GoPutonghua: An Online Learning Platform for Self-learners to Learn Putonghua

Vanessa Sin-Chun Ng, Andrew Kwok-Fai Lui, and Fu-Hong Wong

School of Science and Technology
The Open University of Hong Kong
30 Good Shepherd Street, Homantin, Hong Kong
scng@ouhk.edu.hk

Abstract. Internet has become a part of our life. Online learning offers us convenience, flexibility and collaboration. In this paper, a new online learning platform – GoPutonghua is introduced for self-learners to learn Putonghua. The online platform includes components like learning materials, online exercises, online assessment, learning tools and communications tools. The platform allows learners to conduct online exercises that are related to the lecture contents in e-books and online lectures. An online assessment system is developed for learners to know their performance and seek for improvement. A character dictionary (zidian) and a phrase dictionary (cidian) are built to support mapping from Chinese characters to Pinyin. Learning tools like Pinyin converter and Text-to-Pinyin Speech have been developed to search for the Pinyin (with tone marks) of Chinese characters and produce the correct pronunciations in Putonghua. The two types of communication tools (asynchronous and synchronous) are also included in the platform. GoPutonghua is developed to provide more resources and services for self-learners to learn Putonghua in a hassle-free manner. This paper explains why there is a need of developing the new online platform. The components and the architecture of the platform will be introduced. The implementation details of some components will be discussed in this paper. Finally, an evaluation on the usefulness and the quality of the online platform will also be reported.

1 Introduction

Learning a new language is not an easy task. Learners should expose to the environment with a lot of exercises, practices and training. E-learning systems promote lifelong learning by enabling learners to learn anytime, anywhere and at the learners' pace [1, 2]. Online learning has no time zones, and location and distance are not an issue. Learners can use the Internet to access the learning materials through the online learning platform, and can communicate with the cyber teachers. Situated learning is facilitated, since learners can learn while they are working on the job or in any environment at their own pace.

J. Yong (2005) [3] defines four main functions for e-learning systems based on four participants – lecturers, learners, administrative personnel, and technical staff. In this case, the e-learning system is sub-classified as teaching workflow system,

R. Kwan et al. (Eds.): ICT 2011, CCIS 177, pp. 96–108, 2011.

learning workflow system, administrative workflow system, and infrastructure workflow system. For example, in teaching workflow system, the teaching activities include assessment, materials preparation, student learning services and support, etc. In learning workflow system, the learning activities include assignment, discussion, evaluation, examination, etc.

Based on the principles of e-learning system, an online learning platform namely "GoPutonghua Online Self-Learning Platform" has been developed for self-learners to learn Putonghua. This online learning platform allows learners to conduct online exercises which are related to the lecture contents. An online assessment system is developed for learners to know their performance and seek for improvement. A character dictionary (zidian) and a phrase dictionary (cidian) are built to support mapping from Chinese characters (Hanzi) to Pinyin. Learning tools like Pinyin converter and Text-To-Pinyin Speech has been developed to search for the Pinyin (with tone marks) of Chinese characters and produce with the correct pronunciations in Putonghua. Online lectures and e-books are also included in the platform.

As for the online assessment system, it provides the learners with a reference on how they are acquiring their knowledge and skills, and the topics/areas where they need to seek clarification or invest more effort [4, 5]. Online assessment system has the advantages of instant feedback to students, greater flexibility with respect to location and timing, improved reliability, and enhanced question styles which incorporate interactivity and multimedia. According to R Trillo, et. al (2008) [6], an online assessment system should include the support of automatic corrections, different roles, different types of questions, random generation of tests and questions, etc. The online assessment system of "GoPutonghua Online Self-Learning Platform" allows teachers to set questions, and allows learners to view their results. It supports automatic marking and corrections of online tests.

There are two types of communication tools – asynchronous and synchronous. Synchronous tools allow collaboration and discussion among a group of people at the same time. Asynchronous tools allow communication and collaboration among people at different time. [7, 8] For the synchronous tool, the online platform uses Windows Live Messenger to allow learners to have instant messaging with cyber teachers. For the asynchronous tool, a forum is used for asking and answering questions, and to initiate discussions on topics related to Putonghua. It foments interactivity among learners and opportunities for learners to lead discussions.

2 Existing Online Putonghua Platforms

Two existing Putonghua online learning platforms are available in the market. They are "PLW" CUHK Online Putonghua Platform [9] and TEENS platform [10]. The following table shows a comparison of the two existing platforms and the new one - "GoPutonghua Online Self-Learning Platform". It can be seen that the TEENS platform has learning tools such as Text-To-Pinyin Speech and Pinyin dictionary that "PLW" platform does not have. It is also observed that both of the two existing platforms do not provide communication tools like instant messaging and discussion forum. However, they are important learning tools as learners may encounter problems while learning. The provision of communication tools can facilitate the

Table 1. Comparison of "GoPutonghua Online Self-Learning Platform" with existing platforms

Functions / Features		"PLW" CUHK Online Putonghua Platform	TEENS Platform	GoPutonghua Online Self-Learning Platform
Learning materials		✔	✔	✔
Lab sessions / exercises		✔	✔	✔
Online assessment		✔	✔	✔
Learning tools	TTS	✘	✔	✔
	Pinyin dictionary	✘	✔	✔
Communication Tools	Instant messaging	✘	✘	✔
	Discussion forum	✘		✔

interaction between teachers and learners. In short, the new online learning platform includes the learning materials, lab exercises, online assessment, learning tools and communication tools so as to provide more resources and services for self-learners to learn Putonghua in an easy and convenient way.

3 System Architecture

The architecture of this online learning platform can be visualized by Figure 1.

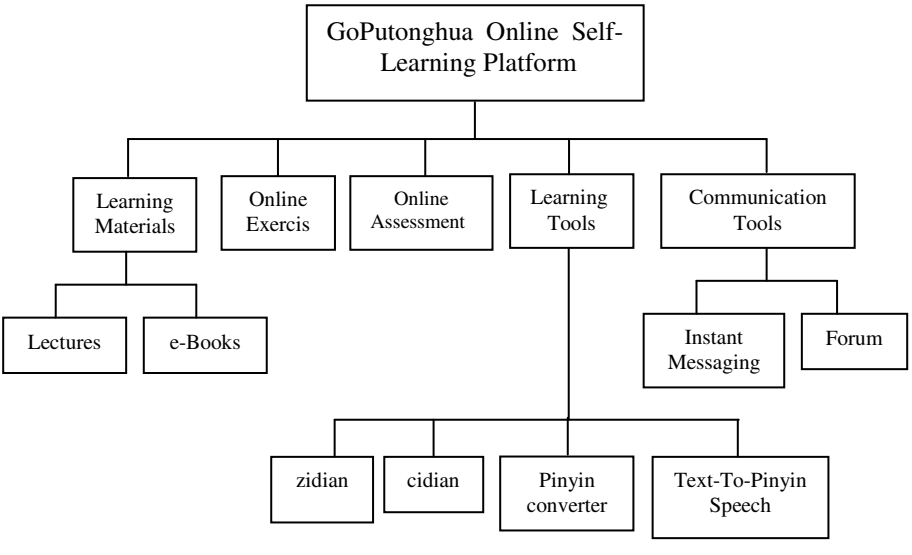

Fig. 1. The platform architecture

4 Methodology

The techniques used in the platform include HTML, CSS, PHP, MySQL. HTML is used to create web pages. CSS is used to define layout of HTML documents. PHP is used to create dynamic web pages and support database. MySQL is used to build a database system used on the platform.

The online platform involves online assessment system, learning tools, learning materials and communications tools.

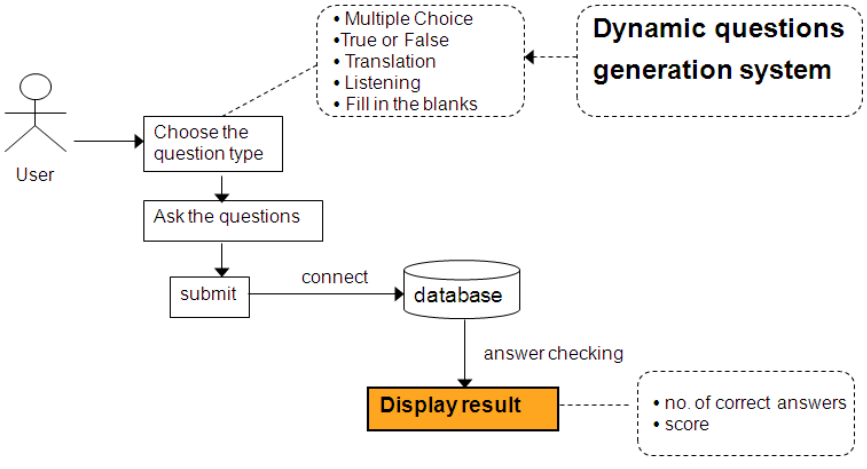

Fig. 2. The implementation of Online Assessment System

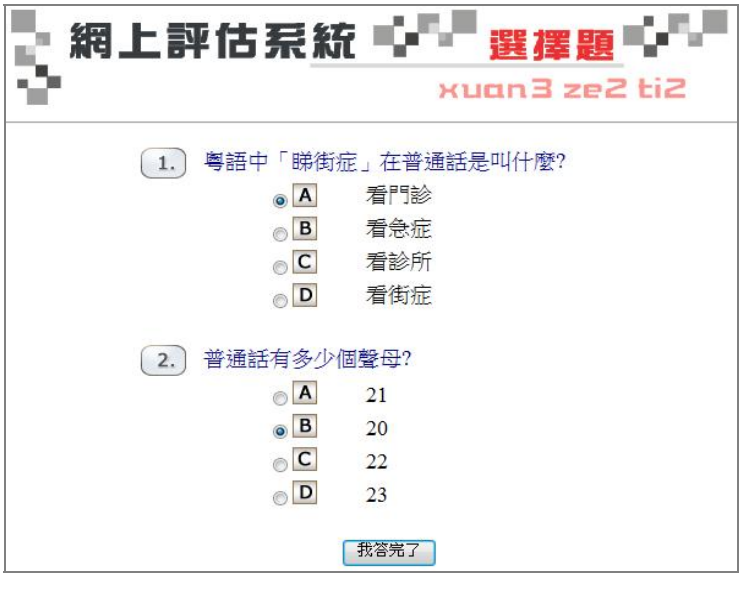

Fig. 3. Multiple Choice Question: before submission

4.1 Online Assessment System (OAS)

The OAS allows cyber teachers to set questions which will be stored in the question database. After submitting the answers, the system will check the answers and calculate the marks through question database. Then the marks will be stored in the marks database. The implementation of the OAS is shown in Fig. 2. Fig. 3 and 4 show the screen shots of different types of questions in OAS.

Fig. 4. Pinyin Translation Question

4.2 Learning Tools

There are four learning tools implemented in this system. They are character dictionary (zidian), phrase dictionary (cidian), Pinyin converter and Text-to-Pinyin Speech. Fig. 5 shows the implementation of a learning tool – character dictionary.

For character dictionary, when the user enters a Chinese character in the textbox and presses the submit button, a connection to the database is established. It will first go through the "zidian" table to search the desired word. If the word is in the database, then the system will find the audio mp3 file for that word. This can be achieved by using the tone number of that word. For example, the word "好" has two pronunciations --- "hao3" or "hao4". The system will go to find the correct audio file for each pronunciation and pass the mp3 file address to the Flash MP3 player. When the user clicks the player, it will play the pronunciation.

When the audio files were found, the system will move to "cidian" table to find all phrases containing the searched word. We can use the Pinyin with tone number of the searched word to find those matched words (obtained previously) having the same tone number. Then, it randomly generates the correct vocabulary. Fig. 6 shows the searching result of the Chinese character "好".

For Text-To-Pinyin Speech, there will be a voice folder containing the pronunciation of Chinese characters. Each time when the learners type in the Chinese characters, the system will look for the Pinyin from the Pinyin database and displays the Pinyin

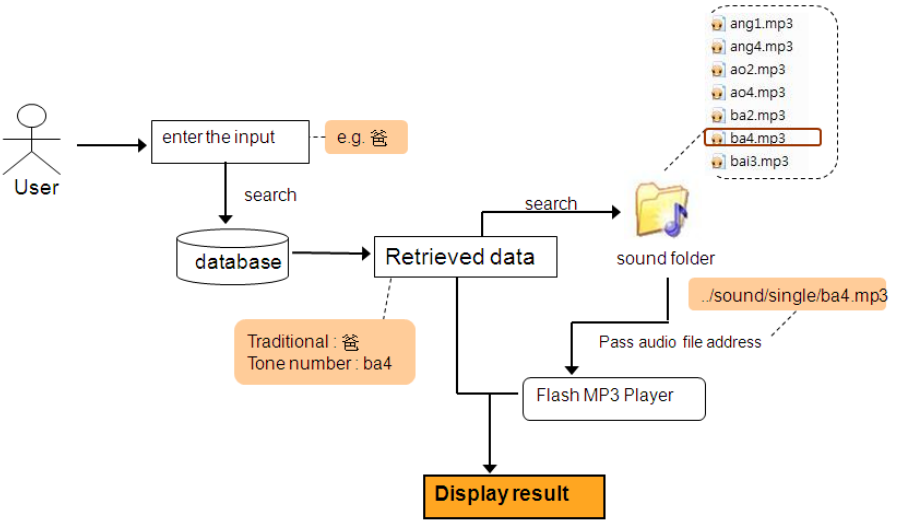

Fig. 5. The implementation of a learning tool – character dictionary (zidian)

Fig. 6. The search result of the word "好" in character dictionary

with correct tone marks, then it picks up the correct mp3 sound file to output the pronunciation.

In order to search the Pinyin and the audio mp3 file of each Chinese word, the inputted content will be first split into an array. For example, the content "今天, 我去了爬山。" will be split as follows:

今	天	,	我	去	了	爬	山	。

After splitting, it needs to check whether the elements in the array are Chinese characters or not. If the element is a single Chinese character, the system will go

through the "zidian" table to search the pinyin with tone mark of that character and output the Pinyin with tone mark and the character. If two or more elements are phrases after combination, the system will search another table called "cidian". After that, it will use the tone number of character /phrase to find the corresponding audio mp3 file. When all the addresses of audio files are collected, they will be sent to the Flash MP3 player. The player will play the pronunciation of the whole content. Fig. 7 shows the result of text to pinyin with spoken speech.

Fig. 7. The result of Text-To-Pinyin Speech

Fig. 8. A screen shot for e-book

4.3 Learning Materials

There are two types of learning materials available in the new online learning platform – e-book and lecture notes. The e-book contains a page turning effect when

the user flips to the next page. The contents in the e-book are mainly in bulletin point format which makes users easy to read and memorize the contents. There is also zoom-in effect in e-book. Fig. 8 shows a screen shot of the e-book. The online lecture notes contain not only text, but also include multimedia elements. Most of the lecture notes include audio effect. When the user clicks the "Play" button, he or she can listen to the pronunciation of the words through the flash mp3 player. Fig. 9 shows a screen shot of the lecture notes.

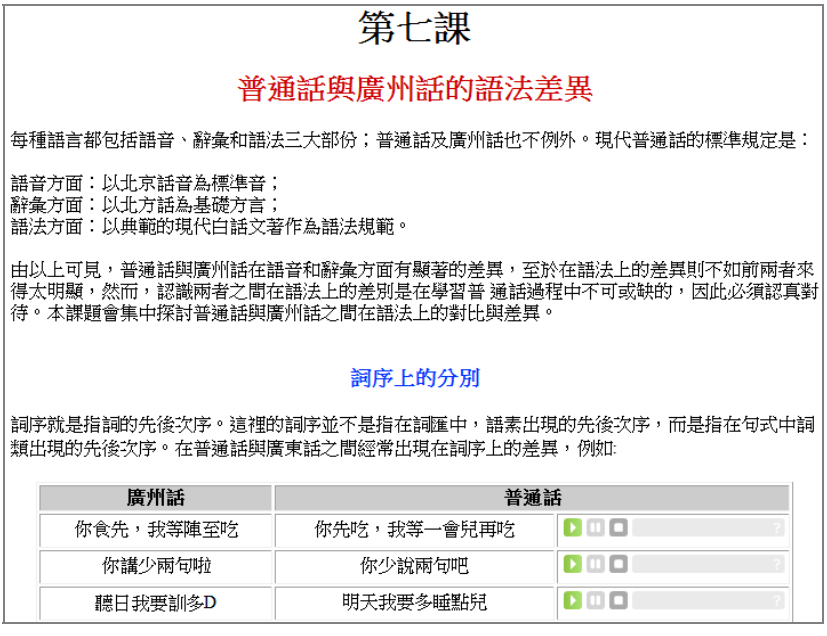

Fig. 9. A screen shot for lecture notes

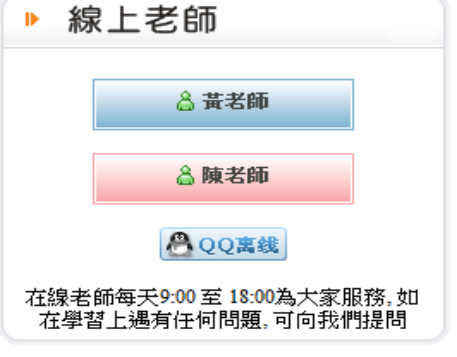

Fig. 10. Instant messaging

4.4 Communication Tools

Instant messaging and discussion forum are provided in communication tools. The new online platform uses Window Live Messenger to allow learners to have instant messaging with cyber teachers. Fig 10 shows a screenshot of instant messaging in GoPutonghua Self-learning Online Platform. It foments interactivity among learners and the cyber teachers. The learners can solve the questions quickly with the help of cyber teachers.

5 Discussion

In Mandarin, there are some rules of pronunciation. The tone of a word can be changed by the word placement. For example, the word "診所" (clinic) represents two characters with third tone distinction. If you pronounce both two characters in third tone, it is quite unnatural and slow in speed. In fact, the tone of the first word "診 "should be pronounced in second tone. However, this change will not affect its meaning. Note that, the examples listed below that has changed the tone marks are just used to show the correct tone when speaking. When writing Pinyin, although there is a change when speaking, the original tone marks are still written.

Rule 1 (Third Tone)
If a 3rd tone is followed by another 3rd tone, the first 3rd tone changes to a 2nd tone.

Examples:

你好 (hello): nǐ hǎo ➔ ní hǎo
保守 (conservative) : bǎo shǒu ➔ báo shǒu
好久 (a long time): hǎo jiǔ ➔ háo jiǔ

Rule 2 (The character "一")
The Chinese character "一" is pronounced in the first tone when it stands alone. But "一" is pronounced in 2nd tone when followed by a 4th tone.

Examples:

一個 (one): yī gè ➔ yí gè
一塊 (one piece): yī kuài ➔ yí kuài
一步 (one step): yī bù ➔ yí bù

Rule 3 (The character "不")
The Chinese character "不" is pronounced in the 4th tone when it stands alone but there is one situation where this changes. "不" is pronounced in 2nd tone when followed by a 4th tone.

Examples:

不變 (not change): bù biàn ➔ bú biàn
不會 (will not, cannot): bù huì ➔ bú huì
不錯 (not bad): bù cuò ➔ bú cuò

If we don't apply the above rules, the content entered by the user will be displayed one character by one character because it only searches the "character database". Also, the flash mp3 player will pronounce one character by one character which is slow in speed and not very natural. In order to make it more natural and increase the speed of pronunciation, one more database will be used that is "phrase dictionary". In addition, the n-gram algorithm will be applied to do the searching.

An n-gram is a sub-sequence of n items from a given sequence. N-grams are used in various areas of statistical natural language processing and genetic sequence analysis. The items in question can be characters, words or base pairs according to the application. For example, if the sequence of Chinese characters "研究生活動中心" uses a 2-gram(bigram) to do word segmentation, the results will be "研究", "究生", "生活", "活動", "動中", "中心". N-gram is determined by the size of N. So it will have unigram (n=1), bigram (n=2), trigram (n=3) and so on. The N-gram algorithm used for Text-To-Pinyin Speech is shown in the Fig. 11. Once a phrase is found in the cidian, the corresponding phrase mp3 file can be found from database, it makes the sound to be more natural.

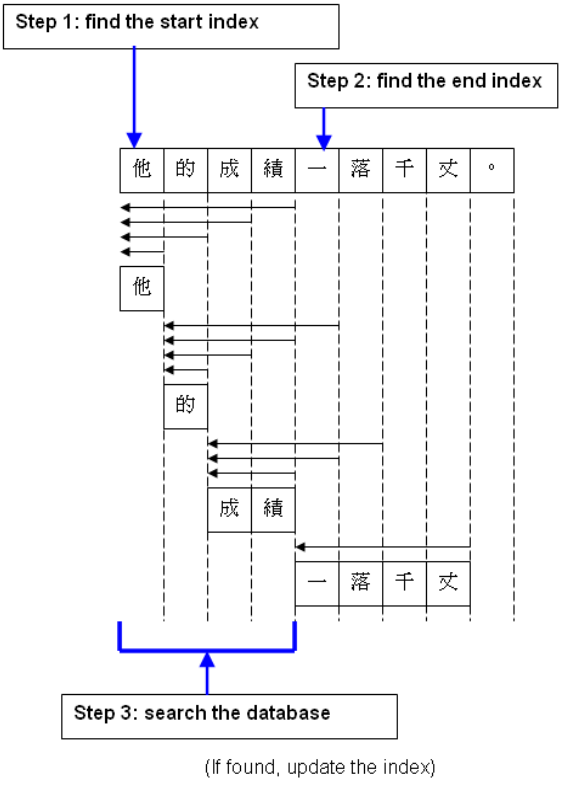

Fig. 11. The search result for using the N-gram algorithm

6 Evaluation of GoPutonghua Online Self-learning Platform

A questionnaire was designed to obtain views and opinions of the GoPutonghua Online Self-Learning Platform. The aim was to identify any areas that could be improved or changed in order to provide better services and resources to the self-learners. Respondents were also asked to rate various components of the platform based on two criteria: usefulness and quality. The total number of people invited to try the platform and fill in the questionnaires was 14, of which the majority of respondents were university students (57.14%), 14.29% were Putonghua teachers, 28.57% were working people.

The respondents were asked to rate the usefulness and quality of each of the components provided in the platform. Components that are high quality doesn't mean that are very useful, and components that are very useful doesn't mean that they are of high quality. Useful components are components that have a valuable purpose and are frequently utilized in learning. Quality components are components that have good graphic design, being free from defects, functioning properly and achieving high end-user satisfaction. Generally speaking, good components should achieve high usefulness and high quality. The results listed in Table 2 show that majority of components is of high usefulness which are over 80%. The most useful component is Text-To-Pinyin Speech (89.29%) while the lowest one is Pinyin Converter (69.64%). Concerning the quality of the platform, majority of components have values nearest to high which is over 70%. The highest quality component is e-Book (78.57%) while the lowest one is Pinyin Converter (60.71%). The result also shows that Text-To-Pinyin Speech has a great difference (21.43%) between its usefulness and quality.

Table 2. Comparison of usefulness and quality of components

	Usefulness	**Quality**
Lecture Notes	80.36%	75%
e-Book	80.36%	***78.57%***
Online Exercise	78.57%	73.21%
Character & Phrase Dictionary	87.50%	73.21%
Pinyin Converter	***69.64%***	***60.71%***
Text-To-Pinyin Speech	***89.29%***	67.86%
Multiple Choice	80.36%	71.43%
True or false	82.14%	66.07%
Listening	83.93%	73.21%
Translation	80.36%	69.64%
Fill in the blanks	78.57%	69.64%
Windows Live Messenger	83.93%	71.43%
Forum	82.14%	75%

Respondents were asked if they had any additional comments or suggestions regarding the platform components, in particular any additional components they would like to see in the platform. The number of responses to this question was 11 and the results are as follows:

1. *"The lecture notes could be more colorful. It is better not to use flash for the e-Book"*
2. *"It is better to press "enter" key to login in instead of pressing the "submit" button"*
3. *"There should an improvement in the quality of pictures.*
4. *"E-book can insert sound and video to enhance its function and effect.*
5. *"For true or false question, it is better to include explanation after submitting the answer. It is better for user to login to the forum without registering it again.*
6. *"It is better to have speaking test for the online assessment system.*
7. *"It is better to have speaking test in the online assessment system.*
8. *"Every word should have sound."*
9. *"It is suggested that to learn Putonghua from life, for example, the system can retrieve a news everyday from the Internet and provide pronunciation and Pinyin in order to attract users to learn Putonghua"*
10. *"E-book can insert sound"*
11. *"Navigation function of the forum is not obvious; there is no button to link to the index page. If the platform is put on the internet, the smoothness and the browsing speed will be affected by bandwidth"*

Respondents thought that users can use the "enter" key to login rather than pressing the "submit" button and the forum's navigation function is not clear. These can be improved in a short period of time so as to provide convenience to the users. Another comment was that the browsing speed may be affected by bandwidth. This can be solved by using the web hosting service as it is more reliable and faster. Other comments suggested inserting sound to e-Book and including speaking tests for the Online Assessment. These were seen as in need of expansion. Learning Putonghua from life through daily news is a new component to the platform. This was seen as in need of expansion and renovation.

Overall, respondents tended to rate the usefulness and quality of the platform as "high". There is also a fairly high percentage of respondents rate it as "very high". There were almost no respondents that rated it as "very low" and "low". To summarize, the results clearly identify that the overall usefulness (81.04%) is higher than the quality (72.1%). In order to minimize the difference between them, the quality of the platform was needed to be improved.

7 Conclusion

In this paper, an online learning platform "GoPutonghua" is proposed to allow self-learners to learn Putonghua online. The online platform contains components like learning materials, online exercises, online assessment, learning tools and communications tools. As compared to the existing online Putonghua platforms, the new learning platform includes more comprehensive functions which address the needs of learners. In order to produce more natural speech for the Text-To-Pinyin speech in the "Learning tools" of the platform, an N-gram algorithm is used to find phrases first in the dictionary. The evaluation results show that the users rated the overall usefulness as 81.04%; while the quality of the platform as 72.1%.

References

[1] Communication Tools, the Digital Education website, http://blogs.ubc.ca/clarkeetec565/communication-tools/
[2] E-assessment, http://en.wikipedia.org/wiki/E-assessment
[3] Yong, J.: Workflow-based e-Learning Platform. In: Ninth International Conference on Computer Supported Cooperative Work in Design, vol. 2 (2005)
[4] Ridgway, J., McCusker, S., Pead, D.: Literature Review of e-Assessment. Nesta Future Lab, Bristol, UK (2004)
[5] Rourke, L., Anderson, T.: Using peer teams to lead online discussions. Journal of Interactive Media en Education (2002), http://www-jime.open.ac.uk/2002/1/rourke-anderson-02-1.pdf
[6] TEENS Platform, http://teens.putonghuaonline.com/
[7] Trillo, R., Ilarri, S., López, J.R., Brisaboa, N.R.: Development of An Online Assessment System to Track the Performance of Students (2008), http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.7210&rep=rep1&type=pdf
[8] Urdan, T.A., Weggen, C.C.: Towards Best Practices in Online Learning and Teaching in Higher Education. MERLOT Journal of Online Learning and Teaching 6 (2000)
[9] Le, W.P.: The Chinese University of Hong Kong, http://www.ilc.cuhk.edu.hk/Chinese/plw/
[10] What is an E-Learning Platform, Anyway?, Australian Flexible Learning Community website, http://community.flexiblelearning.net.au/TechnologiesforLearning/content/article_442.htm