

Article

Exploring the Pronunciation Awareness Continuum through Self-Reflection in the L2 French Learning Process

Camille Meritan 

Modern Languages Department, Bentley University, Waltham, MA 02452, USA; cmeritan@bentley.edu

Abstract: Second language (L2) researchers have established that examining learners' awareness of their own learning process and progress is essential. However, learners exposed to the same input in the classroom may differ in the way that they perform. This difference may be due to the way and depth with which learners process the L2 information. The present study explores self-reflection (i.e., introspective verbal reports) to enhance L2 learning, helping learners develop an awareness of learning as a process. This four-semester-long study investigates whether there is a connection between phonological awareness and self-reflection and explores under which conditions self-reflection could be most beneficial for pronunciation. Sixty learners of French, divided into a Treatment group (with self-reflection) and a Comparison group (without self-reflection), were tracked across semesters. Results on pre/post read-aloud tests surrounding pronunciation lessons—on the vowels /y/-/u/ and the use of liaisons—were contrasted with students' responses to self-reflection questionnaires to explore their learning process. The study revealed that overall, self-reflection led to better learning outcomes, and that a link between attention and understanding may exist, but when this link is absent, learners using self-reflection may not linearly progress.

Keywords: L2 Pronunciation; depth of processing; language awareness; intelligibility; French



Citation: Meritan, Camille. 2021. Exploring the Pronunciation Awareness Continuum through Self-Reflection in the L2 French Learning Process. *Languages* 6: 182. <https://doi.org/10.3390/languages6040182>

Academic Editor: Ronald Leow

Received: 26 July 2021

Accepted: 27 October 2021

Published: 1 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Researchers have recently established a gap in research regarding the development of Second Language (L2) pronunciation and its relationship to phonological awareness (Derwing 2018; Kennedy et al. 2014; Kennedy and Trofimovich 2010; Venkatagiri and Levis 2007). There is, indeed, a lack of efficient pedagogical methods to support and foster learners' processing and phonological awareness. There is also a lack of efficient research methods to assess learners' depth of processing of phonological features and awareness of their own learning process and progress. Yet, since students are the central part in their own learning (Raofi et al. 2013), it is essential to examine this process. One potentially useful tool to understand and operationalize learners' awareness of their learning process and progress, as well as their depth of processing, is self-reflection. Self-reflection, an introspective technique in the form of guided written reports, is fundamentally a learner-centered metacognitive learning strategy. Addressing the existing gap in research, this study investigates the use of guided verbal reports in the form of self-reflection to explore depth of processing in relation to pronunciation and phonological awareness.

Additionally, research on the impact of phonological awareness on the L2 learning process is nascent. Only few studies have investigated awareness through self-reflection, though not longitudinally (Inceoglu 2021; Kennedy et al. 2014; Meritan and Mroz 2019). The current study will quantitatively and qualitatively, as well as longitudinally, explore the role of self-reflection on L2 French learners' pronunciation, their phonological awareness, their learning process, and their learning outcomes (Derwing 2018; Inceoglu 2021; Kennedy et al. 2014). The goal of such investigation is twofold: (1) determine whether there is a connection between phonological awareness and self-reflection, and (2) examine how self-reflection can foster the emergence of awareness of learners' own progress. The

present study contributes to the field of research on L2 pronunciation by extending [Leow's \(2015\)](#) model of the L2 learning process, proposing a Phonological Awareness Continuum framework. This framework posits that, only when attention and noticing ([Schmidt 1990](#)) are combined can optimal learning outcomes be reached.

2. Literature Review

2.1. Foreign Language Pronunciation Instruction

In their 2015 meta-analysis of 86 studies investigating the overall effects of pronunciation instruction, [Lee et al. \(2015\)](#) demonstrated that there was a large effect of pronunciation instruction (Cohen's $d = 0.89$), and that explicit phonological instruction is beneficial for L2 learners. These results are supported in part by [Saito and Plonsky's \(2019\)](#) meta-analysis of 77 studies on L2 pronunciation instruction. They found that instruction was beneficial at a controlled speech level but found less proof when it came to spontaneous speech. They argue, however, that pronunciation instruction has an impact on the development of *controlled pronunciation knowledge* (i.e., declarative knowledge), defined as “what L2 learners learn from explicit phonetic instruction”, into *spontaneous pronunciation knowledge* (i.e., procedural knowledge), defined as “the degree to which L2 learners access explicit pronunciation knowledge [. . .]” (p. 665).

Furthermore, research has shown that pronunciation instruction leads to improved intelligibility by drawing learners' attention to specific pronunciation challenges that may impede their intelligibility ([Derwing 2018](#); [Levis 2018](#); [Meritan and Mroz 2019](#)).

Intelligibility corresponds to “a listener's ability to correctly recognize individual words within an utterance” ([Thomson 2017](#), p. 14) and thus cannot be assessed “without reference to listeners' experience” ([Munro and Derwing 2015](#), p. 382). It is simultaneously impacted by both segmental and suprasegmental features. More specifically, vowel and consonant sounds cannot be omitted in an intelligibility-based approach to L2 and foreign language (FL) pronunciation, and one needs to be sufficiently accurate in producing those sounds to be understood. Furthermore, when taking the message and its interpretation into consideration, suprasegmentals are the most important ([Levis 2018](#); [Thomson and Derwing 2015](#)).

This study thus focused on the explicit instruction of segmental and suprasegmental features hypothesized to impact intelligibility in French: the /y/-/u/ contrastive pair and liaisons, in a classroom setting. It is proposed here that using metacognitive strategies such as self-reflection may help draw learners' attention beyond specific phonological characteristics and raise phonological awareness ([Meritan and Mroz 2019](#); [Inceoglu 2021](#)).

2.2. Phonological Awareness and Depth of Processing

The impact of explicit instruction extends further than helping learners produce studied features more intelligibly. It helps them become more aware of pronunciation challenges they may face, it leads to improved phonological awareness, and finally, it improves pronunciation comprehensibility ([Derwing 2018](#)). Research has shown that developing phonological awareness is key to foster different facets of pronunciation, including comprehensibility, perception, intonation, fluency, self-repairs, and accuracy ([Derwing 2018](#); [Kennedy and Blanchet 2014](#); [Kennedy et al. 2014](#); [Kennedy and Trofimovich 2010](#); [Venkatagiri and Levis 2007](#); [Wrembel 2011, 2015](#)). [Kennedy et al. \(2014\)](#) suggested that “the relationship between awareness and instruction is reciprocal and self-reinforcing” (p. 91): as metacognition develops, learning is reinforced, which in turn furthers the development of metacognition.

In recent years, instructed Second Language Acquisition (ISLA) research has shown that self-assessment can enhance L2 learning, notably by helping learners develop cognitive and metacognitive learning strategies, i.e., an awareness of learning as a process ([de Saint Léger 2009](#); [Glover 2011](#); [Pawlak 2010](#); [Raoofi et al. 2013](#)). [Lennon \(1989\)](#) established that introspective techniques, such as guided reports via prompts, could lead to L2 learners being more aware of their progress, performance, and state of competence. More recently,

researchers have recommended using metalinguistic activities such as self-reflection assessments to allow reflection on language and as a particularly effective learning tool that probes deeper levels of reflection and noticing (Meritan and Mroz 2019; Inceoglu 2021; Moyer 2017; Wrembel 2015). Furthermore, according to Leow (2015), even if learners are exposed to the same input in the language classroom, they may differ in the way that they perform. This difference may be due to the way and depth with which learners process the L2 information, i.e., different depth of processing. Self-reflection may then be a potential practical tool to support learners in processing the L2 information more deeply. However, to date, no study has longitudinally explored the relationship between self-reflection and depth of processing of L2 pronunciation.

According to Foote and Trofimovich (2018), there are a lack of theories pertaining to L2 pronunciation research. Most SLA theories target grammar and/or vocabulary acquisition (see Van Patten's (2007) Input Processing Theory, Swain's (2005) Output Theory, or Leow's (2015) Model of the L2 Learning Process in ISLA), but few have been extended to L2 pronunciation acquisition (see Flege's (1995) Speech Learning Model, or Best's (1995) Perceptual Assimilation Model). Following Foote and Trofimovich's (2018) guidelines on the learning phenomena to be included in a viable L2 Pronunciation Theory, the present study set to explore Leow's (2015) Model of the L2 Learning Process as a viable theoretical framework for L2 pronunciation acquisition and processing. This model is built on "the role of attention in the process of learning an L2 [...] regulated by [...] the depth of processing or the amount of cognitive effort [...], cognitive registration [...], and/or the level of awareness [...]" (Leow 2015, p. 241). It hypothesizes that without minimal attention being paid to the L2 information, there is little chance for foreign language acquisition, especially in a classroom setting.

Depth of processing was first introduced in cognitive psychology by Craik and Lockhart's (1972) framework of levels of processing. In this framework, conceptual processing—the decoding of the word (or sound)—is contrasted with perceptual processing—the encoding of a particular feature, (e.g., encoding the sound of a phoneme). Leow (2015, p. 205) later redefined depth of processing as "if one were to process incoming information at a deeper level, that is, employ greater cognitive effort during processing while using prior knowledge to strengthen the process, the chances of remembering such information are substantially increased".

Leow's (2015) depth of processing (DoP) model was chosen because it fits several of Foote and Trofimovich's (2018) learning phenomena. These include "the importance of input in L2 pronunciation development, the influence of learners' first language (L1), the significant role of individual differences, and the systematicity and variability of pronunciation development".

The model is comprised of both level of processing and cognitive effort. For his model, Leow postulates three processing stages: the input processing stage, the intake processing stage, and the knowledge processing stage. The first stage goes hand in hand with Foote and Trofimovich's (2018) first learning phenomenon: the role of input. Indeed, Leow hypothesizes potential roles of three variables (cognitive registration, awareness, DoP) and assigns them to three types of intake which rely heavily on the level of attention paid to the input (see Figure 1): Attended intake (+attention, −cognitive registration, −awareness, extremely low DoP), Detected intake (+attention, +cognitive registration, −awareness, very low DoP), and Noticed intake (+attention, +cognitive registration, +low level of awareness, low DoP).

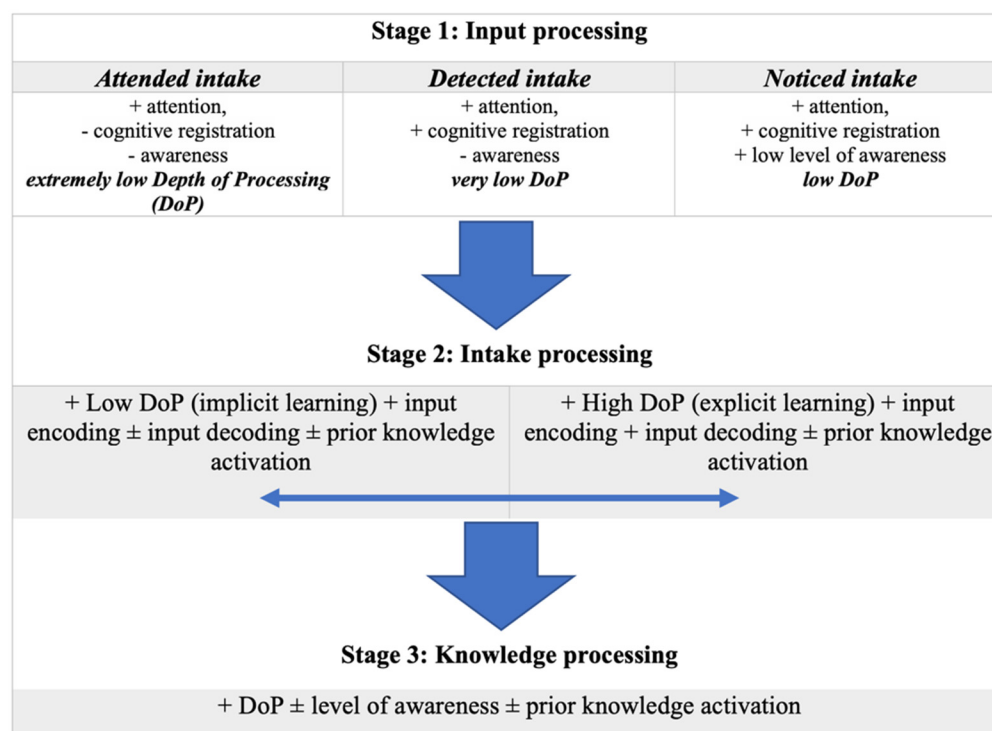


Figure 1. The three processing stages of the L2 learning process (adapted from [Leow 2015](#)).

The second stage demonstrates that intake can be processed in two ways: (1) with low level of processing, encoding the input as a chunk of language, or (2) with high level of processing, i.e., encoding and decoding the input with higher level of awareness. This stage also addresses the potential for type of learning. According to the model, it is possible for both low (implicit learning) and high (explicit learning) Depth of Processing to occur in relation to both encoding and decoding of the incoming L2 data, although the former type of learning is usually not what takes place in the instructed setting. This stage also includes the role of potential activation of old or new prior knowledge and the notion of restructuring. This stage thus can potentially be in accordance with Foote and Trofimovich's "influence of learners' L1" phenomenon as learners could activate their L1 as (prior) knowledge, drawing comparisons between the L1 and the L2. Indeed, as [Leow \(2015, p. 245\)](#) explains: "the combination of prior knowledge [possibly L1] activation, depth of processing, and potential higher level of awareness" allows the data to be restructured and stored.

The third and final processing stage is referred to as the knowledge processing stage and occurs before learners' production. In this stage, depth of processing, levels of awareness and the ability to activate prior knowledge may play a role. [Foote and Trofimovich's \(2018\)](#) "influence of learners' L1" can be assimilated to this stage as well. As prior knowledge potentially plays a role here, the influence of learner's L1 may play a role as well.

Furthermore, depth of processing is closely aligned with awareness ([de la Fuente 2015](#); [Hsieh et al. 2015](#); [Leow 2015, 2019](#); [Rosa and Leow 2004](#)). Of particular importance to this study are two levels of awareness ([Schmidt 1990](#)): awareness at the level of noticing, which is assimilated to low depth of processing and low level of awareness, and awareness at the level of understanding, assimilated to high depth of processing and cognitive effort, i.e., when hypothesis testing and rule formation occur.

However, depth of processing and levels of awareness are not necessarily linear. As depth of processing increases, so does the level of awareness. However, while high depth of processing uses more cognitive effort to arrive at the formation of underlying rules and hypothesis testing, awareness at the level of understanding can only be reached when the correct underlying rule is understood ([Leow 2015, p. 245](#)). Only then, the

need for high depth of processing decreases. Finally, Foote and Trofimovich's last two learning phenomena, namely, "the role of individual differences" and "the systematicity and variability of pronunciation development" can be illustrated in Leow's (2015) model. Indeed, as Leow (2015) explains, learners process the L2 differently, due in part to their individual differences. Pronunciation learning can thus be difficult or incomplete in its outcomes as learners process instruction and the input received differently (Foote and Trofimovich 2018). Furthermore, variability and systematicity of learning can be explained by whether and how the language is processed and manipulated, as learners may monitor their production.

Further research is thus needed to investigate Leow's (2015) model in relation to levels of awareness and depth of processing in pronunciation learning.

2.3. The Phonological Awareness Framework

Leow and Donatelli (2017) warned that research "has yet to establish whether 'awareness' [...] occurs on a continuum" (p. 190). No studies have yet operationalized depth of processing and levels of awareness for phonological items and pronunciation acquisition. This study's contribution to advance research is twofold: (1) to expand on the author's (Meritan and Mroz 2019) Awareness Continuum Theory based on Wrembel's (2013, 2015) studies on metaphonological awareness and (2) to extend Leow's (2015) operationalization of depth of processing of grammatical items to pronunciation. The framework being proposed here (renamed the Phonological Awareness Continuum) thus combines Wrembel's (2013, 2015) research as well as Leow's (2015) model in an attempt to operationalize depth of processing of pronunciation (Table 1).

Table 1. Operationalization of Depth of Processing: Pronunciation/Phonological items (adapted from Wrembel (2013, 2015), and Leow (2015)).

	Low Depth of Processing	Medium Depth of Processing	High Depth of Processing	
Level of Awareness	Attention	Reporting/Noticing	Introspection/Metacognition	+/- Understanding
Description	Signs of active effort but no potential for processing target form.	Signs of pronunciation difficulties regarding production/performance.	Signs of self-reflection on learning process and FL acquisition. Signs of use of prior knowledge (PK)	Signs of hypotheses formation regarding phonological rules, articulatory explanations and examples (correct or not)

FL = foreign language, +/- = correct or incorrect underlying rule.

The framework posits a relationship between attention, noticing, introspection, and understanding in the specific process of raising phonological awareness through self-reflection.

In the proposed framework, attention is a core mental function (Paulson et al. 2013) that allows learners to direct their brain's resources to actively and selectively prioritize some information over other information (Jha 2017; Lo 2018). Directed attention is considered to be a metacognitive strategy as it allows the transformation of received input into detected input (Wrembel 2011). Attention is closely aligned with minimal depth of processing as there is little to no cognitive effort. In the specific case of pronunciation instruction, Guion and Pederson (2007) investigated the impact of attention to sound versus attention to meaning in 76 native speakers of English learning Hindi. After conducting within-groups (from pre- to post-test) and between-groups analyses on a discrimination test and a semantic test, the researchers found that greater attention to specific aspects of pronunciation during instruction improved the learning of those aspects.

Noticing is the "subjective correlate of attention" (Schmidt 1990, p. 6) and represents both detection and registration of the attended input. What learners notice in the input is hypothesized to become intake. Noticing is aligned with medium depth of processing,

and is an upgrade to detected intake (+attention, +cognitive registration, –awareness) and only differs in the sense of a low level of awareness. In the specific case of pronunciation instruction, [Kivistö-de Souza \(2017\)](#) investigated learners' awareness and noticing of contrastive and non-contrastive L2 segments. Participants were 71 advanced L1 Brazilian Portuguese learners of English. They had to complete a perception task to identify if they were able to perceive correct vs. incorrect pronunciation by native speakers of English. The researcher found that the natural L2 input the students received was insufficient for noticing to happen, and that instruction may help learners to notice details of the L2 speech.

Introspection (or metacognition) “assumes that a person can observe what takes place in consciousness in much the same way as one can observe events in the external world” ([Mackey and Gass 2016](#)). It is a mechanism that is sensitive to perceived and processed input—or intake—which produces output based on that specific input ([Goldman 2008](#)). It is defined as a means of “learn[ing] about one's own ongoing (...) mental states and processes” ([Schwitzgebel 2014](#), p. 1), by fostering insight, i.e., gaining cognitive clarity of one's mental states ([Grant et al. 2002](#)) and is a necessary condition for objectively measuring awareness ([Timmermans and Cleeremans 2015](#)). Introspection/metacognition in this study was defined as learners' experiencing insight about their own mental states, learning process, learning progress, and prior knowledge and how they monitored themselves through the learning process ([Cerezo et al. 2016](#); [Leow et al. 2019](#)). It is associated with a high depth of processing as there is cognitive effort to process the target item phonologically. In the specific case of L2 French pronunciation, [Kennedy et al. \(2014\)](#) investigated how learners' awareness of their pronunciation related to their production and found some moderate effect of unguided self-reflections through peer-to-peer journaling on learners' segmental and suprasegmental pronunciation. [Inceoglu \(2021\)](#) conducted a study on 30 advanced learners of L2 French in a pronunciation course and examined the relationship between learners' pronunciation awareness and pronunciation development. It was found that there was a link between pronunciation awareness and pronunciation improvement. More specifically, it was determined that there was an association between learners' progress and their ability to reflect on pronunciation acquisition.

Finally, understanding happens when “attended and noticed instances become the basis for explicit hypothesis formation and testing” ([Schmidt 2010](#), p. 726). Understanding has also been defined as learners' higher-order ability to analyze, compare, and verbalize underlying rules that (re)map language intake ([de la Fuente 2015](#); [Hsieh et al. 2015](#)). [Rosa and Leow \(2004\)](#) conducted a study on 100 advanced learners of Spanish (L1 English). The researchers found that displaying awareness at the level of understanding after receiving instruction represented a significant advantage in producing target structures when compared to learners who only displayed awareness at the level of noticing. In the present study, understanding was defined as indices of learners using their insight to form, articulate, test, or manipulate underlying phonological rules. Awareness at the level of understanding is closely aligned with a high depth of processing as there is a high level of cognitive effort ([Leow 2015, 2019](#)). However, learners may reach a high level of processing without reaching awareness at the level of understanding, i.e., only when the correct underlying phonological rule is given does the learner reach awareness at the level of understanding. Introspection and understanding are considered cognitive learning strategies ([Pawlak 2010](#)).

The Phonological Awareness Continuum framework thus proposes that, subsequent to explicit instruction, awareness at the level of noticing (low depth of processing and low level of awareness) could be the starting point for pronunciation learning in that noticing could foster some additional cognitive effort (medium depth of processing), which in turn could advance phonological awareness at the level of introspection, potentially leading to the level of phonological understanding (high depth of processing) ([Leow 2015, 2019](#); [Leow and Donatelli 2017](#)).

Finally, researchers support the need for more qualitative research, arguing that self-assessment research is primarily quantitative and only explores the validity and reliability of students' self-rating rather than investigating learners' learning process ([de Saint Léger](#)

2009). Kennedy et al. (2014) established a gap in research regarding the development of L2 French pronunciation linked to phonological awareness. They explained that explicit pronunciation instruction and awareness of how L2 pronunciation works helped students pay more attention. However, they called for further investigations featuring a comparison group, with research exploring precisely the relationship between awareness and instruction. Following this call and building on previous research, the current study thus sought to answer the following research questions:

RQ1. *How did self-reflection impact learners' production outcomes on segmental and suprasegmental features, when compared to learners not exposed to self-reflection?*

RQ2. *What did self-reflective statements from a subsample of participants reveal about the role of phonological awareness-raising in the learners' learning process?*

3. Materials and Methods

3.1. Instructional Setting and Recruitment of Participants

According to Derwing and Munro (2013), “researching the longitudinal development of L2 learners is essential to understanding influences in their success” (p. 334). It is additionally necessary to observe the longitudinal outcomes of self-reflection as well as its impact on the learning process. This study thus tracked French learners' learning outcomes, learning growth, and learning process during 4 consecutive semesters at a large Midwestern university.

French 1, 2, 3, and 4 (FR1, FR2, FR3, and FR4) are the four courses that composed the French Foreign Language Requirement sequence at the large Midwestern university where the study was conducted. All instructors were near-native or native speakers of French who followed the same curriculum and used the same instructional materials. The curriculum in these courses was characterized by its integrated approach to explicit pronunciation instruction within an otherwise communicative-language-based approach to teaching and learning (Kennedy et al. 2014; Ranta and Lyster 2018). Following Gordon and Isabelle's (2016) integrated approach, this study's explicit pronunciation lessons were added to the communicative-based curricula, which allowed instructors to introduce the topic, guide learners, and assess their production. Each semester (16 weeks), courses met four times a week for 50 min. Students received a total of six explicit 20 min pronunciation lessons in class followed by practice exercises targeting perception and production. Lessons were based on spelling-to-sound patterns and targeted specific phonological phenomena applied to the lexicon covered in class, with the aim to develop phonological awareness to support the development of their pronunciation. The rest of class-time (roughly 30 min) was dedicated to topics covered in the syllabus. The data reported in this study focus on two of the phonological aspects of French pronunciation covered in instruction and considered important for intelligibility, namely the contrastive vowels /y/-/u/, and liaisons. Students were also exposed to these phenomena in an implicit manner throughout each semester as they were naturally contained in their instructors' input or in the course materials.

Participants were initially recruited from FR1 courses and were then tracked as they pursued FR2, FR3, and FR4. Out of the initial 114 students invited, 101 consented to be part of the study. Data from 41 were discarded as either questionnaires and/or recordings were missing or incorrect, leaving data from 60 students to be used. Following the longitudinal rationale of the study, a recruitment effort was made to retain these 60 participants to track them as they moved up in the FL requirement sequence. As expected though, attrition occurred. Out of the original 60, 34 consented to continue participating in FR2, 25 in FR3, and 13 in FR4.

All participants (N = 60) completed a language background questionnaire. They were all born after 1995 and were 18.75 years old on average (SD = 1.05), with 30 males and 30 females. Seventeen participants reported being non-L1 English speakers speaking various L1s (Mandarin, Spanish, Arabic, Japanese). In the remainder of the paper, the term L2 will be used as an umbrella term to refer to the target language under study.

The treatment group was recruited using convenience sampling. Participants who were enrolled in six sections of French 1 comprised the Treatment group ($n = 41$). There were 22 females and 19 males. They were 18.79 years old on average ($SD = 0.976$). Subsequently, the participants from the initial Treatment group who had elected to participate in the study were contacted and recruited separately to be part of the study in FR2, FR3, and FR4. Of the original 41 in FR1, twenty participants were retained in FR2, twelve in FR3, and four in FR4.

Students in the remaining two sections of FR1 that met earliest in the day comprised the initial Comparison group ($n = 19$). There were 8 females and 11 males. They were 18.71 years old on average ($SD = 0.955$). Of the original 19 in FR1, fourteen participants were retained in FR2, thirteen in FR3, and nine in FR4.

An independent *t*-test and Chi-square tests of independence showed that there was no significant difference between groups at all levels of instruction in terms of age, gender, L1, and language background.

3.2. Procedure

Data collection on /y/-/u/ and liaisons happened at week 10 and 15 in FR 1, at week 13 and 15 in FR 2, at week 4 and 9 in FR 3, and at week 4 and 9 in FR 4.

3.2.1. Read-Aloud Tasks

Read-aloud tasks were chosen since that according to [Saito and Plonsky \(2019\)](#), such controlled speech tasks will consolidate learners' declarative knowledge (i.e., what they learn from instruction). Furthermore, these tasks were found to be the most common form of assessment of pronunciation in [Thomson and Derwing's \(2015\)](#) meta-analysis on the effectiveness of L2 pronunciation instruction. The read-aloud tasks were created and designed for each different level to be incrementally difficult and to balance the number as well as context of tokens per feature (/y/-/u/ and the liaisons).

For /y/ and /u/, the corpus was chosen to target words that appear in as many phonetic contexts as possible, in a variety of places and mode of articulation, so that the results could be generalized over the /y/-/u/ contrast. For the liaisons, the read-aloud tasks mixed contexts so that all possible forms be potentially produced: (1) (mis)categorization of liaisons (e.g., treating a forbidden liaison as a mandatory one), (2) (mis)pronunciation of the liaison consonant due to non-transparent patterns (e.g., an orthographical "d" is pronounced /t/ in liaison), and (3) (lack of) resyllabification across the two words involved (e.g., *les amis* 'the friends' pronounced [lez/a/mi] instead of [le/za/mi]).

Participants in both groups completed the pre- and post-test production at home before and immediately after each lesson. Participants were instructed they could not use any extraneous material or resources to complete the tasks. The audio-recordings of these pre- and post-tests were collected online via the university's Learning Management System (Moodle).

3.2.2. Self-Reflection Questionnaires

According to [Mackey and Gass \(2016\)](#), "one important source of data in the field comes from what learners themselves say about what they know or about how they process their L2, also known as introspective verbal reports" (p. 1). At each pre- and post-test production, immediately after the read-aloud task, the Treatment group were instructed to fill out a self-reflection questionnaire—a type of verbal reports—in English. The questionnaires were available via a secure form sent out via email to the participants. They were instructed to answer a minimum of three sentences for each question.

Two different self-reflection questionnaires—one for the pre-test, and the other for the post-test—were designed with sufficient similarity to foster comparison of data. In designing the questionnaire, the researcher attempted to probe students' self-reflection on their learning growth and learning process. Each questionnaire (pre- and post-) was purposefully designed to be both non-reactive (i.e., self-reflections happened after stu-

dents' performance, to avoid task effect) and veridical (they capture participants' thoughts through guided questions and by encouraging them to write a minimum of three sentences per response) (cf. Egi 2008). Each questionnaire comprised nine closed and open-ended questions, that prompted students to reflect on their current state of learning, their growth, and their ability to perceive and/or produce the phenomenon under study (Raoufi et al. 2013). Through explicit elicitation of attention, noticing, introspection, and understanding (e.g., What have you noticed about your own pronunciation in French as a result of the lesson on obligatory and forbidden liaisons?), the researcher hoped to elicit learners' procedural knowledge (i.e., the extent to which learners were able to access their pronunciation knowledge) (Saito and Plonsky 2019).

3.3. Analyses

3.3.1. Ratings

Raters were selected based on two criteria: (1) demonstrating nativelike command of French, and (2) being accustomed to FL French learners. To better represent the classroom setting in which the study took place, non-native speakers of French were recruited as raters. Indeed, it is often the case that professors, instructors, or teaching assistants (e.g., in large universities) are non-native speakers of French. The chosen raters reflected such setting.

Participants' intelligibility on the production of /y/-/u/ and mandatory liaisons was established through listeners' ratings (Munro et al. 2015). Ten raters evaluated the audio recordings from the production tests. Each sample was rated by 2 independent raters. To ensure reliability, raters underwent one 30-min training session per semester, whether they had previously done ratings for the researcher or not. Each training session consisted of the researcher demonstrating how to rate the audio sample, and raters practicing in front of the researcher. Each audio sample was then randomly assigned to two raters.

Table 2 represents the rating of intelligibility for /y/-/u/. Here, intelligibility was operationalized using words-correct counts, as recognizing words within an utterance is considered to be close to understanding its full meaning (Munro and Derwing 2015; Kang et al. 2018). Raters were instructed to assign of score of 1 or 0 depending on the prompt: Where would your transcription have differed from the script? For liaisons raters had to give a score between 0 and 5 for each liaison taking into account whether the liaison was made or not, the correct or incorrect sound of the liaison, and the resyllabification (or lack thereof) of the liaisons (see rating examples for liaisons in Table 3).

Table 2. Rating instruction for /y/-/u/.

Instructions	As you listen to the recording, please mark on the script any mispronounced 'u' or 'ou': Where would your transcription have differed from the script? E.g. You read "russe", you hear "rousse", your transcription is 'rousse' = mispronounced.	
Rating	Transcription does not match → Mispronounced = 0	Transcription does match → Correct = 1

Due to the ordinal nature of the data, and due to the violations of one or more assumptions of parametric tests, non-parametric statistics were adopted (Plonsky 2015; Sheskin 2011). Following Munro et al.'s (2015) procedure, Mann–Whitney U tests were used to investigate learning outcomes between the two groups, and descriptive statistics were used to investigate learners' growth (or lack thereof) for subsampling purposes. Effect sizes were computed as $d = (M1 - M2)/SD$, with $d = 1.00$ considered a large effect, $d = 0.70$ medium, and $d = 0.40$ small (Plonsky and Oswald 2014).

Following Cumming et al.'s (2012); Larson-Hall's (2016) and Plonsky's (2015) call for changing and advancing the way quantitative research is carried out in second language research, analyses for RQ1 relied on effect sizes and mean differences rather than just p-value statistical significance. Indeed, due to the fact that statistical significance relies mainly on sample size, and that this study had a small sample size ($N = 60$), effect sizes,

which do not change based on sample size, were used to determine “whether there was an effect, how big it was, and how practically important it is” (Larson-Hall 2016, p. 142).

Table 3. Rating instruction for liaisons.

LIAISON HAS TO BE MADE POSSIBLE SCORES:	
0	a liaison that had to be made was not made
2	a liaison that should be made was made BUT: <ul style="list-style-type: none"> • incorrect sound of consonant was produced • AND • Liaison attached to wrong word • e.g., /leʒ a mi/ instead of /le ʒa mi/
3	a liaison that should be made was made AND: <ul style="list-style-type: none"> • correct sound of consonant was produced • BUT • Liaison attached to wrong word • e.g., /leʒ a mi/ instead of /le ʒa mi/
4	a liaison that should be made was made AND: <ul style="list-style-type: none"> • incorrect sound of consonant was produced • BUT • Liaison attached to correct word • e.g., /le ʒa mi/ instead of /le ʒa mi/
5	a liaison that should be made was made AND: <ul style="list-style-type: none"> • correct sound of consonant was produced • AND • Liaison attached to correct word • e.g., /le ʒa mi/

3.3.2. Research Question 1—Learning Outcomes

This study investigated the impact of self-reflection on French learners’ production of the minimal pair of oral vowels /y/-/u/ and liaisons, compared with similar learners who were exposed to explicit instruction only.

Mann–Whitney U tests were conducted on the pre-tests *production* to determine if there were significant differences between the Treatment, and Comparison groups across levels of instruction. In the case of significant differences between the groups at the pre-tests, no further analysis was done. If no significant differences were found at the pre-tests, subsequent post-tests analyses were warranted.

3.3.3. Research Question 2—Learning Process

To explore the role of self-reflections in learners’ phonological awareness process relative to their growth, the second part of the analysis focused exclusively on clusters of the subsample in the Treatment group based on their *production* tests.

Clusters. At any given level of instruction (FR1, FR2, FR3, and FR4), growth and quartiles in each sample were calculated so as to undertake a purposive subsampling. The subsample was chosen if (1) they moved quartiles and (2) if their post-test score was higher than the Treatment group’s mean (based on growth)—high growth makers (HGM), or if their post-test score was lower than the Treatment group’s mean—low growth makers (LGM). For /y/-/u/, the participants with the highest growth rates (HGM) as well as with the lowest growth rates (LGM) were selected to comprise clusters HGMA and LGMA. Similarly, for liaisons, clusters HGMB and LGMB were formed. Due to the longitudinality of the study and the available data from the participants, clusters were not necessarily identical in numbers across levels: there were ten in FR1, ten in FR2, six in FR3. In FR4, four participants completed the study in its entirety.

Coding. Responses to self-reflections accompanying both pre- and post-tests were examined for each participant in each cluster and for each variable.

Holistic analyses of the number of comments. In order to get a clear representation of the number of statements generated by HGMs and LGMs, the number of comments generated per semester, per subgroups and per type of statements (attention, noticing, introspection, understanding) were counted, and the proportions calculated. Data were first segmented by units of meaning (Strijbos et al. 2006), leading to the identification of 330 codable units in FR1, 240 codable units in FR2, 246 codable units in FR3, and 97 codable units in FR4, or a total of 913 codable units.

Thematic qualitative analysis of the content of the comments. Following the holistic analyses of the number of comments generated, the content of the comments was thematically analyzed, and three major themes were identified (1) awareness of progress with or without awareness of process and (2) phonological and linguistic awareness. Two coders independently coded each unit in each dataset, using predetermined codes on attention, noticing, introspection, and understanding based on Wrembel's (2011; 2015) studies, Leow's (2015) model on depth of processing, and the Author's (Meritan and Mroz 2019) Awareness Continuum (Table 4). Inter-coder reliability was determined by the percentage of agreement between the two independent codings. The two coders initially reached 97% agreement, and then met to discuss any disagreement until 100% agreement was reached. For each level of instruction, participants' statements thus coded were then compiled in a table to allow for quantitative as well as qualitative comparisons of HGMs and LGMs, and to foster the emergence of themes of interest related to their awareness process. Following this process, the self-reflective statements of the FR4 participants were also compared through time if and when they were previously in LGM and/or HGM clusters.

Table 4. Operationalization of Depth of Processing and coding procedure for participants' self-reflections (adapted from Wrembel (2013, 2015); Leow (2015), and Meritan and Mroz (2019)).

	Low Depth of Processing	Medium Depth of Processing	High Depth of Processing	
Level of Awareness	Attention	Reporting/Noticing	Introspection/Metacognition	+/- Understanding
Description	Signs of active effort but no potential for processing target form.	Signs of pronunciation difficulties regarding production/performance.	Signs of self-reflection on learning process and FL acquisition. Signs of use of prior knowledge (PK).	Signs of hypotheses formation regarding phonological rules, articulatory explanations and examples (correct or not).
Descriptors	Use senses verbs such as listen, see, or hear; mentioning focus, concentration, and paying attention. Employ low level of cognitive effort.	Use verbs such as distinguish, realize, notice, differentiate, identify, classify, collect, find. Say they are not sure what it is. Still low level of cognitive effort.	Mentions words such as aware, intuition, conscious; and accurate reflection on progress. Makes comments referring to prior knowledge (remember, use, memory, etc.). Some level of cognitive effort to process target item phonologically.	Verbs such as understand, incorporate, know, and <i>potentially</i> give the <u>correct</u> underlying phonological rule. High level of cognitive effort to process target item phonologically.

FL = foreign language, +/- = correct or incorrect underlying rule.

4. Results

The aim of this study was twofold: (1) investigate the impact of self-reflection on French learners' production compared to students without self-reflection, and (2) explore the role of self-reflection in phonological awareness-raising in the learners' learning process.

4.1. Learning Outcomes

Research question 1 investigated the impact of self-reflection (Treatment group) on French learners' production of the minimal pair of oral vowel /y/-/u/ and the use of liaisons, compared to similar learners who did not receive self-reflection questionnaires (Comparison group). Mann–Whitney U tests were run to determine if there were any significant differences between groups. Overall, the Treatment group outperformed the Comparison group on both their production of /y/-/u/ and the correct realization of liaisons.

Table 5 presents the analysis of learning outcomes on production via descriptive statistics, effect sizes, and statistical significance. Effect sizes were computed as Cohen's $d = (M1 - M2) / SD$ with $d = 1.00$ considered a large effect, $d = 0.70$ medium, and $d = 0.40$ small (Plonsky and Oswald 2014).

Table 5. Comparative results on the post-tests production of /y/-/u/ and liaisons across semesters.

Mann–Whitney U Tests on Post-Tests Production of			N	p	Cohen's d	M	SD	Treatment (T)			Comparison (Cmp)			Mean Diff.
								N	Mean	SD	N	Mean	SD	
FR1	Seg	/y/	54	0.718	0.07 (0)	78.70	20.10	38	79.13	20.38	16	77.68	20.05	1.45 (T > Cmp)
		/u/	60	0.002 *	0.95 (L)	91.74	8.93	41	94.21	7.68	19	86.40	9.30	7.81 (T > Cmp)
	Supraseg	Liaisons	51	0.038 *	0.69 (M)	66.94	22.81	38	70.81	23.33	13	55.61	17.41	15.20 (T > Cmp)
FR2	Seg	/y/	32	0.799	0.12 (0)	71.74	18.82	20	72.21	16.08	12	70.14	18.62	2.57 (T > Cmp)
		/u/	32	0.338	0.56 (S)	84.44	10.63	20	86.63	6.06	12	80.77	15.21	5.86 (T > Cmp)
	Supraseg	Liaisons	33	0.319	0.49 (S)	63.71	13.92	20	61.07	16.06	13	67.77	8.87	6.70 (Cmp > T)
FR3	Seg	/y/	17	0.434	0.48 (S)	70.03	15.43	7	74.43	13.87	10	66.96	16.42	7.47 (T > Cmp)
		/u/	17	0.116	0.95 (L)	72.81	22.87	7	61.03	25.24	10	81.05	17.93	20.02 (Cmp > T)
	Supraseg	Liaisons	25	0.644	0.17 (0)	60.05	12.44	12	58.92	14.07	12	61.10	11.19	2.18 (Cmp > T)
FR4	Seg	/y/	11	0.558	0.20 (S)	82.25	18.08	4	79.76	8.13	7	83.67	22.45	3.91 (Cmp > T)
		/u/	11	0.848	0.06 (0)	75.75	17.62	4	75.00	9.62	7	76.19	21.67	1.19 (Cmp > T)
	Supraseg	Liaisons	13	0.315	0.48 (S)	61.01	9.78	4	64.28	9.25	9	59.56	10.17	4.72 (T > Cmp)

* Significant at $p < 0.05$.

Figure 2 represent the post-test mean scores of the Treatment group and the Comparison group across phonological features and semesters.

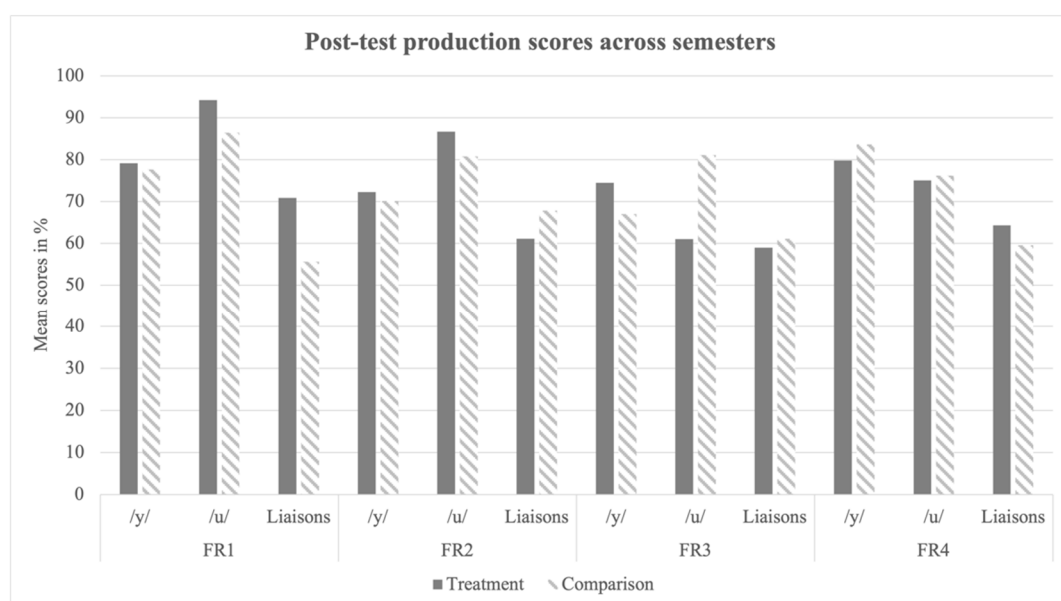


Figure 2. Mean scores on the post-test production tests across four semesters.

In FR1, the Treatment group significantly outperformed the Comparison group on the liaisons ($p = 0.038$, $M_T = 70.81$, $M_{Comp} = 55.61$, $d = 0.69$) with a medium effect size, and on /u/ ($p = 0.002$, $M_T = 94.21$, $M_{Comp} = 86.40$, $d = 0.95$) with large effect size, and outperformed the Comparison on /y/ albeit non-significantly with no effect ($p = 0.718$, $M_T = 79.13$, $M_{Comp} = 77.68$, $d = 0.07$). Here, self-reflection played a significant role on learners' production of the liaisons and /u/, but less so for /y/, where learners seemed to have performed similarly.

In FR2, the Treatment group outperformed the Comparison group on both /y/ ($p = 0.799$, $M_T = 72.71$, $M_{Comp} = 70.14$, $d = 0.12$) with no effect and /u/ ($p = 0.338$, $M_T = 86.63$, $M_{Comp} = 80.77$, $d = 0.56$) with a small effect size, showing the effect—even though minor—of self-reflection on segmentals. Regarding the liaisons, the Comparison group outperformed the Treatment group non-significantly with a small effect size ($p = 0.319$, $M_T = 61.07$, $M_{Comp} = 67.77$, $d = 0.49$).

In FR3, the Treatment group outperformed the Comparison group on /y/ with a small effect size ($p = 0.434$, $M_T = 74.43$, $M_{Comp} = 66.96$, $d = 0.48$). Here, it appears that the Comparison group is catching up to the Treatment group as they outperformed on /u/ with a large effect size ($p = 0.116$, $M_T = 61.03$, $M_{Comp} = 81.05$, $d = 0.95$). They slightly outperformed the Treatment group on the liaisons as well, with no effect (i.e., participants performed similarly) ($p = 0.644$, $M_T = 58.92$, $M_{Comp} = 61.10$, $d = 0.17$). Self-reflection was less impactful in FR3.

In FR4, the results must be interpreted with caution due to the small number of participants. Overall, the Comparison group outperformed the Treatment group on both /y/ ($p = 0.558$, $M_T = 79.7681$, $M_{Comp} = 83.67$, $d = 0.20$) and /u/ ($p = 0.848$, $M_T = 75.00$, $M_{Comp} = 76.19$, $d = 0.06$) with little to no effect. In fact, for /u/, the mean difference between the groups was only 1.19. The Treatment group, however, outperformed the Comparison on the liaisons ($p = 0.315$, $M_T = 64.28$, $M_{Comp} = 59.56$, $d = 0.48$) with a small effect size.

4.2. Learning Process

4.2.1. Subsampling and Coding

The study also explored the role of awareness-raising on the production growth made by a subsample—the highest growth makers (HGMs) and lowest growth makers (LGMs)—from the Treatment group using data drawn from their responses to the open-ended questions to the self-reflection questionnaire. Across levels of instruction (FR1, FR2, and FR3), the participants in the subsample were selected as they met the following criteria: (1) the participants had to have moved up (HGM) or down (LGM) one quartile, and (2) their post-tests scores were above (HGM) or below (LGM) the Treatment group's post-test mean. However, subsampling in FR3 related to /y/-/u/ led the researcher to revise the HGM–LGM subgroups. Indeed, only one participant was considered a *true* HGM and only one was considered a *true* LGM as they met the criteria for selection. Conversely, two participants belonged to the HGM subgroup for /y/ but to the LGM subgroup for /u/, as their scores declined in their production of /u/. Finally, regarding subsampling in FR4, with only four participants retained in the Treatment group at this point in the longitudinal study, all four FR4 participants' statements were analyzed.

Chi-square tests of independence revealed that there were no significant differences between the clusters (high growth makers vs. low growth makers) on /y/-/u/, and on liaisons in terms of gender, L1, and previously learned languages.

4.2.2. Holistic Results on Number of Comments Generated

FR1. Figure 3 represents the proportion of comments generated by FR1 per subgroup (HGM and LGM), per type of statements, and per variables.

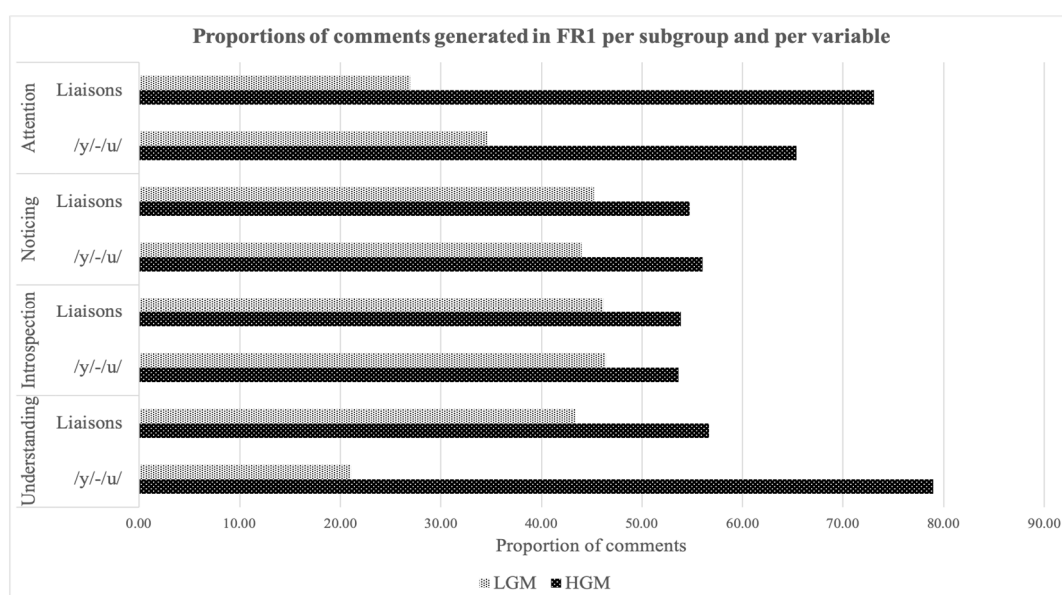


Figure 3. Proportions of comments generated in FR1.

An evaluation of the 330 total statements produced in response to the open-ended questions showed that HGMs generated more comments (58.48%) than LGMs (41.52%), although members of both groups offered similar proportions of comments concerning the two linguistic features under consideration (liaisons vs. /y/-/u/). However, the analysis of the *type* of statements that HGMs and LGMs produced (attention, noticing, introspection, understanding) showed that, whereas both groups offered similar numbers of *noticing* comments (23.64% of all comments, of which 55.13% came from HGMs and 44.87% from LGMs) and *introspection* (45.76% of all comments, of which 54.30% came from HGMs and 45.70% from LGMs), they substantially differed in terms of *attention* (15.76% of all comments, of which 69.23% came from HGMs and 30.77% from LGMs) and *understanding* (14.85% of all comments, of which 65.31% came from HGMs and 34.69% for LGMs).

Overall, HGM participants showed more signs of low, medium, and high levels of processing.

FR2. Figure 4 represents the proportion of comments generated by FR2 per subgroup (HGM and LGM), per type of statements, and per variables.

In total, 240 statements were produced, and overall, HGMs generated more comments (57.08%) than LGMs (42.92%). However, differences in the number of comments generated between the two linguistic features appeared with participants generating more comments concerning the liaisons (60%) than /y/-/u/ (40%). The analysis of the type of statements produced by both groups within each variable also showed differences.

For /y/-/u/, whereas both groups offered similar numbers of comments in terms of *attention* (19.79% of all comments, of which 52.63% came from HGMs and 47.37% from LGMs) and *introspection* (20.83% of all comments, of which 45% came from HGMs and 55% from LGMs), they differed substantially in terms of *noticing* (48.96% of all /y/-/u/ comments, of which 65.96% came from HGMs and 34.04% for LGMs) and *understanding* (10.42% of all comments, of which 60% came from HGMs and 40% from LGMs). Here, LGMs showed more signs of medium depth of processing compared to HGMs but the HGMs showed more signs of low and high depth of processing.

For the liaisons, both groups produced a similar number of statements in terms of *attention*, (14.58% of all comments, of which 52.38% came from HGMs and 47.62% from LGMs), *noticing* (38.19% of all comments, of which 54.55% came from HGMs and 45.45% from LGMs) and *understanding* (22.22% of all comments, of which 53.13% came from HGMs and 46.88% from LGMs), but differed substantially in terms of *introspection* (25.01% of all comments, of which 63.89% came from HGMs and 36.11% from LGMs). Contrary to what

was found for /y/-/u/, here, the HGM participants showed more signs of low, medium, and high depth of processing compared to LGMs.

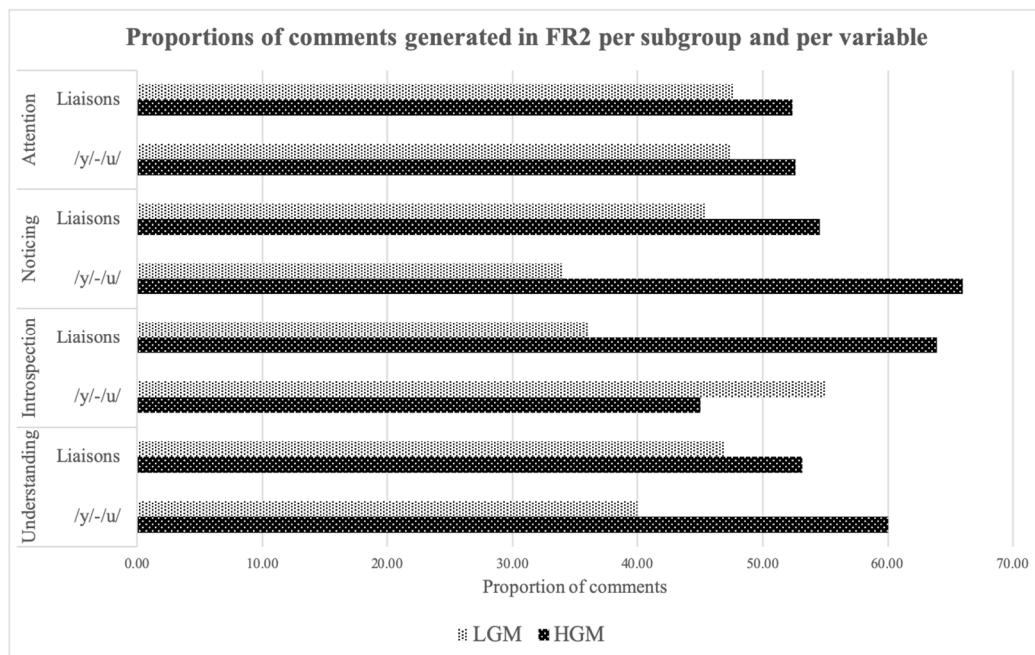


Figure 4. Proportions of comments generated in FR2.

FR3. Figure 5 represents the proportion of comments generated by FR3 per subgroup (HGM and LGM), per type of statements, and per variables.

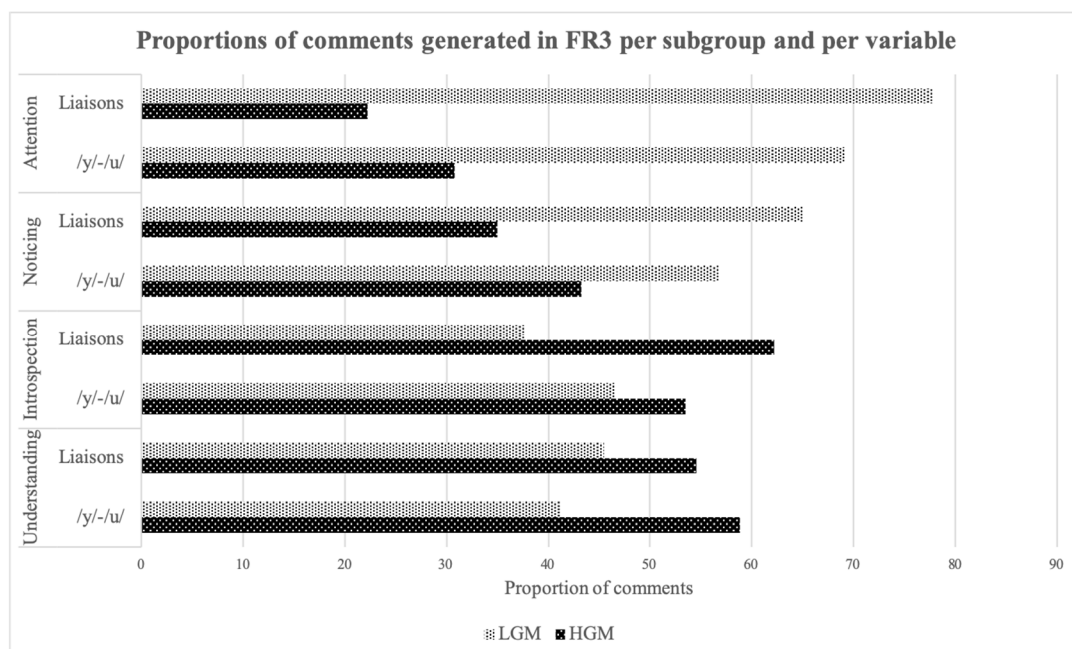


Figure 5. Proportions of comments generated in FR3.

In total, 246 statements were generated by the participants and the evaluation of these statements revealed that contrary to FR1 and FR2, in FR3, HGMs and LGMs produced similar proportions of comments (47.56% for HGMs and 52.44% for LGMs), with LGMs producing slightly more.

Furthermore, members of both groups offered similar proportions of comments on the two linguistic features under consideration. The analysis of the *type* of statements that HGMs and LGMs produced, however, showed that they substantially differed in terms of *attention*, *noticing*, *introspection*, and *understanding*. For both linguistic features, LGMs provided more comments categorized as *attention* (12.60% of all comments, of which 74.19% came from LGMs and 24.81% from HGMs) and *noticing* (31.30% of all comments, of which 61.04% came from LGMs and 38.96% from HGMs).

However, HGMs generated more comments than LGMs in terms of *introspection* (35.77% of all comments, of which 66.23% came from HGMs and 33.77% from LGMs), and *understanding* (20.33% of all comments, of which 56% came from HGMs and 44% from LGMs). Here, LGM showed more signs of low depth of processing, whereas HGMs showed more signs of medium and high depth of processing.

FR4. Figure 6 represents the proportion of comments generated by FR4 per type of statements and per variables.

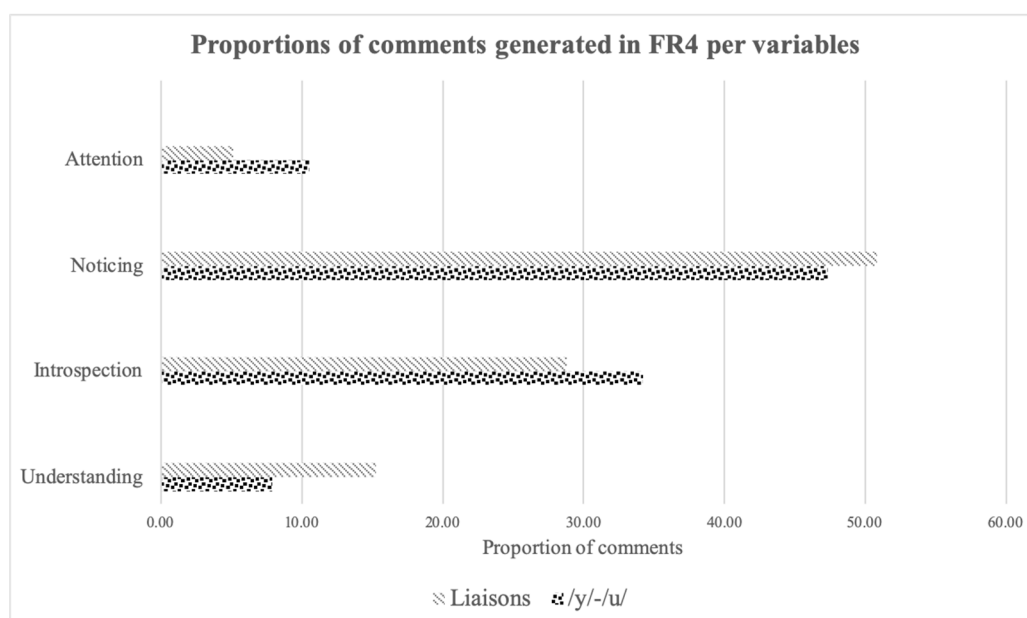


Figure 6. Proportions of comments generated in FR4.

In total, 97 statements were generated by the participants. Since only four participants remained, the distinction between HGM and LGM was not warranted, and a more holistic view is represented.

Differences appeared between the linguistic features under study. There were 22% more comments generated for liaisons (60.82% of all comments) than for /y/-/u/ (39.18% of comments). Comments also substantially differed in terms of *attention*, *noticing*, *introspection*, and *understanding* between liaisons and /y/-/u/.

For /y/-/u/, participants generated more comments regarding *noticing* (49.48% of all comments), and *introspection* (30.93% of all comments). A total of 10.52% of the comments generated were about *attention*, and 7.89% of the comments generated were related to *understanding*.

For liaisons, a similar pattern was found as most comments generated were related to *noticing* (50.85% of all comments), and fewer were related to *introspection* (28.81% of all the comments). A total of 5.09% of all comments generated were related to *attention*, and 15.25% were related to *understanding*. In FR4, there were more signs of low depth of processing (attention + noticing) compared to medium and high depth of processing, confirming that at a certain point, high depth of processing is no longer needed.

The subsequent analysis of these statements across levels revealed three major themes related to phonological awareness: (1) awareness of progress¹ with or without awareness of process, (2) phonological awareness at the level of *noticing* or *understanding* and (3) awareness and use of prior knowledge.

4.2.3. Awareness of Progress with or without Awareness of Process

From FR1 to FR3, all HGMs and LGMs offered qualitative reflections of their progress that were congruent with raters' quantitative evaluation. Tables 6–8 represent participants' growth from pre- to post-test at FR1, FR2, and FR3 for each subgroup on all three phonological features.

Table 6. Participants' pronunciation growth from pre–post-tests in FR1.

	FR1 High Growth Makers (HGM)% Growth from Pre- to Post-Tests					FR1 Low Growth Makers (LGM)% Growth from Pre- to Post-Tests				
Participants #a	111	118	122	135	150	123	131	154	157	159
/y/	21.5	17.8	35.7	7.1	28.6	3.6	−3.6	3.6	7.2	7.1
/u/	8.3	0	0	16.7	0	−16.7	4.1	8.3	0	0
Participants #b	127	131	137	143	152	115	138	141	148	158
Liaisons	40.9	18.2	18.2	22.7	63.6	−18.1	−13.7	−18.1	9	0

Table 7. Participants' pronunciation growth from pre–post-tests in FR2.

	FR2 High Growth Makers (HGM)% Growth from Pre- to Post-Tests					FR2 Low Growth Makers (LGM) % Growth from Pre- to Post-Tests				
Participants #a	119	123	124	135		103	118	122	138	
/y/	16.67	8.33	22.91	16.67		−2.08	8.33	0	4.17	
/u/	17.31	1.93	1.92	5.77		6.73	3.85	5.77	3.85	
Participants #b	111	119	135	142	152	108	123	138	141	151
Liaisons	13.81	11.43	10.48	20.48	10.95	−6.19	2.38	−6.67	2.38	0.95

Table 8. Participants' pronunciation growth from pre–post-tests in FR3.

	FR3 High Growth Makers (HGM)% Growth from Pre- to Post-Tests			FR3 Low Growth Makers (LGM)% Growth from Pre- to Post-Tests		
Participants #a	103	123 (/y/)	135 (/y/)	123 (/u/)	135 (/u/)	146
/y/	26.08	7.97	7.24	7.97	7.24	7.07
/u/	10.53	−41.23	−41.23	−41.23	−41.23	−28.07
Participants #b	111	135	138	103	141	146
Liaisons	9.31	5.86	25.17	−8.28	−3.10	−1.38

Participant 152-HGMb (FR1), for instance, considered that they were “successful”. Several HGMs across semesters (FR1-111-HGMa, FR1-118-HGMa, FR2-135HGMa, FR2-119HGMa, FR3-103HGMa) mentioned improving their pronunciation of /y/ and/or /u/. Several HGMs also discussed their improvement of liaisons. For instance, FR2-135HGMb, and FR2-119HGMb felt “successful”. Similarly, LGMs' self-evaluations that mentioned making small or even no progress were confirmed by raters' evaluations. Participants who made no improvement mentioned their difficulties and lack of progress (FR1-131-LGMa /y/, FR1-158-LGMb, FR2-138-LGMb, FR2-151-LGMb), while participants who showed little improvement discussed their slight but insufficient progress (FR1-157-LGMa /y/, FR1-159-LGMa /y/, FR3-146-LGMa, FR3-146-LGMb), all showing signs of low depth of processing.

Example statements of awareness of progress generated in the self-reflection include the following:

- *I can produce [liaisons] a little better (FR1-148-LGMb).*

- *I have improved the slightest (FR3-146LGMb).*

Table 9 represents participants' scores from pre- to post-test at FR4. In FR4, all four participants accurately assessed their progress on liaisons compared to raters' evaluations.

Table 9. Participants' pronunciation growth from pre-post-tests in FR4.

FR4 High Growth Makers (HGM)% Growth from Pre- to Post-Tests				
Participants #a	122	123	135	146
/y/	45.24	4.76	28.57	47.62
/u/	9.26	11.11	−1.85	3.70
Liaisons	1.32	10.26	2.89	1.84

Example statements of awareness of progress (or lack thereof) generated by FR4 participants in the self-reflection include the following:

- My performance was better than where [it was] last year or even last month" (10.86% growth from pre- to post-test) (FR4-123-liaisons).
- I didn't improve because I am more confused (2.60% growth in FR3, and 1.84% growth in FR4) (FR4-146-liaisons).
- My pronunciation has improved (45.24% growth on /y/ and 9.26% growth on /u/) (FR4-122).
- This form of pronunciation was still very hard for me (18.24% growth on /y/ and 1.85% decline on /u/) (FR4-135).
- I didn't think I improved as much (despite 47.62% growth on /y/ and a 7.01% growth on /u/) (FR4-146).

However, differences appeared in the way the two clusters of participants (HGMs and LGMs) characterized their mental states and processes and their ability to show signs of medium depth of processing. Across semesters, HGMs' comments included an abundance of descriptors such as "intuition", "conscious", "instinctual", "unconsciously", "subconsciously", "instincts", "implicit knowledge", and "aware" (medium depth of processing).

The following comments illustrate selected instances of HGMs self-reflection:

- I am used to producing liaisons unconsciously now (FR1-137HGMb).
- It has become more like language intuition (FR1-122-HGMa).
- Liaisons has become more of a background feature, and I finally hit my threshold of practice where it stopped being difficult (FR2-111HGMb).
- I am more aware of these forms (FR2-135HGMa).
- I can find it without thinking about it as it has become implicit knowledge as I subconsciously know about liaisons (FR3-111HGMb).

Conversely, LGMs' comments showed less signs of medium levels of processing and, in most cases, an absence of introspection concerning the nature of their mental states and processes. FR2-108LGMb, however, described feeling "more confused", explaining that "maybe that mean[t] that [they] were over-applying the liaison before the lesson, and now [they were] more cautious using it".

In FR4, all participants provided comments regarding their mental states and explained how they were more "aware" (135b, 146b, 146a). Participant 123b explained that liaisons came "pretty naturally", and 146 discussed how they went "with [their] instincts and [their] past". Similarly, 123a described being "self-conscious".

4.2.4. Phonological Awareness

Taken as a whole, the *noticing* and *understanding* comments that HGMs and LGMs offered concerning their developing, or stagnant, pronunciation skills differed, particularly in the way they attempted to activate their prior knowledge and formulate phonological rules and articulatory explanations. HGMs' comments revealed an increase in confidence and activation of prior knowledge leading to high levels of processing. These indications of use of prior knowledge reflect a high depth of processing, that is, students showed levels

of cognitive effort and were able to self-report and self-reflect on their learning process. The following comments represent HGMs' statements regarding their ability to activate prior knowledge:

- It's become more of a background feature I don't have to think too hard about it (FR2-111HGMb).
- Going over liaisons helped refresh my memory about them a little bit (FR2-135HGMb).
- I find it without thinking about it as it's become implicit knowledge (FR3-111HGMb).
- I already knew most of the information since we went over liaisons last year in French class (FR3-135HGMb).
- I am more aware since I remember we practiced it last semester too (FR4-146a).
- I go with my instincts and my past (FR4-146b).
- I know that I learned liaisons before, it's language experience (FR4-122b).

Their comments were further embedded into descriptions of an *understanding* process whereby participants tried to verbalize explanations for each variable.

The following examples of self-reflective statements highlight HGMs' higher level of processing:

- I am able to better shape my mouth and pronounce /u/ and /y/ sounds (FR1-118HGMb).
- u as in vous is a little farther back in your mouth (FR1-150HGMb).
- [u] has a lower sound and it is made by rounding our lips (FR2-135HGMA).
- When the letters o and u are together, they create the /u/ sound (FR3-135HGMA).
- The /y/ as in tu is a sound made with the front portion of your mouth whereas the /u/ as in vous is made with the back of your mouth/throat (FR3-123HGMA).
- Liaisons happen when the last letter of a word ends in a S, T, or N and the beginning of the word that follows begins with a vowel or H (FR2-135HGMb).
- There are cases when the liaison is forbidden and knowing the grammatical grouping of two words has an effect on whether or not there is a liaison (FR3-111HGMb).

Furthermore, HGMs' comments made frequent use of verbs such as "understand", "know", "incorporate", and "classify", showing that they focused as much on their phonological understanding as on their performance.

Conversely, LGMs showed little to no signs of activation of prior knowledge or high levels of processing and discussed their underdeveloped pronunciation skills and difficulties recognizing phonological aspects. No real descriptions of phonological patterns were generated for /y/ and /u/, and if they were, they were generally incorrect. Regarding liaisons, there were some discrepancies between participants.

The following are examples of self-reflective statements generated by LGMs and illustrating incorrect rule formation or low levels of processing:

- One of them is more like the English "e" sound (FR2-103LGMA).
- I make liaisons by looking for consonants and vowels (FR2-138LGMb).
- I look at the last letter of each word I've pronounced (FR3-146LGMb).
- Liaisons involve the word ending with a vowel and the next starting with a vowel (FR3-103LGMb).
- There is a liaison when the first word ends in a consonant, and the next words begins with a vowel, but I don't understand when/why there are exceptions (FR2-108LGMb).

Furthermore, LGMs showed no real sign of awareness at the level of *understanding*; rather, they mainly focused on their performance, staying at lower levels of depth processing, mentioning their numerous mistakes and the need for them to practice more, rather than on a learning process that could lead to them understanding the phonological aspects. For instance, FR3-146LGMb and FR3-103LGMb discussed their lack of confidence.

The following comments illustrate selected instances of LGMs' lack of awareness at the level of understanding:

- I am not very fluid when making liaisons (FR1-141LGMb).
- There are some words I struggle with (FR1-123-LGMA).

- I hit a plateau as of late (FR2-103LGMa).
- I don't understand when the exceptions happen (FR2-108LGMb).

In FR4, participants mentioned their ability to identify and distinguish liaisons despite their struggle to precisely “[know] whether there [was] a liaison” (146b). Furthermore, they were all able to give the correct rules for liaisons. For instance, 135b explained that “you either add a Z, or T sound at the beginning of the following word”, and 146b learned that “there are more exceptions and there are more specific words that have liaisons in this lesson [compared to the FR3 lesson]”. Segmentally, despite participants “knowing what to do with [their] tongue to pronounce each sound” (122a) and understanding “the /u/ sound occurs when o and u are next to each other” (135a) or that “[...] the /u/ sound has a lower tone and the /y/ sound has a higher tone” (146a), they also expressed their difficulties and struggle producing these phonemes (123a, 135a).

These findings also illustrate some participants' gap between their *controlled pronunciation knowledge* and *spontaneous pronunciation knowledge*, while others are trying to bridge this gap by using their prior knowledge (Saito and Plonsky 2019).

5. Discussion

The aim of this study was twofold: (1) to longitudinally investigate the impact of self-reflection on learners' production compared to learners not using self-reflection across four semesters of French pronunciation acquisition, and (2) to longitudinally explore the role of self-reflection and phonological awareness-raising on the learning process.

5.1. Impact of Self-Reflection on Learning Outcomes

This study established that across levels of instruction, students who received self-reflection activities were at an advantage on production outcomes compared to those who did not engage in those activities.

Specifically, students in the Treatment group outperformed those who did not engage in self-reflection on the production of /u/ in FR1 and FR2, and on the production of /y/ in FR1, FR2, and FR3. They also outperformed the other group on liaisons in FR1 and FR4. Comparatively, explicit instruction alone led to better performance on liaisons overall in FR2 and FR3. Moreover, students who only received explicit instruction were able to catch up to the Treatment group's performance in FR3 on /u/ and no earlier than FR4 for /y/.

More specifically, in FR1, a large effect was found in favor of the Treatment group on /u/, but no effect was found for /y/ as they outperformed the Comparison group with 95% and 7% of variance explained, respectively, by group membership. There was also a medium effect in favor of the Treatment group on liaisons, with 69% of variance explained by group membership. The medium to large effect sizes found on /u/ and the liaisons suggest that self-reflection was beneficial to pronunciation learning and fostered acquisition of these two variables. However, there was no effect of self-reflection on the acquisition and production of /y/.

In FR2, students in the Treatment group outperformed those who were only exposed to explicit instruction on /u/ with a small effect, with 56% of variance explained by group membership. Despite the Treatment group slightly outperforming the Comparison group, there was no effect for group on /y/ as they were all able to produce /y/ similarly. Comparatively, students exposed to instruction only outperformed the Treatment group on liaisons with a small effect, with 49% of variance explained by group membership. In French 2, overall, self-reflection had less of an effect when compared to participants who did not participate in self-reflective activities. Furthermore, the small effect found on liaisons was achieved by the comparison group, without self-reflection. It seems that from FR1 to FR2, they were able to catch up to the Treatment group.

In FR3, students in the Treatment group outperformed those who were only exposed to explicit instruction on /y/ with a small effect, with 48% of variance explained by group membership. The Comparison group outperformed the Treatment group on /u/ with a large effect, with 95% of variance explained by group membership. Regarding liaisons,

no effect was found, despite the Comparison group slightly outperforming the Treatment group, with 17% of variance explained by group membership. During this third semester of French learning, the large effect size found on /u/ seems to indicate that the Comparison group benefited from not receiving self-reflection. We can assume here that self-reflection impeded the Treatment group's performance, potentially due to a cognitive overload, and/or overgeneralization of /y/.

In FR4, the Treatment group outperformed the comparison group on liaisons with a small effect with 48% of variance explained by group membership. The Comparison group caught up to the Treatment group on /y/ with a small effect, with 20% of variance explained by group membership. No effect was found on /u/, with 6% of variance explained by group membership, in favor of the Comparison group. It can be argued here that in FR4, the Comparison group caught up on all variables, as similar means and small to no effects were found on all post-tests. It took the Comparison group longer than the Treatment group to improve their production of the variables under study.

These findings on self-reflection add to the body of research on the positive role that self-reflection and phonological awareness can play in terms of language learning outcomes (Derwing 2018; Guion and Pederson 2007; Kivistö-de Souza 2017). Indeed, it was found overall that students who engaged in self-reflection had better production outcomes than students who did not. This confirms that self-reflection can help learners produce features under study more intelligibly and can also help them become more aware of their pronunciation difficulties and lead to improved phonological awareness (Derwing 2018).

These results also confirm two recent studies. First, it confirms Kennedy et al.'s study (2014), in which reflective journals were found to positively impact learners' pronunciation of segmental and suprasegmental features. Second, it confirms Inceoglu's (2021) study, which revealed an association between learners' improvement and their reflections on pronunciation learning. To go deeper in the analysis of students' self-reflection related to their improvement, the current study explored students' learning process and their optimal use of self-reflection.

5.2. Learning Process and Use of Self-Reflection

The second goal of this study was to explore Leow's (2015) model of the L2 learning process in ISLA as a potentially viable L2 pronunciation theory by longitudinally examining what self-reflections could unveil about the role of phonological awareness-raising in learners' development of pronunciation skills. It also aimed to determine under which conditions self-reflections were most beneficial.

Across semesters, the consistency between students' observations and expert raters' quantitative evaluations showed that overall, both high- and low-growth students were able to accurately assess their own progress. This supports previous studies on self-reflection, which found an alignment between the use of self-reflection and learners' improvement on pronunciation production (Kennedy et al. 2014; Inceoglu 2021). However, the current study highlights differences with these two studies. Indeed, the first investigated intermediate L2 French learners in a classroom setting in a second language environment (Québec), while the latter investigated intermediate to advanced L2 French learners in a foreign language environment. Furthermore, data were collected over one semester only. The current study investigated novice to intermediate L2 French learners who were tracked across four semesters in a foreign language classroom setting.

The current study also advances research on "one of the more difficult constructs to operationalize and measure in (...) SLA" (Leow and Donatelli 2017, p. 189) by expanding on the Author's (Meritan and Mroz 2019) *Awareness Continuum* and proposing a Phonological Awareness Continuum framework, extending Leow's (2015) model to pronunciation. This framework was supported by observations that the most prominent difference between high- and low-growth makers were related to the volume of their self-reflective statements, especially in the first two semesters of language acquisition and, more importantly, to the content of their self-reflective statements.

Indeed, differences appeared in the way the two clusters discussed their awareness of progress and their awareness of learning process. During the first semester, while evidence of noticing and introspection were represented in similar proportions across all groups of participants in the subsamples, high-growth makers showed substantially more signs of cognitive effort and understanding. During the second semester, differences appeared within each variable. While evidence of cognitive effort and introspection were represented in similar proportions for /y/ and /u/, high-growth makers showed substantially more signs of noticing and understanding. For liaisons, while evidence of cognitive effort, noticing, and understanding were represented in similar proportions, high-growth makers showed substantially more signs of introspection. During the third semester, a shift in proportion occurred where low-growth makers showed substantially more signs of cognitive effort and noticing. High-growth makers, on the other hand, showed substantially more signs of introspection and understanding. We can argue here that low-growth makers were starting to catch up during the 3rd semester in terms of cognitive effort and noticing—low depth of processing—whereas the high-growth makers had been relying on low depth of processing since the 1st semester to reach medium depth of processing by the 3rd semester. This also confirms that once awareness at the level of understanding is achieved, after practice in the L2 output performance, levels of depth of processing logically decrease, and this high depth of processing is unnecessary (Leow 2015). This indicates that learners are now able to access their previous pronunciation knowledge (i.e., their spontaneous pronunciation knowledge) more automatically (Bergsleithner 2019; Saito and Plonsky 2019).

These findings underscore Schmidt's (1990, p. 144) hypothesis that "those who notice the most, learn the most". Indeed, it was found that it is not just noticing L2 information but further processing of that information that leads to higher depth of processing via self-reflection, metacognition, activation of prior knowledge, hypothesis testing, and rule formation. The data reported here advance that those participants who further processed the L2 data—after logically paying attention and encouraged to self-reflect (high depth of processing)—were the ones who performed faster and better. This sheds light on Moyer's (2017) claim that self-reflection is "an especially effective learning tool that require[s] a deeper level of reflection and noticing" (p. 406), particularly since the current findings suggest that the use of self-reflection can be an optimal metacognitive and cognitive pronunciation learning strategy under certain conditions. Specifically, confirming Leow's (2015) claim that "attentional conditions (. . .) all [lead] to better performances" (p. 217), self-reflection should be used early on in the language learning process to foster students' deeper levels of reflection and noticing, and to foster learners' controlled pronunciation knowledge, which would lead to spontaneous pronunciation knowledge later on (Saito and Plonsky 2019).

Qualitatively speaking, explicit instruction was found to be the precondition which allowed noticing (*low DoP and low level of awareness*) to foster additional cognitive effort (*medium DoP*). This combination then evolved into optimal phonological awareness at the level of introspection and led to awareness at the level of understanding (*high DoP*). This was demonstrated by the fact that not only did high-growth makers reached higher levels of processing faster than low-growth makers, but that low-growth makers were limited to low depth of processing, showing an absence of introspection concerning the nature of their mental states, and a lack of spontaneous pronunciation knowledge. This may be explained by the fact that while low-growth makers did reach deeper level of processing with greater cognitive effort, "the potential for cognitive overload (. . .) may [have] occur[ed] and [led] to misunderstanding and confusion of the underlying rule(s)" (Leow 2015, p. 223). These results add to the body of research on depth of processing (de la Fuente 2015; Hsieh et al. 2015; Leow 2015, 2019; Rosa and Leow 2004), as this study confirms that greater depth of processing is closely aligned with higher levels of performance.

Furthermore, this study shed more light on the relationship between depth of processing and the role of prior knowledge. Indeed, as opposed to low-growth makers,

high-growth makers demonstrated an ability to activate their prior knowledge, thus transforming intake into insight (i.e., gaining cognitive clarity), converting their controlled pronunciation knowledge into spontaneous pronunciation knowledge while reflecting on their “own ongoing (. . .) mental states and processes” (Schwitzgebel 2014, p. 1). High-growth makers were able to make use of this newly acquired insight to focus on particular pronunciation features to hypothesize and test pronunciation rules, noticing similarities and differences between what they had previously learned and what they were currently learning (Pawlak 2010). The differences in performance found in this study, paired with awareness at the level of understanding, are commonly found in high depth of processing (Leow and Mercer 2015). In this way, high-growth makers’ enhanced phonological awareness was associated with deeper metaphonological understanding and greater ability to manipulate the language, leading to more noticeable progress in production (Raofi et al. 2013; Wrembel 2015). It can be argued that their additional cognitive effort might have contributed to increased use of the knowledge they already had, offering support for new knowledge (Saito and Plonsky 2019; Bergsleithner 2019). This attention may have also reinforced learners’ higher-order thinking and led to bridging their controlled pronunciation knowledge with their spontaneous pronunciation knowledge (Glover 2011). In turn, it supported more sustainable intake, leading to insight, and resulting in better learning outcomes. Conversely, low-growth makers were not able to reflect as much on their learning process and could not truly formulate phonological rules. They showed signs of low (and sometimes medium) depth of processing, but never high depth of processing, and indicated that, since the development of their controlled pronunciation knowledge was insufficient, they were not yet fully able to transfer it to spontaneous pronunciation knowledge.

Self-reflection thus has the potential to “encourage awareness of learning as a process” (Glover 2011, p. 132), but only leads to optimal learning outcomes if (1) cognitive effort and attention precedes noticing and if the link between cognitive effort (*attention*) and noticing—i.e., detection—is established, and (2) if the link between low depth of processing and medium depth of processing—i.e., insight—is established as well. Finally, the study adds to previous research (Inceoglu 2021; Kennedy and Blanchet 2014; Kennedy and Trofimovich 2010; Venkatagiri and Levis 2007; Wrembel 2015) by suggesting that only phonologically attentive students developed substantially more metaphonological understanding, higher depth of processing, and greater intelligibility.

6. Conclusions

This study demonstrated that Leow’s (2015) model of the L2 learning process in ISLA, originally built to address grammar and vocabulary acquisition and processes, can extend to L2 pronunciation acquisition. Indeed, the model successfully explains and accounts for several of Foote and Trofimovich’s (2018) learning phenomena, namely, “the importance of input in L2 pronunciation development, the influence of learners’ L1, the significant role of individual differences, and the systematicity and variability of pronunciation development” (pp. 176–77), as found in the different processing stages of Leow’s (2015) model. Furthermore, it was found that self-reflection can promote depth of processing for the development of pronunciation in a classroom setting as well as outside of the classroom. These findings go hand in hand with Leow and Mercer’s (2015) argument which stipulates that, in order to promote and foster thorough learning, students in the classroom must be cognitively engaged to attend to and further process L2 *pronunciation data or information*. This study also supports Kennedy and Blanchet’s (2014) findings that students’ phonological awareness fosters processing of the language. This is particularly important since new generations of students tend to prefer intrapersonal, self-directed, independent forms of learning (Leow and Mercer 2015; Seemiller and Grace 2019), making self-reflection an optimal learning strategy that caters to their learning style and their diversity, and efficiently directs their attention to benefit their pronunciation.

Furthermore, when comparing and contrasting students engaging in self-reflection with those who did not, it seems that self-reflection fostered a more stable learning process on the targeted linguistic features, especially during the first few months of acquisition. It is important to note here that non-linear patterns are not uncommon and are to be expected in language acquisition (Leow 2015).

However, it is important to note that there are currently no studies that investigate students' learning outcomes and phonological awareness over more than a 15-week period (see Inceoglu 2021; Kennedy et al. 2014; Meritan and Mroz 2019; Kennedy and Blanchet 2014 for studies 16 weeks or shorter). The current study thus adds to the body of longitudinal research as it presents data from four 15-week semesters.

Thus, given the ease of adding self-reflective activities to existing explicit pronunciation instruction, it is recommended that students' phonological awareness be raised as soon as they start learning a new language using an integrated approach that combines explicit spelling-to-sound instruction with pronunciation learning strategies in the form of guided, open-ended self-reflections in the context of an otherwise communicative-language-based curriculum (see Miller 2012; Darcy 2018; Meritan and Mroz 2019; Martin 2020a, 2020b; Inceoglu 2021, etc., for examples of explicit pronunciation learning techniques, which could integrate self-reflective elements—some do already). These carefully designed activities can then encourage students to identify beneficial cognitive processes while learning the target language (Leow and Mercer 2015).

Finally, future research should investigate the role (or lack thereof) of prior knowledge while taking into account students' L1 and explore the relationship between prior knowledge and cross-linguistic influences (see Bell Philippa and Gauvin 2020).

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board (or Ethics Committee) of The Office For the Protection of Research Subjects (protocol code 18023, 7 November 2017).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to FERPA and IRB regulations.

Conflicts of Interest: The author declares no conflict of interest.

Note

- ¹ The term “progress” will be used throughout to refer to growth, in order to remain true to participants' language.

References

- Bell Philippa, Véronique Fortier, and Isabelle Gauvin. 2020. Using L1 knowledge about language during L2 error correction: Do students make cross-linguistic connections? *Language Awareness* 29: 95–113. [CrossRef]
- Bergsleithner, Joara M. 2019. The role of prior knowledge in depth of processing during written production. In *The Routledge Handbook of Second Language Research in Classroom Learning*. Edited by Ronald P. Leow. New York: Routledge, pp. 104–18.
- Best, Catherine T. 1995. A direct realist perspective on cross-language speech perception. In *Speech Perception and Linguistic Experience: Issues in Cross-language Research*. Edited by Winifred Strange. Timonium: York Press, pp. 167–200.
- Cerezo, Luis, Alison Caras, and Ron P. Leow. 2016. Effectiveness of guided induction versus deductive instruction on the development of complex Spanish “gustar” structures: An analysis of learning outcomes and processes. *Studies in Second Language Acquisition* 38: 265–91. [CrossRef]
- Craik, Fergus I. M., and Robert S. Lockhart. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11: 671–84. [CrossRef]
- Cumming, Geoff, Fiona Fidler, Pav Kalinowski, and Jerry Lai. 2012. The statistical recommendations of the American Psychological Association Publication Manual: Effect sizes, confidence intervals, and meta-analysis. *Australian Journal of Psychology* 64: 138–46. [CrossRef]
- Darcy, Isabelle. 2018. Powerful and effective pronunciation instruction: How can we achieve it? *The CATESOL Journal* 30: 13–45.

- de la Fuente, María José. 2015. Explicit corrective feedback and computer-based, form-focused instruction: The role of L1 in promoting awareness of L2 forms. In *A Psycholinguistic Approach to Technology and Language Learning*. Edited by Ronald P. Leow, Luis Cerezo and Melissa Baralt. Boston: De Gruyter Mouton, pp. 173–99.
- de Saint Léger, Diane. 2009. Self-assessment of speaking skills and participation in a foreign language class. *Foreign Language Annals* 42: 158–78. [CrossRef]
- Derwing, Tracey M. 2018. The role of phonological awareness in language learning. In *The Routledge Handbook of Language Awareness*. Edited by Peter Garrett and Joseph M. Cots. New York: Routledge, pp. 1–20.
- Derwing, Tracey M., and Murray J. Munro. 2013. The development of L2 oral language skills in two L1 groups: A 7-Year study. *Language Learning* 63: 163–85. [CrossRef]
- Egi, Takako. 2008. Investigating stimulated recall as a cognitive measure: Reactivity and verbal reports in SLA research methodology. *Language Awareness* 17: 212–28. [CrossRef]
- Flege, James E. 1995. Second language speech learning: Theory, findings and problems. In *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Edited by Winifred Strange. Timonium: York Press, pp. 233–77.
- Foot, Jennifer A., and Pavel Trofimovich. 2018. Second language pronunciation learning: An overview of theoretical perspectives. In *The Routledge Handbook of Contemporary English Pronunciation*. Edited by Okim Kang, Ron I. Thomson and John M. Murphy. London: Routledge, pp. 75–90.
- Glover, Philip. 2011. Using CEFR level descriptors to raise university students' awareness of their speaking skills. *Language Awareness* 20: 121–33. [CrossRef]
- Goldman, Alvin I. 2008. *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. New York: Oxford University Press.
- Gordon, Joshua, and Darcy Isabelle. 2016. The development of comprehensible speech in L2 learners. *Journal of Second Language Pronunciation* 2: 56–92. [CrossRef]
- Grant, Anthony M., John Franklin, and Peter Langford. 2002. The self-reflection and insight scale: A new measure of private self-consciousness. *Social Behavior and Personality: An International Journal* 30: 821–35. [CrossRef]
- Guion, Susan G., and Eric Pederson. 2007. Investigating the role of attention in phonetic learning. In *Language Experience in Second Language Speech Learning. In Honor of James Emil Flege*. Edited by Ocke-Schwen Bohn and Murray J. Munro. Philadelphia: John Benjamins Publishing Company, pp. 57–77.
- Hsieh, Hui-Chen, Nina Moreno, and Ronald P. Leow. 2015. Awareness, type of medium, and L2 development: Revisiting Hsieh 2008. In *A Psycholinguistic Approach to Technology and Language Learning*. Edited by Ronald P. Leow, Luis Cerezo and Melissa Baralt. Boston: De Gruyter Mouton, pp. 131–50.
- Inceoglu, Solène. 2021. Exploring the relationship between explicit instruction, pronunciation awareness, and the development of L2 French connected speech processes. *Language Awareness* 30: 336–54. [CrossRef]
- Jha, Amishi. 2017. How to Tame Your Wandering Mind. Video File. Available online: https://www.ted.com/talks/amishi_jha_how_to_tame_your_wandering_mind?language=en (accessed on 15 April 2019).
- Kang, Okim, Ron I. Thomson, and Meghan Moran. 2018. Empirical approaches to measuring the intelligibility of different varieties of English in predicting listener comprehension. *Language Learning* 68: 115–46. [CrossRef]
- Kennedy, Sara, and Josée Blanchet. 2014. Language awareness and perception of connected speech in a second language. *Language Awareness* 23: 92–106. [CrossRef]
- Kennedy, Sara, and Pavel Trofimovich. 2010. Language awareness and second language pronunciation: A classroom study. *Language Awareness* 19: 171–85. [CrossRef]
- Kennedy, Sara, Josée Blanchet, and Pavel Trofimovich. 2014. Learner pronunciation, awareness, and instruction in French as a second language. *Foreign Language Annals* 47: 79–96. [CrossRef]
- Kivistö-de Souza, Hanna. 2017. Brazilian EFL learners' awareness about L2 phones: Is “mall” pronounced as ‘mal’? *Trabalhos em Linguística Aplicada* 56: 235–58. [CrossRef]
- Larson-Hall, Jenifer. 2016. *A Guide to Doing Statistics in Second Language Research Using SPSS and R*, 2nd ed. New York: Routledge.
- Lee, Junkyu, Juhyun Jang, and Luke Plonsky. 2015. The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics* 36: 345–366. [CrossRef]
- Lennon, Paul. 1989. Introspection and intentionality in advanced second language acquisition. *Language Learning* 39: 375–96. [CrossRef]
- Leow, Ron P., Luis Cerezo, Alison Caras, and Gorky Cruz. 2019. CALL in ISLA: Promoting depth of processing of complex L2 Spanish “Para/Por” prepositions. In *SLA Research with Implications for the Classroom: Reconciling Methodological Demands and Pedagogical Applicability*. Edited by Robert DeKeyser and Goretta Prieto Botana. Amsterdam: John Benjamins, pp. 155–78.
- Leow, Ronald P. 2015. *Explicit Learning in the L2 Classroom: A Student-Centered Approach*. New York: Routledge.
- Leow, Ronald P., and Jonathan D. Mercer. 2015. Depth of processing in L2 learning: Theory, research, and pedagogy. *Journal of Spanish Language Teaching* 2: 69–82. [CrossRef]
- Leow, Ronald P., and Lucia Donatelli. 2017. The role of (un)awareness in SLA. *Language Teaching* 50: 189–211. [CrossRef]
- Leow, Ronald P., ed. 2019. *The Routledge Handbook of Second Language Research in Classroom Learning*, 1st ed. New York: Routledge. [CrossRef]
- Levis, John M. 2018. *Intelligibility, Oral Communication, and the Teaching of Pronunciation*, 1st ed. Cambridge: Cambridge University Press. [CrossRef]

- Lo, Shih-Yu. 2018. Attention without awareness: Attentional modulation of perceptual grouping without awareness. *Attention, Perception, and Psychophysics* 80: 691–701. [CrossRef] [PubMed]
- Mackey, Alison, and Susan Gass. 2016. *Stimulated Recall Methodology in Applied Linguistics and L2 Research*, 2nd ed. New-York: Routledge.
- Martin, Ines. 2020a. Pronunciation development and instruction in distance language learning. *Language, Learning & Technology* 24: 86–106.
- Martin, Ines. 2020b. Pronunciation can be acquired outside the classroom: Design and assessment of homework-based training. *The Modern Language Journal* 104: 457–479. [CrossRef]
- Meritan, Camille, and Aurore Mroz. 2019. Impact of self-reflection and awareness-raising on novice French learners' pronunciation. *Foreign Language Annals* 52: 798–821. [CrossRef]
- Miller, Jessica. 2012. Teaching French pronunciation with phonetics in a college-level beginner French course. *NECTFL Review* 69: 47–68.
- Moyer, Alene. 2017. Autonomy in second language phonology: Choice vs. limits. *Language Teaching* 50: 395–411. [CrossRef]
- Munro, Murray J., Tracey M. Derwing, and Ron I. Thomson. 2015. Setting segmental priorities for English learners: Evidence from a longitudinal study. *International Review of Applied Linguistics in Language Teaching* 53: 39–60. [CrossRef]
- Munro, Murray J., and Tracey M. Derwing. 2015. A prospectus for pronunciation research in the 21st century: A point of view. *Journal of Second Language Pronunciation* 1: 11–42. [CrossRef]
- Paulson, Steve, Richard Davidson, Amishi Jha, and John Kabat-Zinn. 2013. Becoming conscious: The science of mindfulness: Becoming conscious: The science of mindfulness. *Annals of the New York Academy of Sciences* 1303: 87–104. [CrossRef]
- Pawlak, Mirosław. 2010. Designing and piloting a tool for the measurement of the use of pronunciation learning strategies. *Research in Language* 8: 1–14. [CrossRef]
- Plonsky, Luke, and Frederick L. Oswald. 2014. How big is “big”? Interpreting effect sizes in L2 research. *Language Learning* 64: 878–912. [CrossRef]
- Plonsky, Luke. 2015. *Advancing Quantitative Methods in Second Language Research*. London: Routledge.
- Ranta, Leila, and Roy Lyster. 2018. Form-focused instruction. In *The Routledge Handbook of Language Awareness*. Edited by Peter Garrett and Joseph M. Cots. New York: Routledge, pp. 40–56.
- Raoofi, Saeid, Swee Heng Chan, Jayakaran Mukundan, and Sabariah Md. Rashid. 2013. Metacognition and second/foreign language learning. *English Language Teaching* 7: 36–49. [CrossRef]
- Rosa, Eelena M., and Ronald P. Leow. 2004. Awareness, different learning conditions, and second language development. *Applied Psycholinguistics* 25: 269–92. [CrossRef]
- Saito, Kayuza, and Luke Plonsky. 2019. Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning* 69: 652–708. [CrossRef]
- Schmidt, Richard. 1990. The role of consciousness in second language learning. *Applied Linguistics* 11: 129–58. [CrossRef]
- Schmidt, Richard. 2010. Attention, awareness, and individual differences in language learning. *Proceedings of CLaSLIC*, 721–37. [CrossRef]
- Schwitzgebel, Eric. 2014. Introspection. In *The Stanford Encyclopedia of Philosophy*. Winter 2016. Stanford: The Metaphysics Research Lab Center for the Study of Language and Information, Stanford University. 2014. Available online: <https://plato.stanford.edu/archives/win2016/entries/introspection/> (accessed on 7 February 2019).
- Seemiller, Corey, and Meghan Grace. 2019. *Generation Z: A Century in the Making*. New York: Routledge.
- Sheskin, David J. 2011. Parametric Versus Nonparametric Tests. In *International Encyclopedia of Statistical Science*. Edited by Miodrag Lovric. Berlin: Springer Berlin/Heidelberg, pp. 1051–52. [CrossRef]
- Strijbos, Jan-Willem, Rob L. Martens, Frans J. Prins, and Wim M. G. Jochems. 2006. Content analysis: What are they talking about? *Computers and Education* 46: 29–48. [CrossRef]
- Swain, Merrill. 2005. The output hypothesis: Theory and research. In *Handbook of Research in Second Language Teaching and Learning*. Edited by Eli Hinkel. Mahwah: Lawrence Erlbaum, pp. 471–83.
- Thomson, Ron. 2017. Measurement of accentedness, intelligibility, and comprehensibility. In *Assessment in Second Language Pronunciation*, 1st ed. Edited by Okim Kang and April Ginther. New York: Routledge, pp. 11–29. [CrossRef]
- Thomson, Ron I., and Tracey M. Derwing. 2015. The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics* 36: 326–44. [CrossRef]
- Timmermans, Bert, and Axel Cleeremans. 2015. How can we measure awareness? An overview of current methods. In *Behavioral Methods in Consciousness Research*. Edited by Morten Overgaard. New York: Oxford University Press, pp. 21–46. [CrossRef]
- Van Patten, Bill. 2007. Input processing in adult second language acquisition. In *Theories in Second Language Acquisition*. Edited by B. Van Patten and J. Williams. Mahwah: Lawrence Erlbaum, pp. 115–35.
- Venkatagiri, Horobail S., and John M. Levis. 2007. Phonological awareness and speech comprehensibility: An exploratory study. *Language Awareness* 16: 263–77. [CrossRef]
- Wrembel, Magdalena. 2011. Metaphonetic awareness in the production of speech. In *Speaking and Instructed Foreign Language Acquisition*. Edited by Mirosław Pawlak, Ewa Waniek-Klimczak and Jan Majer. Bristol: Multilingual Matters, pp. 169–82.

-
- Wrembel, Magdalena. 2013. Foreign accent ratings in third language acquisition: The case of L3 French. In *Teaching and Researching English Accents in Native and Non-native Speakers*. Edited by Ewa Waniek-Klimczak and Linda R. Shockey. Berlin: Springer, pp. 31–47. [[CrossRef](#)]
- Wrembel, Magdalena. 2015. Metaphonological awareness in multilinguals: A case of L3 Polish. *Language Awareness* 24: 60–83. [[CrossRef](#)]