, P. (2011). *Second language task complexity: researching the cognition hypothesis of language learning and performance*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *IRAL-International Review of Applied Linguistics in Language Teaching, 45*(3), 161-176.

Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: an instructed SLA approach to L2 phonology. *Language Awareness,* 20 (1), 45-59.

Saito, K. (2013). Reexamining effects of form-focused instruction on L2 pronunciation development: the role of explicit information. *Studies in Second Language Acquisition,* 35, 1-29.

Saito, K. (2015). Communicative focus on second language phonetic form: Teaching Japanese learners to perceive and produce English /ɹ/ without explicit instruction. *Applied Psycholinguistics*, 36, 377-409.

Saito, K., & Akiyama, Y. (2017). Video-based interaction, negotiation for comprehensibility, and second language speech learning: A longitudinal study. *Language Learning*, 67, 43–74.

Saito, K. & Wu, X. (2014). Communicative focus on form and second language suprasegmental learning. *Studies in Second Language Acquisition,* 36, 647-680.

Salaberry, M. R. & López-Ortega, N. (1998). Accurate L2 production across language tasks: Focus on form, focus on meaning, and communicative control. *The Modern Language Journal*, 82.4, 514-532.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning.* New York: Palgrave Macmillan.

Schmidt, R. (1995). Consciousness and foreign language learning: a tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and Awareness* (pp. 1- 63). Honolulu, HI: University of Hawai`i, National Foreign Language Resource Centre.

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

Segalowitz, N. (1997). Individual differences in second language acquisition. In A. de Groot & J. Kroll (Eds.), *Tutorials in Bilingualism* (pp. 85-112). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sicola, L. (2009). "*No, they won't just sound like each other": Attention to phonological form in NNS-NNS inter action on targeted L2 pronunciation tasks*. Landau, Germany: Peter Lang Publishers.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.

Sturm, J. L. (2013). Explicit phonetics instruction in L2 French: A global analysis of improvement. *System*, 41, 654-662.

Syrdal, A. K. and Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, 79, 1086-100.

Szalkowska-Kim, E. (2014). Factors influencing the pronunciation of adult speakers of a foreign language. *Lingua Posnaniensis*, 56.1, 141-158.

Solon, M., Long, A.Y. & Gurzynski-Weiss, L. (2017). Task complexity, language-related episodes, and production of L2 Spanish vowels. *Studies in Second Language* Acquisition, 39, 347-380.

Taylor, D. S. (1991). Who speaks English to whom? The question of teaching English pronunciation for global communication. *System*, 19.4, 425-435.

Trofimovich, P. & Gatbonton, E. (2006). Repetition and focus on form in processing L2 Spanish words: implications for pronunciation instruction. *The Modern Language Journal*, 90 (4), 519-535.

Van den Branden, K. (2006). *Task-based language teaching: from theory to practice.* Cambridge: Cambridge University Press.

UNIVERSITAT DE BARCELONA

Appendix A – Consent form (Experimental group)

Task complexity effects on the acquisition of an L2 vowel contrast:
A task-based pronunciation teaching study

You are welcomed to participate in this MA thesis project about how EFL learners acquire their second language. The aim of this study is to analyse the effect of cognitive complexity and individual differences on the development of pronunciation in a second language.

Procedures:

Data collection will be in a quiet small room and the total amount of time will be around 240 minutes, distributed in 7 sessions on different days in April and May. If you decide to participate in the study, you will do the following tasks:

1. Perception and production tasks on the computer.
2. Four tasks with a classmate where you will have to take decisions about particular situations around a trip to Kenya.
3. Proficiency test: you will need to tell if you know certain English words in the computer.
4. Auditory attention control tasks

Confidentiality and voluntary nature of the study:

All the data collected in this study will be anonymous and private, and your identity will be held in confidence in reports in which the study may be published or databases where it may be stored. Only I will have access to your audio recordings but your data will be de-identified so your identity is not associated with recordings or test scores. Moreover, you may choose not to take part or leave the study at any time and this will not affect your school grades or future relations with the researchers.

Name and Surnames: _____

Yes, I agree to participate:

Signature

Date: ___ /___/ _____

UNIVERSITAT DE BARCELONA

Appendix A – Consent form (Control group)

Task complexity effects on the acquisition of an L2 vowel contrast: A task-based pronunciation teaching study

You are welcomed to participate in this MA thesis project about how EFL learners acquire their second language. The aim of this study is to analyse the effect of cognitive complexity and individual differences on the development of pronunciation in a second language.

Procedures:

Data collection will be in a quiet small room and the total amount of time will be around 60 minutes, distributed in 2 sessions on different days in April and May. If you decide to participate in the study, you will do the following tasks:

1. Perception and production tasks on the computer.
2. Proficiency test: you will need to tell if you know certain English words in the computer.
3. Auditory attention control tasks

Confidentiality and voluntary nature of the study:

All the data collected in this study will be anonymous and private, and your identity will be held in confidence in reports in which the study may be published or databases where it may be stored. Only I will have access to your audio recordings but your data will be de-identified so your identity is not associated with recordings or test scores. Moreover, you may choose not to take part or leave the study at any time and this will not affect your school grades or future relations with the teachers.

Name and Surnames: _____

Yes, I agree to participate:

Signature

Date: ___ /___/ _____

Appendix B – Language background questionnaire

Please give the following information about yourself. This questionnaire will not be shared for privacy issues.

1. Age: _____

2. Sex: Male ☐ Female ☐

3. City and country of birth: _____

4. City and country of residence: _____

5. Mother tongue / Languages from birth. Please write them in order of dominance

    1._____
    2._____
    3._____
    4._____
    5._____

6. Other languages. Please write them in order of proficiency

    1._____
    2._____
    3._____
    4._____
    5._____

7. Please specify the context (natural or instructed), details (school, language academy…) and the number of years and hours learning English

☐ Natural context, details and years: _____

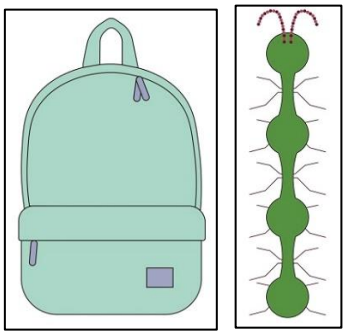
☐ Instructed context, details and years: _____

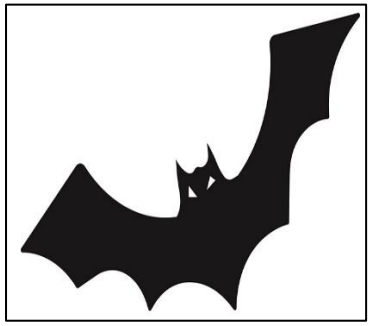
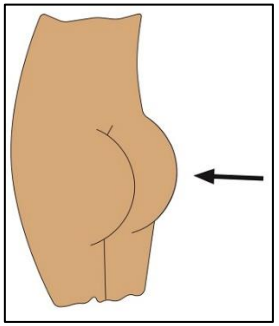
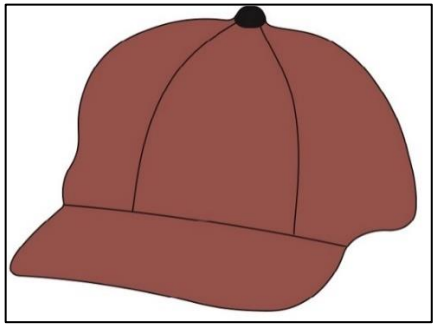
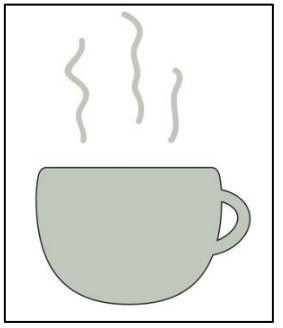
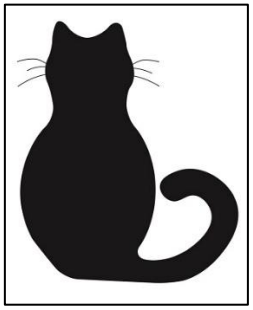
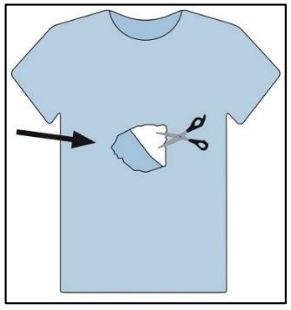
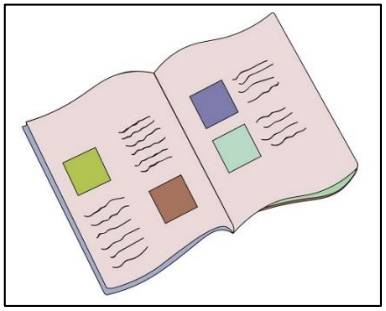
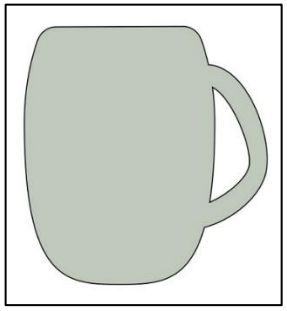
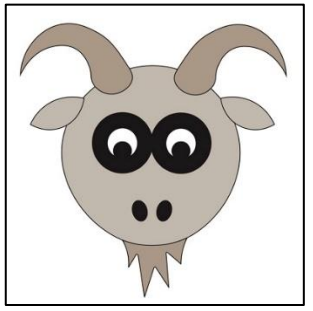
APPENDIX C – Testing and training stimuli

STIMULI: WORDS		
PRE-TEST (Identification/Discrimination)	***PRE-TASK (Words)****	***POST-TEST (Identification/Discrimination)***
Bag/bug	Bag/bug	Bag/bug
Bat/butt	Bat/butt	Bat/butt
Cap/cup	Cap/cup	Cap/cup
Cat/cut	Cat/cut	Cat/cut
Mag/mug	Mag/mug	Mag/mug
Ram/rum	Ram/rum	Ram/rum
Natty-nutty	Natty-nutty	Natty-nutty
Amber-umber	Amber-umber	Amber-umber
Babble-bubble	Babble-bubble	Babble-bubble
Stab-stub	Stab-stub (out)	Stab-stub (out)
	Distractors:	*Non-words:*
	Bean/bin	Gak/guk
	Sheep/ship	Dat/dut
	Teen/tin	Kad/kud
	Feast/fist	Mal/mul
	Weep/whip	Ras/rus
10 pairs - 20 tokens - 40 repetitions	**15 pairs -30 tokens- 30 repetitions**	**15 pairs - 30 tokens - 40 repetitions**

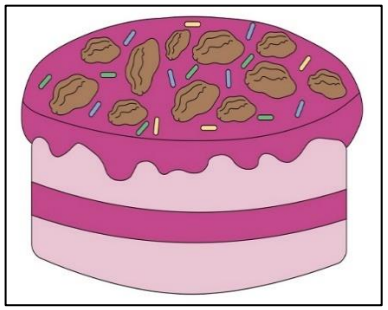
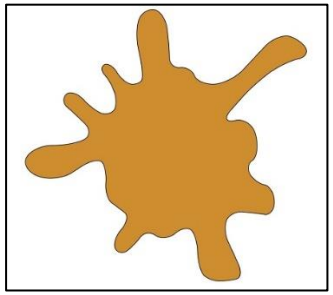
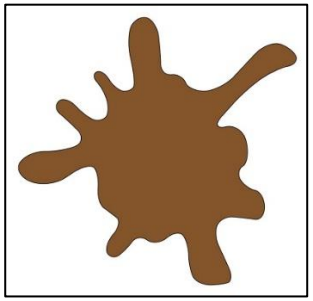
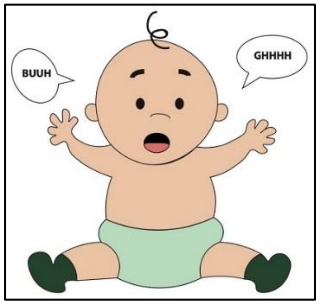
STIMULI: SENTENCES		
PRE-TEST (Delayed sentence repetition task)	***PRE-TASK (Sentences) *****	***POST-TEST (Delayed sentence repetition task)***
1-My BAG has a star. 2-There is a BUG on the table. 3-The BAT can't see anything. 4-The monkey's BUTT is pink. 5-Your CAP is on my head. 6-A CUP of tea, please. 7-This CAT loves dogs. 8-The t-shirt's CUT is big. 9-The weekly MAG is interesting. 10-In a MUG, he drinks coffee. 11-The RAM is in the farm. 12-RUM **is** what they usually drink. 13-The NATTY man went into the bar. 14-He loves the NUTTY flavour. 15-She wears AMBER colour clothes. 16-The UMBER colour chair is broken. 17-Babies BABBLE before they speak. 18-The clown makes a BUBBLE. 19-Killers STAB their victims. 20-I STUB out her smelly cigarette. *Distractors:* Green beans grow from plants These bins are full of rubbish This ship is alone in the sea The sheep are eating flowers	1-These BAGS are bigger than the child. 2-People in Kenia eat BUGS. 3-There was a black BAT in the cave. 4-The chimpanzee is showing its BUTT. 5-My CAP is on the koala. 6-A CUP of tea is what British drink. 7-Dark CATS live next to trees. 8-Her trousers had some CUTS. 9-The best gossiping is in this MAG. 10-This MUG goes in my suitcase. 11-The RAM bit me in the field. 12-RUM is popular in parties. 13-The NATTY woman seduced the man. 14-This cake has a NUTTY flavour. 15-The AMBER colour snake stared at me. 16-The UMBER colour monkey is in love. 17-Babies BABBLE all the time. 18-A huge BUBBLE was flying over the circus. 19-You will have to STAB Tim. 20-I should STUB out her cigarette. 21-Green BEANS grow from plants. 22-These BINS are full of rubbish. 23-The SHEEP are eating flowers. 24-The SHIP is alone in the sea. 25-These TEENS are kissing each other. 26-TINS are full of Kenyan bugs. 27-The FEAST was in the garden. 28-Joe showed his FIST in the picture. 29-The baby WEEPS every night. 30-The woman WHIPS the horse.	1-My BAG has a star. 2-There is a BUG on the table. 3-The BAT can't see anything. 4-The monkey's BUTT is pink. 5-Your CAP is on my head. 6-A CUP of tea, please. 7-This CAT loves dogs. 8-The t-shirt's CUT is big. 9-The weekly MAG is interesting. 10-In a MUG, he drinks coffee. 11-The RAM is in the farm. 12-RUM is what they usually drink. 13-The NATTY man went into the bar. 14-He loves the NUTTY flavour. 15-She wears AMBER colour clothes. 16-The UMBER colour chair is broken. 17-Babies BABBLE before they speak. 18-The clown makes a BUBBLE. 19-Killers STAB their victims. 20-I STUB out her smelly cigarette. 21-This GAK makes nasty bites. 22-The GUK changes its colour. 23-He eats DAT with milk. 24-She drinks DUT everyday. 25-The KAD eats the mouse. 26-The KUD can swim in cold water. 27-That MAL shoots with precision. 28-Her MUL is from this beauty shop. 29-This RAS takes good pictures. 30-They bought a smart RAS. *Distractors:* Green beans grow from plants These bins are full of rubbish This ship is alone in the sea The sheep are eating flowers
10 pairs - 20 tokens - 40 repetitions	**15 pairs - 30 tokens - 30 repetitions**	**15 pairs - 30 tokens - 40 repetitions**

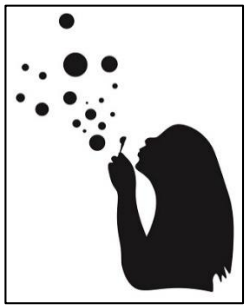
Handmade pictures for words in the pre-task

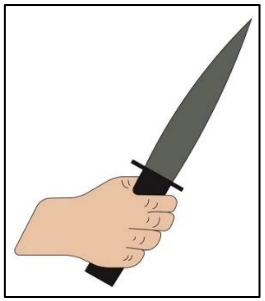
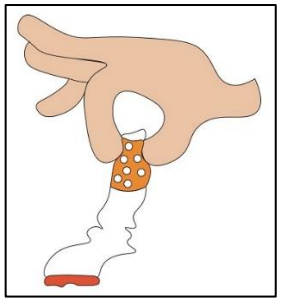
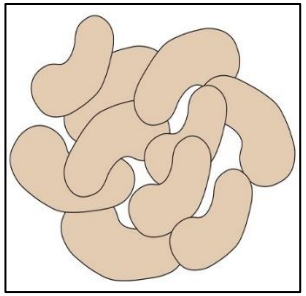
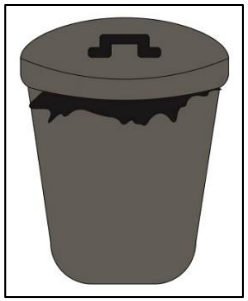
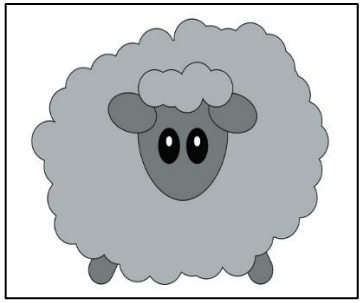
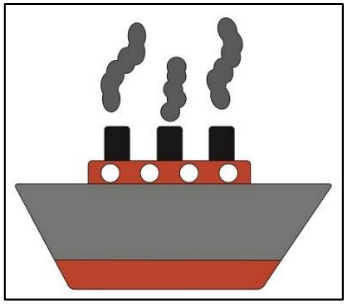


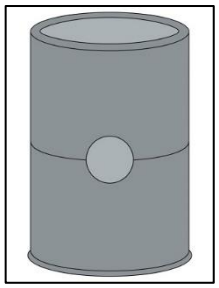
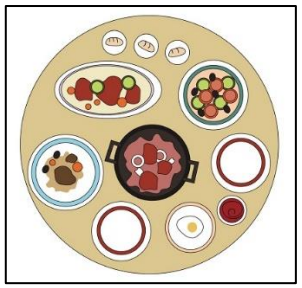
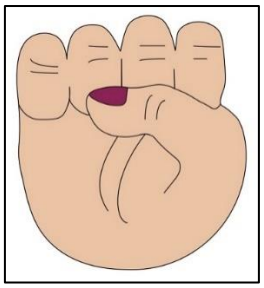










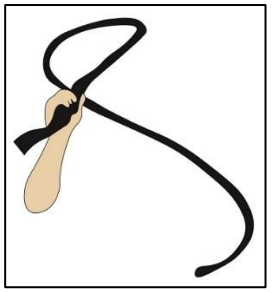


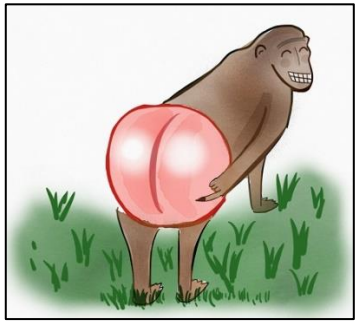


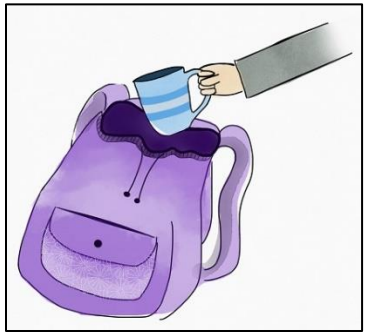







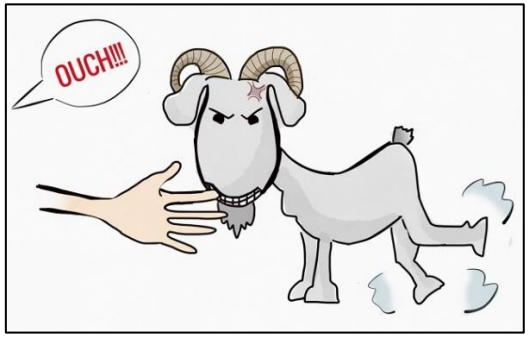
I should study for the exam...
Where's my phone?
I have to post this on Instagram




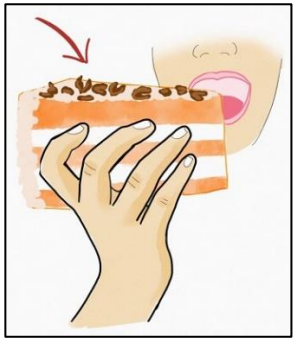


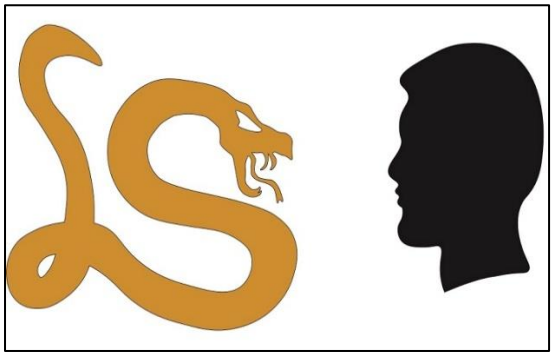
***Handmade pictures for sentences in the pre-task

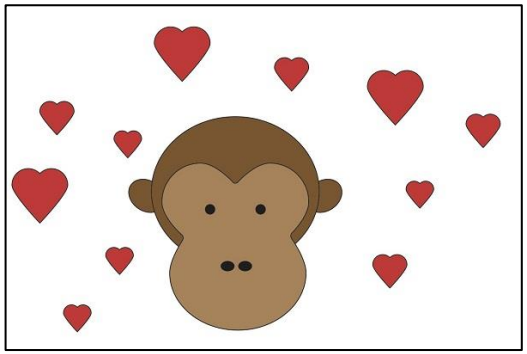






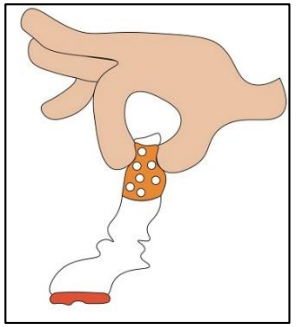


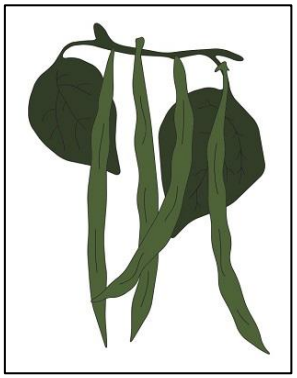


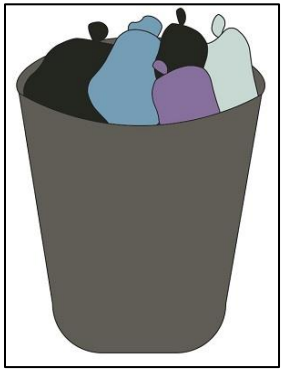


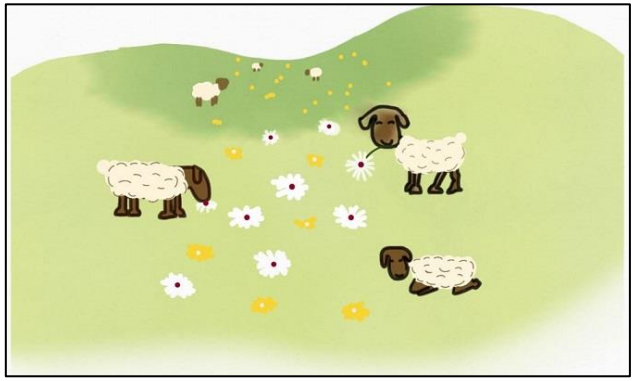

Gugu
Ohh
Ammm

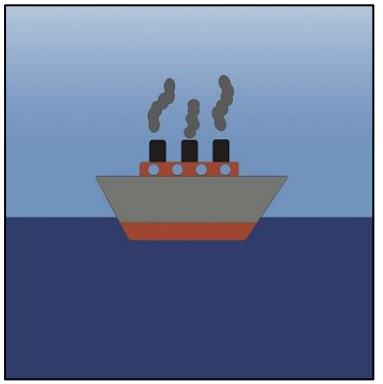

TIM

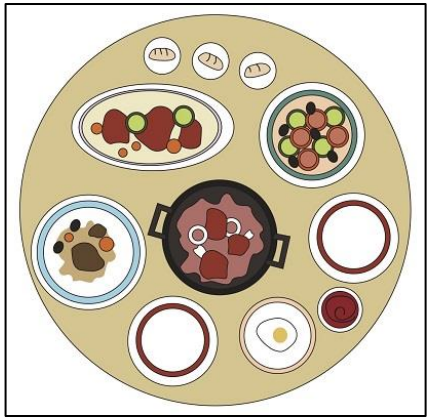


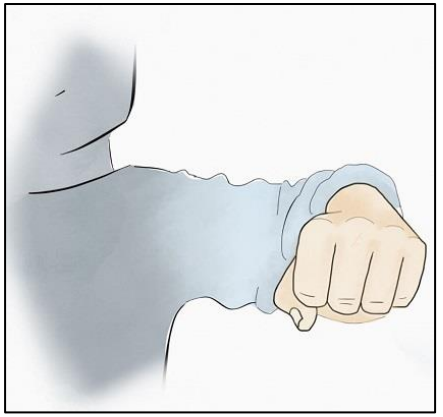




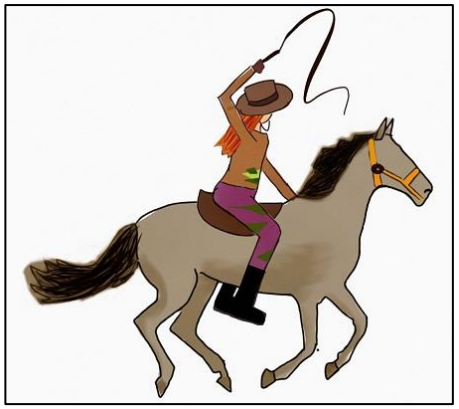









10 11 12 1
2
3
7 6 5 4



Appendix D – Mini pre-tasks and tasks

Mini pre-tasks script

STIMULI: TEXTS	
Mini pre-tasks	
Session 1 *(Listening comprehension)*	A: Good morning Sarah! Today we need to organize activities in the zoo for a big group of Spanish tourists. B: Yes, sure! What kind of animals would they like to see? A: Bats, rams and cats. B: I think they are a good choice. Which objects would they like to buy in our shop? A: Tourist love cups, not caps, and zoo mags. What do you think? B: They may also buy bags. A: Why would they buy bugs? B: Not bugs, bags to put things in! What colour? A: Amber? B: The umber colour? A: No, amber colour, like yellow and orange. B: Alright, I think we have an excellent plan for their visit. A: Yes, they will love the zoo and our activities. B: Sounds good!
Session 2 *(Listening comprehension)*	Barbara and Peter were two farmers who were organizing a trip to Kenya on the first week of August. Their intention was to bring 4 bags as they were going to stay there for the whole month. Barbara wanted to bring her English cup because it was a special present but John told her that there was only space for a cap and they could buy a cup in Kenya as a souvenir. John put his trousers with cuts in the bag but Sarah saw them, took them out and change them for the trousers with cats, which were very funny. They agreed that they would bring some mags to read in the plane because it was a 14-hour-flight. Since the temperatures were cold at night, Barbara decided to take the amber jacket whereas Peter took the umber coat. When they got to the airport they bought a nutty cake and put it in their bags. When they were in the security checkpoint, Peter realised that rum was prohibited in the bags. Barbara told him off but Peter never realised that Barbara's cup was in their bags and that she had brought a picture with her favourite ram from the farm!
Session 3 *(Listening comprehension)*	Margaret and Patrick wanted to organize a roleplay party for their 25th anniversary. First of all, they decided on the roles that people would perform. Margaret preferred a very classic party around a theme such as film characters. However, Patrick liked to combine different themes and create many different profiles. At the end, they agreed that they would have a variety of roles and the main goal was to find the murderer who would silently kill all the characters with a knife. Margaret fancied having a doctor, a zoo keeper, a baby and a journalist. Peter added a chef and a clown. Once the roles were clear, they decided that the zoo keeper would be the murderer and he needed to stab his victims with a dagger. He wore trousers with cuts and he was always in company with his two black cats. Also, he brought two pictures (one of his bat and one of his ram) and hid some rum in his pocket. The clown made bubbles and the child babbled all the time and showed its butt. The crazy doctor had a knife and some bugs. The journalist brought some mags and

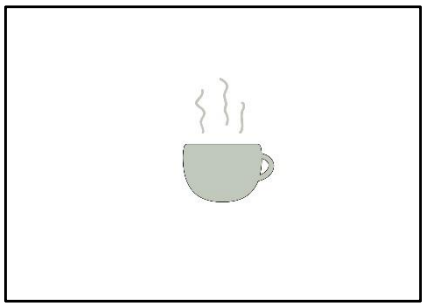
	carried two bags that were secretly kept in the bins. He wore a natty umber suit. Finally, the chef brought some tins with beans and tried to stub out the cigarette of the zoo keeper, who was very angry and whipped him. The feast lasted one hour and a half.
Session 4 *(Listening comprehension)*	A: Hi Ann! I have been told that we need to select some pictures for the school website. B: Do we? Out of these four pictures, which one would you choose? A: We need to follow certain school requirements and classmates opinions. B: OK. David told me that he did not want to appear in the website. A: The school warned me that alcohol or cigarettes mustn't appear either. B: Sandra thought that animals should be part of the pictures. A: Alright, what about this picture of Sandra with the ram? B: Didn't you say that alcohol was forbidden? A: Ram, the male sheep, not rum! B: I see. I think we can include this picture. What about this one? A: A baby making bubbles. Isn't it cute? B: Yes, not like the one where the baby babbles and weeps. A: David is appearing in that one, though. B: What about the group picture wearing the zoo caps? A: You mean the one with the zoo cups? B: No, the picture where caps are on their heads. A: Great! I love this one too. B: Perfect, let's talk to the headmaster and present him our choices. A: Thanks Ann! B: You're welcome!

Mini pre-tasks

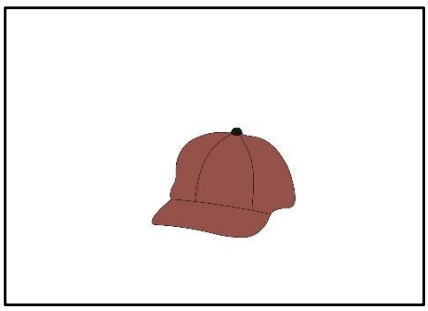
PRE-TASK 1: Session 1

You are going to hear two people organizing zoo activities for a group of Spanish tourists. Please answer the following questions with √ :

1. **What do tourists love?**

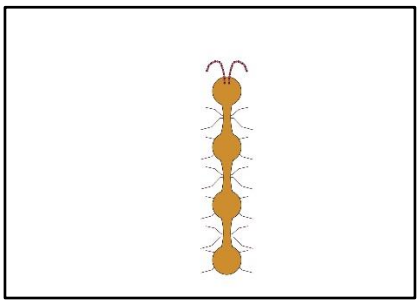


☐

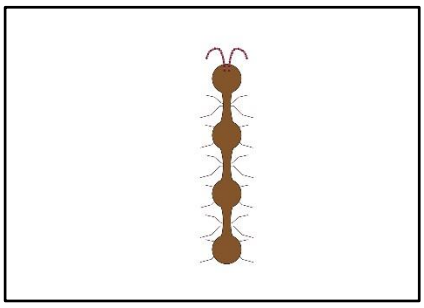


☐

2. **What do tourists want to buy?**



☐



☐

☐

☐

PRE-TASK 2: Session 2

You will hear a short story about a couple of farmers who are organizing a trip to Kenya and they are deciding on the items they want to bring. Please listen to the story and answer if the sentences are TRUE **(T)** or FALSE **(F).**

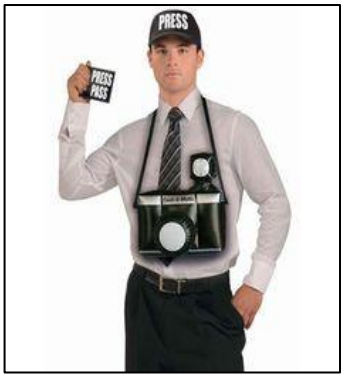
1. Barbara and Peter brought 4 bags. _____
2. Barbara didn't take the English cup. _____
3. They put a cap in the bag. _____
4. Mugs went into the plane. _____
5. John took the amber coat. _____
6. Barbara and Peter bought a nutty cake. _____

PRE-TASK 3: Session 3

You will hear a story about a married couple who wants to organise a roleplay party. They have taken decisions about their roles and how they will be dressed or what they will do in the party. Please connect the OBJECTS/ACTIONS with the different characters.

Be careful, some actions/objects do not correspond to any character!

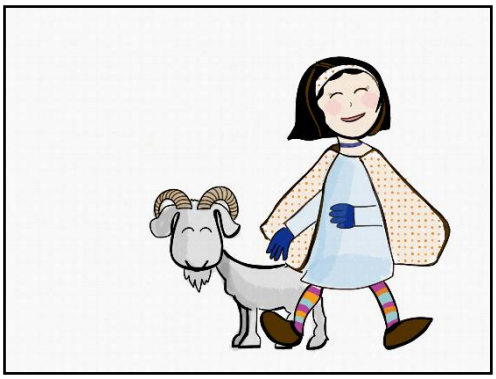
Trousers with cats	Rum in the pocket
Trousers with cuts	Stabs someone
Tins with beans	Whips someone
A picture of my ram	Stubs out the cigarette
Make bubbles	Weeps
Carries mugs	A picture of my bat
Show its butt	Carries mags
Has a knife	Babbles
Carries bags	Carries bugs



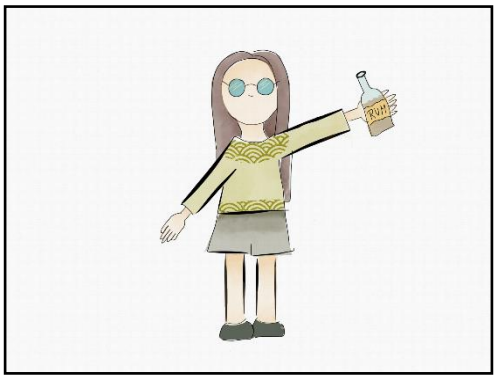
PRE-TASK 4: Session 4

You will hear two classmates deciding on what pictures will appear in the school website.

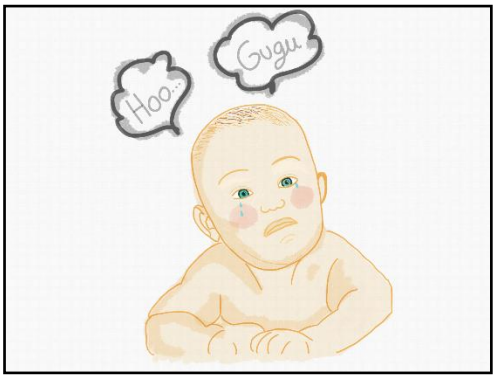
Please write [√] in the SELECTED pictures.



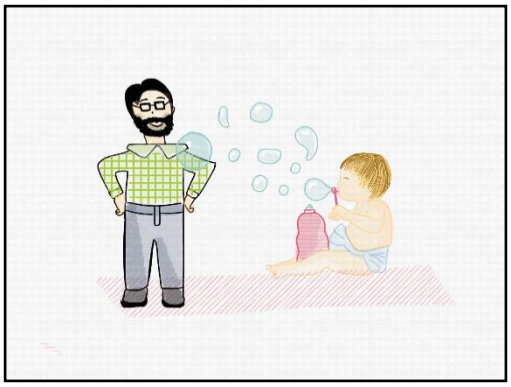
☐



☐



☐



☐

☐

☐

THE DESIGN OF THE TASKS

Task 1: No conditions, 10 words and 2 monosyllabic pairs

Task 2: 2 conditions, 12 words and 3 monosyllabic pairs + 1 disyllabic pair

Task 3: 4 conditions, 18-20 words and 5 monosyllabic pairs + 1 disyllabic pair

Task 4: 6 conditions, 20-22 words and 7 monosyllabic pairs + 3 disyllabic pairs

TASK 1: Session 1

In this task, you have to select what you are going to see and buy in the natural park once you get to Kenya. Your teachers and parents have already made some choices for you. Please agree on the MOST ESSENTIAL 6 items/animals.

1) Look at your list
2) Share your information with your partner
3) Decide on what to see and buy together
4) Write it down in your list
5) Compare lists and check that they are the same

SEE	BUY
1.	1.
2.	2.
3.	3.

LIST OF ITEMS/ ANIMALS

STUDENT A		STUDENT B
Bugs		Cats
Bats		Bags
Caps		Cups
Cats		A sick monkey
Chair		Bats

TASK 2: Session 2

In this task, you need to decide on what things you are going to bring to Kenya. The items that you can see below are your MAIN PREFERENCE.

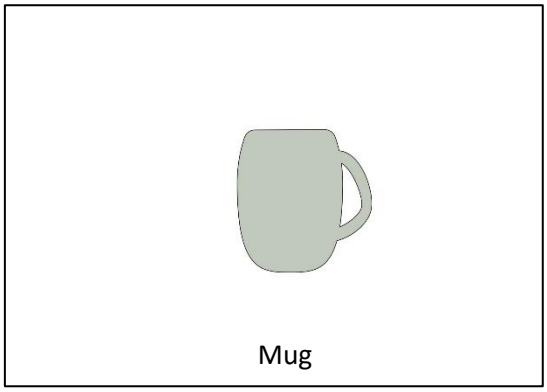
1) Look at your pictures
2) Present your choices to your classmate without showing the pictures
3) Agree on the 6 most essential items
4) Put the selected pictures on the boxes
5) Check the final decision

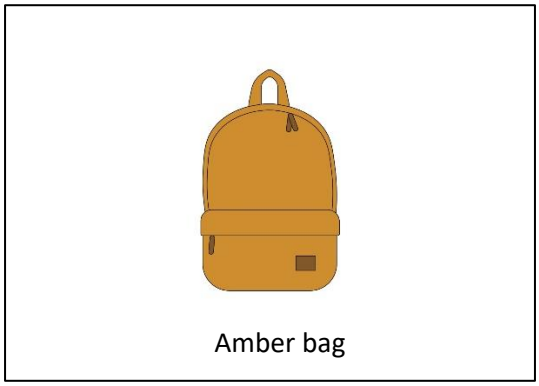
ITEM 1	ITEM 2	ITEM 3
ITEM 4	ITEM 5	ITEM 6

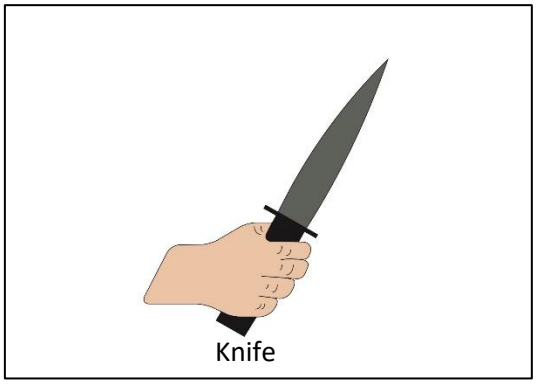
STUDENT A

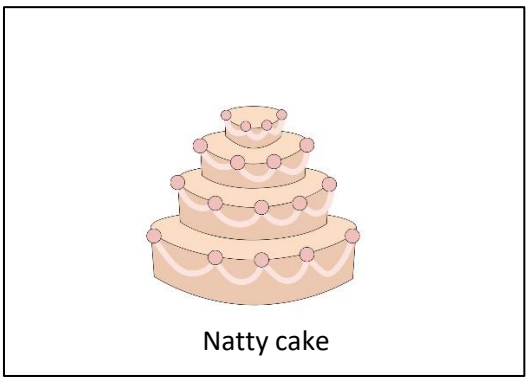
Conditions: You never go anywhere without your coffee

You are allergic to nutty cakes





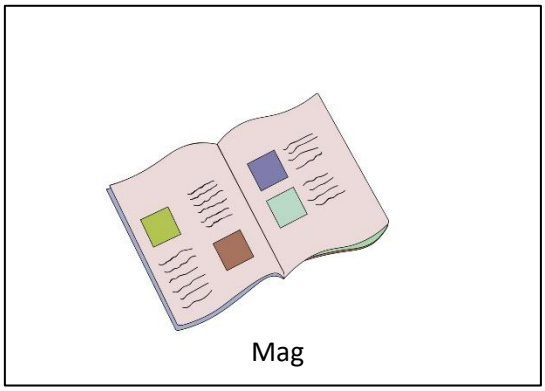


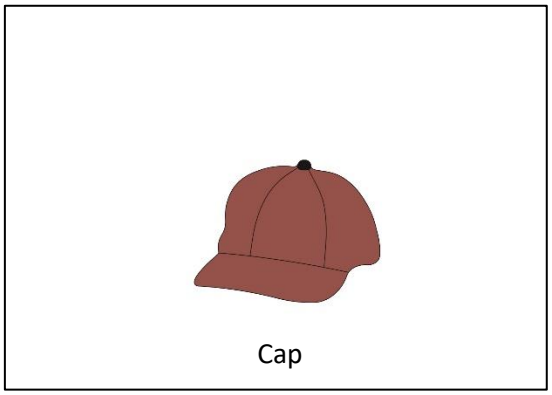


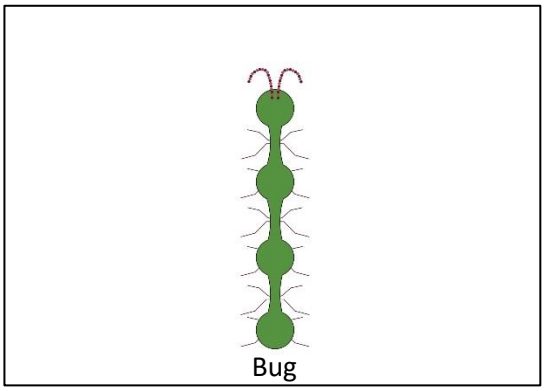
STUDENT B

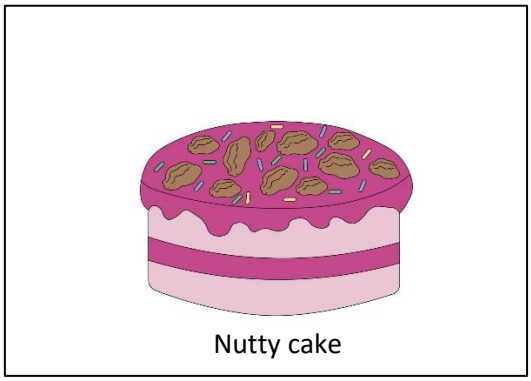
Conditions: You need protection from the sun and you love reading

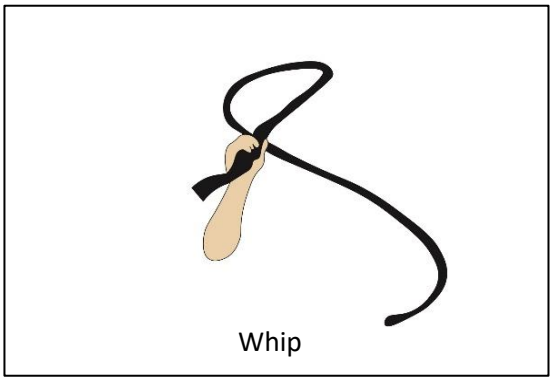
You hate the amber colour











TASK 3: Session 3

You have invited Jack, Ann, Tom and Lucy to a 'roleplay' party. You have already decided on the role that they will play.

1) Read the information about these 4 friends individually
2) Consider the following conditions of the party:

   Alcohol and cigarettes are prohibited

   Animals can come to the party

   People can't wear/bring anything umber

   All of them need to behave with respect

3) Agree with your classmate on one murderer
4) Share information about the characters
5) Discuss what they will wear/ bring/ do and make an agreement

STUDENT A	
JACK: *VET*	**ANN: *POLICEWOMAN***
Has an injection Brings a ram Wears green trousers --- Shows pictures of monkey's butts	Has a gun Brings tins with bugs Wears a blue t-shirt --- Shows her bat
TOM: *CHEF*	**LUCY: *JUDGE***
Has a knife Wears trousers with cats Whips Lucy --- Wants to stab Lucy after her cigarette	Has pills Wears a t-shirt with an amber sheep Holds a glass of water --- Smokes out of the party

You have invited Jack, Ann, Tom and Lucy to a 'roleplay' party. You have already decided on the role that they will play.

1) Read the information about these 4 friends individually
2) Consider the following conditions of the party:

   Alcohol and cigarettes are welcomed

   Anything about animals is NOT allowed

   People can't wear/bring anything green

   All of them need to behave with respect

3) Agree with your classmate on one murderer
4) Share information about the characters
5) Discuss what they will wear/ bring/ do and reach an agreement

STUDENT B	
JACK: *VET*	**ANN: *POLICEWOMAN***
Has an injection Brings rum to drink Wears umber trousers --- Shows pictures of monkey's butts	Has a gun Brings tins and bags Wears an umber t-shirt --- Shows her butt
TOM: *CHEF*	**LUCY: *JUDGE***
Has a knife Wears trousers full of cuts Weeps with Lucy --- Wants to stub out Lucy's cigarette	Has pills Wears a t-shirt with an umber ship Holds a glass of water --- Smokes in the party

PERSON	TO BRING/WEAR/DO
Jack 	
Ann 	
Tom 	
Lucy 	

TASK 4: Session 4

STUDENT A

You are back to Barcelona after your trip to Kenya. Your teachers have told you to decide on 10 pictures that will go to the webpage of the school but you have already chosen some. Remember that you need to follow certain school requirements and classmates' opinions.

1) Look at all your pictures. You can't show them at any point but you can talk about them
2) Consider the following school requirements

   Alcohol is prohibited

   Amber colour items must not appear

   Signs of violence cannot appear

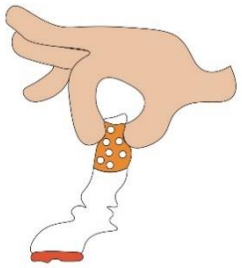
3) Consider the following classmates' opinions

   Sandra doesn't want to appear in 1 picture

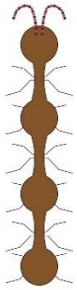
   James prefers individual pictures

   Tom believes animals should appear in the pictures

4) Present your choices and reach an agreement with your classmate
5) Put the selected pictures on the table



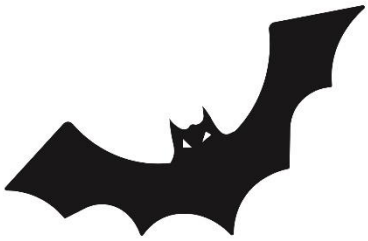
Stub out a cigarette



Umber bug

Mag

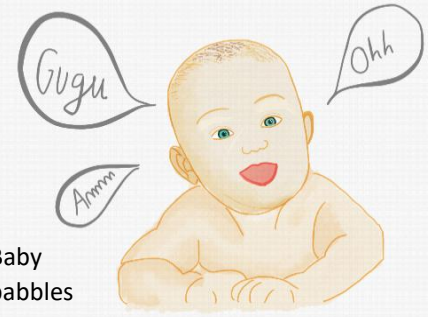
Nutty cake



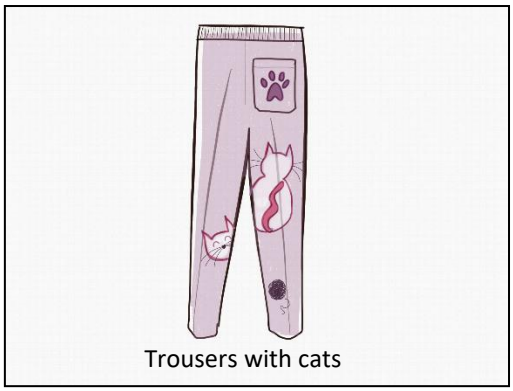
Bat

Sandra with
the ram

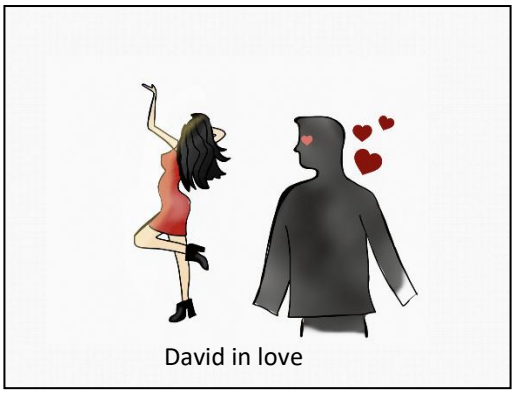
Group with zoo cups



Baby
babbles



Trousers with cats



David in love

STUDENT B

You are back to Barcelona after your trip to Kenya. Your teachers have told you to decide on 10 pictures that will go to the webpage of the school but you have already chosen some. Remember that you need to follow certain school requirements and classmates' opinions.

1) Look at all your pictures. You can't show them at any point but you can talk about them
2) Consider the following school requirements

   Cigarettes are not permitted

   Umber colour items must not appear

   Sexual connotations cannot appear

3) Consider the following classmates' opinions

   David doesn't like his face in some pictures

   Sarah prefers group pictures

   Kim believes animals shouldn't appear in the pictures

4) Present your choices and reach an agreement with your classmate
5) Put the selected pictures on the table

Amber bag

A cup in the bag

Group with zoo caps

Monkey and its butt

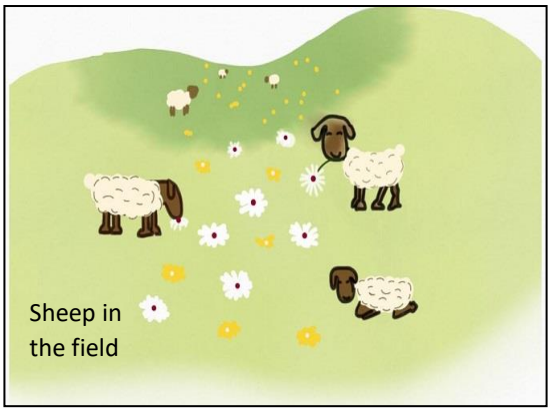
Baby making bubbles

Sandra with
the rum

Trousers
with cuts

Natty cake


TIM
Stab Tim


Sheep in
the field

Appendix E –Teachers and students' evaluations of task complexity

E.1. Teachers' survey about task complexity (4 tasks)

TASKS AND PRONUNCIATION

The goal of this survey is to help us rate the difficulty and mental effort involved in a number of decision-making tasks whose aim is the improvement of English pronunciation.
You will be asked to consider 4 tasks in total. In pairs (Student A & Student B), intermediate/upper-intermediate level students are asked to read certain information related to their trip to Kenya (e.g. things that they would buy and see) and make decisions in order to have a well-organized and enjoyable trip.
Read each of the four tasks and their instructions and rate them. For difficulty, 1 means extremely easy and 9 extremely difficult. For mental effort, 1 means little or no effort and 9 means high mental effort. We would recommend for you to read the four tasks -which are in the email we sent you-before you rate them so that you can compare them.

*Obligatorio

1. Rate the difficulty of TASK 1 *

   Marca solo un óvalo.

   - 1. Extremely easy
   - 2
   - 3
   - 4
   - 5
   - 6
   - 7
   - 8
   - 9. Extremely difficult

2. Rate the mental effort needed to complete TASK 1 *

   Marca solo un óvalo.

   - 1.No mental effort required
   - 2
   - 3
   - 4
   - 5
   - 6
   - 7
   - 8
   - 9. High mental effort required

3. Rate the difficulty of TASK 2 *

Marca solo un óvalo.

- 1. Extremely easy
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9. Extremely difficult

4. Rate the mental effort needed to complete TASK 2 *

Marca solo un óvalo.

- 1.No mental effort required
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9. High mental effort required

5. Rate the difficulty of TASK 3 *

Marca solo un óvalo.

- 1. Extremely easy
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9. Extremely difficult

6. Rate the mental effort needed to complete TASK 3 *

Marca solo un óvalo.

- 1.No mental effort required
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9. High mental effort required

7. Rate the difficulty of TASK 4 *

Marca solo un óvalo.

- 1. Extremely easy
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9. Extremely difficult

8. Rate the mental effort needed to complete TASK 4 *

Marca solo un óvalo.

- 1.No mental effort required
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9. High mental effort required

Remember to rate each task by comparing them to the others!

DIFFERENCES BETWEEN THE FOUR TASKS

Please take a few more minutes to let us know what aspects of the task contributed to your impressions of difficulty and mental effort.

9. In your opinion, what made the tasks more or less difficult? and more or less mentally effortful? *

Thank you!

Thank you for participation in this project. Your help is really valuable to us.

E.2. Students' survey about task complexity (4 tasks)

NAME & SURNAME: _____

TASK NUMBER _____

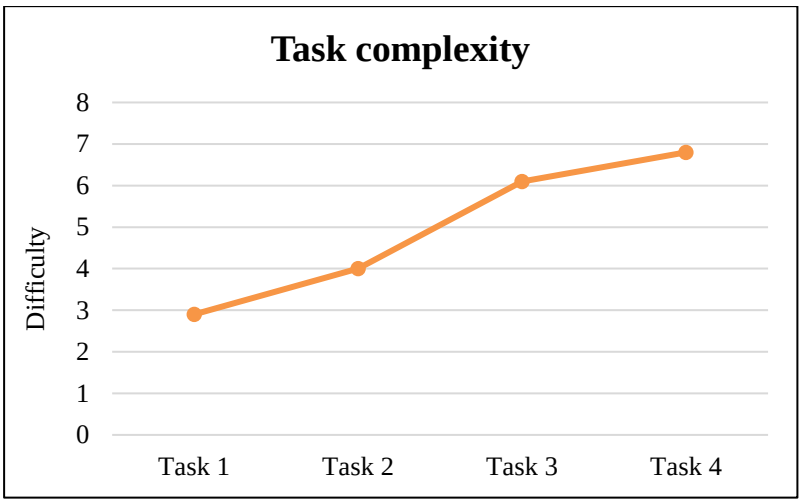
TASK EVALUATION

	Not at all								Very much
The task required mental effort	1	2	3	4	5	6	7	8	9
I want to do more tasks like this in the future	1	2	3	4	5	6	7	8	9
I did well on the task	1	2	3	4	5	6	7	8	9
The task was attractive	1	2	3	4	5	6	7	8	9

How much time do you think you spent on the task? _____

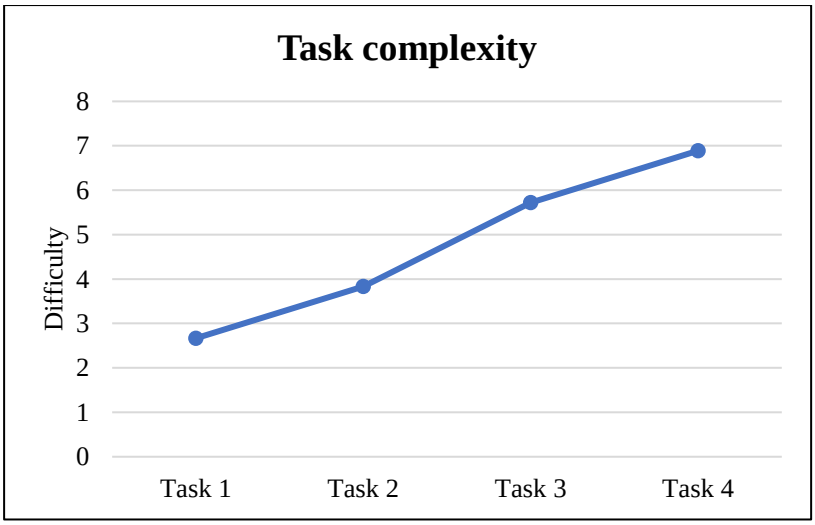
E.3. Teachers and learners' answers about task complexity (4 tasks)

Considering language teachers' answers in the 9-point-scale, task 1 [+S] was significantly less difficult than task 2 [-S], *t* (9) = -3.161, *p*=.012, *r*=.725; task 2 [-S] was significantly less difficult than task 3 [-C], *t* (9) = -3.194, *p*= .011, *r*=.728; nevertheless, there was no statistically significant difference between task 3 [-C] and task 4 [+C], *t* (9) = -1,210, *p*= .257, *r*=.374.



Task complexity ratings by language teachers.

According to the learners' answers in the experiment rated from 1 to 9, task 1 [+S] is significantly less difficult than task 2 [-S], *t* (17) = -4.745, p<.001, *r*=.754; task 2 [-S] is significantly less difficult than task 3 [-C], *t* (17) = -6.776, *p*<.001, *r*=.854; and task 3 [-C] is less difficult in comparison to task 4 [+C], *t* (17) = 4.507, *p*<.001, *r*=.737. As observed in the graph below, there is no great difference between task 3 and task 4 -as found with language learners' ratings-; however, it seems that, according to students, they perceive a difference between the two in terms of difficulty.



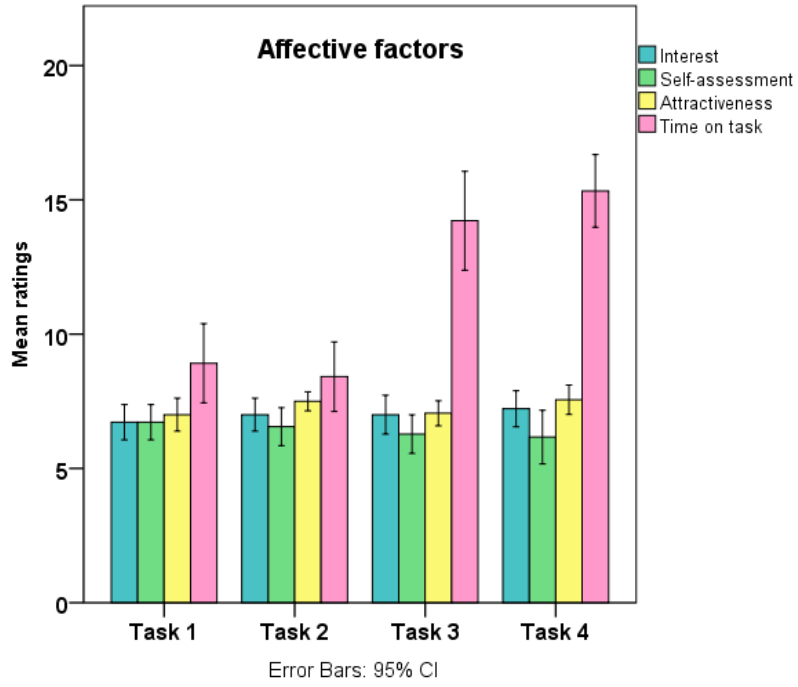
Task complexity ratings by learners.

Apart from task difficulty, the questionnaire asked about learners' interest in the task, self-assessment, attractiveness and perceived time on task in a 9-point-scale.

	Mental effort M/SD	Interest M/SD	Self-assessment M/SD	Attractiveness M/SD	Time on task M/SD
Task 1 (+S)	2.67/1.37	6.72/1.31	6.72/1.31	7.00/1.23	8.92/2.97
Task 2 (-S)	3.83/1.61	7.00/1.23	6.55/1.42	7.50/0.71	8.42/2.61
Task 3 (-C)	5.72/1.22	7.00/1.45	6.27/1.45	7.06/0.94	14.22/3.71
Task 4 (+C)	6.89/1.27	7.22/1.35	6.16/2.01	7.56/1.10	15.33/2.72

Responses to the task difficulty questionnaire by learners.

Given that the data was not normally distributed for any of the variables, a Friedman test was run to observe changes across tasks in terms of affective factors. Although learners' interest in the task increased with complexity; nevertheless, the difference between the tasks did not reach significance (X_2 (3) =7.462, p=.059). Concerning self-assessment, learners perceived that their performance was getting worse when the tasks were more cognitively complex; however, this difference was statistically non-significant (X_2 (3) =1.758, p=.624). When dealing with attractiveness, task 2 and task 4 were more visually attractive than task 1 and 3, which was visible in the learners' responses but, again, this did not reach significance (X_2 (3) =6.197, p=.102). Finally, learners spent significantly less time when the task was simple and more time when it was complex (X_2 (3) =39.705, p<.001).



Learners' responses about affective factors in the task difficulty questionnaire.

Appendix F – Analysis of LREs

Guidelines

PHONETIC LREs:

Phonetic LREs are defined as instances during conversation where students discuss, question or self-repair the pronunciation of words that they are producing.

LREs: sequence longer than 2 turns

Learner A: *ram*

Learner B: *ram with A?*

Learner A: *yes, ram* (LRE)

LRERC: an LRE with a single **recast**[1], 2 turns

Learner A: *cup*

Learner B: *cap?*

Learner A: yes, *cap, cap.* (LRERC)

LRESR: an LRE with a single **self-repair**[2], 1/2 turns

1. Error repair: Learner A: *cap, no cup (LRESR-err)*

2. Non-error repair: Learner B: *amber, amber color (LRESR-nonerr)*

LREREP: an LRE with a single **repetition**[3], 1 turn (except when it is a clear continuation from the previous sentence)

Learner A: *Can we bring a cap?*

Learner B: Yes, a *cap, cap. (LREREP)*

[1] Pronunciation recast: "A correct restatement of a learner's incorrectly formed utterance" (Nichols et al. 2001, p.721).
[2] Pronunciation self-repair: The learners' self-correction of faulty pronunciation.
[3] Pronunciation repetition: The learners' statement of the same word with the same pronunciation.

TASK 1

CODE	**LRE**	**LRERC**	**LRESR**	**LREREP**	**Examples**
*01_1A_*	5	1	2	2	LA: go hmm@p seen bats LB: bats? LA: bats (LREs) LA: I would buy a cap LB: A cap or a cup? LA: A cup (LRE) LB: I also think that we could buy hmm@p bags [/] bags (LREREP) LA: bags? okay hmm@p want hmm@p to [/] to buy a book LB: A book for reading? LA: <No no> [/] no bug or bugs nor nor bags [//] bugs (LRESR-err) (LRE) LB: I still want to buy bags [/] bags LA: bags? Bags okay to buy bags hmm@p okay (LREREP) LB: with a LA: with a bags okay (LRE)
*02_1B_*					LB: What happened to you? LA: You say I say cap nor cup [/] cup nor cap (LRESR-err) LB: Es@s:c cap!!! LA: cap! (LRERC) LB: I said cap LA: No I said cap LB: And I said cap LA: Cap with A! (LRE) LB: Yes! LA: Yes LB: And you with U! LA: Yes buy cup I don't understand you R: You want to say caps or cups? LB: Cups (LRE)

*11_1A_* *12_1B_*	3	2	1	1	LA: What did you say? Cats? LB: Yes cats (LRE) LA: I think that the first thing that I told you about the cups it's a really good idea LB: But why do you want caps [//] cups cups cups (LRESR-err) (LREREP) LA: I prefer cups rather than [/] than chairs LB: But [/] but I don't understand you if <you you can you> [//] you can't sit <in a> [/] in a cup LA: No you can't but you can't drink LA: Let's write cups and that's it LB: caps? R: Not caps! LA: cups LB: yes, <cups cups cups> [/] cups (LRERC) (LRE) LB: I thought about books /bʊks/ LA: What? LB: bugs /bʊks/ LA: books? /bʊks/ LB: yes LA: bugs? /bʌgs/ LB: bugs /bʌgs/ LA: Okey LB: bugs /bʌgs/ (LRERC) (LRE)
*13_1A_*	2		3	1	LA: What do you want to buy in the natural park of Kenya? LB: I am going to buy bags and cups LA: bags? LB: Sí@s:c bags LA: bags (LRE) LB: and cups caps cups cup why? (LRESR-err) LA: because I like it and you? LB: I think it's important to buy caps for their it's sunny in this country LA: Okay

*14_1B_*					LA: I would like to see a bugs (books) and <am um> [//] amber bugs (LRESR-err) LB: hmm@p I would like to buy ai@s:c to see a sick monkey bat and cats LA: Ah cats, yes R: What do you want to buy? LA: Bags for and what else caps [/] caps (LREREP) LA: cups caps (LRESR-err) LB: caps LA: caps (LRE)
*15_1A_OK_* *16_1B_*	2				LA: You are going to buy caps, right? LB: No caps for my head, cups for drink tea or coffee or +/. (LRE) LA: No I am going to buy caps for the head # LB: Maybe we can buy both LB: I would like to buy the bugs LA: Well, I am only going to see bugs, not going to buy it LB: No, not bugs like animals, bags to put in the suitcase and things like that (LRE) LB: No I am not going to buy it too I don't need it.
*17_1A_* *18_1B_*	1		1		LA: In my opinion I think I prefer buy a bags LB: Bags for the school? LA: I don't want to buy bags I prefer ### <I want> [//] I prefer to buy bugs (LRESR-err) LB: hmm@p okay (LRE)

*03_1A_* *04_1B_*	1				LA: I want to buy mmm cups LB: You want to want to buy cups but I want caps LA: No but is <no se que dir@s:c> I want cups not caps (LRE) LB: But for the sun it's better caps than cups because cups hmm@p are [/] are <hottest> [//] hotter LA: Okay [/] okay LB: Okay we go to see sick monkey but I I need a chair to sit LA: A chair? LB: Yes, chair I am tired /ˈtɪred/ LA: tired /ˈtaɪəd/ LB: tired /ˈtaɪəd/ <o com es pronuncii@s:c> [i]
*05_1A_* *06_1B_*	1	1			LA: But what else can we see? Like there are there are some endemic species that are only in Kenia like bugs I don't know we can go and see them LB: No I never see a bag or +/. LA: No bugs LB: Ah bugs LA: Maybe do you feel like seeing them? LB: hmm@p okay. (LRERC) (LRE)
*07_1A_*	2			1	LA: I want to buy a bag and a cup LB: A what? LA: bag LB: vale@s:s LA: yes and LB: no no [/] no LA: bag no? LB: no no LA: I want a bag of Kenia (LRE) with the <como se dice bandera@s:s>? R: with the flag L1: Eso@s:s with the flag of Kenya [ii]

*08_1B_*					LB: I want to buy some caps LA: Me too LB: caps LA: no caps LB: no caps like LA: ah I want to buy cups [/] cups are different (LRErep) LB: We can buy the two things LA: Okay I am agree with you LB: <Entonces ahora la pongo@s:s> LA: caps and cups (LRE)
*09_1A_* *10_1B_*	1		1		LA: We can buy a bag LB: No we can buy a cats a no caps cups (LRESR-err) LA: Yes I agree R: What do you want to buy? LA: A caps LB: okay I'm agree LA: noo! Caps no a cups with U (LRE) R: cups LB: I'm not agree

TASK 2

CODE	**LRE**	**LRERC**	**LRESR**	**LREREP**	**Examples**
*01_1A_*	7	1	1	2	LB: I would like to bring a mag LA: A mag? LB: mag LA: okay (LRE) LB: Because I love reading and it would help me to go pass the travel faster LA: okay LB: Do you accept it? LA: Yes [/] yes <I accept it> [/] I accept it LB: A mag? LB: A mag yes I accept [/] I accept that (LRE) LA: I would like to bring a mug because hmm@p I love coffee and if I don't have the coffee I'm like uhhh you know? LB: okay I agree LA: okay mug (LREREP) LB: hmm@p I would like to take a natty LA: A nu nutty LB: no nutty cake because I love it and there in Kenya there are there aren't this (LRERC) LA: I am sorry but I am allergic to nutty cakes and [/] and I don't accept it (LRE) LA: Okay I would like to bring a natty cake LB: Ah okay LA: natty LB: I like LA: Okay LB: hmm@pI would like to bring a cap LA: A cap? Okay LB: A cap no@s:c? Ah okay because I am really sensitive to the sun and I have to protect myself from the radiation (LRE)

*02_1B_*					LA: Okay [/] okay I accept it LA: Okay I would like to [/] to bring a ram no rum (LRESR-err) because I like drink and I expect that you like it too LB: I don't like drinking but I am not I am fine accepting that you bring the rum LA: okay LB: I would like to bring an umber bag LA: amber or +…? LB: umber LA: umber ok (LRE) LB: Because I don't like amber colur LA: I like amber colour and you like umber we have a problem LB: I hate the amber colour LA: okay LB: What about what about to the bags? LA: To the bugs? What? (LRE) LB: Bags bags (LREREP) LA: But that's it we have six items no? LB: But we have we haven't a bag! LB: So what if we take a don't take the rum with us but we have to decide which colour I hate hate amber LA: I love love umber bag LB: No you have to read what it says here it says that you love um umber LA: I love no no it okay okay LB: So LA: It say it says that that you hate amber bag LB: Amber! LA: Amber colour you hate it LA: Yes LA: Okay I accept it (LRE)

*11_1A_*	2	1	1	1	LA: I want to [/] to bring my collection of mugs because +/. LB: Why do you want mags in a plane? Mugs (LRESR-err) LA: Because <I've always drinking I've always> [//] I'm always drinking coffee I love coffee so I need my mugs LB: But do you want to read a mag? LA: No mugs LB: No yes but it will be great to have a mag with a mug LA: A mag with a mug yes hmm@p but that looks fine (LRE) LB: Okay so ——————— LA: What about hmm@p our bags? LB: Ah yes LA: Our bags I have an amber bag (LREREP) LB: I think that <it's> [//] it could be better an umber bag because +/. LA: Why? LB: Because hmm@p in Kenya the the ground it's brown and< it can you don't> [//] people won'ts look to you if you have an umber bag LA: Okay ——————— LA: And do you think we can bring a natty cake? A natty cake? (LREREP)
*12_1B_*					LB: Or nutty? LA: No a natty cake LB: Natty natty nutty nutty no natty okay(LRE) LA: They are beautiful LB: Beautiful LA: Yes LB: But what do you prefer a nutty cake or a natty girl? LA: I don't like nutty cakes I am allergic to nutty cakes so +/. LB: No? LA: No I prefer a natty cake LB: A natty okay so put it here it will be fine ——————— LA: Or we can bring rum LB: Rum? LA: We can fill our cups with rum and drink drink it in the plane LB: Okay m hmm@p maybe like a botellon@:s or not okay that's great put it here (LRE) LA: Yes? with rum /rʊm/ LB: Rum /rʊm/ no rum /rʌm/ LA: Rum /rʌm/ (LRERC)

*13_1A_*	3		3	4	LA: I think that we can bring a [/] a sun protection and mags LB: Mags? LA: Mags LB: Okay (LRE) LA: I would bring a [/] a natty cake natty LB: Natty? LA: Natty [/] natty because I like cakes and it's a nice cake and I like it and +… (LREREP) R: Do you agree? LB: No a nutty cake LA: Nut cake? Nutty? LB: Nutty LA: Oh no I can't I can't eat nutty cake I am allergic yes [/] yes it's true and I can't eat this please don't bring a nutty cake (LRE) LA: And we we can bring a cup or mug with you prefer if you prefer mug or cup [//] cap (LRESR-err) LB: cap or cup? LA: cup cup (LRE) LB: A cap is better LA: Why? LB: Because you protect of the sun LA: Okay and but I think <I like> I love coffee and I drink coffee every hour after dinner and I need coffee <for life> [//] for live LB: I don't like hmm@p I don't like it
*14_1B_*					R: Do you have anything to drink your coffee? LA: Sí A mag mug [/] mug (LRESR-err) (LREREP) LA: And what else? LB: hmm@ LA: For the sun LB: Umber bag LA: Amber or umber? LB: Umber LA: Umber oh I like more amber (LRE) LB: No LA: No? Why? Amber it's it's perfect

					LA&LB: amber [/] amber (LREREP) LA: Okay put amber [//] umber (LRESR-err) LA: What else? What else? LB: A mags? LA: Mags? LB: Yes LA: Mags okay hmm@p yes (LRE)
*15_1A_OK_* *16_1B_*	3			1	LB: I would like to bring an umber bag to put all my things and +/. LA: Which colour? LB: Umber LA: Well I [/] I like the umber colour but I [/] I prefer an amber colour bag (LRE) LB: No I hate that colour I'm so sorry LA: Doesn't matter LA: Do you think is a good idea to bring the cup from the English or British flag? LB: hmm@p Yes LA: Yes, okay? R: Yes? LB: Ah a cup! I understand a cap like for my head (LRE) LA: No I said a cup a small +/- LB: Yes but I need a cap for my head because I have to <go> hmm@p [//] be careful with the sun LA: Yes but I can't go anywhere without my coffee LA: A natty cake for the plane to eat if we are hungry? LB: Yes well I prefer a nutty a nutty cake (LREREP) LA: I am allergic to nutty cakes and I cannot I can't we take the natty cake okay? LB: You can take some bugs LA: hmm@h bugs the insects? LB: Yes LA: For the plane? To do what with [/] with them? (LRE)

*17_1A_* *18_1B_*	2		2	2	LB: I want to bring an umber [/] umber bag? (LREREP) LA: hmm@p <no se què dir@s:C> R: Ask her LA: <no se@s:c> Do you want to bring a amber bag? LB: No an umber is like brown no? (LRE) LA: No but amber bag is very important to put objects LB: But I hate the umber colour so +/… LA: Okay another LB: What about if we bring a nu [//] nutty cake? (LRESR-err) LA: No because I'm allergic to nutty cakes LB: <en serio@s:c?> I understand R: What is your alternative? LA: hmm@p Natty [/] natty cakes (LREREP) LA: I want to [/] to bring a mag for reading LB: Okay and a ram? LA: Ram? LB: What's ram? LA: A a drink? R: It's not ram it's not ram the male sheep it's it's +…(LRE) LA: Rum rum (LRESR-err)
*03_1A_*	6				LB: Yes, I'm [/] I'm hungry I [/] I propose to put in nutty cake LA: Oh I <espera com era@s:c?> I'm allergic to nutty cakes LB: But I I love this cake with nuts LA: Do you like natty cakes? LB: Natty cake? LA: Yes Maria, natty cakes LB : Yes okay okay it's [/] it's good (LRE) LB: I need a cap too! LA: What? LB: I need a cap too LA: Okay? LB: Because the sun it's dangerous LA: Yes [/] yes (LRE)

*04_1B_*					LB: So <you tell> [//] you told me you want a bag LA: No a cup LB: A cup? LA: Yes a cup oh but we need a bag too (LRE) LB: A bag? LA: Yes a bag (LRE) LB: Umber bag? LA: Yes LB: Yes umber bag? Okay LA&LB: amber/umber? LA: No, vale@s:c LB: Yours is amber and I need umber bag LA: Well LB: Because I [/] I hate umber colour LA: hmm@p well okay LB: Umber bag (LRE) LA: Or rum? LB: Rum? LA: Rum yes LB: But I [/] I don't like drink I think it's it's bad you know? LA: Yes [/] yes (LRE)
*05_1A_*			1	1	LA: And well we'll need +… if we want to [/] to carry all of these things we need a bag right? LB: Yes LA: Maybe I've had a pretty amber bag which can be beautiful LB: I don't like the amber colour (LRE) LA: Why not? LB: I don't know but I don't like this. I prefer the umber [/] umber colour (LREREP) LA: I haven't anything umber but if you have something we can carry it LB: Okay LA: If we go to a party or something we will need to carry some alcoholic drink or something and in Kenya I think there are very expensive maybe we can carry +/.

*06_1B_*					LB: In a party I prefer a nat [//] nutty cake (LRESR-err) LA: I think that we can also go with a cake but I think nutty is a bad idea because I am very allergic to nutty cakes and I want some cake maybe a natty cake? LB: Yes [/] yes LB: In Kenya I need to whip /wɪp/ LA: To weep /wi:p/ ? Why do you need to weep /wi:p/? You are sad? LB: To control it LA: Ah to whip /wɪp/! Okay maybe we find some lions and you can use them and act like Indiana Jones if you feel like doing it good idea [i]
*07_1A_* *08_1B_*	3	1	1	2	LA: I want to bring a cap /kæp/ cop /kɒp/ R: cup /kʌp/ LA: cup /kʌp/ because I I never go anywhere without my coffee I need to bring (LRERC) LB: Okay LA: Okay? LB: Ah okay (LRE) LB: You bring the knife and I bring the nutty cake LA: No no no no because I am allergic LB: No LA: I am allergic hmm@p to natty cakes nutty (LRESR-err) so maybe we can bring a natty cake LB: Com@s:c? LA: Natty nutty <tu has dicho@s:c> LB: Ah <vale vale vale@s:c> LA: <Si no?@s:cz+> I want to bring a natty cake (LRE) LB: Yes yes LA: Perfect LB: Then we can take my umber bag LA: Not my amber bag LB: No LA: Amber no <la mía@s:c?> amber (LREREP) LB: I hate amber LA: Well if if you want don't worry if you want LB: My bag? LA: Yes your umber bag but I prefer the amber amber bag (LREREP) (LRE)

*09_1A_*	2				LB: I think that we can bring a umber bag because I hate the amber colour R: What do you think? LA: I I like amber colour LB: Amber or umber? LA: Amber LB: I hate amber colour LA: Pues@s:c I am I agree LB: Thank you (LRE)
*10_1B_*					LA: hmm@p I think I bring a natty cake LB: A nutty? LA: Nor natty nutty LB: I prefer a nutty cake LA: No I am allergic to nutty (LRE)

TASK 3

CODE	**LRE**	**LRERC**	**LRESR**	**LREREP**	**Examples**
*01_1A_*	7	2	2	3	LB: Jack is going to wear umber jacket LA: Amber or umber? LB: Umber LA: Umber LB: with U LA: with U okay I don't accept it because is [/] is forbidden or prohibited R: forbidden LA: for forbidden (LRE) R: So you said the vet would bring a +…? LA: A ram LB: Rum LA: Ram LB: Ram is a animal rum is alcohol […]
*02_1B_*					LA: But ram /ræm/ is an animal and in your paper says that ram /ræm/ because ram /ræm/ is an animal and you can't come with animals okay in so no in the paper? LB: Yes but rum /rʊm/ R: rum /rʌm/ LB: rum /rʌm/ is alcohol (LRERC) LA: No ram /ræm/ I say ram /ræm/ the animal! LB: What? LA: Ram [/] ram /ræm/ (LRE) LB: Okay okay LA: The police woman wears a blue t-shirt LB: Yes I accept it LA: Okay LB: <a veure@s:c> no no LA: Why? LB: Because she it says here that she wears an umber t-shirt

					LA: Amber with A? LB: with /ʌ/ U! LA: No I don't accept it because in the paper it says that people can't [/] can't wear bring anything umber with u! LB: Umber? LA: Umber okay? LB: This is impossible LA: And that's is why I come with blue t-shirt (LRE) LB: So the policewoman wants to bring tins and bu bags (LRESR-err) LA: Tins with bugs [/] bugs ? (LREREP) LB: bags! LA: Bags okay I accept it LB: Yes LA: okay tins with bugs LB: Not bugs bags! LA: Ah ba okay okay bags okay (LRE) LA: We start with Tom the chef /ʧiːf/ LB: The chef /tʃɛf/ LA: Chef okay sorry /tʃɛf/ [i] LB: So he wants to wear trousers full of cuts LA: Okay I accept it because animals can come to the party you say cats or cuts? LB: Cuts! LA: Cuts ah okay okay LB: Okay (LRE) LA: In in in Jack you said before that he will hmm@p bring rum rum /rʌm/ the like the +/. (LREREP) LB: Rum LA: Rum okay in my paper says that alcohol are prohibit [/] prohibited R: prohibited /prəˈhɪbɪtɪd/ LA: prohibited and that's why we have to [/] to +/. LB: Okay I already cross it out [i] LA: Lucy wears a t-shirt with an amber ship

					LB: With an umber +/.? LA: Amber [/] amber with A with A sheep sheep like the sheep the yes the [/] the animal (LRESR-err) LB: So let's discuss this after I let you that someone bring a blue t-shirt why right now you don't let me that someone bring a +/. LA: Umber with u LB: Umber [/] umber (LREREP) LA: Because in my paper say people can't wear anything umber! […] LA: And in your paper says that the persons can't [/] can't wear an umber LB: No it says that it can [/] can wear anything green green LA: Okay if if in your paper has [/] has <esperate@s:c> LB: Ah okay so we wear amber LA: Amber LB: Ah okay amber ship (LRE) LA: T-shirt with an amber sheep LB: sheep amber +… LA: like the animal okay? LB: No LA: Yes the animal sheep LB: Like the a ship [/] a ship like the a Titanic LA: No LB: Ship /ʃiːp/ ship /ʃɪp/ LA: No <ovella@s:c> in Catalan [i] LA: Shows her bat […] LB: That she shows her butt LA: Her butt? Okay okay I accept it (LRERC) (LRE)

*11_1A_*	9		2	3	LA: Jack has an injection hmm@p and brings a ram okay? He can he can bring a ram LB: A ram? LA: Yes ram LB: My Jack drink drinks at home rum LA: Ah drinks rum? LB: Rum okay so LA: Yes but in my party I do not want my guests to bring alcohol and cigarettes so he can he can't bring the alcohol LB: Okay [/] okay nice(LRE) LB: But could he wear umber colours [//] umber trousers LA: Amber? LB: Umber LA: Amber or umber? LB: Umber (LRE) LA: Umber okay hmm@p no I want Jack to bring wear trousers because he is the vet and you know green trousers and hmm@p and I want him to show pictures of monkey's butts /bæts/ LB: Bat butt? LA: To show pictures of monkey's butts /bʌts/ (LRESR-err) LB: Of monkey's butts okay nice it will be good # LB: I'm sorry about green [/] about green trousers maybe it could be better hmm@p like hmm@p brown trousers LA: Brown? LB: Yes brown like umber trousers LA: Amber or umber? LB: Umber you know like brown LA: I can't that… No people can't bring anything umber LB: So amber [/] amber could be great for both (LREREP) LA: Okay (LRE) LA: What else? LB: And then bring tins and bags LA: With bugs LB: Bags yes

*12_1B_*					LA: Bags or bugs? LB: bags LA: Bags? LB: Bags! Yes you know LA: Ah okay tins with bags LB: But that doesn't make it sense (LRE) LB: She probably will wear [/] wear an umber t-shirt [/] umber t-shirt [/] umber t-shirt you know umber (LREREP) LA: Umber? LB: Umber t-shirt LA: Umber t-shirt what about a blue t-shirt? LB: Okay LA: Because she is a police woman LB: Okay blue t-shirt okay (LRE) LA: I think hmm@p she might to want to show her butt LB: LOL? LA: Her bat (LSRESR-err) LB: So she has a bat LA: Yes she has a bat a bag ai@s:c a bat that flies LB: But she isn't batman she? Just have +… LA: No no because can't come cannot to our party LB: But if she shows her butt? LA: Butt? LB: Yes butt LA: Hmm@p no because they have to behave with respect a bat would be better (LRE) LA: Wears trousers with cats LB: Yes <no no no no> [/] no full of cats LA: Cats LB: Meow No no LA: Wears trousers full of cuts (LRE) LA: And whips Lucy? LB: No it would be better +… he weeps [/] weeps with Lucy LA: Weeps?

					LB: Yes we have to behave with respect LA: So weeps with Lucy [i] LB: Yes and he wants to stub [/] stub out Lucy's /lu:si/ cigarette (LREREP) LA: She and he wants to stub out Lucy's /lusi/ cigarette LB: Yes Lucy /lusi/ cigarette LA: Yes but he also wants to stab Lucy after her cigarette so he can stub out her cigarette and then stab her LB: Both LA: In the back LB: Yes it would be nice so stab and stub (LRE) LA: Brings no [//] wears a t-shirt with an amber sheep LB: Amber? Okay LA: Amber LB: Yes (LRE) LA: Wears a t-shirt with an amber sheep LB: No it doesn't +… amber? What the fuck? A sheep [/] a sheep the [/] the animal it they are white LA: Well but the +/. LB: It's better a ship [/] a ship! It's better a ship R: And? LB: And I don't like to show animals so it's better a ship LA: You don't like that okay [i]

*13_1A_* *14_1B_*	7		3	8	LA: A ram [/] ram ? LB: Rum LA: Ram LB: Rum LA: No ram [/] ram LB: Rum LA: No ram no I I don't +… Jack the vet <he can't> [/] he can't bring alcohol and cigarettes are prohibited R: Prohibited LA: For forbidded [ii] (LRE) LB: Vet wears an amber trousers LA: Amber? LB: Amber LA: Or umber? LB: Amber LA: Amber? LB: Umber LA: Amber am am um am (LRESR_err) LB: Umber LA: With U? LB: Yes LA: No LB: Why? LA: People can [//] can't bring anything um umber it's forbidden it's a lot of conditions R: So maybe LA: Green [/] green trousers LB: People can't wear or bring anything green R: So maybe LA: Amber [/] amber (LREREP) LB: Ah clar LA: No amber! LB: Amber LA: With U is forbidden forbidden [/] forbidden LB: Okay (LRE) LA: Shows pictures of monkey's butts /bæts/

					LB: Bats? LA: Butt /bʌt/ LB: Butt (LRESR-err) (LRE) LB: She brings tins and bags LA: Bags no LB: Bags LA: With [/] with +… Tins <que es tins@s:c> LB: Bags [/] bags (LREREP) […] LA: Brings tins with bugs no bags bugs bugs[/] bugs (LREREP) LB: No no no no no LA: Oh why? LB: Anything about animals is not allows LA: Okay <tins with bags> [/] bags tins with bags (LREREP) LB: Bags LA: <Yes yes yes yes yes [/] yes> (LRE) LA: And show her [/] her bat? LB: Bat? LA: Yes her [/] her bat LB: Okay LA: Bat [/] bat (LREREP) LB: Bat LA: No butt LB: Bat LA: Bat [/] bat (LREREP) LB: Shows her butt (LRESR-err) LA: No with respect please LB: Bat [/] bat (LREREP) LA: Bat LB: Bat LA: Bat (LRE) LB: And wants +/. LA: Stub [//] stab (LRESR-err) LB: Stab?

					LA: Stab Lucy after [/] after the cigarette stab [/] stab (LREREP) LB: Stub LA: Stab no stab no no violence please […] LB: No violence no LA: Stub no cigarrettes are prohibited are forbidden (LRE) LA: Lucy has pills and wears a t-shirt with an amber ship [/] ship LB: No Lucy can't wear a green +/. LA: Amber no green amber! LB: Amber LA: Amber not too? Amber! LB: Umber LA: No umber <can bring> [/] can't [/] can't bring (LRE) LA: But sheep or ship? Sheep ship? Sheep ship? Sheep? LB: <Què es?@s:c> LA: Sheep wears una@s:c amber sheep or ship LB: No LA: Ship LB: Sheep LA: Sheep no ship LB: Sheep LA: Ship […] LA: <Ship ship ship> [/] ship [i]

*15_1A_OK_*	7		1	3	LB: Do you agree to wear some green trousers? LA: No LA: Why? LB: Because I don't want someone going with green things or [/] or anything green LA: Okay LB: What [/] what do you think? LA: I think the umber colour is better LB: What colour? LA: Umber (LRE) LB: No sorry anyone can come with something or with umber [/] umber colour (LREREP) LB: Maybe wear trousers full of cuts LA: Yes I think it's +/. R: Full of… LB: Cuts LA: Repeat? LB: Cuts LA: The animal or the +/.? LB: No [/] no the action (LRE) LB: Wants to stub out Lucy's cigarettes LA: Hmm@p okay LB: Yes? LA: No what did you say? Repeat please? LB: Stub out Lucy's cigarette […] LB: First we have to stub out the cigarettes and then stab with the knife Lucy well we agree no? R: Do you agree to stab and stub? LB: No I think not stab we have to have respect LA: Okay R: So only+… LB: Stub out! (LRE) LB: What about wears an umber t-shirt? LA: Repeat? LB: Umber t-shirt
*16_1B_*					

					LA: No I said before the umber colour is prohibited we cannot wear anything with that colour so the woman can wear a blue t-shirt LB: uhhuh (LRE) LB: What about if Ann brings tins and bags? LA: Yes I think the animals can come and we +/. LB: Bags [/] bags not the insect bags (LREREP) (LRE) LA: What do you think if Lucys shows her bat? LB: Who? LA: Bat the +/. LB: Lucy Ann sorry LA: Oh yes R: Bat? LA: The animal I think LB: No butt [/] butt (LREREP) LA: Never mind and it doesn't matter but well we have to respect no? (LRE) LB: What about if she wears a t-shirt with an umber sheep? LA: Eh can you repeat? LB: With an umber sheep LA: Amber or umber? I don't+/. LB: Umber (LRE) LA: No no? I think no [/] no? We cannot bring anything with umber colour hmm@p <espera@s:c> what sheep repeat please? LB: What? LA: What does the t-shirt has? LB: Ship (LRESR-err)

*17_1A_* *18_1B_*	4			1	LB: What does Jack wants to bring to Kenya? LA: Jack brings a ram LB: A rum to drink? LA: No a ram LB: But animals are not allowed in Kenya (LRE) LB: Ann shows her butt LA: No her butt <sino@s:c> her bat LB: Bat? LA: But animals are not allowed! LB: Okay a butt (LRE) LA: And wears trousers with cuts LB: Okay it's good R: With what? LA: Cuts LB: Cats the animals? LA: Cuts […] LB: But animals are not allowed! So he can bring trousers with cuts! (LRE) LB: What does Lucy wants to do? LA: hmm@p Lucy wears a t-shirt with an amber sheep LB: Amber like orange? Or umber? LA: Amber amber (LREREP) LB: But it's better the umber colour LA: No because people can't wear anything umber LB: So amber (LRE)

*03_1A_* *04_1B_*	5			1	LB: What do your vet bring? LA: Hmm@p he brings a ram and your +/. LB: Ram or rum? LA: Ram LB: Okay my vet brings rum to drink LA: But in the party well alcohol and cigarettes are pro+… R: Prohibited LA: Prohibited [ii] (LRE) LB: My vet wears umber trousers I like this colour LA: Umber? Mine green LA: Green? No no [/] no impossible I don't [/] I don't like it [//] it's a horrible colour LB: Amber or umber? LA: Umber LB: People can't bring or wear anything umber in my +/. LA: And in my role party people can't bring or wear anything green R: Decide on another colour LB: Okay you +… LA: Blue LB: Or amber? LA: Perfect (LRE) LB: My chef wear trousers full of cuts you +… LA: My chef too LB: Yes? LA: Yes LB: Trousers with cuts LA: Ah cuts or cats? LB: Cuts [/] cuts LA: <No tía@s:c> LB: Yes but LA: Well I prefer cuts than cats LB: Because she have [//] she has a knife and cuts +… (LRE) LA: I think that he wants to stab Lucy after her cigarette and you? LB: Yes stub out Lucy's cigarettes because probably he don't like cigarette and smokes and in a party I think cigarettes and alcohol is [//] are welcome

					LA: Yes he wants to stab Lucy! Stab Lucy [/] stab (LREREP) LB: Ah okay! LA:< Ah havies en@s:c>… No no stab Lucy LB: Yes yes […] LA: No [/] no stab Lucy LB: Tom is the murderer (LRE) LA: She wears a t-shirt with an amber sheep you? LB: Umber? LA: Amber LA: sheep LB: Amber ship you tell me? LA: Sheep [/] sheep or no [i] LB: Ship or sheep? LA: Sheep LB: I tell you ship LA: No the other LB: Why not? LA: Sheep LB: Wears t-shirt with an amber ship! I prefer that LA: Okay (LRE)
*05_1A_*	8		2	1	LA: I think that Jack will bring a ram is it okay right? LB: No LA: Why not? LB: Ah yes yes [/] yes Jack bring no! He bring rum to drink yes LA: Ah rum okay but not a ram LB: No LA: No okay but I think that rum is not a good idea because we shouldn't drink alcohol in the party (LRE) LA: But what if she shows a picture of a monkey's butt /bat/? LB: Hmm@p What? LA: What if she shows a picture of a monkey's butt /bʌt/ (LRESR-err). That's okay?

*06_1B_*					LB: Yes LA: Yes? Picture of a monkey's butt /bʌt/ (LRESR-err) (LRE). LA: But she also wanted to bring a tin with [/] with bugs LB: Tin? LA: No, like a tin with bugs. Do you think +/. LB: Yes she brings a bags LA: Yes as animals are welcomed yes LB: No animals +… R: What do you have? Tins and +… LB: bags LA: Maybe it's not a good idea LB: No LA: We shouldn't let her right? Okay so not bugs (LRE) LA: And finally I think that she planned to [/] to show a bat LB: Bat or butt? LA: Yes a bat I think it can be strange but +/. LB: No hmm@p is shows her butt LA: No I don't think she will show her butt because I need that people should behave better than this so maybe not we shouldn't let her show her butt because it can be offensive you know (LRE) LA: But he planned that's what really scared me to whip Lucy LB: weep with Lucy? LA: Whip Lucy [/] whip Lucy (LREREP) but I think it's a bad idea maybe he shouldn't do it LB: it's weeps with Lucy LA: Ah that she weep yeah I think that can be okay yeah why not? (LRE) LA: But I've also heard that instead <of weeping> [/] of weeping with Lucy he wanted to stab Lucy after her cigarette. Did you hear so or did you hear another thing? LB: Yes LA: Yes? LB: Yes LA: With the knife? LB: No! She wants to stub out her cigarette LA: Ah okay stub! Okay yes that's probably yes I think that's okay right? (LRE) LA: And I think that she finally wanted to wear a trousers with cuts which must be okay right?

					LB: okay LA: Yes? LB: Cats or cuts? LA: Cats [/] cats the animals (LREREP) LB: No LA: Why not? LB: Hmm@p in the party I don't have hmm@p about animals LA: nothing about animals! LB: # trousers full of cats [//] cuts! (LRESR) (LRE). LA: But she also was wearing [//] was going to wear a t-shirt with an amber sheep LB: Sheep? LA: Yes LB: Sheep or ship? LA: Sheep[/] Sheep [i] LB: Sheep okay LA: So do you agree? LB: Can you repeat please? LA: That she was going to wear a t-shirt with an amber sheep LB: An amber sheep LA: Yes? Do you agree? LB: Yes yes (LRE)
*07_1A_*	4		1		LB: Jack wants to bring rum to the party LA: No LB: Why? LA: Rum? LB: Rum LA: Rum not this is not possible because alcohol are prohibited so he can't bring rum so he can bring some <ram ram> [/] ram yes why not LA: No because anything about animals is not allowed in the party (LRE) LB: Jack wants to wear umber trousers LA: Not this is not possible LB: Why? LA: Because people can't wear anything umber because hmm@p umber <era un así marron no@s:c?>

*08_1B_*					LB: uhhuh (LRE) LA: because the the animals can enter to the party so if they see amber[//]no umber so they can eat this so he can wears green trousers because it's like the jungle you know green the jungle it's the same so the animals are going to be like in his house (LRESR-err) LB: No one in the party would wear anything green LA: Why? LB: No LA: Give me a reason [/] give me a reason LB: Kenyan people hate green too LA: I don't know that why? R: Choose another colour LB: Amber? LA: Oh an amazing colour this is the colour that Jack is going to wear so amber trousers no? Yes LA: And he is going to bring tins with bugs LB: Bugs? LA: Bitxo@s:c LB: No no [/] no anything I repeat you anything about animals is not allowed in our party […] LB: But she can bring tins and bags LA: Tins and +/. LB: Bags LA: Perfect (LRE) LA: So she can show her butt LB: <Que era butt?@s:c> LA: Butt LB: El culo? Ah don't worry she can show the [/] the butt and +… (LRE) LA: He wants to wear a t-shirt with amber sheep LB: Perfect LA: Perfect? Perfect? You said perfect eh! LB: uhhuh LA: Sheep vee@s:c LB: No <a ver a ver@s:c> LA: But you said yes! […]

					LA: But he can [/] she can bring an amber t-shirt with ship LB: Okay! LA: Amazing incredible [i]
*09_1A_*	4			1	LB: He can wear amber [/] amber trousers (LREREP) LA: Amber or umber? LB: Umber LA: No I didn't like umber (LRE) LB: She can also bring tins and bags LA: No animals are prohibit […] LA: <No no no no no> no bags with U LB: I said bags with A LA: Ah okay (LRE) LB: He wear trousers that are full of cuts LA: He is dangerous LB: Why is danger LA: Cuts no? LB: The trousers that girls wear nowadays are full of cuts LA: Ah okay [/] okay (LRE) LA: Amber or umber? LB: Amber and ship (LRE)

TASK 4

CODE	**LRE**	**LRERC**	**LRESR**	**LREREP**	**Examples**
01_1A_	5			5	LA: Okay I think that we can choose for example a pink trousers with cats LB: Hmm@p Can you repeat? LA: Trousers with cats [/] cats (LREREP) LB: Ah okay hmm@p yes yes I think we could (LRE) LA: What about a picture of a baby babbles LB: What? LA: A picture of a baby babbles LB: A baby <un bebé burbuja@s:c?> LA: Babbles babbles babbles (LREREP) LB: No LA: No why? LB: Because I think that I <o sea@s:s>I have a baby making bubbles LA: Okay I accept it # (LRE) LB: What about a picture of a cup in a bag? LA: A cap or a cup? LB: A cup LA: cup in a bag hmm@p okay (LRE)
02_1B_					LB: What about a picture of an amber bag? LA: An amber? LB: With +… Of an <amber bag> [/] amber bag (LREREP) LA: Amber colour items must must not appear (LRE) LA: Okay what about a picture of a person who is eating a nu nutty [/] nutty cake nor a natty nutty okay? (LREREP) LB: Yes I have a natty cake but if you prefer nutty LA: No no <we can> [/] we can choose the two LB: Both? LA: Both LB: Ah okay (LRE) LA: I have a picture of bat […] bat (LREREP)

11_1A_ 12_1B_	11	1	2	7	LB: So we can put this a trouser with cuts [/] a trouser with cuts [/] with cuts /kæt/ (LREREP) LA: A trouser with cuts /kʌts/ LB: Yes cuts /kʌts/ LA: Why not? (LRERC) (LRE) LA: Do you want to choose the picture of the natty cake? LB: Ah okay yes great yes LA: Yes? A natty cake (LREREP) (LRE) LB: And if I put a group with a picture of with zoo caps LA: Zoo cups? LB: If we can put here a picture with a group with zoo caps (LREREP) LA: Okay LB: A picture in group (LRE) LA: And a picture of a mag? LB: Of a mag? But what's about the mag? LB: About gossip LA: Gossip? LB: Gossip okay (LRE) LB: And [/] and a picture with Sandra with rum rum [/] rum? (LREREP) LA: And Sandra with rum? But alcohol is not allowed LB: No? And +… LA: No but maybe Sandra with the ram? LB: Ram? LA: Yes Sandra with the ram (LREREP) LB: I don't like it (LRE) LA: Ah okay LB: I don't like animals LA: You don't like animals LB: And a natty cake? LA: A natty cake? LB: But we already have a cake (LRE) LA: And a picture of a baby making bu [/] bubbles?

					LB: Hmm@p okay LA: Yes? LB: Yes bubbles no babbles bubbles LA: Bubbles yes but he could also be making [//] babbling LB: Yes babbling yes (LRE) LA: And a cup in the bag [/] a picture of a cup in the bag? LB: What [/] what? LA: A cup in the bag (LREREP) LB: No no no okay I understand LA: There is a man putting inside a cup into a # LB: Yes why not? (LRE) LA: But a group with [/] with zoo cups /kæps/ [/] with [/] with zoo cups /kʌps/ (LRESR-err) Do you remember a picture of that? LB: Yes [/] yes I [/] I remember it do you want it? LA: Because James doesn't appear in this picture LB: Zoo cups yes? LB: Zoo [/] zoo cups (LRE) LB: Caps not cups [/] caps not cups [/] cups not (LREREP) LA: Do you want a group with zoo caps? LB: Yes LA: Well that would be fine (LRE) LB: I have sheep in [/] in the field but I can't put it sheep sheep [/] sheep in the field I can't put it there […] [i] LA: Hmm Sandra with the ram? But they don't want animals LB: Yes LA: And an [/] an umber bag [//] no an umber bug? (LRESR-err) LB: No I don't like um umber colour

13_1A_	5		2	8	LA: A natty [//] nutty cake nutty [/] nutty (LRESR-err) (LREREP) LB: Nutty? LA: Nutty cake LB: Okay (LRE) LB: An umber [/] umber bag (LREREP) LA: Umber? LB: Uhhuh LA: No it's +/. LB: Why? LA: Mustn't umber color mustn't appear in the pictures it's +/- LB: Um umber LA: Umber [/] umber (LREREP) LB: Mustn't appear LA: Okay (LRE) LA: And the picture there [/] there is a mag LB: Hmm@p LA: Mag! (LREREP)
14_1B_					LA: Sandra with the ram [/] ram (LREREP) LB: Rum? LB: And a cup in the bag? LA: Cup? [/] Cup yes (LREREP) LB: Yes? LA: Yes (LRE) LA: Trousers with cats LB: Cats? LA: No but cats paint in the trousers no really cats LB: Okay (LRE) LB: And trousers with cuts [/] cuts? (LREREP) LA: Yes! Why not? LA: The picture who there's a group with zoo caps /kæps/ cups /kʌps/ (LRESR-err) LB: Yes

					LA: A natty cake? R: Natty? LA: Natty [/] natty (LREREP) LB: Yes! (LRE)
15_1A_OK	6			2	LB: Maybe we can put a photograph of a amber [/] amber colour bag (LREREP) LA: Repeat please? LB: umber LA: Hmm@p I think you said umber colour I cannot +… a colour more brown or yellow? LA: Repeat LB: Amber LA: Sorry but the school thinks that umber colour is prohibited and well we cannot appear anything with that colour (LRE) LA: What about a natty cake? LB: Repeat please? LA: Natty cake na na natty cake (LREREP) (LRE) Lb: And what about a trousers with cuts? LA: Repeat? LB: Cuts LA: A trousers with cuts […] yes it's okay because I prefer [//] well I have trousers with cats if you prefer LB: Well no animals should not appear LA: I forgot it then okay the trousers with cuts (LRE) LA: What about a group a photo of a group with zoo cups? LB: Repeat? LA: cups LB: But a group of +/. LA: Yes with zoo cups (LRE) LB: And what about a cup in the bag? LA: Repeat? LB: Cup LA: of a cup of what [/] what do you said? LB: A cup in the bag
16_1B_					

					LA: hmm@p okay I +… (LRE) LB: What about hmm@p stab Tim? LA: hmm@p? LB: Stab Tim a photo of +/. (LRE) LA: Well I think the signs [/] signs of violence cannot appear and I think have not to put a picture of Tim stab stabbing someone
17_1A_ 18_1B_	6		1	3	LB: What kind of pictures do we need put in the webpage? LA: A trousers with cats for example LB: With cats or cuts? LA: Cats [/] cats (LREREP) LB: Hmm@p Okay (LRE) LB: I think we can [/] we can put a the [/] the picture of the amber bag LA: Amber or umber? LB: Amber LA: Amber colour items must not appear and umber bug? (LRE) LB: And a [/]] the picture of the [/] a sheep in the field sheep? (LREREP) LA: Yes it's good LB: We can put a [/] the picture of Sandra with the rum LA: Ram or rum? LB: Rum LA: But alcohol is prohibited and I prefer Sandra with the ram LB: Okay (LRE) LB: We can put a baby making bubbles LA: Hmm@p a babble or bubble? LB: Bubble LA: I prefer a baby babbles LB: The two! (LRE) LA: Okay LA: And a natty /næti/cake? LB: Natty /næti/or nutty /nʌti/?

					LA: Nutty /nʌti/ LB: I prefer the natty cake LA: The two? LB: Okay (LRESR-err) (LRE) LA: And a mag? LB: Mag? Mag or mug? LA: Mag a gossip time a magazine LB: Okay (LRE) LB: We can put a the picture of the trousers with cuts LA: Cat /kæt/ or cut /kæt/? (LRE) LB: Cut [/] cut /kʌt/ LA: Okay (LREREP)
03_1A_ 04_1B_	6		1	1	LB: You say ram? LA: Yes Sandra with the ram LB: Okay (LRE) LB: You and you would like a natty cake? LA: What? LB: Natty cake LA: And a nutty cake? LB: I [/] I don't have problems with nutty cake but I like much natty cake LA: Okay natty cake! LB: Both too? LA: Perfect (LRE) LA: A group with zoo cups? LB: Cups? LA: Yes LB: Why cups? LA: So +/. LB: They are normal in the zoo and have have hot LA: Well in [/] in my picture people have caps and cups both LB: Okay caps and cups (LRE) LA: I have a cup in the bag

					LB: What? LA: A cup in the bag for an emergencies for example right now I need a cup [/] a cup LB: Of coffee perfect I like this picture LA: Yes? You need a cup too? LB: Yes LA: Okay LA: And a mag? Picture of a mag? Do you like mags? LB: Ah a mag! LA: Yes LB: <I I> [/] I don't have any picture with mag or mug but I love read so it's good (LRE) LB: I have amber bag! LA: No amber? LB: Amber bag (LREREP) LA: Oh no amber colour items must not appear in my pictures [//] our pictures so umber bug /bæg/ bug /bʌg/? (LRESR-err) LB: No I don't accept umber colour in my pictures it's not possible (LRE) LA: And a bat? LB: Bat? LA: Do you like it? LB: Monkey and its butt you know LA: Yes I like bats LB: Ah bats okay it's better than butt is hmm@p little bit sexual (LRE)

05_1A_	5		2	5	LA: So what else? I think that Sandra wanted to appear with a ram do you agree? LB: A ram? LA: Yes LB: Hmm@p o I think that bueno@s:c I believes that animals shouldn't appear in the pictures LA: And a bat neither? LB: What? LA: And a bat [/] maybe a bat? (LREREP) LB: Ah bat no LA: No right? LB: No (LRE) LB: What do you think about trousers with cuts [/] cuts? (LREREP) LA: I think it can be violent LB: Cats? The animal or +/.? LA: The trousers with designs of cats LB: Ah designs okay LA: Do you agree? Because it's friendly and funny! LB: uhhuh (LRE)
06_1B_					LA: So what else? Do you have any other photo which do you think we can +…? LB: A cap /kæp/ in the bag? Cap? /kæp/ Cup? /kʌp/ (LRESR-err) (LREREP) LA: Yes why not? But hmm@p what colour is it? LB: Blue LA: Ah okay yes (LRE) LA: So maybe another one of a [/] of a girl with a mag? LB: Can you repeat please? LA: Another photo I took of a girl with a mag (LREREP) LB: With a mag? LA: Yes LB: Why not? LA: Yes do you agree? LB: Yes (LRE) LB: What do you think about the amber bag LA: I don't think +… excuse me, did you say amber?

					LB: Amber (LREREP) LA: No I don't really like amber because umber is much better LB: No LA: No but you don't have +… so maybe not to put any bug LB: Okay (LRE) LA: And we can put as we are a very healthy class or group we can we can put a photo of a [//] of ourselves like with stabbing stubbing up a cigarette stubbing up [//] out a cigarette (LRESR-err) LB: Yes LA: Because we are very healthy and +/. LB: Yes
07_1A_ 08_1B_	3	1		4	LB: How about putting a natty cake of the photo on the website natty cake? (LREREP) LA: Yes yes natty cake […] LA: I have nutty /næti/ [//] nutty /nʌti/ cake do you? (LRESR-err) LB: okay LA: Do you like? Perfect LA: Okay so we can put a bat LB: Bat? LA: Bat [/] bat (LREREP) LB: Bat like murciélago@s:s ah no no no animals must not appear (LRE) LA: Animals with the animals you [/] you have a problem with the animals LA: So we can put a picture about a mag mag [/] mag (LREREP) LB: Mag? LA: Mag mag yes LB: What kind of mag? […] LA: Look me MAG! LB: Okay (LRE) LB: How about trousers with cuts? LA: Trousers with cuts? LB: Okay it's okay but if we can put a trousers with cats

					LA: Sergio animals +/. LB: Yes [/] yes I know that I know that and I understand but are not a real animal it's a picture LA: You have a problem with trousers with cuts? LB: No I don't have a problem LA: I have a problem with trousers with cats LB: But it's not a real animal it's a picture in the trousers LA: No anything LB: But Tom believes animals should appear in the [/] in the pictures LA: Yes but Kim believes that animals shouldn't appear (LRE) LA: Yes but we can put baby babbles [/] babbles (LREREP) LB: Babbles yes
09_1A_	9		2	4	LA: I choose Sandra with the ram LB: Ram or rum? LA: Ram LB: Ram? LA: Ram [/] ram (LREREP) (LRE) LA: A nutty /næti/ cake? LB: Okay I have a nutty cake […] natty or nutty? LA: Nutty /nʌti/ (LRESR-err) LB: Ah I have natty cake R: Do you have any problem with this? LA & LB: No (LRE) LB: So another option? LA: Hmm@p a mag LB: What? LA: Mag LB: Mag or mug? LA: Mag (LRE) LB: And a cup in the bag? LA: Que@s:s? LB: A cup in the bag [/] a cup in the bag (LREREP)

					LA: A cup in the bag? LB: Yes LA: Baby babbles LB: Babble or bubble? LA: Babble LB: Yes okay I have the same LA: No LB: Yes ba with babbles babbles or bubbles? LA: Bubbles LB: Si@s:s yes (LRE) R: You have you have open your mouth LA: Babbles (LRESR-err) LA: Trousers with cats LB: With cats? Hmm@p with cats or with cuts? LA: (silence) LB: I prefer trousers with cuts LA: Signs of violence cannot appear LB: So trousers with cats (LRE) LB: And a amber bag? LA: Amber? LB: Amber bag LA: or umber? LB: No amber bag! (LREREP) LA: No LB: Why not? LA: I didn't like amber color LB: But umber colour items must not appear (LRE) LA: I choose a bat LB: Bat? No LA: Why not? LB: Because Kim believes animals shouldn't appear in the pictures (LRE)

					LA: A group with zoo cups cups [/] cups (LREREP) LB: And with zoo caps? And caps and not cups? I prefer caps LA: The two options LB: Okay (LRE)

[i] Pronunciation LRE recast with a non-target form (not included in the count).
[ii] Lexical LRE (not included in the count).

Appendix G – Extra tables and figures

G.1. Identification test results

Report				
Group		**Accuracy_T1**	**Accuracy_T2**	**Accuracy_T3**
Experimental	N	18	18	18
	Mean	,4632	,9444	,9444
	Std. Deviation	,09260	,04817	,04582
Control	N	18	18	
	Mean	,6514	,6861	
	Std. Deviation	,11896	,12164	

Table 8.1. *Descriptives (means and SD) for word accuracy (pre-test, post-test and delayed post-test).*

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Time	Pillai's Trace	,974	301,306[b]	2,000	16,000	,000	,974
	Wilks' Lambda	,026	301,306[b]	2,000	16,000	,000	,974
	Hotelling's Trace	37,663	301,306[b]	2,000	16,000	,000	,974
	Roy's Largest Root	37,663	301,306[b]	2,000	16,000	,000	,974
a. Design: Intercept							
b. Exact statistic							

Table 8.2. *One-way repeated measures ANOVA for the experimental group (accuracy T1 – T2 – T3[1]).*

Pairwise Comparisons						
Measure: Accuracy						
					95% Confidence Interval for Difference[b]	
(I) Time		**Mean Difference (I-J)**	**Std. Error**	**Sig.[b]**	**Lower Bound**	**Upper Bound**
1	**2**	-,481*	,023	,000	-,542	-,420
	3	-,481*	,019	,000	-,532	-,431
2	**1**	,481*	,023	,000	,420	,542
	3		,013	1,000	-,035	,035
3	**1**	,481*	,019	,000	,431	,532
	2		,013	1,000	-,035	,035
Based on estimated marginal means						
*. The mean difference is significant at the ,05 level.						
b. Adjustment for multiple comparisons: Bonferroni.						

Table 8.3. *Bonferroni pairwise comparisons for the experimental group (accuracy T1- T2 - T3).*

[1] T1: Pre-test
T2: Post-test
T3: Delayed post-test

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	Interval of the Lower	Upper			
Pair 1	Accuracy_T1 - Accuracy_T2	-,03472	,08164	,01924	-,07532	,00587	-1,804	17	,089

Table 8.4. *Paired samples t-test for the control group (accuracy T1 - T2).*

A mixed design ANOVA determined that there was a significant effect of time on the groups (F (1,34) = 295.297, p<.001, η^2= .897) but a non-significant difference between the experimental and control groups (F (1,34) = 1.395, p=.246, η^2= .039), see tables 8.5 and 8.6.

Multivariate Tests[a]							
Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Pillai's Trace	,897	295,297[b]	1,000	34,000	,000	,897
	Wilks' Lambda	,103	295,297[b]	1,000	34,000	,000	,897
	Hotelling's Trace	8,685	295,297[b]	1,000	34,000	,000	,897
	Roy's Largest Root	8,685	295,297[b]	1,000	34,000	,000	,897
Time * Group	Pillai's Trace	,867	221,158[b]	1,000	34,000	,000	,867
	Wilks' Lambda	,133	221,158[b]	1,000	34,000	,000	,867
	Hotelling's Trace	6,505	221,158[b]	1,000	34,000	,000	,867
	Roy's Largest Root	6,505	221,158[b]	1,000	34,000	,000	,867
a. Design: Intercept + Group b. Exact statistic							

Table 8.5. *Mixed design ANOVA: Tests of within-subjects effects (T1 - T2 - T3) and interaction between time and group.*

Tests of Between-Subjects Effects Measure: Accuracy Transformed Variable:						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	33,911	1	33,911	2137,528	,000	,984
Group	,022	1	,022	1,395	,246	,039
Error	,539	34	,016			

Table 8.6. *Mixed design ANOVA: Tests of between-subjects effects (experimental – control groups).*

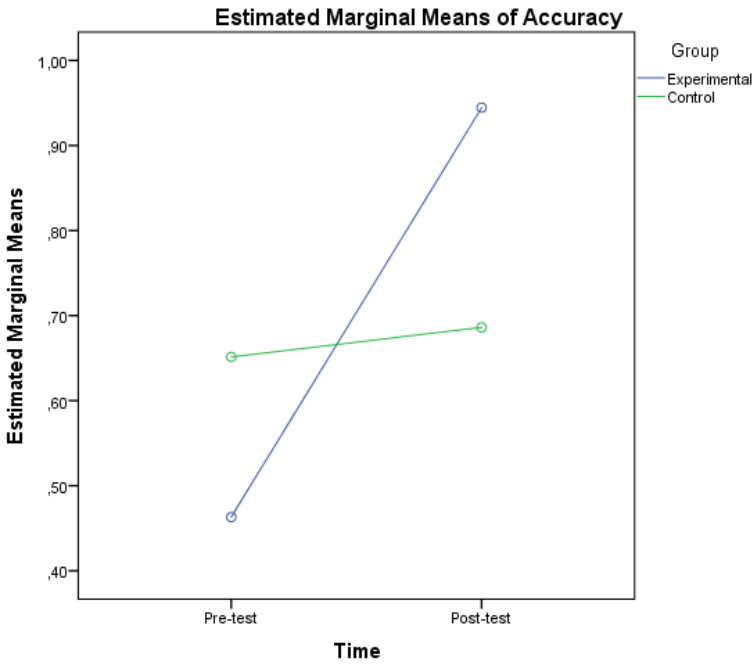


Figure 8.1. *Disordinal interaction between time (T1 – T2 – T3) and groups (experimental - control groups).*

A one-way ANCOVA was used to determine whether the different interventions (treatment vs. no treatment) were statistically significantly different having adjusted for the covariate variable. When the means were adjusted for pre-test in the experimental (*M*= .9444, *SD*=.04817) and control group (*M*= .6861, *SD*= .12164), the one-way ANOVA revealed that there was a significant effect of groups on time, after controlling for differences in the pre-test (*F* (1,33) = 110.545, *p*<.001, η^2= .770). The pairwise comparisons confirmed significant differences between experimental and control groups in terms of accuracy scores (*p*<.001). See table 8.7 and 8.8.

Tests of Between-Subjects Effects

Dependent Variable: Accuracy

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	,705[a]	2	,352	62,303	,000	,791
Intercept	,332	1	,332	58,753	,000	,640
Accuracy_T1	,104	1	,104	18,434	,000	,358
Group	,625	1	,625	110,545	,000	,770
Error	,187	33	,006			
Total	24,820	36				
Corrected Total	,892	35				

a. R Squared = ,791 (Adjusted R Squared = ,778)

Table 8.7. *One-way ANCOVA for T2 accuracy scores (experimental and control group).*

Pairwise Comparisons						
Dependent Variable: Accuracy						
(I) Group		Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
Experimental	**Control**	,356*	,034	,000	,287	,425
Control	**Experimental**	-,356*	,034	,000	-,425	-,287
Based on estimated marginal means						
*. The mean difference is significant at the ,05 level.						
b. Adjustment for multiple comparisons: Bonferroni.						

Table 8.8. *Bonferroni pairwise comparisons for T2 accuracy scores (experimental and control groups).*

Moreover, in order to exclude the possibility of ceiling effects in the control group, the one-way between-groups ANOVA showed that low proficiency participants' accuracy in the post-test differed significantly as a function of the treatment (F (1,16) = 127.92, p<.001, η^2= .88). See table 8.9.

ANOVA					
Accuracy_T2					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	,432	1	,432	127,920	,000
Within Groups	,054	16	,003		
Total	,486	17			

Table 8.9. *One-way ANOVA for low proficiency learners in the experimental and control groups.*

G.2. Discrimination test results

Before assessing learners' performance in the pre-test, post-test and delayed post-test, the by-item dataset was used to analyse the effects of item type (test vs. control) on accuracy scores and RT; in other words, to get a general impression that learners were understanding the task in the correct way by performing better in the control trials. Averages confirmed that learners were, in general, faster and more accurate in the control items than the test items in the pre-test, post-test and delayed post-test (see table 8.10). Taking into consideration that the data was substantially positively skewed (p<.001), the Kruskal-Wallis Test proved a statistically significant difference between test items and control items in terms of accuracy (H (1) = 85.533, p<.000); however, differences were not significant for response latency (H (1) = 2.705, p=.100).

Report				
Test			**Accuracy**	**RT**
Pre-test	**Test**	N	2880	1873
		Mean	,66	879,0178
		Std. Deviation	,472	320,27749
	Control	N	296	251
		Mean	,85	870,5345
		Std. Deviation	,356	367,10064
Post-test	**Test**	N	2880	2081
		Mean	,75	743,5640
		Std. Deviation	,434	239,02061
	Control	N	144	136
		Mean	,96	671,6199
		Std. Deviation	,201	246,99694
Delayed post-test	**Test**	N	1440	1112
		Mean	,79	687,0217
		Std. Deviation	,405	229,85633
	Control	N	144	135
		Mean	,94	747,9553
		Std. Deviation	,243	309,85442

Table 8.10. *Descriptives (means and SD) for accuracy and RT in test and control items (experimental and control groups).*

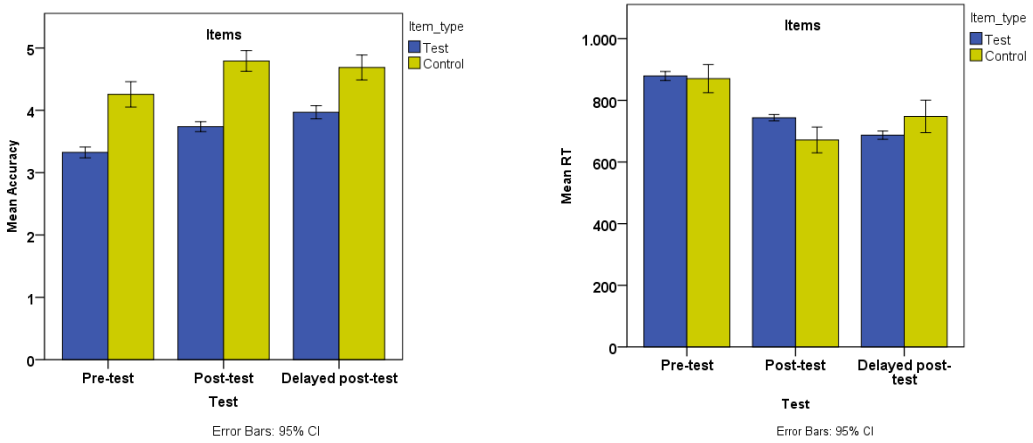


Figure 8.2. *Bar graphs for mean accuracy and RT scores in the discrimination of test vs. control items across time (experimental and control groups).*

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Accuracy is the same across categories of Item_type.	Independent-Samples Kruskal-Wallis Test	,000	Reject the null hypothesis.
2	The distribution of RT is the same across categories of Item_type.	Independent-Samples Kruskal-Wallis Test	,100	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Table 8.11. *Kruskal-Wallis Test between test and control items for accuracy and RT (experimental and control groups).*

Tests of Normality						
	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Accuracy_T1	,162	18	,200*	,967	18	,736
Accuracy_T2	,121	18	,200*	,969	18	,773
RT_T1	,230	18	,013	,803	18	,002
RT_T2	,114	18	,200*	,971	18	,811
*. This is a lower bound of the true significance. a. Lilliefors Significance Correction						

Table 8.12. *Tests of normality of distribution for accuracy and RT (pre-test and post-test control group).*

Extreme Values				
			Case Number	Value
RT_T1	**Highest**	1	30	1646,54
		2	14	1191,45
		3	4	1118,53
		4	26	1035,25
		5	34	1017,22
	Lowest	1	10	690,86
		2	24	700,08
		3	20	740,10
		4	2	745,92
		5	18	811,78

Table 8.13. *Extreme values table to recognize outliers.*

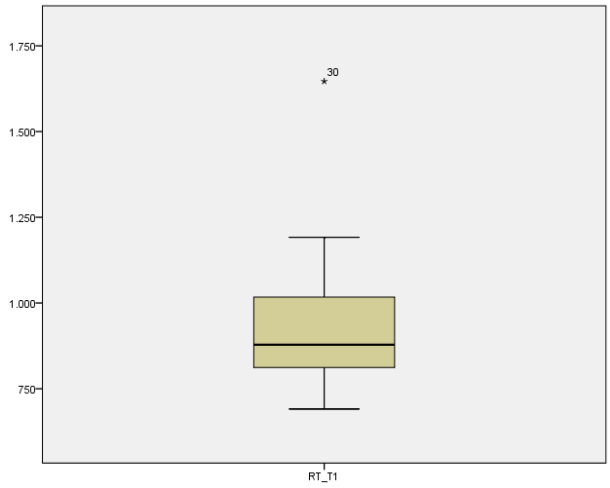


Figure 8.3. *Boxplot for normality of distribution and identification of outliers.*

Report							
Group		**Accuracy_T1**	**Accuracy_T2**	**Accuracy_T3**	**RT_T1**	**RT_T2**	**RT_T3**
Experimental	N	18	18	18	18	18	18
	Mean	,6653	,8139	,8361	817,1984	707,3938	663,1967
	Std. Deviation	,09409	,12042	,13699	194,83901	152,15643	148,50860
Control	N	17	17		17	17	
	Mean	,6684	,7176		882,7532	753,3602	
	Std. Deviation	,09592	,14163		141,01500	155,02380	

Table 8.14. *Descriptives (means and SD) for word accuracy and RT (pre-test, post-test and delayed post-test).*

A mixed design ANOVA with 'time' as the within-group factor and 'group' as the between-group factor informed that there was a statistically significant effect of *time* on the two groups (F (1,33) = 37.331, p<.001, η^2= .531) but both experimental and control groups were not statistically different, (F (1,33) = 1.753, p=.195, η^2= .050) at pre-test (M=.6653, SD=.0940 vs. M=.6684, SD=.0959, respectively) and post-test (M=.8139, SD=.1204 vs. M=.7176, SD=.1416, respectively). See table 8.15 and 8.16.

Accuracy results:

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Time	Pillai's Trace	,746	23,478[b]	2,000	16,000	,000	,746
	Wilks' Lambda	,254	23,478[b]	2,000	16,000	,000	,746
	Hotelling's Trace	2,935	23,478[b]	2,000	16,000	,000	,746
	Roy's Largest Root	2,935	23,478[b]	2,000	16,000	,000	,746
a. Design: Intercept b. Exact statistic							

Table 8.15. *One-way repeated measures ANOVA for the experimental group (accuracy T1- T2 - T3).*

Pairwise Comparisons Measure: Accuracy						
					95% Confidence Interval for Difference[b]	
(I) Time		**Mean Difference (I-J)**	**Std. Error**	**Sig.[b]**	**Lower Bound**	**Upper Bound**
1	2	-,149*	,022	,000	-,208	-,089
	3	-,171*	,027	,000	-,242	-,099
2	1	,149*	,022	,000	,089	,208
	3	-,022	,019	,809	-,074	,029
3	1	,171*	,027	,000	,099	,242
	2	,022	,019	,809	-,029	,074
Based on estimated marginal means *. The mean difference is significant at the ,05 level. b. Adjustment for multiple comparisons: Bonferroni.						

Table 8.16. *Bonferroni pairwise comparisons for the experimental group (accuracy T1- T2 - T3).*

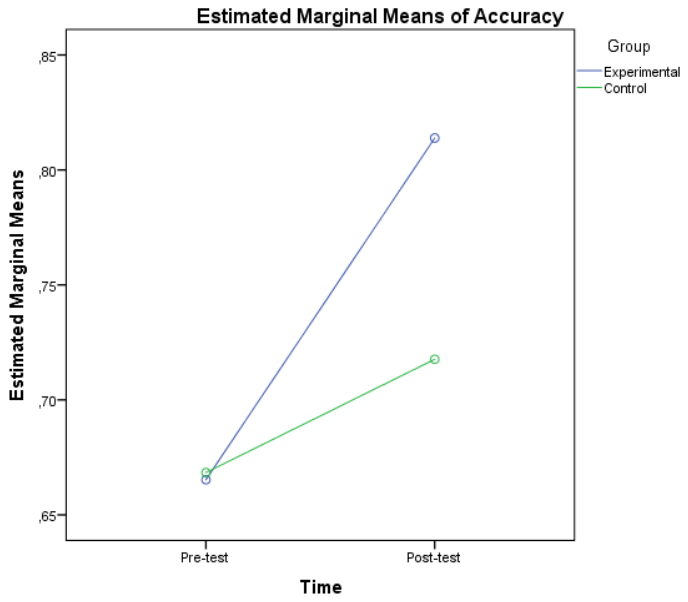


Figure 8.4. *Line graph for ordinal interaction between accuracy (T1 – T2 – T3) and groups (experimental and control groups).*

Reaction time results:

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Time	Pillai's Trace	,558	10,101[b]	2,000	16,000	,001	,558
	Wilks' Lambda	,442	10,101[b]	2,000	16,000	,001	,558
	Hotelling's Trace	1,263	10,101[b]	2,000	16,000	,001	,558
	Roy's Largest Root	1,263	10,101[b]	2,000	16,000	,001	,558
a. Design: Intercept b. Exact statistic							

Table 8.17. *One-way repeated measures ANOVA for the control group (RT T1- T2 - T3).*

Pairwise Comparisons Measure: RT						
					95% Confidence Interval for Difference[b]	
(I) Time		**Mean Difference (I-J)**	**Std. Error**	**Sig.[b]**	**Lower Bound**	**Upper Bound**
1	2	109,805*	36,800	,025	12,101	207,508
	3	154,002*	34,387	,001	62,704	245,300
2	1	-109,805*	36,800	,025	-207,508	-12,101
	3	44,197	21,570	,169	-13,072	101,466
3	1	-154,002*	34,387	,001	-245,300	-62,704
	2	-44,197	21,570	,169	-101,466	13,072
Based on estimated marginal means *. The mean difference is significant at the ,05 level. b. Adjustment for multiple comparisons: Bonferroni.						

Table 8.18. *Bonferroni pairwise comparisons for the control group (RT T1- T2 - T3).*

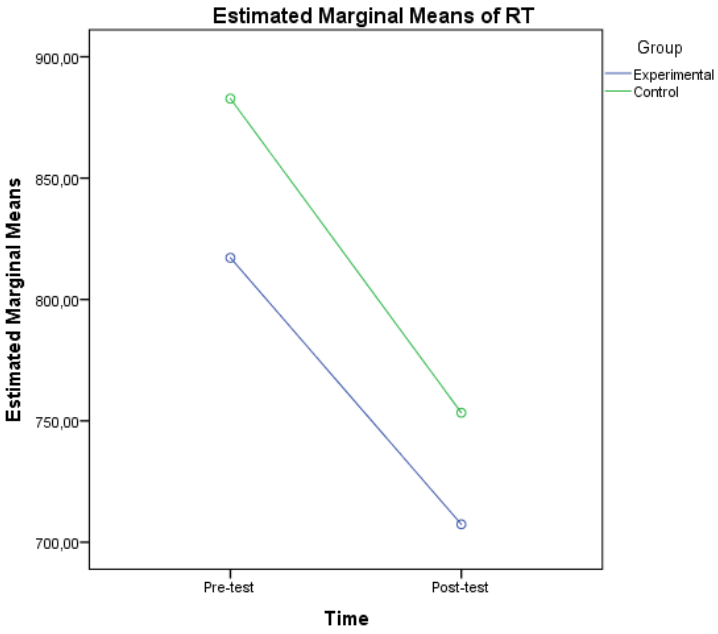


Figure 8.5. *Line graph for no interaction between RT (T1 – T2 – T3) and groups (experimental and control groups).*

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Accuracy_T1 - Accuracy_T2	-,04926	,09697	,02352	-,09912	,00059	-2,095	16	,052
Pair 2	RT_T1 - RT_T2	129,39295	112,94348	27,39282	71,32277	187,46313	4,724	16	,000

Table 8.19. *Paired samples t-test for accuracy (T1- T2) and reaction time (T1 -T2).*

Multivariate Tests[a]							
Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Time	Pillai's Trace	,531	37,331[b]	1,000	33,000	,000	,531
	Wilks' Lambda	,469	37,331[b]	1,000	33,000	,000	,531
	Hotelling's Trace	1,131	37,331[b]	1,000	33,000	,000	,531
	Roy's Largest Root	1,131	37,331[b]	1,000	33,000	,000	,531
Time * Group	Pillai's Trace	,222	9,410[b]	1,000	33,000	,004	,222
	Wilks' Lambda	,778	9,410[b]	1,000	33,000	,004	,222
	Hotelling's Trace	,285	9,410[b]	1,000	33,000	,004	,222
	Roy's Largest Root	,285	9,410[b]	1,000	33,000	,004	,222
a. Design: Intercept + Group b. Exact statistic							

Table 8.20. *Mixed-design ANOVA: Tests of within-subjects effects (T1 - T2 - T3) for the experimental and control group.*

Tests of Between-Subjects Effects						
Measure: Accuracy Transformed Variable:						
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
Intercept	35,887	1	35,887	1658,884	,000	,980
Group	,038	1	,038	1,753	,195	,050
Error	,714	33	,022			

Table 8.21. *Mixed-design ANOVA: Tests of between-subjects effects (experimental vs. control group).*

A one-way ANCOVA served to predict whether it was the treatment that differentiated the groups, after modifying the pre-test value. Having adjusted the post-test mean scores of the experimental (*M*=.815, *SE*=.023) and control (*M*=.716, *SE*= .024) groups, the one-way ANCOVA revealed that there was a statistically significant difference between experimental and control groups at time 2 (F (1,32) = 9.118, p=.005, η^2= .222), regardless of differences in time 1.

Tests of Between-Subjects Effects						
Dependent Variable: Accuracy						
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
Corrected Model	,347[a]	2	,173	18,389	,000	,535
Intercept	,012	1	,012	1,283	,266	,039
Accuracy_T1	,266	1	,266	28,189	,000	,468
Group	,086	1	,086	9,118	,005	,222
Error	,302	32	,009			
Total	21,246	35				
Corrected Total	,648	34				
a. R Squared = ,535 (Adjusted R Squared = ,506)						

Table 8.22. *One-way ANCOVA: Tests of between-subjects effects (experimental and control groups).*

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Time	Pillai's Trace	,447	26,690[b]	1,000	33,000	,000	,447
	Wilks' Lambda	,553	26,690[b]	1,000	33,000	,000	,447
	Hotelling's Trace	,809	26,690[b]	1,000	33,000	,000	,447
	Roy's Largest Root	,809	26,690[b]	1,000	33,000	,000	,447
Time * Group	Pillai's Trace	,005	,179[b]	1,000	33,000	,675	,005
	Wilks' Lambda	,995	,179[b]	1,000	33,000	,675	,005
	Hotelling's Trace	,005	,179[b]	1,000	33,000	,675	,005
	Roy's Largest Root	,005	,179[b]	1,000	33,000	,675	,005
a. Design: Intercept + Group b. Exact statistic							

Table 8.23. *Mixed-design ANOVA: Tests of within-subjects effects (T1 - T2 - T3) for the experimental and control group.*

Tests of Between-Subjects Effects						
Measure: RT Transformed Variable:						
Source	**Type III Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**	**Partial Eta Squared**
Intercept	43670834,313	1	43670834,313	1006,121	,000	,968
Group	54367,408	1	54367,408	1,253	,271	,037
Error	1432369,621	33	43405,140			

Table 8.24. *Mixed-design ANOVA: Tests of between-subjects effects (experimental vs. control group).*

Correlations			
		ID_gains_acc_pre_post	**DIS_gains_acc_pre_post**
ID_gains_acc_pre_post	Pearson Correlation	1	-,366
	Sig. (2-tailed)		,136
	N	18	18
DIS_gains_acc_pre_post	Pearson Correlation	-,366	1
	Sig. (2-tailed)	,136	
	N	18	18

Table 8.25. *Pearson correlations between identification and discrimination gains (experimental group).*

G.3. Delayed sentence repetition task results

Male and female English native speakers word productions were analysed independently because, for B1 and B2 (for /æ/ & /ʌ/), they were shown to be significantly different (t (58) = -7.627, p<.001, r=.707 and t (58) = -3.491, p=.001, r=.416; respectively). Moreover, non-word productions were also tested for gender differences in the first and second formant (Barks). The independent samples t-test showed that there were also statistically significant differences in terms of B1 (t (18) = -3.731, p=.002, r=.660); however, the differences in B2 did not became statistically significant (t (18) = -1.729, p=.101, r=.377), even if the tendency was the same (males: 11.00 vs. females: 11.65). Secondly, male and female non-native word productions were compared for Z_1 and Z_2 (/æ/ & /ʌ/). In the experimental group, the independent samples t-test reported that gender made significant differences in the two formants, t (1427) = -13.121, p<.001, r=.328. and t (1427) = -19.005, p<.001, r=.449. Female non-native speakers also had higher values than male speakers in the control group, making the differences for B1 (t (1077) = -8.750, p<.001, r=.257) and B2 (t (1077) = -11.552, p<.001, r=.332) statistically significant. Concerning the experimental group productions of non-words, males and females behaved significantly different for B1 (t (677) = -8.142, p<.001, r=.298) and B2 (t (677) = -10.560, p<.001, r=.376). In relation to the control group, the independent samples t-

test also reported differences between the two normalized formants: *t* (338) = -6.918, *p*<.001, *r*=.352 and *t* (338) = -7.145, *p*<.001, *r*=.362.

In terms of height, the two male speakers had a Bark measure of 7.24 (807.91 Hz) for vowel /æ/ and 5.57 (579.74 Hz) for vowel /ʌ/. In contrast, the two female speakers produced a Bark measure of 8.71 (1036.32 Hz) for vowel /æ/ and 7.53 (846.23 Hz) for vowel /ʌ/. In relation to advancement, the two male native speakers had a Bark measure of 11.63 (1631.79) for vowel /æ/ and 10.02 (1280.57 Hz) for vowel /ʌ/ but the two female native speakers produced 12.06 (1743.33 Hz) for vowel /æ/ and 11.09 (1503.46 Hz) for vowel /ʌ/. Finally, the Euclidean distance between male speakers and female speakers was 1.55 barks vs. 3.03; respectively). See figure 2.

The same procedure was carried out with native speakers' productions of non-words. In terms of height, the two male speakers had a Bark measure of 6.31 (674.49 Hz) for vowel /æ/ and 5.43 (563.18 Hz) for vowel /ʌ/. In contrast, the two female speakers produced a Bark measure of 7.52 (853.41 Hz) for vowel /æ/ and 7.24 (803.31 Hz) for /ʌ/. In relation to advancement, the two male native speakers had a Bark measure of 11.61 (1629.38) for vowel /æ/ and 10.39 (1353.15 Hz) for vowel /ʌ/. Nevertheless, the two female native speakers produced a Bark measure of 12.07 (1744.80 Hz) for vowel /æ/ and 11.23 (1537.20 Hz) for vowel /ʌ/. Finally, the Euclidean distance got reduced for non-words: 1.83 barks vs. 6.71; respectively), see figure 3.

Report										
Speakers						F1_mean	F2_mean	B1_mean	B2_mean	Euclidean_ distance
NSs		Male	Pre-test	æ	N	1	1	1	1	1
					Mean	793,8950	1606,2480	7,1862	11,5289	2,4546
					Std. Deviation					
				ʌ	N	1	1	1	1	
					Mean	557,8020	1281,8620	5,3870	10,0452	
					Std. Deviation					
			Post-test	æ	N	2	2	2	2	2
					Mean	814,9308	1644,5675	7,2792	11,6881	1,1063
					Std. Deviation	172,04733	25,98264	1,14342	,08603	,24465
				ʌ	N	2	2	2	2	
					Mean	590,7133	1279,9317	5,6666	10,0201	
					Std. Deviation	17,97937	69,27054	,12941	,29396	
		Female	Pre-test	æ	N	1	1	1	1	1
					Mean	1089,5370	1741,4430	9,0399	12,0298	1,8635
					Std. Deviation					
				ʌ	N	1	1	1	1	
					Mean	915,1660	1576,2640	7,9994	11,3925	
					Std. Deviation					
			Post-test	æ	N	2	2	2	2	2
					Mean	1009,7233	1744,2800	8,5545	12,0850	3,6262
					Std. Deviation	173,32130	91,64104	1,03476	,35801	1,22479
				ʌ	N	2	2	2	2	
					Mean	811,7642	1467,0617	7,3034	10,9394	
					Std. Deviation	117,40683	35,85267	,80385	,16857	
NNs	Experimental	Male	Pre-test	æ	N	18	18	18	18	18
					Mean	682,5588	1396,9277	6,3654	10,5660	1,4518
					Std. Deviation	57,40713	80,72300	,42330	,34689	,65291
				ʌ	N	18	18	18	18	
					Mean	615,6292	1286,6580	5,8314	10,0370	
					Std. Deviation	59,36137	95,08398	,47036	,44606	
			Post-test	æ	N	36	36	36	36	36
					Mean	714,4030	1364,7924	6,6023	10,4193	1,4765
					Std. Deviation	86,71585	137,80757	,63413	,63357	,64761
				ʌ	N	36	36	36	36	
					Mean	614,5259	1223,1461	5,8187	9,7151	
					Std. Deviation	108,25027	135,73047	,85675	,65221	
			Delayed post-test	æ	N	36	36	36	36	36
					Mean	688,9513	1404,8142	6,4065	10,5971	1,6745
					Std. Deviation	81,01741	204,98363	,62509	,82754	,94883
				ʌ	N	36	36	36	36	
					Mean	592,0532	1261,5917	5,6504	9,8949	
					Std. Deviation	72,84825	150,17283	,58986	,72193	
		Female	Pre-test	æ	N	18	18	18	18	18
					Mean	782,6307	1619,3346	7,0847	11,5591	1,4528
					Std. Deviation	59,90445	86,07499	,43189	,36132	,45781
				ʌ	N	18	18	18	18	
					Mean	705,7919	1503,0096	6,5123	11,0565	
					Std. Deviation	79,57437	77,45656	,60039	,34382	
			Post-test	æ	N	36	36	36	36	36
					Mean	817,4478	1586,4678	7,3005	11,4326	1,5404
					Std. Deviation	112,50497	108,33832	,82653	,45489	,66178
				ʌ	N	36	36	36	36	
					Mean	707,2881	1426,4914	6,5378	10,7180	
					Std. Deviation	86,86424	120,34497	,64275	,55054	
			Delayed post-test	æ	N	36	36	36	36	36
					Mean	829,6701	1583,8891	7,4041	11,4114	2,0583
					Std. Deviation	93,63076	111,08810	,66634	,48429	1,03089
				ʌ	N	36	36	36	36	
					Mean	647,0954	1380,7436	6,0560	10,5066	
					Std. Deviation	117,64000	134,98153	,94162	,63261	
	Control	Male	Pre-test	æ	N	14	14	14	14	14
					Mean	683,5129	1404,1419	6,3798	10,6080	1,3770
					Std. Deviation	49,61208	95,57738	,37053	,41545	,35041
				ʌ	N	14	14	14	14	
					Mean	633,6958	1384,0855	5,9846	10,4733	
					Std. Deviation	45,04980	111,74775	,35318	,47398	
			Post-test	æ	N	28	28	28	28	28
					Mean	651,1308	1422,4714	6,1207	10,6679	1,2485
					Std. Deviation	89,04257	183,16576	,69312	,76070	,71604
				ʌ	N	28	28	28	28	
					Mean	622,1646	1354,7585	5,8781	10,3567	
					Std. Deviation	81,39329	169,98479	,64441	,70295	
		Female	Pre-test	æ	N	22	22	22	22	22
					Mean	762,8273	1628,2264	6,9299	11,5874	1,6642
					Std. Deviation	100,64594	98,95943	,73134	,39531	,41849
				ʌ	N	22	22	22	22	
					Mean	697,7480	1491,3878	6,4615	10,9772	
					Std. Deviation	67,37605	78,82911	,50561	,34432	
			Post-test	æ	N	44	44	44	44	44
					Mean	769,7992	1594,7657	6,9893	11,4429	1,6533
					Std. Deviation	102,47607	130,79411	,73031	,55822	,89843
				ʌ	N	44	44	44	44	
					Mean	659,2492	1496,5539	6,1462	11,0026	
					Std. Deviation	111,64722	176,95092	,88305	,73621	

Table 8.26. *Descriptives (means &SD) for NSs/NNSs productions of vowels /æ/ & /ʌ/ (T1 -T2 - T3) in words.*

Native speakers' productions of vowels /æ/ and /ʌ/ were compared for B1 and B2 values. The one-way ANOVA revealed that male native speakers make a significant distinction between /æ/ & /ʌ/ for B1 (F (1,4) = 12.178, p=.025, η^2 =.752) and B2 (F (1,4) = 139.366, p<.001, η^2 =.971). The female English speakers did not show statistically significant differences for B1 (F (1,4) = 3.809, p=.123, η^2 =.487) but they did for B2 (F (1,4) = 19.349, p=.012, η^2 =.828). To sum up, native speakers showed significant differences for both B1 (F (1,10) = 5.860, p=.042, η^2 =.327) and B2 (F (1,10) = 21.040, p=.001, η^2 =.677) in terms of the two target vowels, see table 8.27 and 8.28.

ANOVA							
Gender			**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Male	**B1_mean**	Between Groups	4,208	1	4,208	12,178	,025
		Within Groups	1,382	4	,346		
		Total	5,590	5			
	B2_mean	Between Groups	3,871	1	3,871	139,366	,000
		Within Groups	,111	4	,028		
		Total	3,983	5			
Female	**B1_mean**	Between Groups	2,092	1	2,092	3,809	,123
		Within Groups	2,197	4	,549		
		Total	4,289	5			
	B2_mean	Between Groups	1,429	1	1,429	19,349	,012
		Within Groups	,295	4	,074		
		Total	1,725	5			

Table 8.27. *One-way between-subjects ANOVA for vowel differences (/æ/ & /ʌ/) in B1 and B2 by gender (NSs).*

ANOVA						
		Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
B1_mean	Between Groups	6,116	1	6,116	5,860	,042
	Within Groups	12,586	10	1,259		
	Total	18,702	11			
B2_mean	Between Groups	5,003	1	5,003	21,040	,001
	Within Groups	2,378	10	,238		
	Total	7,381	11			

Table 8.28. *One-way between-subjects ANOVA for vowel differences (/æ/ & /ʌ/) in B1 and B2 (NSs).*

ANOVA								
Gender	Test			Sum of Squares	df	Mean Square	F	Sig.
Male	Pre-test (T1)	B1_mean	Between Groups	1,283	1	1,283	6,506	,021
			Within Groups	3,156	16	,197		
			Total	4,440	17			
		B2_mean	Between Groups	1,252	1	1,252	8,086	,012
			Within Groups	2,477	16	,155		
			Total	3,729	17			
	Post-test (T2)	B1_mean	Between Groups	2,838	1	2,838	14,248	,002
			Within Groups	3,188	16	,199		
			Total	6,026	17			
		B2_mean	Between Groups	2,091	1	2,091	11,664	,004
			Within Groups	2,869	16	,179		
			Total	4,960	17			
	Delayed post-test (T3)	B1_mean	Between Groups	2,841	1	2,841	17,556	,001
			Within Groups	2,589	16	,162		
			Total	5,430	17			
		B2_mean	Between Groups	2,048	1	2,048	14,864	,001
			Within Groups	2,204	16	,138		
			Total	4,252	17			
Female	Pre-test (T1)	B1_mean	Between Groups	1,474	1	1,474	5,776	,029
			Within Groups	4,084	16	,255		
			Total	5,558	17			
		B2_mean	Between Groups	1,137	1	1,137	9,748	,007
			Within Groups	1,866	16	,117		
			Total	3,003	17			
	Post-test (T2)	B1_mean	Between Groups	2,998	1	2,998	8,655	,010
			Within Groups	5,542	16	,346		
			Total	8,539	17			
		B2_mean	Between Groups	2,233	1	2,233	17,289	,001
			Within Groups	2,066	16	,129		
			Total	4,299	17			
	Delayed post-test (T3)	B1_mean	Between Groups	7,926	1	7,926	27,805	,000
			Within Groups	4,561	16	,285		
			Total	12,487	17			
		B2_mean	Between Groups	3,397	1	3,397	22,771	,000
			Within Groups	2,387	16	,149		
			Total	5,783	17			

Table 8.29. *One-way between-subjects ANOVA for vowel differences (/æ/ & /ʌ/) in B1 and B2 (NNSs).*

Multivariate Tests[a]									
Gender				Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Male	Experimental	Time	Pillai's Trace	,361	1,977[b]	2,000	7,000	,209	,361
			Wilks' Lambda	,639	1,977[b]	2,000	7,000	,209	,361
			Hotelling's Trace	,565	1,977[b]	2,000	7,000	,209	,361
			Roy's Largest Root	,565	1,977[b]	2,000	7,000	,209	,361
Female	Experimental	Time	Pillai's Trace	,721	9,033[b]	2,000	7,000	,012	,721
			Wilks' Lambda	,279	9,033[b]	2,000	7,000	,012	,721
			Hotelling's Trace	2,581	9,033[b]	2,000	7,000	,012	,721
			Roy's Largest Root	2,581	9,033[b]	2,000	7,000	,012	,721
a. Design: Intercept									
b. Exact statistic									

Table 8.30. *One way repeated-measures ANOVA for the experimental group (T1 – T2 – T3) by gender.*

Pairwise Comparisons								
Measure: Euclidean Distance								
Gender				Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
							Lower Bound	Upper Bound
Male	Experimental	1	2	,006	,214	1,000	-,640	,652
			3	-,178	,157	,866	-,651	,294
		2	1	-,006	,214	1,000	-,652	,640
			3	-,184	,119	,483	-,543	,175
		3	1	,178	,157	,866	-,294	,651
			2	,184	,119	,483	-,175	,543
Female	Experimental	1	2	-,073	,177	1,000	-,606	,461
			3	-,556*	,182	,047	-1,104	-,009
		2	1	,073	,177	1,000	-,461	,606
			3	-,484*	,114	,009	-,829	-,138
		3	1	,556*	,182	,047	,009	1,104
			2	,484*	,114	,009	,138	,829
Based on estimated marginal means								
*. The mean difference is significant at the ,05 level.								
a. Adjustment for multiple comparisons: Bonferroni.								

Table 8.31. *Bonferroni pairwise comparisons for the experimental group (T1 – T2 – T3) by gender.*

Multivariate Tests[a]								
Gender			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Male	**Time**	Pillai's Trace	,311	2,713[b]	1,000	6,000	,151	,311
		Wilks' Lambda	,689	2,713[b]	1,000	6,000	,151	,311
		Hotelling's Trace	,452	2,713[b]	1,000	6,000	,151	,311
		Roy's Largest Root	,452	2,713[b]	1,000	6,000	,151	,311
Female	**Time**	Pillai's Trace	,008	,081[b]	1,000	10,000	,782	,008
		Wilks' Lambda	,992	,081[b]	1,000	10,000	,782	,008
		Hotelling's Trace	,008	,081[b]	1,000	10,000	,782	,008
		Roy's Largest Root	,008	,081[b]	1,000	10,000	,782	,008
a. Design: Intercept								
b. Exact statistic								

Table 8.32. *One way repeated-measures ANOVA for the control group (T1 – T2 – T3) by gender.*

Pairwise Comparisons							
Measure: Euclidean Distance							
Gender			Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
						Lower Bound	Upper Bound
Male	1	2	,157	,095	,151	-,076	,390
	2	1	-,157	,095	,151	-,390	,076
Female	1	2	,054	,190	,782	-,368	,476
	2	1	-,054	,190	,782	-,476	,368
Based on estimated marginal means							
a. Adjustment for multiple comparisons: Bonferroni.							

Table 8.33. *Bonferroni pairwise comparisons for the control group (T1 – T2 – T3) by gender.*

Multivariate Tests[a]								
Group			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Experimental	Time	Pillai's Trace	,599	11,194[b]	2,000	15,000	,001	,599
		Wilks' Lambda	,401	11,194[b]	2,000	15,000	,001	,599
		Hotelling's Trace	1,493	11,194[b]	2,000	15,000	,001	,599
		Roy's Largest Root	1,493	11,194[b]	2,000	15,000	,001	,599
	Time * Gender	Pillai's Trace	,613	11,872[b]	2,000	15,000	,001	,613
		Wilks' Lambda	,387	11,872[b]	2,000	15,000	,001	,613
		Hotelling's Trace	1,583	11,872[b]	2,000	15,000	,001	,613
		Roy's Largest Root	1,583	11,872[b]	2,000	15,000	,001	,613
a. Design: Intercept + Gender b. Exact statistic								

Table 8.34. *Mixed design ANOVA: Tests of within-subjects effects for males and females (experimental group).*

Pairwise Comparisons							
Measure:							
			Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
Group						Lower Bound	Upper Bound
Experimental	1	2	,428*	,092	,001	,184	,673
		3	,383*	,084	,001	,159	,607
	2	1	-,428*	,092	,001	-,673	-,184
		3	-,045	,057	1,000	-,197	,107
	3	1	-,383*	,084	,001	-,607	-,159
		2	,045	,057	1,000	-,107	,197
Based on estimated marginal means *. The mean difference is significant at the ,05 level. b. Adjustment for multiple comparisons: Bonferroni.							

Table 8.35. *Bonferroni pairwise comparisons for /æ/ B1 value in time (T1 – T2 – T3).*

A mixed design ANOVA revealed that learners approached native speakers' B1 values for /æ/ significantly from time 1 to time 3 (F (2,15) = 11.194, p=.001, η^2 =.599) and, specifically, from time 1 to time 2 and time 3 (p<.05), as shown in the pairwise comparisons (see table 8.34 and 8.35). The between-groups test revealed that there was a statistically significant difference between males and females, as previously observed, F (1,16) = 8.529, p=0.10, η^2 =.348.

Tests of Between-Subjects Effects							
Measure: B1_A Transformed Variable:							
Group		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Experimental	Intercept	67,421	1	67,421	95,621	,000	,857
	Gender	6,014	1	6,014	8,529	,010	,348
	Error	11,281	16	,705			

Table 8.36. *Mixed design ANOVA: Tests of between-subjects effects for males and females for /æ/ B1 value (experimental group).*

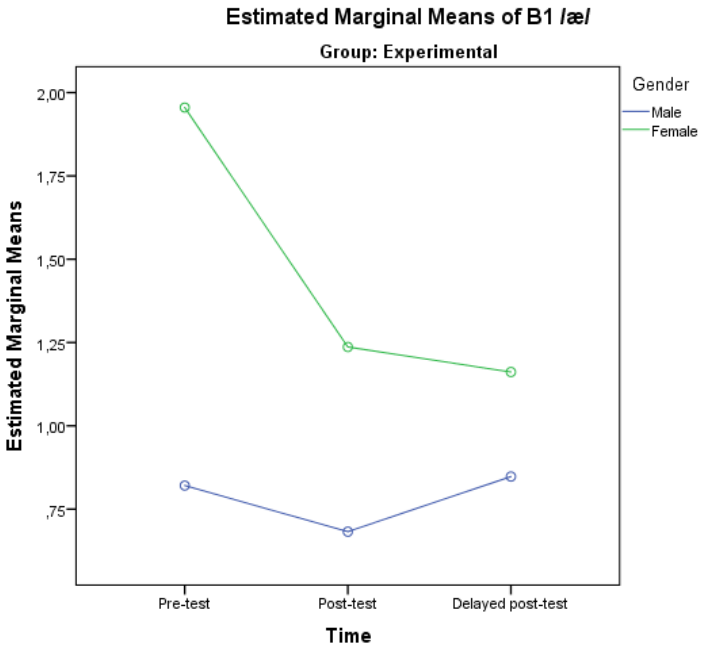


Figure 8.6. *Ordinal interaction between time and gender for /æ/ B1 value.*

A mixed-design ANOVA confirmed that learners distance between time 1 and time 3 got significantly reduced (F (2,15) = 8.139, p=.004, η^2 =.520), especially from time 2 to time 3 (p<.05) and showed differences between males and females (F (1,16) = 58.020, p<.001, η^2 =.784).

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Time	Pillai's Trace	,520	8,139[b]	2,000	15,000	,004	,520
	Wilks' Lambda	,480	8,139[b]	2,000	15,000	,004	,520
	Hotelling's Trace	1,085	8,139[b]	2,000	15,000	,004	,520
	Roy's Largest Root	1,085	8,139[b]	2,000	15,000	,004	,520
Time * Gender	Pillai's Trace	,607	11,566[b]	2,000	15,000	,001	,607
	Wilks' Lambda	,393	11,566[b]	2,000	15,000	,001	,607
	Hotelling's Trace	1,542	11,566[b]	2,000	15,000	,001	,607
	Roy's Largest Root	1,542	11,566[b]	2,000	15,000	,001	,607
a. Design: Intercept + Gender b. Exact statistic							

Table 8.37. *Mixed design ANOVA: Tests of within-subjects effects for males and females for /ʌ/ B1 value (experimental group).*

Pairwise Comparisons						
Measure: B1_A						
(I) Time		Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	,189	,104	,269	-,090	,468
	3	-,112	,127	1,000	-,451	,226
2	1	-,189	,104	,269	-,468	,090
	3	-,301*	,078	,004	-,510	-,092
3	1	,112	,127	1,000	-,226	,451
	2	,301*	,078	,004	,092	,510
Based on estimated marginal means *. The mean difference is significant at the ,05 level. b. Adjustment for multiple comparisons: Bonferroni.						

Table 8.38. *Bonferroni pairwise comparisons for /ʌ/ B1 value in time (T1 – T2 – T3).*

Tests of Between-Subjects Effects						
Measure: B1_A Transformed Variable:						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	13,277	1	13,277	30,888	,000	,659
Gender	24,939	1	24,939	58,020	,000	,784
Error	6,877	16	,430			

Table 8.39. *Mixed design ANOVA: Tests of between-subjects effects for males and females for /ʌ/ B1 value (experimental group).*

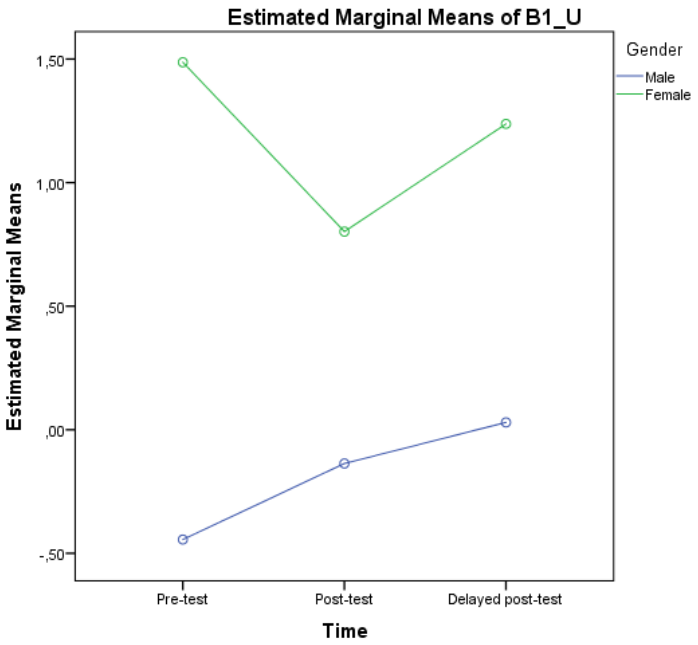


Figure 8.7. *Ordinal interaction between time and gender for /ʌ/ B1 value.*

G.4. Generalization results for perception and production

Identification

Given that the data was normally distributed for voices in the pre- and post-test gains (p>.05), a paired samples t-test revealed that there were no significant differences between identifying words in the male voice or female voice (M= .4361, SD= .1105 vs. M=.4597, SD= .1358, respectively; t (17) = -.541, p=.596, r=.130).

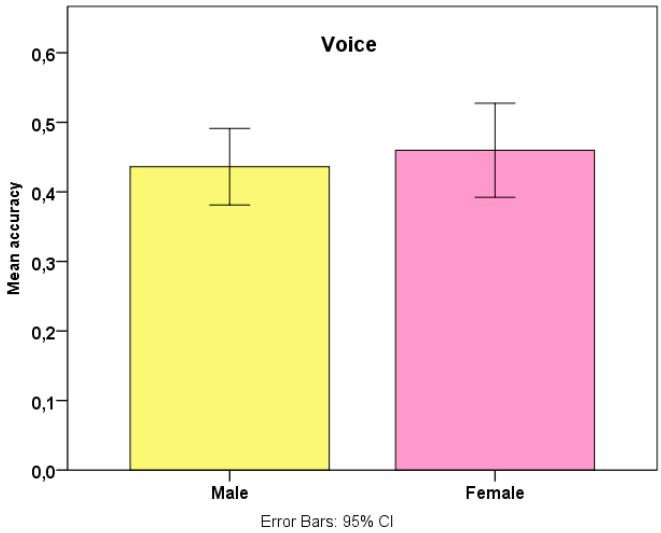


Figure 8.8. *Bar graph for accuracy gains in the identification of words uttered by male and female voices.*

Concerning accuracy gains in the identification of vowels /æ/ (M= .4361, SD= .1022) and /ʌ/ (M=.4597, SD= .1228) in the pre- and post-test, a parametric paired samples t-test failed to show any significant differences, t (17) = -.647, p=.527, r=.155).

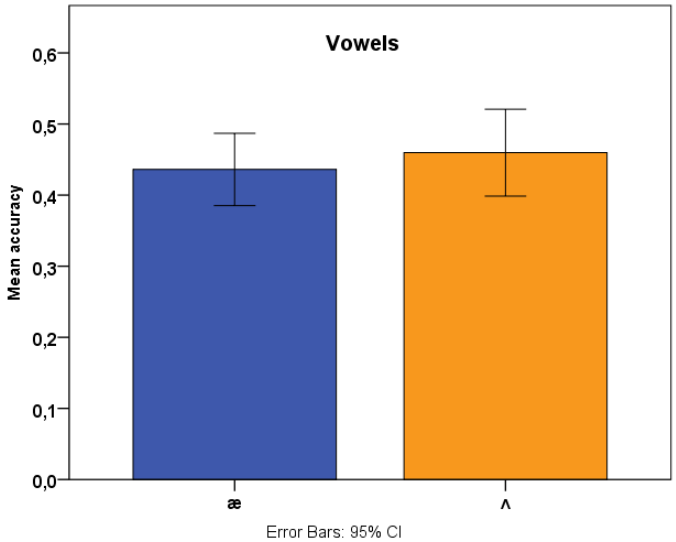


Figure 8.9. *Bar graph for accuracy gains in the identification of words containing /æ/ and /ʌ/.*

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval Lower	95% Confidence Interval Upper			
Pair 1	Gains_male - Gains_female	-,02361	,18520	,04365	-,11571	,06849	-,541	17	,596
Pair 2	Gains_æ - Gains_ʌ	-,02361	,15494	,03652	-,10066	,05344	-,647	17	,527
Pair 3	Accuracy_T2_old_speakers - Accuracy_T2_new_speakers	-,03056	,03694	,00871	-,04892	-,01219	-3,510	17	,003
Pair 4	Accuracy_T1_old_speakers - Accuracy_T2_new_speakers	-,46319	,08049	,01897	-,50322	-,42317	-24,414	17	,000
Pair 5	Accuracy_T2_word - Accuracy_T2_non-word	,06667	,05423	,01278	,03970	,09364	5,215	17	,000
Pair 6	Accuracy_T1_words - Accuracy_T2_nonwords	-,41458	,07380	,01739	-,45128	-,37789	-23,835	17	,000

Table 8.40. *Summary paired samples t-test between different independent variables of the identification test (experimental group).*

Discrimination

The Shapiro-Wilk test of normality showed that gains in accuracy and RT were normality distributed for male and female voices, $p>.05$; therefore, a paired samples t-test was run to check for differences between the Male-Male-Female sequence ($M= .1431$, $SD= .1100$) and the Female-Female-Male sequence ($M= .0861$, $SD= .0896$) in accuracy scores. In addition, RT mean scores for the M-M-F sequence ($M= -111.85$, $SD= 168.87$) and the F-F-M sequence ($M= -76.91$, $SD= 158.08$) were calculated. Despite differences in accuracy may be understood from figure 8.44, participants did not behave significantly different when the sequence was M-M-F or F-F-M in terms of accuracy t (17) = 1.827, p=.085, r=.405) and reaction time t (17) = -1.297, p=.212, r=.300), see table 8.41.

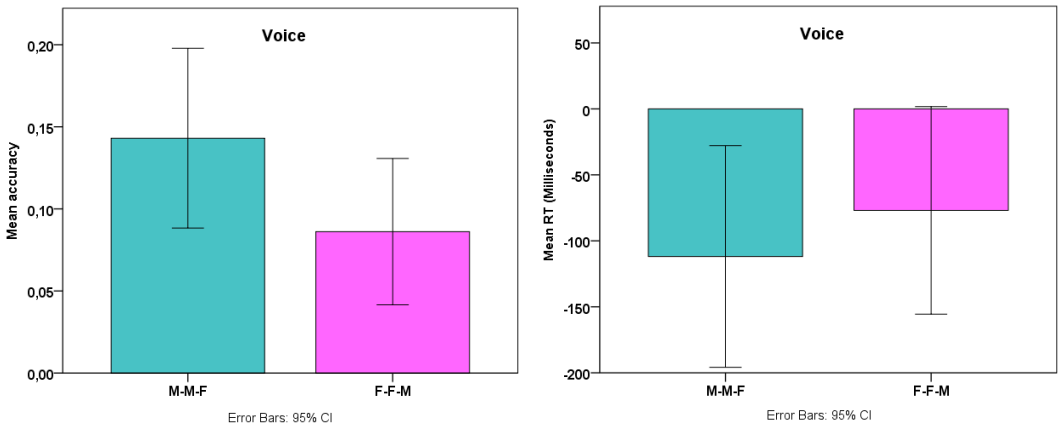


Figure 8.10. *Bar graphs for accuracy and RT gains in the discrimination of MMF and FFM trials.*

Apart from male and female voices, the ABX discrimination test included the following four sequences: ABB, ABA, BAA, BAB. Given that the data was normally distributed

for the four orders (p>.05), a one way repeated-measures ANOVA was run to see whether there were significant changes for gains in accuracy and RT.

Concerning accuracy, learners did not perform differently as a function of sequencing, (F (3,15) = 2.132, p=.139, η^2= .299) and, as observed in the pairwise comparisons, there was not an advantage of one sequence over the others, p>.05 (see table 2.20 and table 2.21). Similarly, response latencies did not vary depending on the order of the target items (F (3,15) =1.815, =.188, η^2= .266) and the pairwise comparisons indicated no changes among ABB, ABA, BAA, BAB sequences, p>.05 (see tables 8.41, 8.42, 8.43 and 8.44).

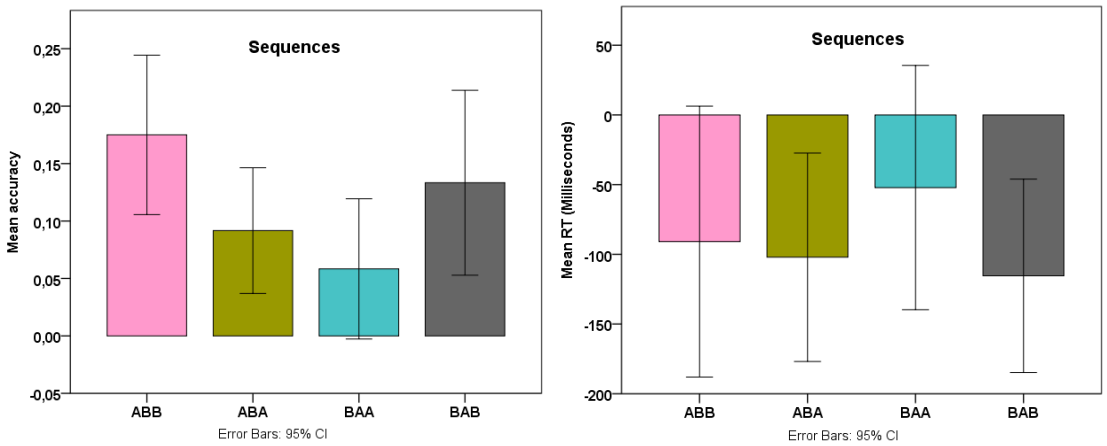


Figure 8.11. *Bar graphs for accuracy and RT gains in the discrimination of ABB, ABA, BAA, BAB trials.*

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	the Difference Lower	the Difference Upper			
Pair 1	Gains_acc_MMF - Gains_acc_FFM	,05694	,13223	,03117	-,00881	,12270	1,827	17	,085
Pair 2	Gains_rt_MMF - Gains_rt_FFM	-34,93648	114,27698	26,93534	-91,76509	21,89212	-1,297	17	,212
Pair 3	Accuracy_T2_old_speakers - Accuracy_T2_new_speakers	,04537	,13453	,03171	-,02153	,11227	1,431	17	,171
Pair 4	RT_T2_old_speakers - RT_T2_new_speakers	-20,42555	68,73051	16,19994	-54,60443	13,75333	-1,261	17	,224
Pair 5	Accuracy_T1_old_speakers - Accuracy_T2_new_speakers	-,10324	,08433	,01988	-,14518	-,06131	-5,194	17	,000
Pair 6	RT_mean_T1_old_speakers - RT_mean_T2_new_speakers	88,51880	145,08686	34,19730	16,36880	160,66880	2,588	17	,019
Pair 7	Accuracy_T2_word - Accuracy_T2_nonword	,06806	,12060	,02843	,00808	,12803	2,394	17	,028
Pair 8	RT_T2_word - RT_T2_nonword	-33,51355	41,24226	9,72089	-54,02284	-13,00426	-3,448	17	,003
Pair 9	Accuracy_T1_words - Accuracy_T2_nonwords	-,08056	,09863	,02325	-,12960	-,03151	-3,465	17	,003
Pair 10	RT_T1_words - RT_T2_nonwords	76,29100	150,49235	35,47139	1,45292	151,12909	2,151	17	,046

Table 8.41. *Summary paired samples t-test between different independent variables of the discrimination test (experimental group).*

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Sequence	Pillai's Trace	,299	2,132[b]	3,000	15,000	,139	,299
	Wilks' Lambda	,701	2,132[b]	3,000	15,000	,139	,299
	Hotelling's Trace	,426	2,132[b]	3,000	15,000	,139	,299
	Roy's Largest Root	,426	2,132[b]	3,000	15,000	,139	,299
a. Design: Intercept b. Exact statistic							

Table 8.42. *One-way repeated-measures ANOVA for discrimination accuracy scores (ABB, ABA, BAA, BAB).*

Pairwise Comparisons Measure: Accuracy						
					95% Confidence Interval for Difference[a]	
(I) Sequence		**Mean Difference (I-J)**	**Std. Error**	**Sig.[a]**	**Lower Bound**	**Upper Bound**
1	2	,083	,036	,210	-,025	,192
	3	,117	,046	,124	-,020	,253
	4	,042	,044	1,000	-,090	,173
2	1	-,083	,036	,210	-,192	,025
	3	,033	,035	1,000	-,072	,138
	4	-,042	,048	1,000	-,186	,103
3	1	-,117	,046	,124	-,253	,020
	2	-,033	,035	1,000	-,138	,072
	4	-,075	,047	,758	-,214	,064
4	1	-,042	,044	1,000	-,173	,090
	2	,042	,048	1,000	-,103	,186
	3	,075	,047	,758	-,064	,214
Based on estimated marginal means a. Adjustment for multiple comparisons: Bonferroni.						

Table 8.43. *Bonferroni pairwise comparisons for discrimination accuracy scores (ABB, ABA, BAA, BAB).*

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Sequence	Pillai's Trace	,266	1,815[b]	3,000	15,000	,188	,266
	Wilks' Lambda	,734	1,815[b]	3,000	15,000	,188	,266
	Hotelling's Trace	,363	1,815[b]	3,000	15,000	,188	,266
	Roy's Largest Root	,363	1,815[b]	3,000	15,000	,188	,266
a. Design: Intercept b. Exact statistic							

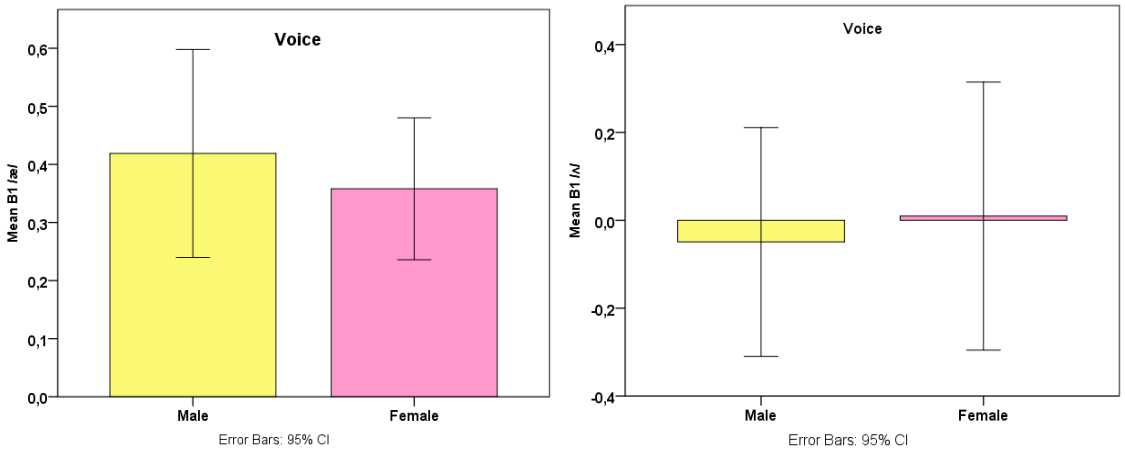
Table 8.44. *One-way repeated measures ANOVA for discrimination RT scores (ABB, ABA, BAA, BAB).*

Pairwise Comparisons						
Measure: RT						
(I) Sequence		Mean Difference (I-J)	Std. Error	Sig.[a]	95% Confidence Interval for Difference[a]	
					Lower Bound	Upper Bound
1	2	11,228	22,507	1,000	-55,933	78,389
	3	-38,730	17,703	,258	-91,556	14,097
	4	24,572	32,066	1,000	-71,114	120,258
2	1	-11,228	22,507	1,000	-78,389	55,933
	3	-49,957	27,424	,517	-131,793	31,878
	4	13,344	22,137	1,000	-52,715	79,403
3	1	38,730	17,703	,258	-14,097	91,556
	2	49,957	27,424	,517	-31,878	131,793
	4	63,301	33,873	,474	-37,777	164,380
4	1	-24,572	32,066	1,000	-120,258	71,114
	2	-13,344	22,137	1,000	-79,403	52,715
	3	-63,301	33,873	,474	-164,380	37,777
Based on estimated marginal means a. Adjustment for multiple comparisons: Bonferroni.						

Table 8.45. *Bonferroni pairwise comparisons for discrimination RT scores (ABB, ABA, BAA, BAB).*

Delayed sentence repetition task

Given that the data was normally distributed (p>.05), three paired samples t-tests were run to see whether being exposed to a male or female voice played a role in learners' performance of the delayed-sentence repetition task. Concerning /æ/, learners in the experimental group performed equally when the voice was a male or a female in terms of height (*t* (17) = .549, *p*=.590, *r*=.131) and advancement (*t* (17) = .141, *p*=.890, *r*=.034). As for /ʌ/, voice did not have any implications in the results of the delayed sentence repetition task. No distinction was shown either for B1 values (*t* (17) =-.334, *p*=.742, *r*=.080) or B2 values (*t* (17) = .581, *p*=.569, *r*=.034). Finally, the Euclidean distance did not change as a result of the voice they were exposed to (*t* (17) =-1.477, *p*=.158, *r*=.337). See table 8.46 and figure 8.12.



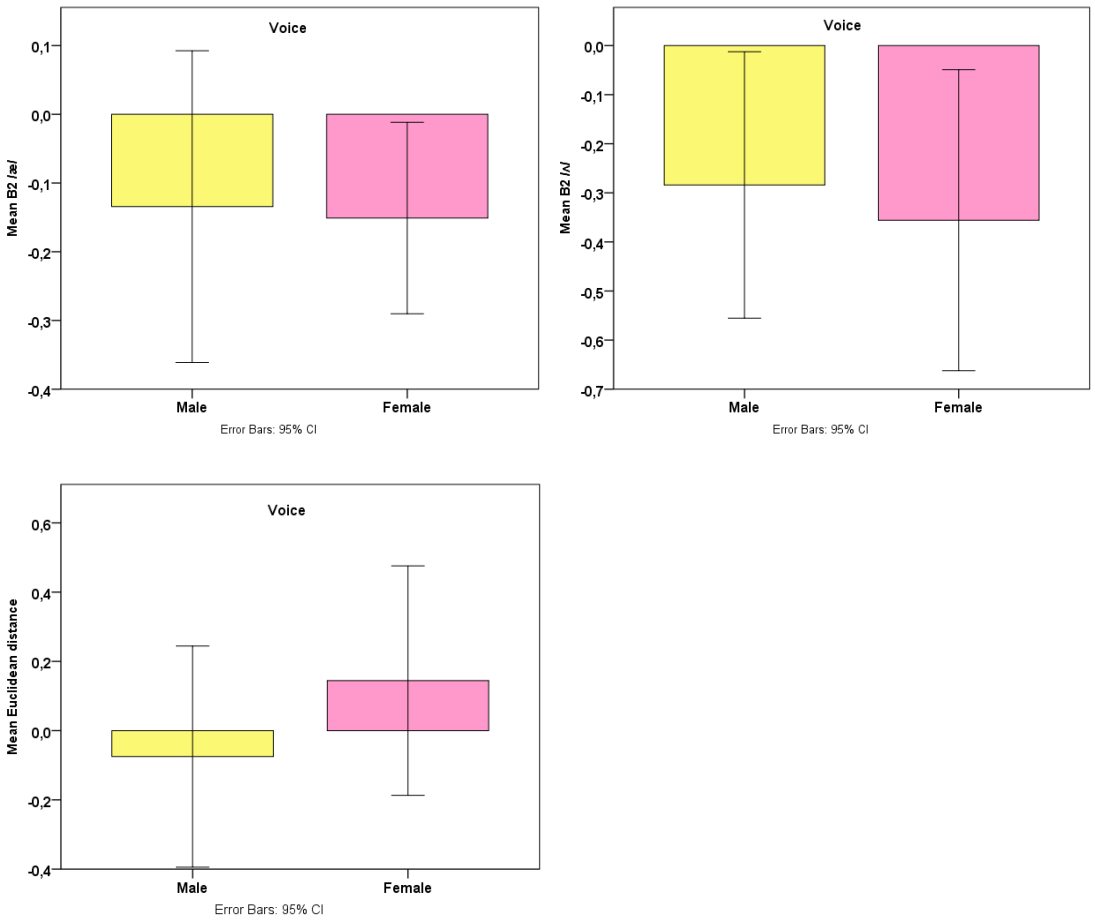


Figure 8.12. *Bar graphs for B1, B2 and E.D. gains in the production of sentences uttered by male and female voices.*

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	Interval of the Lower	Upper			
Pair 1	B1_Gains_A_Male - B1_Gains_A_Female	,06088	,45688	,11081	-,17403	,29579	,549	16	,590
Pair 2	B1_Gains_U_Male - B1_Gains_U_Female	-,05891	,74787	,17627	-,43081	,31300	-,334	17	,742
Pair 3	B2_Gains_A_Male - B2_Gains_A_Female	,01666	,50231	,11840	-,23314	,26645	,141	17	,890
Pair 4	B2_Gains_U_Male - B2_Gains_U_Female	,07176	,52374	,12345	-,18869	,33221	,581	17	,569
Pair 5	Eucldiandist_Gains_Male - Eucldiandist_Gains_Female	-,21941	,63003	,14850	-,53271	,09390	-1,477	17	,158
Pair 6	B1_T1_A_Old - B1_T2_A_New	-,09431	,37899	,08933	-,28278	,09416	-1,056	17	,306
Pair 7	B1_T1_U_Old - B1_T2_U_New	,08845	,51562	,12153	-,16796	,34487	,728	17	,477
Pair 8	B2_T1_A_Old - B2_T2_A_New	-,04139	,39044	,09203	-,23555	,15277	-,450	17	,659
Pair 9	B2_T1_U_Old - B2_T2_U_New	,12414	,59187	,13950	-,17019	,41847	,890	17	,386
Pair 10	Euclidean_distance_T1_Old - Euclidean_distance_T2_New	,09824	,56366	,13286	-,18206	,37854	,739	17	,470
Pair 11	B1_T1_A_word - B1_T2_A_nonword	-,45190	,54693	,12891	-,72388	-,17991	-3,505	17	,003
Pair 12	B1_T1_U_word - B1_T2_U_nonword	-,35701	,63749	,15026	-,67403	-,04000	-2,376	17	,030
Pair 13	B2_T1_A_word - B2_T2_A_nonword	-,02838	,31900	,07519	-,18701	,13026	-,377	17	,711
Pair 14	B2_T1_U_word - B2_T2_U_nonword	-,03490	,37890	,08931	-,22332	,15352	-,391	17	,701
Pair 15	Euclidean_distance_mean_T1_word - Euclidean_distance_mean_T2_nonword	-,52438	,46912	,11057	-,75767	-,29109	-4,742	17	,000

Table 8.46. *Summary paired samples t-test between different independent variables (experimental group).*

Report								
Speakers						**B1_mean**	**B2_mean**	**Euclidean_distance**
NSs		**Male**	**Post-test**	**æ**	N	2	2	2
					Mean	6,2594	11,6104	1,1597
					Std. Deviation	,41230	,05709	,38970
				ʌ	N	2	2	
					Mean	5,4820	10,3636	
					Std. Deviation	,31966	,19989	
		Female	**Post-test**	**æ**	N	2	2	2
					Mean	7,7239	12,1062	2,8763
					Std. Deviation	1,37913	,22437	1,09354
				ʌ	N	2	2	
					Mean	7,1123	11,1688	
					Std. Deviation	,91455	,44252	
NNSs	**Experimental**	**Male**	**Post-test**	**æ**	N	9	9	9
					Mean	6,5611	10,6183	1,6986
					Std. Deviation	,45421	,42208	,57702
				ʌ	N	36	36	
					Mean	5,9626	10,1158	
					Std. Deviation	,41579	,64421	
			Delayed post-test	**æ**	N	9	9	9
					Mean	6,2522	10,5823	1,3995
					Std. Deviation	,39883	,45650	,95945
				ʌ	N	36	36	
					Mean	5,9762	10,4976	
					Std. Deviation	,47760	,58348	
		Female	**Post-test**	**æ**	N	9	9	9
					Mean	7,7929	11,5636	1,8128
					Std. Deviation	,36759	,31058	,75671
				ʌ	N	36	36	
					Mean	7,0952	11,0490	
					Std. Deviation	,56029	,37750	
			Delayed post-test	**æ**	N	9	9	9
					Mean	7,3474	11,2078	1,3262
					Std. Deviation	,43032	,40203	,73136
				ʌ	N	36	36	
					Mean	6,6598	11,1760	
					Std. Deviation	,52919	,34132	
	Control	**Male**	**Post-test**	**æ**	N	7	7	7
					Mean	5,9360	10,6703	1,2003
					Std. Deviation	,32421	,60708	,81316
				ʌ	N	28	28	
					Mean	5,6982	10,5824	
					Std. Deviation	,47096	,51941	
		Female	**Post-test**	**æ**	N	11	11	11
					Mean	6,6705	11,5594	1,2725
					Std. Deviation	,47213	,31948	1,07704
				ʌ	N	44	44	
					Mean	6,3845	11,3229	
					Std. Deviation	,46411	,49738	

Table 8.47. *Descriptives for native and non-native productions of vowels /æ/ & /ʌ/ (T1 -T2 - T3) in non-words.*

G.5. Task complexity and language-related episodes results

In order to have reliable numbers of language-related episodes, three coders analysed 100% of the data collected. In task 1, coder 1 agreed 66.7% and 77.8% with coders 2 and 3, respectively, and coder 2 and 3 agreed on 85.7%. Coder 1 in task 2 reached an agreement of 100% and 96.7% with coders 2 and 3, respectively, and coders 2 and 3 also agreed on 96.7%. As regards task 3, coder 1 agreed on 98.2% and 94.5% on coders 2 and 3, respectively, and coder 2 agreed on 92.9% with coder 3. Finally, in task 4, coder 1 agreed on 94.9% with coders 2 and 3, and coder 2 agreed on 100% with coder 3. See figures 9.1 and 9.2).

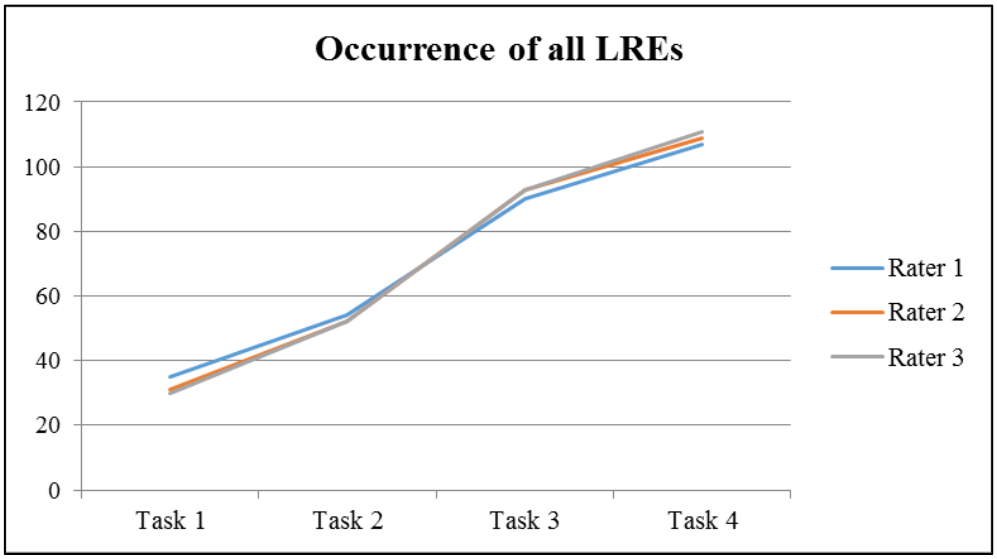


Figure 9.1. *Inter-rater results for the occurrence of LREs (General LREs, LRERC, LRESR, LRERC).*

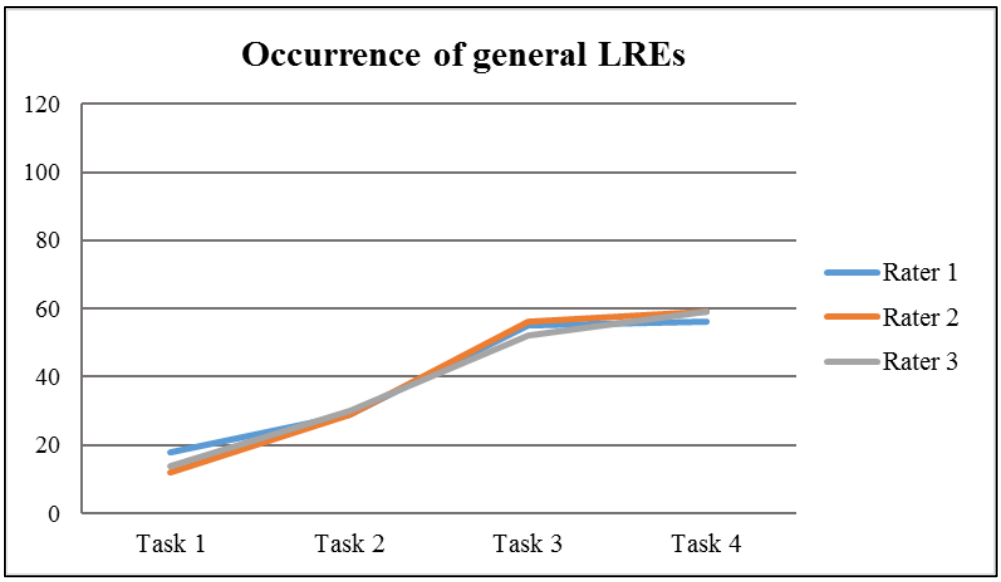


Figure 9.2. *Inter-rater results for the occurrence of only general LREs.*

Report					
Analysis		**Task_1**	**Task_2**	**Task_3**	**Task_4**
All LREs	**N**	18	18	18	18
	Mean	3,89	6,00	10,00	13,11
	Std. Deviation	2,968	2,910	4,826	4,562
General LREs	**N**	18	18	18	18
	Mean	2,00	3,22	6,11	6,22
	Std. Deviation	1,283	1,927	1,844	2,315
Rate LREs	**N**	18	18	18	18
	Mean	,8311	1,2129	1,2265	1,6034
	Std. Deviation	,63959	,44522	,59580	,62131

Table 9.1. *Descriptives (means and SD) across tasks (task 1 – task 2 – task 3 – task 4) for all LREs, GenLREs and LREs/minute*

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Tasks	Pillai's Trace	,895	42,630[b]	3,000	15,000	,000	,895
	Wilks' Lambda	,105	42,630[b]	3,000	15,000	,000	,895
	Hotelling's Trace	8,526	42,630[b]	3,000	15,000	,000	,895
	Roy's Largest Root	8,526	42,630[b]	3,000	15,000	,000	,895
a. Design: Intercept							
b. Exact statistic							

Table 9.2. *One-way repeated measures ANOVA across tasks (task 1 – task 2 – task 3 – task 4) for all LREs.*

Pairwise Comparisons						
Measure: Tasks						
					95% Confidence Interval for Difference[b]	
(I) tasks		**Mean Difference (I-J)**	**Std. Error**	**Sig.[b]**	**Lower Bound**	**Upper Bound**
1	2	-2,111*	,529	,006	-3,691	-,531
	3	-6,111*	,771	,000	-8,411	-3,811
	4	-9,222*	1,165	,000	-12,698	-5,747
2	1	2,111*	,529	,006	,531	3,691
	3	-4,000*	,970	,004	-6,895	-1,105
	4	-7,111*	1,453	,001	-11,446	-2,776
3	1	6,111*	,771	,000	3,811	8,411
	2	4,000*	,970	,004	1,105	6,895
	4	-3,111	1,140	,086	-6,513	,291
4	1	9,222*	1,165	,000	5,747	12,698
	2	7,111*	1,453	,001	2,776	11,446
	3	3,111	1,140	,086	-,291	6,513
Based on estimated marginal means						
*. The mean difference is significant at the ,05 level.						
b. Adjustment for multiple comparisons: Bonferroni.						

Table 9.3. *Bonferroni pairwise comparisons across tasks (task 1 – task 2 – task 3 – task 4) for all LREs.*

When analysing language-related episodes exclusively, the data was not normally distributed. As a result, the Friedman two-way ANOVA reported that general LREs were produced significantly different across the four tasks (X^2 (3) = 33.671, p<.001), see table 9.4. Nevertheless, Bonferroni pairwise comparisons indicated that learners engaged in more LREs from task 1 to task 3 and 4 (p<.05) but it was not the case between task1 and task 2 (p=121) and from task 3 to task 4 (p=1.000). See figure 9.3 for differences across tasks.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distributions of Task_1, Task_2, Task_3 and Task_4 are the same.	Related-Samples Friedman's Two-Way Analysis of Variance by Ranks	,000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Table 9.4. *Friedman test across tasks (task 1 – task 2 – task 3 – task 4) for general LREs.*

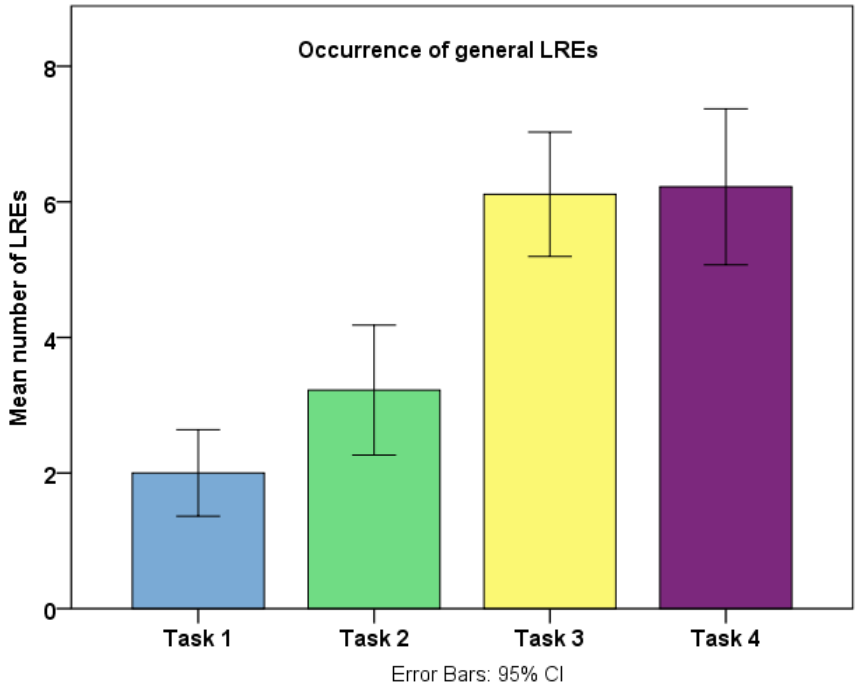


Figure 9.3. *Mean scores for the occurrence of general LREs (only interactional moves).*

Report				
	Task_1	**Task_2**	**Task_3**	**Task_4**
N	18	18	18	18
Mean	,8311	1,2129	1,2265	1,6034
Std. Deviation	,63959	,44522	,59580	,62131

Table 9.5. *Descriptives (means and SD) across tasks (task 1 – task 2 – task 3 – task 4) for LREs rate.*

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
Tasks	Pillai's Trace	,608	7,747[b]	3,000	15,000	,002	,608
	Wilks' Lambda	,392	7,747[b]	3,000	15,000	,002	,608
	Hotelling's Trace	1,549	7,747[b]	3,000	15,000	,002	,608
	Roy's Largest Root	1,549	7,747[b]	3,000	15,000	,002	,608
a. Design: Intercept b. Exact statistic							

Table 9.6. *One-way repeated measures ANOVA across tasks (task 1 – task 2 – task 3 – task 4) for LREs rate*

Pairwise Comparisons Measure: LREs rate						
					95% Confidence Interval for Difference[b]	
(I) Tasks		**Mean Difference (I-J)**	**Std. Error**	**Sig.[b]**	**Lower Bound**	**Upper Bound**
1	**2**	-,382*	,099	,008	-,678	-,086
	3	-,395*	,088	,002	-,658	-,133
	4	-,772*	,199	,007	-1,365	-,179
2	**1**	,382*	,099	,008	,086	,678
	3	-,014	,108	1,000	-,337	,309
	4	-,390	,165	,183	-,884	,103
3	**1**	,395*	,088	,002	,133	,658
	2	,014	,108	1,000	-,309	,337
	4	-,377	,155	,159	-,840	,086
4	**1**	,772*	,199	,007	,179	1,365
	2	,390	,165	,183	-,103	,884
	3	,377	,155	,159	-,086	,840
Based on estimated marginal means *. The mean difference is significant at the ,05 level. b. Adjustment for multiple comparisons: Bonferroni.						

Table 9.7. *Bonferroni pairwise comparisons across tasks (task 1 – task 2 – task 3 – task 4) for LREs rate.*

The mixed-design ANOVA revealed that there were statistically significant differences between General LREs, LRE recasts, LRE self-repairs (errors) and LRE repetitions, F (3,10) = 41.467, p<.001, η^2 =.926, see table 9.8 (appendix I.5). The descriptives in table

9.7., showed that learners produced more general LREs than repetitions, followed by self-repairs and recasts. Bonferroni pairwise comparisons showed that differences were found across all categories ($p<.05$) except for LRERC and LRESR-err ($p>.05$), see table 9.10. Moreover, the between-subjects effects indicated that there were statistically significant differences across tasks (F (1,3) = 29.694, $p<.001$, $\eta^2= .867$), see table 9.11. Finally, a significant disordinal interaction between LREs and tasks (F (9,36) = 3.305, $p=.005$, $\eta^2= .452$) confirmed that the production of the different types of LREs depended on the tasks the learners were involved in (see figure 9.4).

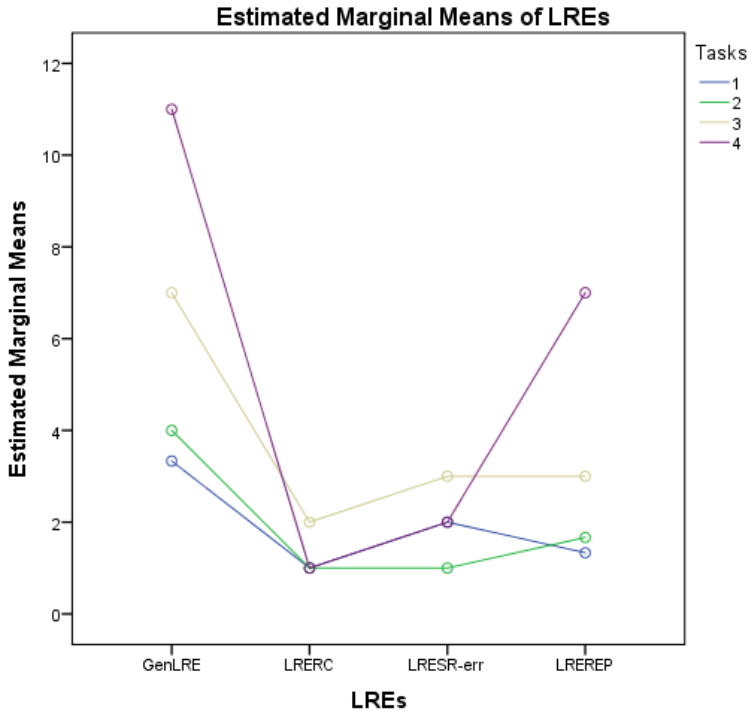


Figure 9.4. *Disordinal interaction between tasks (Task 1 – 2 – 3 – 4) and LREs (GenLRE, LRERC, LRESR-err, LREREP).*

Report					
Tasks		**GenLRE_mean**	**LRERC_mean**	**LRESR_mean**	**LREREP_mean**
1	**N**	18	8	10	8
	Mean	2,0000	1,0000	1,6000	1,2500
	Std. Deviation	1,28338	,75593	,84327	,46291
2	**N**	18	6	12	14
	Mean	3,2222	1,0000	1,5000	1,8571
	Std. Deviation	1,92676	0,00000	,79772	1,02711
3	**N**	18	2	12	16
	Mean	6,1111	2,0000	2,0000	2,6250
	Std. Deviation	1,84355	0,00000	,85280	2,30579
4	**N**	18	2	14	18
	Mean	6,2222	1,0000	1,5714	4,3333
	Std. Deviation	2,31505	0,00000	,51355	2,16930

Table 9.8. *Descriptives (means and SD) for mean scores of different types of LREs*.

Multivariate Tests[a]							
Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**	**Partial Eta Squared**
LREs	Pillai's Trace	,926	41,467[b]	3,000	10,000	,000	,926
	Wilks' Lambda	,074	41,467[b]	3,000	10,000	,000	,926
	Hotelling's Trace	12,440	41,467[b]	3,000	10,000	,000	,926
	Roy's Largest Root	12,440	41,467[b]	3,000	10,000	,000	,926
LREs * Tasks	Pillai's Trace	1,357	3,305	9,000	36,000	,005	,452
	Wilks' Lambda	,025	9,762	9,000	24,488	,000	,709
	Hotelling's Trace	25,497	24,552	9,000	26,000	,000	,895
	Roy's Largest Root	24,980	99,921[c]	3,000	12,000	,000	,962

a. Design: Intercept + Tasks
b. Exact statistic
c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Table 9.9. *One-way repeated measures ANOVA for all kinds of LREs (GenLRE, LRERC, LRESR-err, LREREP) and interaction between LREs and tasks.*

Pairwise Comparisons						
Measure: LREs						
(I) LREs		Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
1	2	5,083*	,509	,000	3,478	6,689
	3	4,333*	,561	,000	2,564	6,102
	4	3,083*	,419	,000	1,761	4,406
2	1	-5,083*	,509	,000	-6,689	-3,478
	3	-,750	,333	,264	-1,801	,301
	4	-2,000*	,215	,000	-2,678	-1,322
3	1	-4,333*	,561	,000	-6,102	-2,564
	2	,750	,333	,264	-,301	1,801
	4	-1,250*	,215	,001	-1,928	-,572
4	1	-3,083*	,419	,000	-4,406	-1,761
	2	2,000*	,215	,000	1,322	2,678
	3	1,250*	,215	,001	,572	1,928
Based on estimated marginal means						
*. The mean difference is significant at the ,05 level.						
b. Adjustment for multiple comparisons: Bonferroni.						

Table 9.10. *Bonferroni pairwise comparisons for different kinds of LREs.*

Tests of Between-Subjects Effects						
Measure: Tasks						
Transformed Variable:						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Intercept	494,083	1	494,083	433,829	,000	,973
Tasks	89,083	3	29,694	26,073	,000	,867
Error	13,667	12	1,139			

Table 9.11. *One way-repeated measures ANOVA for different tasks (task 1 – task 2 – task 3 – task 4).*

Correlations								
Gender			ID_gains_acc_pre_post	DIS_gains_acc_pre_post	PRO_B1_A_gains	PRO_B1_U_gains	PRO_ED_gains	All_LREs
Male	**ID_gains_acc_pre_post**	Pearson Correlation	1	-,516	,074	,040	-,244	-,269
		Sig. (2-tailed)		,155	,851	,918	,526	,484
		N	9	9	9	9	9	9
	DIS_gains_acc_pre_post	Pearson Correlation	-,516	1	,211	-,386	,286	,400
		Sig. (2-tailed)	,155		,585	,304	,455	,286
		N	9	9	9	9	9	9
	PRO_B1_A_gains	Pearson Correlation	,074	,211	1	,125	,879**	,605
		Sig. (2-tailed)	,851	,585		,748	,002	,085
		N	9	9	9	9	9	9
	PRO_B1_U_gains	Pearson Correlation	,040	-,386	,125	1	-,034	,163
		Sig. (2-tailed)	,918	,304	,748		,932	,676
		N	9	9	9	9	9	9
	PRO_ED_gains	Pearson Correlation	-,244	,286	,879**	-,034	1	,704*
		Sig. (2-tailed)	,526	,455	,002	,932		,034
		N	9	9	9	9	9	9
	All_LREs	Pearson Correlation	-,269	,400	,605	,163	,704*	1
		Sig. (2-tailed)	,484	,286	,085	,676	,034	
		N	9	9	9	9	9	9
Female	**ID_gains_acc_pre_post**	Pearson Correlation	1	-,077	-,206	-,634	,245	,277
		Sig. (2-tailed)		,844	,596	,067	,525	,470
		N	9	9	9	9	9	9
	DIS_gains_acc_pre_post	Pearson Correlation	-,077	1	-,221	,197	-,567	-,842**
		Sig. (2-tailed)	,844		,567	,611	,112	,004
		N	9	9	9	9	9	9
	PRO_B1_A_gains	Pearson Correlation	-,206	-,221	1	-,174	,442	-,229
		Sig. (2-tailed)	,596	,567		,654	,234	,554
		N	9	9	9	9	9	9
	PRO_B1_U_gains	Pearson Correlation	-,634	,197	-,174	1	-,709*	-,104
		Sig. (2-tailed)	,067	,611	,654		,032	,789
		N	9	9	9	9	9	9
	PRO_ED_gains	Pearson Correlation	,245	-,567	,442	-,709*	1	,338
		Sig. (2-tailed)	,525	,112	,234	,032		,373
		N	9	9	9	9	9	9
	All_LREs	Pearson Correlation	,277	-,842**	-,229	-,104	,338	1
		Sig. (2-tailed)	,470	,004	,554	,789	,373	
		N	9	9	9	9	9	9

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Table 9.12. *Pearson correlations between LREs and gains in perception and production (experimental group).*

G.6. Learner factors

Correlations									
Group				Proficiency_acc_scores	ID_gains_acc_pre_post	DIS_gains_acc_pre_post	PRO_B1_A_gains	PRO_B1_U_gains	PRO_ED_gains
Experimental	Male	Proficiency_acc_scores	Pearson Correlation	1	-,237	,588	,308	-,491	,428
			Sig. (2-tailed)		,539	,096	,419	,179	,250
			N	9	9	9	9	9	9
		ID_gains_acc_pre_post	Pearson Correlation	-,237	1	-,516	,074	,040	-,244
			Sig. (2-tailed)	,539		,155	,851	,918	,526
			N	9	9	9	9	9	9
		DIS_gains_acc_pre_post	Pearson Correlation	,588	-,516	1	,211	-,386	,286
			Sig. (2-tailed)	,096	,155		,585	,304	,455
			N	9	9	9	9	9	9
		PRO_B1_A_gains	Pearson Correlation	,308	,074	,211	1	,125	,879**
			Sig. (2-tailed)	,419	,851	,585		,748	,002
			N	9	9	9	9	9	9
		PRO_B1_U_gains	Pearson Correlation	-,491	,040	-,386	,125	1	-,034
			Sig. (2-tailed)	,179	,918	,304	,748		,932
			N	9	9	9	9	9	9
		PRO_ED_gains	Pearson Correlation	,428	-,244	,286	,879**	-,034	1
			Sig. (2-tailed)	,250	,526	,455	,002	,932	
			N	9	9	9	9	9	9
	Female	Proficiency_acc_scores	Pearson Correlation	1	,195	-,152	,321	-,285	,439
			Sig. (2-tailed)		,615	,695	,400	,457	,237
			N	9	9	9	9	9	9
		ID_gains_acc_pre_post	Pearson Correlation	,195	1	-,077	-,206	-,634	,245
			Sig. (2-tailed)	,615		,844	,596	,067	,525
			N	9	9	9	9	9	9
		DIS_gains_acc_pre_post	Pearson Correlation	-,152	-,077	1	-,221	,197	-,567
			Sig. (2-tailed)	,695	,844		,567	,611	,112
			N	9	9	9	9	9	9
		PRO_B1_A_gains	Pearson Correlation	,321	-,206	-,221	1	-,174	,442
			Sig. (2-tailed)	,400	,596	,567		,654	,234
			N	9	9	9	9	9	9
		PRO_B1_U_gains	Pearson Correlation	-,285	-,634	,197	-,174	1	-,709*
			Sig. (2-tailed)	,457	,067	,611	,654		,032
			N	9	9	9	9	9	9
		PRO_ED_gains	Pearson Correlation	,439	,245	-,567	,442	-,709*	1
			Sig. (2-tailed)	,237	,525	,112	,234	,032	
			N	9	9	9	9	9	9

Table 10.1. *Pearson correlations between proficiency and gains in perception and production by gender (experimental group).*

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	RT_mean_TC -	-45,07700	70,31712	16,57390	-80,04488	-10,10912	-2,720	17	,015
Pair 2	RT_mean_FC -	13,65013	39,73393	9,63689	-6,77917	34,07943	1,416	16	,176

Table 10.2. *Paired-samples t-test about differences in congruency with target and filler words (experimental group).*

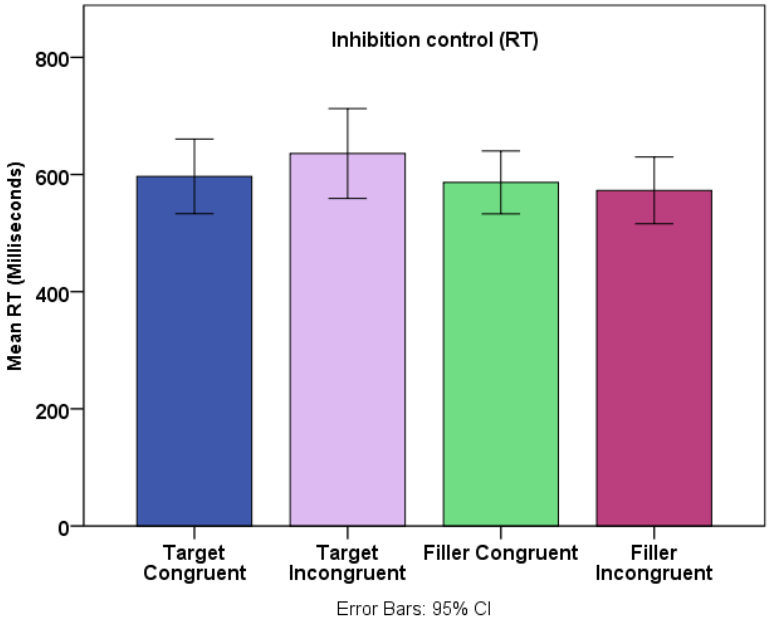


Figure 10.1. *Bar graph on RT mean scores for target and filler congruent and incongruent conditions (experimental group)*

Correlations

Group				Inhibition_RT_scores	ID_gains_acc_pre_post	DIS_gains_acc_pre_post	DIS_gains_RT_pre_post	PRO_B1_A_gains	PRO_B1_U_gains	PRO_ED_gains
Experimental	Male	Inhibition_RT_scores	Pearson	1	-,356	-,277	-,291	-,348	,419	-,072
			Sig. (2-tailed)		,346	,471	,447	,359	,262	,853
			N	9	9	9	9	9	9	9
		ID_gains_acc_pre_post	Pearson	-,356	1	-,516	,147	,074	,040	-,244
			Sig. (2-tailed)	,346		,155	,706	,851	,918	,526
			N	9	9	9	9	9	9	9
		DIS_gains_acc_pre_post	Pearson	-,277	-,516	1	,102	,211	-,386	,286
			Sig. (2-tailed)	,471	,155		,794	,585	,304	,455
			N	9	9	9	9	9	9	9
		DIS_gains_RT_pre_post	Pearson	-,291	,147	,102	1	-,550	,016	-,744*
			Sig. (2-tailed)	,447	,706	,794		,125	,968	,022
			N	9	9	9	9	9	9	9
		PRO_B1_A_gains	Pearson	-,348	,074	,211	-,550	1	,125	,879**
			Sig. (2-tailed)	,359	,851	,585	,125		,748	,002
			N	9	9	9	9	9	9	9
		PRO_B1_U_gains	Pearson	,419	,040	-,386	,016	,125	1	-,034
			Sig. (2-tailed)	,262	,918	,304	,968	,748		,932
			N	9	9	9	9	9	9	9
		PRO_ED_gains	Pearson	-,072	-,244	,286	-,744*	,879**	-,034	1
			Sig. (2-tailed)	,853	,526	,455	,022	,002	,932	
			N	9	9	9	9	9	9	9
	Female	Inhibition_RT_scores	Pearson	1	,222	-,418	-,155	-,505	,090	-,068
			Sig. (2-tailed)		,565	,262	,690	,165	,817	,861
			N	9	9	9	9	9	9	9
		ID_gains_acc_pre_post	Pearson	,222	1	-,077	,162	-,206	-,634	,245
			Sig. (2-tailed)	,565		,844	,677	,596	,067	,525
			N	9	9	9	9	9	9	9
		DIS_gains_acc_pre_post	Pearson	-,418	-,077	1	,313	-,221	,197	-,567
			Sig. (2-tailed)	,262	,844		,412	,567	,611	,112
			N	9	9	9	9	9	9	9
		DIS_gains_RT_pre_post	Pearson	-,155	,162	,313	1	-,559	-,277	-,215
			Sig. (2-tailed)	,690	,677	,412		,118	,471	,579
			N	9	9	9	9	9	9	9
		PRO_B1_A_gains	Pearson	-,505	-,206	-,221	-,559	1	-,174	,442
			Sig. (2-tailed)	,165	,596	,567	,118		,654	,234
			N	9	9	9	9	9	9	9
		PRO_B1_U_gains	Pearson	,090	-,634	,197	-,277	-,174	1	-,709*
			Sig. (2-tailed)	,817	,067	,611	,471	,654		,032
			N	9	9	9	9	9	9	9
		PRO_ED_gains	Pearson	-,068	,245	-,567	-,215	,442	-,709*	1
			Sig. (2-tailed)	,861	,525	,112	,579	,234	,032	
			N	9	9	9	9	9	9	9

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
c. Cannot be computed because at least one of the variables is constant.

Table 10.3. *Correlations between inhibitory control and gains in perception and production (experimental group).*

Paired Samples Test									
		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Colour_mean - Digit_mean	-,00868	,06591	,01553	-,04145	,02409	-,559	17	,584

Table 10.4. *Paired samples t-test about differences in colour and digit (experimental group).*

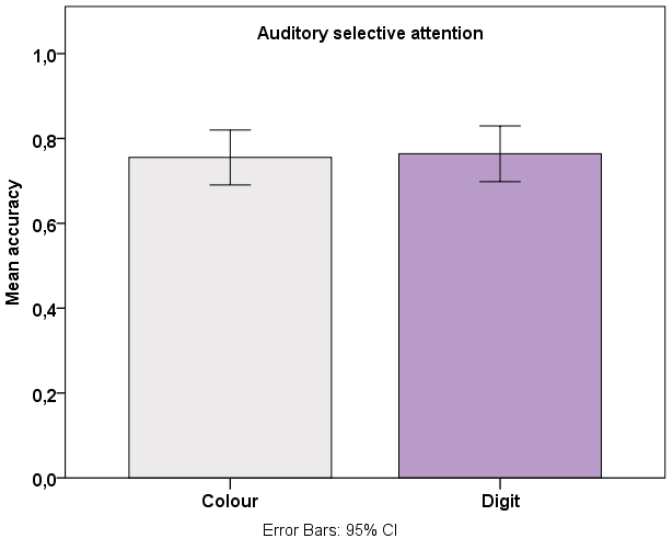


Figure 10.2. *Bar graph on accuracy scores for colour and digit (experimental group).*

Correlations										
Group				Attention_acc_scores	ID_gains_acc_pre_post	DIS_gains_acc_pre_post	DIS_gains_RT_pre_post	PRO_B1_A_gains	PRO_B1_U_gains	PRO_ED_gains
Experimental	Male	Attention_acc_scores	Pearson Correlation	1	-,370	,413	-,283	,265	-,654	,431
			Sig. (2-tailed)		,328	,270	,461	,490	,056	,247
			N	9	9	9	9	9	9	9
		ID_gains_acc_pre_post	Pearson Correlation	-,370	1	-,516	,147	,074	,040	-,244
			Sig. (2-tailed)	,328		,155	,706	,851	,918	,526
			N	9	9	9	9	9	9	9
		DIS_gains_acc_pre_post	Pearson Correlation	,413	-,516	1	,102	,211	-,386	,286
			Sig. (2-tailed)	,270	,155		,794	,585	,304	,455
			N	9	9	9	9	9	9	9
		DIS_gains_RT_pre_post	Pearson Correlation	-,283	,147	,102	1	-,550	,016	-,744*
			Sig. (2-tailed)	,461	,706	,794		,125	,968	,022
			N	9	9	9	9	9	9	9
		PRO_B1_A_gains	Pearson Correlation	,265	,074	,211	-,550	1	,125	,879**
			Sig. (2-tailed)	,490	,851	,585	,125		,748	,002
			N	9	9	9	9	9	9	9
		PRO_B1_U_gains	Pearson Correlation	-,654	,040	-,386	,016	,125	1	-,034
			Sig. (2-tailed)	,056	,918	,304	,968	,748		,932
			N	9	9	9	9	9	9	9
		PRO_ED_gains	Pearson Correlation	,431	-,244	,286	-,744*	,879**	-,034	1
			Sig. (2-tailed)	,247	,526	,455	,022	,002	,932	
			N	9	9	9	9	9	9	9
	Female	Attention_acc_scores	Pearson Correlation	1	,040	-,280	-,862**	,402	,134	,215
			Sig. (2-tailed)		,918	,466	,003	,283	,731	,579
			N	9	9	9	9	9	9	9
		ID_gains_acc_pre_post	Pearson Correlation	,040	1	-,077	,162	-,206	-,634	,245
			Sig. (2-tailed)	,918		,844	,677	,596	,067	,525
			N	9	9	9	9	9	9	9
		DIS_gains_acc_pre_post	Pearson Correlation	-,280	-,077	1	,313	-,221	,197	-,567
			Sig. (2-tailed)	,466	,844		,412	,567	,611	,112
			N	9	9	9	9	9	9	9
		DIS_gains_RT_pre_post	Pearson Correlation	-,862**	,162	,313	1	-,559	-,277	-,215
			Sig. (2-tailed)	,003	,677	,412		,118	,471	,579
			N	9	9	9	9	9	9	9
		PRO_B1_A_gains	Pearson Correlation	,402	-,206	-,221	-,559	1	-,174	,442
			Sig. (2-tailed)	,283	,596	,567	,118		,654	,234
			N	9	9	9	9	9	9	9
		PRO_B1_U_gains	Pearson Correlation	,134	-,634	,197	-,277	-,174	1	-,709*
			Sig. (2-tailed)	,731	,067	,611	,471	,654		,032
			N	9	9	9	9	9	9	9
		PRO_ED_gains	Pearson Correlation	,215	,245	-,567	-,215	,442	-,709*	1
			Sig. (2-tailed)	,579	,525	,112	,579	,234	,032	
			N	9	9	9	9	9	9	9

Table 10.5. *Pearson correlations between auditory selective attention and gains in perception and production (females/males of the experimental group).*

Correlations									
Proficiency_median			Attention_acc_scores	ID_gains_acc_pre_post	DIS_gains_acc_pre_post	DIS_gains_RT_pre_post	PRO_B1_A_gains	PRO_B1_U_gains	PRO_ED_gains
Low	Attention_acc_scores	Pearson Correlation	1	-,567	,223	-,461	,726*	,430	,432
		Sig. (2-tailed)		,112	,564	,212	,027	,248	,246
		N	9	9	9	9	9	9	9
	ID_gains_acc_pre_post	Pearson Correlation	-,567	1	-,396	,443	-,745*	-,624	-,313
		Sig. (2-tailed)	,112		,292	,233	,021	,073	,412
		N	9	9	9	9	9	9	9
	DIS_gains_acc_pre_post	Pearson Correlation	,223	-,396	1	,080	,233	,699*	-,204
		Sig. (2-tailed)	,564	,292		,839	,547	,036	,598
		N	9	9	9	9	9	9	9
	DIS_gains_RT_pre_post	Pearson Correlation	-,461	,443	,080	1	-,430	-,425	-,226
		Sig. (2-tailed)	,212	,233	,839		,248	,254	,559
		N	9	9	9	9	9	9	9
	PRO_B1_A_gains	Pearson Correlation	,726*	-,745*	,233	-,430	1	,292	,599
		Sig. (2-tailed)	,027	,021	,547	,248		,446	,088
		N	9	9	9	9	9	9	9
	PRO_B1_U_gains	Pearson Correlation	,430	-,624	,699*	-,425	,292	1	-,235
		Sig. (2-tailed)	,248	,073	,036	,254	,446		,544
		N	9	9	9	9	9	9	9
	PRO_ED_gains	Pearson Correlation	,432	-,313	-,204	-,226	,599	-,235	1
		Sig. (2-tailed)	,246	,412	,598	,559	,088	,544	
		N	9	9	9	9	9	9	9
High	Attention_acc_scores	Pearson Correlation	1	-,289	,085	-,404	,158	-,093	-,036
		Sig. (2-tailed)		,451	,827	,281	,684	,812	,927
		N	9	9	9	9	9	9	9
	ID_gains_acc_pre_post	Pearson Correlation	-,289	1	-,432	-,003	,323	-,080	-,071
		Sig. (2-tailed)	,451		,246	,993	,397	,838	,855
		N	9	9	9	9	9	9	9
	DIS_gains_acc_pre_post	Pearson Correlation	,085	-,432	1	,581	-,191	-,511	-,061
		Sig. (2-tailed)	,827	,246		,101	,622	,160	,876
		N	9	9	9	9	9	9	9
	DIS_gains_RT_pre_post	Pearson Correlation	-,404	-,003	,581	1	-,414	-,647	-,035
		Sig. (2-tailed)	,281	,993	,101		,268	,060	,929
		N	9	9	9	9	9	9	9
	PRO_B1_A_gains	Pearson Correlation	,158	,323	-,191	-,414	1	,101	,515
		Sig. (2-tailed)	,684	,397	,622	,268		,795	,156
		N	9	9	9	9	9	9	9
	PRO_B1_U_gains	Pearson Correlation	-,093	-,080	-,511	-,647	,101	1	-,339
		Sig. (2-tailed)	,812	,838	,160	,060	,795		,372
		N	9	9	9	9	9	9	9
	PRO_ED_gains	Pearson Correlation	-,036	-,071	-,061	-,035	,515	-,339	1
		Sig. (2-tailed)	,927	,855	,876	,929	,156	,372	
		N	9	9	9	9	9	9	9
*. Correlation is significant at the 0.05 level (2-tailed).									

Table 10.6. *Pearson correlations between auditory selective attention and gains in perception and production (low-/ high-proficiency learners of the experimental group).*

When comparing these results on individual differences with the control group, it appears that inhibition control was not significantly related to any test of perception and production (p>.05), just like the experimental group. Similar to the results from the experimental group, there was a significantly strong correlation between selective attention and production gains for /ʌ/, *r*=-919, *p*=003 in male learners. Nevertheless, no correlations were found in relation to female learners.

Group			Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Experimental	**Male**	Pearson Chi-Square	,900[a]	1	,343		
		Continuity Correction[b]	,056	1	,813		
		Likelihood Ratio	,908	1	,341		
		Fisher's Exact Test				,524	,405
		Linear-by-Linear Association	,800	1	,371		
		N of Valid Cases	9				
	Female	Pearson Chi-Square	,225[a]	1	,635		
		Continuity Correction[b]	0,000	1	1,000		
		Likelihood Ratio	,228	1	,633		
		Fisher's Exact Test				1,000	,595
		Linear-by-Linear Association	,200	1	,655		
		N of Valid Cases	9				
Control	**Male**	Pearson Chi-Square	1,215[c]	1	,270		
		Continuity Correction[b]	,109	1	,741		
		Likelihood Ratio	1,243	1	,265		
		Fisher's Exact Test				,486	,371
		Linear-by-Linear Association	1,042	1	,307		
		N of Valid Cases	7				
	Female	Pearson Chi-Square	2,396[d]	1	,122		
		Continuity Correction[b]	,883	1	,347		
		Likelihood Ratio	2,516	1	,113		
		Fisher's Exact Test				,242	,175
		Linear-by-Linear Association	2,178	1	,140		
		N of Valid Cases	11				

a. 4 cells (100,0%) have expected count less than 5. The minimum expected count is 1,33.
b. Computed only for a 2x2 table
c. 4 cells (100,0%) have expected count less than 5. The minimum expected count is 1,29.
d. 4 cells (100,0%) have expected count less than 5. The minimum expected count is 2,27.

Table 10.7. *Chi square test between auditory selective attention and inhibition control (experimental and control groups).*