PROTUTOR: A PRONUNCIATION TUTOR THAT USES HISTORIC OPEN LEARNER MODELS

A Thesis Submitted to
the College of Graduate Studies and Research
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in the Department of Computer Science
University of Saskatchewan
Saskatoon

By

Carrie Demmans Epp

DISCLAIMER

The ProTutor system was exclusively created to meet the thesis and/or exhibition requirements for the degree of Master of Science at the University of Saskatchewan. Reference in this thesis/dissertation to any specific commercial products, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement, recommendation, or favoring by the University of Saskatchewan. The views and opinions of the author expressed herein do not state or reflect those of the University of Saskatchewan, and shall not be used for advertising or product endorsement purposes.

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
S7N 5C9

ABSTRACT

Second language learners face many challenges when learning a new language. To determine which challenges learners needed additional support in overcoming, we conducted a needs assessment of the Russian language program at the University of Saskatchewan and found that their students needed the most help with speaking in Russian. As a result, we designed an Intelligent Tutoring System (ITS) to help students learn how to pronounce Russian properly. We hoped to alleviate some of the challenges that learners face when learning to pronounce words in a second language by building an ITS that uses a Historic Open Learner Model (HOLM) to encourage learner reflection and to help maintain learner motivation. We designed, built, and performed a formative evaluation of a system, called ProTutor, using beginner learners of Russian as a second language at the University of Saskatchewan. This evaluation showed that learners have a positive perception of HOLMs and of the system as a whole. However, ProTutor needs further evaluation in order to determine its effectiveness as a learning aide.

ACKNOWLEDGMENTS

This work is largely due to the guidance provided by Professor Gord McCalla. Gord was kind enough to keep me exploring and point me in interesting directions when I was far too focused on my goals. I would also like to thank Chris Brooks, Greg Logan, and Clayton Epp for all of the technical advice that they provided while I was designing and developing ProTutor.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Second language learners face many challenges. A lack of time to learn the language and decreasing motivation because of a lack of visible progress are among those problems (Johnson, Wu, & Nouhi, 2004). The problem of decreasing motivation because of a lack of visible progress is especially obvious in language pronunciation classes. Without motivation, students withdraw from educational activities and are unsuccessful in their attempts to acquire skills in their chosen second language (Archibald & O'Grady, 2008). This makes it especially important to support and encourage students so that they remain motivated, since motivation is required for continued learner involvement and the acquisition of new language skills (Vosniadou, 2001).

Second language students must learn how to produce and recognize the individual phones (individual unique sounds made by alphabet characters) that are found in the language that they are learning. They must also learn how to combine these phones so that they can properly pronounce them in combinations in order to say words and later sentences. An inability to recognize and pronounce the language's phones can lead to miscommunications that can increase the risk that a learner is exposed to when s/he is immersed in a second language environment. This inability could result in the learner taking a wrong turn and winding up in an unsafe area or insulting someone who may do him or her harm.

Furthermore, adult second language learners rarely receive adequate training in pronunciation and oral interaction. They often spend less than one hour a week speaking in the foreign language that they are learning. This creates a challenge when the learner is immersed in a foreign language environment since learners cannot control the immersion environment; they can only control their reaction to it. Learners exercise this control through their ability to communicate and the environment responds accordingly. In everyday activities, communication

is performed orally since it is difficult to write a note when you run into someone on the street or need help at the grocery market. While the immersion environment may provide the fastest and toughest way to learn the language, it provides the learner with little support.

We created a second language pronunciation tutor, called ProTutor, that supports the learner and reduces the risks associated with miscommunication by giving the learner additional practice before s/he is immersed in a second language environment. This also reduces the risks associated with errors by providing a sheltered environment where the learner can practise without having to worry about the penalties, such as a loss of grades or embarrassment, that would typically be administered in a second language immersion environment or classroom (Archibald & O'Grady, 2008).

ProTutor has two types of Open Learner Model to help students see their progress when learning to speak a second language. The first type of Open Learner Model (OLM) is the standard one that allows the learner to see a representation of his or her current performance and the second type of OLM creates a Historic Open Learner Model (HOLM) by adding a historic component to the OLM. The creation of the HOLM in this way communicates learner progress by allowing the learner to see his or her model at different times in his or her training. By adding previous performance information to create the Historic Open Learner Model, we have allowed learners to more clearly see how their model has changed so that they can see their progress, which helps them remain motivated (Vosniadou, 2001). It is also hoped that this will encourage learner reflection about which activities have contributed to performance changes so that the learner can make more informed decisions about his or her learning activities.

ProTutor's OLM and HOLM visually represent the learner's abilities by using text and images. The learner is informed and encouraged to reflect and compare the system's visualization of his or her abilities to that of an expert's through a visual representation of both the learner's pronunciation and a native speaker's pronunciation. This representation is shown in the models and exhibits the learner's strengths and weaknesses through a short piece of text. Both of the learner models in ProTutor also display a group model that serves to inform and motivate learners by allowing them to see a representation of the whole group's pronunciation strengths and weaknesses alongside their own. The HOLM adds to this by showing the learner

parts of his or her model from a previous time as well as his or her current model so that learners can clearly see their progress.

In addition to the use of the above mentioned open learner modeling, ProTutor hopes to help learners overcome motivational challenges through the use of a variety of activities and positive reinforcement (Vosniadou, 2001). Learners are given the opportunity to practise their newly acquired language pronunciation skills in a low-pressure environment since their use of ProTutor will not be a measured part of their grade. Rather, ProTutor provides carefully designed supplementary instruction and practice that will help the learner acquire oral proficiency in the new language.

ProTutor was built as a proof-of-concept web-application and released in stages throughout the fall of 2009. Its design and implementation were limited in scope because of time and technological constraints. However, this staggered release of the features being studied allowed us to perform a small formative evaluation with learners from introductory Russian as a second language classes at the University of Saskatchewan (U of S). This evaluation, while limited, indicated that the addition of a HOLM was well received by learners and that the ProTutor system when used with the HOLM was perceived to have helped learners. Furthermore, it appears to have kept learners working on their pronunciation of the second language.

The above mentioned motivations, system development, and evaluation are further developed in subsequent chapters. In Chapter 2, we discuss the relevant previous work and how it relates to ProTutor. We follow this with a discussion of ProTutor's design and architecture in Chapter 3. The formative evaluation that we performed and the evaluation's results are described in Chapter 4. This is followed with a discussion of the implications of and potential contributing factors to these results in Chapter 5, and in Chapter 6 we finish our discussion by tying everything together and indicating the future directions that ProTutor could take.

CHAPTER 2

RELATED WORK

There has been much work in the area of Intelligent Tutoring Systems (ITS), especially in the domain of language learning. Many of these systems create models of the learner, regardless of their domain of application (Johnson, Wu, & Nouhi, 2004) (Burston, 2008) (Pantelia, 2007) (Tao, 2007) (Beck, Jia, & Mostow, 2003) (Blair, Schwartz, Biswas, & Leelawong, 2006) (Martin & Nicholas, 2007) (Price, McCalla, & Bunt, 1999) (Tan & Biswas, 2006) (Tsiriga & Virvou, 2002) (Vocab Box, 2000) (Wang & Seneff, 2007) (Extemporel Inc., 2008). There are also commercially available language ITSs that do not create learner models (Lord, 2008) (Collentine, 2002) (Papadopoulou, 2008) (Pardo Ballester, 2009) (Salmerón & Burston, 2008) (Teach 2000, 2010) (Hofs, 2002). When ITSs use Open Learner Models (OLM) they do so for a variety of reasons; the primary reasons are to give learners the opportunity to inform, self-regulate, and reflect about the current state of their knowledge which helps promote the development of accurate knowledge among learners (Vosniadou, 2001) (Tan & Biswas, 2006) (Bull & Kay, 2007).

In the case of the Tactical Language Training System (TLTS), BonPatron, Filoglossia+, Chuala, Eyespeak, and Microsoft's Reading Tutor, the system allows the learner to playback what s/he has said so that the learner can determine how s/he is progressing (Johnson, Wu, & Nouhi, 2004) (Burston, 2008) (Pantelia, 2007) (Tao, 2007) (Extemporel Inc., 2008) (Li, Ju, Deng, & Acero, 2007). For the purposes of this discussion, we will call this feature simple playback. Some of these simple playback systems only allow the learner to listen to an utterance immediately following its initial recording while others, like the TLTS, allow the learner to replay their utterances at a later time. In one way, this might be viewed as providing the learner with historic information about his or her performance. However, it is only historic in the sense that the recording is from an earlier portion of the same training session. We mean for data in the learner model to be considered historic when it is from previous training sessions and is

capable of showing the learner how his or her performance has changed rather than what his or her current performance is like. In addition to not providing historical information to the learner, the systems that provide simple playback do not provide learners with supplementary information about the quality of their pronunciation while they listen to their speech. This provides them with no more additional functionality than a tape recorder would, and it means that inexperienced learners may be unable to identify their mispronunciations (Archibald & O'Grady, 2008).

While some systems, like the TLTS, do not provide supplementary information along with their learner model, others do. Betty's Brain uses learner models and allows the learner to test Betty, who is a representation of the learner's knowledge (Tan & Biswas, 2006). This provides the learner with supplementary information that reveals Betty's misconceptions. This supplementary information models Betty's knowledge and reveals his or her misconceptions, which allows learners to identify and correct shortcomings in their knowledge by correcting Betty's misconceptions. ProTutor uses a similar approach by providing learners with a visualization of both their and an expert's pronunciation so that they can identify their strengths and weaknesses by comparing their performance to the revealed ideal performance. While this does not reveal the supplementary information in the same way, it still provides it to the learner so that s/he can more easily identify what s/he needs to improve upon.

2.1. PEDAGOGICAL APPROACHES AND DELIVERY TECHNOLOGIES

Various pedagogical approaches and environments are used to help learners acquire knowledge. The Betty's Brain team uses the learning by teaching approach to achieve its pedagogical goals. They do this because teaching requires the structuring and explanation of information. This approach also requires the teacher to be responsible for ensuring that the agent being taught learns the subject matter (Blair, Schwartz, Biswas, & Leelawong, 2006) (Tan & Biswas, 2006). However, this approach is uncommon in language learning and was not used by a single one of the language ITSs that we studied (Table 2.1). ITSs for language learning or Computer Assisted Language Learning (CALL) tools tend to use one or both of two pedagogical approaches:

- drill and practice
- simulation

Drill and practice CALL tools are well suited to beginner learners, they tend to follow more traditional pedagogical approaches to language learning, and they are likely the simplest to implement. Historically, many of the drill and practice CALL systems have been text based (Martin & Nicholas, 2007) (Vocab Box, 2000) (Teach 2000, 2010) (Pumpkimber, 2006). This is partly because text is the most appropriate medium for what these ITSs were trying to achieve. Like most text based CALL, they were trying to teach grammar, spelling, reading comprehension, or composition. This makes text an appropriate medium since these three aspects of language learning are by their very nature well suited to textual and graphical representations. An example of this type of CALL would be Martin and Nicholas's use of a constraint-based model to teach learners how to decline German adjectives (Martin & Nicholas, 2007). Even though many ITSs use the drill and practice pedagogical approach, many other CALL tools implement drill and practice exercises through games (Pumpkimber, 2006) (International Cafe, 2007) (Mich, Betta, & Giuliani, 2004) (Word Confusion, 2008) (Translator Alligator, 2008). A classic example of a text based version of this is Word Munchers, which was commonly used in North American schools during the 1980s and 1990s to help children improve their vocabulary (Pumpkimber, 2006).

2.1.1. Speech Recognition

This text only based approach, while common, is being replaced by multimodal approaches as it becomes easier for other technologies to be integrated into an ITS. One technology that is becoming more common in CALL systems is speech recognition. This technology is especially useful for CALL applications that want to help learners with oral aspects of language acquisition.

INTELL uses Carnegie Mellon's Sphinx III speech recognition system, along with speech synthesis, to help people learn how to read and write in an Indian language called Telegu (Nagamani, Narendra prasad, & Girija, 2005). Like INTELL, project LISTEN uses speech recognition software and reading practice to help learners improve their oral reading proficiency in their native tongue (Beck, Jia, & Mostow, 2003). The focus of project LISTEN is to help

primary school aged children improve their oral reading skills in English, and it is only one of the projects with this goal. There is also a group out of Microsoft that uses speech recognition software in an Automatic Reading Tutor system for children (Li, Ju, Deng, & Acero, 2007). Both of these groups have seen relative success at helping children improve their native tongue oral reading proficiency by using speech recognition technologies with the foreknowledge of what the learner should be saying because s/he is reading a text. We are hoping that by having ProTutor use a similar approach to the above oral reading projects, we will be able to help adult second language learners increase their oral proficiency.

While speech recognition is useful in helping native language speakers (Beck, Jia, & Mostow, 2003) (Li, Ju, Deng, & Acero, 2007) (Mich, Betta, & Giuliani, 2004), it can benefit second language learners even more since they have fewer opportunities to practise their language skills. Wang and Seneff have created a spoken language translation game where learners are given a sentence in English and must then say the equivalent sentence in Chinese (Mandarin) (Wang & Seneff, 2007). Their system has a limited domain and application, but the learners found it useful since it helped them acquire the oral language necessary to work in the airline industry. Jianguo and Xiaozhen help learners of Mandarin at a different level of granularity than Wang and Seneff do. Both of their systems use speech recognition to teach learners how to speak in Mandarin, but Jianguo and Xiaozhen focus on individual characters, whereas Wang and Seneff focus on the learner's ability to speak in complete sentences. Furthermore, Jianguo and Xiaozhen's system teaches learners how to write Mandarin characters by using text and images (Jianguo & Xiaozhen, 2004).

Another system that uses speech recognition to help its learners is pARLiNG, which was created by Mich, Betta, and Giuliani. It uses multimedia tools in order to get children to practise reading and listening to a second langauge, English, that they are learning. pARLiNG also creates a model of the learner based on his or her interactions with the system so that it can adapt the vocabulary found in the games that the learner plays in order to increase his or her vocabulary (Mich, Betta, & Giuliani, 2004). ProTutor uses a similar approach to personalizing a learner's activities, but rather than selecting words that the learner does not know from stories that s/he's read, ProTutor selects words based on the sounds that the learner is having trouble pronouncing.

While both pARLiNG and ProTutor use multimedia activities and speech recognition, there are many commercial CALL systems that do not use speech recognition software. These systems have added other output modes to their systems in order to make their systems more engaging and useful to the learner. The addition of richer graphics and audio examples, whether provided through recordings or text to speech, have become more common. Fifteen of the twenty-four CALL applications that I analyzed use graphics in their materials, and twelve of the nineteen commercially available software packages provide audio examples to the learner, usually through the use of recordings (Table 2.1). In contrast to this, three of the five research based CALL applications use text to speech to provide an audio mode to the learner, the fourth system provides an audio mode through the use of recordings, and the fifth system has no audio output mode (Table 2.1).

2.1.2. Learning by Testing

While most of the research-based drill and practice language ITSs use text to speech and speech recognition software to help their learners acquire language proficiency, it is far more common for commercially available drill and practice software packages to require that learners acquire knowledge and skills through the use of tests. Vocatude is one of these learning by testing software packages. It allows learners to listen to the correct pronunciation of words and test their knowledge of pronunciation and spelling through a dictation feature (Hofs, 2002). Vocatude's use of Text to Speech (TTS) is well suited to training and improving a beginner learner's listening comprehension and it is one of the few commercially available learning by testing software packages that does this. Some other CALL software packages that use learning by testing are Teach2000, Vocab Box, RosettaStone, BiLingo Kidz, and Filoglossia+ (Pantelia, 2007) (Vocab Box, 2000) (Lord, 2008) (Teach 2000, 2010) (RosettaStone, 2008). Teach2000 and Vocab Box allow learners to test their second language vocabulary knowledge in a text-based environment. However, Vocab Box also builds a learner model by tracking learner activities and creating a model of the learner's knowledge from those activities. It then uses this model to train areas where the learner has exhibited weak vocabulary skills.

2.1.3. Instruction

There are also commercially available language ITSs that exclusively use instruction to impart knowledge to the learner. An example of this approach to language education is evident in Chuala. It uses language lessons along with an open model of an expert speaker's pronunciation, in the form of recordings, so that learners can compare their pronunciation to that of an expert (Extemporel Inc., 2008). Another system, en una palabra, also exclusively uses instruction to help its learners acquire listening skills, vocabulary, and cultural knowledge of an area in Argentina (Pardo Ballester, 2009). While this form of CALL is uncommon, it is easier to build and can later be expanded by adding other pedagogical approaches and interactive activities.

2.1.4. Combined Pedagogical Approaches

Unlike Chuala and en una palabra, the majority, of the studied commercial CALL systems, do not adhere to any one approach, but use a combination of approaches (Table 2.1). RosettaStone is one of these systems. It is the only one that combines multimedia activities, tests, simulation, and instruction to achieve its goal of helping the learner obtain language proficiency (RosettaStone, 2008). Other systems that combine a subset of the pedagogical approaches used by RosettaStone are BiLingo Kidz, Bon Patron, Filoglossia+, Eyespeak, los aztecas, TESOROS, International Café, Teach2000, TLTS, Vocab Box, Vocatude, and Read, Write & Type!. Bilingo Kidz and Teach2000 combine games, with learning by testing, and drill and practice in order to get their younger target audiences to learn basic language skills. While Teach2000 is only a vocabulary acquisition tool (Teach 2000, 2010), Bilingo Kidz focuses on getting children to learn vocabulary and develop their Spanish reading and listening skills (Lord, 2008).

Both Bilingo Kidz and Teach2000 forgo instruction while trying to achieve their goals. Other systems, like Bon Patron, EyeSpeak, and INTELL, combine drill and practice with instruction to meet their learners' needs. Bon Patron is a French composition aide that gives learners feedback about any compoisition that they enter; it uses a non-persistent learner model that contains information about the learner that is obtained from him or her via a survey every time that s/he uses the system. BonPatron then uses the survey responses to analyze the learner's composition

and provide him or her with feedback and instruciton that is similar to that provided by Microsoft Word when performing a grammar check (Burston, 2008).

One of the other commercially available CALL systems, Eyespeak, that combines instruction with drill and practice focuses on the learner's oral production of the language rather than his or her ability to compose text in the language. Eyespeak uses speech recognition to analyze learners' oral production of the language. It then provides learners of English as a second language with feedback about their pronunciation as they perform drills and receive instruction. It does all of this while providing an expert model and maintaining a learner model that allows for simple playback (Tao, 2007). INTELL is the third system that combines instruction with drill and practice. It is unique among the systems that use this combined approach because it uses speech recognition software and a learner model while helping learners of Telegu, but it is not commercially available (Nagamani, Narendra prasad, & Girija, 2005).

Drill and practice has also been combined with learning by testing in several commercially available systems. Some of the systems we found that did this are: Filoglossia+, Vocab Box, and Vocatude. Filoglossia+ uses this combination of pedagogical approaches to increase beginner to intermediate learners' listening and reading comprehension of modern Greek. However, it only provides information about the correctness of a learners' input and no feedback as to the cause of the error or how to fix the error. So, the learner must keep trying until s/he happens upon the correct answer. In this way, Filoglossia+'s developers hoped that learners would acquire knowledge through the tests and practice that they perform (Pantelia, 2007).

Vocab Box adds a basic closed learner model to the combined drill and practice and learning by testing pedagogical approaches. It uses flashcards to test the learner's knowledge, creates a model of the learner's abilities, and then recommends flashcards to practice based on the learner's performance (Vocab Box, 2000). The third system that uses this pedagogical appraoch, Vocatude, does not include a learner model, but it does allow the learner to customize the system so that s/he can learn any vocabulary in any language. This is because Vocatude requires that the learner enter the vocabulary that s/he wants to learn before s/he is given activities for and tested on that vocabulary (Hofs, 2002).

Several of the commercially available CALL systems use instruction, drill and practice, and simulation or games to achieve their goals. TESOROS, TLTS, and Read, Write, and Type! are among these tools. TLTS uses a combination of drill and practice alongside instruction and simulated conversation in a game based environment to teach learners how to speak (compose sentences and pronounce words) in a foreign language (Johnson, Marsella, & Vilhjalmsson, 2004). TESOROS tries to help learners acquire language skills at a lower level than the TLTS does. TESOROS is mainly focused on basic vocabulary and grammar acquisition, and it has a secondary focus of developing the learner's listening and reading comprehension. It does this by sending learners on an investigative quest, similar to the TLTS's mission based approach, but it does not require the learner to interact verbally. Rather, it requires the learner to read and listen to comic strips in order to obtain information and solve the mystery (Collentine, 2002). Read, Write & Type! has an even more basic goal: to teach children how to sound out, read, and spell regularly spelled English words through the use of games and by providing instruction and drills to learn and to practise their newly acquired skills (Papadopoulou, 2008).

While these systems exemplify the recent increase in simulation and game based educational systems for language learning, there are others that use a different subset of pedagogical approaches. There are two such CALL systems that allow learners to order food in a second language. One is a game and the other is a simulation. The game is called International Cafe and it allows learners to read along with a simulated native speaker when s/he orders the food that the learner has selected (International Cafe, 2007). International Cafe allows learners to practise their reading and listening comprehension, whereas the other system, L2tutor, allows learners to interact with a restaurant server in order to practise their second language composition skills and receive feedback about their performance (Price, McCalla, & Bunt, 1999). It focuses on the learner's ability to interact by writing responses in the form of sentences that are appropriate given the context of the situation. Some newer systems like the TLTS, which also teaches learners how to interact in a foreign language, focus on getting learners to speak in the second language while using grammatically correct sentences (Johnson, Marsella, & Vilhjalmsson, 2004), whereas others, like los aztecas, are meant to teach narration and the use of several forms of the past tense through games and instruction (Salmerón & Burston, 2008).

In the same vein as los aztecas, Teach2000 uses games rather than simulation to help learners acquire second language vocabulary (Teach 2000, 2010). Teach2000 is not unique in its approach to helping language learners. There are many other commercially available games that can help language learners. Flip Words, Translator Alligator, and Word Confusion are among them. FlipWords combines the concept of crossword puzzles and Wheel of Fortune to help its learners identify words and phrases (Flip Words, 2008). This game exposes learners to idioms and cultural information as well as helps them learn how to spell. The Translator Alligator game helps learners acquire second language vocabulary by asking them to select the correct translation of a displayed word (Translator Alligator, 2008), and Word Confusion helps with vocabulary acquisition and understanding by asking the learner to select the correct word to complete the sentence (Word Confusion, 2008). While both Translator Alligator and Word Confusion are drill and practice based, the drills that they expect the learner to do are all in the form of games.

2.2. SUMMARY

While all of the above mentioned systems and pedagogical approaches help learners acquire second language skills, not all of them are appropriate for use in pronunciation tutoring and few of them provide this functionality. The learning by teaching and learning by testing approaches used by many systems (Tan & Biswas, 2006) (Vocab Box, 2000) (Teach 2000, 2010) (Hofs, 2002) will be difficult to use for pronunciation tutoring with beginner language learners because these learners do not have the necessary background knowledge to teach one-another and only telling learners that they mispronounced something is insufficient to produce a change in their behaviour; they need additional information about how to fix the error, what error they made, or what the expected performance is.

The learning by testing approach provides no instruction to the learner so that s/he can learn how to fix mispronunciations and the beginner learner does not yet possess this knowledge from other sources. Furthermore, since the beginner level of learner may not possess the procedural knowledge necessary to properly pronounce the phones, s/he cannot teach another learner how to form the sounds. In addition to these limitations, it would be inappropriate to have a learner

who is trying to learn how to perform a physical action, such as producing a phone, teach someone else how to perform the task since the other person's ability to perform the task does not reflect upon the instructor's ability to perform the task, nor does it help the learner to be able to properly pronounce the language's phones, which is ProTutor's primary educational goal. A teacher who successfully gets a student to properly pronounce a phone may not be able to pronounce that phone himself or herself, but s/he can use audio and video clips to model the correct performance and provide procedural instruction to achieve the goal, much as a CALL system could.

Regardless of the pedagogical approach used by each of the described systems, none of them seem to use Open Learner Models (OLM) with the goal of encouraging learner motivation. While many of the described systems use learner models, only some of them open their model up to the user in an attempt to inform him or her, and none of those systems provide the learner with historical information about his or her performance or access to a group model (Table 2.1). Even within the field of open learner modeling, no one is known to have studied the use of historical open learner modeling for the purpose of maintaining learner motivation, although some believe that it has potential (Bull & Kay, 2007). ProTutor's use of OLM and HOLM aims to address the use of learner models for informing learners of their abilities as well as their classmate's abilities and of motivating them by allowing them to reflect over changes in their performance and how their performance compares to that of the group.

You can see how ProTutor's system characteristics compare to the above discussed CALL applications in Table 2.1, which summarizes the learner modeling characteristics, pedagogical approaches, modes of interaction, and adaptivity of the studied systems whether they be commercially available or only used within the research community.

Table 2.1: A summary of the characteristics of different CALL applications

✓ - has this property
? - presence of property is unknown

Software	Learner Model						Modes				Adaptive	Pedagogy					Commercially Available
	Constraint-Based	*Simple Playback*	*Complex Representation*	*Open*	*Expert*	*History*	*Text*	*Graphics*	*Audio Input*	*Audio Output*		*Learning by Teaching*	*Drill & Practice*	*Games & Simulation*	*Learning by Testing*	*Instruction*	
BiLingo Kidz							✓	✓		✓			✓	✓	✓		✓
BonPatron	✓	✓		✓			✓						✓			✓	✓
Chuala	?	✓		✓	✓		✓			✓						✓	✓
En una palabra							✓	✓		✓						✓	✓
Eyespeak		✓	✓	✓	✓	✓	✓	✓	✓	✓			✓			✓	✓
Filoglossia+		✓		✓	✓		✓	✓	✓	✓			✓		✓		✓
FlipWords							✓	✓						✓			✓
INTELL			✓				✓	✓	✓	✓	?		✓			✓	
International Cafe	?						✓	✓	✓	✓			✓	✓			✓
LISTEN			✓			✓	✓	✓	✓	✓	✓		✓				
L2Tutor			✓	✓			✓				✓			✓		✓	
los aztecas							✓	✓		✓				✓		✓	✓
Microsoft's Reading Tutor	?	✓	✓	✓			✓	✓	✓	✓	✓		✓				
pARLiNG			✓		✓		✓	✓	✓	✓	✓		✓	✓			
Read, Write & Type!							✓	✓	✓	✓			✓	✓		✓	✓
RosettaStone							✓	✓	✓	✓	?		✓	✓	✓	✓	✓
Word Munchers							✓						✓				✓
Teach2000							✓						✓	✓	✓		✓
TESOROS							✓	✓		✓			✓	✓		✓	✓
TLTS	?	✓	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓		✓	✓
Translator Alligator							✓						✓	✓			✓
Word Confusion							✓						✓	✓			✓
Vocab Box	?		✓				✓				✓		✓		✓		✓
Vocatude		✓					✓		✓	✓			✓		✓		✓
ProTutor	✓		✓	✓	✓	✓	✓	✓	✓		✓		✓	✓		✓	

CHAPTER 3

PROTUTOR DESIGN

ProTutor was designed to be an easily accessible and usable educational tool to support language learners while they were taking a second language class. To determine what type of ITS learners needed, we performed a needs assessment and based our design on the results of this assessment. We then built a distributed web-application, called ProTutor, to provide the functionality that was required by our design. The process of designing ProTutor, its architecture, and its methods of interacting with the learner and giving the learner feedback are described throughout this chapter.

3.1. NEEDS ASSESSMENT

We set out to build an ITS to help people learn languages, but we were uncertain what second language learners need the most help with. There were suspicions that students of second language classes need significant help with learning to speak and pronounce the language that they are learning. These suspicions were based on how little time is spent on pronunciation and speaking in university level second language classes. In order to focus the design of the system that we were planning to build and to see if these suspicions were valid, a needs assessment of the Russian language program at the University of Saskatchewan (U of S) was undertaken in early 2008. This helped ensure that the tool being built would be of use to someone and it helped to secure a group of potential participants to evaluate ProTutor.

We developed a survey, which consisted of two parts, that was used to perform the needs assessment (see Appendix A). The first part used a five point Likert-scale (1 indicated agreement and 5 disagreement) to assess what skills the U of S Russian language students have in each of the four language components: speaking, listening, reading, and writing. The second part of the survey ranked the importance of the four language components and the sub-skills needed to perform well in each of the language components. It was designed to be done on paper because this would allow for the instructors to provide thoughtful and reflective responses

and because it would allow the instructor whose second language was English the time necessary and a medium in which she could accurately and effectively communicate her opinions.

The survey was given to all three of the Russian Language instructors at the U of S: each of the instructors has many years of experience teaching Russian as a foreign language. Two of the instructors completed and returned the survey. The surveys were distributed with instructions to complete one for each level of Russian that they have instructed at the U of S. The first or beginner level of Russian was all courses below RUS217, which is equivalent to the first three language classes, and the second or intermediate level consisted of RUS217 and all language courses that follow it, of which there are three.

The instructors indicated that all of their students need the most help with their speaking skills (See Table 3.1). They also indicated that listening was the second most important skill that their intermediate level students need help with and that the beginner level students need more help with writing than they do with listening.

The beginner Russian students at the U of S are good at asking simple questions orally, summarizing audio materials in text form, and answering simple fact retrieval questions about audio materials. When they are reading they recognize the words that they know, understand single simple sentences, and can answer simple fact-based questions about the text that they have read. When the students are writing they select the correct words and use the correct word order.

Table 3.1: The median of the skill rankings for all levels of student

	Speaking	**Writing**	**Reading**	**Listening**
Skill Importance	1	3	3.5	2.5
Skills Students Need The Most Help With	1	1.5	2	1.5

All of the U of S Russian students need to improve their pronunciation, intonation, and grammar when speaking. At the upper year levels, they also need to increase their rate of speech. These students have poor comprehension of complex audio texts, do not recognize the

words that are spoken, and cannot transcribe audio sessions. They need to improve their ability to summarize and comprehend texts when they are reading, and the beginner level students need to improve their ability to recognize grammatical patterns in the texts that they have read. When writing, these students also need to improve their grammar and increase their rate of composition.

This core need of improving oral production might be met by building a pronunciation tutor, called ProTutor, that also allows students to read Russian aloud. This speaking practice with feedback from the tutoring system should help students identify phones, which is the basic building block of listening comprehension. Along with helping the students identify phones, ProTutor has the primary goal of helping students learn how to pronounce Russian phones and words by using pronunciation explanations, exercises, and games.

3.2. SYSTEM DESIGN

ProTutor is designed to support how people learn, and more importantly how people learn languages. In children, language is developed by first listening to the language, then imitating the language, and later innovating and inventing language (Palmer, 2007). Adults are not in a position to learn languages in this way. However, one of the basic guidelines for developing multimedia CALL systems states that learners need to have opportunities to produce target language output with the expectation that they will be understood if they perform the task properly (Chapelle, 1998). This is why we will still expect learners to produce oral language and ProTutor will understand the learners' utterances through the use of speech recognition software.

While the traditional second language classroom does provide the opportunity for learners to speak in the target language, this opportunity has not been exploited to its full potential; in the traditional second language classroom, adults have usually been taught to read and write in the second language before they are taught to speak in the language. This may be because it usually takes years of listening to a language before a person can begin to speak in that language using complete sentences. However, one's ability to speak can grow quickly when given many opportunities to converse with expert speakers (Palmer, 2007). This means that a learner must be given the opportunity to speak and to hear expert speakers, which is supported by the

multimedia CALL development guideline that states that learners need to have opportunities to produce target language output (Chapelle, 1998). In addition to having learners produce the language, they can be provided with examples from simulated expert speakers, much like the television can provide a useful expert model for children over the age of three (Palmer, 2007).

The feedback provided by the system will model the correct pronunciation of the language much as an expert tutor would (Lu, Di Eugenio, Kershaw, Ohlsson, & Corrigan-Halpern, 2006). It will do this by giving procedural instruction that models the expected behaviour and by opening up a model of an expert's pronunciation. This model will be shown as part of the learner's model and will show how a native Russian speaker or expert would pronounce the same thing that the learner's model is pronouncing.

ProTutor aims to have learners practise their skills by performing drills that reinforce existing knowledge, and by having learners read text that is within their knowledge range and abilities, even if the words they are reading are unfamiliar. To ensure that the text is within the learner's expected knowledge range, our primary source of material is the learner's course textbook, Troika (Nummikoski, 1996). However, we added materials when we found resources that completed or complemented the textbook's material. For example, the textbook chapter that focuses on food only includes a small subset of the vocabulary for fruit and vegetables. So, we added additional fruit and vegetables to the available vocabulary for food themed activities.

The addition of this vocabulary along with the use of several different activities helps provide the learner with a variety of activities, which helps meet the recommendation of educational researchers to use variety in order to maintain learner motivation and to account for individual differences and development (Vosniadou, 2001). To do this, educational researchers recommend using a wide range of materials while providing learners with time to practise their newly acquired and developing skills (Vosniadou, 2001). The variety of vocabulary and the variety of activities that can be performed with text or with text and images will give learners ample opportunity to practise their skills.

Each activity has been designed in such a way that it can easily be added to or removed from the system. The modular design of these activities and the use of the model view controller design pattern allows for new types of activities to be quickly developed and integrated should

the need or desire arise. The only limitation to the addition of new activities is that all of the activity resources be in the database so that the diagnosis of learner errors does not require changes whenever a new activity is added.

The instructional materials were also designed in a modular fashion so that they could easily be added, removed, or updated without requiring major system changes. These materials were designed to help learners by providing procedural instruction about the pronunciation of individual phones that the the learner is having trouble pronouncing. This procedural instruction should help learners to solve future pronunciation problems by increasing learner reflection and teaching the learner how to properly perform the task (Lu, Di Eugenio, Kershaw, Ohlsson, & Corrigan-Halpern, 2006). Lu et al further refined this approach to combining different pedagogical strategies when they found that expert tutors use procedural instruction following student reflection (Lu, Di Eugenio, Kershaw, Ohlsson, & Corrigan-Halpern, 2006). As a result, ProTutor is designed to use procedural feedback immediately after the learner has viewed the open portions of his or her learner model and has chosen to work on improving his or her pronunciation of a specific phone; this also gives learners the needed opportunity to correct their errors (Chapelle, 1998) .

Within ProTutor, learners can move through the sequence of recommended activities at their liberty. This should benefit ProTutor's adult learners, since the ability to exercise control by selecting learning materials is more beneficial to older learners than it is to young learners (Beck J. , 2007). It should also help account for the learner's affective state by giving learners control over the difficulty level of the work they are to perform, which allows them to manage their emotions in the way that they feel works best for them (Picard, 2007). However, a planned path, based on learner progress is delivered to the learner if s/he does not request a modification to the plan. The recommended activities that are found along this path are sequenced and based on the 21 themes found in the chapters of the learner's course textbook. Each of these themes is based on topics that learners encounter in everyday life, such as education, cooking, shopping, scheduling, and gift giving.

The system continually provides feedback to the learner about different aspects of his or her interaction with ProTutor. Each activity that the learner completes provides him or her with

feedback about how s/he performed on game-play aspects of the activity. The learner also receives immediate and delayed feedback about his or her pronunciation. The immediate feedback is provided through his or her computer's speakers as the learner records his or her voice while s/he is performing activities. The delayed feedback is provided through the open portions of the learner model where the learner can see what has been pronounced properly and improperly. The use of Open Learner Models to provide feedback should encourage learner reflection and motivation by allowing students to see a representation of their pronunciation skills. Furthermore, the design of ProTutor's open models provides an important utility that should be included in all multimedia CALL systems: it helps learners identify their own errors (Chapelle, 1998).

Once a learner has identified his or her errors and has chosen to work on one of them, ProTutor provides additional feedback. We wanted to ensure that this feedback, which is most likely given when learners are experiencing difficulties, is both constructive and helpful. To do this we have the system describe the desired behaviour and actions rather than have the system describe the learner's error.

3.3. EXPECTED COMMON USE CASES

The following use cases illustrate how a learner might interact with the system in order to acquire increased proficiency in his or her spoken language skills.

3.3.1. Use Case 1: Practice Speaking

User: Learner

Goal: To practise his or her pronunciation of Russian words and phrases.

Scenario: If the learner is entering for the first time, s/he fills out a survey about his or her language proficiencies (languages known and fluency levels for reading, writing, and speaking), and experience using the language. The system then recommends activities to the learner.

If the learner has previously entered the system and completed the survey, then s/he is immediately taken to the list of recommended activities. The learner selects an activity and then

performs the assigned tasks. The system then analyzes the learner's performance, stores the results of this analysis in the learner model, and recommends activities and instruction that it thinks will benefit the learner.

3.3.2. Use Case 2: Self-monitoring (Viewing the OLM and HOLM)

User: Learner

Goal: To motivate the learner, encourage learner reflection, and inform learners of their progress.

Scenario: Either the learner chooses to view a visualization of his or her learner model (LM) or the system displays the LM's visualization to the learner because the availability of the model needs to be brought to the learner's attention.

The LM visualization selects a sentence to display to the learner. The sentence that is selected for display in the learner model is designed to highlight the learner's strengths and or weaknesses, and its phone transcription represents how that learner would pronounce the sentence if the learner were reading the sentence aloud (see the Representative Sentence subsection of the Open Learner Model section on page 48 for details of how this is done). The learner inspects his or her representative sentence's phone transcription alongside the expert speaker's phone transcription. S/he compares the two in order to see how they differ. If the learner has access to the HOLM s/he will also compare his or her current phone transcription to his or her past phone transcription to look for changes.

The system summarizes the learner's strengths and weaknesses and makes recommendations to the learner based on this information. When the learner is having a bad day s/he may choose to work on a sound that s/he rarely mispronounces, whereas on a good day, the learner may choose to try to master a phone that s/he is having difficulty pronouncing. The learner views instructional material about that phone and performs many exercises to master the phone. S/he then reviews the LM visualization (OLM or HOLM) for changes in his or her performance.

3.4. SYSTEM ARCHITECTURE

ProTutor consists of several components that must work together in order to help the learner properly pronounce words when speaking in a foreign language. There are six basic components, some of which have subcomponents:

- Speech Recognition Engine
- Diagnosis Engine
- Pedagogical Delivery Subsystem
    - Activity Delivery Engine
    - Feedback Engine
    - Pedagogical Planning Engine
- Constraint Based Language Model
- Learner Model
- Libraries
    - Pedagogical Activities
    - Language Resources

All of the above components interact in order to analyze the learner's speech input so that mispronunciations can be identified. Once this is done, the system helps the learner improve by creating and delivering an instructional plan that allows the learner to practise his or her pronunciation of sounds. The manner in which these components interact is described below and it can be seen in Figure 3.1.

3.4.1. Speech Recognition Engine

The Speech Recognition Engine (SRE) processes the learner's speech so that the learner's utterance can be identified. This identification allows ProTutor to determine which sounds that a learner is having trouble pronouncing. The determination of pronunciation errors is done in the Diagnosis Engine (DE), after the SRE has processed the signal, identified the utterance, and passed it to the DE.

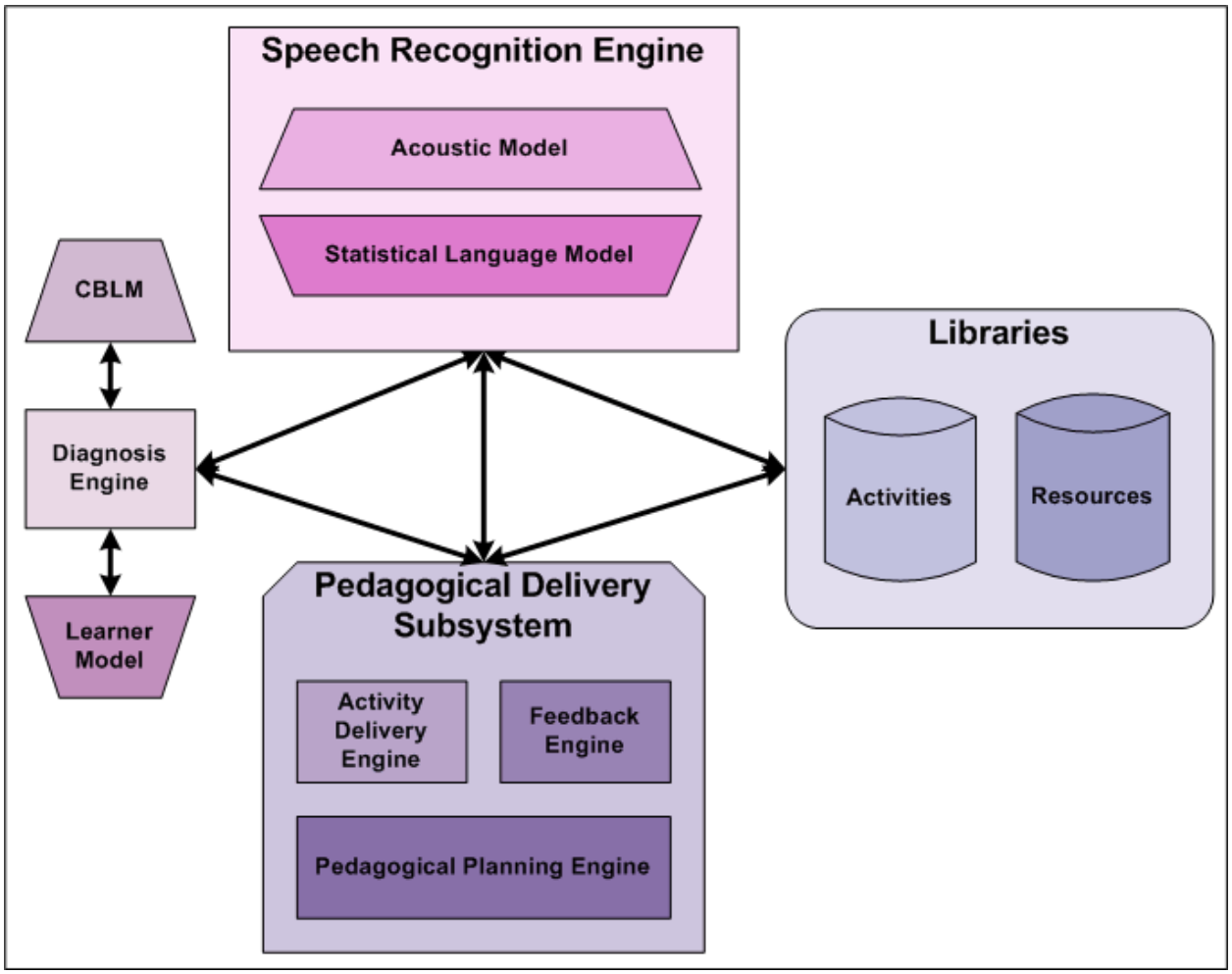

Figure 3.1: ProTutor's high-level architecture

3.4.2. Diagnosis Engine

The DE takes the results from the SRE and compares them to the expected phones for words that were recorded during the performance of the pedagogical activity that was assigned by the Pedagogical Planning Engine (PPE). The DE uses a Constraint Based Language Model (CBLM) to compare the learner's results to the results that were expected for that word (Barrow, Mitrovic, Ohlsson, & Grimley, 2008). Based on that comparison the DE identifies which phones were correctly and incorrectly pronounced. The number of incorrect and correct pronunciations of each phone that is part of the word is then updated in that learner's Learner

Model. In order to see what errors are being made, each mispronunciation is also recorded in the raw diagnosis results portion of the learner model (See Table 3.2).

Table 3.2: Raw diagnosis results: A subset of what is stored in the database for each phone that the learner has recorded. In this case the learner has pronounced 2 phones correctly and 1 phone incorrectly.

AnalysisID	AttemptID	PhoneID	Recognized PhoneID	Correct	Learner
1137378	1307	17	17	1	Learner0
1137379	1308	11	4	0	Learner0
1137390	1309	1	1	1	Learner0

3.4.3. Pedagogical Delivery Subsystem

The Pedagogical Delivery Subsystem (PDS) takes the pedagogical plan and delivers it to the learner. This subsystem interacts with the learner graphically and textually. It reacts to the learner's input and delivers pedagogical activities as needed through the Activity Delivery Engine. The Feedback Engine (FE) generates motivational feedback for the learner based on game-play aspects of the pedagogical activities. The FE later gives the learner feedback about his or her pronunciation performance through a visualization of the underlying learner model.

3.4.4. Pedagogical Planning Engine

The Pedagogical Planning Engine (PPE) is a subcomponent of the PDS and it has two main components: a prescribed path through the available activities and personalized additions to that path. The prescribed path through the activities is the same for each learner. However, each learner also receives several sets of personalized modifications to their pedagogical plan, in the form of added personalized activities. To provide these modifications, the PPE interacts with the Libraries and the Learner Model in order to plan targeted activities for the learner. The targeted activities consist of recommended instructional materials and personalized activities based on the results of the diagnosis that are stored in the learner's model. Learners are encouraged to further practise their most frequently mispronounced phones and they are not given the option of practising the phones for which they have a moderate frequency of mispronunciation.

3.4.5. Learner Model

The Learner Model must track the learner's errors and the frequency with which each phone is mispronounced. The Learner Model also stores the learner's mother tongues and a list of all of the other languages that the learner speaks along with his or her level of proficiency in each language. Learner honesty and accuracy about his or her skills in each language is required since this information is obtained from the learner via a survey when learners first log on to ProTutor. If learners can be trusted to be honest about their abilities and experiences, then this information can later be used to increase the system's adaptivity.

We can only measure the learner's proficiency in Russian once s/he has been using ProTutor for some time and must, to an extent, trust the learner's self-assessment until we have sufficient evidence to the contrary. Unfortunately, we cannot ensure that learner self-assessments are completely accurate without the use of language proficiency exams. Even though they may be inaccurate, we could use the information that learners enter to fine tune which activities a learner is initially recommended and to provide more personalized instructional materials based on the learner's other language experience.

In addition to tracking the learner's pronunciation accuracy, the Learner Model indirectly tracks how long it took the learner to fix his or her errors by tracking changes in the learner's pronunciation accuracy. The Learner Model also tracks which pedagogical activities were performed by the learner and when s/he performed each activity while trying to improve his or her pronunciation. This can later be used to make better activity recommendations based on how the learner's performance changes following the completion of different activities.

3.4.6. Constraint-based Language Model

The Constraint-based Language Model (CBLM) contains information about how the target second language is properly pronounced. Since learners receive little to no phonetic instruction in early language courses, we hybridized the CBLM so that it contains a list of exceptions to the model and a simplified list of the language's pronunciation rules. ProTutor will use this set of rules as constraints against which a learner's pronunciation will be compared during diagnosis and violations of these constraints will be used to identify pronunciation errors, rather than what

is causing the error (Barrow, Mitrovic, Ohlsson, & Grimley, 2008). The simplified list of constraints and list of exceptions that are found in the CBLM are also used to create the expert model's pronunciation so that the behavior that ProTutor models is consistent with the other feedback that will result from use of the CBLM during diagnosis and later pedagogical planning.

3.4.7. Libraries

The system has two types of library. The first library, Language Resources, contains different cultural, language, literary, and artistic media. The Language Resources library holds tongue twisters, compositions (stories, poetry, sentences, and excerpts from novels), and dictionaries (visual and textual). The second library, Pedagogical Activities, contains lessons of varying sizes to help the learner achieve accurate pronunciation of the target language. It holds pronunciation lessons, drills, and phone explanations that can help the learner obtain accurate pronunciation in a foreign language.

3.5. THE PATH OF AN UTTERANCE THROUGH PROTUTOR

When a learner is performing activities s/he records many utterances. These utterances are captured and processed by ProTutor in order to build the learner model and provide the learner with materials to help him or her improve.

Once the learner has indicated that s/he would like to record an utterance, the application server, protutor, connects the learner to the streaming server, protutorstreamer, which captures his or her utterance by streaming and saving the learner's audio input. This file is then converted to the appropriate format for the speech recognition engine, Sphinx-3 (Carnegie Mellon University, 2007). The Sphinx-3 speech decoder processes the recorded utterance and outputs the word that it recognizes to a text file. Protutor processes the speech recognition output files and enters the results of the speech recognition into the database so that the diagnosis process can begin.

The diagnosis engine then uses the CBLM to identify which phones the learner has pronounced correctly and incorrectly. For diagnosis purposes, the expected phones are derived using information in the CBLM in combination with the knowledge of which word the learner

was supposed to be saying. This expected or expert pronunciation model is then compared against the results of the speech recognition. An example of this is provided in Table 3.3, where ProTutor has given the learner the Russian word for actor, а к т ё р.

Table 3.3: Creating the phone mapping of a word using the CBLM and then diagnosing the learner's utterance by comparing the phone mapping to the learner's utterance.

Phone Mapping Steps		**Expected Word: актёр**			**Stress: Character 4 (ё)**	
Break word up into its parts		а	к	т	ё	р
Check to see if it's an exception		It's not an exception.				
Apply constraints						
	Stress contraints	Unstressed а remains а. Stressed ё remains ё.				
	Intermediate phone mapping	а	к	т	ё	р
	Consonant agreement between words constraints	Nothing happens here since this is only 1 word				
	Intermediate phone mapping	а	к	т	ё	р
	Consonant agreement within word constraints	None of the consonants are voiced (the unvoiced т is preceded by another unvoiced consonant к) so nothing needs to be done here.				
	Intermediate phone mapping	а	к	т	ё	р
	Consonant softening constraints	т precedes a stressed soft vowel. So it is softened to т'				
	Intermediate phone mapping	а	к	т'	ё	р
Final Phone Mapping Result		**а**	**к**	**т'**	**ё**	**р**
Student Utterance		а	к	т	ё	р
Diagnosis Results		**Correct**	**Correct**	**Incorrect**	**Correct**	**Correct**

The diagnosis engine will first break the word actor up into individual characters. It then applies the pronunciation constraints from the CBLM. It starts by applying the constraints for stressed vowels; in English, this is similar to indicating which vowels are emphasized and if the vowel is long or short. It then applies the constraints that affect the pronunciation of consonants to ensure that voiced consonants become unvoiced and unvoiced consonants become voiced, when appropriate. The final set of constraints that is applied, to find the word's phone mapping, is the consonant softening constraints; these ensure that consonants that precede a softening character are modified appropriately. The result of applying all of these constraints is the final or expert phone mapping. The final phone mapping is then compared, phone by phone, to what the learner said. If the phones match then the learner has pronounced that phone correctly. Otherwise, the learner has made an error. In either case the result is recorded.

The diagnosis engine updates the learner model, by changing the accuracy of the learner's current pronunciation for each phone that was in the expected utterance (Figure 3.2); it also stores which phones were pronounced for each of the expected phones (Table 3.2). This provides the learner model with information about which pronunciation errors are made, the situation in which they are made, and when the learner correctly pronounces phones.

Learner	phoneID	accuracy	numGood	numBad	Current Attempt Diagnosis
Learner0	14	0.499998815325892 ⇩ 0.499997630657398	211028	211029 ⇨ 211030	incorrect
Learner0	1	0.500000244170789 ⇩ 0.500000488341339	1023874 ⇨ 1023875	1023873	correct
Learner0	10	0.500000496365118 ⇩ 0.500000992729251	503662 ⇨ 503663	503661	correct
Learner0	19	0.500000500605232 ⇩ 0.500000000000000	499396	499395 ⇨ 499396	incorrect

Figure 3.2: How the model is updated following diagnosis. The arrows point from the old value to the new value based on the results of the current attempt's diagnosis.

3.6. ACTIVITY RECOMMENDATIONS

ProTutor provides each learner with a series of activity recommendations whenever they log on to the system. In order to give learners a sense of accomplishment, ProTutor also shows learners the recommended activities that they have most recently completed beside the list of recommended activities that learners get when they log on (Figure 3.3).

At the beginning, each learner is given the same learning path and set of recommended activities; this is called the Planned Common Learning Path. Once a learner has completed one of his or her recommended activities, it is moved to the recently completed activities list and a new recommended activity appears. The learner can choose to perform the displayed

recommended activities in any order, which gives learners a sense of control and choice while still requiring that each activity be performed at some point.

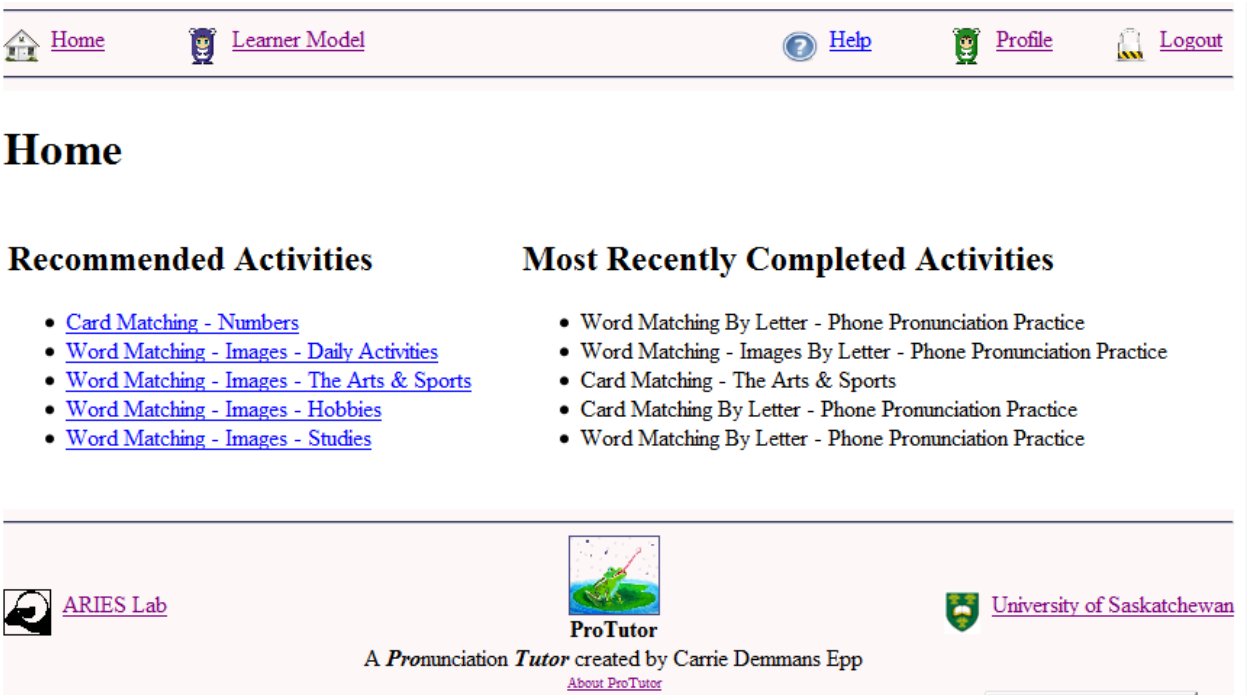

Figure 3.3: A screenshot of the homepage, which is where activities are recommended to the learner and where the learner can see which activities s/he has recently completed.

Once a learner has completed enough activities to populate the pronunciation performance portion of his or her learner model s/he receives personalized activity recommendations through his or her learner model. The learner will still follow the Planned Common Learning Path. However, every time that the learner views the learner model s/he will be given access to his or her Personalized Recommendations. These personalized activity recommendations act much like extra lessons or homework do in a regular classroom. The learner can perform them in order to improve a specific skill, in this case his or her pronunciation of a phone, before returning to performing activities from the class' regular curriculum.

3.6.1. Planned Common Learning Path

The Planned Common Learning Path is the set of activity recommendations that are sequenced based on the sequencing of the chapters and themes found in the course textbook,

Troika (Nummikoski, 1996). There are three basic activities that have variants and each of those is initially recommended once for each theme.

In order to encourage the review and retention of newly learned vocabulary, activities for each theme are later re-recommended. The random selection of the vocabulary used in each activity helps ensure that the exact same set of words is not given to the learner multiple times so that learners get a different activity each time.

The basic activity types are card matching, word matching, and flash cards. Their descriptions and the descriptions of their variants can be found in the ACTIVITIES sub-section later in this chapter. The activities that make-up the Planned Common Learning path are recommended in the order presented in Table 3.4.

Table 3.4: The Sequencing of Activities within the Planned Common Learning Path. The activities are recommended in the order listed from top to bottom, left to right.

ThemeName	Card Matching	Word Match (with Images)	Word Match	Flash Cards
Introductions	✓	✓	✓	✓
Language Skills	✓	✓	✓	✓
Possessions	✓	✓	✓	✓
Navigation	✓	✓	✓	✓
Locations	✓	✓	✓	✓
Vocations	✓	✓	✓	✓
Daily Activities	✓	✓	✓	✓
The Arts & Sports	✓	✓	✓	✓
Hobbies	✓	✓	✓	✓
Studies	✓	✓	✓	✓
Shopping	✓	✓	✓	✓
Introductions	✓			
Language Skills	✓			
Possessions	✓			
Navigation	✓			
Locations	✓			
Vocations	✓			
Daily Activities	✓			
Numbers	✓			
Introductions		✓		
Language Skills		✓		
Possessions		✓		
Navigation		✓		
Locations		✓		

ThemeName	Card Matching	Word Match (with Images)	Word Match	Flash Cards
Vocations		✓		
Daily Activities		✓		
Numbers		✓		
Making Plans	✓	✓	✓	✓
Gifts	✓	✓	✓	✓
Weather	✓	✓	✓	✓
Travel	✓	✓	✓	✓
Eating	✓	✓	✓	✓
Introductions	✓			
Language Skills	✓			
Possessions	✓			
Navigation	✓			
Locations	✓			
Vocations	✓			
Daily Activities	✓			
Numbers	✓			
Introductions				✓
Language Skills				✓
Possessions				✓
Navigation				✓
Locations				✓
Vocations				✓
Daily Activities				✓
Numbers				✓
Future Plans	✓	✓	✓	✓
Feelings	✓	✓	✓	✓
The Arts & Sports	✓	✓	✓	✓
Hobbies	✓			
Studies	✓			
Shopping	✓			
Hobbies		✓		
Studies		✓		
Shopping		✓		
Hobbies			✓	
Studies			✓	
Shopping			✓	
Hobbies				✓
Studies				✓
Shopping				✓
Numbers	✓	✓	✓	✓

3.6.2. Personalized Recommendations

From the information presented in the model, the learner can choose to work on any of the selected phones for which they receive feedback. This means that learners can try to improve phones that they are already good at or that they are bad at. This gives the learner additional control in the hope that it will help keep them motivated and address some of their emotional needs (Beck J. , 2007) (Picard, 2007).

For each phone that the learner chooses to work on, the learner receives instructional material that tells him or her how to properly pronounce that sound or when a character should be pronounced as a certain phone so that s/he can focus on what s/he should do while performing the personalized activities that are recommended alongside the instructional material (Figure 3.4). The personalized activities randomly select words based on the phone that is being trained (see Random Word Selection on page 41 for details). This means that the words for each of the recommended activities could come from any theme. However, each of the words will be guaranteed to have at least one instance of the desired phone.

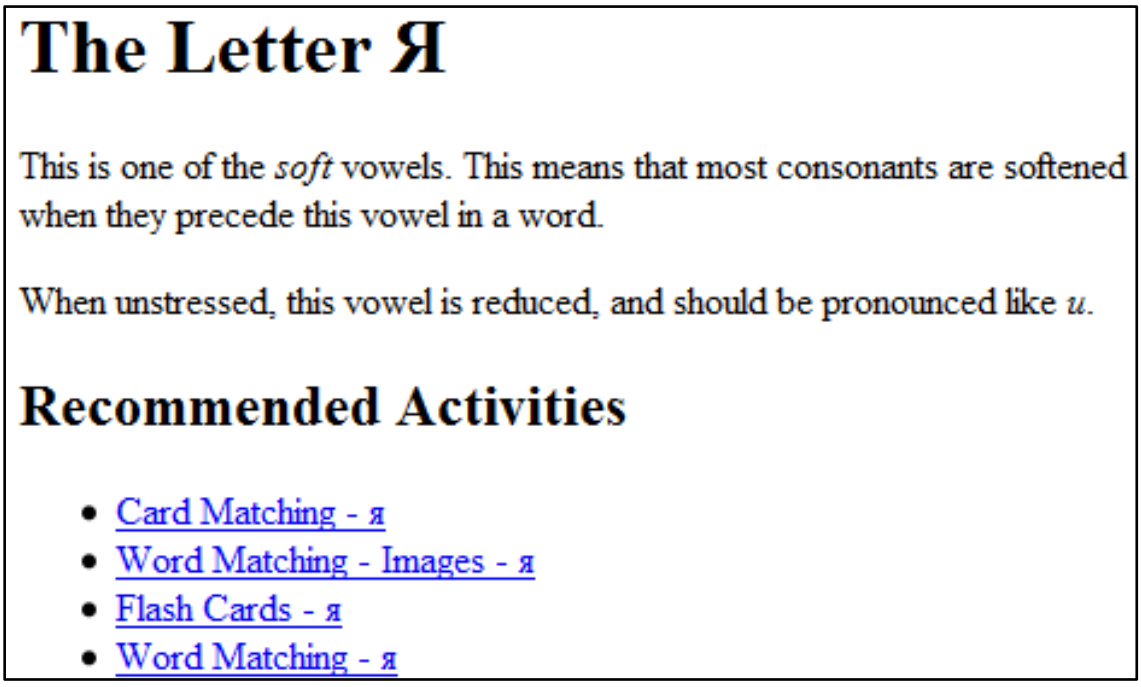

Figure 3.4: A screenshot of the instructional materials and recommended activities for a phone, я, that the learner chose to work on.

3.7. ACTIVITIES

ProTutor has 3 basic activities and variants on those activities that it recommends to learners. The basic activity types are the memory card game, word match, and flash cards. Each of these activities and its variants are recommended to the learner for each theme within the course textbook, Troika. These recommendations appear when the learner first logs on to the system and each time that s/he returns to ProTutor's homepage (Figure 3.3). For each activity, the words that it trains are randomly selected (see Random Word Selection on page 41).

In each activity the learner must record himself or herself saying the word (Figure 3.5). Learners can record themselves saying a word as many times as they would like, but cannot delete previous recordings (Figure 3.6). In order to complete an activity the learner must record all of the words that are in the activity at least once. At which point, the learner is given feedback about the game-play aspect of the activity.

All of the activities are accompanied by instructions and a phone chart. The instructions ensure that learners know how to proceed. The phone chart is available to learners who need additional support or who may have forgotten how an infrequently used character is pronounced.

Figure 3.5: ProTutor's Recording Screen

Figure 3.6: ProTutor's Re-recording Screen

3.7.1. Memory Card Game

The memory card game only has one version, which is based on the real-world memory game that is played using a deck of cards (Concentration (game), 2010). In contrast to the game played with a deck of cards (Figure 3.7), ProTutor populates the cards that are used with images of the randomly selected vocabulary based on the current theme or phone (see Random Word Selection on page 41). The learner must then flip up to two cards, one at a time, until s/he finds two cards that have the same image (Figure 3.8).

If the two cards that are flipped do not match (Figure 3.9), ProTutor flips them back over so that the vocabulary image is no longer visible (Figure 3.10). If the two cards match (Figure 3.11), the learner records himself or herself saying the Russian word that is represented by the image on the cards.

For the memory card game, learners are told how many cards they flipped over while trying to find all of the pairs (Figure 3.12). They can compare this against the number of cards that were in the game to see how well they did in the game or they can keep track of this to see if they are improving at the game.

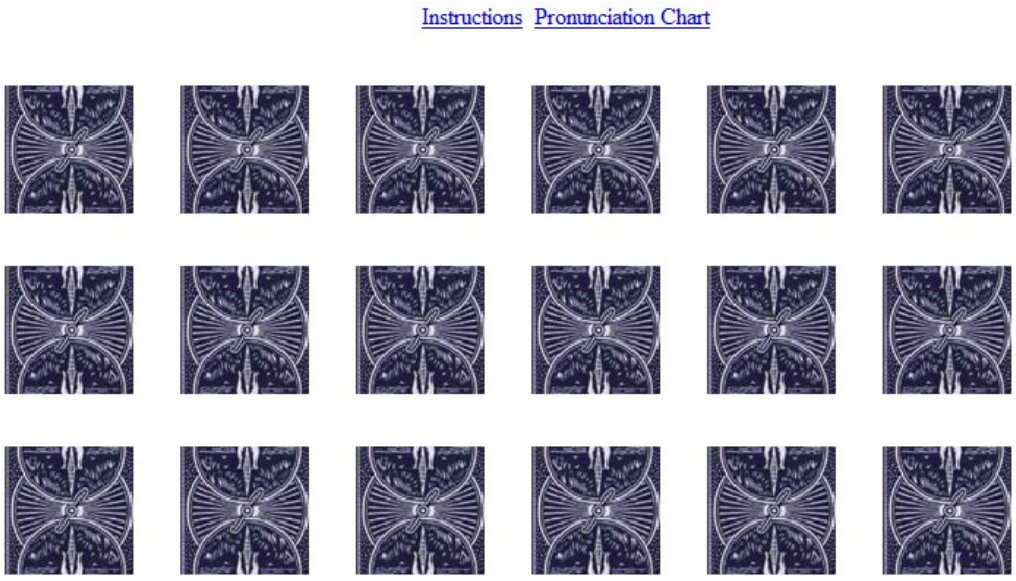

Figure 3.7: Memory Card Game before game-play has begun.

Figure 3.8: Memory Card Game with one card flipped.

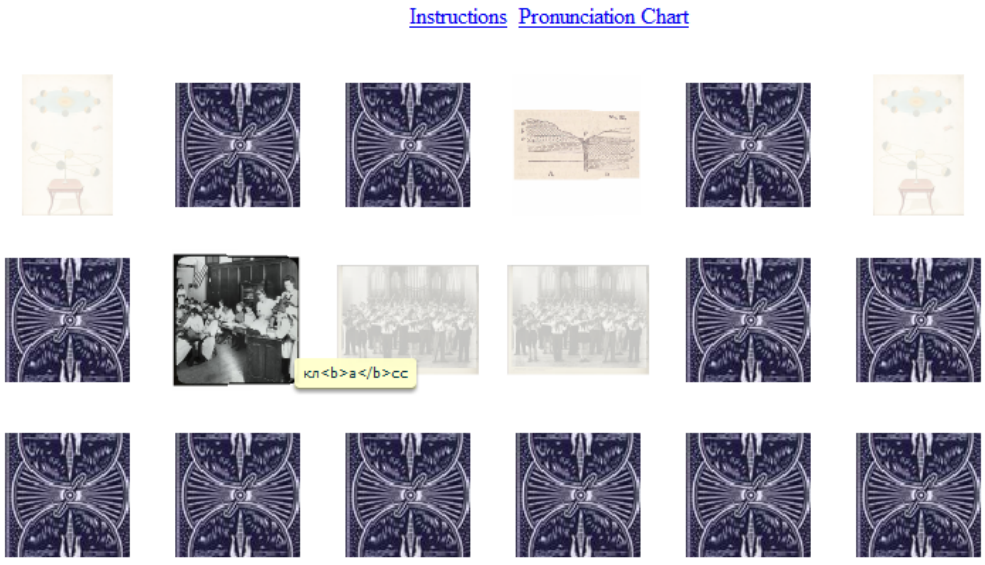

Figure 3.9: Memory Card Game when the two flipped cards do not match.

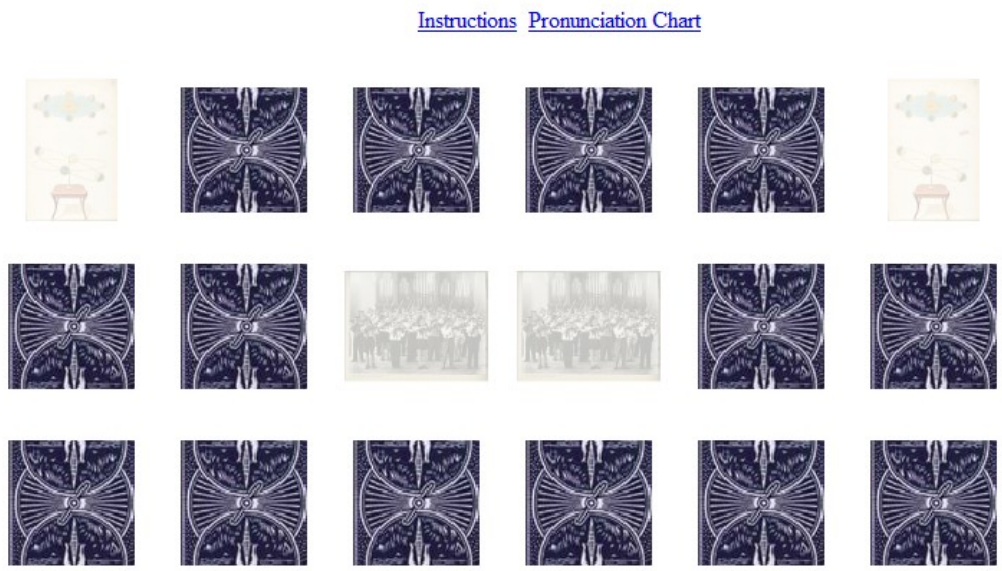

Figure 3.10: Memory Card Game after mismatched cards are flipped back to their initial, face-down, position.

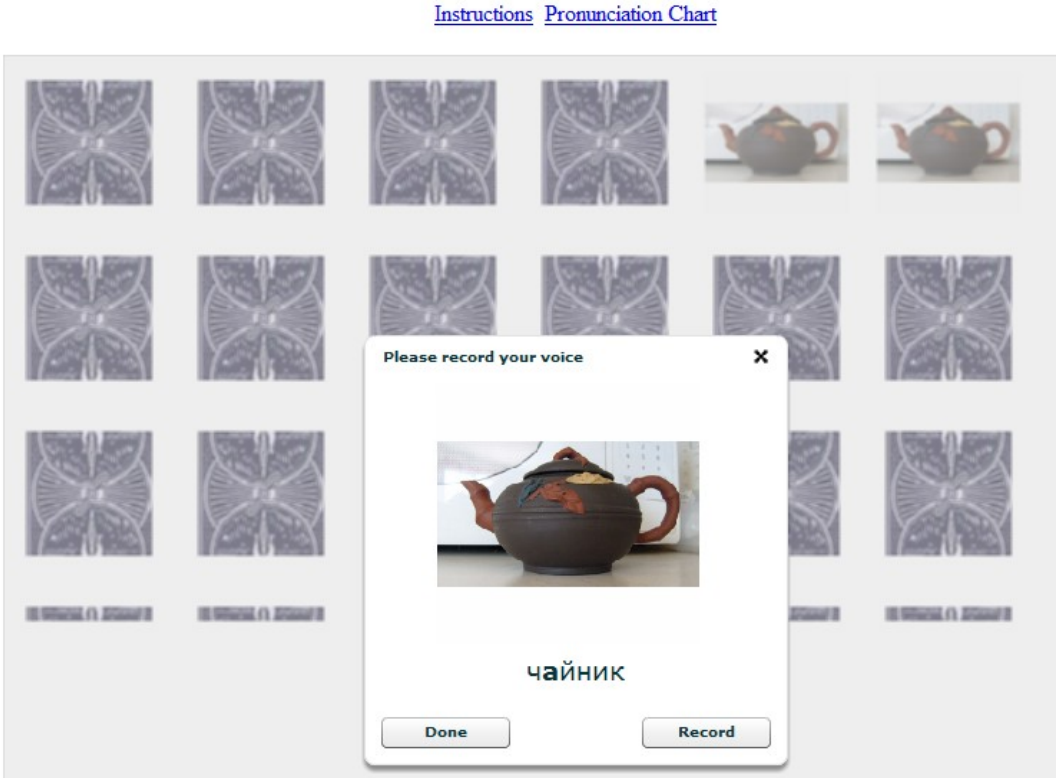

Figure 3.11: Memory Card Game when the two flipped cards match.

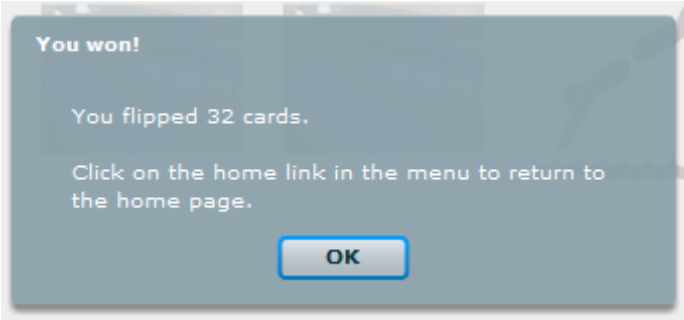

Figure 3.12: Game play feedback for the Memory Card Game. 8 extra cards were flipped.

3.7.2. Word Match

The word match activity has two versions: one where the learner matches the Russian word to its representative image (Figure 3.13) and another where s/he matches the Russian word to its English equivalent (Figure 3.14). For this activity, the learner drags the Russian word from the list of words on the left to its matching card on the right. If the word matches the card that the learner has dragged it to (Figure 3.15), then the learner gets to record himself or herself saying the word. If the learner incorrectly matches a word to a card, then the word automatically returns to the list of words that is found on the left.

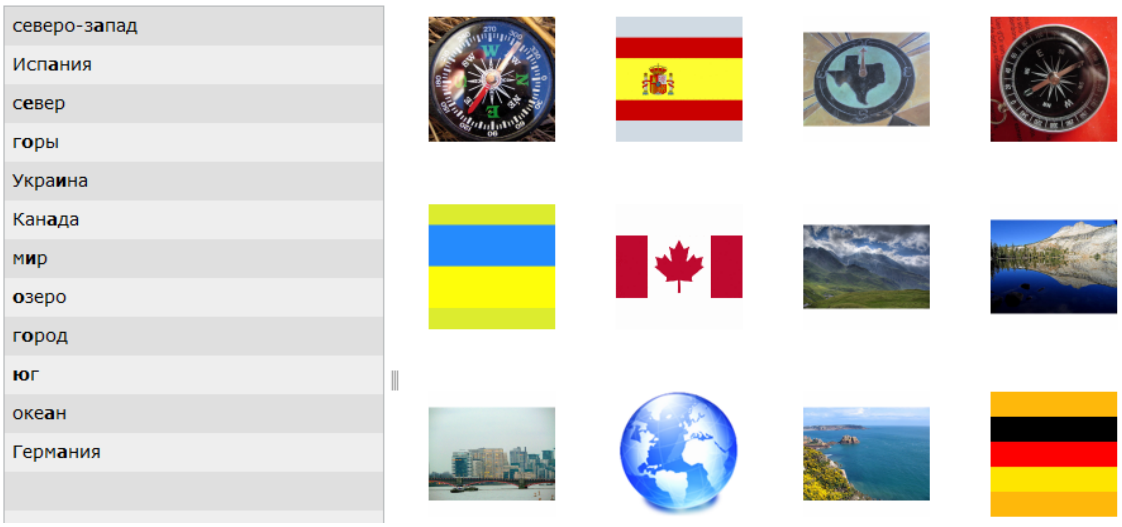

Figure 3.13: Initial word Match with Images screen, before game-play has started.

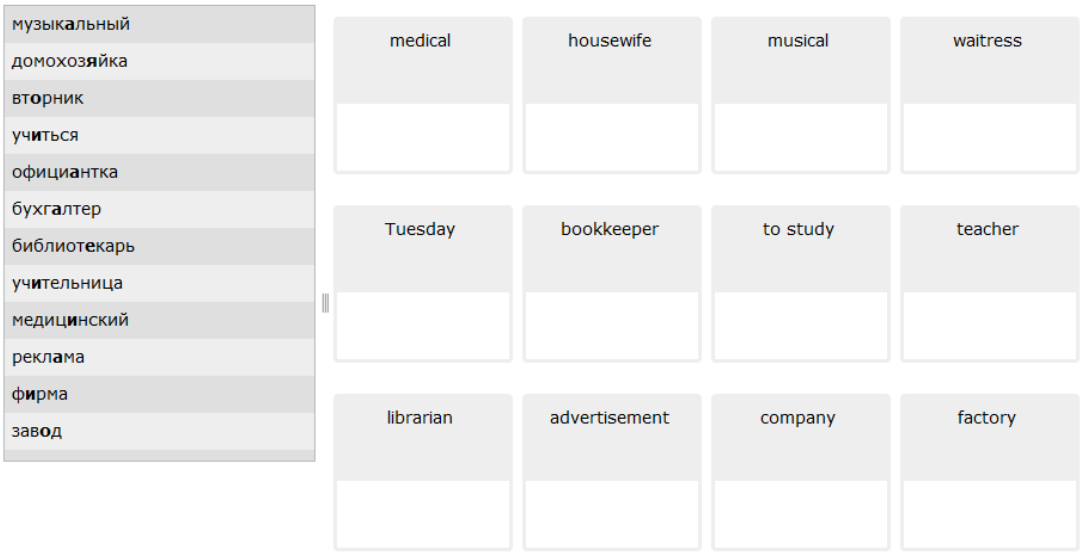

Figure 3.14: Initial word Match without images screen.

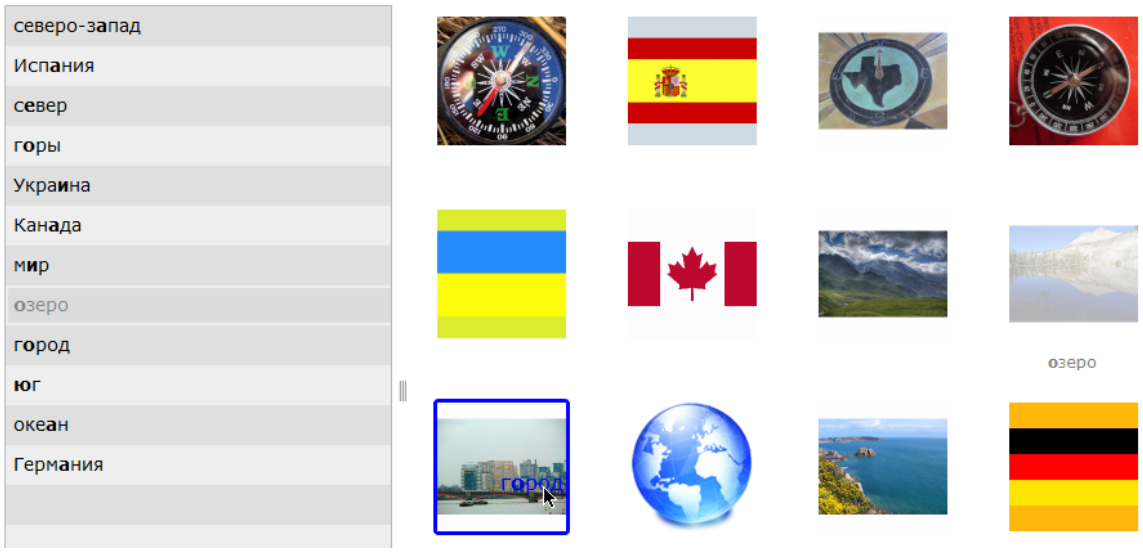

Figure 3.15: Word Match when a match is made, just before the recording screen pops up. The image of the lake is greyed out because it was previously matched and the word recorded.

Once the learner has matched and recorded all of the words, the activity is considered complete and s/he receives feedback about the number of matches that s/he attempted to make (Figure 3.16).

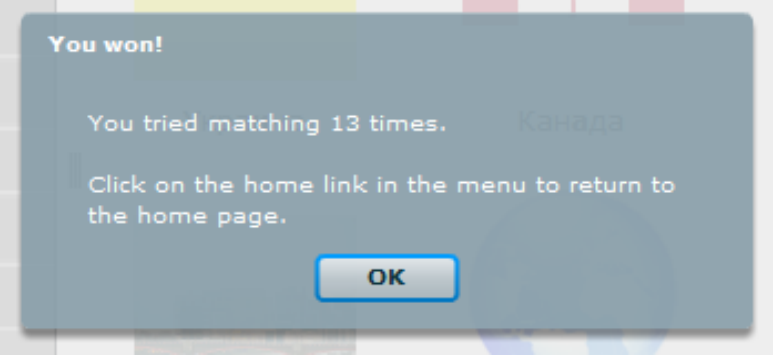

Figure 3.16: Word Match game-play feedback. Since there were 12 words, one mismatch was made.

3.7.3. Flash Cards

The flash cards activity was inspired by the use of real flash cards, where the word is in English on one side of the card and in Russian on the other side of the card (Figure 3.17). For this activity there is also a variant, where the Russian word is on one side of the card and an image that represents that word is on the other side of the card. However, we only used the image-free version of the game during our study. By default, the cards are traversed in order, but a learner can skip a card by clicking on the next card in the sequence (Figure 3.18).

Learners have the option of peeking at the word by flipping a card over to see the Russian word before they record the word, or they can record the word without peeking (Figure 3.19). Once the learner has completed the activity by recording all of the words, s/he is shown how many times s/he flipped a card to peek at the Russian word before s/he successfully completed the activity (Figure 3.20). This gives the learner a sense of how s/he is doing in his or her vocabulary acquisition and of how s/he might do on an upcoming vocabulary test.

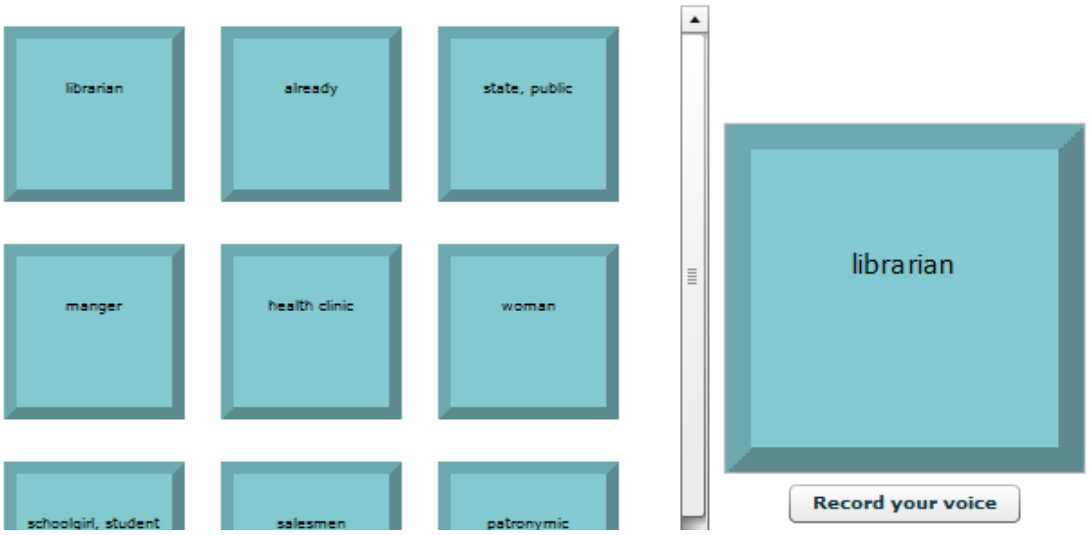

Figure 3.17: Flash Card initial screen before anything has been recorded.

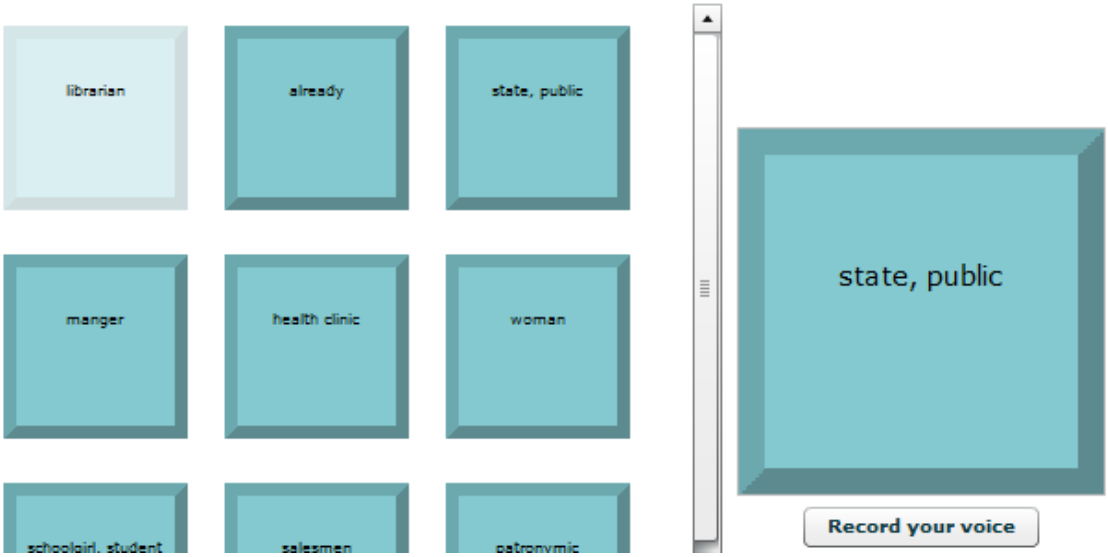

Figure 3.18: Flash Card with the first card recorded, and the second card skipped.

Figure 3.19: The Flash Card is flipped to peek at the Russian word.

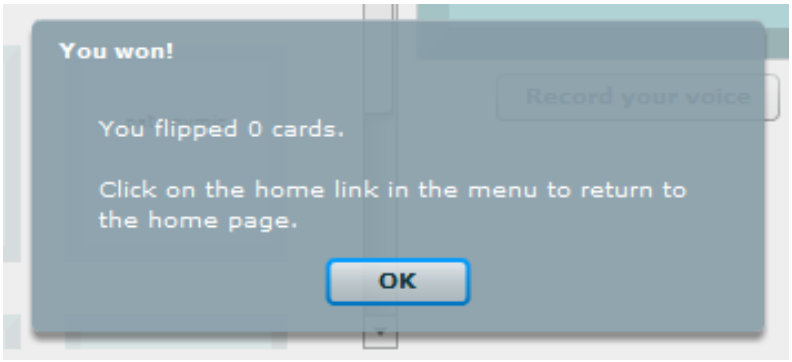

Figure 3.20: Flash Card game-play feedback. This time the learner did not peek at any of the Russian words.

3.7.4. Random Word Selection

Words are randomly selected for each of the activities in order to ensure that the learner gets a variety of experience. Words can be selected randomly based on two criteria:

- theme
- phone

All of the words associated with a theme can be selected for activities that do not require images. When an activity needs images, the set of possible words is reduced to those that have images. Activities that require that words are selected for personalized training of a specific phone do not take the theme into account. In this case, only the phone and the necessity of images for the activity are considered important. So, the words are randomly selected based on the presence of the phone and whether or not the word has an associated image. When words are being chosen based on a phone rather than a theme, the desired phone can be anywhere in the word since the main focus is to provide the learner with additional opportunities to practice saying that phone.

In all of the stored procedures that select the vocabulary to be used for activities, the query results are first ordered by the result of the built-in Transact-SQL function NEWID() (Microsoft Corporation, 2010). This ensures that the same set of words is not selected twice in a row as is demonstrated in Table 3.5 and Table 3.6. However, it is possible for the same word to appear in subsequent executions of the stored procedures that are used to randomly select vocabulary. Even though words can reappear in subsequent calls to the stored procedure, it would be unlikely to get the exact same set of words returned in subsequent executions of the stored procedure.

Table 3.5: Results of selecting 10 words based on theme when images are not needed for the activity.

RusWord	EngWord	ThemeID	ImgLoc	VocabID	Stress
горячий	hot (temperature of food)	16	NULL	1991	4
булочка	bun	16	NULL	1903	2
мёд	honey	16	NULL	1922	2
квас	Russian homemade beer	16	NULL	1911	3
закуска	appetizer, snack	16	NULL	1954	4
заняты	occupied, busy (pl)	16	NULL	1998	2
закрыт	closed	16	NULL	1994	5
стакан	glass (cup)	16	NULL	1965	5
вилка	fork	16	Dictionary/VocabImages/Troika16/fork.png	1951	2
нами	us (instr)	16	NULL	2015	2

Table 3.6: Results of re-selecting 10 words based on theme when images are not needed for the activity

RusWord	EngWord	ThemeID	ImgLoc	VocabID	Stress
встреча	meeting	16	NULL	1976	5
голоден	hungry	16	NULL	1990	2
заказать	to order (pf)	16	NULL	2027	6
вами	you (instr - pl)	16	NULL	2016	2
двойной	double	16	NULL	1993	6
блины	Russian pancakes	16	NULL	1901	5
заказывать	to order	16	NULL	2026	4
стакан	glass (cup)	16	NULL	1965	5
дуршлаг	colander	16	Dictionary/VocabImages/Troika16/Colander_3.png	2199	6
заняты	occupied, busy (pl)	16	NULL	1998	2

3.8. LEARNER MODEL

ProTutor provides feedback to the learner through the use of two different visualizations of the learner model. One is an Open Learner Model (OLM) and the other is a Historic Open Learner Model (HOLM). In order to inform the learner, both of the visualizations show him or her a representation of his or her current abilities and highlight the learner's strengths and weaknesses. However, the HOLM adds historic information about learner performance in order to maintain learner motivation.

Learners only have access to one visualization of the underlying model at a time. They are given access to their OLM until their HOLM is ready, at which point they are only given access to their HOLM. In both cases, the learner's entire model is not opened up to him or her. To keep the feedback tractable, we chose to limit how much of the learner's underlying model is shown.

3.8.1. Learner Model (Data)

The underlying learner model holds data about the learner's perceived proficiency in each of the languages that s/he speaks, amount of experience with those languages, his or her mother tongue, and information about his or her pronunciation proficiency in Russian. We collect information about the learner's proficiency and experience with each of the languages that the learner speaks from him or her when s/he first logs in. We do this through an initial survey where the learner must enter the requested information (see LEARNER INITIALIZATION SURVEY in APPENDIX C).

The majority of the information that is stored about a learner's Russian pronunciation proficiency is derived from the learner's interactions with ProTutor. It is the result of the learner's performance and it shows how well s/he pronounces each phone that s/he should have said as well as how often s/he was supposed to have said that phone (see Table 3.7 for sample data).

Table 3.7: A sample of a learner's phone pronunciation proficiency.

Learner	phoneID	accuracy	numGood	numBad
Learner3	3	0.500067540186411	4512	4511
Learner3	4	0.500202675314147	1504	1503
Learner3	5	0.500081050413357	3760	3759

The sum of the numGood column and the numBad column is the total number of times that the learner was expected to attempt pronouncing that phone, and the accuracy column is the percentage of times that the learner pronounced the phone correctly. The learner's current phone pronunciation accuracy is cumulative over all of the utterances that s/he has performed. Once a week, the historic portion of the learner model is created by taking and storing the data that is in the current pronunciation accuracy portion of the leaner model. This creates a weekly snapshot of the data contained in the pronunciation proficiency portion of the learner model (Table 3.8). It also means that the historic portion of the model's data for the learner's phone pronunciation accuracy is cumulative over all of the utterances that s/he has performed up to the date on which the snapshot of his or her abilities was taken.

Table 3.8: Sample data from the historic learner phone pronunciation table.

Learner	date	phoneID	accuracy	numGood	numBad
Learner3	2009-11-05	29	0.499976697581209	10728	10729
Learner3	2009-11-12	29	0.50001256944619	19890	19889
Learner3	2010-02-11	4	0.500334448160535	748	747
Learner3	2010-02-18	4	0.500260552371027	960	959
Learner3	2009-11-19	10	0.499994260855592	43560	43561
Learner3	2009-11-26	10	0.500004656837635	53685	53684

Not only does the learner model contain aggregations of the learner's pronunciation proficiency, it also stores a record of the learner's raw pronunciation results. This keeps a record of each phone that was pronounced by the learner when s/he pronounced a phone correctly and when s/he pronounced a phone incorrectly (Table 3.2). This could later be used for the recommendation of more detailed and specialized instructional materials and for refining the selection of words for the personalized activity recommendations.

The final part of the pronunciation accuracy portion of the learner model, is the aggregation of the entire group's pronunciation results into a group model (Table 3.9). This aggregation is cumulative over the entire set of utterances for all users. The group's pronunciation accuracy is calculated by averaging the accuracy of a phone for all learners.

Table 3.9: Sample data from the group model

phoneID	accuracy	phone
1	0.477779949617044	a
3	0.428597000675664	в
4	0.500075147115495	г

In addition to tracking the learner's pronunciation proficiency, the learner model tracks information about the pedagogical activities that each learner attempts (Table 3.10) and each activity that s/he completes (Table 3.11). The learner model records when each activity recommendation is made, when the learner follows the recommendation, how many times the learner attempts the activity before completing it, and when each of these actions is performed. If we look at the first row of Table 3.11, we can see that the learner completed the activity on his or her eighth attempt. This means that s/he failed to record all of the words that were given to

him or her in the activity the first seven times through it. If we look at the second row, we can see that the learner completed the activity on his or her first attempt, and the table's third row shows that the learner has yet to complete the activity.

Table 3.10: Sample activity recommendation data from the learner model

initRecTime	Learner	numRecs	dateFollowed	followed	ARPairID	completed
2009-11-12 00:00:00.000	Learner1	1	2009-11-12 17:35:17.310	1	82	1
2009-10-04 00:00:00.000	Learner1	1	2009-12-03 19:51:28.583	1	63	1
2009-12-03 00:00:00.000	Learner1	1	2009-12-03 19:47:37.997	1	46	0

Table 3.11: Sample completed activity data from the learner model.

firstAttempt	mostRecentAttempt	numAttempts	completed	ARPairID	Learner	ActivityID
2009-10-04 13:44:25.167	2009-12-03 19:38:22.350	8	1	2	Learner1	255
2009-11-12 17:32:54.617	2009-11-12 17:32:54.617	1	1	81	Learner1	425
2009-12-10 11:23:45.230	2010-01-24 16:48:57.647	2	0	82	Learner1	455

3.8.2. Open Learner Model

The OLM consists of a limited visualization of the data contained in the learner model; it has four parts (Figure 3.21):

- a representative sentence
- a list of phones
- a group model of phone pronunciation
- a list of recently completed activities

These four parts are meant to motivate the learner and to inform him or her so that the learner can reflect on his or her performance. Each part's role in this is described in further detail below.

Carrie's Learner Model

Здравствуйте Меня зовут Алексей. Это моя жена Наталья, а это её мама и бабушка.
Здрафствуйти Миня завут Аликс'ей. Эта мая жина Натали а ета её мама и бабушка.

Здравствуйте Меня зовут Алексей. Это моя жена Наталья, а это её мама и бабушка.
Здрафствуйти Миня завут Аликс'ей. Эта мая жина Натали а эта её мама и бабушка.

Individual Phones (Sounds)

Your best 3 Phones:	The 3 Phones You Should Work on More:
ё е ж	ю э г
The class is best at these phones: **е ш м** The class most needs to work on these phones: **э б ц**	**Legend:** Your classmates are good at this phone Your classmates need to work on this phone

A ' that follows a letter indicates that the preceding letter should be softened. If you click on one of the linked phones. You will receive some information about that phone as well as links to activities, that will help you to improve your pronunciation of that phone.

Most Recently Completed Activities

- Word Matching - Images - Introductions
- Card Matching - Language Skills
- Card Matching - Possessions
- Card Matching - Introductions

Figure 3.21: Screenshot of a ProTutor OLM

3.8.2.1. Representative Sentence

The representative sentence has three parts: the original sentence, the learner's phone transcription, and the expert's (native Russian speaker's) phone transcription (Figure 3.22). The three parts of the representative sentence allow the learner to see both his or her model and the ideal model in order to compare them and understand which phones s/he is pronouncing well and what s/he can improve upon. The recasting of the learner's pronunciation that ProTutor's OLM facilitates by providing the expert's phone transcription will benefit learners more as their pronunciation proficiency increases and it provides an ideal model for them to work towards (Archibald & O'Grady, 2008).

The representative sentence is selected from a collection of sentences based on the learner's best and worst pronounced phones. The stored procedure that selects the sentence tries to find a sentence that contains all of the learner's current best and worst phones. If that sentence does not exist, then a sentence that contains a subset of those phones is selected. Two sets of potential representative sentences for a learner are extracted from the set of all possible sentences. The first set, is the set of sentences that contain the learner's good phones. The second set, is the set of sentences that contain the learner's bad phones. If an intersection exists between these two sets, then one of the sentences that is in the intersection is chosen as the representative sentence for that learner. If no sentences are found in the intersection of the good and bad sets, then one of the sentences from the good set is shown to the learner. This prevents learner discouragement by giving them positive reinforcement rather than only showing them their mistakes.

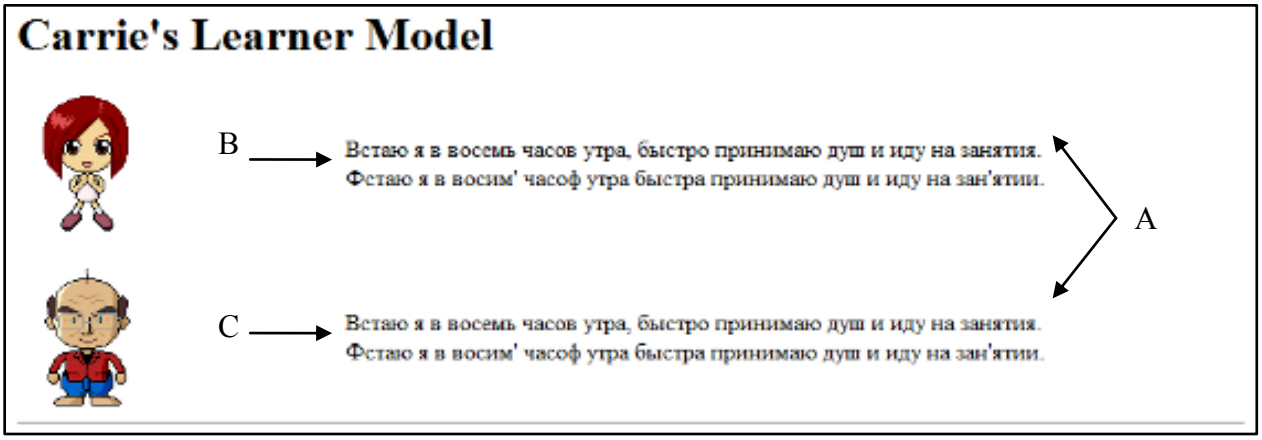

Figure 3.22: The representative sentence portion of the OLM. **A** points to the original sentence, **B** points to the learner's phone transcription, and **C** points to the native Russian speaker's transcription.

The expert's phone transcription of the representative sentence is stored in the database along with the sentence (Table 3.12). The learner's phone transcription is dynamically generated based on his or her three most commonly mispronounced phones. The learner's phone transcription is built by taking the expert's phone transcription and replacing the phones that the learner mispronounces with the most common error type for second language learners when pronouncing that phone. For example, if the representative sentence selected for the learner has sentenceID 1 and the learner's worst phones are *o*, *p*, and *г*, then his or her phone mapping might look like "Здрафствуйти Миня зовут Аликс'ей. Это моя жина Натали а это её мама и бабушка" instead of " Здрафствуйти Миня завут Аликс'ей. Эта мая жина Натали а эта её мама и бабушка ". The learner can then compare the two phone transcriptions to look for differences, which should inform him or her of the differences between his or her performance and the expected ideal performance.

Table 3.12: Two possible representative sentences and their expert phone mapping (sentenceMapping).

sentenceID	sentence	sentenceMapping
1	Здравствуйте. Меня зовут Алексей. Это моя жена Наталья, а это её мама и бабушка.	Здрафствуйти. Миня завут Аликс'ей. Эта мая жина Натали а эта её мама и бабушка.
2	Меня зовут Володя. А это наша семья: мой отец Сергей, моя мама Татьяна и мой дедушка Иван.	Мин'я завут Валод'и. А эта наша симья мой ат'ец Сиргей, мая мама Тат'яна и мой д'едушка Иван.

To prevent confusion over which transcription is the learner's, each transcription is placed beside an avatar. The expert's avatar is always that of a teacher (Figure 3.23), but learners can select or create an avatar to represent themselves. This selection of an avatar by the learner is done when s/he first logs into ProTutor. The learner can then change his or her avatar at any time by visiting his or her profile page.

Figure 3.23: ProTutor's teacher avatar.

3.8.2.2. Phone List

In the OLM the list of phones consists of two groups of three phones (Figure 3.24). This means that the learner is receiving feedback on six phones. The first group of phones is the set of phones that the learner most often pronounces correctly (best) and the second group is the set of phones that the learner most often mispronounces (worst) (Figure 3.24). The groups of best and worst pronounced phones are based on all learner performance up to the current time. Learners are presented with both their best and worst phones because as previous work has suggested learning and motivation can be best maintained by balancing positive and negative feedback (Barrow, Mitrovic, Ohlsson, & Grimley, 2008).

3.8.2.3. Group Model of Phone Pronunciation

Individual learner models are closed to the group, but a collective group learner model is open to all learners. The group model, which is an aggregation of all of the individual learners' models, will show how the other learners are progressing through the various phones as a group. More specifically, the group model shows the group's three best and worst pronounced phones.

In Figure 3.24, what the group is good at is indicated textually and by using emoticons. Regardless of which of the learner's phone lists that a phone from the group's open model appears in, emoticons appear beside the phones displayed in the learner's open model when the learner has a phone in common with the group. The textual version of the group open model is displayed beside the legend that explains the emoticons, so that the learner can see the open portion of the group model even if s/he has nothing visibly in common with the group.

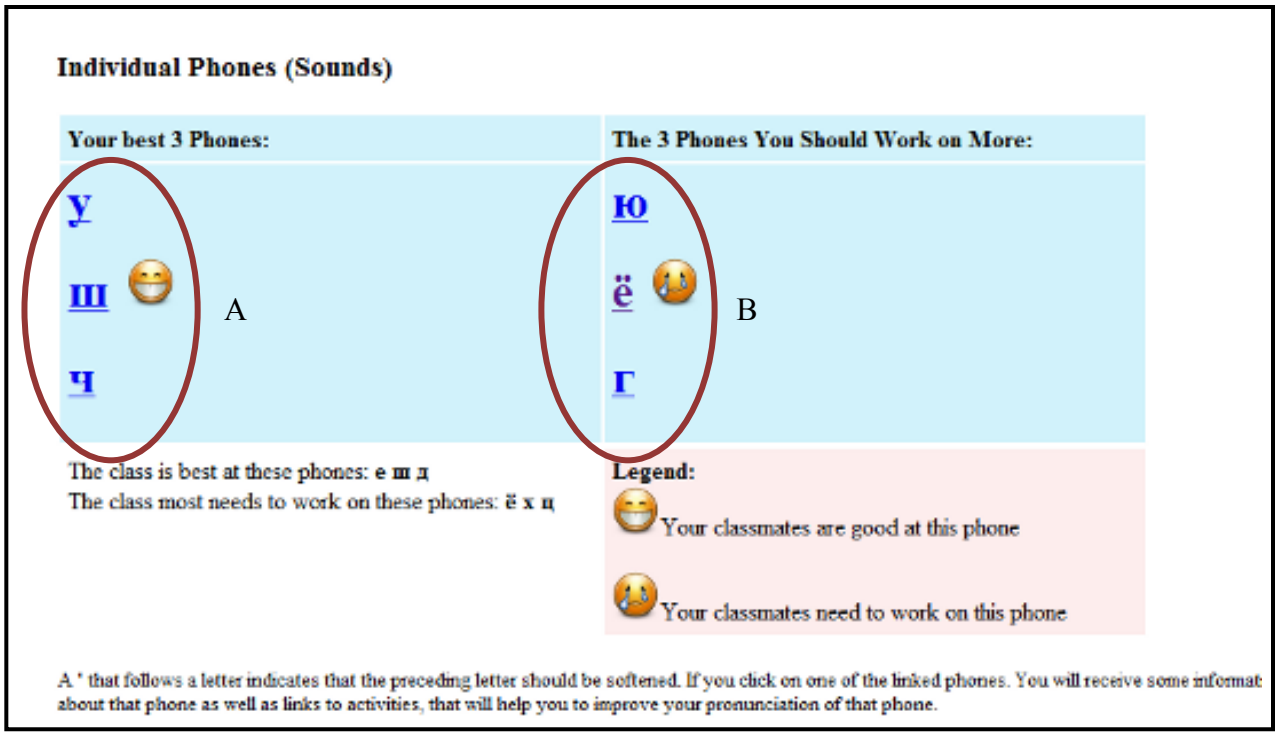

Figure 3.24: The best and worst phones section of the OLM. The phones circled beside the A are the learner's best and those circled beside the B are the learner's worst. This learner has one good phone in common with the rest of the class and one bad phone in common with the rest of the class.

The use of an open group model protects the individual learner's privacy while allowing learners to see how they are performing relative to the rest of their class. This should help maintain learner motivation by allowing learners to see that everyone still needs to improve their pronunciation. The open group model can also help motivate competitive learners by showing them that the class is better than they are at particular phones (Bull & Kay, 2007).

3.8.2.4. Recently Completed Activities

To encourage learner reflection on the causes of changes in their model we included a list of recently completed activities at the bottom of the OLM (Figure 3.25). This information when combined with reflection, also allows learners to make decisions about what has positively and negatively affected their performance so that learners can make better choices about how to proceed with their training. The list of recently completed activities can also help maintain learner motivation by showing them what they have done, since it allows learners to see their five most recently completed activities. This can give learners a sense of accomplishment even if they have yet to see an improvement in their model.

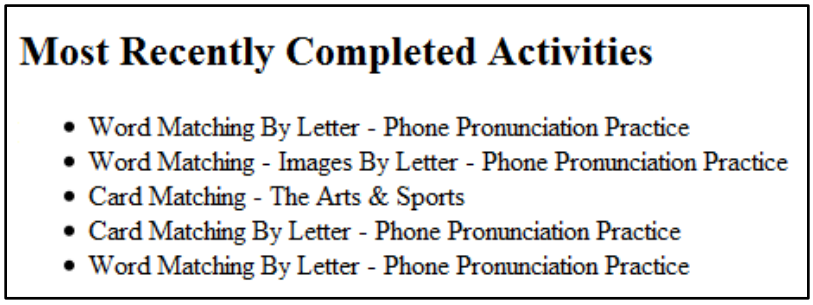

Figure 3.25: The portion of the OLM that shows the learner his or her list of most recently completed activities.

3.8.3. Historic Open Learner Model

Much like the OLM, the HOLM consists of its underlying data and the visualization of parts of that data. The HOLM has the same basic parts as the OLM with some modifications and additions. These modifications and additions result in the following list of components:

- a representative sentence
    - current
    - historic
- a list of phones
    - current
    - historic

- a group model of phone pronunciation
- a pronunciation visualization
- a list of recently completed activities

The modifications and additions that provide historical information are provided to help motivate learners and encourage reflection, both of which can help them acquire new knowledge and improve their current skills (Vosniadou, 2001). These modifications and additions should also better facilitate the learner's reflection of what caused changes in his or her performance so that s/he can make better decisions about his or her learning path.

3.8.3.1. Representative Sentence

In the HOLM, the learner's phone transcription of the representative sentence is modified in two ways. The first is exactly the same as in the OLM, by using the learner's current set of three worst and best pronounced phones to create his or her current phone transcription. The second modification is a historic phone transcription, which uses the learner's previous three worst and best phones to create his or her phone transcription (Figure 3.26). By showing these two transcriptions one above the other, we hope to highlight how the learner's model has changed and allow him or her to reflect on his or her performance and experiences. S/he can still compare his or her model to that of an expert native speaker, but since the expert's, native Russian speaker, model is always perfect it shows only one phone transcription and appears exactly as it does in the OLM.

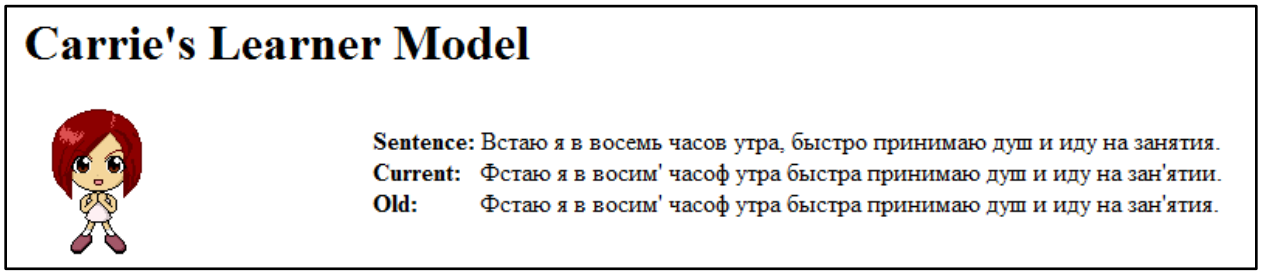

Figure 3.26: The learner's Representative Sentence as it is shown in the HOLM. This example shows that the learner's pronunciation of his or her current representative sentence is the same as it would have been previously.

3.8.3.2. Phone List

In order to highlight changes in the learner's performance, the HOLM modifies the OLM's phone list by adding two more phone groups: the learner's previous three best and worst pronounced phones based on all of the learner's sessions leading up to the date selected for the historic view (Figure 3.27). The date selected for the historic information is the oldest of the historic models that are stored in the database. The addition of these two historic phone groups means that the learner could receive feedback on up to twelve phones in the HOLM, rather than the six phones for which they receive feedback in the OLM. This is because the learner's current and previous best and worst phone lists could be exactly the same, entirely different, or somewhere in between.

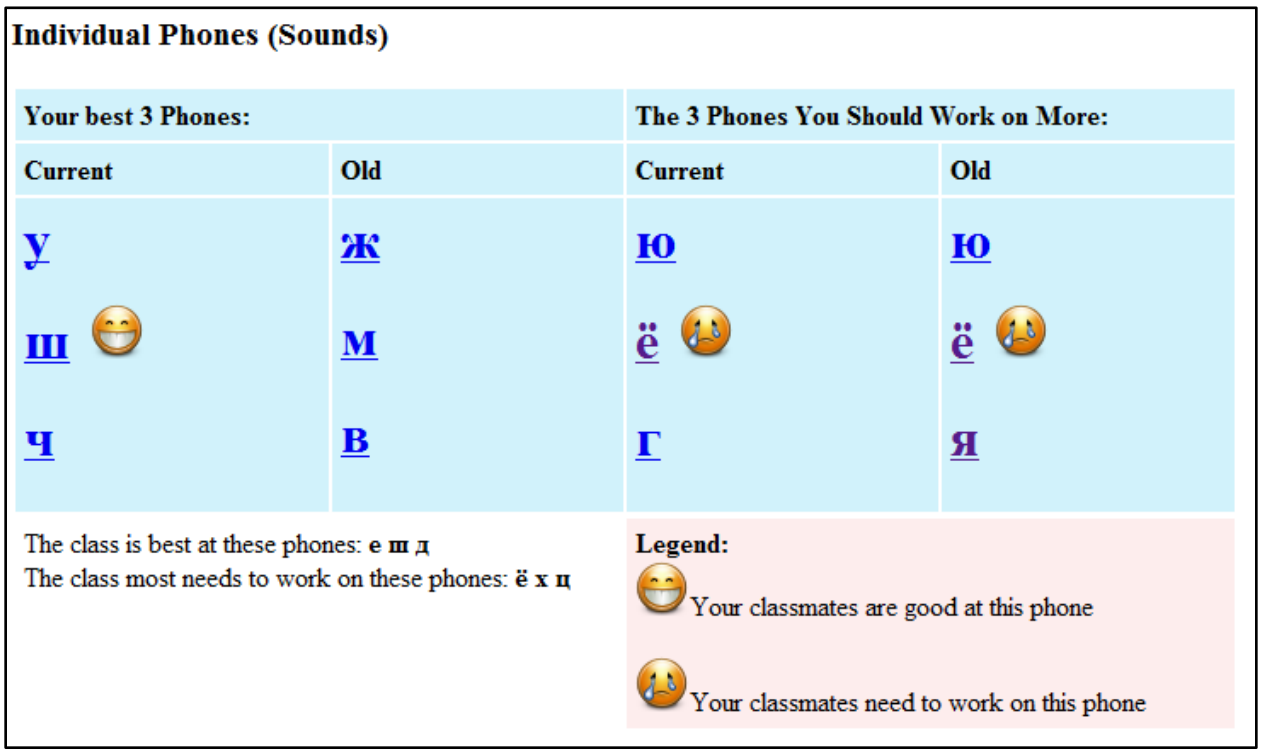

Figure 3.27: The HOLM's best and worst phone lists. This learner has improved his or her pronunciation of at least 4 phones.

3.8.3.3. Group Model of Phone Pronunciation

The HOLM also opens up the group model. However, it only shows the group's current best and worse phones. This still informs the learner about how s/he is doing compared to the larger group and allows him or her to reflect and make choices based on this information. However, it strategically limits the amount of information that the learner has access to so that his or her attention is focused on how s/he is doing, since it is less important for learners to see how the group's performance is changing than it is for them to see how their performance has changed.

3.8.3.4. Pronunciation Visualization

The HOLM displays a line-chart that shows how the accuracy of the learner's pronunciation of their fourth through eighth most accurately pronounced phones has changed (Figure 3.28). This is meant to help encourage learners to keep working whenever a phone drops off of their best pronounced list. The addition of the line-chart was needed to help clarify that the removal of a phone from a learner's best list did not necessarily mean that s/he had gotten worse at pronouncing that phone. Rather, it could mean that s/he had improved another phone to a point where s/he pronounced it better than the phone that was previously one of his or her most accurately pronounced. It is also important to make it clear when the phone is removed from the best phone list because the learner's pronunciation of that phone has worsened, even if this information might be demotivating to the learner. In this case, the learner needs to know that the change in his or her focus has harmed his or her performance so that s/he can make appropriate decisions about how to proceed in the future.

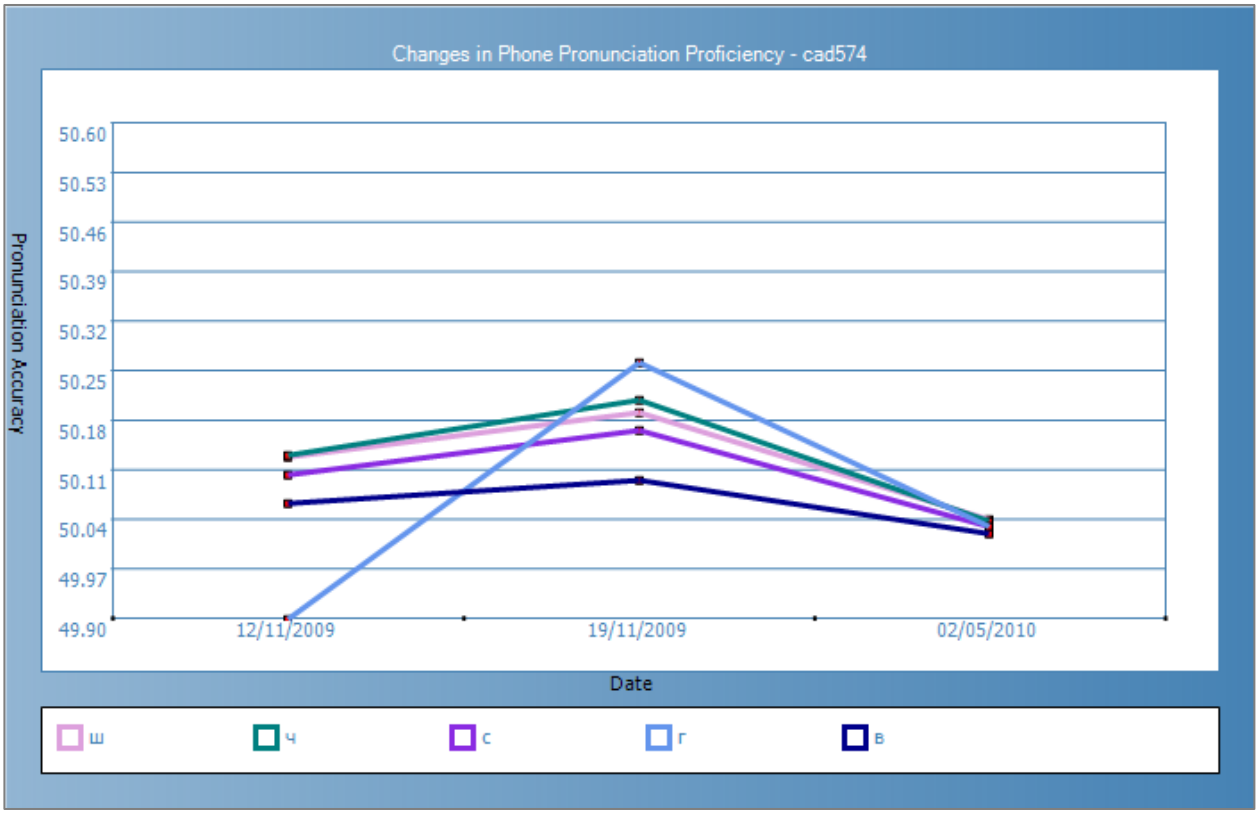

Figure 3.28: The HOLM pronunciation visualization.

3.9. PROTUTOR SYSTEM SUMMARY

The addition of a historic component, expert model, and three best and worst phones to the learner model will provide enough summative and detailed information to help encourage learner reflection, focus the learner's attention on important features of his or her performance, and maintain his or her motivation to keep working. The game-play and varied nature of ProTutor's activities should also help encourage learner motivation throughout the training process. In addition to this, the procedural instruction that accompanies personalized activities should give learners the targeted help and practice that they need to improve their pronunciation skills.

CHAPTER 4

SYSTEM EVALUATION

We performed a formative evaluation of ProTutor and its learner models using students from beginner and intermediate Russian second language classes at the U of S. The goal of this evaluation was to determine if learners found ProTutor and its learner models helpful, motivating, and useful. This involved having students use ProTutor for extended periods of time. Five students completed the study (see APPENDIX B - STUDY APPROVAL), which consisted of three phases.

Phase one involved performing activities in ProTutor while receiving no feedback about pronunciation quality. This helped us to overcome the cold-start problem, by giving us the opportunity to collect data so that a learner model could be built. During this phase of the experiment, each learner followed a prescribed set of activities. The learner could choose to perform the recommended activities in any order by skipping a number of activities, but the activities that they chose to skip remained in the recommended activities list until they were completed. This eventually forced learners to do each of the recommended activities. During this phase of the study, the skipping of activities was the only control that the learner could exercise. Skipping activities did not affect the learner's ability to transition into the next phase of the study, provided that the learner had completed enough activities to populate his or her learner model.

In the second phase, learners were given access to their OLM along with personalized activity recommendations that might be helpful to them. Learners could choose to perform any (or even none) of these recommended activities. Learners could access their personalized activities by clicking on one of the phones in their best or worst list. This would redirect them to instructional material about that phone and activities that were meant to train that phone. The personalized activities that were not completed would later re-appear in their main list of

recommended activities; learners could also choose to skip the personalized recommendations in order to perform activities that were recommended from the common learning path. All other areas of the system behaved the same as in phase 1.

In the third phase, learners were given access to their HOLM. All other aspects of the system, with respect to learner control, remained the same. However, because of the nature of the HOLM, learners had more personalized activity recommendations to choose from since the HOLM's list of best and worst phones could contain up to 12 phones, whereas the OLM's could only contain 6 phones.

Following each phase of the study, learners completed a survey. These surveys evaluated participant attitudes towards ProTutor and the learner models that it employed. Some of the survey's questions were open ended, but the majority of them were statements that were rated on a seven point Likert-scale, where 1 was agree (usually meaning a positive comment about ProTutor) and 7 was disagree (usually implying a negative comment). Screenshots of each of the surveys can be found in SURVEYS USED FOR THE STUDYwithin APPENDIX C.

Even though all of the participants spoke at most one additional language to their mother tongue, I still consider them to be highly motivated language learners because they all attended class and performed at least one extracurricular Russian language learning activity other than using ProTutor. Furthermore, three of those students were doing at least three other extracurricular Russian language learning activities.

While the sample size was too small to ensure statistical significance, it did show some positive changes in learner perception of ProTutor and its learner models that are supported by learner actions. In order to clarify the discussion, I have subdivided the analysis of the results based on learner perceptions of the system as a whole, their perceptions of areas of the system, and their behaviour within the system. We are interested in the learner's overall impressions of ProTutor and in the learner's perception of the following areas of ProTutor: changes in pronunciation quality, parts of the learner models, and activity recommendations.

4.1. LEARNER IMPRESSIONS OF PROTUTOR'S EFFECT ON PRONUNCIATION

Following phase 3, the participants believed that ProTutor helped them to identify when they were mispronouncing something, which was a considerable improvement in their perception over their phase 1 responses (Table 4.1). They also felt that the HOLM version of ProTutor did a better job of helping them to identify what they were pronouncing correctly (Table 4.2). In contrast to their responses while using the HOLM, participants did not feel that the OLM helped identify their correct phones or their incorrect phones.

Table 4.1: Participant response to "ProTutor helped me to identify when I was mispronouncing something".

	Study Phase		
	1	2	3
Mean	6.3	5.8	3.4
Standard Deviation	0.5	1.6	1.8

Table 4.2: Participant response to "ProTutor helped me to identify when I was pronouncing something correctly".

	Study Phase		
	1	2	3
Mean	6.3	5.6	3.6
Standard Deviation	0.5	1.5	1.8

The results of the survey indicated that participants found all of the learning activities helpful in improving their Russian when they were used in combination with the HOLM, whereas only two of the activities were considered to be helpful when used in combination with the OLM. Learners felt that ProTutor helped them work on improving their weaknesses (Table 4.3) and tried to help them improve their Russian pronunciation (Table 4.4). This was an improvement over phase 1 and phase 2, where learners did not feel that ProTutor helped them work on improving their weaknesses or their pronunciation.

Table 4.3: Participant response to "ProTutor helped me to work on my weaknesses".

	Study Phase		
	1	2	3
Mean	5.0	5.2	3.6
Standard Deviation	1.9	1.1	1.9

Table 4.4: Participant response to "ProTutor tried to help me improve my Russian pronunciation".

	Study Phase		
	1	2	3
Mean	4.7	4.2	2.8
Standard Deviation	2.0	0.8	1.3

Even though we could not accurately measure changes in learner pronunciation due to limitations in the accuracy of our speech recognition software, participants felt that their pronunciation of individual phones (Table 4.5) and that of Russian words (Table 4.6) improved while using the HOLM version of ProTutor. Interestingly, when asked to again rate their Russian oral proficiency, at the end of the study, only two learners chose a higher language proficiency category, from the available categories of poor, mediocre, good, excellent, and fluent, than they had when they initially rated their oral abilities (Table 4.7). Even though the learners felt that their ability to pronounce parts of Russian words and the words themselves had improved, they did not perceive an improvement in their overall pronunciation ability. This may be due to how each of the learners perceives the proficiency categories or because of an increase in awareness of their abilities. It may also mean that learners improved the accuracy of their self-perception through the reflection that was encouraged by the OLM and HOLM.

Table 4.5: Participant response to "I feel that my pronunciation of individual Russian sounds has improved".

	Study Phase		
	1	2	3
Mean	4.5	3.6	2.6
Standard Deviation	1.6	1.1	1.3

Table 4.6: Participant response to "I feel that my pronunciation of Russian words has improved".

	Study Phase		
	1	2	3
Mean	4.3	3.8	2.4
Standard Deviation	1.8	0.8	1.1

Table 4.7: Participant self-ratings of their Russian pronunciation proficiency.

	Oral Proficiency	
Learner	Start	End
1	mediocre	good
2	good	good
3	poor	mediocre
4	poor	poor
5	good	mediocre

4.2. LEARNER IMPRESSIONS OF THE LEARNER MODELS' PARTS

Regardless of which learner model was used, participants liked to see how they pronounced the different phones (Table 4.8), which phones they pronounced well (Table 4.9), and the phones for which they need to improve their pronunciation (Table 4.10).

Table 4.8: Participant response to "I liked being able to see how I pronounced different sounds when I looked at the learner model".

	Study Phase	
	2	3
Mean	2.6	1.8
Standard Deviation	1.9	0.8

Table 4.9: Participant response to "I liked how the learner model showed which sounds I was good at pronouncing".

	Study Phase	
	2	3
Mean	2.0	1.6
Standard Deviation	1.2	0.5

Participants also demonstrated some preference, when using the HOLM, for being able to see how the rest of the class was doing compared to themselves (Table 4.11), and they liked being able to see the expert model of a native Russian speaker pronouncing the representative sentence that was part of their learner model (Table 4.12). Participants also liked to compare their phone transcription of the representative sentence to the phone transcription of the expert (Table 4.13).

Table 4.10: Participant response to "I liked how the learner model showed which sounds I should work on improving my pronunciation of".

	Study Phase	
	2	3
Mean	1.4	1.8
Standard Deviation	0.5	0.8

Table 4.11: Participant response to "I liked being able to see how the rest of the class was doing in comparison to me".

	Study Phase	
	2	3
Mean	4.0	3.2
Standard Deviation	2.0	1.1

Table 4.12: Participant response to "I liked how the learner model showed how a native Russian speaker would say the same sentence".

	Study Phase	
	2	3
Mean	2.4	2.4
Standard Deviation	0.9	1.3

Table 4.13: Participant response to "I liked to compare my model to the native speaker's model".

	Study Phase	
	2	3
Mean	2.6	2.2
Standard Deviation	1.1	1.1

Regardless of the model used, participants felt that the learner model accurately reflected their pronunciation of Russian words (Table 4.14). One participant explained her comfort with the learner model because she "felt that the learner model accurately reflected [her] ability to pronounce Russian words". Four of the five learners also stated that being able to see their previous performance was motivating. One of those learners commented that "seeing that my weaker sounds were no longer my weakest made me feel that I was improving". Interestingly, when a learner felt that his model was inaccurate he ignored it.

Table 4.14: Participant response to "I felt that the learner model accurately reflected my ability to pronounce Russian words".

	Study Phase	
	2	3
Mean	2.2	2.6
Standard Deviation	0.8	1.3

There did not seem to be a difference in preference for having access to the group aspect of the model based on the model type. Four of the five learners liked to see how their classmates were doing in comparison with them. One of those four wanted to be able to hear her classmates and the fifth participant did not like the group model because she likes "to have an individual approach" and because s/he dislikes it when people are compared. It seems that adding a partial group model made it difficult for that participant to see that the learner model and recommendations were personalized.

4.3. LEARNER IMPRESSIONS OF THE RECOMMENDED ACTIVITIES

For the most part, learners felt that the activities reinforced what they were learning in class (Table 4.15). The neutral result for this during phase two may be the result of a synchronization problem that was created because learners had slowed the pace of their activities or been doing so many activities that they got ahead of the class.

Participants felt that the exercises allowed them to both practise reading Russian aloud (Table 4.16) and speaking in Russian (Table 4.17). It appears that learners did not make a significant distinction between the two actions.

Table 4.15: Participant response to "The exercises reinforced what I was learning in class".

	Study Phase		
	1	2	3
Mean	2.2	4.4	1.8
Standard Deviation	1.2	1.1	0.4

Table 4.16: Participant response to "The exercises allowed me to practise reading aloud in Russian".

	Study Phase		
	1	2	3
Mean	1.5	2.8	1.6
Standard Deviation	0.8	1.9	0.5

Regardless of whether the learners felt that they were reading aloud or speaking in Russian, they felt that the activities helped them to focus on their pronunciation, even without the use of a learner model (Table 4.18).

Table 4.17: Participant response to "The exercises allowed me to practise speaking in Russian".

	Study Phase		
	1	2	3
Mean	2.0	2.6	1.8
Standard Deviation	0.9	0.9	0.4

Table 4.18: Participant response to "The exercises helped to focus my pronunciation when saying Russian words".

	Study Phase		
	1	2	3
Mean	2.8	3.6	2.0
Standard Deviation	1.5	1.7	0.7

Independent of the learner model type, learners sometimes chose activities based on the information shown to them in their learner model (Table 4.19). This is supported by the majority of the learners, accessing instructional material and their personalized activity recommendations. The one learner who viewed instructional material based on the feedback presented in his model, but did not attempt the personalized activity recommendations said that "I thought it was

interesting how [the model] was able to identify which sounds I wasn't pronouncing correctly. I didn't realize that they were areas of weakness. I haven't done them yet, but I plan to do the activities that are recommended for those sounds."

Learners also believed that they followed ProTutor's activity recommendations when they had access to the HOLM, but not when they had access to the OLM (Table 4.20).

Table 4.19: Participant responses to "I chose activities based on the information presented to me in my learner model".

	Study Phase	
	2	3
Mean	3.0	3.0
Standard Deviation	1.0	1.7

Table 4.20: Participant responses to "I followed ProTutor's activity recommendations".

	Study Phase	
	2	3
Mean	3.0	2.2
Standard Deviation	2.3	0.8

4.4. SYSTEM TRACKING OF LEARNERS' PROTUTOR USE

Four of the five participants chose to continue using ProTutor after the third survey, which was when the study's participation requirements were considered complete. Of these learners, all completed activities leading up to their December final exam. One learner even continued to use ProTutor into term 2.

For the purposes of discussing learner system use, we would like to use the term session to mean a period of time where the learner uses the system without considerable interruption. This means that they would have been actively performing activities or looking at other pages within ProTutor. If there is a considerable gap of time, more than 10 minutes, between activities then we consider the previous session to have ended. Keeping this in mind, the mean number of sessions across all learners was 4.60 with a standard deviation of 2.79, where 8 is the most sessions that any one learner performed and 2 is the least (Table 4.21). The learners' session size, measured in number of completed activities during a session, was also highly variable with a mean of 5.26 and a standard deviation of 2.69, and there seems to be no pattern or consistency in the session size for individual learners. This can be seen in Table 4.21, where we can see the

number of activities that each of the learners performed in each of their sessions. We can also see, in Table 4.22, how many activities each learner attempted and completed while using ProTutor.

Table 4.21: A summary of learner session sizes.

	Session Size				
Session	Learner1	Learner2	Learner3	Learner4	Learner5
1	7	5	6	9	2
2	6	2	1	4	1
3	3		3		2
4	13		2		10
5	3				4
6	22				3
7	12				7
8					2
Average	9.43	3.5	3	6.5	3.89
Standard Deviation	6.80	2.12	2.16	3.54	3.09

Table 4.22: The number of activities that each learner attempted and completed.

	Number of Activities	
Learner	Attempted	Completed
Learner 1	66	36
Learner 2	7	6
Learner 3	12	11
Learner 4	13	13
Learner 5	31	25

All of the learners completed the majority of the activity recommendations that they followed. However, there is one learner who only completed just over half of the activities that s/he attempted even though s/he attempted and completed by far the most activities of any of the learners. This may be because all of the activities that she did not complete had been previously completed by her and had been re-recommended as review material or because s/he disliked one of the activity types and may not have been paying attention to which activity recommendations

s/he was following. Overall, 91% of all attempted activities were completed, and of those completed activities, 82% were completed on the first attempt, an additional 8% were completed on the second attempt, and only 10 % took more than two attempts to complete.

All of the learners had visible changes in their learner models between the beginning and the end of the study. The changes in pronunciation accuracy of individual phones were visible because even small changes resulted in changes in the ranking of phones within the learner's model; a typical change in a learner's phone pronunciation accuracy would be a change of less than one percent or going from 0.499 to 0.503 and even if this change was insignificant, it still resulted in visible changes to the learner's model.

The visibility of these changes to the learner could not be measured because we could not measure the learner's ability to identify differences between models even though we were tracking the number of times each learner viewed his or her OLM and HOLM. It is possible that some of the learners' repeated views of the HOLM or OLM, that were temporally close, could have been the learners' looking for differences and changes in their models. In situations like this (Table 4.23), where the model could not have changed because there were no activities performed in between views of their learner models or where learners had not logged out for more than a few minutes, I counted multiple learner model views as one view since they did not need to be reminded of what was in their model and learners were not shown new information. Keeping this in mind, on average learners viewed their OLM 1.4 times, with a standard deviation of 0.55, and their HOLM an average of 3.6 times, with a standard deviation of 1.82, over a similar period of time. If I do not place this restriction on learner model views then learners viewed their OLM an average of 2.0 times with a standard deviation of 1.0 and they viewed their HOLM an average of 4.4 times with a standard deviation of 2.51. In either case, learners used their HOLM more than their OLM over a similar period of time (Table 4.24), which supports their statements regarding the HOLM version of ProTutor being more helpful to them.

Table 4.23 A sample of the logfile data for an insignificant view of the HOLM. The logfile lines in between the two HOLM Views show that the learner briefly (12 secs) viewed his profile page in between looking at his HOLM. This is likely because he saw the top of the HOLM and either forgot which avatar he had picked for himself or wanted to see if there was one he liked better for representing himself so he went to his profile page to see.

Learner	Log Row	Date	Time	URI	Action
Learner4	187	02/12/2009	21:30:55	/LearnerModel/HistoricLM.aspx	HOLM View
Learner4	188		21:30:55	/WebCharts/a4e70055-b272-4ef2-8487-cc59ab6d63f9.Png	Request for chart that is part of the HOLM
Learner4	189	02/12/2009	21:30:57	/LearnerProfile.aspx	Profile View
Learner4	190	02/12/2009	21:30:57	/Avatars/female_short_chestnut.png	Image request
Learner4	. . .	02/12/2009	. . .	. . .	Various image requests. One per row of the log
Learner4	212	02/12/2009	21:30:59	/Avatars/male_light_brunnette.png	Image Request
Learner4	213	02/12/2009	21:31:07	/LearnerModel/HistoricLM.aspx	HOLM View

Table 4.24 Learner Model views by model type and learner. The restricted views are controlled for the amount of time between views, so that there needed to be a significant amount of time or an activity done in between individual views.

		Learner					Statistics	
	Views	1	2	3	4	5	Mean	Standard Deviation
OLM	Restricted	2	1	1	1	2	1.4	0.55
	Non-restricted	3	2	1	1	3	2.0	1.0
HOLM	Restricted	6	1	2	4	3	3.2	1.92
	Non-restricted	8	1	4	5	4	4.4	2.51

In addition to viewing their learner models, four of the five learners chose to view instructional material based on their learner model by clicking on one of the phones from their phone list. Three of those learners attempted the recommended personalized activities, and two of them completed many of their personalized activity recommendations. Some of the personalized activities were attempted multiple times, but were not completed; this was done by learners who completed other activities as well as the learner who attempted but did not complete his or her personalized recommended activities. The lack of completion could have

been due to several reasons, like a learner running out of time or being more interested in reviewing vocabulary related to a theme that the learner would soon be tested on.

Three of the participants that looked at the instructional material and received personalized activities looked at the instructional material from phones on both their best and worst list. All of the personalized activities that these learners completed were based on phones in that learner's worst list. The learner who only looked at the instructional material for her best phones chose to work on the two phones that were heavily emphasized in class. The emphasis placed on these phones, by this learner's instructor, because of their difficulty for native English speakers (о and л) may have influenced which phones she chose to work on more than her model's feedback did. Three of those four learners also found the personalized activities to be helpful in some way, which is supported by their completion of many of the phone based personalized activity recommendations.

4.5. OVERALL IMPRESSIONS OF PROTUTOR

Generally, participants found ProTutor to be useful, which is indicated by 80% of them continuing to use ProTutor after completing the study requirements. Learner 2 stated that ProTutor "… is motivating, but more in a real-world sense. I'll probably think more about this when I speak in class", and learner 3 appreciated having the opportunity to practise in an environment where no-one else could listen, while at the same time receiving feedback about his pronunciation. Further to this, the open-ended survey questions revealed that each participant found ProTutor to have different positive qualities, some of which are summarized in Table 4.25.

Table 4.25: Some of the learner identified positive qualities of ProTutor

	Learner				
	1	2	3	4	5
Learner thought ProTutor was helpful		✓	✓	✓	
Learner thought ProTutor was easy to use	✓		✓		
Learner thought ProTutor was fun to use	✓				✓

The majority of the negative feedback about ProTutor had to do with the visual design of the system and the quality of the images used in the activities. Some participants also requested the addition of new features such as audio playback, the ability to pick any activity and theme, and more difficult activities where they could try to pronounce larger sections of text.

Overall these results show that learners viewed ProTutor in a positive light and indicate that ProTutor should be further developed and studied.

CHAPTER 5

DISCUSSION

This was a small initial study of the potential of ProTutor and its HOLM. We wanted to see if the use of gamelike activities, personalized recommendations, an open group model, and a HOLM would affect learner motivation and language acquisition. While the activities and personalized recommendations were performed and recommended in the same way throughout the study, the learner model treatments were not balanced for effect due to the formative nature of the study, the extensive time required to populate the HOLM, and the small number of participants that we were able to recruit. However, since we were looking to study the potential value of a HOLM in motivating learners, our study design was sufficient to determine if the HOLM and its parts should be further studied as a motivational tool in highly technical and difficult skill based domains. Future studies can be performed to see if there is a statistically significant difference between the use of the OLM and the HOLM and whether either of these affects learner outcomes.

The study's design may mean that the learner perception that ProTutor was better at helping identify learner strengths and weaknesses when using the HOLM than it was when using the OLM could be the result of increased system familiarity. Even though this interpretation of the results is supported by one learner's comment about his comfort with seeing a model of his performance, I do not believe that it is the sole reason for the difference since system familiarity did not change learner perceptions as to the usefulness of other portions of ProTutor, such as the recommended activities. That said, it is possible that the activities were initially perceived as more natural because they were based on familiar real world games and activities, and the learner's statement about the learner model that "It was weird at first. But now I think it's helpful. It's really beneficial to be able to compare and contrast", more accurately reflects why learner perceptions of the usefulness of the OLM were less positive than those of the HOLM.

The study and system design were limited in scope because of time constraints and our inability to rely on the accuracy of our speech recognition engine. Since the use of ProTutor as a support tool for a class was a primary goal of developing the system, that course's timeline prescribed when we had to release the software. This limited the amount of time that we had to develop certain features and required us to ensure a decent level of stability which meant that some features were cut and others were released iteratively, as they were needed.

Many factors, such as possible noise in the learners' environment, low quality microphones, inadequacies in the corpus used to train the speech recognition engine (VoxForge Russian Speech Corpus, 2007), and inadequacies in the recognition software itself, may have contributed to the lack of reliability of our speech recognition engine. A tight timeline also prevented us from discovering some of the problems with the speech recognition engine and modifying the configuration to maximize the speech recognition's accuracy before the study began. In practice the accuracy of the speech recognition may have been barely good enough to reliably identify the learner's top and bottom three phones.

Even though we were limited by the reliability of the speech recognition software, we were able to maintain learner trust in ProTutor and its HOLM by only showing learners their three best and worst phones. This only allowed learners to see overall changes in the ranking of how well they pronounced each phone without giving learners any information about the accuracy of their pronunciation for each of those phones. By limiting the model's feedback in this way, we seem to have succeeded in maintaining learner trust in the HOLM because participants responded positively to ProTutor and felt that the HOLM accurately reflected their pronunciation abilities. In addition to this, one of the learners commented that the feedback s/he received "wasn't too big of a surprise" and was consistent with her professor's feedback. Furthermore, by only showing learners what they were good at and what they needed to improve we were able to provide them with enough information to make informed decisions about what they wanted to work on without showing them the detailed information that might have made them question the validity of the feedback that we were giving them.

Even with this approach, learners who had a relatively consistent quality of pronunciation for the phones in their model could see some potentially confusing information. This was shown by

the comments of a participant who felt that the model accurately reflected his pronunciation most of the time. At one point he noted that "a couple of [his] previous best pronunciations became [his] weakest. And [he] doubt[ed] that [his] pronunciation of some of them [had] changed at all".

In addition to this problem, it was also possible that misdiagnoses led to learners trying to improve phones that they were pronouncing accurately and believing that they were properly pronouncing phones that they consistently mispronounced. In the first case, learners are in no way harmed because they are continuing to practice and re-inforce a skill that they already have. In the second case, ignoring the problem and continuing to practice a mispronunciation of a phone could result in the learner ingraining that error and the mistake could become a habit. While ProTutor does not protect against this, it is no worse than any of the other technologies and strategies that learners currently have available to them. The tapes that they listen to and repeat without supervision do not guard against this and the often unsupervised practice that they get in class and tutorial cannot effectively guard against this problem either.

The restriction of the amount of feedback provided to learners in the learner model may have helped guard against this problem. It also helped reduce the amount of information that learners were given so that they did not feel overwhelmed by the amount of information provided to them in the HOLM. We do, however, believe that learners would benefit from knowing how they are performing on certain phones that are of interest to them, even if those phones are not in the top or bottom three. For that reason, it might be interesting to allow learners to choose to view their performance of a specific phone. While I do not believe that the current speech recognition engine is accurate enough to support this, I think that this could easily be added whenever the reliability of the engine is substantially increased by controlling some of the factors that limit its reliability and by tweaking the Sphinx-3 decoder until we discover the best possible configuration given our constraints and needs or by requiring the learners to perform the activities in a more restricted environment.

However, it would be difficult to control for other factors such as input noise and the need to convert the audio files between different formats given the constraints on ProTutor's system use. Since we wanted learners to be able to use ProTutor from anywhere, we were limited in the technologies that we could use to capture their utterances. This meant that we were required to

convert the audio files between different formats and down-sample them so that they could be processed by the speech recognition engine. This conversion process may have introduced noise into the audio files, which could have negatively affected the speech recognition accuracy. We were also unable to control the quality of the learner's input because they were working from various locations that were potentially noisy and they could use any type of microphone. This added another possible error source if the audio recordings were very quiet or noisy because of the learner's choice of microphone and environment. We can require that they use a specific type of microphone, but we cannot ensure that their environment will be free from background noise.

Learners expressed that they found using ProTutor with the HOLM to be a positive experience and we believe that it helped maintain their motivation to continue improving their pronunciation since most of the learners indicated this in their long answers and they continued to use the system after the study's completion. We believe that the HOLM helped learners to maintain their motivation because it made it easier for them to identify their strengths and weaknesses and because they felt that ProTutor was helping them to improve their Russian pronunciation through the HOLM. This is supported by a learner's statement that being able to see his pronunciation changes in the HOLM was motivating and that he would carry his attention to pronunciation into the classroom setting.

The HOLM may have also increased learner reflection and affected the accuracy of each learner's self-perception. Based on the proficiency categories used in the surveys (poor, mediocre, good, excellent, fluent), most of the learners did not feel that their Russian pronunciation ability could be recategorized into a higher grouping even though they felt that their pronunciation had improved. It is possible that learners felt that their pronunciation had improved, just not to the point where they transcended their current pronunciation proficiency category. However, I believe that the HOLM also helped increase the accuracy of the learner's self-perception and I think that this is a potential benefit to learners that should be further studied.

The addition of the expert model to the OLM and HOLM may have been the portion of the model that helped learners to increase their self-awareness and the accuracy of their self-

perception. Even though learners may not have seen small differences between the expert's and their models, they still liked to compare the models. It is possible that the representative sentence portion of the model that contains the expert model may become more useful to learners as they become more practised at using their model and better trained. At this point it is not possible to tell how long or how frequently they compared their representative sentence phone mapping to that of the expert. However, tracking of this could be easily added when an audio component is added to the model, since we could track when and how often each of the sentences is listened to. Alternatively, we could add an interactive text based tool, which the learner could use to help him or her compare the two phone transcriptions; its use could be tracked in more detail because of its interactive nature.

Regardless of the nature of the model's interaction, open learner modeling is an especially useful tool for providing feedback in skill-based learning like second language pronunciation tutoring. Learners liked being able to see not only how they were doing through the learner model but also how their model had changed, which was made easier for them with the addition of historic information to their OLM. The addition of the open group model, balanced list of best and worst phones, expert model, and the representative sentence to the OLM and HOLM benefited learners by giving them enough summary and specific information to make decisions about their learning, and it helped maintain their motivation. The addition of the HOLM to the tutoring system was well-received by learners and helped keep them motivated throughout the learning process, which is especially challenging when working on difficult technical skills like pronunciation.

CHAPTER 6

CONCLUSION AND FUTURE DIRECTIONS

6.1. CONCLUSION

Second language learners face motivational challenges and time constraints when learning to speak another language. We performed a needs assessment of the U of S Russian language program and designed ProTutor, a CALL system, which will help learners remain motivated while learning to pronounce a language. ProTutor uses a combination of pedagogical activities, positive reinforcement, and a Historic Open Learner Model to motivate learners.

We have designed ProTutor to include a HOLM, which will encourage learner reflection and increase learner motivation by allowing learners to see the progress that they are making. ProTutor uses a constraint-based language model to detect errors and prescribe activities to help the learner overcome errors. ProTutor's HOLM limits its feedback to the user by only showing him or her, his or her current and previous three best and worst pronounced phones alongside the group's best and worst phones. This prevents the HOLM from overwhelming the learner, while balancing positive and negative feedback in order to maintain learner motivation (Barrow, Mitrovic, Ohlsson, & Grimley, 2008). It also draws the learner's attention to the most important aspects of his or her abilities. The addition of the group model allows the learner to see how the rest of his or her classmates are doing as a group, which helps maintain learner motivation since weaker students will see that they are not alone, and competitive students will be able to see roughly where they rank within the class. ProTutor's HOLM also displays an expert speaker's model so that learners know what they are working towards and can compare their performance to that of a native Russian speaker.

We performed a formative evaluation of ProTutor and its HOLM using students from beginner and intermediate Russian language classes at the University of Saskatchewan. The

results of this early evaluation indicate that the use of a HOLM can help motivate learners and should be further evaluated to determine its effect on learning outcomes, especially in highly technical skill-based domains like pronunciation tutoring. If we look at the success of ProTutor in getting students to continue working on their pronunciation, we can also see that the accuracy of the model and feedback is not the most important componenent of the system, provided that the turoring system is also motivating and helpful to learners. If we want to take this even further, we can compare ProTutor to human tutors; this shows that ProTutor is no worse than a benevolent error-prone human tutor since learners remained motivated and continued to use the system even though the feedback given by ProTutor, as a result of inaccuracies in the diagnosis process, was often incorrect.

Following this formative evaluation, we integrated ProTutor into the beginner and intermediate Russian language classes at the University of Saskatchewan where we can further evaluate the motivational and educational effects of the HOLM. In the future, we plan to add and evaluate an audio component to the HOLM so that learners can hear a representation of both their and the expert's pronunciation and compare the audio representations of the model in order to learn how to self-identify and self-correct errors in their pronunciation.

Through this preliminary, formative, evaluation we have shown the potential for:

- ProTutor and its HOLM to affect learner motivation
- the usefulness of the open group model
- the usefulness of the expert model
- the usefulness of the best and worst 3 phones at helping to focus the learner

Since ProTutor has shown its potential and will continue to be used, expanded, and studied, we hope to prove the effectiveness of the tools and approaches that it uses, especially with respect to learning outcomes. It needs to have more activities of different difficulty levels made available to the learner. The HOLM can be further expanded through the use of text to speech software that would add an audio model, and the addition of text to speech software could also provide an audio model for learners to hear and mimic. These additions and other future directions are discussed in the remainder of this chapter.

6.2. FUTURE DIRECTIONS

6.2.1. Activities and Pedagogical Planning

The version of ProTutor that was built and released to our participants is a proof of concept web-based platform that can be easily expanded and built upon. It does not contain all of the designed activities or detailed pedagogical planning that had been initially wanted. However, because of how we built the system the previously designed activities and the pedagogy that surrounds them can be added iteratively.

The drill and practice pedagogical activities were successfully released to learners with a focus on making them as game-like as possible. We can now add more complicated activities, which will fulfill some of the learners' desire for more challenging activities. The released version of the Pedagogical Planning Engine was far simpler than the initial design. The Planned Common Learning Path and Personalized Recommendations can now be modified to be more dynamic and adaptive. They can also be modified to use pedagogical conceptual frameworks, like Gagne's events of instruction (Kruse, 2006), rather than solely relying on expert tutoring strategies.

6.2.1.1. Activity Types

We can add several activity types by adding to the Language Resources library. It currently contains some compositions, but it would be capable of supporting a wider variety of activities if we were to add additional compositions (stories, poetry, sentences, and excerpts from novels), tongue twisters, songs, audio clips, video clips, dictionaries (audio, visual, and textual), maps, names, and descriptions of cultural events. The addition of these resources would allow us to use the four activity types found below instead of only having the learner record individual words:

- syllable add-ons: start with a single syllable and add syllables until the sounds can be pronounced in a full word.
- word add-ons: start with a single word and add words on until a sentence can be properly pronounced.

- compositions (paragraphs, poems, or songs): these will help to reinforce the rhythm of the language along with pronunciation, and learners can be rewarded with the use of audio or video clips. Furthermore, songs and poems expose a learner to rhyming, which helps learners identify phones and distinguish between them (Palmer, 2007).
- tongue twisters: these are especially challenging combinations of sounds. They also reinforce the rhythm of the language.

The addition of songs and tongue twisters is particularly important because they reinforce the rhythmic use of the second language and help develop auditory memory, which is an important skill for language learning (Palmer, 2007). Rhythmic tasks also expose learners to the natural flow of the language, which learners can then mimic and later master with additional practice.

6.2.1.2. Pedagogical Plan

Some of the added pedagogical activities would require that the learner read parts of the text, while the system reads the remaining parts of the text that belong to the story or article. This read-along approach would help to train sounds that the learner needs to improve as well as provide the learner with an accomplishable task so that learner motivation is maintained (Casey, 2007). With the addition of audio and video materials to the Language Resources Library, we would be able to have some texts followed by video clips of a segment of the same story that is presented in the text that the learner has just completed reading. This would help motivate the learner by providing positive reinforcement and rewarding the effort that the learner has made (Johnson, Wu, & Nouhi, 2004) (Vosniadou, 2001) (Lu, Di Eugenio, Kershaw, Ohlsson, & Corrigan-Halpern, 2006). Furthermore, viewing video clips following their reading practice serves to reinforce the knowledge that learners acquired while reading a section of text and it allows them to practise listening to an accurate model of the foreign language.

The integration of the above mentioned activity types would allow learners to progress through a series of tasks of increasing difficulty for sounds that they are having trouble mastering. They would be able to start at any point in the activity progression. Learners would have to practise their pronunciation in increasingly difficult environments until they encounter

difficulty with a phone, at which point the tasks assigned to learners for practice would be simplified so that learners can improve their performance.

The first task assigned to the learner from the progression would depend on the learner's previous performance and abilities, all of which are tracked in the learner model. Learners could also be given the chance to change the difficulty level of the tasks that they are performing, making the system both adaptive and adaptable. If a learner feels like s/he is having a bad day, the learner could lower the difficulty level of the tasks that s/he is performing and if s/he is feeling brave, s/he could jump right to the most difficult level.

Learner effort and progress through the assigned and chosen tasks could be rewarded with a navigation simulation environment where the learner plays a game in which s/he either gives or follows directions through a city (Lu, Di Eugenio, Kershaw, Ohlsson, & Corrigan-Halpern, 2006). This game would be used to support learning by coupling it with the previously mentioned educational tasks and procedural phone instruction (Conati, 2007). It is hoped that this navigation activity would provide enough of a simulation based environment to benefit learners, since simulation based games allow learners to assimilate previously acquired knowledge with the experiences that they have within the simulation environment (Tan, Skirvin, Biswas, & Catley, 2007)

The final addition that we would like to make to the pedagogical plan is to update some of the instructional material by adding a graphical representation (images and video) of a simulated expert speaker. This would model the formation of the mouth for the phones that need to be produced, much like watching a person speak on the television does. A strategically chosen set of phones could be targeted for this improvement and then the instructional material for other phones could later be upgraded as time permits.

The process of planning the learner's path can also be updated based on the learner's self-assessment of their skills in Russian and the learner model's knowledge of the learner's mother tongue and other language abilities. This information could inform the system as to which activities might be most useful in helping the learner to overcome his or her pronunciation errors. Furthermore, ProTutor could use the stored information about learner backgrounds and changes in learner performance to recommend activities that resulted in improved pronunciation for a

learner with a similar background. This would provide additional personalization and the possibility for creating and using learner stereotypes that could be used to provide additional help to learners as their model is being populated.

6.2.2. Text to Speech

We want ProTutor to allow students to listen to spoken Russian since this will help them practise their listening skills. This listening practice will address the need to help learners identify phones, which is the basic building block of listening comprehension. For this reason, we could augment ProTutor by using Text to Speech (TTS) software both to give the learner feedback and to provide a model of a native Russian speaker for the learner to listen to because TTS of a native speaker can help reinforce correct pronunciation by providing accurate pronunciation examples to the learner (Vosniadou, 2001). The TTS could also provide feedback through the learner models by emulating how an expert speaker would pronounce something. We believe that the addition of this functionality is both potentially beneficial to learners and desired by them.

6.2.2.1. TTS Software

Using a TTS engine that allows the reader's speed to be adjusted would allow ProTutor to have the learner listen to texts and other system feedback at different speeds, which would more closely simulate the variety of native speakers that a learner might encounter if s/he were immersed in a second language environment. This would also allow the learner to be challenged by increasing the rate of the simulated speech since this requires that the learner remember more and that the learner quickly identifies the phones and words that s/he is hearing. In contrast, a carefully pronounced slow rate of speech may be easier for the learner to understand (Archibald & O'Grady, 2008).

6.2.2.2. Audio Open Learner Model

To create an Audio Open Learner Model (AOLM) we could add TTS functionality to the representative sentence portion of the learner models. This would allow the learner to hear how

s/he would typically pronounce his or her representative sentence and compare that to how a native Russian speaker sounds saying the same sentence. Even though learners did not express a need to hear the model, one of them expressed a desire to hear his fellow classmates' models. However, he did not want his classmates to listen to him because they would be able to recognize his voice. He then stated that "maybe if it was confidential it wouldn't be so weird, but there are so few people I'm sure I would recognize their voices and vice versa". If listening to fellow learners' models was found to be beneficial, the addition of TTS could allow for this while maintaining learner confidentiality. The addition of TTS would also accommodate the learner's desire to see how s/he is doing in comparison to his or her classmates while protecting the other person's privacy since that person's voice would not be used and therefore could not be recognized.

6.2.2.3. Expert Modeling

TTS as well as audio and video clips, which model the correct pronunciation of words in the target second language, fulfill the learner's need to hear the language spoken, and learners could be able to request that the system read them a segment of text, which would help reinforce their listening and reading skills (Casey, 2007). The use of audio and video clips would also serve to model the correct pronunciation of the language much as an expert tutor would (Lu, Di Eugenio, Kershaw, Ohlsson, & Corrigan-Halpern, 2006).

Within the pedagogical activities, the reading of text by an expert speaker could be performed using TTS because that simulates expert speakers and would allow learners to have the pronunciation of words modeled when they want and need them modeled. Learners would simply select the word or words that they would like to hear and ProTutor could read the words aloud. This means that learners can easily listen to just one word or a series of words that they select, and that learners do not need to listen to disconnected audio clips to hear a complete phrase or sentence.

6.2.3. HOLM Expansion

6.2.3.1. Diagnosis

The diagnosis engine can be updated so that we can track and use information about learner performance on sequences of phones rather than only tracking performance information for individual phones. This would allow us to have the representative sentence more accurately reflect the type of errors that the learner makes in different situations.

6.2.3.2. Visualization

The visualizations within the HOLM can be improved through minor changes and additions that would help clarify how the learner's pronunciation proficiency has changed. An example of one of these changes would be adding an arrow beside phones in the best and worst phones list so that learners could see how each phone in the list has changed rather than just see changes in the phones' rankings.

The HOLM can also be expanded so that learners can select the dates for which they would like to see their model. This would allow them to view old versions of their model so that they could see their progression over time rather than just compare their performance at two points in time. Once the speech recognition has been improved, the HOLM can be further expanded by increasing the accuracy of the phone transcription of the learner's representative sentence so that it shows the frequency and type of error that the learner makes.

6.2.4. HOLM Evaluation

The HOLM needs to be further evaluated with a study that is balanced for learning effects of the various learner models. Alternatively, a study could be run with two treatment groups one of which is given the OLM and the other the HOLM. In addition to balancing the treatments, the HOLM needs to be evaluated for its effects on the rate at which learners skills change when using the HOLM as compared to using the other treatment options (the OLM and no model).

The validity of the system's diagnoses and pedagogical task assignment should be evaluated using expert analysis where experts code the speech that was recorded from an exercise performed by the learner so that we can compare the system's coding of the recorded exercise speech to that of the experts. Since ProTutor identifies mispronunciations of specific phones, the human coder would do the same. For example, if the learner pronounces an "o" when s/he should have pronounced an "a", the coder would record an error in that location of the transcribed speech. Comparing ProTutor's error identification with the coder's analysis of the same segment of speech would allow us to see how accurately the system can identify mispronunciations.

Coding the learner's speech at the beginning and end of a pedagogical exercise could be used to evaluate the effectiveness of the assigned tasks. The end coding could be compared to the initial coding to find differences and changes in learner performance. If the learner corrects the pronunciation of the phones, then the exercise may have been effective. Furthermore, the system could record sessions at the beginning and end of a classroom theme, without providing any language tutoring, so that the progress made using only classroom instruction can be compared to the progress made when the system tutors the learner. This would help to show how ProTutor affects learning outcomes, by showing the changes in learner performance while they are using the system.

In addition to determining how the HOLM and OLM affect learning outcomes, these two models need to be better evaluated for how they affect learner reflection about changes in their models and how this affects activity choices. This might be initially studied using a diary study or through several observed think-aloud sessions. If we had a larger pool of participants who had used ProTutor more, data mining tools and techniques could be used to identify patterns between changes in a learner's model and changes in his or her activity selection behaviours, but these tools would not be able to identify why the learner made those changes, which is why we need to collect information about the learner's motivation when performing tasks and making decisions.

LIST OF REFERENCES

Archibald, J., & O'Grady, W. (2008). *Contemporary Linguistic Analysis.* Pearson.

Barrow, D., Mitrovic, A., Ohlsson, S., & Grimley, M. (2008). Assessing the Impact of Positive Feedback in Constraint-Based Tutors. *Intelligent Tutoring Systems (ITS)*, (pp. 250--259). Montreal.

Beck, J. (2007). Does learner control affect learning? *13th International Conference on Artificial Intelligence in Education (AIED)*, (pp. 135-142). Los Angeles.

Beck, J. E., Jia, P., & Mostow, J. (2003). Assessing student proficiency in a Reading Tutor that listens. *9th International Conference on User Modeling (UM)*, (pp. 323 - 327). Johnstown.

Blair, K., Schwartz, D., Biswas, G., & Leelawong, K. (2006). Pedagogical Agents for Learning by Teaching: Teachable Agents. *Educational Technology & Society, Special Issue on Pedagogical Agents* , 56-61.

Bull, S., & Kay, J. (2007). Student Models that Invite the Learner In: The SMILI:() Open Learner Modelling Framework. *Int. J. Artif. Intell. Ed. , 17* (2), 89-120.

Burston, J. (2008). *CALICO Software Review - BonPatron: An Online Spelling, Grammar, and Expression Checker.* Retrieved March 28, 2010, from CALICO Journal: https://www.calico.org/p-381-BonPatron%20%28012008%29.html

Carnegie Mellon University. (2007). *CMU Sphinx*. Retrieved April 2, 2008, from SourceForge: http://cmusphinx.sourceforge.net/html/cmusphinx.php

Casey, B. (2007). *Reading with Under Fives.* Retrieved October 18, 2007, from BBC: http://www.bbc.co.uk/cbeebies/grownups/about/howchildrenlearn/reading_with_under_fives.shtml

Chapelle, C. (1998). MULTIMEDIA CALL: LESSONS TO BE LEARNED. *Language Learning & Technology , 2* (1), 22-34.

Collentine, J. (2002). REVIEW OF TESOROS: A MULTIMEDIA-BASED SPANISH COURSE ON CD-ROM. *Language Learning & Technology , 6* (1), 33-39.

Conati, C. (2007). Supple Interfaces for Education: can Affective Interactions improve Learning via Educational Games? *Computer/Human Interaction (CHI) Supple Interfaces Workshop.* San Jose.

Concentration (game). (2010). Retrieved February 2, 2010, from Wikipedia: http://en.wikipedia.org/wiki/Concentration_%28game%29

Extemporel Inc. (2008). *Chuala*. Retrieved April 2, 2008, from Chuala: http://www.chuala.com/

Flip Words. (2008). Retrieved September 4, 2008, from Primary Games: http://www.primarygames.com/langarts/flipwords/index.htm

Hofs, D. (2002). *Vocatude.* Retrieved April 2, 2008, from http://www.oriente.nl/voca/index.php

International Cafe. (2007). Retrieved August 25, 2008, from http://www.transparent.com/games/cafe/flash.htm

Jianguo, L., & Xiaozhen, Z. (2004). Design and implement of multimedia intelligent tutoring system for Chinese character. *IEEE International Conference on Multi-Media Engineering Education*, (pp. 459-463).

Johnson, W. L., Marsella, S., & Vilhjalmsson, H. (2004). The DARWARS Tactical Language Training System. *Interservice/Industry Training, Simulation and Education Conference.* Orlando.

Johnson, W. L., Wu, S., & Nouhi, Y. (2004). Socially intelligent pronunciation feedback for second language learning. *International Conference on Intelligent User Interface (IUI) Workshop in Modeling Human Teaching Tactics and Strategies.* Island of Madeira.

Kruse, K. (2006). *Gagne's Nine Events of Instruction: An Introduction.* Retrieved June 27, 2010, from e-learningguru.com: http://www.e-learningguru.com/articles/art3_3.htm

Li, X., Ju, Y., Deng, L., & Acero, A. (2007). Efficient and robust language modeling in an automatic children's reading tutor system. *2007 ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing Proceedings 4*, (pp. IV193-V196).

Lord, G. (2008). CALICO Software Review - BiLingo Kidz. *CALICO* , 204-219.

Lu, X., Di Eugenio, B., Kershaw, T. C., Ohlsson, S., & Corrigan-Halpern, A. (2006). Tutorial Dialogue Patterns: Expert vs. Non-expert Tutors. *3rd Midwest Computational Linguistics Colloquium*. Urbana.

Martin, B., & Nicholas, A. (2007). Studying Model Ambiguity in a Language ITS. *11th International Conference on User Modeling (UM)*, (pp. 425-429). Corfu.

Mich, O., Betta, E., & Giuliani, D. (2004). PARLING: e-Literature for Supporting Children Learning English as a Second Language. *9th international Conference on Intelligent User Interfaces (IUI)* (pp. 283-285). Funchal, Madeira, Portugal: ACM.

Microsoft Corporation. (2010). *NEWID (Transact-SQL)*. Retrieved February 6, 2010, from msdn SQL Server Developer Center: http://msdn.microsoft.com/en-us/library/ms190348.aspx

Nagamani, M., Narendra prasad, M., & Girija, P. N. (2005). Improving reading and writing skills with intelligent tutor for Telegu language learning - INTELL. *2005 International Conference on Intelligent Sensing and Information Processing Proceedings*, (pp. 426-49b).

Nummikoski, M. (1996). *Troika A Communicative Approach to Russian Language, Life, and Culture (Тройка)*. Toronto: John Wiley & Sons Inc.

Palmer, S. (2007). *Listening, Speaking and Literacy*. Retrieved October 18, 2007, from BBC: http://www.bbc.co.uk/cbeebies/grownups/about/howchildrenlearn/listening_speaking_literacy_02.shtml

Pantelia, M. (2007, Sept.). *CALICO Software Review - Filoglossia+*. Retrieved March 27, 2010, from CALICO Journal: https://www.calico.org/p-341-Filoglossia%2B%20%2892007%29.html

Papadopoulou, Y. (2008, Jan.). *CALICO Software Review - Read, Write & Type! Learning System, Wordy Qwerty, and Spaceship Challenge.* Retrieved March 27, 2010, from CALICO Journal: https://www.calico.org/p-382-ReadWrite&amp%3BType%20%28012008%29.html

Pardo Ballester, C. (2009). CALICO Software Review - En una palabra, Córdoba, Argentina: A CD-ROM for Exploring Culture in Spanish. *CALICO , 26* (3), 674-683.

Picard, R. (2007). *Affective Computing.* Cambridge, Massachusetts: MIT Press.

Price, C., McCalla, G., & Bunt, A. (1999). L2tutor: a mixed-initiative dialogue system for improving fluency. *Computer Assisted Language Learning Journal 12(2)* , 83-112.

Pumpkimber. (2006). *Super Munchers*. Retrieved September 4, 2008, from Classic-PC-Games: http://www.classic-pc-games.com/pc/educational/super_munchers.html

RosettaStone. (2008). Retrieved August 25, 2008, from RosettaStone: http://www.rosettastone.com

Salmerón, F., & Burston, J. (2008). *CALICO Software Review - Viaje al pasado: los aztecas.* Retrieved March 27, 2010, from CALICO Journal: https://www.calico.org/p-380-Viaje%20al%20pasado:%20los%20aztecas%20%28012008%29.html

Tan, J., & Biswas, G. (2006). The Role of Feedback in Preparation for Future Learning: A Case Study in Learning by Teaching Environments. *Intelligent Tutoring Systems (ITS)*, (pp. 370-381). Jhongli, Taiwan.

Tan, J., Skirvin, N., Biswas, G., & Catley, K. (2007). Providing Guidance and Opportunities for Self-Assessment and Transfer in a Simulation Environment for Discovery Learning. *29th Annual Meeting of the Cognitive Science Society*, (p. 1539). Nashville.

Tao, R. (2007). *CALICO Software Review - Eyespeak.* Retrieved March 28, 2010, from CALICO Journal: https://www.calico.org/p-340-Eyespeak%20%2892007%29.html

Teach 2000. (2010). Retrieved June 23, 2010, from http://www.teach2000.org/

Translator Alligator. (2008). (Pearson Education, Inc.) Retrieved September 4, 2008, from FunBrain: http://www.funbrain.com/lang/index.html

Tsiriga, V., & Virvou, M. (2002). Dynamically initializing the student model in a Web-based language tutor. *1st International IEEE Symposium on Intelligent Systems*, (pp. 138-143). Varna.

Vocab Box. (2000). Retrieved April 2, 2008, from Free Downloads Center: http://www.freedownloadscenter.com/Business/Educational_Tools/Vocab_Box.html

Vosniadou, S. (2001). *How children learn. Educational Practices Series-7.* The International Academy of Education (IAE) and the International Bureau of Education (UNESCO).

VoxForge Russian Speech Corpus. (2007). Retrieved January 15, 2010, from VoxForge: http://www.dev.voxforge.org/projects/Russian/browser/Trunk/AcousticModels/README

Wang, C., & Seneff, S. (2007). A Spoken Translation Game for Second Language Learning. *Artificial Intelligence in Education (AIED).* Marina del Rey.

Word Confusion. (2008). (Pearson Education, Inc.) Retrieved September 4, 2008, from Fun Brain: http://www.funbrain.com/whichword/index.html

APPENDIX A

NEEDS ASSESSMENT SURVEY

Hello,

Please fill out one of these surveys for each level of Russian that you have taught at the U of S. For the purposes of this study the first level consists of students in Rus114, Rus117, and Rus214, and the second level consists of students in Rus217 or higher. There is plenty of room for you to add additional information in the form of comments. However, you are not required to comment.

Sincerely,

Carrie Demmans

Speaking						
My students can form sentences within a conversation						
Agree	1	2	3	4	5	Disagree
My students can ask questions about course material						
Agree	1	2	3	4	5	Disagree
My students have the vocabulary necessary to converse about the topics we have covered						
Agree	1	2	3	4	5	Disagree
My students have the grammatical knowledge necessary to converse about the topics we have covered.						
Agree	1	2	3	4	5	Disagree
My students can form sentences at a reasonable speed						
Agree	1	2	3	4	5	Disagree
My students pronounce known words properly						
Agree	1	2	3	4	5	Disagree
My students use the correct stresses when speaking						
Agree	1	2	3	4	5	Disagree
My students use the correct intonation when speaking						
Agree	1	2	3	4	5	Disagree
Comments:						

Listening						
My students have good listening comprehension						
Agree	1	2	3	4	5	Disagree
My students recognize the words that they hear						
Agree	1	2	3	4	5	Disagree
My students can transcribe speech						
Agree	1	2	3	4	5	Disagree
My students spell the words correctly when they transcribe speech						
Agree	1	2	3	4	5	Disagree
My students write the correct words when transcribing speech						
Agree	1	2	3	4	5	Disagree
My students transcribe the grammatical constructs as spoken by the speaker						
Agree	1	2	3	4	5	Disagree
My students can summarize stories that they listen to						
Agree	1	2	3	4	5	Disagree
My students can answer questions about stories that they listen to						
Agree	1	2	3	4	5	Disagree
Comments:						

Reading						
My students have good reading comprehension						
Agree	1	2	3	4	5	Disagree
My students recognize the words that they read within a text						
Agree	1	2	3	4	5	Disagree
My students recognize the grammatical constructs used within a text						
Agree	1	2	3	4	5	Disagree
My students understand individual sentences within a text						
Agree	1	2	3	4	5	Disagree
My students can summarize the texts that they read						
Agree	1	2	3	4	5	Disagree
My students can answer questions about the texts that they read						
Agree	1	2	3	4	5	Disagree
Comments:						

Writing						
My students use correct grammar in their compositions						
Agree	1	2	3	4	5	Disagree
My students use the correct words in their compositions						
Agree	1	2	3	4	5	Disagree
My students use the correct word order in their compositions						
Agree	1	2	3	4	5	Disagree
My students use an appropriate format for their compositions						
Agree	1	2	3	4	5	Disagree
My students can compose a text within a reasonable time period						
Agree	1	2	3	4	5	Disagree
Comments:						

Please rank the following (1 - most important, 5 - least important):
My students need help with:
___ Speaking
___ Writing
___ Reading
___ Listening
___ Other (_____)

It is essential that my students can:
___ Speak
___ Write
___ Read
___ Listen
___ Other (_____)

When speaking students need correct:
___ Grammar
___ Pronunciation
___ Intonation
___ Word choice
___ Other (_____)

When writing students need correct:
___ Grammar
___ Spelling
___ Word choice
___ Other (_____)

When listening students need to:
___ Recognize and understand grammatical constructs
___ Recognize and understand vocabulary
___ Recognize and understand intonation
___ Recognize and understand idioms
___ Other (_____)

When reading students need to:
___ Recognize and understand grammatical constructs
___ Recognize and understand vocabulary
___ Recognize and understand idioms
___ Other (_____)

APPENDIX B

ETHICS APPROVAL

STUDY APPROVAL AND CONSENT FORM

The Recommending Learning Activities and Representing Learner Performance in ProTutor, an Electronic Second Language Pronunciation Tutor study protocol was approved by the University of Saskatchewan's Behavioural Research Ethics Board on 29 April 2009. The certificate of approval can be found on page 96. A copy of the consent form begins on page 97.

UNIVERSITY OF SASKATCHEWAN

Behavioural Research Ethics Board (Beh-REB)

Certificate of Approval

PRINCIPAL INVESTIGATOR	DEPARTMENT	BEH#
Gordon McCalla	Computer Science	09-69

INSTITUTION(S) WHERE RESEARCH WILL BE CONDUCTED
University of Saskatchewan
Saskatoon SK

STUDENT RESEARCHERS
Carrie Demmans Epp

SPONSOR
NATURAL SCIENCES & ENGINEERING RESEARCH COUNCIL OF CANADA (NSERC)

TITLE
Recommending Learning Activities and Representing Learner Performance in ProTutor, an Electronic Second Language Pronunciation Tutor

ORIGINAL REVIEW DATE	APPROVAL ON	APPROVAL OF:	EXPIRY DATE
05-Apr-2009	29-Apr-2009	Ethics Application Consent Protocol	28-Apr-2010

Full Board Meeting ☐ Date of Full Board Meeting:

Delegated Review ☒

CERTIFICATION
The University of Saskatchewan Behavioural Research Ethics Board has reviewed the above-named research project. The proposal was found to be acceptable on ethical grounds. The principal investigator has the responsibility for any other administrative or regulatory approvals that may pertain to this research project, and for ensuring that the authorized research is carried out according to the conditions outlined in the original protocol submitted for ethics review. This Certificate of Approval is valid for the above time period provided there is no change in experimental protocol or consent process or documents.

Any significant changes to your proposed method, or your consent and recruitment procedures should be reported to the Chair for Research Ethics Board consideration in advance of its implementation.

ONGOING REVIEW REQUIREMENTS
In order to receive annual renewal, a status report must be submitted to the REB Chair for Board consideration within one month of the current expiry date each year the study remains open, and upon study completion. Please refer to the following website for further instructions: *http://www.usask.ca/research/ethics_review*

John Rigby, Chair
University of Saskatchewan
Behavioural Research Ethics Board

Please send all correspondence to:

Research Ethics Office
University of Saskatchewan
Box 5000 RPO University, 1602-110 Gymnasium Place
Saskatoon SK S7N 4J8

DEPARTMENT OF COMPUTER SCIENCE

INFORMED CONSENT FORM

This consent should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information is not included here, please contact one of the study investigators that are listed below. Please take the time to read this form carefully and to understand any accompanying information.

Research Project: Recommending learning activities and visualizing learner performance in ProTutor

Investigators: Dr. Gordon McCalla, Professor, Department of Computer Science (966-4902), mccalla@cs.usask.ca

Carrie Demmans Epp, Masters Student, Department of Computer Science (966-2676), cad574@mail.usask.ca

*Purpose and Procedure:*In this study you will have learning material recommended to you based on various personal characteristics that you have. Some of these characteristics will be explicitly asked from you (e.g. what languages you speak), while others will be derived from actions that you perform within the learning environment, which is called ProTutor (e.g. what type of learning activities help you the most). In addition, you will be asked to fill out a short questionnaire before, during, and after using ProTutor. We hope to help you improve your ability to properly pronounce Russian by using ProTutor, and want to determine if our approach to pronunciation tutoring can do this.

The completion of all of the questionnaires should take approximately an hour and a half. The amount of time that you use the system before filling out the end questionnaire is up to you. You may only perform a few pronunciation activities, or you might perform them all. This depends on your desire to use the system and improve your ability to speak Russian.

The data collected from this study will be used in thesis work and in articles for publication in journals and conference proceedings. Any write-ups of the data will not include any information that can be linked directly to you. As one way of thanking you for your time, we will be pleased

to make available to you a summary of the results of this study once they have been compiled. This summary will outline the research and discuss our findings and recommendations. Our appreciation for your participating in this study will also include remuneration of $50. If you choose to withdraw from the study early, your remuneration will be prorated based on the length of your involvement in the study.

Potential Benefits: Participation in this study may result in a better understanding of how you learn and an improvement in your ability to pronounce words and speak in Russian. It will also give you the opportunity to practice reading in Russian while receiving feedback without having to worry about how your classmates will respond to your efforts.

Potential Risks: You may continue to practice pronunciation errors, which may cause you to develop a habitual mispronunciation of certain Russian sounds. This is the same risk that you take when repeating the phrases on the practice tapes that accompany your class workbook or when you read texts aloud without an instructor's feedback.

If at any time, we become aware of additional risks, we will postpone the study and notify you immediately so that these risks can be minimized before we continue with research activities or, if necessary, stop the study altogether.

Confidentiality: All of the information we collect from you (data logged by the computer, observations made by the experimenters, and your questionnaire responses) will be stored so that your name, student number, nsid, or email address is not associated with it.

None of the information collected will be shared with your instructor. None of the information collected will influence your grade in any of your courses.

Storage of Data: The anonymized research materials will be stored with complete security, by Professor Gordon McCalla in the ARIES research laboratory, throughout the entire investigation. The data will be kept for a minimum of 5 years after which it may be destroyed. If after 5 years we no longer need the data and choose to destroy it, we will delete all of the files, reformat the disk on which they were stored, and overwrite the disk with new data a minimum of two times.

Right to Withdraw: Your continued participation should be as informed as your initial consent, so you should feel free to ask for clarification or new information throughout your participation. If you have further questions concerning matters related to this research, please contact one of the investigators listed above.

Your participation is voluntary, and you can answer only those questions that you are comfortable with. There is no guarantee that you will personally benefit from your involvement in this study. The information that you share will be held in strict confidence and discussed only with the research team. You may withdraw from the research project for any reason, at any time, without penalty of any sort.

If you do not participate in this study or choose to withdraw from this study, the data that was automatically collected by the learning environment, ProTutor, will be discarded once you have finished using the system.

Questions: At any time during the study you may contact those listed below to have any questions concerning the research project answered.

- Dr. Gordon McCalla, Professor Dept. of Computer Science (306) 966-4902, mccalla@cs.usask.ca
- Carrie Demmans Epp, MS Student, Dept. of Computer Science (306) 966-2676, cad574@mail.usask.ca

This study has been approved by on ethical grounds by the University of Saskatchewan Behavioural Research Ethics Board on 29 April 2009. Any questions you have regarding your rights as a participant may be addressed to that committee through the Office of Research Ethics at the University of Saskatchewan (306) 966-2084. If you live outside of Saskatoon, and are participating in this study you may call the Research Ethics Office collect.

If you would like a copy of this consent form please print it through your browser, or contact one of the investigators listed above. This research has the ethical approval of the Office of Research Services at the University of Saskatchewan.

Consent to Participate: By choosing the "yes" option below, you indicate that you have understood to your satisfaction the information regarding participation in the research project and agree to participate. In no way does this waive your legal rights nor release the investigators, sponsors, or involved institutions from their legal and professional responsibilities.

I have read and understood the description provided; I have had an opportunity to ask questions and my questions have been answered. I consent to participate in the research project, understanding that I may withdraw my consent at any time. A copy of this consent form has been made available to me for my records.

Date: 29/04/2010

Yes No

APPENDIX C

LEARNER SURVEYS

The Learner Initialization Survey (Figure C.1) is the only survey whose sections are unique. The other surveys build on each other.

LEARNER INITIALIZATION SURVEY

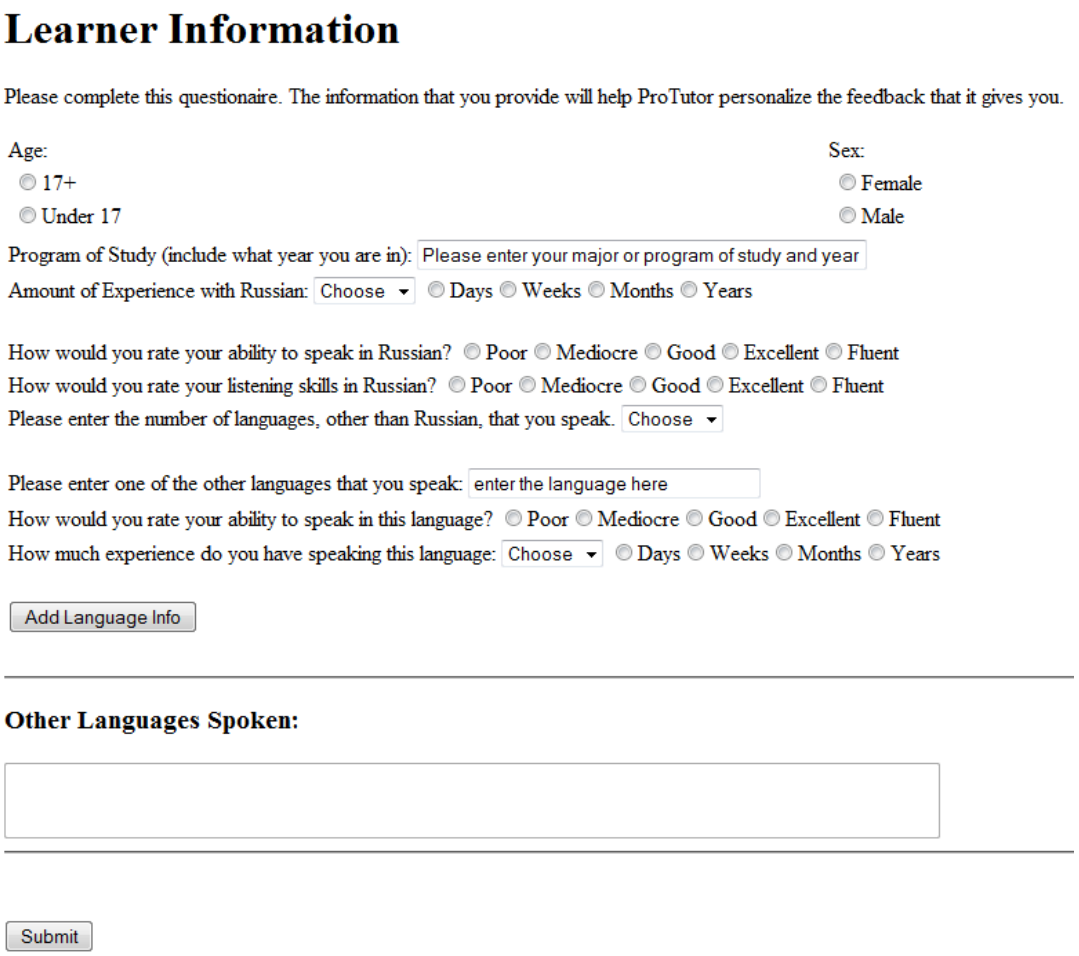

Figure C.1: A screenshot of the Learner Initialization Survey.

SURVEYS USED FOR THE STUDY

PHASE 1 SURVEY

This survey consists of two parts and is given to the learner after the first phase of the study. It is the base upon which the other surveys are built. When this survey is administered, the learner has yet to see a model of his or her abilities or receive feedback about his or her pronunciation.

Activities Section

This section of the survey collects information about how learners perceived the recommended activities and how they feel that those activities affect them. It also collects information about their perceived changes in performance as a result of using ProTutor. The below list contains the statements that the learner is asked to agree or disagree with in the order that the learner sees them. An excerpt of the survey is shown in Figure C.2.

1. The exercises reinforced what I was learning in class.
2. The exercises allowed me to practise reading aloud in Russian.
3. The exercises allowed me to practise speaking in Russian.
4. The exercises helped to focus my pronunciation when saying Russian words.
5. ProTutor helped me to identify when I was pronouncing something correctly.
6. ProTutor helped me to identify when I was mispronouncing something.
7. ProTutor helped me to work on my weaknesses.
8. ProTutor tried to help me improve my Russian pronunciation.
9. I feel that my pronunciation of individual Russian sounds has improved.
10. I feel that my pronunciation of Russian words has improved.

Self-reporting Section

In this section learners report on the other language activities that they participate in (Figure C.3). They also perform a self-rating of their Russian pronunciation skills, much like they did in the LEARNER INITIALIZATION SURVEY.

.

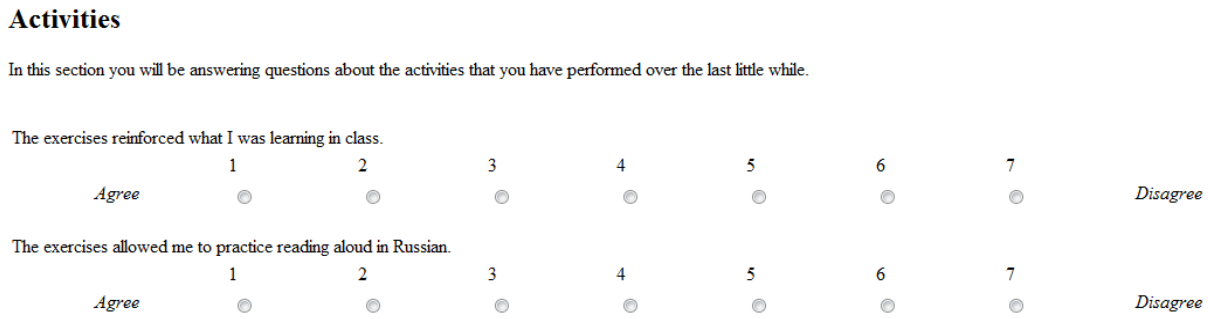

Figure C.2: A screenshot of the first part of the activities section of the survey.

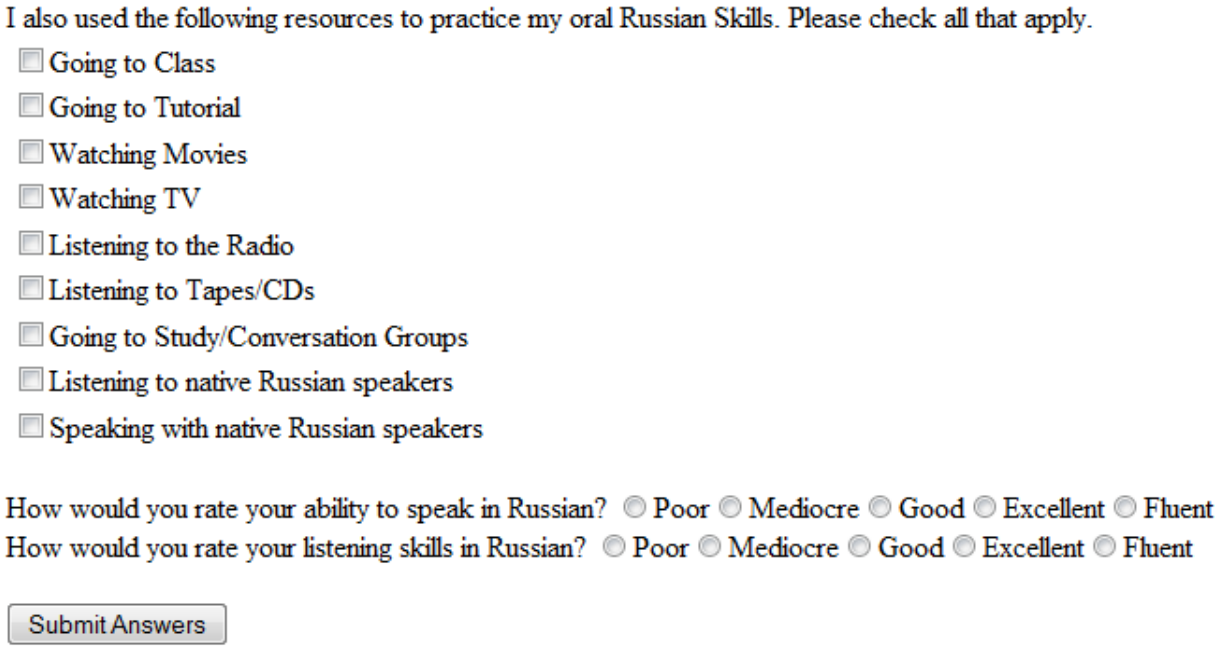

Figure C.3: A screenshot of the self-reporting section of the survey.

PHASE 2 SURVEY

The survey administered after the OLM phase starts with the Activities Section from phase 1. This is followed by an Activity Characteristics Section and then the Self-reporting Section from phase 1. The survey ends by collecting information about the learner's perceptions of characteristics of the OLM.

Activity Characteristics Section

In this section, the learner is asked to identify which activities were easy, helpful, and difficult as well as the reason that the learner believes that the identified activities have this characteristic. An excerpt of the survey can be seen in (Figure C.4)

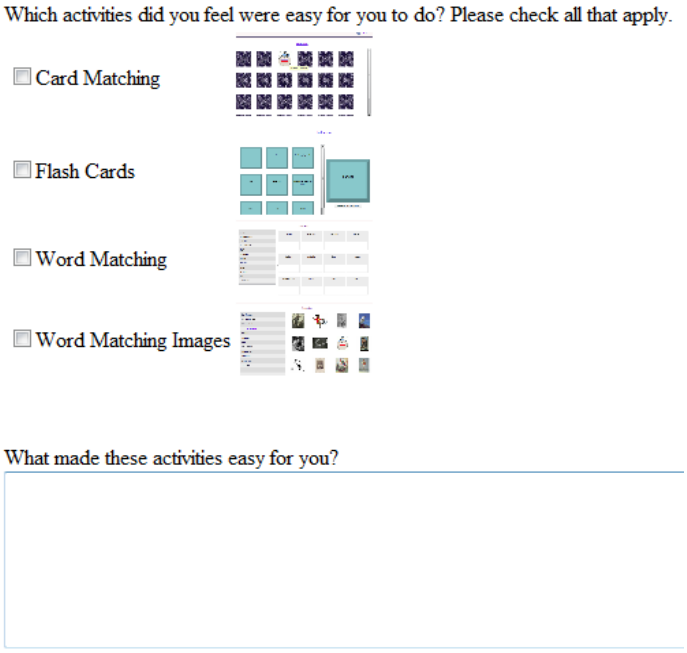

Figure C.4: A partial screenshot of the Activity Characteristics Survey.

OLM Section

This section asks questions about specific aspects of the OLM. The Likert-scale statements are in the below list, and an excerpt of them can be seen in Figure C.5.

1. I liked being able to see how I pronounced different sounds when I looked at the learner model.
2. I liked how the learner model showed which sounds I was good at pronouncing.
3. I liked how the learner model showed which sounds I should work on improving my pronunciation of.
4. I felt that the learner model accurately reflected my ability to pronounce Russian words.
5. I liked how the learner model showed how a native Russian speaker would say the same sentence.
6. I looked for differences between my model's pronunciation and the native speaker's model pronunciation.
7. I liked to compare my model to the native speaker's model.
8. I could see differences between how I said a sentence and how a native Russian speaker would say the same sentence.
9. I liked being able to see how the rest of the class was doing in comparison to me.
10. I would have liked to be able to see my classmates' models.
11. I chose activities based on the information presented to me in my learner model.
12. I followed ProTutor's activity recommendations.

The open ended questions that are part of this section of the survey are in the below list. They address issues related to activity recommendations and openness of the model.

1. How did your learner model affect which activities you chose to do?
2. Why did you follow or ignore ProTutor's activity recommendations?
3. If you followed the system's activity recommendation, did you find that they helped you?
4. How comfortable were you with seeing a model of your performance? Why?

5. How comfortable were you with the idea that others could see how they were performing in relation to the rest of the class, which includes you? Why?

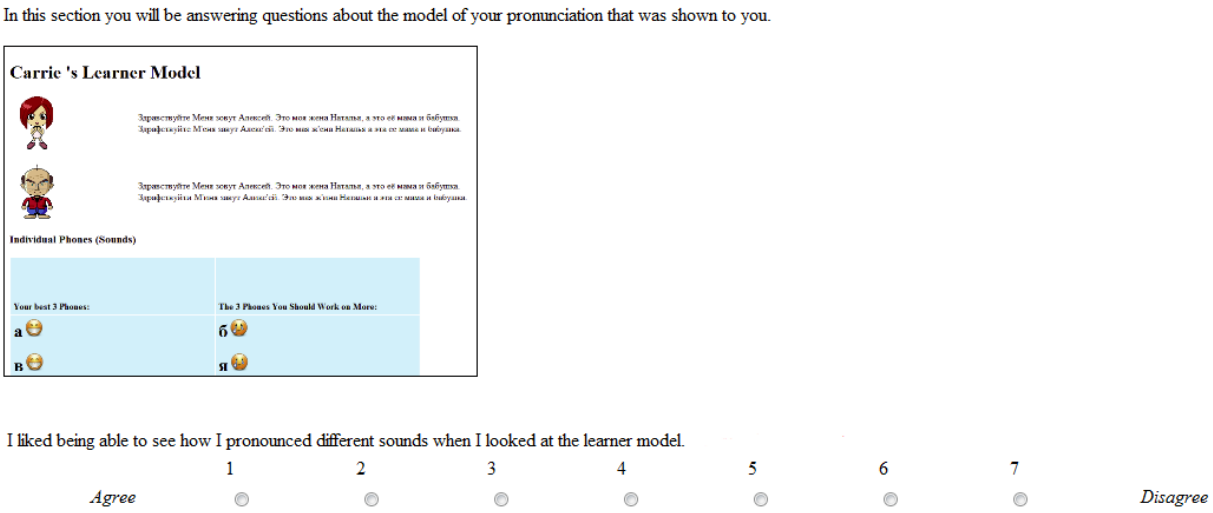

Figure C.5: A partial screenshot of the beginning of the Likert-scale portion of OLM survey.

PHASE 3 SURVEY

The phase 3 survey starts with the Activities Section, which is followed by the Activity Characteristics Section and then the Self-reporting Section. This is followed by the OLM section of the survey and finally, the HOLM Section and the General Impressions Section.

HOLM Section

This section collects information about the learner's perceptions of the HOLM. It consists of two parts: a Likert-scale portion and a set of open-ended questions. A partial screenshot of the Likert-scale portion of the survey is shown in Figure C.6. The statements that the learner is asked to agree or disagree with are in the list below.

1. I frequently looked at a previous version of my learner model.
2. I saw a difference between my old learner model and my current learner model.
3. I liked seeing how my pronunciation had changed.
4. I was motivated by seeing how my pronunciation had changed.

5. The model showed that my pronunciation had improved over time.
6. I found it useful to see how my pronunciation had changed.
7. Seeing both my old and new learner model changed which activities I chose to do.

The open ended portion of this section asked the following questions:

1. How did seeing a model of your previous performance motivate you?
2. How did seeing how your pronunciation was changing affect what activities you chose to do?
3. How did seeing how your pronunciation was changing affect your motivation to learn how to speak Russian?

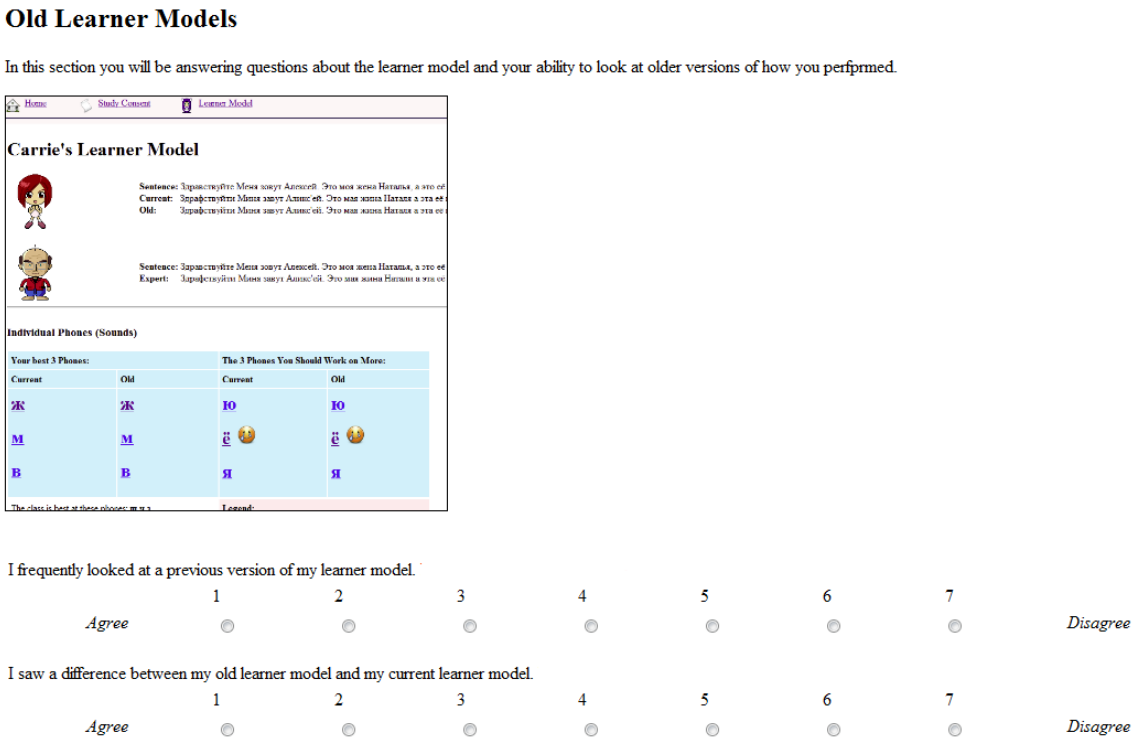

Figure C.6: A partial screenshot of the HOLM portion of the survey.

General Impressions Section

We collected information about the learner's general impressions of ProTutor using the following open-ended questions.

1. What did you like about the system?
2. What would you like to see improved?
3. What would you like to see removed?

APPENDIX D

SYSTEM DEVELOPMENT

ProTutor is a web-application that has been distributed over three servers: one streaming server (protutorstreamer), one application server (protutor), and one database server (protutordb). The protutor and protutorstreamer machines are running Windows Server 2008 and protutordb is running Microsoft SqlServer 2008 on a Windows XP machine. Protutorstreamer is only responsible for the streaming, recording, and storage of the learner's speech. Protutorstreamer uses Wowza Media Server Pro 1.5.0 to provide this service and protutor uses IIS 7 to serve the ProTutor .NET application.

We used Carnegie Mellon University's Sphinx-3 decoder for our Speech Recognition Engine because it is freely available and has been shown to work for multiple languages (Beck, Jia, & Mostow, 2003) (Nagamani, Narendra prasad, & Girija, 2005) (Carnegie Mellon University, 2007). The configuration of Sphinx-3 that we are using has a trigram based language model that is used in combination with an acoustic model in order to perform the decoding aspects of speech recognition. To train the decoder we used VoxForge's Russian corpus and language model, which has reported word error rates of 28.195% (VoxForge Russian Speech Corpus, 2007). While this was less than ideal, it was the best corpus that we could find to successfully train Sphinx-3.

ProTutor was built and its parts were integrated using Python, XML, ASP.NET, TSQL, C#.NET, HTML, and Flex. Each of ProTutor's components was tested before it was integrated with the larger system. Following integration of all of ProTutor's parts, a team of volunteers helped manually test ProTutor prior to its release. They also helped re-test ProTutor whenever functionality was added or bugs were fixed. Those testers were:

- Wojtek Quibell
- Clayton Epp

- Adam Mckenzie
- Dmytro Dyachuk
- Mayya Sharipova
- Veronika Makarova
- Julita Vassileva
- Greg Logan