Moving Beyond the Plateau with Computer-Assisted Pronunciation Mediation

Evgeny Pyshkin[1,*], John Blake[1], and Natalia Bogach[2]

[1]School of Computer Science and Engineering, University of Aizu, Aizu-Wakamatsu, 965-8580, Japan
[2]Institute of Computer Science and Technology, Peter the Great St. Petersburg Polytechnic University, St. Petersburg, 195251, Russia
Email: pyshe@u-aizu.ac.jp (E.P.); jblake@u-aizu.ac.jp (J.B.); bogach_nv@edu.spbstu.ru (N.B.)
[*]Corresponding author
Manuscript received November 8, 2023; revised December 8, 2023; accepted December 15, 2023; published June 21, 2024

***Abstract*—This paper examines the plateau effect and the role of Computer-Assisted Pronunciation Training (CAPT) in enabling learners to move beyond the plateau. Specifically, the samples and empirical evidence for this study is provided by StudyIntonation, a top-down CAPT environment supporting visual display of phrasal intonation in the form of pitch graphs presented for both the model pitch of native speakers and attempts of learners. Utterances of model speakers are grouped by situational context, prosodic forms and mode of feedback. The system largely relies upon concepts of sociocultural theory, such as mediation through the zone of proximal development. With the help of an experimental group of high school students, we use mathematical models to pinpoint the Zone of Proximal Development (ZPD), allowing for delivery of tailored task classification based on syntax and discourse tags. Longitudinal pronunciation assessment supported by dynamic time warping and cross-recurrence quantification analysis is used to locate the ZPD, and provide personalized practice. Preliminary experiments for the current study demonstrate that this approach can help learners in overcoming the plateau effect experienced during the learning process. The reported new findings create further grounds for improving and targeting the CAPT system feedback to learners with respect to their individual achievements and differences in learning styles.**

***Keywords*—iCAPT, L2 pronunciation instruction, dynamic assessment, scaffolding, Vygotsky sociocultural theory, zone of proximal development**

I. Introduction

Language plateau is used to describe the absence of progress, encountered by language learners, typically at rather high levels of language proficiency. Despite ongoing effort and practice, learners feel that they are not making any significant improvements [1, 2]. A holistic plateau effect unfolds through the following external manifestations: 1) Unbalanced perceptive and productive skills; 2) Sacrifice of complexity for fluency; 3) Reliance on known vocabulary; 4) Unnatural (marked) speech; and 5) Fossilized errors of pronunciation and language use.

The larger part of research for the plateau phenomenon in language learning is focused on the origins of plateau and forms of its manifestation in language practice of learners; meanwhile, there is only partial understanding of plateau overcoming [3], and what are the drivers that may head learners out of plateau. Nonetheless, Acton [4] contends that fossilized pronunciation imperfections may be mitigated through appropriate training including noticing pronunciation mistakes and readiness for daily self-training. He also contends that significantly more progress can be made with seemingly fossilized speech than is generally acknowledged.

Following Acton's assertion, we examine the plateau effect using our Computer-Assisted Pronunciation Training (CAPT) framework StudyIntonation [5] and seek to understand whether CAPT for phrasal intonation exercises may create a sufficient impetus to moving beyond the plateau both in the sense of mastering language skills and increasing students' motivation. We also show that the issue of conveying proper CAPT feedback [6], may be, at least partially, solved within Sociocultural Theory (SCT) of Vygotsky and regard feedback as object-regulation type of mediation through the Zone of Proximal Development (ZPD).

Intelligent CAPT (iCAPT) tools are complex instruments integrating automatic speech recognition, pitch visualization, and pitch evaluation algorithms into environments accessible through web interfaces and mobile applications. Advancing web and mobile applications to develop pronunciation is currently one of the leading trends in pronunciation teaching [6].

iCAPT systems actively involve approaches bridging language-related linguistic theories to human-machine interaction mechanisms, the latter, in turn, connected to advanced multimodal software interfaces harnessing the recent achievements of multimedia technology. These software interfaces become particularly important while supporting individual learning when the teacher is not present; however, they can also be efficiently used to improve traditional classroom scenarios.

The following section introduces the software platform and theoretical assumptions which conduct the present research; Section III addresses the specific plateau effect aspects discusses the opportunities which CAPT framework may offer; Section IV presents the preliminary experiments providing evidence for soundness of theoretical discourse; Section V discusses the major effects observed.

II. Pronunciation Training Based on Vygotskian Developmental Model

A. StudyIntonation CAPT Platform

The software platform for this research StudyIntondation [5] is a free multimodal pronunciation iCAPT framework with audiovisual feedback built on the grounds of a top-down pronunciation training paradigm, and accessible primarily through the interface of mobile applications (Fig. 1).

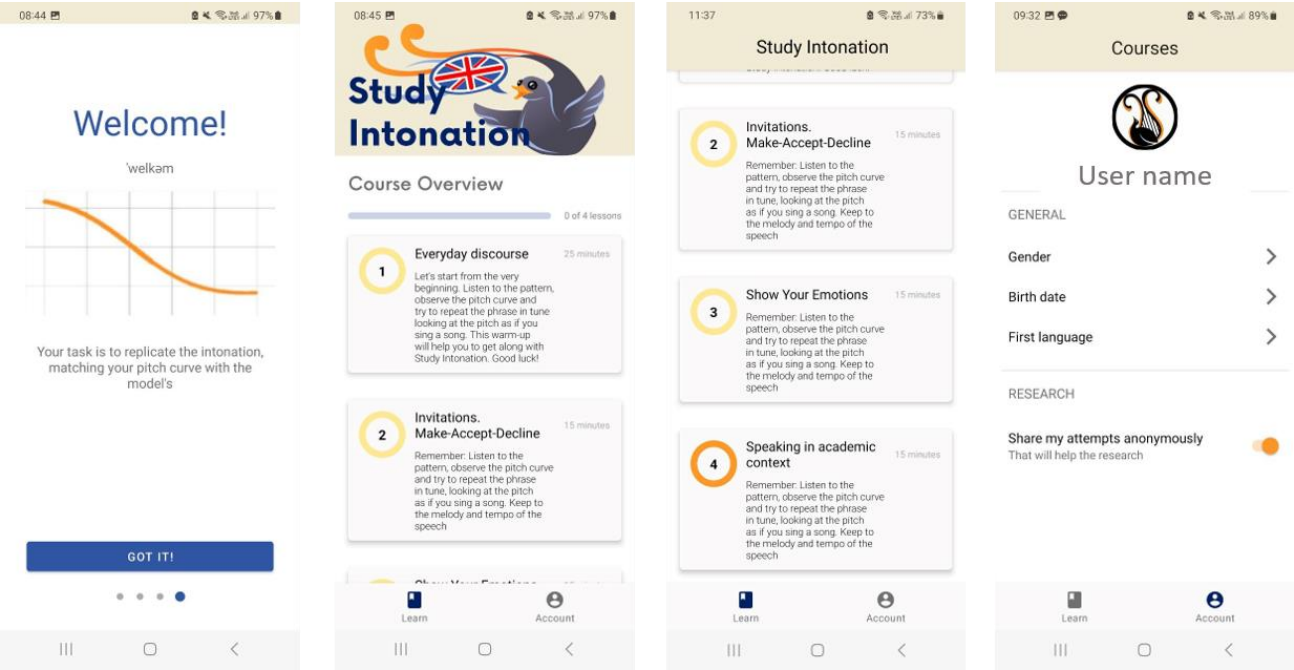

Fig. 1. StudyIntonation user interface.

From a speech technology standpoint, this system presents the contour of the fundamental frequency f_0 extracted from the audio track. This offers a visual representation of the phrasal intonation inherent in pre-recorded patterns of native speech. Despite their intuitiveness and utility, such pitch graphs are not inherently designed to offer immediate instructional or corrective feedback. Contemporary CAPT tools do not automatically guarantee pedagogical and linguistic soundness [7] that is one of the capstone problems of this class of systems. Consequently, in order to estimate learning progress in prosody production and to enable more objective judgements on learning dynamics, further research is needed.

Since 2017, we have developed a number of research versions aimed at examining the specific features addressing the problem of generating instructive and personalized feedback supported by the model pronunciation courses developed by language teachers (Fig. 2).

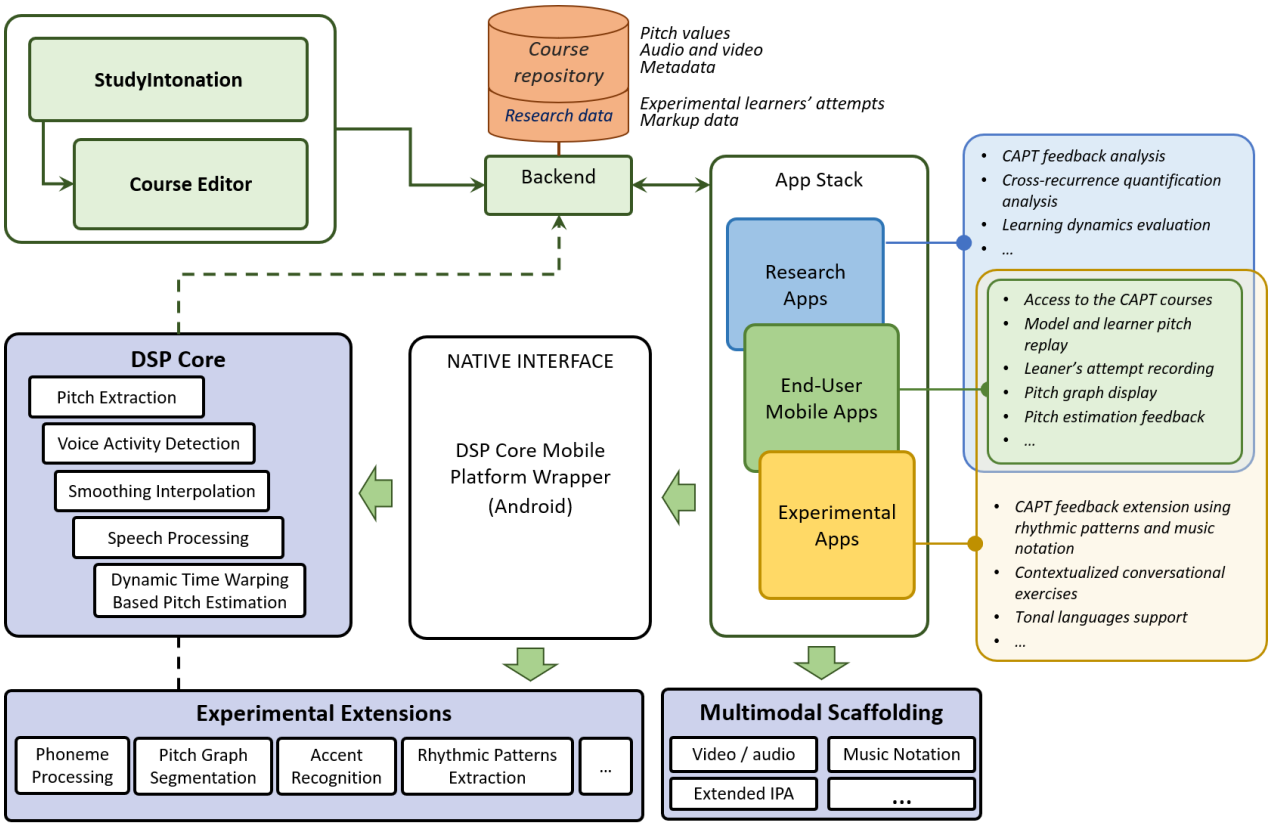

Fig. 2. System architecture.

The early designs of StudyIntonation [5, 8–10] were English only and applied Pearson Correlation Coefficient

(PCC) and Root Mean Square/Mean Squared Error (RMS/MSE) metrics for measuring the distortion between the model and the learner pitch graphs, while in the recent versions [11, 12], we extended the set of target languages and metrics used.

B. Sociocultural Approach to CAPT Feedback as a Way Towards Better Pedagogical Soundness

For our CAPT system project, we drew on established developmental theories for pedagogical insights, specifically incorporating Vygotsky's Sociocultural Theory (SCT) [13, 14] to design an internal mediation mechanism [12]. The increasing acknowledgement of SCT underscores the growing acceptance of the progressive idea individuals are not repositories of knowledge, but develop through holistic and social experiences. Based on Vygotsky's argument that learning leads development [15], certain types of instruction can play a significant role in developing additional language proficiency. A key tenet within SCT is that the human mind is mediated [16], with mediation taking the form of object-, other- and self-regulation.

In the context of additional language learning, object regulation refers to inanimate objects, e.g. online translation tools [17]. Other-regulation refers to regulation stemming from interactions with other people and self-regulation refers to a learner no longer needing object- or other-regulation to complete a task due to internalizing mediation [18]. Therefore, CAPT is a form of object-regulation, which learners may internalize to shift towards self-regulation of pronunciation features and other linguistic knowledge.

How mediation is delivered is particularly important. Dynamic Assessment (DA) is a teaching method in which assessment and mediation (intervention) are conducted dynamically, focusing on "what learners will be able to do in the future" rather than evaluating "what they can do" at a certain point in time. In DA, the learner's abilities are brought out within the ZPD. This view has attracted attention in multiple fields, including the education of children with disabilities [19, 20], and second language acquisition [21, 22].

While there are examples of DA achieving results within larger classrooms [23], DA has tended to be conceptualized mostly as a one-to-one activity [24]. This is because the amount of mediation a learner needs at any one time is constantly changing, including what may appear to be backward steps [17].

According to Devrim [25], feedback in SCT based approaches is not concerned with providing the correct form or bringing errors to students' better attention; but serves as a dialogic exchange between more knowledgeable others and learners [15].

Effective error correction and language learning depend to a considerable degree on mediation which co-constructs a zone of proximal development in which feedback as regulation becomes relevant and can therefore be appropriated by language learners to modify their interlanguage systems [25].

C. Dynamic System Theory (DST) as a Way Towards Better Observability and Research Rigor

We showed earlier [12] that assessment of learning dynamics when interacting with CAPT system can be described in terms of Dynamic System Theory (DST), with the understanding that the agents of language instruction, namely the learner and the CAPT software, can both be considered dynamic systems [26–29]. Hence, language learning should be examined with respect to learners' dynamic behavior: indeed, mastering a language assumes the development of language abilities over time, and not just a process of acquiring abstract rules or learning static patterns [26, 27, 30]. We observed that DST applied to L2 pronunciation teaching and longitudinal L2 pronunciation assessment can be an insightful tool to explore the individual progress of the learners as well as the processes occurring during L2 phonological development [31, 32].

According to Geert [33], the process of transition of a dynamic system from the actual state to its novel state driven by a rule of change (often referred as an "evolution rule") can be described as a sequence in a state space:

$$x_{t+1} = f\left(x_t\right) \rightarrow x_{t+2} = f\left(x_{t+1}\right) \rightarrow \ldots$$

Given that learning is a precursor to development, a CAPT system may guide learners to attain the next level of language proficiency or provide an impetus to jump off the "learning plateau"? One of the outcomes of SCT is that challenging tasks offered to learners maximize their learning opportunities and may accelerate their progress [14]. While the benefits of CAPT practices are clearly evident; to the best of our knowledge, existing research has not conclusively demonstrated that CAPT practices not only enhance pronunciation but also contribute holistically to learners' overall language development and increase their illocutionary force.

The experimental observations illustrate the known phenomenon that learners' performance does not increase linearly, but passes through periods of progression and regression alternatively [34–36]. These are not isolated jumps, but the stages of a continuous developmental process; each individual learner demonstrated unique patterns of this developmental trajectory.

DST and Vygotskian SCT may seem conceptually distant, nonetheless, Van Geert showed that DST may be used as a mathematical model for developmental theories by incorporating the Vygotsky dialectical mechanism of development into the dynamic system model and articulated such concepts as mediation, Zone of Proximal Development (ZPD), etc. as per dynamic equations. Van Geert argues the equations can identify features in the process of maturing [33, 37].

Subsequently, we could spot the specific tasks, where performance metrics produce oscillatory behavior, and which offer valuable insights into which exercises could be of maximum usefulness due to heightened sensitivity, responsiveness and perception of learners while being within their ZPD [28, 38, 39].

The mathematical model for the Vygotskian concept of learning includes evolution rules for actual and potential developmental levels [37]. This model describes the reciprocal interaction between the actual and potential development zones with the support of control variables. The effect of this dynamic model is that it allows for observing the learner's gradual transition from one developmental level to another. The existence of content that is more sensitive to experience than the others is, according to Geert [40], based

on two opposing tendencies that are likely to occur in learning and developing systems: preference for novelty and preference for familiarity.

These tendencies can be expressed in the form of a pair of exponential functions [37, 40]:

$$f_{familar}(i) = ab^{ci}$$
$$f_{novel}(i) = dg^{fi}$$

where $a, d, f > 0, b, g \in (0,1), c < 0$ and i is the distance between learning contents in the array I. Van Geert [40] names the most preferred content for both functions as the one at the cross-section point of $f_{familar}(i)$ and $f_{novel}(i)$.

D. Research Questions

Rooted in our previous work [12], the current study is guided by the following research questions:

1) How developmental processes of learners observed in the course of user interaction with L2 suprasegmental CAPT system correlate to the plateau effect?
2) What activities commonly advised for surmounting learning plateaus can be incorporated into StudyIntonation?
3) Is there a discernible positive trend in learner outcomes and attitudes which may be attributed to CAPT interaction experience?

III. Beyond-the-Plateau Effects of Top-Down Pronunciation Exercise Mediated by a CAPT System

In this section, we consider the five manifestations of the language plateau effect, and analyze to what extent CAPT system pronunciation activities may help to overcome them.

A. Unbalanced Perceptive and Productive Skills

Recent theories of second language acquisition no longer assume that productive skills arise naturally from comprehension skills [1, 41]. To reduce the gap between receptive and productive competence it is necessary to be aware of two hypotheses, which presume that comprehensible input is noticing (the noticing hypothesis), and focused output (the output hypothesis).

In a nutshell, the noticing hypothesis states that noticing is an essential precursor to learning. According to the noticing hypothesis, there is a difference between input and intake [1]. Intake is the part of the input that learners notice, which may serve as the basis for language development and later emergence of new features in speech.

Among the features which are likely to contribute to the extent to which learners will notice features of input the nature of the activity, learners are taking part in, plays an important role. These exercises are text-based and largely rely upon intermittent focusing on attention, memorizing and production activities such as analytical reading, Q&A sessions, text discussion and retelling. The output hypothesis suggests that practice in using target language forms is necessary for learners to acquire new target language. In context of CAPT training we observed that memorizing speech constructions performed together with training to produce them with correct intonation have mutually enhancing effect both for noticing and output [5].

B. Sacrifice of Complexity for Fluency

The development of fluency may mean greater ease of use of known language forms, but it does not necessarily imply development in complexity. To enhance the sophistication of learners' language, it is essential for them to acquire and incorporate new linguistic forms into their active language use, i.e. their productive linguistic repertoire. This involves processes that mediate the incorporation of intake into the developing system.

For example, prior to completing a role-play task that practices a particular speech situation, students can first perform some preliminary activities to learn key vocabulary they will use in the role play.

They could also listen to and practice a model dialog, which serves both to display different communication strategies and to model the kind of task the students will perform. Listening to a dialog supported by intonation visualization, provided by a CAPT tool, allow focusing on intonation and speakers' emotions that helps to capture learner's attention.

C. Reliance on Known Vocabulary

Vocabulary development plays a vital role in making the transition from an intermediate to a more advanced level of language proficiency, but many learners appear to have reached a learning plateau in relation to vocabulary. This may be seen in the overuse of lower-level vocabulary and failure to acquire more advanced-level vocabulary, as well as limited awareness of collocational usage. One of the key problems in helping learners improve their vocabulary is finding effective ways to help them remember words they have encountered.

Our mental lexicon is highly organized and efficient, and that items that are related semantically are stored together. Apart from generally acknowledged effective teaching and learning strategies like guided discovery, contextual guesswork, and mastering dictionary working out the correct tone movement for new words intensifies their memorization and automates usage.

Another dimension of vocabulary development, which is essential if students are to make a successful transition to the advanced level, is to expand what has been called their collocational competence [1]. Collocation refers to restrictions on how words can be used together, such as which prepositions are used with particular verbs, or which verbs and nouns are used together. Vocabulary development does not only involve acquiring new words, but also involves expanding knowledge of the collocational patterns that known words can enter into. Top-down approaches, which operate with sentences, have a positive reinforcement effect on learners' collocational experience.

D. Unnatural (marked) Speech

Learners' spoken English may be accurate and fluent but marked, which means the language used is non-typical or natural. There are many factors that can contribute to the naturalness of speech. One important factor is the extent to which the learners are using what are sometimes called multi-word "chunks", as well as conversational routines or fixed expressions. Random patterns of words do not occur together in speech, but often occur as multi-word chunks. These may be two-, three-, four-, five-, or even six-word chunks.

In CAPT sessions, learners engage in dual training that targets both cognitive skills and procedural memory.

This approach compels learners to internalize the language's formal structures and semantic nuances, cultivating unconscious habits that facilitate spontaneous communication.

E. Fossilized Errors of Pronunciation, Language Use, Grammar

Fossilization refers to the persistence of errors in learners' speech despite progress in other areas of language development. One feature of fossilized language items that suggests a partial explanation for this phenomenon is that fossilized errors tend not to affect our understanding of the speaker, although they might be irritating and may also be stigmatized, since they often reflect errors that are typical of elementary level learners (such as omission of third-person "s" [1]). Since fossilized errors do not generally trigger misunderstanding and hence do not prompt a clarification request from listeners, learners may remain oblivious to their existence.

IV. Preliminary Results and Observations

A. Experiment Setup

A cohort of 60 master's students consented to participate in this research study examining the plateau effect in language learning. Among the participants, three have Arabic as their first language (L1), five have Chinese, and the remainder have Russian. The participants were interviewed before and after a series of experiments, where they were practicing with a subset of 20 StudyIntonation tasks to imitate such aspects of pronunciation as sounds, stress, rhythm and intonation.

B. Learners' Pre-Experiment Interview

In order to address the question of how EFL learners describe their perception of the plateau phenomenon, some qualitative information on the views of English language learners toward the learning curve they experienced during their language studies was needed [42]. The primary instrument utilized was a semi-structured interview.

This conceptualization of the plateau phenomenon as a process was used to design the interview questions. Prior to conducting the interviews, the concept of plateau was fully explained to the interviewees in order to get the most accurate data. The questions for the interviews were as follows:

1) Have you ever experienced a plateau?
2) If possible, read your essays written some time ago. What do you feel about them?
3) What enables plateau?
4) What hinders plateau?
5) Would you like to jump off the plateau? Do you believe it is a plausible scenario for you?

The results were predictable fairly enough, as all the respondents answered positively; moreover, the most frequent comment was that 'it couldn't be the other way'.

C. Experiment

All the tasks refer to speech situations in academic environments, but differ in complexity. Each task was repeated by a learner until one of 3 scenarios of performance dynamics could be considered plausible, i.e. each task could be attributed to learner's one of 3 development zones:

1) Actual Development Zone (ADZ), where learner/model distortion steadily decreases, performance metrics improve and do not degrade over time, synchronization is observable.
2) Far Development Zone—the opposite situation, when the task is too difficult yet, learner/model distortion remains significant and no positive dynamics are present.
3) Zone of Proximal Development (ZPD), where performance is unstable, distortion metrics oscillates in the vicinity of two levels—one shows the actual performance, while the other delineates the area where performance metrics may migrate under mediation.

StudyIntonation task screen shown in Fig. 3 illustrates a typical ADZ scenario. The performance metric diminishes smoothly, showing that the learner's production gradually converges towards the model and a kind of synchronization of two systems occurs.

Microgenetic effects, which could be attributed to ZPD scenario, cover all attempts for a specific task within one training session of one learner (Fig. 4). Transition to ZPD was located by oscillation of Dynamic Time Warping (DTW) metric between two levels.

D. Learners' Post-Experiment Interview

Informal follow-ups indicate that students continue to use the techniques they have learned in the course and perceive their improvement to be relatively permanent.

1) From 0 to 2, please, rate your impression about the impact of training onto 5 aspects of plateau: 0–hardly any impact; 1–moderate; 2–strong enough.
2) Have you got acquainted with some novel facts about pronunciation or got any motivation uplift on pronunciation training (please answer "yes" or "maybe", or "no" both for novelty and motivation).

In Table 1, the numbers indicate how many participants have given a particular answer during the post-interview.

Table 1. Impact of pronunciation training over the specific aspects of language plateau

Aspects	No impact	Moderate impact	Strong impact
Unbalanced Skills	10	86	4
Fluency for Complexity	8	35	57
Poor Vocabulary	0	45	55
Unnatural Speech	0	43	57
Fossilized Errors	20	75	5
Novelty	0	0	100
Motivation	2	80	18

It is difficult to quantify accurately the degree of improvement, in part because their enhanced intelligibility is a function of more than just their pronunciation [25], in a traditional sense. Whereas pronunciation of individual phonological segments may not improve radically in spontaneous speech, the change in overall intelligibility is unmistakable. Evidence for the effectiveness is of three types. First, by the midpoint of the course, supervisors and peers are normally reporting to participants that they can, indeed, hear a difference in their speech (perhaps to some extent due to improved delivery and pacing or successful context rehabilitation). Second, independent judges assessing pre-

and post-audiotapes rarely fail to note significant improvement in those who complete all assignments satisfactorily. And third, student response to the program has been enthusiastic.

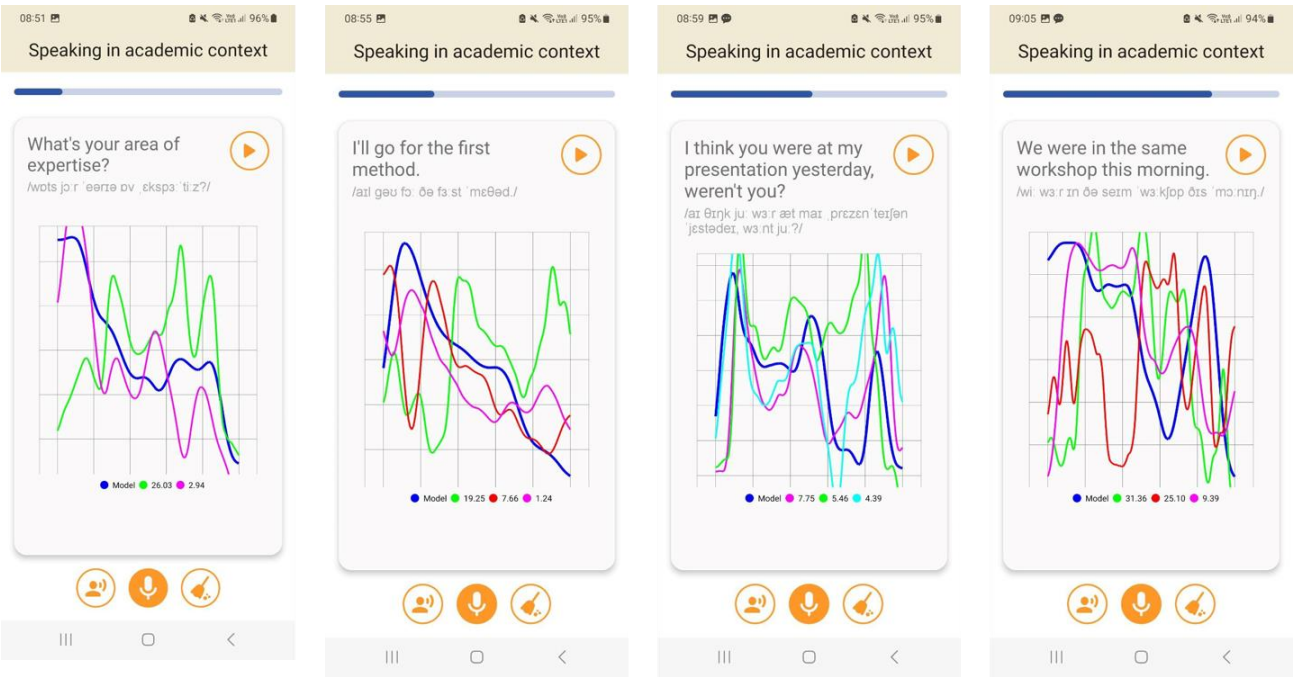

Fig. 3. Typical ADZ scenario.

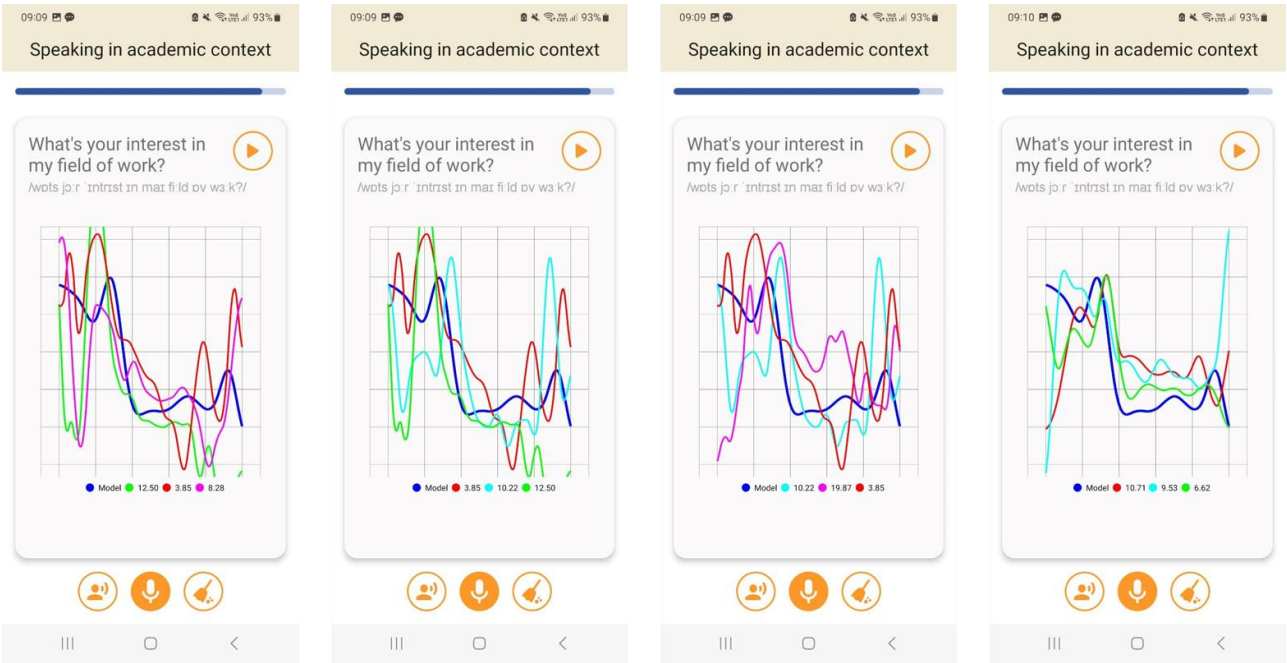

Fig. 4. Successful ZPD scenario.

V. Discussion

In previous study [12] we made an attempt to take advantage of a threefold view upon interaction of learner and CAPT: not only a learner and a CAPT are both dynamic and linguistic systems in contact; but also these are a learner, who develops, and a teacher, who is supposed to mediate this development (and, by and large, develops too). This view leads us to the technique of locating learner's ZPD with the support of formal mathematical models. The question is what to do if we are able to successfully locate the learner's ZPD. We performed some work to find the way towards DST-grounded scaffolding through ZPD, where the exercises can be classified based not only on conversational situations but on syntactic or discourse tags, such as open question, exclamation, and long affirmative sentence.

In this work we elaborate on this approach and show that performing task classification based on the located ZPD can help language learners move beyond plateau. The phenomenon of the language learning plateau is understood as a temporary stagnation in the development of second language (L2) learners. This typically occurs as learners are transitioning to a more advanced level after achieving a certain degree of proficiency. Under certain pedagogical procedures, such as via our CAPT system, the language learning plateau in foreign language learning development can be overcome [1, 3, 42–45].

Since the Zone of Proximal Development (ZPD) is as limitless as a sky without a ceiling, setting a definitive goal becomes an elusive endeavor. However, employing the sandwich paradigm of test-teach(mediation)-test, as described by Kozulin [21], allows us to gauge the learner's relative altitude from the baseline.

The experimental setup and observations reported in our previous study [12] demonstrate good correlation with the known phenomenon that learners' progress and performance does not necessarily increase linearly, but passes through periods of progression and regression alternatively [34–36].

VI. Conclusion

The preliminary experiments within the scope of this and previous studies confirm the possibility to locate and quantitatively evaluate the zones of proximal development of individual learners using longitudinal pronunciation assessment supported by DTW and cross-recurrence quantification analysis. This research is a step towards better understanding how dynamic assessment can be performed during the process of phrasal intonation training with the help of a CAPT system, and how to the learners' movement from one developmental level to another can be determined and supported. One of major challenges we observed is that while working on exercises through learner's ADZ, DTW (or any other possible metrics) are either immediately good or rapidly and steadily converge to good values. Thus, it indicates that two phonological systems (L1 and L2) are synchronizing with each other. As learners transition to a new level of proficiency and enter the ZPD, they may encounter tasks where their performance, as measured by DTW is not strong. However, with targeted instruction that supports their prosodic memory—even for a short period—they might achieve a high DTW score on a single attempt. This means a student can produce a good result once, but may not be able to maintain this level of performance. Instead, learners might hit a plateau in their learning progress, which is the main focus of this paper.

Dynamic assessment involves two crucial elements: (1) the learner's assessment and (2) tailored intervention, which is the stimulation toward challenging yet achievable objectives formulated based on the assessment. While one can assess a learner's developmental progress by determining their elevation from the ground level, it is essential to adjust the extent of mediation dynamically, adapting to the learner's current developmental stage.

The issue concerning the most desirable attentional states in CAPT remains unresolved. Lacking sufficient empirical evidence, researchers still differ in their views on whether learners' attention should be explicitly or implicitly directed to computerized visualization cues.

To conclude, let is point out that the focus of this particular contribution is not on the experiments, but on the theoretical and preliminary empirical grounds for using a CAPT system for moving beyond the learning plateau. Further extensive experiments with involving participants with a diverse range of their L1 are surely required.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

N.B. coordinated the development of StudyIntonation CAPT system, E.P. supervised the software design and the project on CAPT feedback tailoring, J.B. contributed to the assessment process and formulation of linguistics and language teaching grounds for the system. The authors jointly wrote the paper. All authors have approved the final version.

Funding

This research was funded by the Japan Society for the Promotion of Science (JSPS), grant number 23K00679 "Theory, Methodology and Tools for Tailoring CAPT Feedback".

Acknowledgment

We would like to express our sincere gratitude to Prof. Nicholas Carr for his very helpful comments in the margins of the initial manuscript and for his valuable suggestions contributing to better positioning of this paper in scope of SLD, SCT and DA theories and educational practices.

We also highly appreciate our colleagues Elena Boitsova, Akemi Kusakari, Iurii Lezhenin, Roman Svechnikov, Andrey Ostapchuk, Dmitrei Efimov, Veronica Khaustova, Victor Khaustov, as well as many involved students for their contribution to conceptualization, design, and implementation of the core components of *StudyIntonation* environment at its different stages. Their efforts created necessary practical grounds for the current study.

References

[1] J. C. Richards, *Moving beyond the Plateau*, Cambridge University Press New York, NY, 2008.

[2] L. J. Murphy, "Why some ESL students experience a language learning plateau while others do not," Ph.D. dissertation, Teachers College, Columbia University, 2017.

[3] D. Utegulova, "Overcoming language learning plateau: EFL teachers' perceptions and practical recommendations," M.S. Thesis, Suleyman Demirel University, 2022.

[4] W. Acton, "Changing fossilized pronunciation," *TESOL Quarterly*, vol. 18, no. 1, pp. 71–85, 1984.

[5] Y. Lezhenin, A. Lamtev, V. Dyachkov, E. Boitsova, K. Vylegzhanina, and N. Bogach, "Study intonation: Mobile environment for prosody teaching," in *Proc. 2017 3rd IEEE International Conference on Cybernetics (CYBCONF)*. IEEE, 2017, pp. 1–2.

[6] M. C. Pennington, "Teaching pronunciation: The state of the art 2021," *RELC Journal*, vol. 52 No.1, 2021, pp. 3–21.

[7] M. C. Pennington and P. Rogerson-Revell, "Using technology for pronunciation teaching, learning, and assessment," *English Pronunciation Teaching and Research*, Springer, 2019, pp. 235–286.

[8] E. Boitsova, E. Pyshkin, Y. Takako, N. Bogach, I. Lezhenin, A. Lamtev, and V. Diachkov, "StudyIntonation courseware kit for EFL prosody teaching," in *Proc. 9th International Conference on Speech Prosody 2018*, 2018, pp. 413–417.

[9] N. Bogach, "Languages and cognition: towards new call," in *Proc. 3rd International Conference on Applications in Information Technology*. ACM, 2018, pp. 6–8.

[10] E. Pyshkin, J. Blake, A. Lamtev, I. Lezhenin, A. Zhuikov, and N. Bogach, "Prosody training mobile application: Early design assessment and lessons learned," in *Proc. 2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, vol. 2. IEEE, 2019, pp. 735–740.

[11] N. Bogach, E. Boitsova, S. Chernonog, A. Lamtev, M. Lesnichaya, I. Lezhenin, A. Novopashenny, R. Svechnikov, D. Tsikach, K. Vasiliev *et al.*, "Speech processing for language learning: A practical approach to computer-assisted pronunciation teaching," *Electronics*, vol. 10, no. 3, p. 235, 2021.

[12] V. Mikhailava, J. Blake, E. Pyshkin, N. Bogach, S. Chernonog, A. Zhuikov, M. Lesnichaya, I. Lezhenin, and R. Svechnikov, "Dynamic assessment during suprasegmental training with mobile CAPT," in *Proc. Speech Prosody*, vol. 2022, 2022, pp. 430–434.

[13] L. S. Vygotsky, *Thought and Language*, MIT press, 2012.

[14] J. M. Alharbi, "Insight into the role of interaction in language acquisition: Vygotsky's interactionist theory of language," *Arab World English Journal (AWEJ)*, vol. 14, no. 2, pp. 281–294, 2023.

[15] J. P. Lantolf and M. E. Poehner, *Sociocultural Theory and the Pedagogical Imperative in L2 Education: Vygotskian Praxis and the Research Practice Divide*, Routledge, 2014.

[16] L. S. Vygotsky and M. Cole, *Mind in Society: Development of Higher Psychological Processes*, Harvard university press, 1978.

[17] J. P. Lantolf, S. L. Thorne, and M. E. Poehner, "Sociocultural theory and second language development," *Theories in Second Language Acquisition*, Routledge, 2014, pp. 221–240.

[18] J. P. Lantolf and G. Appel, "Theoretical framework: An introduction to Vygotskian perspectives on second language research," *Vygotskian Approaches to Second Language Research*, Ablex Publishing Corporation, 1994.

[19] C. S. Lidz, *Practitioner's Guide to Dynamic Assessment*, Guilford Press, 1991.

[20] R. Feuerstein, Y. Rand, and J. E. Rynders, *Don't Accept Me as I Am: Helping "Retarded" People to Excel*, Springer, 2013.

[21] A. Kozulin and E. Garb, "Dynamic assessment of EFL text comprehension," *School Psychology International*, vol. 23, no. 1, pp. 112–127, 2002.

[22] M. E. Poehner, "Both sides of the conversation: The interplay between mediation and learner reciprocity in dynamic assessment," *Sociocultural Theory and the Teaching of Second Languages*, pp. 33–56, 2008.

[23] A. Kusakari, "Effect of dynamic assessment on japanese EFL learners' comprehension and production of counterfactual conditionals," *Tohoku English Education Society Research Bulletin*, no. 39, pp. 129–144, 2019.

[24] A. Nicholas, "Dynamic assessment and requesting: Assessing the development of Japanese EFL learners' oral requesting performance interactively," *Intercultural Pragmatics*, vol. 17, no. 5, pp. 545–575, 2020.

[25] D. Y. Devrim, "Theorizing written feedback as a mediation tool within the Sydney school's genre pedagogy: a focus on ZPD and scaffolding," *Functional Linguistics*, vol. 1, pp. 1–13, 2014.

[26] D. Larsen-Freeman, "Chaos/complexity science and second language acquisition," *Applied Linguistics*, vol. 18, no. 2, pp. 141–165, 1997.

[27] D. Larsen-Freeman, "On language learner agency: A complex dynamic systems theory perspective," *The Modern Language Journal*, vol. 103, pp. 61–79, 2019.

[28] M. Koopmans, "Education is a complex dynamical system: Challenges for research," *The Journal of Experimental Education*, vol. 88, no. 3, pp. 358–374, 2020.

[29] P. Hiver, A. H. Al-Hoorie, and R. Evans, "Complex dynamic systems theory in language learning: A scoping review of 25 years of research," *Studies in Second Language Acquisition*, pp. 1–29, 2021.

[30] D. LaScotte, C. Meyers, and E. Tarone, "Voice and mirroring in SLA: Top-down pedagogy for L2 pronunciation instruction," *RELC Journal*, vol. 52, no. 1, pp. 144–154, 2021.

[31] C. L. Nagle, "Assessing the state of the art in longitudinal l2 pronunciation research: Trends and future directions," *Journal of Second Language Pronunciation*, vol. 7, no. 2, 2021, pp. 154–182.

[32] N. Esteve-Gibert and B. Guellaï, "Prosody in the auditory and visual domains: A developmental perspective," *Frontiers in Psychology*, vol. 9, p. 338, 2018.

[33] P. L. Geert, "Dynamic systems, process and development," *Human Development*, vol. 63, no. 3-4, pp. 153–179, 2019.

[34] P. Chang and L. J. Zhang, "A CDST perspective on variability in foreign language learners' listening development," *Frontiers in Psychology*, vol. 12, p. 21, 2021.

[35] M. Verspoor and K. Bot, "Measures of variability in transitional phases in second language development," *International Review of Applied Linguistics in Language Teaching*, vol. 60, no. 1, pp. 85–101, 2022.

[36] J. Hodgetts, *Pronunciation Instruction in English for Academic Purposes: An Investigation of Attitudes, Beliefs and Practices*, Springer, 2020.

[37] P. Geert, "Vygotskian dynamics of development," *Human Development*," vol. 37, no. 6, pp. 346–365, 1994.

[38] J. P. Lantolf, L. Kurtz, and O. Kisselev, "Understanding the revolutionary character of L2 development in the ZPD: Why levels of mediation matter," *Language and Sociocultural Theory*, vol. 3, no. 2, pp. 153–171, 2016.

[39] R. Ableeva, "Dynamic assessment of listening comprehension in second language learning," Ph.D. dissertation, The Pennsylvania State University, College of the Liberal Arts, 2010.

[40] P. Geert, "A dynamic systems model of basic developmental mechanisms: Piaget, Vygotsky, and beyond," *Psychological Review*, vol. 105, no. 4, p. 634, 1998.

[41] S. Krashen, "The case for comprehensible input," *Language Magazine*, vol. 7, no. 1, pp. 1–6, 2017.

[42] M. Mirzaei, M. Zoghi, and H. Davatgari Asl, "Understanding the language learning plateau: A grounded-theory study," *Teaching English Language*, vol. 11, no. 2, pp. 195–222, 2017.

[43] M. Q. Al-Shormani, "Fossilization and plateau effect in second language acquisition." *Language in India*, vol. 13, no. 2, 2013.

[44] Q. Xu, "Moving beyond the intermediate EFL learning plateau," *Asian Social Science*, vol. 5, no. 2, pp. 66–68, 2009.

[45] F. Yi, "Plateau of EFL learning: A psycholinguistic and pedagogical study," 2011.

Copyright © 2024 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).