10 Pronunciation instruction in ESP teaching to enhance students' prosody

Leticia Quesada Vázquez[1]

Abstract

This study investigates the efficacy of explicit rhythm instruction to improve engineering students' prosody in English. A pronunciation module of ten weekly sessions of 30 minutes held within the class schedule was designed for a technical English course at Rovira i Virgili University. Sessions were outlined using a communicative framework. Two hundred and ninety eight Spanish/Catalan students were divided into three experimental groups receiving rhythm instruction, and three control groups which did not. Students were recorded before and after the training. Six native American English speakers were also recorded as a reference point. VarcoV values were measured and compared using PRAAT and the data were analysed using mixed analysis of variances (ANOVAS) and t-tests. Results reveal that the experimental group tends to increase in VarcoV after training, approaching English rhythm, while the control group presents incongruences. Despite results not always being significant, an analysis of the effect sizes for the t-tests comparing before and after VarcoV values for the experimental vs. the control groups shows significance. These results support the hypothesis that rhythm instruction can be beneficial to improve English for Specific Purposes (ESP) students' prosody.

Keywords: ESP, rhythm instruction, VarcoV.

1. Rovira i Virgili University, Tarragona, Spain; leticia.quesada@urv.cat; https://orcid.org/0000-0002-7422-6251

How to cite this chapter: Quesada Vázquez, L. (2019). Pronunciation instruction in ESP teaching to enhance students' prosody. In S. Papadima-Sophocleous, E. Kakoulli Constantinou & C. N. Giannikas (Eds), *ESP teaching and teacher education: current theories and practices* (pp. 163-176). Research-publishing.net. https://doi.org/10.14705/rpnet.2019.33.932

1. Introduction

Effective communication has always been at the core of ESP teaching (Dudley-Evans & St John, 1998). Since its inception in the 1960s, ESP has aimed at meeting students' needs for competent communication in their professional environments, especially of those tertiary learners who live and study in non-English speaking countries (Tzoannopoulou, 2015). Moreover, the launching of the Bologna Process in European universities has fostered the need for more ESP courses which advocate for enhancing communicative skills that favour the international market and the mobility derived from it (Räisänen & Fortanet-Gómez, 2008; Wilkinson, 2008).

It has been observed that ESP students may have the knowledge to face a communicative situation but, according to Douglas (2000), they often fail to effectively transmit and interact in the target language. Walker and White (2013) argue that, in contexts where learners have to speak, the practice of language skills that ensure communication can reduce students' anxiety and improve their intelligibility and fluency. Active listening, questioning, spoken interaction, and oral presentation rehearsal are some of the tools recommended so as to achieve this purpose (Dudley-Evans & St John, 1998). Nevertheless, not much attention has been paid to the application of these tools and others to pronunciation teaching in order to improve ESP students' intelligibility, comprehensibility, and fluency.

Pronunciation has often been neglected within English as a Second Language (ESL) and English as a Foreign Language (EFL) classrooms for several reasons, such as lack of time, teachers' limited training, or insufficient guarantee of lasting results (Derwing & Munro, 2015). Besides, the dichotomy between the nativeness and intelligibility principles (Levis, 2005) directly affects pronunciation teaching. Pronunciation instruction has long been associated with an ideal of nativeness, i.e. achieving a fully L1 English speaker accent. Consequently, the more native-like an ESL learner sounds when speaking English, the better he/she will be understood. However, such an ambitious goal is rarely achieved and students can become highly demotivated in their way to succeed. This frustrating attempt to reach the perfect pronunciation is more evident in adults, who have passed the critical period and may suffer from fossilisation (Levis, 2005). As for ESP students for whom English becomes just another working tool, sounding like a native speaker does not tend to be an appealing aim to achieve, but they would rather work at being understandable when communicating in the target language. Hence, the intelligibility principle, which claims to focus on just the aspects that guarantee communication and comprehension, seems to adjust better to ESP students' expectations. Nevertheless, determining which pronunciation features are more useful to teach for communicative purposes is not always an easy task and requires a deep analysis of the students' needs.

Experts agree that suprasegmental features play a more important role in global prosody than segmental ones. Suprasegmentals help the speech sound coherent and connect concepts that go beyond the meaning of isolated words (Gilbert, 2008). However, time is always tight in the ESL classroom, since several skills have to be taught. Therefore, deciding which is the feature that best meets the needs of each particular group of students becomes essential to ensure the proper functioning of the classroom (Basturkmen, 2010). ESP students are not language-oriented, so it is better to opt for simple and practical features that are easy to understand. Besides, their main aim is to be able to communicate, so fluency and comprehensibility issues are key in their learning process. Some studies have proved, by manipulating second language learners' speeches artificially, that the more L2-like the rhythm of their speech is, the more intelligible the speech becomes (Quene & Van Delf, 2010; Tajima, Port, & Dalby, 1997). As a consequence, language rhythm postulates as a plausible candidate.

Rhythm is found in the foundations of speech, organising thoughts, and connecting ideas. When the rhythm of a language is modified, the speech does not meet the listeners' expectations: it does not anticipate the lexical and syntactic information needed for an effective comprehension of the message (Derwing & Munro, 2015). Hence, both production and perception become compromised and misunderstandings and communication breakdowns arise. When speaking a second language, students tend to adopt their mother tongue rhythm, since they are not aware of the differences in rhythm among languages.

Consequently, students alter the rhythm of the target language, making the speech difficult to follow. Syllable-timed languages, on the one hand, and stress-timed languages, on the other hand, are placed at the extremes of the rhythm continuum: Spanish and Catalan are syllable-timed languages, so their rhythm is based on syllables that have approximately the same duration when pronounced. Lloyd James (1940) compared Spanish to a machine gun shot. English, however, is a stress-timed language, basing its rhythm on two different beats defined by stress: stressed syllables, which are pronounced longer, and unstressed syllables, which are pronounced shorter. In this case, Lloyd James related it to a morse code message. For this reason, Spanish and Catalan students are often said to sound chopped and pause wrongly when speaking English, a fact that negatively affects their fluency and comprehensibility in the second language.

Several studies have proved that the introduction of rhythmic cues in the EFL classroom can improve learners' intelligibility, fluency, and comprehensibility (Chela-Flores, 1997; Hahn, 2004; Tsiartsioni, 2011). Chela-Flores (1997) designed word-decontextualised patterns to be taught to Spanish students at the university of Zulia, in Venezuela, for a semester. Results revealed that students improved in both perception and recognition under controlled circumstances. Hahn (2004) investigated primary stress with international teaching assistants in the US. Three different versions of the same speech were created based on primary stress: a version where primary stress was correctly placed, a second one where it was incorrectly placed, and a third one where it was missing. American university students assessed the intelligibility of the speeches and results showed that the version where primary stress was correctly placed was considered more intelligible. Tsiartsioni (2011) worked with three age groups of EFL Greek students (6, 12, and 16 years old). Each age group was further divided into an experimental group that received rhythm instruction and a control group that did not. When calculating vocalic and consonantal Pairwise Variability Indexes (vocPVI and consPVI, respectively), it was observed that the experimental group improved its rhythm while the control group failed to do so. Little research on rhythm instruction has been conducted with ESP students. Chela-Flores (1993) introduced rhythm training within an ESP reading course in a nonnative environment obtaining encouraging results on listening discrimination (as cited in Chela-Flores, 1993). The current study investigates the effectiveness of rhythm instruction within an ESP course to improve the students' prosody in the target language. For this purpose, the following hypotheses have been formulated:

- the introduction of a pronunciation module based on Celce-Murcia, Brinton, and Goodwin's (1996) steps to teach communicatively will help students improve their L2 global prosody;

- students' L1 *negative transfer* (Celce-Murcia et al., 1996, p. 20) will decrease more when receiving explicit rhythm instruction; and

- students' rhythm will approach more that of L1 English speakers' when receiving explicit rhythm instruction.

2. Method

The present study is associated with an extended project that focusses on the effectiveness of explicit rhythm instruction within the EFL classroom. For this purpose, a classroom-based pronunciation research study (Derwing & Munro, 2015) was conducted. This empirical longitudinal study examined the progress of first-year undergraduate engineering students attending a compulsory B2 technical English course, which took place from February to May 2017 at Rovira i Virgili University. A pronunciation module was designed and embedded as part of the course. It consisted of thirty-minute sessions taught for ten weeks within regular classes. Two hundred and ninety-eight students were randomly divided into three experimental groups which received rhythm instruction during the sessions, and three control groups which did not. All the students were recorded before (pre-test) and after (post-test) training. The test consisted of four exercises: reading ten sentences and a text aloud, introducing themselves, and giving their opinion on social media. They were recorded individually in three isolated rooms at the university library using two Sony PCM-M1O and a Zoom H4nsp recorders. Besides, six native American English speaking visiting students from Bates College, (Maine, USA), were also recorded taking the test as a reference point.

This study concentrates on the acoustic analysis of the rhythm of the sentences uttered by those students who completed the treatment, and compares them with those of the natives. For this purpose, VarcoV values were measured. VarcoV is a rhythmic measure that estimates "the standard deviation of vocalic interval duration divided by mean vocalic duration, multiplied by 100" (White & Mattys, 2007, p. 508). After comparing several rhythmic measures, White and Mattys (2007) concluded that VarcoV was the most reliable measure in order to assess rhythm within the second language field. Due to its wider variation in syllable length, English shows higher VarcoV values than Spanish/Catalan. Hence, this study aims to examine, on the one hand, if ESP students' values increase after treatment and if there is a difference in results depending on the instruction received and, on the other hand, how the results obtained approach natives' values. The values of each of the sentences were obtained using PRAAT (www.praat.org). The boundaries of vocalic and consonant interval clusters were marked for each sentence, and Ordin and Polyanskaya's (2014, 2015) script[2] was run. Pauses were not considered in the analysis.

The data were analysed statistically with mixed repeated measures ANOVAS and t-tests. First of all, the effect of each instruction was examined by carrying out a mixed ANOVA with time (before and after instruction) and sentence as within-subjects factors, group (experimental or control) as a between-subjects factor, and VarcoV values as the dependent variable. Then, we were interested in studying the degree of variation between utterances before and after treatment for each group. Hence, a second mixed ANOVA was performed, this time with the difference in VarcoV values between the sentences before and after training as the dependent variable, sentence as the within-subjects factor and group as the between-subjects factor. Next, four t-tests were run for each sentence in order to analyse the impact of the difference depending on the instruction received: two paired-samples t-tests that compared the groups' learning process, and two-

2. The VarcoV values obtained from Ordin and Polyanskaya's (2014, 2015) script do not include the multiplication by 100, unlike specified in White and Mattys (2007).

independent-samples t-tests comparing the initial and final performance of the two groups. Finally, the effect sizes of the difference between sentences before and after instruction were further studied by performing one more independent-samples t-test. For all the tests, the alpha value was always set at .05[3].

2.1. Participants

Only the students who attended 90% of the sessions were considered to have fulfilled the treatment. Unfortunately, a lot of them dropped out or skipped sessions, so they had to be dismissed. In the end, 42 students could be used as subjects, 21 per group. The profile of these students was quite homogeneous: they were all between 18 and 20 years old, mostly balanced bilinguals of Spanish and Catalan; there were two students who were dominant in Spanish and three students whose mother tongue was a different language. However, all of them pursued primary and secondary education in Catalonia and, consequently, their command of both languages was high. The main difference between students was their initial level of English: the control group started with a slightly higher level of English (four low-intermediate, eleven intermediate, and six advanced students) than the experimental group (ten low-intermediate, six intermediate, and five advanced students). However, this variability had already been predicted because of the different educational backgrounds of ESP students (some of them coming from high school, others from vocational training), so level was not considered a determinant factor affecting the outcome of the treatment.

2.2. The pronunciation module

Sessions followed Celce-Murcia et al.'s (1996) steps to teach communicatively so as to guarantee a communicative framework. They all started with a description and analysis of the aspect to be taught. Next, students listened to some podcasts in order to distinguish and get familiar with the feature. Finally, they practised the item at three different levels: controlled practice, doing activities like reading aloud, and *listen and repeat*; guided practice, playing

3. Statistics have been verified by Prof. Urbano Lorenzo, Rovira i Virgili University.

guessing games, and doing gap-filling exercises; and communicative practice by participating in group discussions and debates.

The module was scheduled according to the syllabus of the course, i.e. the grammar and vocabulary used had already been taught in class. This way, students could concentrate better on practising their speaking skills and, at the same time, their motivation was guaranteed because they continued working on concepts directly related to their disciplines (Anderson-Hsieh, 1990). Regarding materials, activities were adapted from several pronunciation books and research papers by mainly modifying rhythm activities to use technical vocabulary, or they were created from scratch by using online resources such as videos or images.

3. Results and discussion

VarcoV means were calculated for each of the sentences. Table 1 compiles the means and standard deviations obtained for the experimental and the control group, both before and after treatment, and for the natives.

Table 1. VarcoV means per sentence

Group	**Sentence**				
	1	**2**	**3**	**4**	**5**
Control pre-test	0.37 (0.08)	0.40 (0.09)	0.39 (0.12)	0.54 (0.09)	0.45 (0.09)
Control post-test	0.36 (0.11)	0.40 (0.09)	0.38 (0.10)	0.52 (0.08)	0.47 (0.14)
Experimental pre-test	0.39 (0.11)	0.38 (0.07)	0.38 (0.10)	0.50 (0.09)	0.41 (0.09)
Experimental post-test	0.40 (0.08)	0.39 (0.08)	0.39 (0.13)	0.51 (0.09)	0.50 (0.12)
Natives	**0.33* (0.08)**	**0.48 (0.03)**	**0.36* (0.06)**	**0.56 (0.07)**	**0.57 (0.06)**
Group	**Sentence**				
	6	**7**	**8**	**9**	**10**
Control pre-test	0.48 (0.12)	0.55 (0.10)	0.42 (0.07)	0.49 (0.10)	0.48 (0.11)

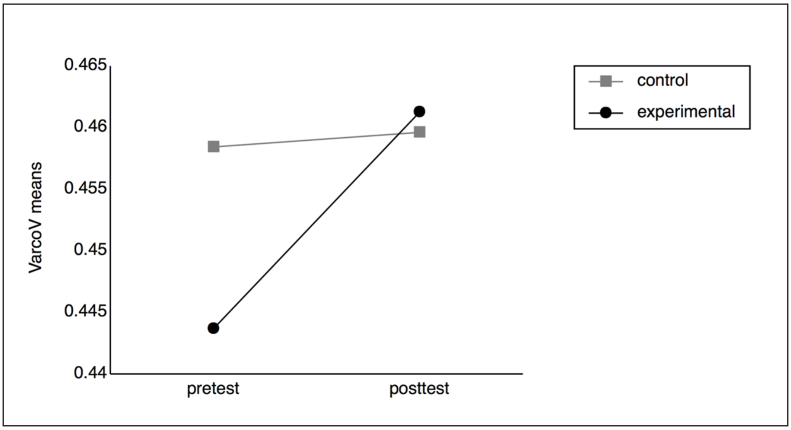
Control post-test	0.49 (0.10)	0.53 (0.09)	0.43 (0.08)	0.51 (0.13)	0.52 (0.11)
Experimental pre-test	0.45 (0.10)	0.53 (0.13)	0.41 (0.08)	0.52 (0.11)	0.48 (0.05)
Experimental post-test	0.45 (0.09)	0.56 (0.09)	0.43 (0.06)	0.48 (0.09)	0.51 (0.08)
Natives	**0.53 (0.06)**	**0.64 (0.05)**	**0.50 (0.05)**	**0.47* (0.07)**	**0.52 (0.11)**

By examining the means, several observations could be made. Firstly, as expected, VarcoV values were generally higher for native speakers than for second language learners. There were three exceptions, though: Sentences 1, 3, and 9 (marked with an asterisk). In these cases, natives' values were noticeably lower than those of the ESP students. By examining the sentences in depth (see supplementary materials), it could be observed that Sentences 1 and 3 were much shorter than the rest. Natives tended to speak more slowly in these sentences, vocalizing excessively, and even pausing in unnatural places, showing an intentional will to sound clear. They seemed to be more aware of the need to make themselves understood. On the other hand, they uttered longer sentences in a more natural way, maintaining a more constant rhythm. This phenomenon could explain why natives' VarcoV values are lower in these sentences. However, Sentence 9 was not short and showed lower VarcoV values too. ESP students struggled a lot when reading this sentence, leading to many mispronunciations, a fact that could have affected results. However, further research should be conducted to prove this point.

Secondly, the experimental group tended to increase its VarcoV values after treatment, approaching native performance, while the control group behaved incongruently. Shadowed in grey, the increase in VarcoV for the experimental group was discernible in eight out of the ten sentences while the control group's figures only rose in half of them. Thus, numbers suggested that explicit rhythm instruction helped students to adopt the rhythm of the target language. It is true that the control group showed higher values in some of the sentences, closer to native figures. However, control group students started with a higher initial English level, so higher values were expected regardless of instruction.

As previously mentioned, in order to determine the significance of the effect of the instruction, a mixed repeated measures ANOVA was carried out with time and sentence as within-subjects factors and group as a between-subjects factor. Non-significant results were found for both time $F(1,40)=2.006$, $p=.164$, and group $F(1,40)=.267$, $p=.608$, but they were significant in regards to sentence, $F(9,32)=29.172$, $p=.01$. However, the time*sentence interaction did not show significance, $F(9,32) = 1.185$, $p = .313$. Despite a lack of statistical significance, a clear impact of explicit rhythm instruction could be observed in Figure 1.

Figure 1. VarcoV progress after instruction



While the control group remained almost stable after treatment, the experimental group underwent a sharp increase after instruction. Hence, a positive effect of explicit rhythm instruction was still detected.

So as to examine the disparity in figures before and after treatment based on the instruction received, the second mixed ANOVA was run. The differences in VarcoV values before and after training were the dependent variable, sentence was the within-subjects factor, and group was the between-subjects factor. No significance was shown for either group, $F(1,40)=1.532$, $p=.223$, or sentence, $F(9,32)=1.185$, $p=.313$. The sentence*group interaction was also non-significant $F(9,32)=.961$, $p=.456$.

T-tests were further performed to look for variations within each sentence. For each sentence, four different t-tests were performed: two paired-samples t-tests comparing each group and two independent-samples t-tests comparing both groups before and after instruction. Neither the t-tests ($p>.05$) nor the corresponding effect sizes ($d<2$) showed significance. Nevertheless, a difference between the experimental and the control group effect sizes was observed when analysing paired-samples t-tests for each of the sentences. While the experimental group tended to display a positive difference, the control group shows more negative ones and, when the difference was positive, it was still smaller than for the experimental one (Table 2).

Table 2. Effect sizes of control and experimental paired-sample t-tests per sentence

Effect sizes difference per sentence										
Group	**Sentence**									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Control	-1.28	0.06	-0.12	-0.23	0.13	0.04	-1.73	0.12	0.11	0.31
Experimental	1.77	0.08	0.06	0.17	0.65	-0.05	0.24	0.25	-0.26	0.39

For this reason, another independent-sample t-test was performed to analyse the relevance of the effect size. This time, results were statistically significant $T(18)=-2102$, $p=.05$. Findings revealed a better performance when teaching rhythm explicitly. Nevertheless, further research needs to be conducted to reach more conclusive results.

4. Conclusions

By comparing students' progress after pronunciation instruction and correlating it to the native speaker counterparts, this study examines the effectiveness of specific pronunciation teaching (in this case, explicit rhythm instruction) on ESP students' prosody. Several limitations such as a dramatic decrease of the population under study, overcrowded classes, or limited time for instruction should be taken into account when interpreting the statistical non-significance of the results. Still, explicit rhythm instruction is shown to increase students' VarcoV values in most of the sentences analysed, enhancing students' acquisition of the rhythm of the target language, and consequently decreasing the negative transfer from their mother tongue. Besides, figures tend to get closer to the native equivalents. On the contrary, a lack of this kind of training results in inconsistent behaviour. Findings, hence, suggest that ESP students' prosody can improve by means of explicit rhythm instruction, but more research has to be conducted in order to reach statistical significance. On the other hand, as the control group also shows signs of improvement in some sentences, pronunciation seems to arise as a beneficial aspect to teach within the ESP classroom.

Acknowledgements

This research was funded by a Martí i Franquès doctoral grant from Universitat Rovira i Virgili and the project FFI2017-84479-P from the Spanish Ministry of Economy and Business.

Supplementary materials

https://research-publishing.box.com/s/kqyiwxmatusypidyrfawezo6uw3z1zl3

References

Anderson-Hsieh, J. (1990). Teaching suprasegmentals to international teaching assistants using field-specific materials. *English for Specific Purposes, 9*(3), 195-214. https://doi.org/10.1016/0889-4906(90)90013-3

Basturkmen, H. (2010). *Developing courses in English for specific purposes*. Palgrave Macmillan.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching pronunciation: a reference for teachers of English to speakers of other languages*. Cambridge University Press.

Chela-Flores, B. (1993). On the acquisition of English rhythm: theoretical and practice issues. *Lenguas Modernas, 20,* 151-164.

Chela-Flores, B. (1997). Rhythmic patterns as basic units in pronunciation teaching. *Onomázein, 2*, 111-134. http://onomazein.letras.uc.cl/Articulos/2/4_Chela.pdf

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals. Evidence-based perspectives for L2 teaching and research.* John Benjamins.

Douglas, D. (2000). *Assessing language for specific purposes*. Cambridge University Press.

Dudley-Evans, T., & St John, M. (1998). *Developments in English for specific purposes: a multidisciplinary approach.* Cambridge University Press.

Gilbert, J. B. (2008). *Teaching pronunciation: using the prosody pyramid.* Cambridge University Press.

Hahn, L. (2004). Primary stress and intelligibility: research to motivate the teaching of suprasegmentals. *TESOL Quarterly, 38*(2) 201-223. https://doi.org/10.2307/3588378

Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. *TESOL Quarterly, 39*(3) 369-378. https://doi.org/10.2307/3588485

Lloyd James, A. (1940). *Speech signals in telephony.* Pitman & Sons.

Ordin, M., & Polyanskaya, L. (2014). Development of timing patterns in first and second languages. *System, 42*, 244-257. https://doi.org/10.1016/j.system.2013.12.004

Ordin, M., & Polyanskaya, L. (2015). Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. *The Journal of the Acoustical Society of America, 138*(2), 533-544. https://doi.org/10.1121/1.4923359

Quene, H., & Van Delft, L. E. (2010). Non-native durational patterns decrease speech intelligibility. *Speech Communication, 52*(11-12), 911-918. https://doi.org/10.1016/j.specom.2010.03.005

Räisänen, C., & Fortanet-Gómez, I. (2008). The state of ESP teaching and learning in Western European higher education after Bologna. In I. Fortanet-Gómez & C. Räisänen (Eds), *ESP in European higher education. Integrating language and content* (pp. 11-51). John Benjamins. https://doi.org/10.1075/aals.4.03rai

Tajima, K., Port, R., & Dalby, J. (1997). The effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics, 25*(1), 1-24. https://doi.org/10.1006/jpho.1996.0031

Tsiartsioni, E. (2011). Can pronunciation be taught? Teaching English speech rhythm to Greek students. In E. Kitis, N. Lavidas, N. Tpointzi & T. Tsangalidis (Eds), *Selected papers from the 19th International Symposium on Theoretical and Applied Linguistics (ISTAL 19) (pp.* 447-458*)*. Aristotle University of Thessaloniki.

Tzoannopoulou, M. (2015). Rethinking ESP: integrating content and language in the university classroom. *Procedia - Social and Behavioral Sciences, 173*, 149-153. https://doi.org/10.1016/j.sbspro.2015.02.045

Walker, A., & White, G. (2013). *Technology enhanced language teaching: connecting theory and practice.* Oxford University Press.

White, L., & Mattys, S. (2007). Calibrating rhythm: first language and second language studies. *Journal of Phonetics, 35*(4), 501-522. https://doi.org/10.1016/j.wocn.2007.02.003

Wilkinson, R. (2008). Locating the ESP space in problem-based learning: English-medium degree programmes from a post-Bologna perspective. In I. Fortanet-Gómez & C. A. Räisänen (Eds), *ESP in European Higher Education: Integrating Language and Content* (pp. 55-73). John Benjamins. https://doi.org/10.1075/aals.4.05wil

Research-publishing.net

Published by Research-publishing.net, a not-for-profit association
Voillans, France, info@research-publishing.net

ESP teaching and teacher education: current theories and practices
Edited by Salomi Papadima-Sophocleous, Elis Kakoulli Constantinou, and Christina Nicole Giannikas

Publication date: 2019/06/22

Typeset by Research-publishing.net
Cover layout by © 2019 Raphaël Savina (raphael@savina.net)

ISBN13: 978-2-490057-45-0 (Ebook, PDF, colour)
ISBN13: 978-2-490057-46-7 (Ebook, EPUB, colour)
ISBN13: 978-2-490057-44-3 (Paperback - Print on demand, black and white)
Print on demand technology is a high-quality, innovative and ecological printing method; with which the book is never 'out of stock' or 'out of print'.

British Library Cataloguing-in-Publication Data.
A cataloguing record for this book is available from the British Library.

Legal deposit, UK: British Library.
Legal deposit, France: Bibliothèque Nationale de France - Dépôt légal: juin 2019.