博士論文

Instructional Design and Effects of Mandarin Tone Learning Tool for Acquiring Pronunciation Ability

（中国語声調発音能力を向上させるための
教材デザインと効果について）

2021 年 3 月

康 茗淞

宇都宮大学大学院工学研究科
システム創成工学専攻

内容梗概

第二言語学習において，大人は子供と比べ，新しい言語に関する知識を記憶する能力や，知識を統合し運用するための学習の戦略，経験，知能などでは優れていると思われるが，相違点が多い「発音」は子供のうちに習得しないと，母語話者レベルになることがほぼ不可能と言われている．日本語と中国語は共に漢字を利用するものの，その発音は大きく異なり，それが，日本人が中国語を学習する際の大きな障壁となっている．中国語学習教材は多数存在するが，それらを使って，日本人が正しい発音を身につけるのは容易ではない．そこで，我々は，音声の認識能力の向上に着目し，中国語の発音を効率よく学ぶことができる教材の開発に着手した．

中国語発音学習の一般的な学習方法では，簡単な単語を覚えること，モデル音声を模倣することが行われるが，日本人による中国語発音学習には次のような三つの課題がある；(1)中国語には日本語にない発音，ならびに声調要素があること，(2)これらの習得速度には個人差があるため教室での一斉授業は生徒全員の発音を正すことが困難であること，(3)中国語初心者は，発音練習の中で自分の発音の誤りに気づき修正することが難しいことである．しかし，中国語の意味伝達において重要な「声調」という要素の訓練は疎かにされがちである．このため，不正確な発音が身に付いてしまい，学習の進捗が滞ることが多いと思われる．

そこで本研究では，まず，声調習得のための基本能力である聴覚の認知能力向上を目指して，中国語に初めて触れる日本語母語話者の，中国語声調に対する感知状態を調査した．その結果，ほとんどの中国語初心者は，四音節の中国語声調の聞き分けができないことが分かった．しかし，四音節声調課題に次の三つのルール(1)四音節に必ず四種類の声調を含めること，(2)一問に使用する音素は一種類であること，(3)スピードを大幅に落とすこと（一音節あたり２秒），を追加することで，声調のカテゴリー化知覚を促進できることが明らかとなった．

そこで次に，インプット仮説と認知負荷理論を重視した感知状態調査実験をもとに，声調のカテゴリー化知覚の促進を目指して，声調の聞き分けを行う多種の音声課題を含む感知訓練問題集を作成した．この問題集の設計においては，異なる声調を強調した多音節感知問題を中心に，一番易しい問題から，声調の並び，音域，音源或いは音素などの変化を少しずつ加えて，適切な認知負荷を学習者に与えることをコンセプトとした．設計した声調感知問題集の有効性を評価するため，感知問題集と声調の特徴を示す映像説明を教育コンテンツにしたアプリを作成し，既存の音声分析機能付きの発音練習アプリと比較するユーザー実験を行った．中国語に初めて触れる者に対して声調に関する知識を与えた後に，実験参加者を，既存の中国語学習アプリを使用するグループと我々が開発したアプリを使用するグループの２つに分け，それぞれのアプリを用いて３

時間の学習を行った後，発音の聞き取りテストを行い，効果を測定した．その結果，我々が開発した感知訓練アプリを使用した参加者の平均声調正解率は，二つの基本発音課題において，既存の発音練習アプリと比較してそれぞれ 20%および 16%向上した．しかし，声調の感知にはかなりの個人差が見られ，今回の感知課題では上手く声調を認識できず，学習効果が表れなかった参加者も存在した．

そこで，さらなる声調学習の効率向上や学習時の負担軽減と学習のモチベーションの向上を目指して，認知負荷理論を応用し，学習内容自体と関係する内在性認知負荷（学習のコンテンツ）はそのままに，操作の手間など，感知訓練アプリ使用時に発生する外在性（余計な）認知負荷を最小限に抑えてアプリのユーザーインターフェース(UI)の改良を行うとともに，訓練方法の見直しも行った．また，改良前の感知訓練アプリを用いた実験では，アプリ使用前に発音聞き取りテストを予習することで，アプリ使用後のテスト成績が約 20%上昇したことから，発音の聞き取り練習を感知訓練アプリに組み込むことによる学習効果向上も試みた．

改良により，感知訓練アプリにおける感知問題の回答操作と画面遷移の回数を，先述のプロトタイプの半分に削減した．さらに画面遷移ごとのローディング時間を無くし，訓練時に必要な操作と映像説明などの声調情報を獲得する手間を最低限にした．加えて，アプリ起動後に発音聞き取り練習画面が表示される仕組みを作成し，感知訓練の前にテストの簡単な予習を行うことが学習効果へ与える影響も調査した．その結果，UI 改良後の感知訓練アプリを使用した参加者の平均声調正解率は，二つの基本発音課題において， プロトタイプよりも 10%および 20%向上し，フレーズの声調発音正解率も 13%上昇した．しかし，感知訓練の前に発音聞き取りテストの予習を組み込むことによる学習効果の向上は本実験では見られなかった．

以上のことから，インプット仮説と認知負荷理論に基づいて学習教材を設計し，外在性負荷を抑えて UI を実装した感知訓練アプリは，声調学習に有効であると結論付けられる．また，予習，学習，テストという流れの効果については今後の検証が必要である．

今後の課題として，自主的な発音練習時に生じた声調の間違いを，学習者にフィードバックするための情報提供法の検討が挙げられる．その一つとして，発音訓練に特化した機械学習モデルを構築し，話者の声調認識を行い，発音学習をサポートする方法が考えられる．さらに，学習モデルの予測結果に加えて，機械学習の学習過程や用いたデータの可視化情報をフィードバックとして提示するなども考えられる．アプリ使用者の発音誤り傾向や苦手な声調パターンや音素の解消に向けた発音課題の設計は，中国語発音教育における大きな課題である。

Instructional Design and Effects of Mandarin Tone Learning Tool for Acquiring Pronunciation Ability

MING SUNG KANG

ABSTRACT

In the 2nd language learning, adults are superior in their ability to memorize knowledge of new languages and have better learning strategies, experience, and intelligence to integrate knowledge. However, unless one learns pronunciation in childhood, it is almost impossible to reach a native-level accent, especially foreign languages. Although both of Japanese and Chinese use Chinese character in their language, their pronunciation differs completely. This is a major problem when Japanese learn Chinese. Although there are many Chinese teaching materials, it is still not easy for Japanese to acquire correct pronunciation. In this thesis, we aim to develop a learning method and a tool of Chinese pronunciation for learners of Chinese as a foreign language.

The general method for learning Chinese tone is focused on performing exercises of simple conversation. However, there are three challenges in learning Chinese pronunciation by Japanese: (1) Part of elements of pronunciation that does not exist in Japanese, such as tones. (2) There is a large individual difference in tone learning speed, which places a heavy burden on educators. (3) There are learners who practicing pronunciation of words, phrases or conversation even if they cannot recognize it. Furthermore, a lack of guidance and learners not being conscious of tones cause fossilized errors and learning stagnation.

To implement the learning method, we carried out the perceptual experiments to investigate the tone-understanding ability of university students who does not have Chinese learning experience. The experiment results showed that the students could not initially distinguish tone correctly. To promote the categorization perception of tones of learners, we designed the four-syllable tone perception questions using the following three rules: (1) One question includes all four types of tone. (2) Only one pronunciation used for each question. (3) Significantly slow down audio speed (2 seconds per syllable). These rules help students identify pitch changes in tone.

Referring to the input hypothesis and the results of perceptual experiments, we made a collection of the perceptual training questions, composed of various training tasks. These tasks led learners to start from the simplest perception question and gradually change the pitch ranges, sound phoneme, or speaker as the germane cognitive load. Furthermore, showing the answer status of each question and task as learning

feedback, the users could improve their perceptual ability and select perceptual training tasks that matched their abilities according to the feedback.

To verify the effect of the perceptual training with tasks designed, we compared tone pronunciation abilities after using our perceptual training application with an imitation practice application with voice analysis function. After use of the application for a total of 3 hours, users of the perceptual training application that focuses on cognitive ability of tone had an average correct answer rate 20% and 16% higher in two basic tone pronunciation tasks, respectively. However, though the perceptual training app is more effective for improving the tone pronunciation ability compared to existing apps, there were users who could not recognize the tone and had a low correct answer rate in the pronunciation test.

To improve the learning effect of the perceptual training application, we optimized the interface design, operation design, and loading time, which aimed to minimize the extraneous cognitive load without changing the intrinsic cognitive load (contents of learning). Furthermore, according to the experiment conducted using the prototype of the perceptual training app, the pronunciation test after the app was used increased about 20% by preparing the pronunciation test before using the app. Therefore, we also tried to improve the learning effect by incorporating pronunciation practice into the perceptual training application.

According to the tone pronunciation correct answer rates of users after using the application, the implementation of the tone pronunciation practice function did not help perceptual training app users to improve their tone pronunciation ability. However, we considered the shortness of sensory memory and optimized the interface and operation design of the perceptual training application, compared with the prototype. As a result, the tone pronunciation correct answer rate of users was 10% and 20% higher in two basic pronunciation tasks and 13% higher in the phrase pronunciation task after using for a total of 3 hours. The experimental results show the learning effectiveness and potential of perceptual training. We considered that this research can provide an empirical foundation for improving Chinese tone ability in the future.

As a future task, we will study on the feedback method of pronunciation errors. Practicing pronunciation could be the cause of tonal pronunciation errors. Therefore, we considered using machine learning to support tone pronunciation practice. We aimed to use machine learning to design pronunciation tasks rather than recognize tones correctly, as is the case in existing research. Furthermore, in addition to the prediction results of the trained model, the percentage of classification, the training data set, and the visualized data should be presented to the learner as feedback for learning. It is also necessary to consider the instructional design of pronunciation practice based on difficulty of tone combination, error patterns, phonemes, etc. Classifying learners automatically according to their phonological ability and providing navigation considering their individual differences using machine learning are also future issues.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Language learning may be faster and easier for younger people, and it seems to be due to the differences in effort to acquire native and second languages. However, there is a significant difference in language cognition between native language and second language learning, such as the learning environment and knowledge of the learner.

According to related works [1], [2], language learning speed of adolescents and adults in the same learning environment is higher than that of children and younger children, who increase their language learning ability as their age increases. However, with the speed of language learning, pronunciation is different from learning words and grammar as well as constructing and translating documents [2]. Though older people have more knowledge, learning experience, and memory strategies, it is almost impossible for adult learners of a second language to pronounce with a native-level accent. The existing research was mainly focused on achieving “native-like” fluency or accent in foreign languages [3].

Second language learning refers to interlanguage theory. A systematic error that occurs during language learning is mainly due to negative transfer of the

native language. Learning to listen and pronounce a language is different from learning to read and write as they involve remembering knowledge (letters, phonetic transcription, words, grammar, phrasing, etc.). The environment in which only the native language is used and age (neurological constraints [3]) make it difficult to cognize elements of pronunciation in a foreign language (phonemes, mora, accent, syllable, tones, etc.) that are not in the learners' mother tongue. To learn the pronunciation of a foreign language, generally learners need to be able to cognize elements of pronunciation and combine them with knowledge of letters and notation.

1.1. Background of This Research

Studies on language production [4], [5] show that pronunciation of the second language is quite distinct from grammatical encoding, and another study on bilingual language processing [6] shows that bilingual speakers have a knowledge of two phonological systems and can activate both during language processing. However, the difficulty of pronunciation is that there are many learners who cannot pronounce or recognize the second language even if they have knowledge of it. The pronunciation errors of learners are caused by the native-language transfer or part of elements of pronunciation that do not exist in the learners' native language. Though it is important to teach pronunciation properly, it is difficult for all learners to cope with the difficulties and correct pronunciation errors in the second language class, which aims to make learners gain basic conversation skills in a short period of time (one or two semesters). In addition, teachers that instruct language learners may be non-native speakers, and it is possible that the teacher's own pronunciation accuracy and the ability to correct the learner's pronunciation are insufficient.

1.1.1. Studies about Mandarin Chinese Tone Pronunciation

There are several languages called tonal languages [7] that distinguish the meaning of spoken words based on pitch change patterns included in each syllable. It takes a great deal of effort and time for second language learners to recognize tone correctly and learn to pronounce the words with the right tone, especially for the non-tone native-language speakers. Furthermore, there is a large individual difference in tone learning speed, which places a heavy burden on educators who must help learners deal with tone learning difficulties and keep learners motivated.

We targeted Mandarin Chinese, which has four main tones, and investigated the study of Mandarin Chinese tone pronunciation. The four main tones used in Mandarin Chinese conversation are the flat-high tone (tone 1), medium-rising tone (tone 2), low-dipping tone (tone 3), and high-falling tone (tone 4) as shown in Figure 1-1 using Zhao's five-level tone mark [8]. Compared with languages (like Vietnamese) that have six or more tones, the features of Mandarin Chinese tones are more obvious and seems easy. However, acquisition of Mandarin Chinese tone is difficult for non-Mandarin Chinese speakers, and many studies on Mandarin Chinese tones and tone pronunciation have been conducted to solve this issue.

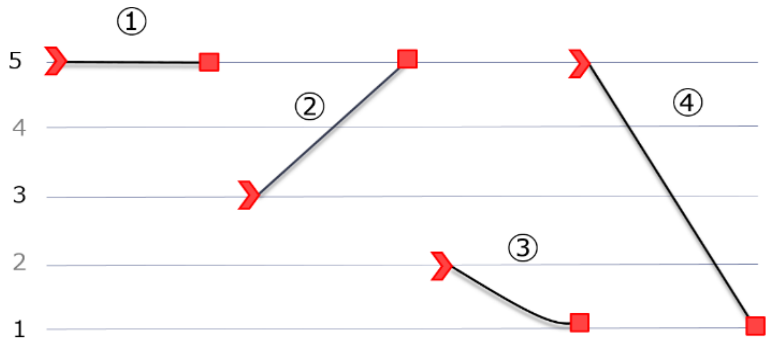


Figure 1-1. Four tones in Mandarin Chinese using five-level tone mark (the numbers on the left side show the relative pitch: 1. Bottom, 2. Low, 3. Middle, 4. High, 5. Top).

Speaking with a correct tone is crucial in Chinese conversation. In Chinese primary education, many expressions place importance on pronunciation and tone learning [9], [10]. However, their pronunciation varies widely and there are substantial individual differences in learning tone, which causes problems when learning Chinese. Furthermore, the task of tone acquisition is a burden on both the teacher and the learner. In Chinese classes and materials for quick raise the level of Chinese conversational skills that learners are most interested in and to avoid frustration, the training of tone tends to be neglected. A study on Mandarin Chinese tone perception has shown that even the advanced second-language learners who have sufficient knowledge of words have difficulty recognizing tones included in polysyllables and sentences [11].

The research on Chinese tone is mainly based on phonetic experiments and analyzing cause, error, or acoustic features, and there are few contents related to training methods. The results of the research are principally tone difficulty order, the error rate of combinations, and cause analysis results (negative transfer of native language, insufficient guidance, etc.). The remedial measures or opinions on tone learning are mainly provisions of model voice, pronunciation practice by imitation, to be conscious of tone, and performing exercises designed to focus on tone combination and error trends. In addition, the research direction of related works on the acquisition of tone in Japan and Chinese are addressed by Ding's research [12]. According to the research, the main idea in related works is designing phonetic experiments based on the tone error trend and identifying error patterns through experimental data analysis. However, there are few learning method contents in the related works, and research on a new learning method is needed. In many tone studies, even with learners who have studied Chinese for more than two years, the ability to pronounce Chinese tones is quite low. One of the studies analyzed the tone errors of Japanese learners in disyllabic

words [13]. It found that the tone error rate of disyllabic words of seven participants who studied Chinese for two to three years was an average of 35.7%. Another study analyzed the tone errors of polish speakers who have studied Chinese for more than sixty hours (half of them learn over two hundred hours) [14]. Their disyllabic tone pronunciation error rate was an average of 43.3%.

Figure 1-2 shows the results of classifying the research on tone learning [13]–[27] that we surveyed in Taiwan according to the research target and controlled variables.

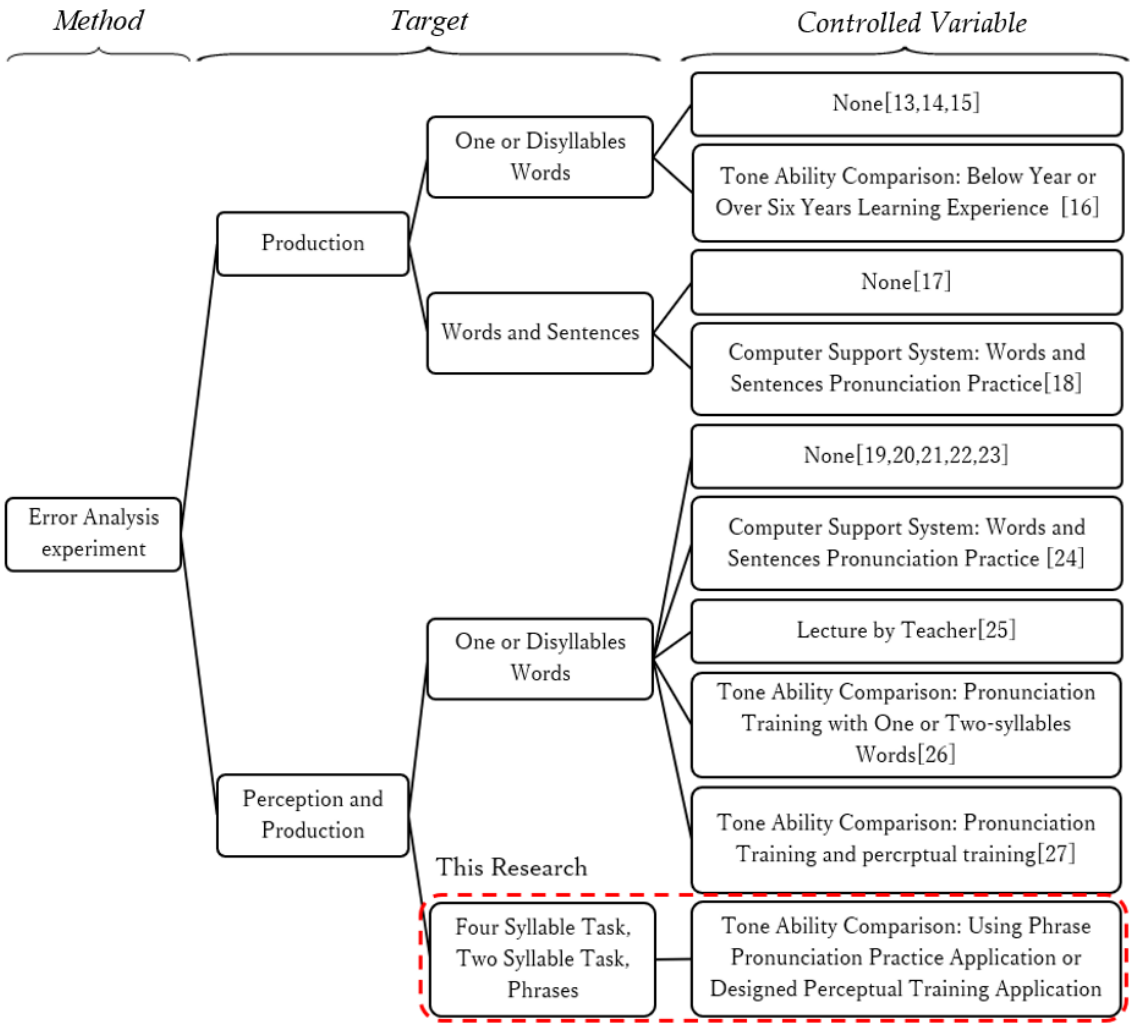


Figure 1-2. Classifying the research on tone learning [13]–[27] and the standpoint of this research.

There are several studies on Mandarin Chinese tonal pronunciation targeting learners whose native language is Vietnamese, which has more kinds of tones [15], [23], [24]. Vietnamese speakers, whose native language is a tone language, have the knowledge of tone language, and they master the Vietnamese tones that seems to be more difficult than Mandarin tones. However, due to the negative transfer of the mother tongue, the Vietnamese tones that learners are used to speaking may cause pronunciation errors. In addition, a study target on English speaking learners who have studied Chinese for more than six months and study abroad in Taiwan (70% of the participants learn Chinese over two years) [22]. As a result, the tone pronunciation error rate of the participants was 29% on average.

Though the majority of the research on pronunciation of Mandarin Chinese tones we investigated focused on error analysis and error corrections by phonetic experiments, several studies have examined the differences in performance due to controlled variables. A study by Korean learners [16] compared the tone pronunciation ability of six learners with a Chinese learning experience of less than one year and learning experience of more than six years. As a result, the error rate of learners with a learning experience of less than one year was 30.6% and 38.9% for learners with a learning experience of more than six years. In this study, participants with more learning experience show lower performance in tone pronunciation. And the study investigating the tone errors of Hong Kong students showed that with experience studying in Taiwan for six months, one year, two years and three years, the error rates of students were 38.5%, 28.8%, 16% and 12.1% on average [17], respectively. Another study by Vietnamese learners [23] show that the error rate of learners with a learning experience of less than one year was 40%, less than two years was 31%, more than four years was 17%.

Regarding tone learning, a lack of guidance and learners not being conscious of tones, especially fossilized errors due to neglect of tone learning in initial education, are the reasons that even advanced learners have such a high tone error rate. Some studies [18], [24] have compared abilities before and after using a computer support system that provides a voice model of Chinese phrases, which represents pronunciation scoring and pitch trajectory. Other studies have compared the ability of two years American learners before and after reinforcement training by teachers [25]. The results show that the average pronunciation and perception error rates of two-syllable words are 53.3% and 60.6%. And they reduced to 26.3% and 45.3% after training by teacher for one month (A total of about 10 hours). However, these studies showed that the tone of the tone improved by the learning method performed, and that it was effective, but could not show the superiority of the learning method performed. Though the participants' performance improved, evaluation of training methods and judgment of superiority were difficult, and training by teachers should consider teachers' personalities.

A study compared two training methods to estimate the effects of teaching methods, one uses a monosyllabic tone, and another uses disyllable tone combinations for the initial training [26]. As a result, the pronunciation error rate of monosyllabic words of the group trained for the monosyllabic tone was 45.3%. And the error rate of the monosyllabic words from the group trained for disyllabic tone pairs perception and pronunciation tasks designed was 10.9%. The pronunciation error rate of the disyllabic words from the group trained for monosyllabic tone was 61.7%. And the error rate of the disyllabic words of the group trained for disyllabic tone pairs was 16.7%. However, the participants' learning experiences and nationalities were different in this research. In related works, the learning experience and nationality of participants is rarely consistent.

There are studies that have conducted perceptual training to improve tone ability in Chinese perception and production; however, the content of the learning is mainly monosyllabic and disyllabic. A study that described the effects of perceptual training showed that after two weeks of perceptual training using disyllable words, the tone perceptual ability improved by 21% and lasted for over six months [28]. Studies have also divided participants into groups to measure the learning effect of the perceptual training and pronunciation practice, showing that when using monosyllabic words to perform perceptual and pronunciation training, perceptual training has a significant effect on improving perceptual ability, but the improvement in pronunciation ability is much less than that achieved by pronunciation practice [29]. Similarly, pronunciation practice is effective to improve monosyllable tone pronunciation, but the effect to improve monosyllable tone perception is lower than that of perceptual training. However, a study using the application to practice disyllable word perception and pronunciation by imitating model voice with pitch tracking shows that it is effective to improve pronunciation ability only with perceptual training [27]. Training only in perception is sufficient for improving the tone perception and pronunciation in disyllabic words. And the disyllabic words tone error rate of participants who had studied Chinese for over eighty hours before training was 44.5% on average. The participants after perceptual training with 1280 disyllabic words in the application for eight hours, their error rate reduced 24%.

Therefore, we aim to improve learners' tone cognition and pronunciation ability of tone by using different learning methods and strategies. For this purpose, we researched theories and experimental results of tone cognition and human processing. In language learning, we focused on the difference between pronunciation learning and acquisition of knowledge, like learning words or grammar, and considered cognitive and pronunciation learning support with the

application. This study aims to improve the pronunciation learning efficiency of tones, which is a major issue for foreign language learners, especially those who do not speak tonal languages, and conduct surveys and evaluation experiments on Japanese learners.

1.2. Tone Cognition and Human Processing

A study [11] has shown that even advanced second-language learners have difficulty recognizing tones. Differences between target language and native languages and individual cognitive ability determine difficulty in learning tones. However, educational policies and awareness of tone and learning strategies can be modified, which is a major factor in improving the efficiency of tone acquisition. A study was conducted to compare the learning effects of using monosyllable words with using disyllable words on the initial education of Mandarin Chinese tones [26]. The results show that using disyllable words in the initial education is much more effective in acquisition of tone perception and pronunciation of monosyllable words and disyllable words (18.7% to 45% improvement). Regarding tone error of Mandarin Chinese learners, a study shows that teaching the third tone as "dipping-rising tone" may cause tone pronunciation errors [30]. Furthermore, there are tone errors due to wrong learning strategies of learners, such as learners who are confident or good at recognizing and imitating pitch learning tones as melody or learners who concentrate on reading and writing without noticing the importance of tone that may lead to fossilized errors.

Therefore, we focused on tone recognition based on information processing in cognitive psychology [31], [32], which uses computer information processing to approach or explain human information processing.

Cognitive psychology aims to elucidate the human information processing (perception, memory, cognition, etc.). Human information processing has stages and includes the following three processes:

1. Discovery—Be aware of information in the environment that were previously overlooked;
2. Memory—Make the information meaningful and store in short-term or long-term memory;
3. Modify—Modify long-term memory (re-learn).

It is assumed that tone learning is recognizing audio information based on existing knowledge of learners [33]. The knowledge of tone, especially the visual symbols as shown in Figure 1-1, can greatly help learners to discover and memorize tones. Furthermore, it is considered that recognizing tone requires considering the shortness of sensory memory (about 1 second or 0.25 to 4 seconds) among three stages of human memory: sensory memory, short-term memory, and long-term memory.

In the theory of cognitive psychology and education, based on the fact that the mental resources that can be used for human information processing are quite finite, cognitive load theory [34] for effective use of human cognitive resources has been developed. Cognitive load can be classified into the following three categories in relation to learning.

1. Intrinsic cognitive load—Related to the learning content (difficulty of learning content).
2. Germane cognitive load—Related to learning effect (learning tasks, function, etc.).
3. Extraneous cognitive load—Related to how to present the learning content (interface, operation design, etc.).

Even learners who are unable to recognize the tone are likely to be able to pronounce tone correctly in pronunciation practice by imitation of words and

phrases due to the help of sensory memory. Imitation pronunciation practice seems to be effective; however, we consider that for learners who are unable to recognize the tone, pronunciation practice by imitation improves speech memory and imitation ability rather than acquisition of tone. In addition, to avoid fossilized errors caused by pronunciation practice without correction, it is important to support learners' language recognition with different learning contents and methods or voice analysis.

1.3. Research Method

The general learning method and many teaching materials emphasize on imitation practice. Furthermore, in Chinese classes that are offered as second or foreign language education, students are required to learn basic knowledge of Chinese and to develop basic speaking skills and listening comprehension in a rather limited time. This means teachers could not take a biased educational policy in class aimed at overcoming specific difficulties. Especially acquisition of tone varies considerably from person to person. There is a strong need for support through electronic teaching materials that can provide voice stimulation leading to the improvement of students' perceptual and pronunciation abilities.

To develop students' sense of a second language, teachers must be conscious of the students' recognition of the phonetic composition of the target language. However, it is liable to neglect the ability of Chinese learners to recognize the tone. For this reason, incorrect pronunciation wears, the progress of learning is often stalled, especially for learners who are not good at dealing with tone. Training and confirming ability about pronunciation by supported with teaching materials in this background is considered necessary. Therefore, we considered that a mobile app offering high teacher availability and freedom would be most suitable.

Table 1-1. Comparison of Feature between Related Works and This Research.

Features of research	Related works (Figure 1-2)	This research
Purpose	Investigating how to overcome the difficulties of Chinese tone pronunciation by finding out error trends among learners. Few studies considered the differences in performance due to controlled variables like learning experience or teaching strategies.	Discovering an appropriate learning method based on the perceptual state of Chinese learners. Focusing on polysyllabic perceptual training tasks design and verify the effectiveness of training by solving perception questions.
Research method	Investigate error trends through phonetic experiments refers to an interlanguage hypothesis, contrast analysis, and error analysis. The contents of the phonetic experiments are mainly disyllabic words.	Investigate tone-understanding ability by perceptual experiments to design learning tasks. Carried out the evaluation experiment with the disyllabic, four-syllable task and phrases to examine tone pronunciation ability.
Challenges	Influence of native language, Insufficient guidance.	Inappropriate learning method or tasks.
Learning methods	Imitation practice focusing on tone combination or lecture by the teacher (mainly on disyllabic words and phrases).	Task design that aims to improve perceptual ability; Polysyllabic perceptual training tasks that make learners feel the difference of tone.

We consider that acquiring tone should start from recognition, the input of learners. Therefore, we emphasize a learning method that aims to make learners feel the difference of each tone. Furthermore, providing the model voice thought to be useful for learners to grasp the acoustic feature of Chinese tone is also our purpose.

Table 1-1 shows the features of the related works in described in Section 1.2 and this research.

1.3.1. Perception of Information about Tone

Recognition of tone is based on auditory sense. First, learners store voice information in working memory, and then, based on existing knowledge (the

definition of tone), they make the voice information meaningful and store it in long-term memory. It is speculated that tone is often a difficult feature to learn because students cannot distinguish pitch difference, which is a characteristic of tone, or they are unable to link the information with their existing knowledge.

Abilities that necessary for learning Chinese tone are correct recognition of tone based on knowledge acquisition and perceptual ability in human information processing. The proposed method emphasizes the perceptual ability of the acoustic features of tone. To move from the first stage of information processing, discovery and recognition of information processing, to the next stage, coding and storage, it is important to enhance the impression by tone imaging and encoding.

1.3.2. Program Instruction Design on Perceptual State

A stepwise learning method such as program instruction starts with simplified items and gradually adds factors to approach high difficulty objects little by little. For validation of the characteristics of the program instruction, the item contents need to be designed based on experimental data on tone learning. In accordance with the tone learning procedure proposed, for improving perceptual tone ability, we have to investigate the factors that make tone difficult to learn and designed training items based on a learner's tone perception survey.

1.3.3. Support through Electronic Teaching Materials

Because of large individual difference in tone learning speed, which places a heavy burden on educators, a material that provide voice stimulations to help learners deal with tone learning difficulties is important. In order to develop students' sense of a second language, teachers must be conscious of the students' recognition of the phonetic composition of the target language. Therefore, we

considered that a mobile app offering high teacher availability and freedom would be most suitable.

Based on the related research and theory in the previous section, we considered the intrinsic cognitive load (difficulty of speech recognition, complexity of information received by learners), germane cognitive load (the instructional design and feedback to promote tone understanding) of the general learning method, and extraneous cognitive load (information that is not expected to be effective in learning tones, teaching material interface and operation design, etc.). The research procedure is shown in Figure 1-3.

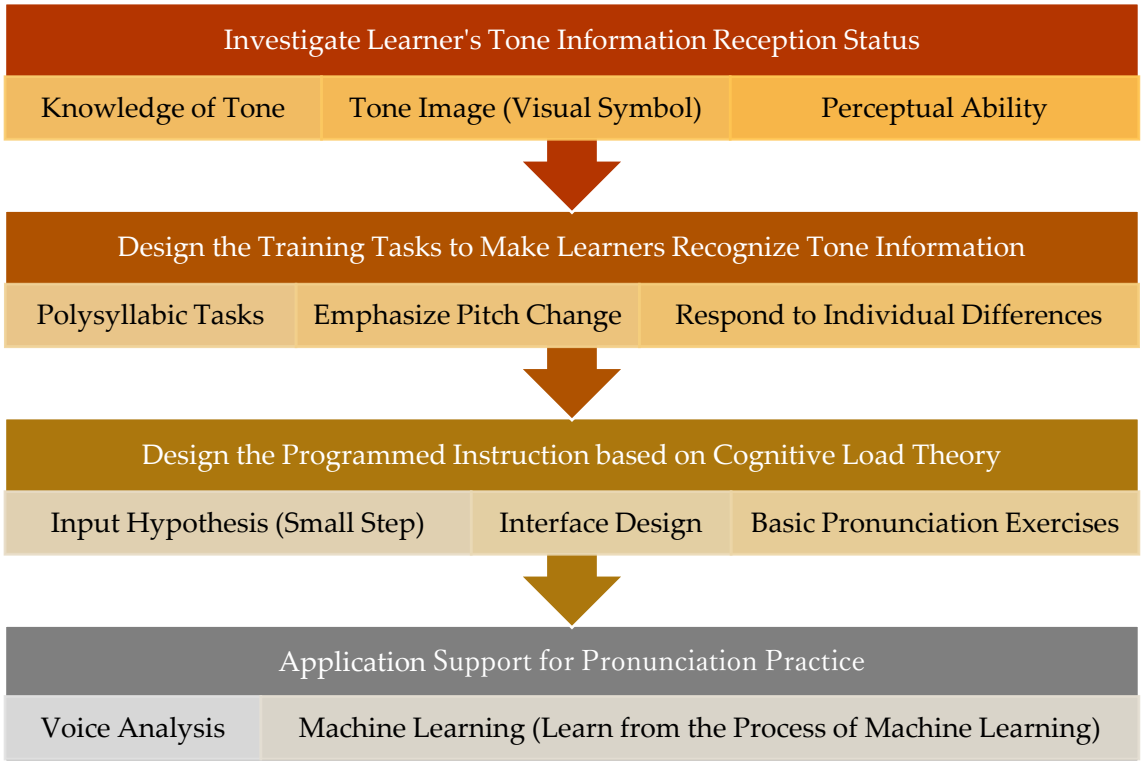


Figure 1-3. Research method.

This thesis is organized as follows. Chapter 1 describes the related works on Mandarin Chinese tone learning and the cognitive abilities of learners that are tend to be neglected in Mandarin Chinese education and theories about human cognition. Based on the learning theory which is thought to be appropriate for

improving tonal skills, we focused on a new learning method from a different viewpoint and aimed at developing a tone perceptual training application.

In Chapter 2, as a first step in realizing the proposed method, we conducted tone perceptual experiments and describe the results of investigating the tone recognition state of learners, the individual difference and the effects of training tasks. Based on the results, learning contents of the tone perceptual training application designed and the shortcomings of general learning methods and the improvement plans based on cognitive load theory are described.

In Chapter 3, the effect of the improved learning method is shown by the evaluation experiments results of the user study of the tone learning applications. We compared the effect of the imitation practice application and the perceptual training application we designed. The application usage status of participants, results of pronunciation tests and questionnaire are described.

In Chapter 4, based on the theories about human cognition, we attempted to improve the tone perceptual training app by optimizing the user interface and adding pronunciation practice function. The results of evaluation experiment that compared the effect of the optimization are described.

Chapter 5 presents the general discussion and the fundamental study for future tasks designed to make pronunciation practice more effective and feedbacks on the pronunciation of learners provided by machine learning. In addition, the conclusion and future tasks of this research are described.

Chapter 2

Development of a Tone Perceptual Training Tool

Based on the relevant research and the theories in the previous section, this research aims at developing a self-learning material with multi-syllable perceptual training contents. We targeted on Japanese learners who have no Chinese learning experience and investigated the educational contents that help learners acquire categorical perception of Mandarin Chinese tone. According to the results of the investigation, we made a prototype of tone perceptual training material with the educational contents.

2.1. Educational Contents Design

In the theory of cognitive psychology and education, to make full use of learner's cognitive resources, we focused on difficulty of speech recognition, complexity of tone information received by learners, instructional design and feedback to promote tone understanding.

2.1.1. Intrinsic Cognitive Load (Contents of Learning)

The model voices used for imitation pronunciation practice are mainly words and phrases. However, a Chinese word or phrase is composed of many elements

that could be not exist in learners' native language and learners should need to deal with phonemes, tones, phonetic notations, characters and meaning during practicing pronunciation of words and phrases. It is assumed that the intrinsic cognitive load appears to be very high for most Chinese beginners. Even learners aim to learn tone by phrases pronunciation practice, the intrinsic cognitive load would increase because of the other elements in Chinese phrases.

First, we focused on the first stage of human information processing, discovery, and aimed to make learners cognizant of the tone. An experiment was conducted to investigate the relationship between knowledge and tone perception using visual symbols of tone and tone perception tests. The experiment was planned to conduct with native speakers of Japanese that use Chinese characters as well as Mandarin Chinese as participants and investigated how beginners can recognize tones, how designed perception tests can help improve cognitive abilities, and individual learner differences in tonal cognitive abilities. We aimed to create a collection of perceptual training questions that are unproblematic, emphasizing the comparison of each tone, and with many variations of model voice (multiple pitch ranges, sounds, phonemes and speakers) to make learners recognize the differences of tone in relative pitch patterns.

2.1.2 Germane Cognitive Load (Instructional Design and Feedback for Learning)

The general teaching materials are a provision of model voice (manly words, phrases, and sentences) and provide visual feedback using pitch tracking function for learners to check their pronunciation. However, lack of cognitive ability, especially for beginners of Chinese, means that categorization perception of tone has not been advanced, and it is unlikely that learners can improve or

evaluate their own pronunciation correctly by listening or checking the results of pitch tracking. Furthermore, the results of pitch tracking are different from human perception, and acoustic analysis results vary for visual symbols designed based on the perception of native Chinese speakers.

We considered that the tasks design of tone learning aims at improving the cognitive ability, which is the basic ability for tone acquisition, and speech information provided should focus on promoting the categorization perception of tones. Therefore, referring to the input hypothesis [35], we made the perceptual training questions collection, that led learners to start from the simplest perception question and gradually change the pitch ranges, sound phoneme, or speaker as the germane cognitive load. Furthermore, showing the answer status of each question and task as learning feedback, the users could improve their perceptual ability and select perceptual training tasks that matched their abilities according to the feedback.

2.2. Perceptual State of Chinese Tone

The learning method we prefer emphasizes step by step acquisition with polysyllabic perceptual training. To implement the learning method, we carried out the perceptual experiments to investigate the tone-understanding ability of university students who have not had previous Chinese learning experience. Before experimenting on the perceptual state of the Chinese tone, it is necessary to let learners acquire a basic knowledge of Chinese pronunciation first. The introduction about tone is performed using the tone image seen in the previous section. In addition, we let participants hear models that combined voice and tone images, such as a four-syllable template of tones, disyllabic words, four syllables, and number counting in Chinese.

2.2.1. About Tone

In Chinese pronunciation, each syllable has a tone that distinguishes meaning through pitch change. The four kinds of Chinese tones are the first tone (high tone), the second tone (mid-rising tone), the third tone (low-dipping tone or falling-rising tone) and the fourth tone (high-falling tone). The requirement that each tone holds is the pitch height relate to the others. We considered that it is necessary to emphasize the training of relative pitch height perception for the task design.

2.2.2. Tone Image

The image of the tone uses Yuen Ren Chao's five level tone marks [8] (Figure 2-1). Then, in consideration of the actual conversation state and the burden and efficiency of learning, the third tone (Figure 2-2) of the perceptual experiment used the "low-dipping tone" [36], [37].

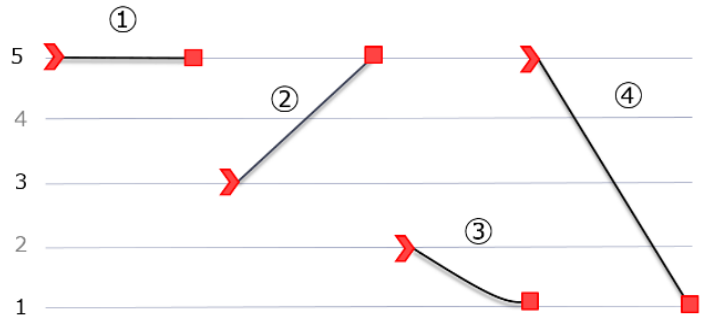


Figure 2-1. Images of tones used in the perceptual experiment.

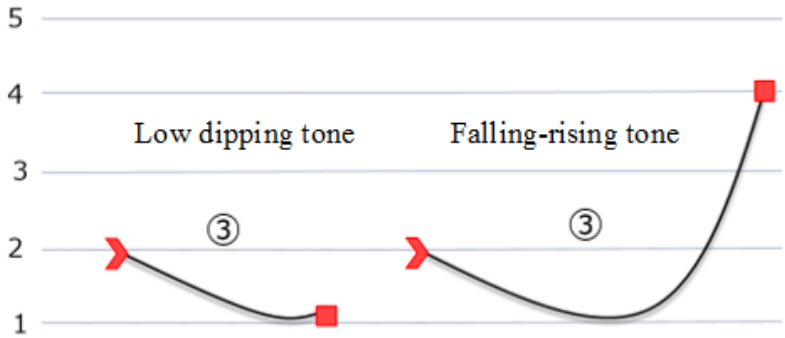


Figure 2-2. Two images of the third tone.

Zhang's disyllabic word pronunciation research [13] shows that there is a tendency to speak the third tone with a "Falling-rising tone" as one of the reasons for pronunciation errors of the third tone. Furthermore, the frequency of the "Low-dipping tone" is higher than the "Falling-rising tone" in Chinese, and it is easier to express the third tone without misusing it.

2.2.3. Four-character Idioms

In order to investigate the Chinese tone perceptual state of Japanese native speakers who have no Chinese learning experience, we used the four-character idioms in the first experiment. The way the experiment progresses is to gather the participants in the classroom, distribute sheets for recording the interview results, and let them write the tone of the heard voice. Audio of each question is presented twice and the participant fills answers in the sheet by drawing the image of tone shown in Figure 2-1. In order to investigate the Chinese tone perceptual state of Japanese native speakers who have had no Chinese learning experience, we used four-character idioms in the first experiment. After explaining the language composition, we let the participants hear the four-character idioms and fill in answers on the answer sheet (Appendix 1).

In this experiment, we chose four-character idioms that exist in both Japanese and Chinese, such as "臥薪嘗膽" and "森羅萬象," and the four-character idioms were presented using a model voice. After hearing the model voice for each question twice at normal speed, we let participants hear the model voice twice at a low speed and answer the question again to investigate the state of tone perception for low-speed Chinese.

The experiment was conducted using a total of eight participants: one female in her 20s, six males in their 20s, and one male in his 50s on 17th April 2017.

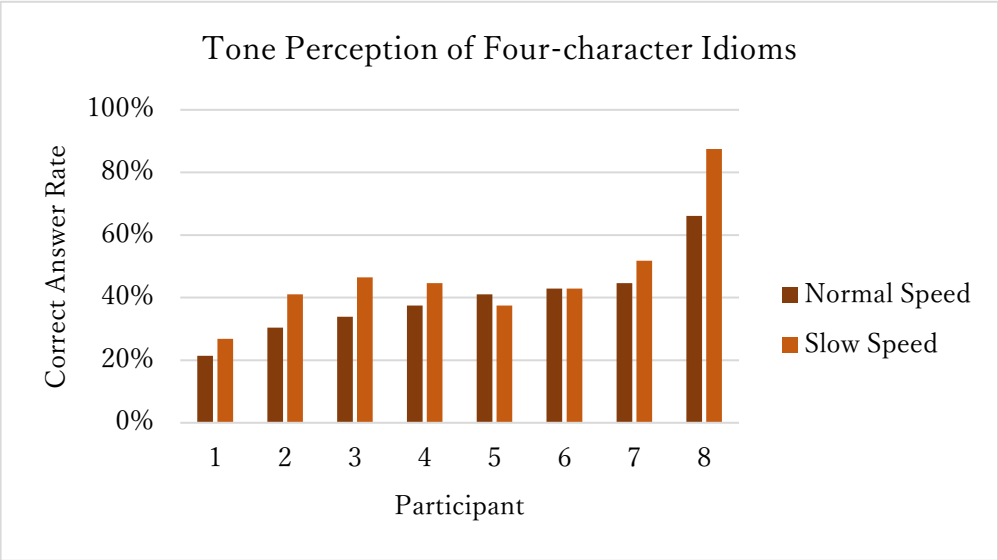


Figure 2-3. Tone perceptual experiment 1.

Except for one participant, the correct answer rate of the four-character idioms (a total of 56 syllables) was less than 50%. Figure 2-3 shows the experimental results. According to Figure 2-3, it appears that improving the rate of correct answers by listening at a lower speed isn't significant. Sometimes participants made a mistake when answering the slow speed task even if they answered correctly in the normal speed task.

The experimental results show that beginners of Chinese could not recognize tone correctly. Therefore, repeating the model voice of Chinese phrases, which is a general Chinese pronunciation exercise, is performed in a state in which the tone can hardly be recognized, and we speculate that there is a high possibility that tone learning does not advance in this case.

2.2.4. Tone Perceptual Experiment with Rules

In response to the results of 2.2.3, to realize the stepwise learning method of program instruction, we designed a tone classification task, which seems to be effective in the early stages of learning, and carried out the following experiment two weeks later on 1st May 2017 (Appendix 2).

In the experiment (Figure 2-4), we aimed to create tasks that simplified the four-character idioms and carried out perception experiments of four syllables using the following three rules.

(1) One question includes all four types of tone
(2) Only one pronunciation used for one question
(3) Increase the speed reduction rate (2 seconds per syllable)

The first rule was used to make participants concentrate on the difference of each tone; the second rule was to eliminate voice information that may confuse participants, and according to the results of experiment 1, speed reduction was considered to have had an effect on the improvement of tone perception; thus, we used the third rule to increase the effects of the method. We surmised that individual differences in tone perception of participants would appear; thus, three males in their 20s were added to the participant group in this experiment.

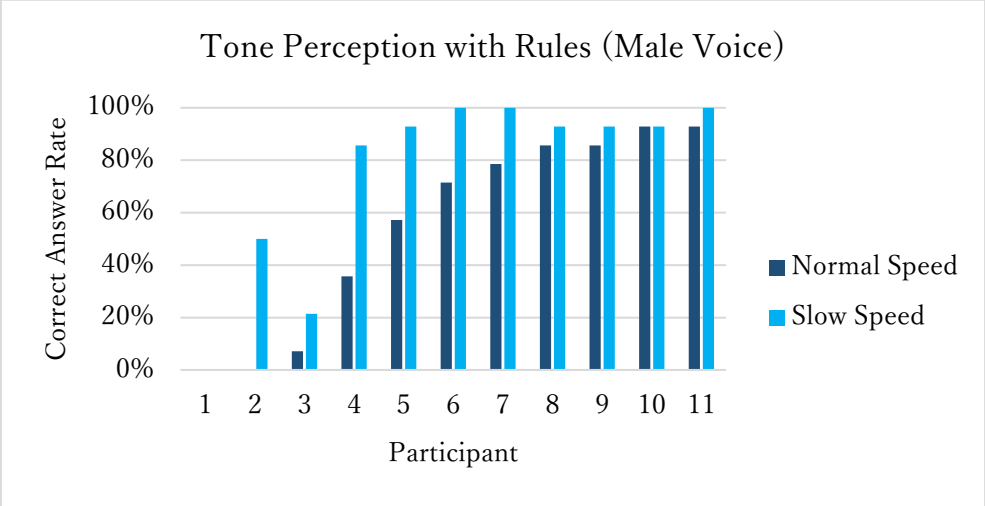


Figure 2-4. Tone perceptual experiment 2.

As shown in Figure 2-4, the four kinds of tones are classified by the relative height of the pitch, and by using these three rules, it is easier to feel the difference and the pitch change regarding tone. Three trends were found in the tone learning results from the experiment.

(1) 3 of the 11 participants found it difficult to recognize the tone, even with this method.

(2) 3 participants were able to recognize tones at normal speed by hearing them again at low speed.
(3) 5 participants had a favorable recognition status irrespective of speed.

In this experiment, we decided to use four kinds of phonemes for each of the three questions according to participants who demonstrated trend (2) and had a 0%, 0%, 25% correct answer rate for normal speed questions when changing the phoneme. After hearing the slow stimulus, the correct answer rate was 50%, 87.5%, and 87.5%, respectively. Therefore, lowering the speed can be considered an effective method to improve the tone perception of different phonemes. The phonemes used were two vowels similar to the Japanese "i" and "u" and two vowels that are not used in Japanese. Then, we changed the speaker of the questions from male to female and carried out this tone perception experiment again two weeks later on 15th May 2017 (Appendix 3).

According to the results of this experiment (Figure 2-5), we created two groups of tendencies: difficulty in recognition and favorable recognition state. The classification ability of the tone and different phoneme perceptual experience acquired by the training can also be applied to the tone with the changed pitch range.

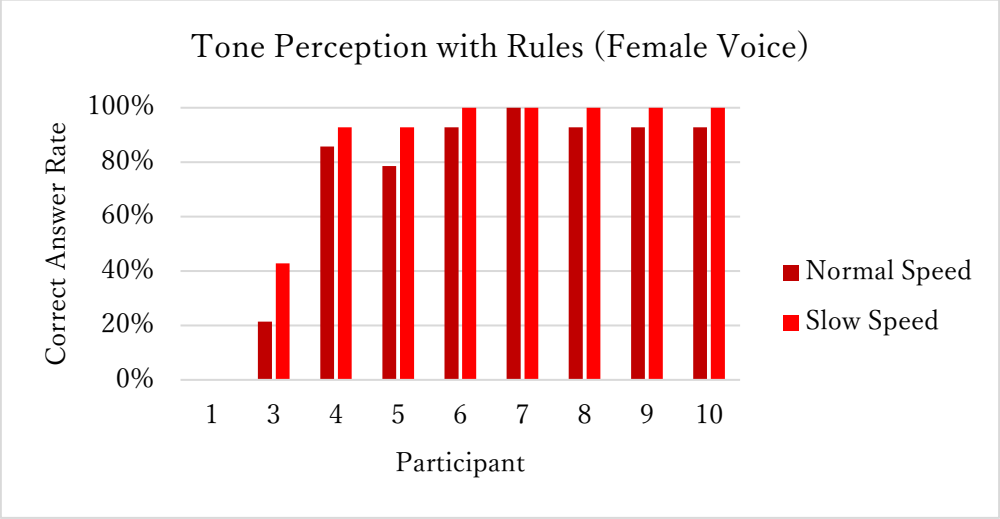


Figure 2-5. Tone perceptual experiment 3.

After conducting these two experiments, we have concluded that by using the above three rules, 73% of participants were able to improve their ability to classify tones by answering tone perceptual questions.

2.2.5. Relationship Between Phoneme and Tone Perception

We speculated that the perceptual state of phonemes may be related to tone perception. To confirm this, we designed a phoneme and tone perceptual experiment and carried it out a week later on 22th May 2017 (Appendix 4).

In this experiment, we used six four-character idioms (24 syllables) and let participants mark the phonemes and tones on the answer sheet. There were 14 phonemes tested in the perceptual test. If the phoneme "su" was tested, participants would hear six four-character idioms in order once and mark the sign of the phonemes that they thought was included in the syllable. The 14 phonemes were presented with Zhuyin fuhao (BoPoMoFo), the Chinese transliteration system used in Taiwan. To test 14 phonemes and tone, participants would hear the 6 four-character idioms 11 times. The perceptual test of tone was the last one in the experiment.

Figure 2-6 shows the correct answer rate of 10 phonemes (Zhuyin fuhao) used more than three times in the six four-character idioms and tone perception. As shown in Figure 2-6, the relationship between phonemes and tone perception is weak.

Figure 2-7 shows the number of incorrect answers filled in for each phoneme. The data shown in Figure 2-7 does not include incorrect answers filled in due to retroflex false perception; the perception of retroflex had an 87.5% error rate in 24 syllables of seven participants.

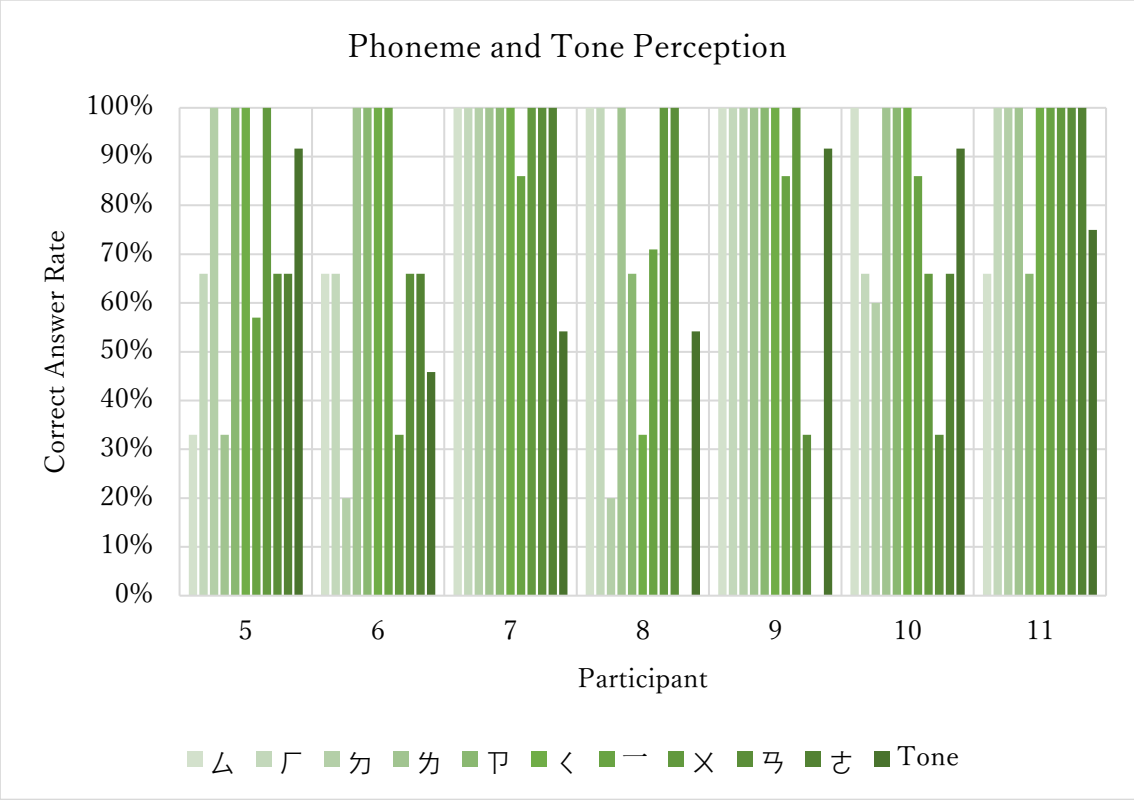


Figure 2-6. Correct answer rate of phonemes and tone perception of 7 participants (experiment 4)

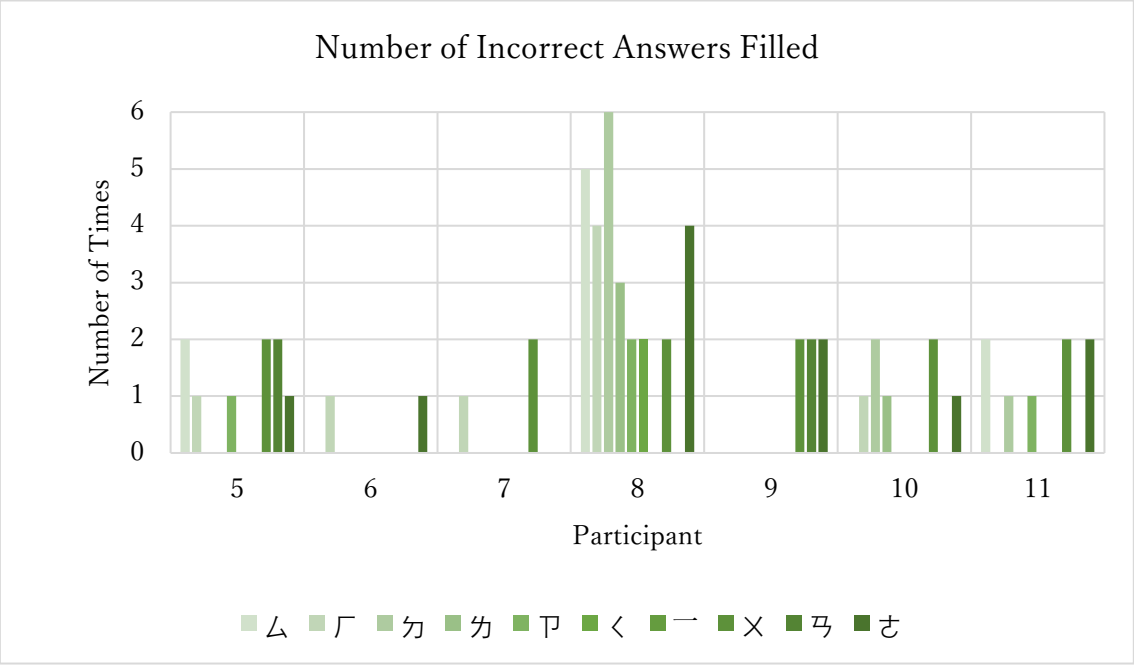


Figure 2-7. Number of incorrect answers filled with each phoneme in experiment 4.

2.2.6. Learning Effect of Tone Perceptual Questions with Rules

In order to investigate the learning effect of answering rule-attached tone perceptual questions in more detail, the fifth experiment was performed with perceptions of four-character idioms after the rule-attached tone perception on 5th June 2017. (Appendix 5)

This experiment did not use the speed reduction method, and 22 questions of rule-attached tone perception using 22 phonemes and four-character idioms perceptual questions were carried out. The results are shown in Figure 2-8. The 22 phonemes consisted of the vowels and the combination of vowels with pronunciation changes.

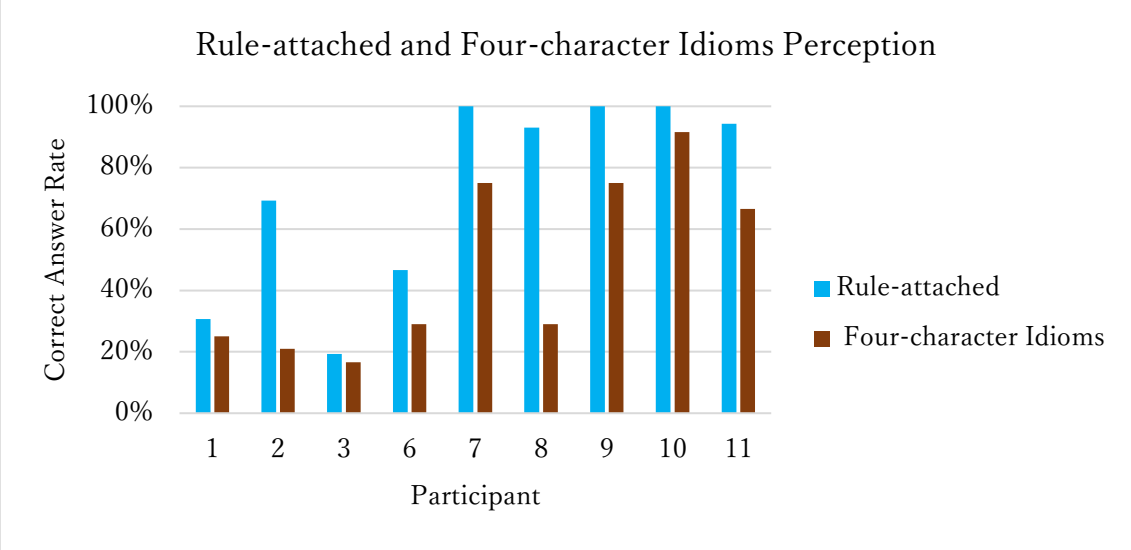


Figure 2-8. Tone perceptual experiment 5.

According to the 4th experiment's results, after conducting the rule-attached tone perceptual experiment, there were five participants with a correct answer rate of more than 80% in the four-character idioms, significantly higher than the 1st experiment's results. However, in the 5th experiment's results, although there was a participant who had a correct answer rate of more than 90% by tone classification of the rule-attached tone perceptual questions, they obtained a correct answer rate of less than 30% in the four-character idioms section. The

results are considered to be related to the number of times the model voice is heard.

2.2.7. Tone Perception from One Syllable

In the previous section, 27% of participants could not correctly classify tone in the rule-attached tone perceptual experiments, and the improvement of the correct answer rate was insignificant. Therefore, to discover the perceptual question with the learning effect, we conducted the 6th, 7th, and 9th experiments on 19th June, 3rd July and 31st July 2017. (Appendix 6,7,9) In this case, perceptual experiments were carried out by gradually changing the method from the questions considered to be the easiest, and the change in the correct answer rate of perception was investigated. In each experiment, five patterns, as shown in Table 2-1, were also compared, in addition to one syllable, rule-attached, and four-character idioms tone perceptual questions.

Table 2-1. Question patterns added to perceptual experiments

x, y in question patterns indicates the composition of the four-character idioms

Question patterns	Detail
xxxx four-character idioms	The four-character idioms perception with the same tone in 4 syllables
xxyy four-character idioms	The four-character idioms that the first two syllables are the same, and the last two syllables are the same tone.
wxyz four-character idioms	The four-character idioms that 4 syllables all different tone
xxyy four-character idioms x,y=1 or 4	The four-character idioms that the first two syllables are the same, and the last two syllables are the same tone being 1st or 4th tone.
xxyy four-character idioms（x,y=4）	The four-character idioms that the first two syllables are the same, and the last two syllables are the same tone. ＊The answers were all 4th tone.

Although the results of the survey (Figures 2-9, 2-10, and 2-11) showed an increase in the correct answer rate for rule-attached tone recognition, some participants had a correct answer rate of less than 60%. From these experiments,

it was found that there was a large individual difference in the perceptual ability of tone among participants whose native language is Japanese. It is considered important to acquire the basic ability of perceptual and pronunciation ability of tone with suitable training tasks.

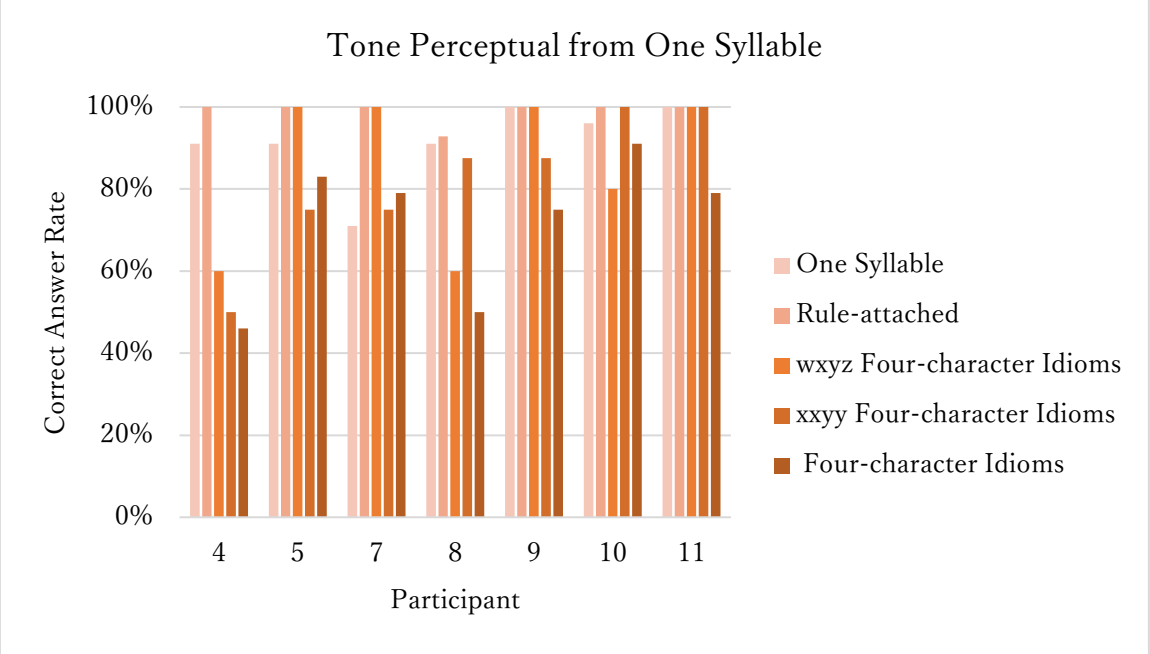


Figure 2-9. Tone perceptual experiment 6.

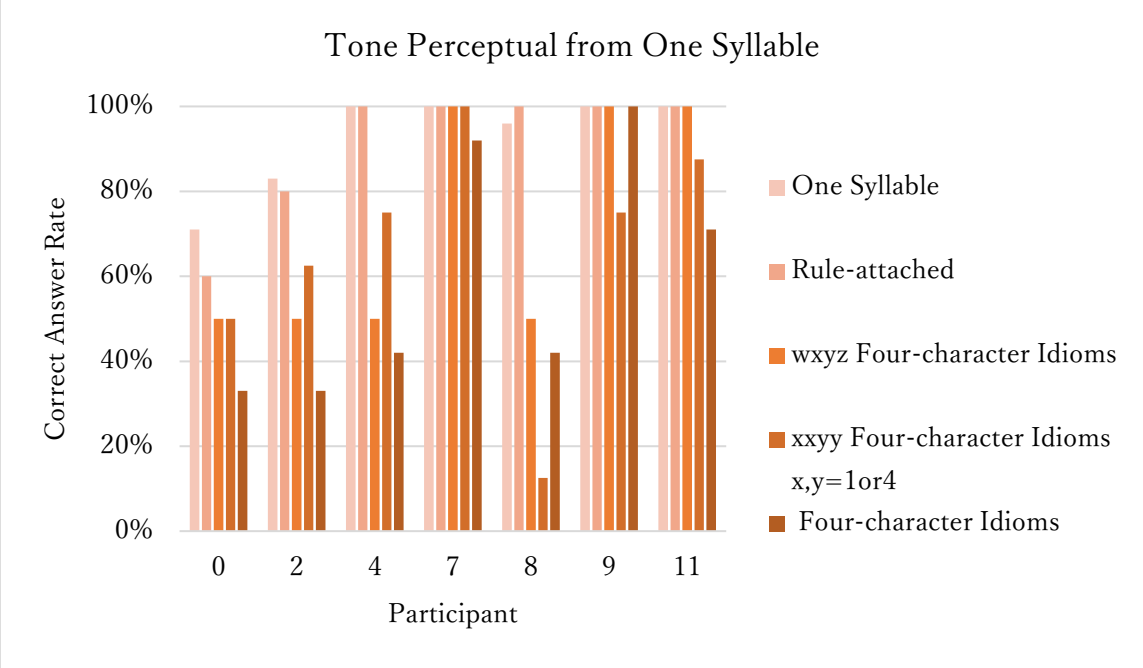


Figure 2-10. Tone perceptual experiment 7.

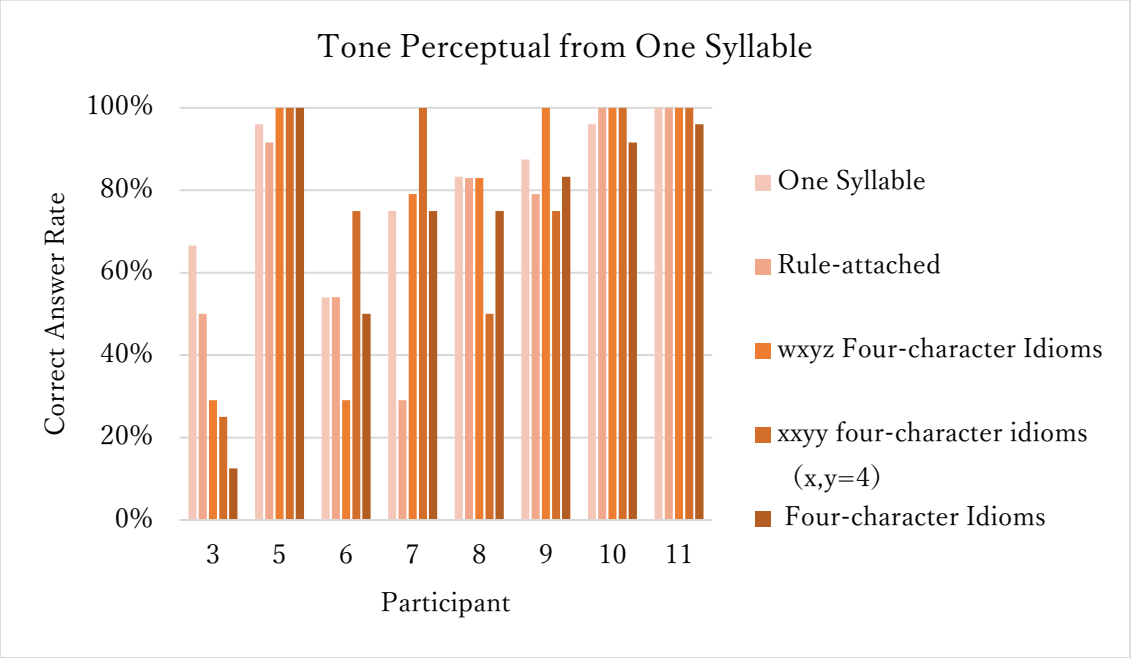


Figure 2-11. Tone perceptual experiment 9.

2.2.8. Change of Correct Answer Rate at Perceptual Tasks - Phoneme and Speed

We confirmed the correct answer rate for different tasks in the experiments we conducted to investigate the learners' tone perceptual state regarding different voice element changes.

The 8th experiment was executed by adding the following 2 rules to the four-character idioms and carried out on 18th July 2017 (Appendix 8). We carried out and compared the correct answer rate with the normal (3 rules) rule-attached tone perception (Figure 2-12).

(1) One question includes all four tones

(2) Increase the speed reduction rate (2 seconds per syllable)

As shown in Figure 2-12, the correct answer rate of some participants at normal speed decreased in the four-character idioms composed of unknown or complex phonemes. However, we found that the correct answer rate regarding the normal rule-attached questions can be obtained by lowering the speed.

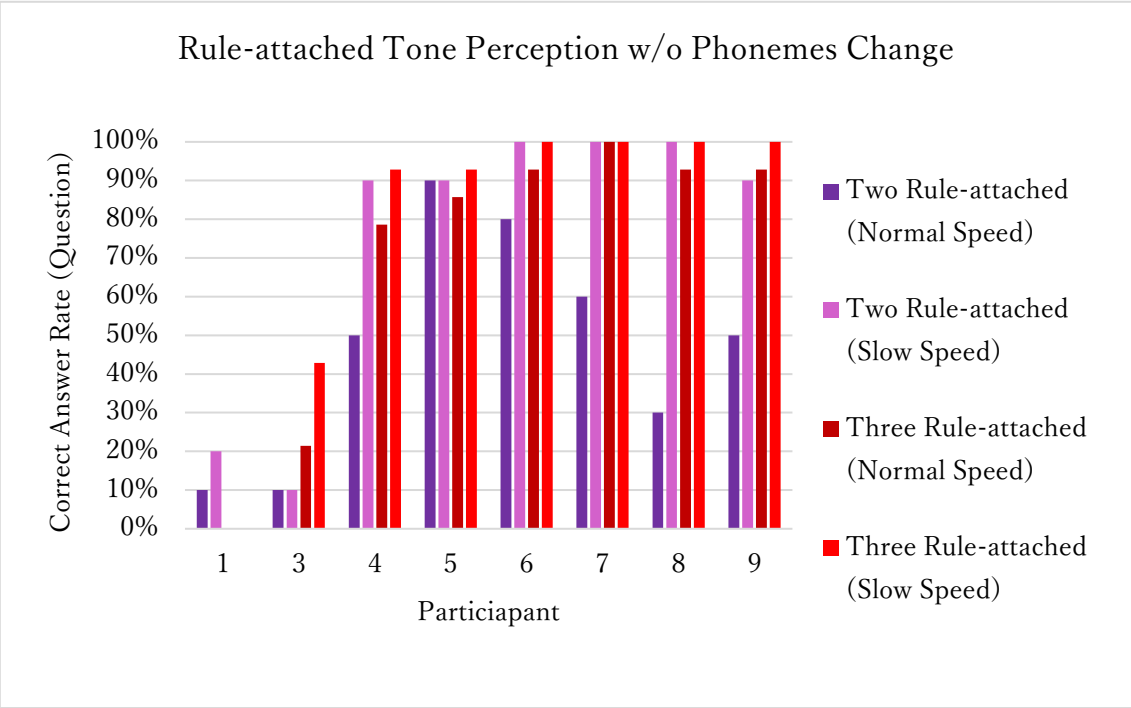


Figure 2-12. Compare of tone perceptual experiment 3 and 8.

2.2.9. Error Analysis in Perceptual Experiments

A survey of difficulties for users who advanced their training using the technique designed by the authors found that they successfully improved their perceptual ability of tone. According to this survey, we analyzed the error rate of each tone in the four-character idioms perceptual experiments of participants whose correct answer rate for syllables was 85% or higher in the simplified tone perceptual experiment. We show the error tendency caused by the complication of phonemes and tone combinations. For the analysis, we used the data from experiments that carried out both simplified tone perception and four-character idioms in the experiment conducted by the authors in Chapter 3.

Figures 2-13 and 2-14 show the error rates of the four tones and the ratios of confusing them with other tones. As shown in Figures 2-13 and 2-14, the tone with the highest error rate was the 4th tone with 38.4%, and the others showed an error rate of 20.6% to 24%. The 1st tone tended to be confused with the 2nd

and 4th tone. Furthermore, the 2nd and 3rd tone were easily mistaken for each other, and the 4th tone was easily mistaken for the 1st tone.

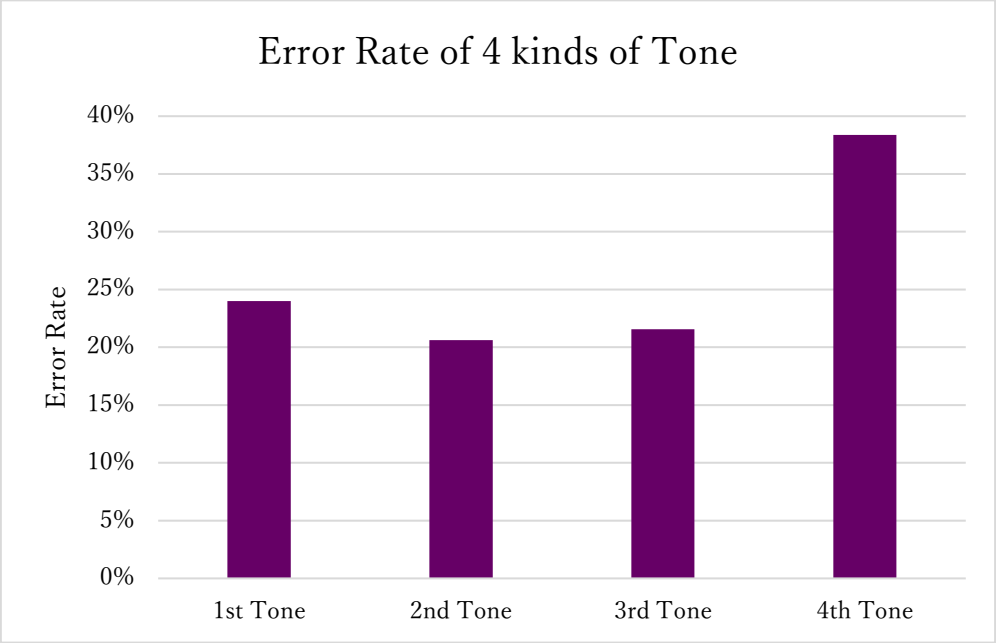


Figure 2-13. Tone perception error rate of 4 tones.

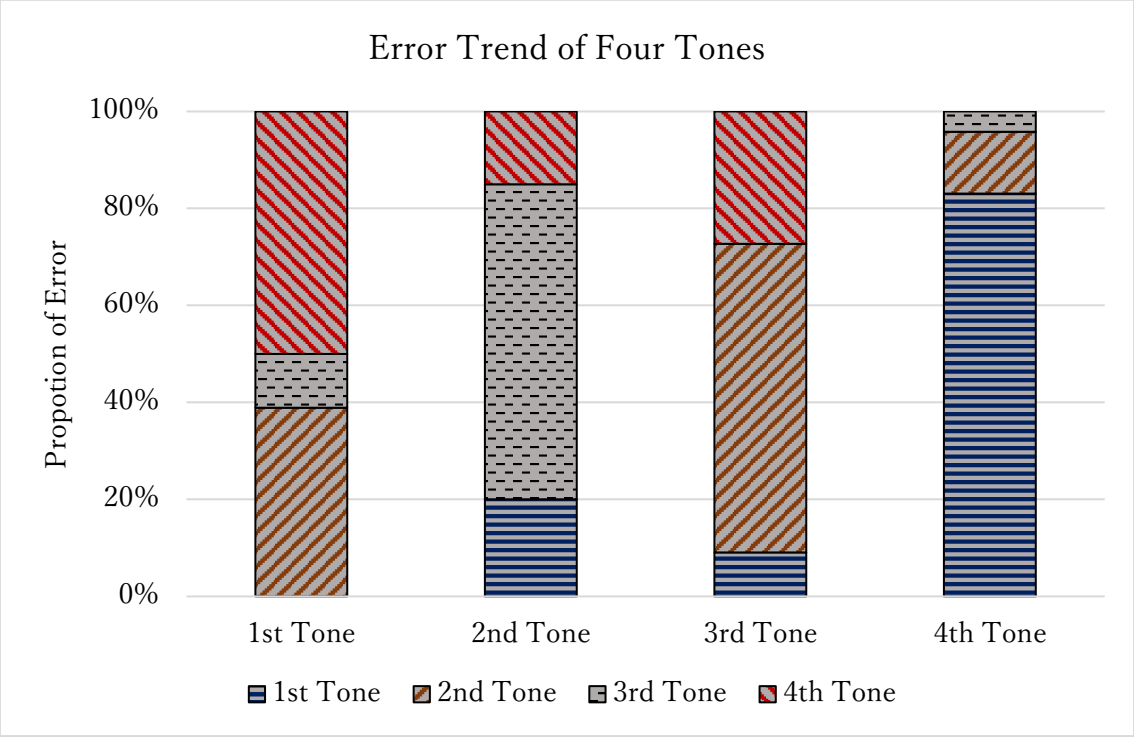


Figure 2-14. Tone perception error trend of 4 tones.

The error rate of error patterns is shown in Figure 2-15 by multiplying the error rate of each tone and the ratios of being mistaken for the other tones. As shown in Figure 2-15, the highest probability was found with the 4th tone being mistaken for the 1st tone with 31.9%, and the second highest was mistaking the 3rd tone as the 2nd tone (13.7%).

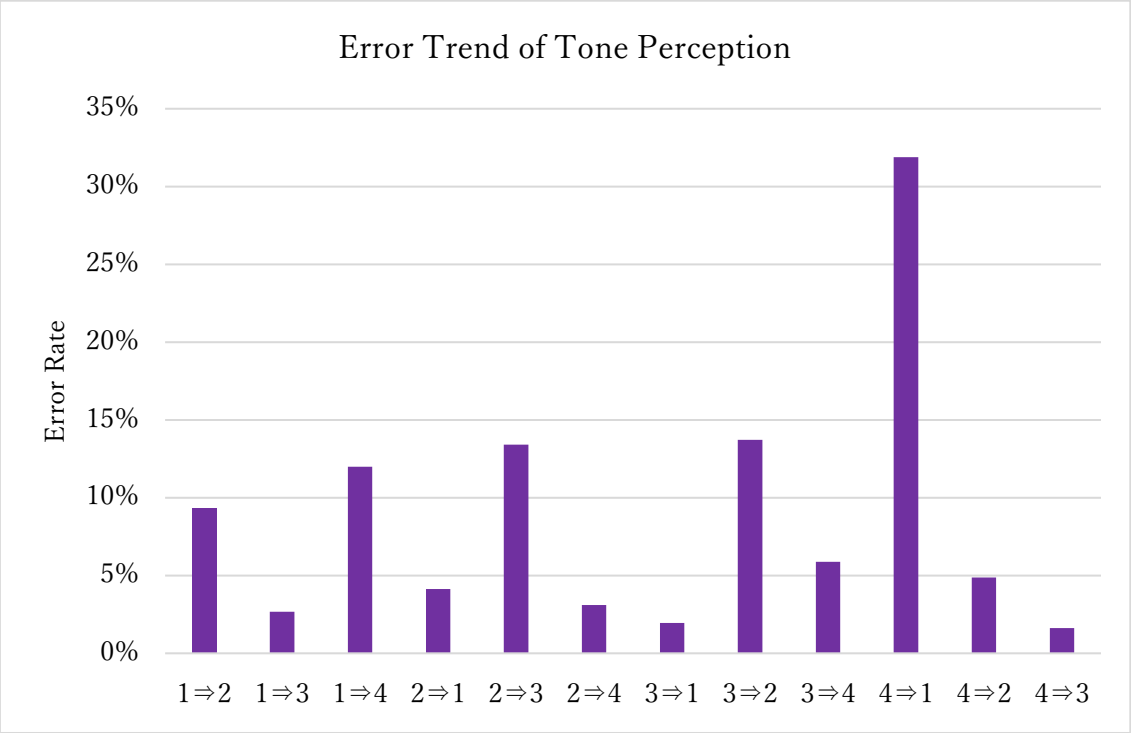


Figure 2-15. Error rate of tone error patterns.

We found that it is necessary to consider the training measures of the error patterns and implement them in items after the simplified task in the application.

2.3. Learning Through Perceptual Questions

Based on the results of the tone perceptual experiment, the perceptual questions using the three rules in 2.2.4 were designed to help students identify pitch changes in tone.

Furthermore, we prepared the tone perception questions for the learner who needs more manageable tasks that consist of pure tones with different pitch range and pitch height. The change in model voice was gradually added to improve students' perceptual ability by answering several perception questions. The tone perception questions consist of pure tones using the rule (1) One question includes all four types of tone in section 2.2.4, and we designed them with four pitch range: 1. High bandwidth (150Hz-600Hz) 2. Low pitch (80Hz-140Hz) 3. Medium pitch (160Hz-280Hz) 4. High pitch (200Hz-360Hz). Except for adjusting the width of the frequency change and pitch range of pure tones, provision of model voice with both male and female voices, adding questions with different phonemes etc. were used to make learners training step by step.

The tone learning tool for basic learning of tone was constructed with the following perceptual training tasks.

(1) Introduction to Chinese Tones-explanation by texts and tone animations
(2) Pure Tone Perception–high bandwidth (10 questions)
(3) Pure Tone Perception– low pitch (10 questions)
(4) Pure Tone Perception– medium pitch (10 questions)
(5) Pure Tone Perception– high pitch (10 questions)
(6) Male voice perception of tone– Hai (10 questions)
(7) Male voice perception of tone– Ha, I (12 questions)
(8) Female voice perception of tone– Hai (10 questions)
(9) Female voice perception of tone– Ha, I (12 questions)
(10) Ability test (20 questions)

The task voice can be replayed to promote active learning of learners. In addition to the above-mentioned tone perception questions (104 questions), there are 15 rule-attached tone perception questions and 20 single syllable tone

perception questions added for supplementary training. For some questions, the four-syllable voice template animation of the tone presented as a tip.

Furthermore, the tasks for middle or advanced learners were created, such as monosyllabic words (20 questions), four-character idioms consisted of all four kinds of tones (10 questions) and four-character idioms (28 questions). The tasks aim to let learners have experience dealing with polysyllable perception questions. The four-character idioms were considered to make learners get training to deal with tone in the process of learning Chinese words, phrases, conversation, etc. The tone perceptual training questions designed as education contents are totally 197 questions.

Chapter 3

Comparison of Effectiveness of Applications to Improve Chinese Tone Pronunciation Ability

To verify the effect of the perceptual training with tasks designed in Chapter 2, we compared tone pronunciation abilities before and after using our perceptual training application with an imitation practice application. We conducted two experiments on February and November 2018 (Appendix 10), one of the experiments carried out a pronunciation pre-test for grouping the participants based on the tone ability, the other experiment did not carry out the pronunciation pre-test so that participants would not know the contents of the pronunciation test.

3.1. Experimental Method

We gathered 20 Japanese college students who had never learned Chinese and divided them into 4 groups (depends on their age and gender, or the results of the pre-test) and to measure the learning effect of the applications. Each group used the application with different learning methods or experimental methods. We considered that practice every day in a short period is more effective for acquisition of tone, and the procedure of the experiment we designed was as follows:

(0) Pre-test: Basic pronunciation ability test after listening to the tone template model voice three times

(1) Acquisition of knowledge

(2) Training with the app (first session- 3 days, 30 minutes per day)

(3) Post-test: Basic pronunciation ability test with the same contents as the pre-test

(4) Training with the app (second session- 3 days, 30 minutes per day))

(5) Final-test: Pronunciation ability test

(6) Questionnaire

3.1.1. Acquisition of Knowledge

The participants of both groups learned the basic knowledge about Chinese pronunciation before using the application. The contents explained the importance of the Chinese tone. The image of tone was expressed by Yuen Ren Chao's five-level tone mark [16] (Figure 2-2), and the phonological features of each tone were used. Specifically, the different impressions of the third tone and the neutral tone were introduced.

3.1.2. Application to Compare with

The target of the perceptual training application consisted of the perceptual training questions designed in Chapter 2 to compare with was the Chinese pronunciation learning tool "声調確認くん" (Seicho Kakunin Kun) provided by NHK (Japan Broadcasting Corporation), which includes a function to display model voices of Chinese words, phrases, and sentences. The users of the NHK's tone learning application can record and play their pronunciation, track voice pitch, and compare to the model voice. The participants can practice pronunciation by imitation for a range of more than 160 model voices that include monosyllabic, disyllabic and polysyllabic words, four-character idioms and

simple sentences ("你好", "学以致用 ", "我是大学生", "你为什么学日语", "能帮我把我的行李放上去吗", etc.).

Figure 3-1. NHK's tone learning application.

3.1.3. Training with the Application

The period of the training for each session is three days. On two sessions, a total of six days, the participants were told that they must use the app for 30 minutes a day. After a session, they took a pronunciation test. The participants were instructed to meet the schedule for the tone pronunciation test, and the purpose was to master tone.

3.1.4. Pronunciation Test

The contents of the test were designed focusing on polysyllabic pronunciation, correspondence ability of phonemes and tone combination change to examine the tone pronunciation ability. The basic pronunciation tasks designed is considered that help learners practice pronouncing the acoustic feature of tone.

(1) Template: Present the phoneme (e.g., Hai) and the tone images (Figure 2-2) and request the pronunciation of the application users.
(2) Decomposition of the template: Present the phoneme and the tone images of the four kinds of tones individually and request the pronunciation of the application users.
(3) Reordering of template: Present a moving image to change the order of the template tones and request the pronunciation of the application users (a total of 96 syllables; 24 patterns).
(4) Two-syllable Phoneme Combination Change: Present two phonemes (e.g., Ha, Si) and tone images, and request the pronunciation of the application users (a total of 60 syllables; 15 patterns * 2 kinds of phoneme combination) to examine the ability to deal with phoneme change.
(5) Phrase repeating: Ask the applications users to imitate the model voice, and then present the reading of phrases (e.g., Ni Hao, Huan Ying Guan Lin), images of tone, and the model voice (12 syllables on the pre-test and post-test).
(6) Phrase: Present the reading of phrases (e.g., Ni Hao, Huan Ying Guan Lin) and images of tone. And then request the pronunciation of the application users (12 syllables on pre- and post-test; 12 or 24 syllables on final-test).

Scoring for tone pronunciation is based on discrimination by Chinese native speaker using five-grade evaluation. (1) Mistake: a wrong tone, (2) Unknown: a tone that Chinese-speaking listeners cannot recognize, (3) Recognizable: Chinese-speaking listeners can judge the pronunciation as the required tones, but it is an extremely unnatural pronunciation, (4) Easy to understand: Chinese-speaking listeners will judge the pronunciation as the required tones, and will not cause information transmission error in conversation, but the pronunciation may be

flawed in syllable length, pitch, or pitch change, (5) Can be used as an educational model: Chinese-speaking listeners could not find the flaw in the tone pronunciation.

1 or 2 points of pronunciation may cause considerable information transmission error in conversation. 3 points are unnatural pronunciation, and information transmission errors may occur. 4 or 5 points can be considered as having no possibility of information transmission error in the conversation.

3.1.5. Questionnaire

We use a user questionnaire to investigate the demographic data of the participants, such as age, language, comments on application, application usage status and the pronunciation task, and impression of Chinese, tone learning, and application.

3.2. Experiment Results

3.2.1. Application Usage Status

The participants of the groups did not carry out the pronunciation pre-test, two imitation practice application users practiced the 18 phrases 5 to 6 times. The other three imitation practice application users are actively practicing different phrases (totally 132, 161 and 186 phrases). These five participants used 62.6 items on average in the first session (1.5 hours), 76.8 items on average in the second session, and an average of 139.4 items for a total of 3 hours in six days.

The number of items (perceptual questions) used in the perceptual training app designed in this study was 54, 84, 115, 148 and 168 items for the first three days (first session), and due to the possibility of user records of the second session were missed, reliable data cannot be provided. According to the results, the number of items used are vary widely among participants.

In additions, the participants of the groups carried out the pronunciation pre-test, the number of phrases used by the users of the imitation practice app was 32, 54, 69, 94 and 100 for the first session, and 71, 86, 109, 139 and 163 for the second session. A participant chose to practice phrases repeatedly each day.

The number of items (perception questions) used by the users of the perceptual training app was 64, 118, 162, 166 and 212 for the first session and 48, 112, 183 and 212 for the second session (excluding the participant who did not use the application).

The usage status of apps is showed in Figure 3-2. As shown in Figure 3-2, the imitation practice app users w/o taking the pronunciation pre-test practiced a total of 69.8 and 62.6 phrases on average for three days (first session), about 23.3 and 20.9 phrases a day (30 minutes).

The total number of items used of the perceptual training app users w/o taking the pronunciation pre-test were 144.4 and 113.8 items over three days, about 48.1 and 37.9 items a day.

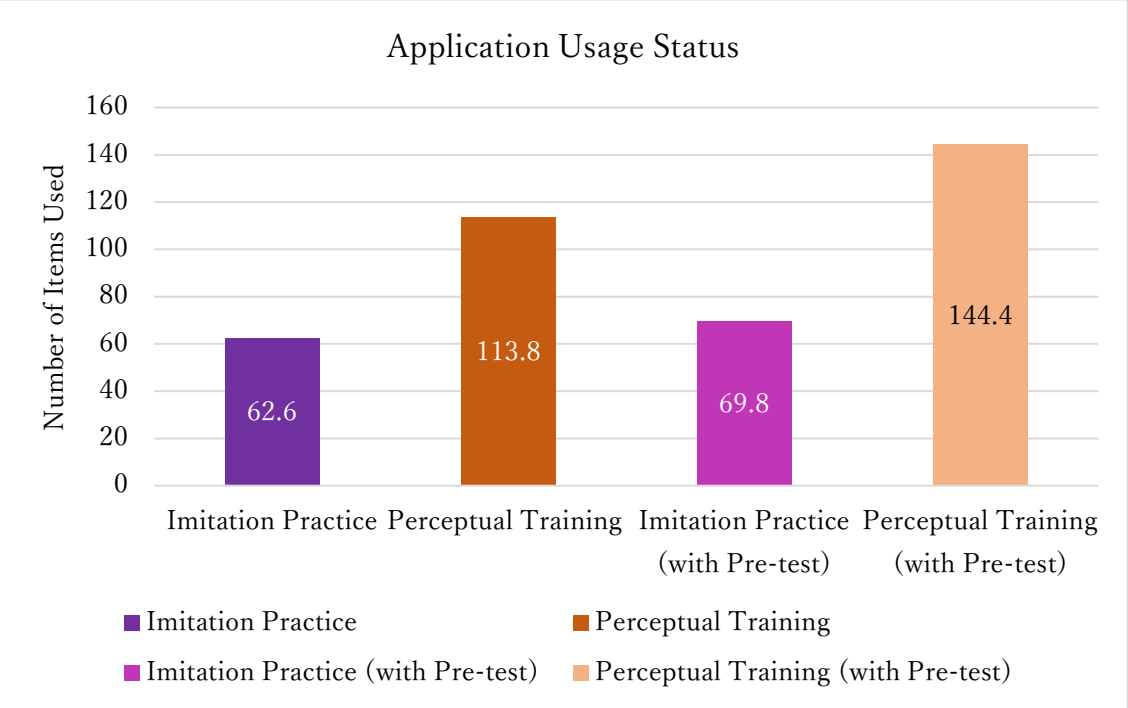


Figure 3-2. Usage status of different applications and experimental design.

3.2.2. Pronunciation Test Results

Four groups of participants trained with different learning apps and w/o took a pronunciation pre-test. Tone pronunciation without information transmission errors in the conversation was regarded as a correct pronunciation. The average correct answer rate and variance of two basic pronunciation tasks (reordering of four tones and two-syllable phoneme combination change) of different groups are presented in Figure 3-3 and Figure 3-4. The average correct question-answer rate of two-syllable phoneme combination change is presented in Figure 3-5.

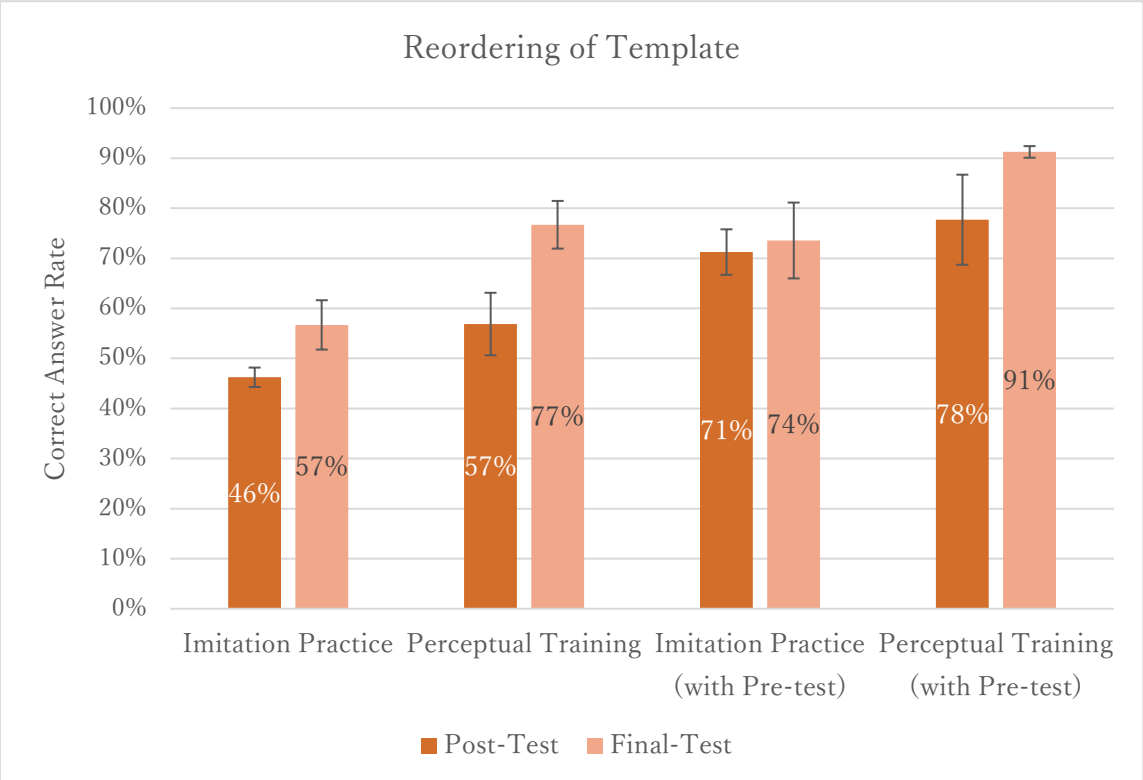


Figure 3-3. Correct answer rate of four syllable pronunciation task of four groups of participants.

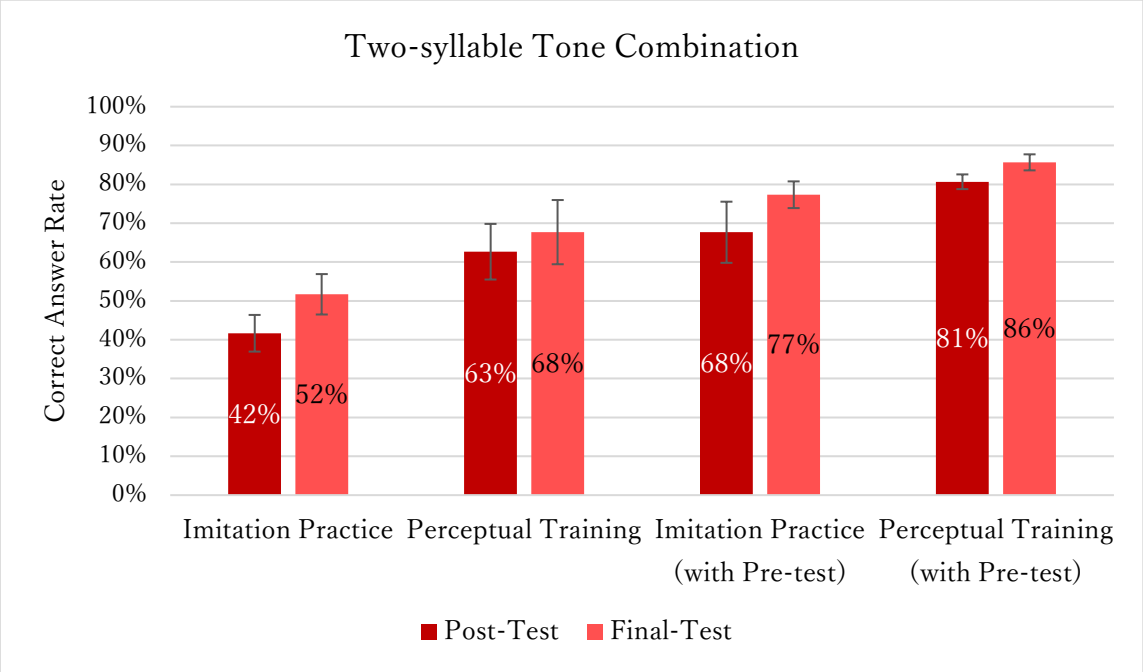


Figure 3-4. Correct answer rate of two syllable pronunciation task of four groups of participants.

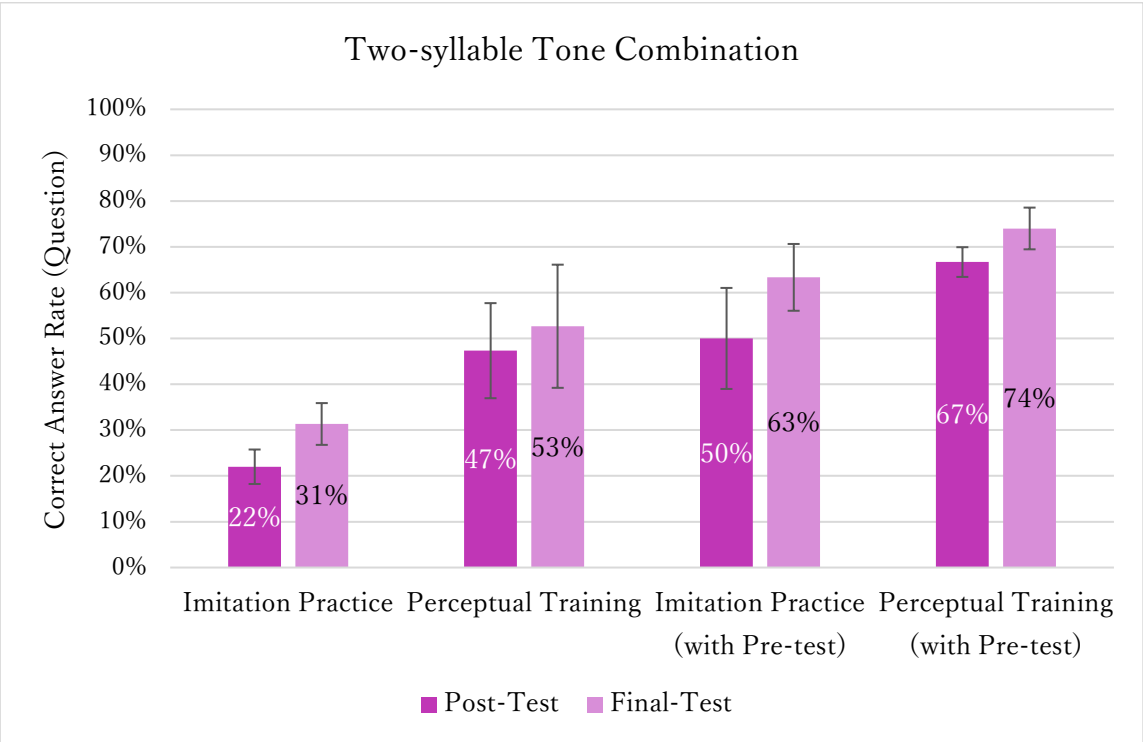
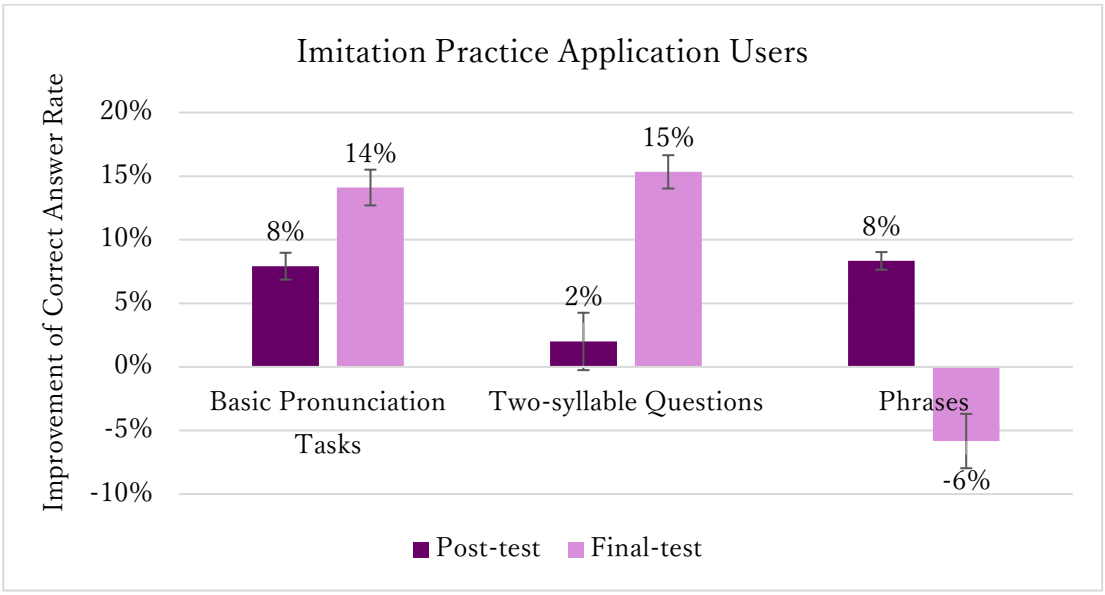
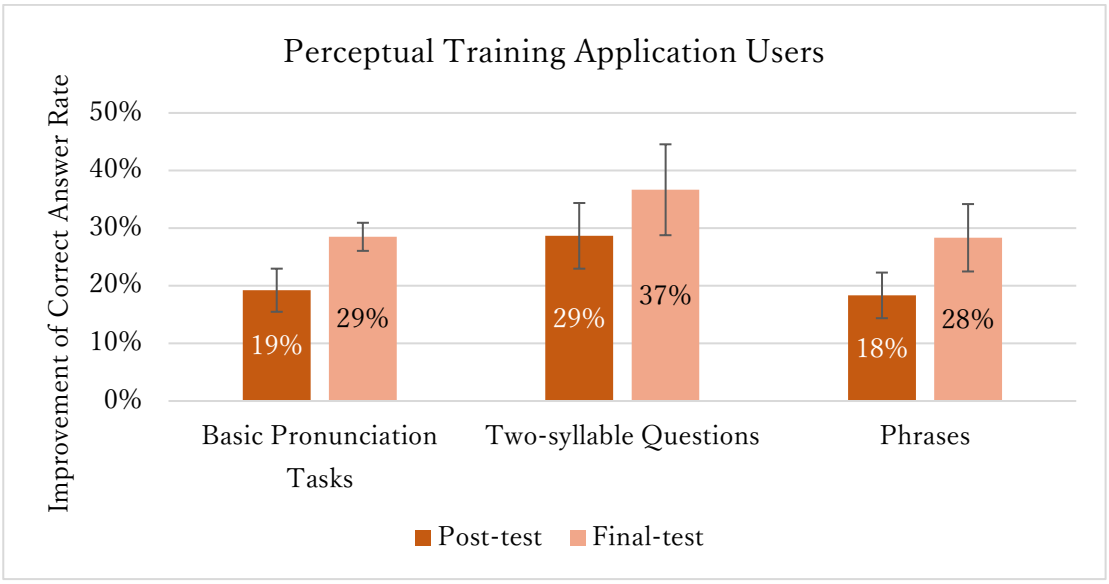


Figure 3-5. Correct question answer rate of two syllable pronunciation task of four groups of participants.

To further distinguish tone pronunciation ability, we regarded the grades of the two-syllable phoneme combination change task as the correct question-answer rate of two-syllable words as related works. The improvement of the correct answer rate that was obtained by calculating the difference compared to the pre-test is shown in Figure 3-6, including two basic tasks, two-syllable questions, and phrase pronunciation.



(a)



(b)

Figure 3-6. Improvement of correct answer rate of (a) Imitation practice application users; (b) Perceptual training application users.

As shown in Figure 3-6, the perceptual training app users' improvement regarding the correct answer rate was more than twice that of the imitation practice app users. For the phrase pronunciation task, the correct answer rate of imitation practice app users declined due to not doing the repeating task beforehand. However, the correct answer rate of the perceptual training app users improved even final-test that was not included repeating task.

Table 3-2 shows the results of the t-test, which included the correct answer rate of four pronunciation tasks of imitation practice and perceptual training app users before and after training with the app. There were a statistically significant difference and a large effect in the correct answer rate of the reordering of four tones task and two-syllable questions from learning with the imitation practice app and the perceptual training app ($p < 0.05$). However, for the phrase pronunciation task, only the correct answer rate of the perceptual training app users met the significance level of 0.05.

Table 3-2. T-Test between correct answer rate of imitation practice and perceptual training app users before and after training with app.

Users	Paired Samples T-Test (Correct answer rate of pronunciation tasks: second test and pre-test)	df	p	Cohen's d
Imitation Practice App	Reordering of Four Tones	4	0.048	1.26
	Two-syllable Phoneme Combination Change		0.178	0.730
	Two-syllable Questions		0.040	1.34
	Phrases		0.423	-0.399
Perceptual Training App	Reordering of Four Tones		0.022	1.64
	Two-syllable Phoneme Combination Change		0.057	1.18
	Two-syllable Questions		0.043	1.31
	Phrases		0.025	1.57

Pronunciation tasks that include more complicated phonemes are expected to be advantageous for application users focusing on phrase repeating exercises.

However, results exceeding expectations have come out, practicing phrase pronunciation in such a state that the basic practice is deficient is not efficient to improve the ability to pronounce a phrase that has never been heard.

Table 3-3 shows the results of the t-test, which included correct answer rate of imitation practice and perceptual training app users w/o taking a pre-test before training with the app. There was a statistically significant difference in the correct answer rate of the two-syllable questions from learning with the imitation practice app and the perceptual training app or taking and not taking a pre-test before training with app. ($p < 0.05$). However, for the two basic pronunciation tasks, the correct answer rate of the imitation practice and perceptual training app users taking a pre-test did not meet the significance level of 0.05.

Table 3-3. T-Test between correct answer rate of imitation practice and perceptual training app users w/o taking a pre-test before training with app.

Pronunciation Tasks	Paired Samples T-Test		df	p	Cohen's d
Reordering of four tones	Imitation Practice App with pre-test	Imitation Practice App	9	**<0.001**	**2.018**
	Perceptual Training App with pre-test	Perceptual Training App		**0.017**	**0.918**
	Perceptual Training App	Imitation Practice App		**0.026**	**0.839**
	Perceptual Training App with pre-test	Imitation Practice App with pre-test		0.167	0.476
Two-syllable Phoneme Combination Change	Imitation Practice App with pre-test	Imitation Practice App		**<0.001**	**2.135**
	Perceptual Training App with pre-test	Perceptual Training App		**0.036**	**0.777**
	Perceptual Training App	Imitation Practice App		**0.005**	**1.164**
	Perceptual Training App with pre-test	Imitation Practice App with pre-test		0.116	0.550
Two-syllable Questions	Imitation Practice App with pre-test	Imitation Practice App		**<0.001**	**2.001**
	Perceptual Training App with pre-test	Perceptual Training App		**0.005**	**1.150**
	Perceptual Training App	Imitation Practice App		**0.004**	**1.208**
	Perceptual Training App with pre-test	Imitation Practice App with pre-test		**0.021**	**0.879**

Taking a basic tone pronunciation test before training with the app would improve tone pronunciation ability and make the perceptual training has no significant impact on improving the correct answer rate of two basic pronunciation tasks. However, for the pronunciation tasks with relatively high discrimination, there was still a statistically significant difference in the correct answer rate of different apps users.

3.2.3. Results of Questionnaire

Table 3-4 shows the average score of the nine questionnaire items related to tone learning using a 6-point Likert scale. As shown in Table 3-4, users of the imitation practice app used the application more thoroughly and were more focused on learning tone.

Table 3-4. Questionnaire Results.

Item/Average Score	**Imitation Practice App**	**Perceptual Training App**
I focus on learning the tone	**5.2**	4.5
(On first session) I use the application thoroughly	**5.4**	4.6
(On second session) I use the application thoroughly	**5.7**	4.4
This application helps me to understand tone	**5.4**	5
The pronunciation test helps to train tone ability	5.1	4.8
Tone learning is interesting with this application	4.5	4.3
The learning method of this application is troublesome	2.8	3
Tone perception is difficult	4.5	4.6
Tone pronunciation is difficult	4.8	5

According to the results of the questionnaire on each training task in perceptual training apps (Table 3-5), which is useful for understanding the tone, the process that seems to be the most useful is decreasing the speed (5.6 points), followed by the male/female voice change (5.5 points), four tones in four syllables (5.4 points) and change phoneme (5.3 points). These are the same methods used in the survey experiments to help Japanese native speakers who are learning Chinese to recognize tone, as seen in tone perceptual experiment with rules. However, the textual explanation and explanation by animations get the lowest scores (4.2, 4.3 points). It is considered that we should change the way to provide the explanation, especially the animations that aimed to emphasize the visual symbols for helping learners memorize the features of tones.

Table 3-5. Questionnaire results of items design in perceptual training app.

Item that helps to understand the tone	Average Score
The Perceptual Training App	5
Textual Explanation	4.2
Explanation by Animations	4.3
Pure Tone	4.9
Decrease the Speed	**5.6**
Fixing the Phoneme	5.2
Four Tones in Four Syllables	**5.4**
Pure Tone Different Pitch Range	4.9
Male/Female Voice Change	**5.5**
Change Phoneme	**5.3**

Table 3-6 shows the average score of functions in imitation practice app. The items that get the lowest scores were male/female voice provision (4 points) and the record/play function (4.9 points). Compare to the questionnaire results of items design in perceptual training app, the score of male/female voice in the application got the exact opposite results.

Table 3-6. Questionnaire results of functions in imitation practice app.

Item that helps to understand the tone	Average Score
The Imitation Practice App	5.4
Listen to the Phrases in the App	5.5
Imitate the Phrase Pronunciations Provided in the App	5.5
Record the Pronunciation Practice and Play it	4.9
Track the Voice Pitch and Provide the Image of it	5.4
Male/Female Voice Provision	4

3.2.4. Error Analysis of Pronunciation Test

We investigated the tendencies of users of different applications using the pronunciation test data of the evaluation experiment.

Figure 3-7 shows the error rate of each tone in two basic tasks. As shown in Figure 3-7, the result of the error analysis of four tones was 3 (third tone), 2 (second tone), 1 (first tone), and 4 (fourth tone) in descending order of error rate. However, before training with the application, the error analysis results of the four-syllable task, reordering of four tones, was 3, 1, 2, 4. The error rate of the first tone was about twice that of the second tone. From these results, we concluded that the first tone (Figure 2-1), which is high and flat, is more likely to cause errors than other tones when there are a large number of syllables to pronounce. However, the two application users' error rates of the first tone in two pronunciation tasks are obviously decreased after training.

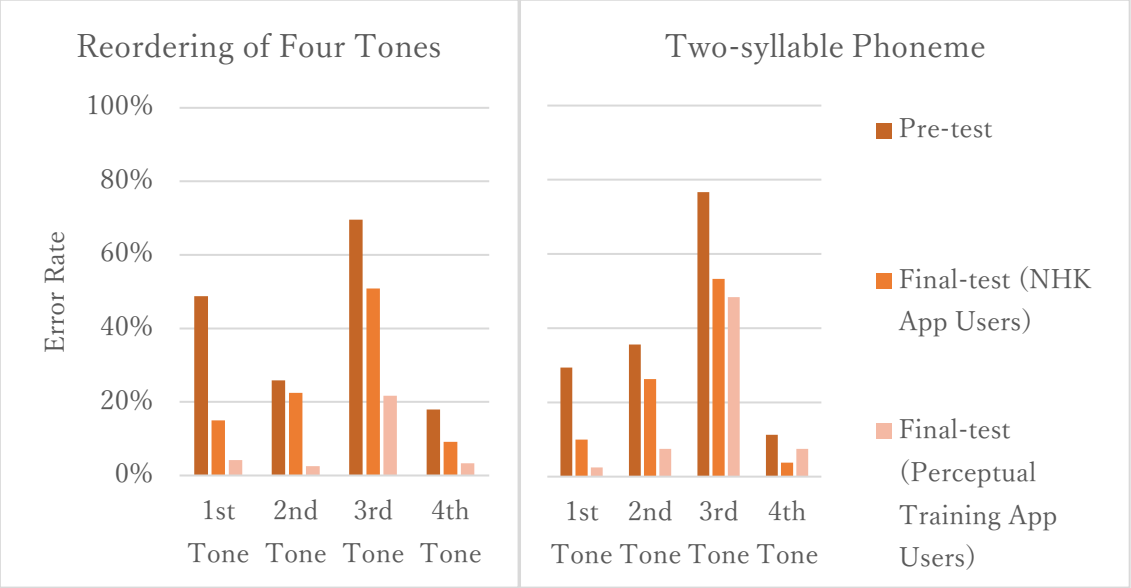


Figure 3-7. Error rate of four tones in two basic pronunciation tasks before and after training with app.

Furthermore, it is important to pronounce the third tone in a lower voice compared to speaking Japanese; thus, even when participants can recognize the relative pitch height, there is a high possibility that a pronunciation error will occur. We speculated that feedback on pronunciation is more important for the pronunciation of the third tone.

In addition to the error analysis of the fourth tone, we considered tone combinations of the two-syllable and four-syllable tasks. The error rate of two-syllable combinations is shown in Figure 3-8. As shown in Figure 3-8, there is no change in the tendency of errors before and after training, and the combination including the third voice is relatively difficult to improve by training, especially in 4+3, 3+2, and 3+4. The error rate did not go below 50% on average for both groups.

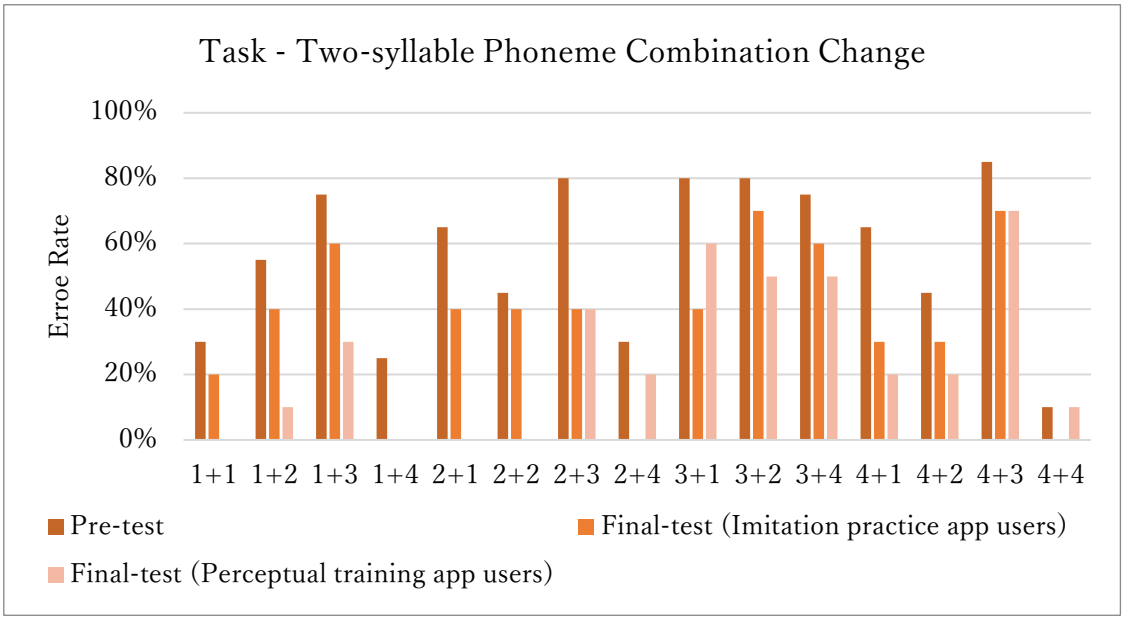


Figure 3-8. Error rate of tone error patterns in two syllable pronunciation tasks.

The error rates of tone combinations in the four-syllable task (a total of 24 patterns) are shown in Figure 3-9. According to Figure 3-9, the tone combination with the lowest error rate before training is "1234," which was presented three times before the pre-test. However, the error rate of the tone combination "1234" did not decrease after training and had the highest error rate in the final-test. Furthermore, before training, five tone combinations had relatively low error rates (the lowest and the second-lowest), but only two of them were in the same position after training.

As a result of the error analysis, we have concluded that it is relatively difficult to pronounce the combination "1234" used in the experiment as a template of tone in the pre-test and animation explanation. In addition, except for the tone combination "1234," the combination "1423" had the lowest error rate before and after training. Furthermore, there are participants that tend to pronounce the third tone as the fourth tone because of the image of the third tone, the "Low-dipping tone" (Figure 2-2), used in the experiments. They tried to express "Low-dipping", however, it sounds the same as "High-falling".

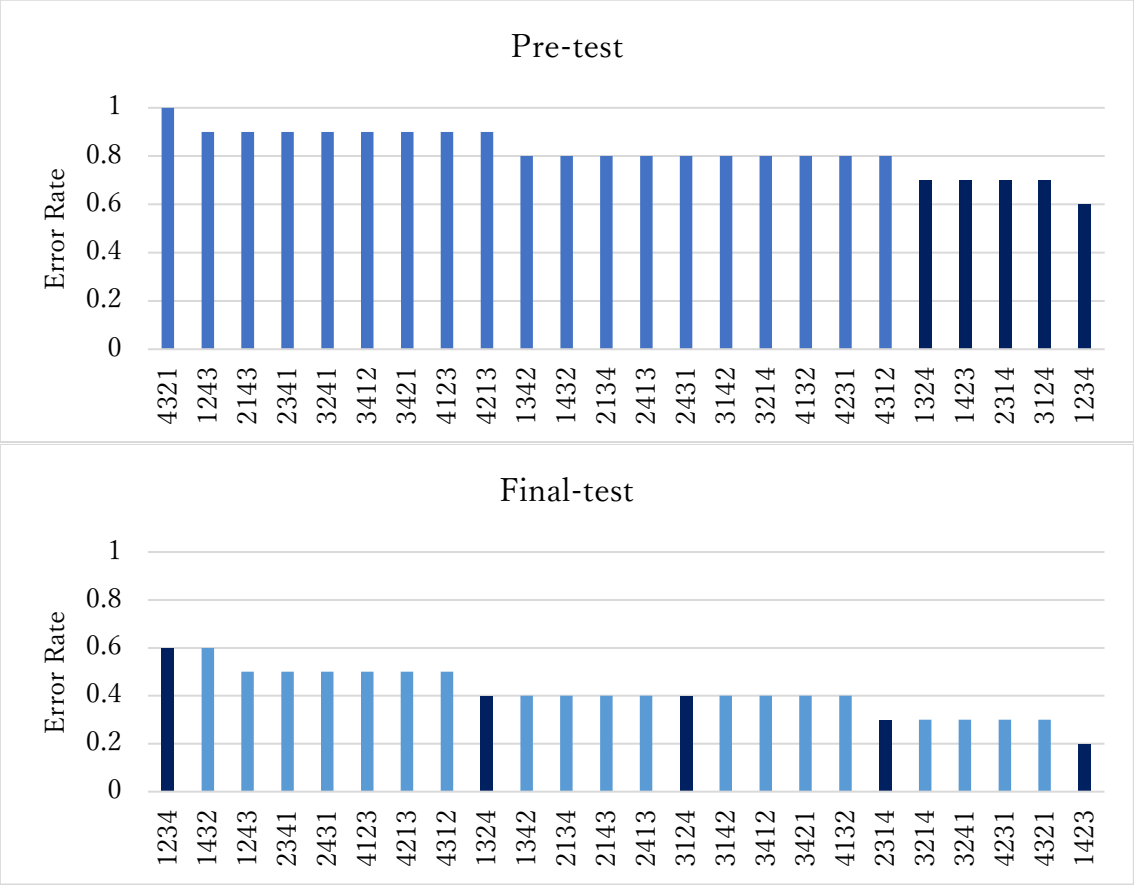


Figure 3-9. Error rate of tone error patterns in four syllable pronunciation tasks. Dark blue in the figure shows the tone combinations that participants had higher performance in the pre-test.

3.3. Discussion

According to the pronunciation test results of the evaluation experiment, perceptual training instead of phrases imitation practice, taking the basic pronunciation pre-test, and the optimization of the information presentation and operation of the perceptual training application are helpful to improve the application users' the correct answer rate of tone pronunciation.

In the usage status of the applications, the users of the perceptual training application listened to more model voice than the imitation practice application users. Also, taking the basic pronunciation pre-test have a positive impact on the number of items used of the application. We consider that in addition to the

designed perceptual questions of the perceptual training application, making the users concentrate on perception is also an important factor in improving the learning effect. We speculate that the reason why the users has taken the pronunciation pre-test had a higher tone pronunciation ability and the increase in the number of items used is because of the preview effect and the tone pronunciation preview increased learning motivation of the users.

The error that pronounce the third tone like the fourth tone also happened to two imitation application users. This error caused these users have an error rate of over 25% at the reordering of four tones task, and this error is slightly less at the two-syllable task and phrase pronunciation task. Furthermore, there is an imitation application user pronounce all four kinds of tone as fourth tone at two basic tone pronunciation tasks in the post-test and the final-test even if he could imitate the phrases pronunciation pretty well (Correct answer rate of 100% in the post-test, 92.3% in the final-test). However, his correct answer rate of the phrase pronunciation task he took right after the imitation practice increased from 30% in the post-test to 61.5% in the final-test. We speculate that he didn't actually learn how to pronounce the tone or realize the feature and difference between each tone. He only improved his imitation and memory ability of Chinese pronunciation by imitation practice application. In addition, the results of improvement of phrases pronunciation (Figure 3-7) show that faced to the phrases never heard, the correct answer rate of imitation practice app users declined. On the contrary, the correct answer rate of perceptual training app users kept increasing after training, even if they didn't do the repeating task beforehand.

Chapter 4

Optimization of the Perceptual Training Application

4.1. Introduction

In Chapter 3, we showed that our perceptual training app is more effective than the existing imitation pronunciation practice app. However, as a prototype of the perceptual training app to verify the effect of the educational content designed, there are some issues about teaching material design still remain, such as 1. The way to show the educational contents, 2. The time required to load animations and questions, 3. The operations of the perceptual training application required to get explanations of tone and train through the perception questions, 4. Combine the perceptual training with pronunciation practice.

4.1.1 Extraneous Cognitive Load (Perceptual, Visual, and Operation Load)

The intrinsic cognitive load of phrase pronunciation practice seems to be too high for Chinese beginners described in the previous chapter, though it can be reduced by having preliminary knowledge. However, information unrelated to tone would lower learners' concentration on tone and affect learning effectiveness. In addition, the pitch tracking function often used in teaching materials and pronunciation practice applications takes time to record, analyze,

and confirm, which may reduce the pace of learning and increase cognitive load. The results of audio analysis could not faithfully reproduce what humans feel and may not be effective for learning.

Using teaching materials and applications for learning can be further classified into three categories: perceptual load, visual load, and operation load. Regarding the perceptual load, we set the learning target only to tone recognition, and the audio task was designed based on the learner's cognitive ability. The audio task was provided as a self-learning material in the order of task difficulty. To minimize the visual load of learning contents, only the images of the four tones (Figure 2-1) that show relative pitch patterns of Mandarin tones were used as visual symbols in the perceptual training application. The operation load of using the application is manly related to how to select the perceptual training question sets and answer the perception questions.

It is assumed that the application interface and operation design are related to the perceptual load, visual load, and movement load during learning and affect learning effectiveness. To investigate the effects of interface and operation design, we focused on the burden and efficiency of learning and optimized the interface design, the loading time, and the operation method of the perceptual training app.

4.2. Application Interface and Operation Design

Though the perceptual training application was shown to be more effective than the imitation pronunciation practice application, according to the questionnaire results after the experiment, the users of the NHK's application focused more on learning and found it harder to learn tones. In addition, NHK's app is more interesting for tone learning, and the perceptual training app users felt that it was more troublesome.

Therefore, we optimized the interface design, operation design, and loading time, which aimed to minimize the extraneous cognitive load without changing the intrinsic cognitive load (contents of learning).

4.2.1. The Perceptual Training Application Prototype

To carry out the experiment in Chapter 3, we adopt the platform "啟発くん" to provide the educational contents designed by us. On this platform, we can edit text, pictures, images and test questions, so that users can freely use the content of our designed teaching materials through the Internet. We put our educational contents on this platform as the prototype of tone perceptual training materials.

The main screen and the image of the perception question in the application prototype are shown in Figure 4-1. The perceptual training items are presented as a list. Users have to click on the item and wait for the loading time of the explanation page, the question and the answer page to complete a perception question. Furthermore, after confirming the answer of the question, user have to wait for the loading time again to go back to the question answered or to go to the next question.

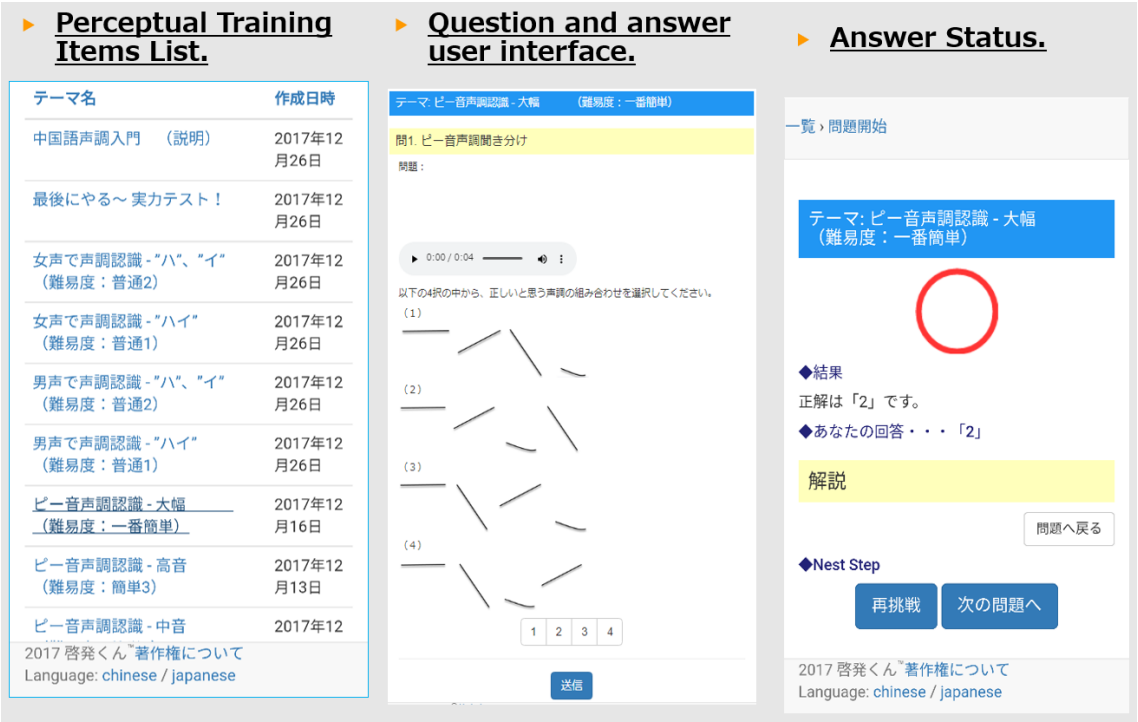


Figure 4-1. Screenshots of the perceptual training application prototype.

According to the related research on brief auditory storage [38], humans' auditory sensory memory can only last for a few seconds. Therefore, we considered that the biggest burden or hindrance for users is the operations of the perceptual training application prototype required to answer the questions. The operations are: 1. Wait for the application to load (0.8-1 second), 2. Click the button to play the sound of the perception question, 3. Click a selection to answer the question, 4. Click the button to send out the answer, 5. Wait for the application to load (0.8-1 second), 6. Confirm the answer of the question and click the button to move to another question.

After choosing and clicking a selection, user have to click a button to send out the answer. Furthermore, from entering a question to getting a feedback from the application, user have to wait for the application to load the question and the answer status. We speculate that in the process of repeatedly listening to the tonal audio and obtaining feedback, the redundant operations and waiting time will

increase the cognitive load, reduce the user's willingness and make it more difficult for the learner to concentrate on the tone perception.

4.2.2. The Optimized Tone Perceptual Training Application

Figure 4-2 shows the main screen and image of the perception question in the optimized tone perceptual training application. We emphasized the category and level of the perception questions, simplified the name of the items, and changed the interface color and the way to play the media files.

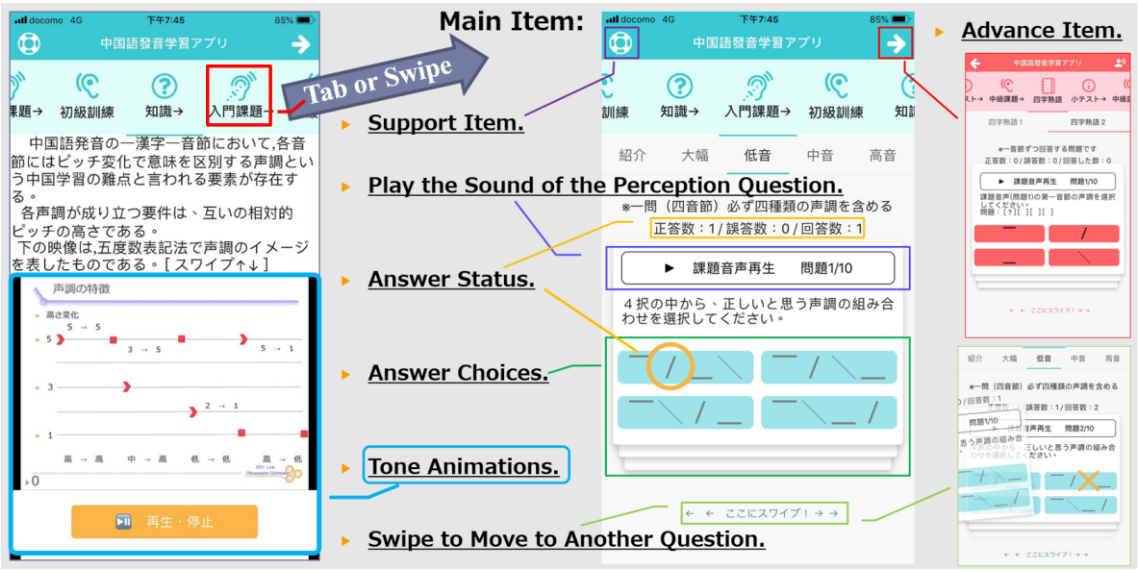


Figure 4-2. Images of the optimized perceptual training application.

The optimized perceptual training application loads several animations and questions in advance. These data will be utilized when users start them. Users do not have to do anything to find the tone animation and they also can switch the item by swiping.

The most important optimization for tone recognition is the effort required to answer the perceptual training questions. Therefore, we eliminated loading time (0.8 to 1 second in the prototype) required to move to each question, task, and confirmation screen of answer status. The operations of the optimized perceptual training application required to answer the questions are: 1. Click the

button to play the sound of the perception question, 2. Click the button to answer the question, 3. Swipe to move to another question.

With the perceptual training app after optimization, it was possible to move to the next question within the length of sensory memory (about 1 second) and challenge the next question. Table 4-1 is a comparison between the application prototype and the optimized version. Obviously, the optimized version has become simpler in operating procedures, screen transitions, and provisions of feedback.

Table 4-1. Comparison of application interface and operation

Item / (time or second)	The Perceptual Training App	
	Prototype	Optimized
The number of operations required to view the animation	2	1
The number of screen transition required for view the animation	1	0
Loading time required for each animation	1	0
The number of operations required to move to another item	2	1
Number of screen transition required to move to another item	2	1
Loading time required to move to another item	2	0
The number of operations required to answer a perception question	6	3
The number of screen transition required to answer a perception question	2	1
Loading time required for each perception question	1	0
Loading time required to show answer status	1	0

4.2.3. Practicing Pronunciation in the Perceptual Training Application

As described in Chapter 3, we conducted the pronunciation pre-test before training with the app, and the correct answer rate after training with the application was significantly improved by about 20%. It is considered that do the

basic pronunciation pre-test after listening to the tone template model voice is effective to improve the tone ability of application users.

Therefore, to further verify the effect of basic pronunciation pre-exercises, we added a page for instructing the tone pronunciation pre-exercise in the application (Figure 4-3). In the pronunciation pre-exercise page, light practicing content (12 syllables) was used to support the perceptual training after enabling the app. Users will view the pronunciation pre-exercise page that provides the template model animation and request the pronunciation of the application users right after they start the perceptual training application. Users have to click the button at the end of the pronunciation pre-exercise page to go to the main screen of the perceptual training application. In addition, for the users who can quickly solve the perception questions, the basic pronunciation tasks and the model voice of the four tones were presented at the end of the perceptual training questions for users to practice pronunciation (Figure 4-3).

(a) (b)

Figure 4-3. The pronunciation pre-exercise page in the perceptual training application. (a) Pre-exercise; (b) Extended training.

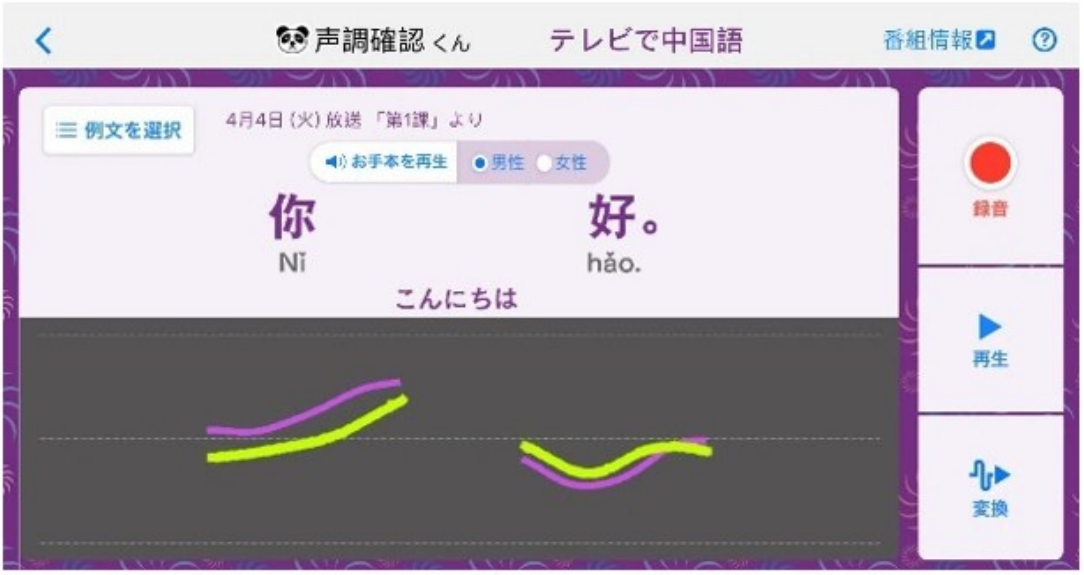
4.3. Effectiveness of Optimized Applications to Improve Chinese Tone Pronunciation Ability

To confirm the effectiveness of the optimized perceptual training, we compared tone pronunciation abilities of perceptual training application users before and after optimizing. We conducted the experiment on June 2019 (Appendix 11).

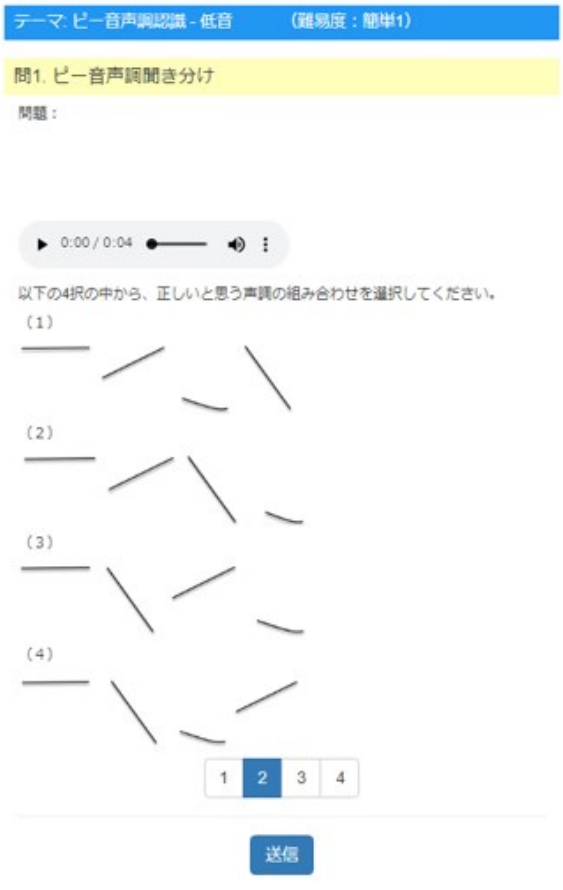
4.3.1. Experimental Method

As the experiments conducted in Chapter 3, we divided 10 Japanese college students who had never learned Chinese into 2 groups. Groups used the perceptual training application with the pronunciation pre-exercise page, and the other used the perceptual training application without the pronunciation pre-exercise page. After training with the application, they took pronunciation tests with several pronunciation tasks as the experiment in Chapter 3.

We compared the users' tone pronunciation ability after using four kinds of applications abovementioned, (a) NHK's app (the imitation practice app), (b) prototype app (the perceptual training app prototype), (c) Optimized app (optimized perceptual training app) with pronunciation pre-exercise page, (d) Optimized app without pronunciation pre-exercise page. Figure 4-4 shows the user interface of applications' main function.



(a)



(b)

(c)

Figure 4-4. User interface of the (a) Imitation practice app; (b) Perceptual training app prototype; (c) Optimized perceptual training app for pronunciation or perception practice.

4.3.2. Application Usage Status

The five NHK's app users used 62.6 items on average in the first session (1.5 hours). The number of perceptual questions used in the prototype app was 113.8 on average for the first three days (first session). In additions, the number of perception questions used by the users of the optimized app with pronunciation pre-exercise page was 134 on average for the first session and 132 for the second session. In addition, the optimized app users used 157.6 perception questions on average for the first session and 152.5 for the second session (excluding one of the five participant who did not use the perception questions).

The usage status of apps is showed in Figure 4-5. As shown in Figure 4-5, the number of items used of prototype app is about 1.8 times the number of NHK's app users, which means that it takes more time to complete a phrase pronunciation practice using NHK's app than to answer a perception question in the prototype app.

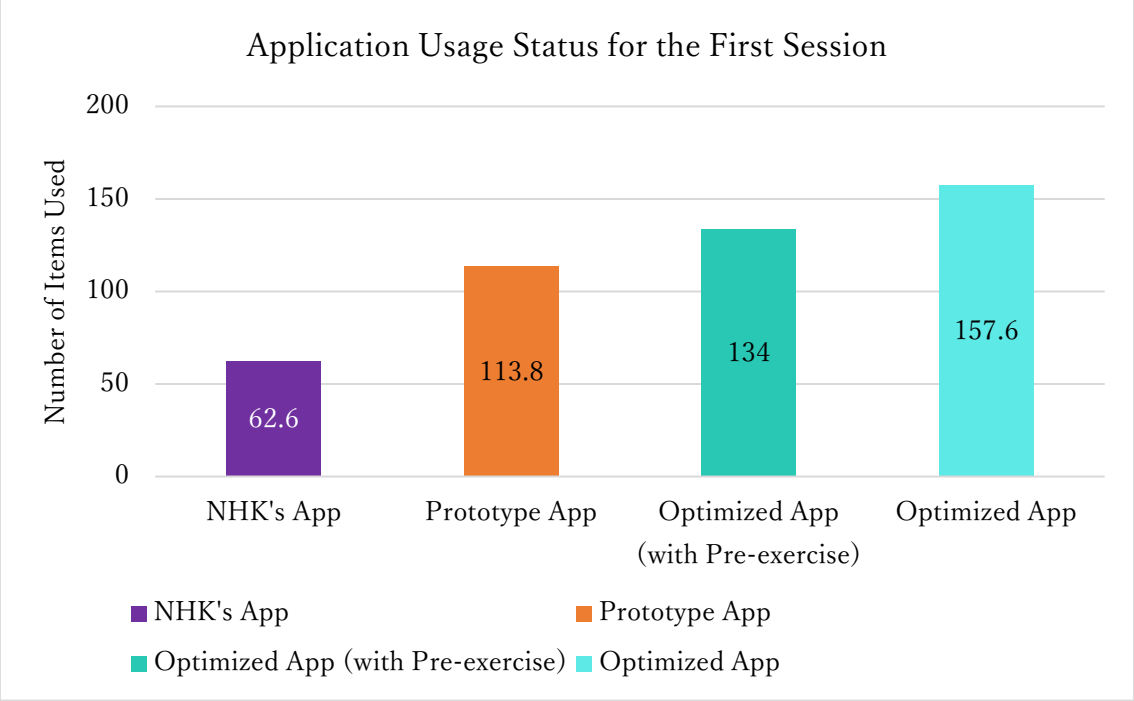


Figure 4-5. Number of items used of four groups of participants.

Furthermore, after optimizing the prototype app, the average number of perception question used is about 1.4 times that of the prototype app. However, the number of perception question used of optimized app with pronunciation pre-exercise page is 0.85 times of optimized app without pronunciation pre-exercise page. The pronunciation pre-exercise page makes users have less time for answering the perception questions.

In addition, for the tone perceptual questions using pure tone produced as an introductory item, the average correct answer rate of each prototype and optimized app users was 73.3% to 100%, and the average correct answer rate of them was 89.4%. For the tone perceptual task, the pure tone had a high correct answer rate even though the students had no previous experience learning Chinese. In the tone perceptual questions that used a human voice instead of a pure tone, a considerable individual difference was confirmed in the tone perceptual questions that used human voice. For this task, the average correct answer rate of each user was 34% to 94%, and they had a correct answer rate 71.7% on average.

The result is similar to the results of the perception experiments in Chapter 2. Even tone perception questions with human voice simplified, some learners still cannot figure out the tones smoothly. There were two prototype app users had lower correct answer rates of 34% and 45% in the three-rule attached perception (the correct answer rates of the other users were 82.3%, 86% and 94%). On the contrary, there were six optimized app users had lower correct answer rates of 38.6%, 60%, 68.2%, 72.7%, 72.7% and 73% in the three-rule attached perception (the correct answer rates of the other users were 88.6%, 93.2% and 100%). However, there were three optimized app users were able to challenge the three-rule attached perception questions again and the correct answer rates of them improved from 38.6%, 60% and 72.7% to 73.8%, 81% and 93.2%,

respectively. We speculate that the optimization of app operations and eliminating loading time required to answer perception questions not only allow users to train with more perception questions within the same time, but also improve the effectiveness of training.

4.3.3. Pronunciation Test Results

Four groups of participants trained with different mobile apps. Tone pronunciation without information transmission errors in the conversation was regarded as a correct pronunciation. The average correct answer rate and variance of two basic pronunciation tasks (reordering of four tones and two-syllable phoneme combination change) and phrase pronunciation task of different groups are presented in Figure 4-6, Figure 4-7 and Figure 4-8.

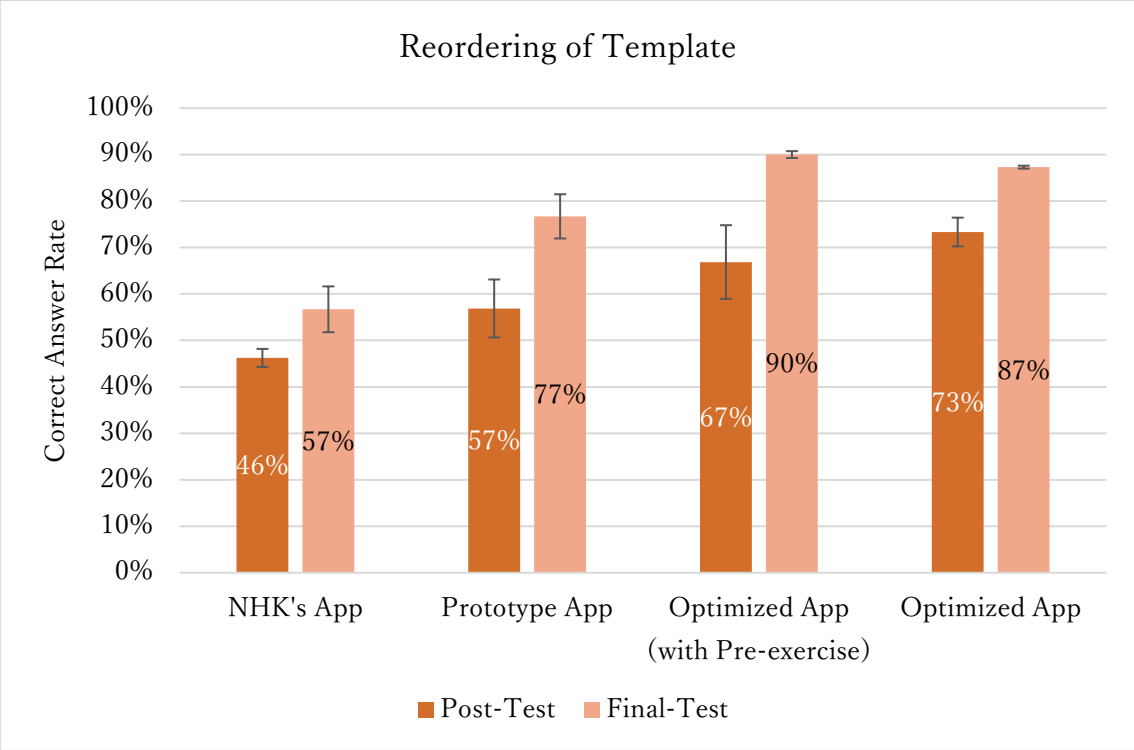


Figure 4-6. Apps users' correct answer rate of four syllable pronunciation task.

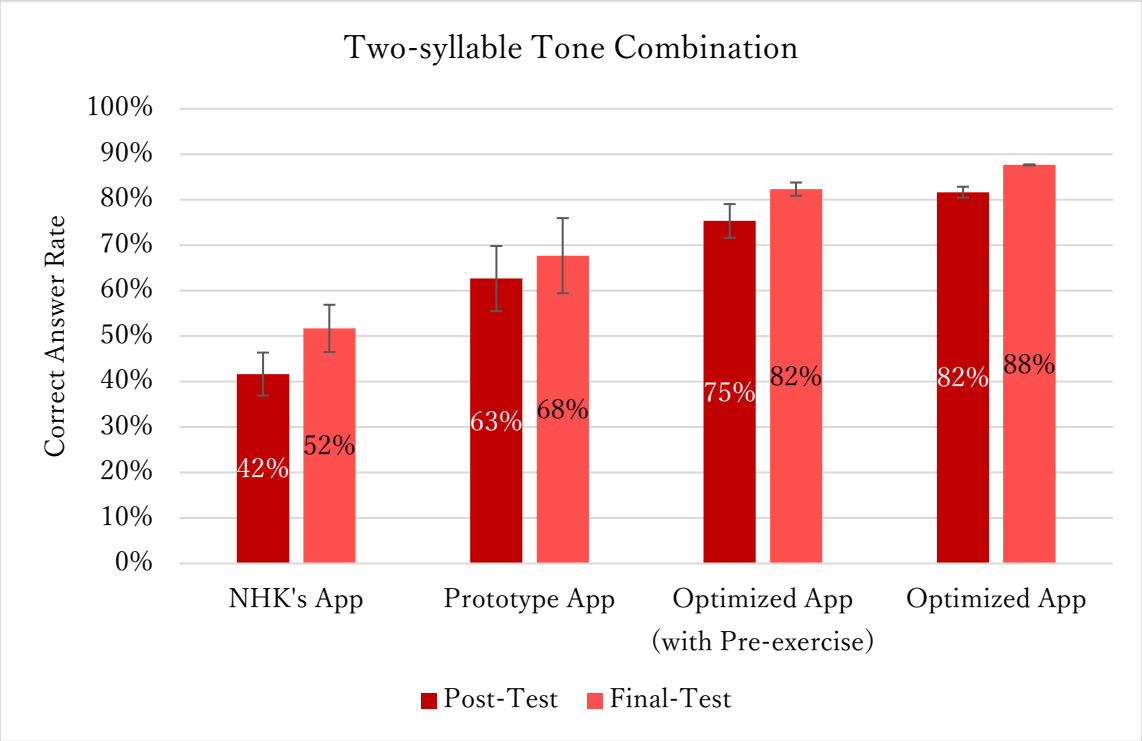


Figure 4-7. Apps users' correct answer rate of two syllable pronunciation task.

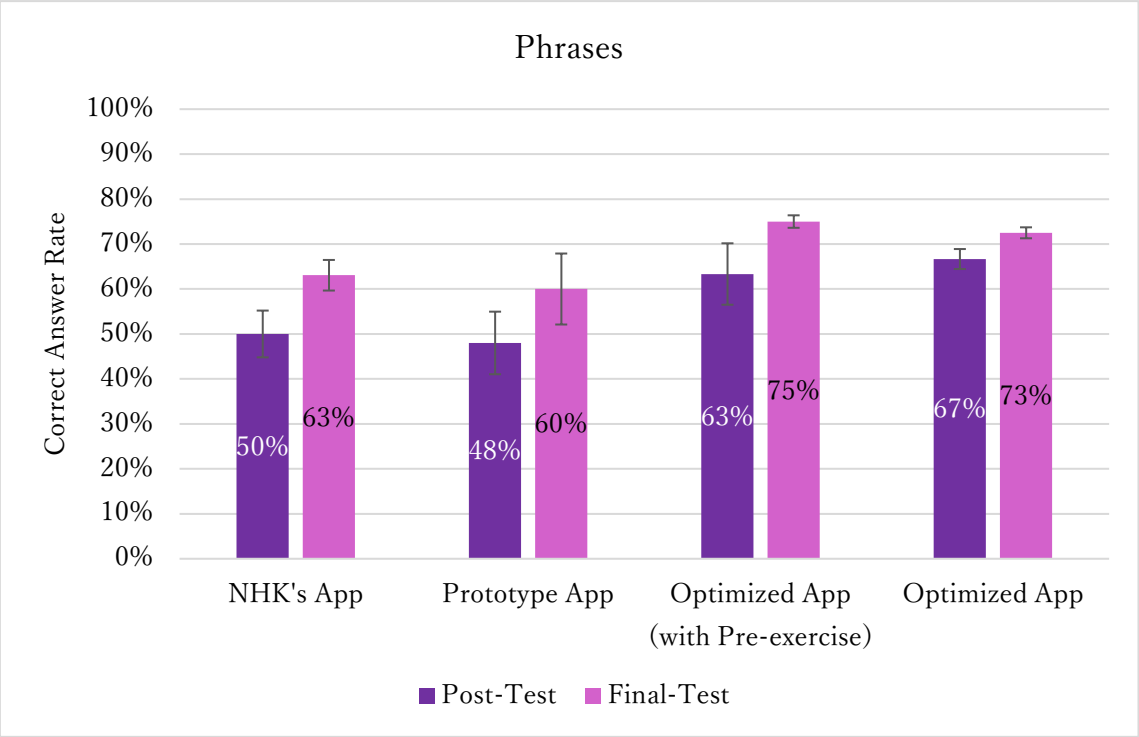


Figure 4-8. Apps users' correct question answer rate of phrase pronunciation task.

The prototype app and optimized app users' average correct answer rates of pronunciation tasks that require tone-based ability (a) four-syllable tone combinations, (b) two-syllable phoneme combination change, are more than 10% higher than those of NHK's app users. However, the results of (c) the phrases pronunciation task performed after imitating showed that the correct rates of the NHK' app and the prototype app users was very similar. Figure 4-8 shows that NHK's app users that focused on phrase pronunciation gives their advantage on phrase pronunciation task. However, even if the phrases pronunciation test is performed after imitating pronunciation practice, the correct answer rate of NHK's app users is still 10% lower than those of optimized app users.

To further distinguish tone pronunciation ability, we regarded the grades of the two-syllable phoneme combination change task as the correct question-answer rate of two-syllable words as related works. The average correct question-answer rate of two-syllable phoneme combination change is presented in Figure 4-9. And the correct question-answer rates of each app user are presented in Figure 4-10 and Figure 4-11.

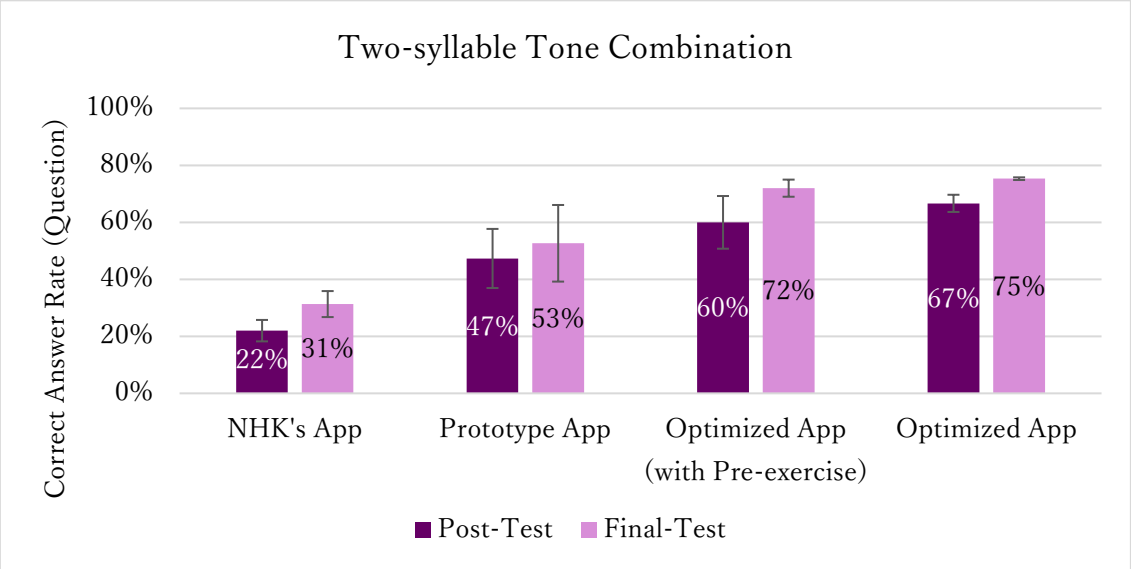
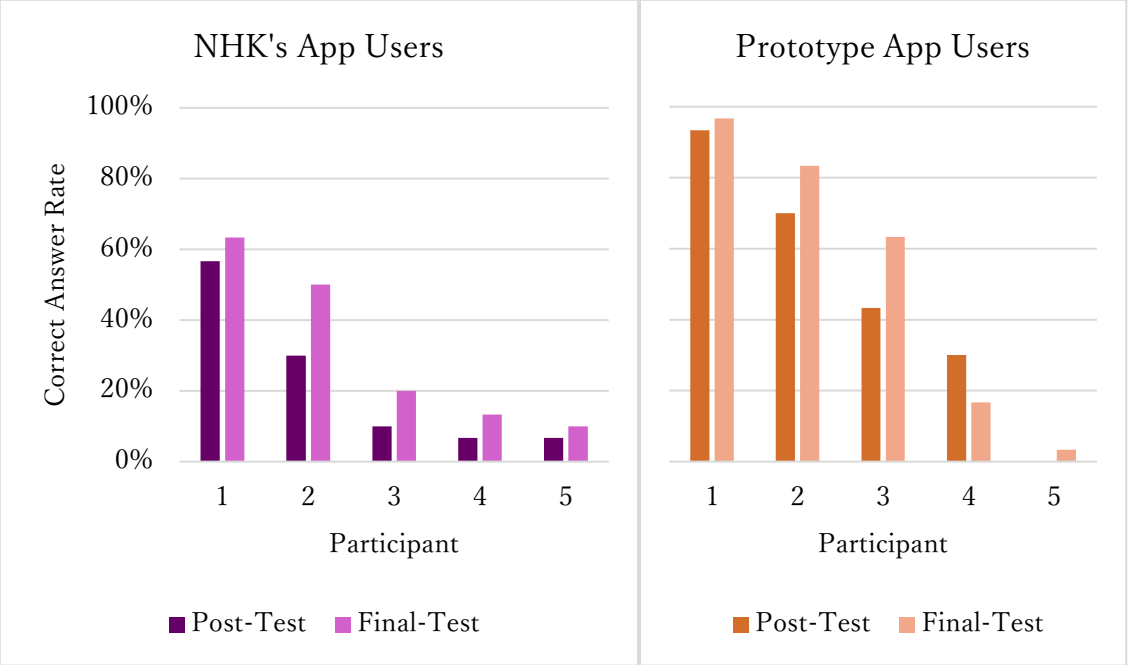
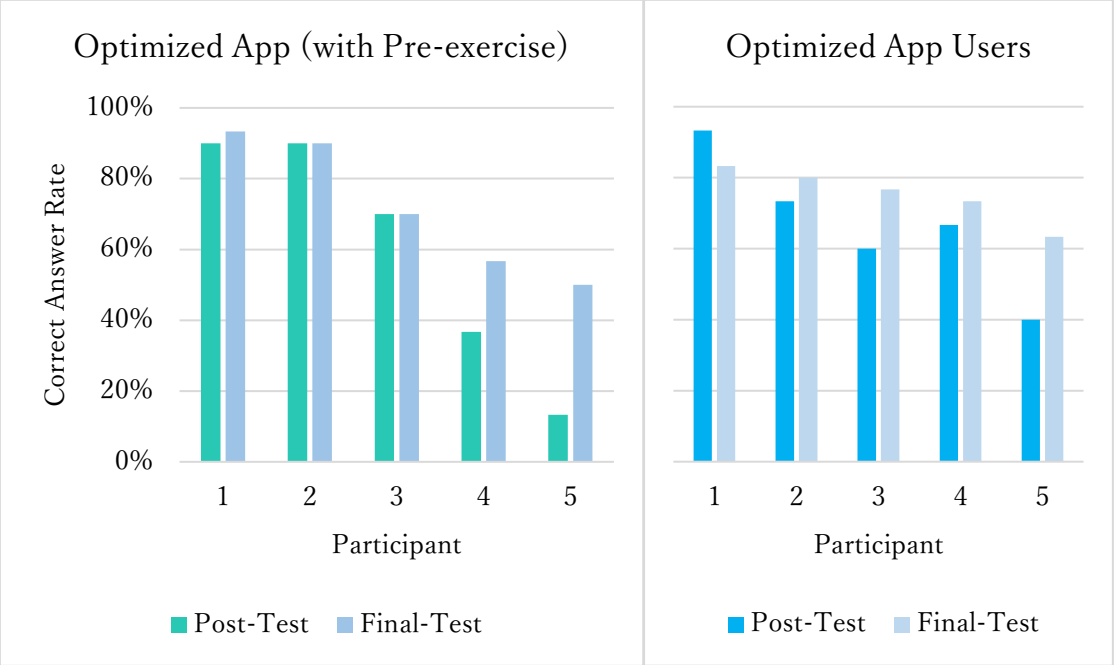


Figure 4-9. Apps users' question correct answer rate of two syllable pronunciation task.



(a) (b)

Figure 4-10. (a) NHK's app; (b) Prototype app users' correct question answer rate of two syllable pronunciation task.



(a) (b)

Figure 4-11. (a) Optimized app (with pre-exercise page); (b) Optimized app users' correct question answer rate of two syllable pronunciation task.

Figure 4-9 shows the two syllable pronunciation results of users obtained by calculating the correct answer rate of two syllable questions that both syllables in the question are pronounced correct. According to the results, we can more clearly find that tone pronunciation ability obtained after using different apps is different. After using the app for a total of 3 hours, the correct answer rate of the prototype app users is 22% higher than the correct answer rate of NHK's app users, and the difference of the correct answer rate between the optimized app users and NHK's app users is up to about 42%. In addition, there are huge individual differences in tone pronunciation between users as the results of tone perception. Half of the NHK's app and prototype app users had the correct answer rate less than 50%. However, all the optimized app users had a correct answer rate at least higher than 50%.

Furthermore, the three participants mentioned in the previous section whose correct answer rate of perception questions improved by over 20% in the second session (participants 5 in (a) of Figure 4-11, and participants 3 and 5 in (b) of Figure 4-11) also improved their pronunciation correct answer rate by 37%, 17% and 23% on the final-test.

For three pronunciation tasks, the results of the t-test between correct answer rates of different apps users are shown in Table 4-2. As shown in Table 4-2, using the perceptual training app has a significant impact on improving tone pronunciation ability. Furthermore, optimizing the application design to reduce the extraneous cognitive load greatly increase the learning effect of perceptual training app. However, the pronunciation pre-exercise function has no significant impact on improving tone pronunciation ability.

Table 4-2. T-Test between correct answer rate of apps users.

Pronunciation Tasks	Paired Samples T-Test		df	p	Cohen's d
Reordering of Four Tones	Prototype App	NHK's App	9	**0.026**	**0.839**
	Optimized App	Prototype App		**0.043**	**0.745**
		NHK's App		**<0.001**	**1.829**
		Optimized App with Pre-exercise		0.628	0.159
Two-syllable Phoneme Combination Change	Prototype App	NHK's App		**0.005**	**1.164**
	Optimized App	Prototype App		**0.028**	**0.827**
		NHK's App		**<0.001**	**1.968**
		Optimized App with Pre-exercise		0.106	0.567
Two-syllable Questions	Prototype App	NHK's App		**0.004**	**1.208**
	Optimized App	Prototype App		**0.028**	**0.826**
		NHK's App		**<0.001**	**3.235**
		Optimized App with Pre-exercise		0.317	0.335

4.3.4. Results of Questionnaire

Table 4-3 shows the average score of the ten questionnaire items related to tone learning using a 6-point Likert scale.

Table 4-3. Questionnaire results.

Item/Average Score	**NHK's App**	**Prototype App**	**Optimized App**
I focus on learning the tone	**5**	4.2	4.8
(On first session) I use the application thoroughly	**5.4**	4	5.2
(On second session) I use the application thoroughly	**5.4**	4.4	5.2
This application helps me to understand tone	**5.4**	5	5
The pronunciation test helps to train tone ability	5.2	5	**6**
Pronunciation tests are interesting	**5.4**	5.2	4.8
Tone learning is interesting with this application	**5.2**	4.4	4.8
The learning method of this application is troublesome	2.4	**2.8**	2.4
Tone perception is difficult	4.4	**5**	4.2
Tone pronunciation is difficult	4.2	**5**	3.8

As shown in Table 4-3, except for the two items of "Tone perception is difficult" and "Tone pronunciation is difficult", NHK's app got higher scores than prototype app. Even though the tone pronunciation performance of NHK's app users is worse than the users of the prototype app, in the aforementioned two items, NHK's app users scored the difficulty of tone perception and pronunciation 0.6 and 0.8 lower than that of the prototype app users. A questionnaire survey reports on the motivation of Japanese Chinese learners and factors that makes learning difficult [39]. According to the research, it was mainly pronunciation, tone, syllable rhythm, intonation, and reading methods (81.1%) that seemed to be the most difficult for Chinese language learners, and the ability

to engage in conversation is the greatest motivation for Japanese college students to learn Chinese (67.7%). The results show that the educational contents of NHK's app focusing on phrases imitating match with the main learning motives of Japanese Chinese learners, and it make users liked NHK's app more.

However, among the perceptual training application, the optimized version has relatively advantages, and the optimized app users scored the difficulty of tone perception and pronunciation are lower than that of the NHK's app users. In the following, we will make detailed comparisons based on the design of the perceptual training applications before and after optimization.

According to the results of the questionnaire on each training task in perceptual training apps (Table 4-4), which is useful for understanding the tone, the process that seems to be the most useful are decreasing the speed (5.45 points) and the male/female voice change (5.45 points), followed by, change phoneme (5.4 points) and four tones in four syllables (5.2 points). These are the same methods used in the survey experiments to help Japanese native speakers who are learning Chinese to recognize tone, as seen in tone perceptual experiment with Rules. Textual explanations and explanations by animation got the lowest scores by the prototype app users (4.2 and 4 points, respectively), however, they greatly improved and explanations by animation got the highest scores by the optimized app users. We consider that optimizing the operation design and the loading time of the perceptual app also made users easier to get or memorize information provided by the app, especially the textual and the animation explanation.

Table 4-4. Questionnaire results of items design in perceptual training app.

Task of Perceptual Training/ Score (It helps to understand the tone)	**Prototype**	**Optimized**	**Average**
Textual Explanation	4.2	**5.4**	4.8
Explanation by Animation	4	**6**	5
Pure tone	4.6	**5.4**	5
Decrease the Speed	**5.4**	**5.8**	**5.6**
Fixing the Phoneme	4.8	5.2	5
Four Tones in Four Syllables	5	**5.8**	**5.4**
Pure Tone Different Pitch Range	4.6	**5.4**	5
Male/Female Voice Change	**5.4**	**5.8**	**5.6**
Change Phoneme	5.2	**5.8**	**5.5**

4.3.5. Error Analysis of Pronunciation Test

We investigated tone errors of the two-syllable and four-syllable tasks after using optimized app. The error rate of two-syllable combinations is shown in Figure 4-12. The tendency of errors is the same as the results in Chapter 3, the combination including the third tone is relatively difficult for learners, especially tone 4+3, and 2+3. The error rate is over 50% on average.

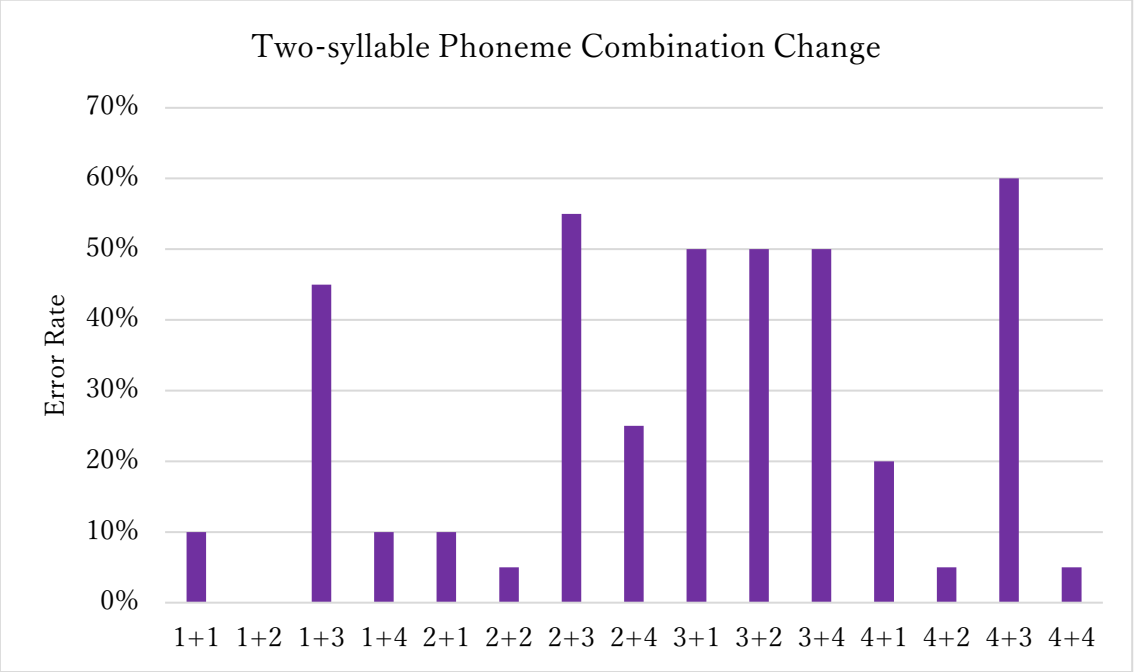


Figure 4-12. Error rate of tone error patterns in two syllable pronunciation tasks.

The error rates of tone combinations in the four-syllable task (a total of 24 patterns) are shown in Figure 4-13. The tone combination with the lowest error rate is "1234," which was the contents of the animation presented when users start the optimized app. It is greatly different from the result in Chapter 3 (Figure 3-9), the error rate of the tone combination "1234" did not decrease after using prototype app and had the highest error rate in the final-test. Furthermore, according to the questionnaire using a 6-point Likert scale, explanation by animation is the most improved item that of score increased 1.7 point after optimization. It is considered that the app optimization not only improved app users' pronunciation ability but also change the error trends of app users.

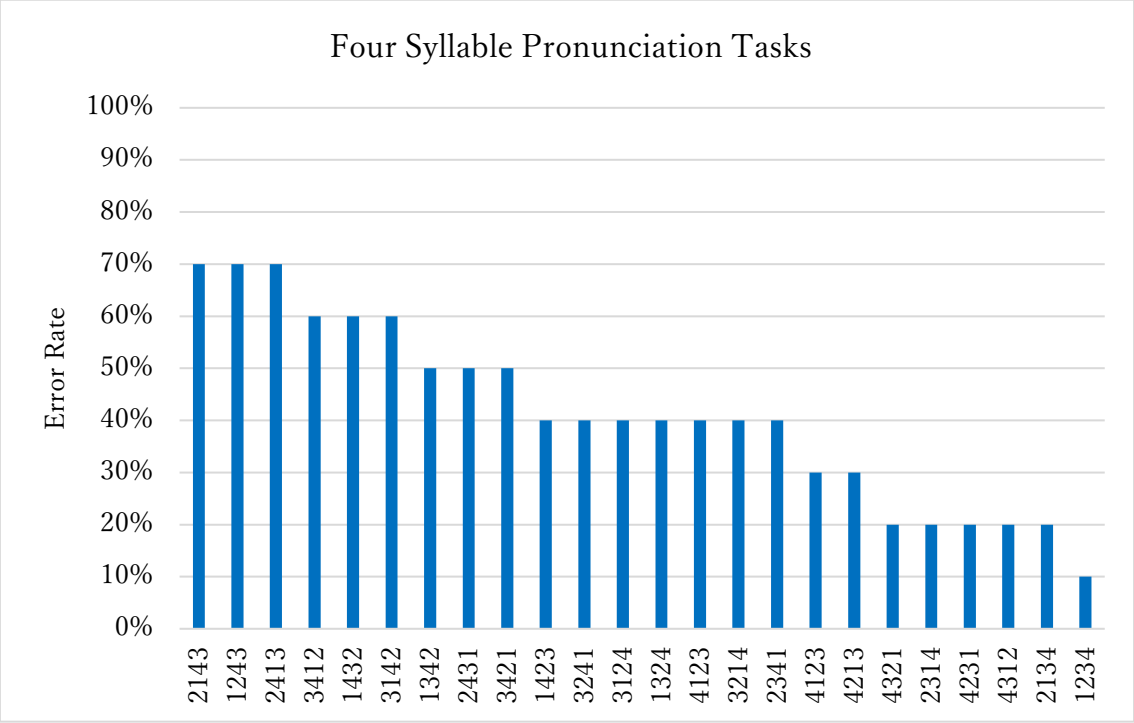


Figure 4-13. Error rate of tone error patterns in four syllable pronunciation tasks.

4.4. Discussion

There are various factors that can affect the results of the tone pronunciation experiment, such as training time, individual differences among learners, learning experience of participants, etc. However, considering the reasons why the related works [27], [29] that compared perception and pronunciation training described above led to different conclusions, we speculate that using monosyllable or disyllable words and training with an application or a teacher is key. Using monosyllable words for training does not focus on the relative pitch and categorization perception of tones. Though the participants in the study could generally distinguish four tones in monosyllables, recognizing and imaging the relative pitch change of Mandarin Chinese tone correctly was still difficult for them [26].

Furthermore, practicing pronunciation with a teacher can give the learners a proper opinion and help them to avoid pronunciation error, and as shown in our experimental results, pronunciation practice with the application may lead to fossilized errors. On the other hand, the feedback learners can get from perceptual training with a teacher is similar to perceptual training with an application (the answer of the perception question). However, the time that learners have to wait for getting the audio of perception question and feedback from teacher will increase the extraneous cognitive load just like the prototype app. The perceptual training with an application that can let learners feel free to choose the perception question and based on cognitive psychology theories, the perceptual training app we developed examines the learning content, conducts experiments, and optimizes perceptual training tasks and application interfaces that help users to learn Mandarin tone. It is assumed that the best way to implement the learning method, perceptual training, is to train with the app and the polysyllabic tasks designed aim to make learners feel the difference and the pitch change regarding tone.

The results of the evaluation experiment showed that the optimized perceptual training application user's cognitive ability was improved, and tone pronunciation errors were reduced. However, as a result, the implementation of the tone pre-exercise page did not help perceptual training app users to improve their tone pronunciation ability. On the other hand, the number of the perception questions answered by users was reduced because of pronunciation practice, which may adversely affect the tone learning effect of the perceptual training app.

In addition, a participant who had a high correct answer rate (93.3%) in the disyllabic pronunciation task after training with perception questions performed worse (83.3%) in the disyllabic pronunciation task after practicing pronunciation instead of perception. One of the reasons is the low-dipping image of the third

tone led him to pronounce the third tone like the fourth tone (high-falling), however, the pronunciation after first session didn't cause the error. It is considered that perceptual training is effective in avoiding pronunciation errors or pronunciation practice may cause errors in pronunciation. Therefore, to support pronunciation practice to improve tone production, tasks design and feedback of pronunciation practice with the application should be based on considering the tone understanding and perception, error correction, vocal cords exercises, etc.

Chapter 5

Conclusion and Future Tasks

This paper discusses the specialty of pronunciation learning in language learning and emphasizes the importance of auditory cognitive ability in the second language. It takes Chinese tone learning as an example and considers the cognitive load and the shortage of the general learning method's pronunciation practice by imitating the model voices of words, phrases, or sentences. Against this background, we focused on polysyllable perceptual training, emphasized comparison of each tone, and designed tasks commensurate with the learner's cognitive ability using the small step strategy. Furthermore, we considered the shortness of sensory memory and optimized the interface and operation design of the perceptual training application, compared with the imitation practice app, and as a result the tone pronunciation correct answer rate of users was 29.7% to 36% higher in two basic pronunciation tasks and 9.4% higher in the phrase pronunciation task after using for a total of 3 hours. In the results of the paired sample t-test, there was a statistically significant difference in the correct answer rate of the pronunciation tasks mentioned above from learning with the imitation practice app and the perceptual training application ($t(9) = 5.78$, $p < 0.001$, Cohen's $d = 1.83$, $t(9) = 6.22$, $p < 0.001$, Cohen's $d = 1.97$ and $t(9) = 3.33$, $p = 0.009$, Cohen's $d = 1.05$).

However, the implementation of the tone pronunciation practice function did not help perceptual training app users to improve their tone pronunciation ability. Furthermore, practicing pronunciation could be the cause of tonal pronunciation errors. Therefore, we considered using machine learning to support tone pronunciation practice. We aimed to use machine learning to design pronunciation tasks rather than recognize tones correctly, as is the case in existing research.

5.1. Learning Mandarin Tone with Machine Learning

In many cases, AI refers to machine learning or neural networks that refer to human cognitive structures. For foreign language pronunciation learning, there are some cases that recommend using AI, such as speech recognition systems, which can be used for practicing conversation in foreign languages. Speech recognition can grade learners' pronunciation through speech analysis by acoustic features, etc. However, the answers calculated using these technologies and AI may be wrong, or the accuracy rate may be significantly reduced depending on the purpose of the system. Using a speech recognition conversational system to practice tone pronunciation, it is likely that the correct tones and tonal errors will be ignored to make a conversation. In the studies of Chinese tone recognition using machine learning, the tone of monosyllable words was used as training data, and the method and accuracy of tone recognition were examined.

The highest accuracy (95.5%) was obtained using the convolutional neural network (CNN) and the mel-frequency cepstral coefficients (MFCCs) for tone pronunciation of 4500 syllables by 125 children aged 3 to 10. This study [40] showed that tone recognition by machine learning is possible; however, there are some shortcomings in learning tone with the tone recognition system. This study

used children's speech, which had limitations on the pitch range, and the spectrogram and MFCCs of pronounced monosyllabic words in the paper showed that the third tone was a dipping-rising tone (Figure 5-1). In related works, as mentioned in Chapter 1, training with disyllabic words is more effective than training with monosyllabic words [26] and teaching the dipping-rising tone may cause pronunciation errors [31]. Furthermore, as the tones are made up of differences in relative pitch patterns, we considered that it is necessary to design pronunciation tasks that promote tone categorization of learners.

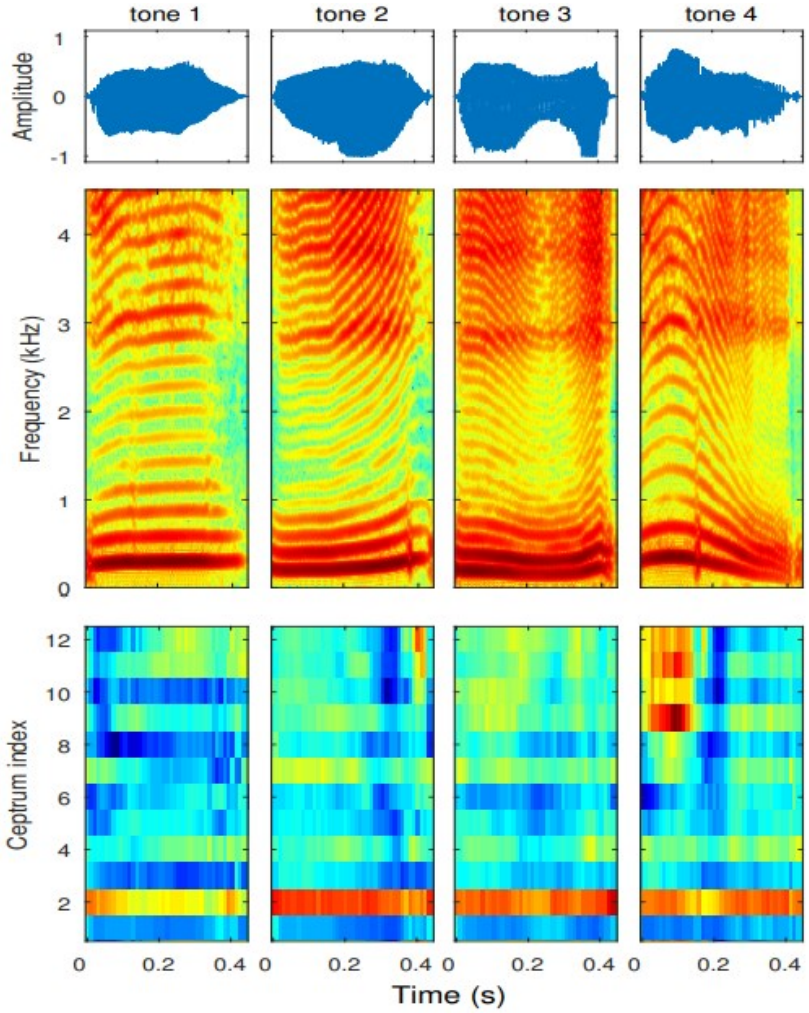


Figure 5-1. Machine learning data from previous studies on tone recognition [40]. Top: Time waveforms; Middle: Spectrograms; Bottom: Mel-frequency cepstral coefficients (MFCCs).

A study [41] that used the training data fetched and selected from the pronunciation of 590 short sentences by two pairs of men and women obtained an accuracy of 87.6% using a CNN. The accuracy of tone recognition is strongly related to the data used for machine learning, and we aimed to design pronunciation tasks that improve the abilities required for tone pronunciation rather than correctly recognizing tones. At the first stage of learning tone, we consider that learners should acquire the cognitive ability of tones, then learners should practice pronunciation with vocal cords and their tonal cognition. We used the designed voices as training data and made a pronunciation exercise using a CNN. The designed task for the pronunciation exercise was a disyllabic task that requested the pronunciation of tones 1, 2, 3, and 4 in first syllable combined with tone 1 in the second syllable within 1.1 seconds and the length of the first syllable in less than 0.55 seconds. The data used for machine learning (CNN) were pronunciation data (a total of 547 items of audio, 60% for training and 40% for validation) of four men and three women who followed the rules of the pronunciation task for the training and test set. The classification accuracy was 98.2%, which was high because of data pattern selection by rules and the features of the two-syllable tone. Figure 5-2 shows a part of the data used for machine learning.

It was considered that we should make tone learners from the machine learning aware of the processes of how a CNN classifies different tone pairs and the results of the classification by percentage. Learners can obtain key information to acquire tones from the process of machine learning, learning with the training data (pronunciations follow rules of pronunciation task) visually (Figure 5-2) and through hearing and practicing with the visual form of their pronunciations and the classification results by percentage by using the CNN.

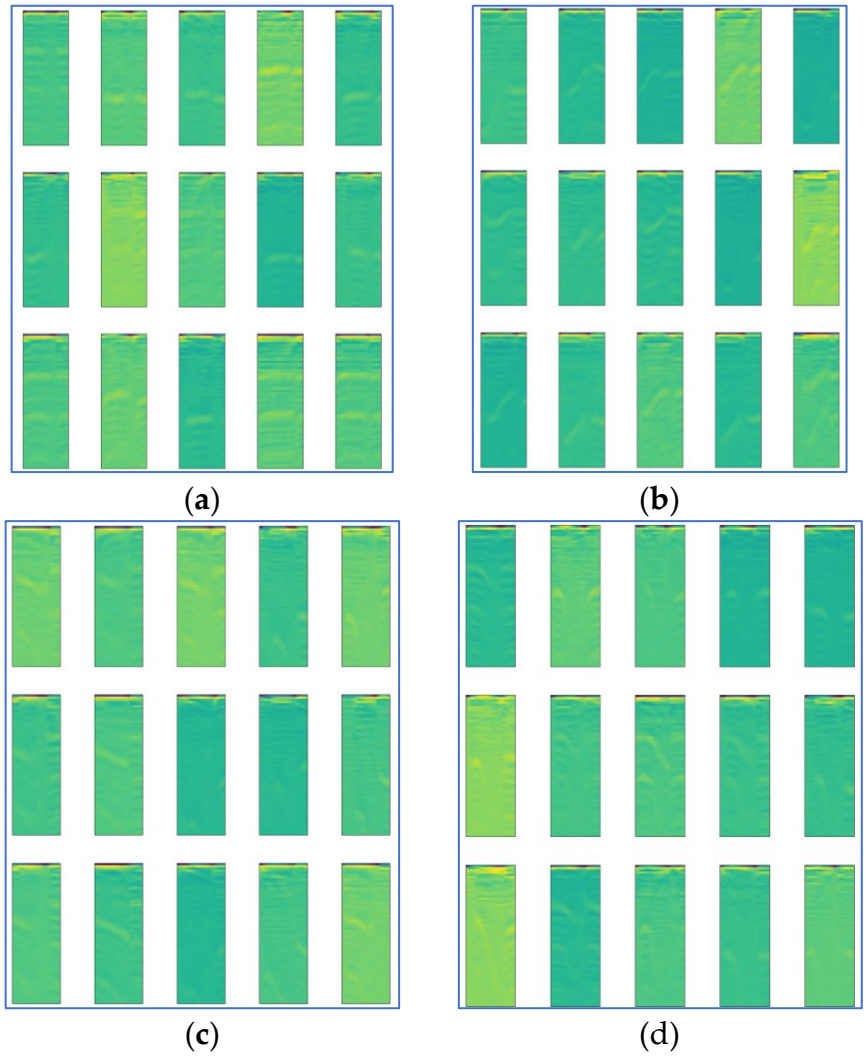


Figure 5-2. Training data for pronunciation task designed in this study. (**a**) Combination of tone 1 and tone 1; (**b**) Combination of tone 2 and tone 1; (**c**) Combination of tone 3 and tone 1; (**d**) Combination of tone 4 and tone 1.

The design of the pronunciation task aimed to enable learners to be able to pronounce tones 1, 2, 3, 4 combined with tone 1 correctly and individually with the speed of ordinary conversation (about 1 second). Even for native speakers, to get a high percentage of classification by a CNN with a pronunciation following the rules that are strictly judged requires practicing a little. It is assumed that practicing with the designed tone pronunciation task using a CNN helps to promote the tone understanding and categorization of learners and makes learners focus on controlling vocal cords.

5.1.1. Discussion

According to the results, using AI for pronunciation training seems to be helpful for tone learning. However, factors such as the user interface, loading time and individual differences of learners will greatly make a difference to the effect of pronunciation training function. Though the feedback that learners obtain from AI helps them to judge their own pronunciation, the effect of pronunciation practice is highly dependent on learner's own pronunciation and recognizing ability. In the acquisition of basic tonal cognitive and pronunciation ability, perceptual training is an effective method.

According to the experimental in Chapter 4, the use of apps for perceptual training has advantages in obtaining information and feedback smoothly over perceptual training with teachers. In addition, using apps to improve students' tonal recognition and pronunciation ability can effectively reduce teachers' burden on pronunciation teaching. However, the accuracy, immediacy, and convenience of AI are still a major issue in terms of the ability to correct errors and provide opinions for pronunciation practice.

In the field of pronunciation teaching, teachers' error correction and opinion-providing abilities are also different. It is even generally believed that native Chinese speakers are more suitable for teaching. Therefore, we draw preliminary conclusions from the experimental results that there will be a higher risk of repeated pronunciation errors in the process of self-learning and obtaining negative feedback when using AI for pronunciation practice. Though using AI for pronunciation training does have a positive effect on tonal learning, there are still some risks and the pronunciation guidance of native-speaking teachers still cannot be replaced by AI.

5.2. Conclusion and Future Tasks

This research shows that in the context of large individual differences in the ability to acquire tones, tone learning cannot only involve phrases pronunciation practice, but must pay attention to the tone perceptual ability of learners and the cognitive load during learning.

In this research, we used conventional methods of related works and carried out phonetic experiments and error analysis. However, we focused on the ability of polysyllabic tone perception, performance distribution, and improvement of ability in different perceptual tasks and designed training tasks thought to be effective for tone acquisition.

To verify the training effect, we compared the tone production ability before and after use of the existing phrases pronunciation practice app and a question collection system consisting of the designed perceptual training tasks. In addition, we considered the learning effect of applications based on the different experimental procedure that w/o taking a pronunciation pre-test and different operation design of perceptual training application at several pronunciation tasks.

With a short-term application (a total of about three hours), the perceptual training app had a significant impact on improving tone pronunciation ability and taking a basic tone pronunciation test designed by us before training with the app would improve tone pronunciation ability of imitation practice and perceptual training app users.

Furthermore, the pre-exercise function that instructs perceptual training application users to practice pronunciation lightly after enabling app has no learning effect we expected. However, optimize the perceptual training app aim

to reduce the extraneous cognitive load to the limit greatly increase the learning effect of the perceptual training app.

The results from the evaluation experiment shown that training by solving perceptual questions designed is more efficient than imitation practice to learning Chinese tone. In addition, we considered that perceptual training with an app is the most efficient way for learners to receive tonal information and to reduce the cognitive load during learning. It is assumed that the best way to implement the learning method, perceptual training, is to train with the app and the polysyllabic tasks designed aim to make learners feel the difference and the pitch change regarding tone.

Furthermore, in addition to the prediction results of the trained model, the percentage of classification, the training data set, and the visualized data were presented to the learner as feedback for learning. Machine-learning-assisted tone pronunciation practice is feasible.

In the future, it is necessary to consider the instructional design of pronunciation practice based on difficulty of tone combination, error patterns, phonemes, etc. Furthermore, it is also necessary to optimize the interfaces and operation design or use cloud computing to minimize the computing time and extraneous cognitive load. On the other hand, there are studies showing that the individual differences between people and the data of language processing (in Mandarin Chinese) can be used to predict people's ability, which helps in understanding human cognition and creates navigation by machine learning [42], [43]. Classifying learners automatically according to their phonological ability and providing navigation considering their individual differences using machine learning are also future issues.

Acknowledgments

I wish to express his deep thanks to professor Madoka Hasegawa, for her valuable advice and helpful discussions during the course of this study.

I would also like to thank my advisor, professor Atsushi Itou, who has influence me most in this research. Without his support and guidance, this thesis would not have been written in the current form.

I would like to show my gratitude to professor Mie Sato, professor Masahiro Fujii, professor Kanemitsu Ootsu, professor Tetsu Yajima and professor Masafumi Koike for sharing their pearls of wisdom with us during the course of this research, and I thank professor Shinichi Watanabe, professor Tomoharu Ishikawa and professor Hiroyuki Hatano for their so-called insights.

In addition, I wish to thank graduate students, Yuya Kiryu, Hayato Aratame, Akimasa Kimura, Yusuke Otsuka and other member of Ito Lab, for their helpful suggestions during the course of this study.

I would also like to express my gratitude to the students who cooperation in the experiments and Prof. Matsukane who provided valuable advice to proceed the research.

References

[1] C. E. Snow and M. Hoefnagel- Höhle, “The critical period for language acquisition: Evidence from second language learning,” Child Dev., vol. 49, no. 4, pp. 1114-1118, Dec. 1978.

[2] B. McLaughlin. “Myths and Misconceptions about Second Language Learning: What Every Teacher Needs to Unlearn,” eScholarship: Center for Research on Education, Diversity and Excellence.

https://escholarship.org/uc/item/1t55s0tc (accessed Sept. 9, 2020).

[3] T. D. Keeley, “Is a Native-like Accent in a Foreign Language Achievable? Examining Neurological, Sociological, Psychological, and Attitudinal Factors,” *Keieigaku Ronshu*, vol. 26, no. 4, pp. 59-92, Mar. 2016.

[4] K. Bock and W. J. M. Levelt, “Language production: Grammatical encoding,” *Handbook of Psycholinguistics*, Academic (Press Inc.), San Diego, CA, USA, 1994, pp. 945-984.

[5] Z. M. Griffin and V. S. Ferreira, “Properties of Spoken Language Production,” Handbook of Psycholinguistics, Dec. 2006, pp. 21-59. [Online]. Available: 10.1016/B978-012369374-7/50003-1

[6] E. P. Altenberg and H. S. Cains, “The effects of phonotactic constraints on lexical processing in bilingual and monolingual subjects,” *J. Verbal Learning Verbal Behav.*, vol. 22, no. 2, pp. 174-188, Apr. 1983.

[7] D. Deutsch, “Speaking in Tones,” *Sci. Am. Mind*, vol.21, Jan. 2007. [Online]. Available:10.1038/scientificamericanmind0710-36

[8] Y. R. Chao, “Four tones,” *Lang. Problems*, The commercial Press, Beijing, China, 1980, ch. 5, pp. 59-81.

[9] S. Aihara, *Learn Chinese from basic pronunciation*. Tokyo, Japan: Asahi Press, 2003.

[10] T. Yuyama and N. Takeda. “Use and the education problem of the function that displays the tone wave in the study of Chinese pronunciation,” Application of the e-Learning system to the basic education of the Chinese language.

http://chinese-you.net/pdf/e-Learning08-03.pdf (accessed Sept. 9, 2020).

[11] E. Pelzl, E. F. Lau, T. Guo and R. DeKeyser, “Advanced second language learners’ perception of lexical tone contrasts,” *Stud. Second Lang. Acquis.*, vol. 41, no. 1, pp. 59-86, Mar. 2019.

[12] R. Ding, “Talking about the Recognition of the Problem with "Tone is difficult" for Japanese Learners of Chinese as the Second Language,” *J. Center Foreign Lang. Educ.*, Shimane Univ., vol. 12, pp. 75-87, Apr. 2017.

[13] K. J. Chang, L. M. Chen and N. C. Lee, “Tonal errors of Japanese students learning Chinese: A study of disyllabic words,” *Comput. Ling. Chin. Lang. Process.*, vol. 11, no. 3, pp. 281-296, Sept. 2006.

[14] H. W. Wu, “A Study of Mandarin Chinese Tone Production by Polish Speakers,” M.S. thesis, Grad. Inst. Teaching Chin. Second/Foreign Lang., Nat. Kaohsiung Normal Univ., Kaohsiung, Taiwan, 2011.

[15] H. Y. Chen, “The Study of Mandarin Chinese Phonetic Errors and Teaching Strategies of Vietnamese Spouses in Taiwan,” M.S. thesis, Depart. Lang. Lit. Educ., Nat. Taitung Univ., Taitung, Taiwan, 2007.

[16] S. Kim, “Analysis and Teaching Research on Mandarin Chinese Tone Errors of Korean Students,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2005.

[17] L. Lin, “An Analysis of the Tone Errors of Hong Kong Students Who Are Studying in Taiwan,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2017.

[18] K. Takamatsu, “A Case Study of Mandarin Chinese Tone Combination Training: Error Corrections in Pronunciation for an Aged Japanese Learner,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2018.

[19] L. Y. Cheng, “The Perception and Production of Mandarin Tones by L2 learners,” M.S. thesis, Depart. Foreign Lang. Lit., Nat. Chiao Tung Univ., Hsinchu, Taiwan, 2014.

[20] H. Y. Lin, “Mandarin Tonal Acquisition of Novice Japanese Learners,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2007.

[21] R. H. Jiang, “A Study on Mandarin Disyllabic Tones of German Learners,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2012.

[22] P. L. Tsai, “Perception and Production of Mandarin Chinese Lexical Tone by Adult English Speaking Learners,” M.S. thesis, Inst. Ling., Nat. Tsing Hua Univ., Hsinchu, Taiwan, 2008.

[23] K. L. Tran Thi, “Error Analysis of Mandarin Tones from Vietnamese Learners,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2005.

[24] N. T. Nguyen Thi, “A Case Study of Mandarin Chinese Tone Training: Error Corrections in Pronunciation for beginning Vietnamese learners,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2016.

[25] C.C. Li, “Error Analysis and Remedial Instruction of Mandarin Chinese Tones --- A study on American Learners of Mandarin,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2010.

[26] N. H. Wu, “Initiating foreign students into Mandarin Chinese tone pairs and its effects on tone acquisition,” M.S. thesis, Grad. Inst. Teaching Chin. Second/Foreign Lang., Nat. Kaohsiung Normal Univ., Kaohsiung, Taiwan, 2015.

[27] A. W. Hsia, “Training Non-tonal Speakers in the Perception and Production of Mandarin Tones in Disyllabic Words,” M.S. thesis, Depart. Eng., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2010.

[28] Y. Wang, M. M. Spence, A. Jongman and J. Sereno, “Training American listeners to perceive Mandarin tone,” *J. Acoust. Soc. Am.*, vol. 106, no. 6, pp. 3649-58, Dec. 1999.

[29] M. Li and R. DeKeyser, “Perception Practice, Production Practice, and Musical Ability in L2 Mandarin Tone-word Learning,” *Stud. Second Lang. Acquis.*, vol.39, no.4, pp.593-620, Jan. 2017.

[30] P. S. Zhang, “An Analysis of Tone Errors Made by Japanese Learners of Mandarin in Disyllabic Words,” M.S. thesis, Depart. Chin. Second Lang., Nat. Taiwan Normal Univ., Taipei, Taiwan, 2013.

[31] T. Umemoto, *Course psychology 7 Memory*. Tokyo, Japan: Univ. Tokyo Press, 1969.

[32] T. Iwamoto, “Behavioral Psychology and Cognitive Psychology (V),” *Annu. Rep.*

Cult. Sci., vol. 41, no. 3, pp. 105-141, Feb. 1993.

[33] L. Yin, W. Li, X. Chen, R. Anderson, J. Zhang, H. Shu and W. Jiang, "The role of tone awareness and pinyin knowledge in Chinese reading," *Writ. Syst. Res.*, vol. 3, pp. 59-68, Jan. 2011.

[34] J. Sweller, "Cognitive Load During Problem Solving: Effects on Learning," *Cog. Sci.*, vol. 12, pp. 257-185, Apr. 1988.

[35] S. Krashen, "We acquire vocabulary and spelling by reading: Additional evidence for the input hypothesis," *Mod. Lang. J.*, vol. 73, pp.440-464, Dec. 1989.

[36] Y. Matsumoto, "A Proposal for the Method of Teaching Dynamic Tones in Chinese Intonation," *J. Waseda Univ. Soc. Chin. Lit.*, vol. 40, pp.59-77, Feb. 2014.

[37] J. Shigematsu, "How Japanese Speakers Learn the Chinese Intonation", *J. Waseda Univ. Soc. Chin. Lit.*, vol. 22, pp.58-66, Dec. 1996.

[38] J. C. Darwin, T. M. Turvey and G. R. Crowder, "An auditory analogue of the sperling partial report procedure: Evidence for brief auditory storage," *Cog. Psychol.*, vol. 3, no. 2, pp. 255–67, Apr. 1972.

[39] L. Tao, "Investigation and Research on Learning Motives of Japanese Chinese Learners," *Forum Lang. Inst.*, vol. 8, pp.81-90, Mar. 2014.

[40] C. Chen, B. Razvan, X. Li and C. Liu, "Tone Classification in Mandarin Chinese Using Convolutional Neural Networks," *Conf. Int. Speech Comm. Assoc.*, Sept. 2016.

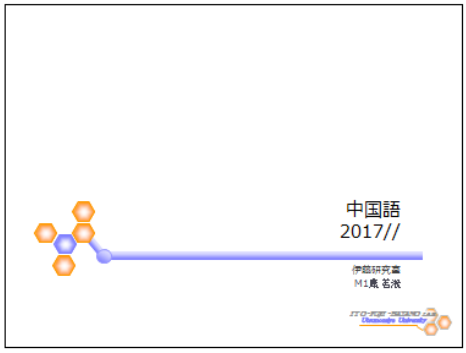
[41] L. J. Shen and W. Wang, "Fusion Feature Based Automatic Mandarin Chinese Short Tone Classification," *Tech. Acous.*, vol. 37, no. 2, pp. 167-174, Apr. 2018.

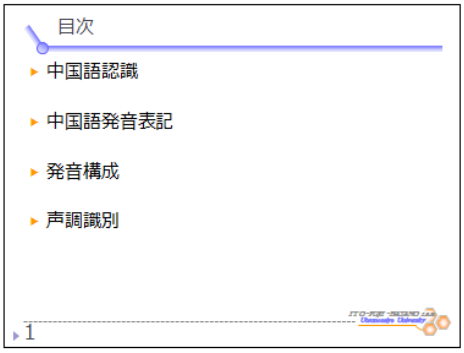
[42] M. Stella and Y. N. Kenett, "Viability in Multiplex Lexical Networks and Machine Learning Characterizes Human Creativity," *Big Data Cog. Comput.*, vol. 3, no. 3, pp. 45, July. 2019.

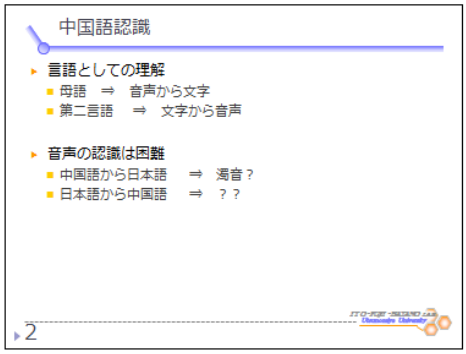
[43] K. D. Neergaard and C. R. Huang, "Constructing the Mandarin Phonological Network: Novel Syllable Inventory Used to Identify Schematic Segmentation," *Complexity*, pp.1-21, Apr. 2019.

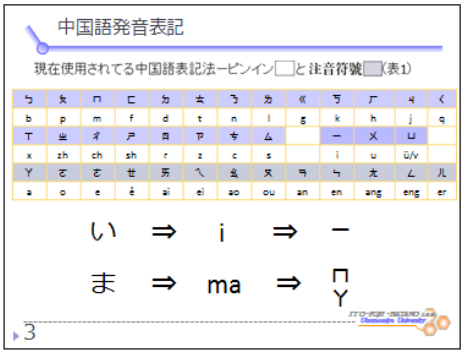
Appendix 1. Contents of experiment 1

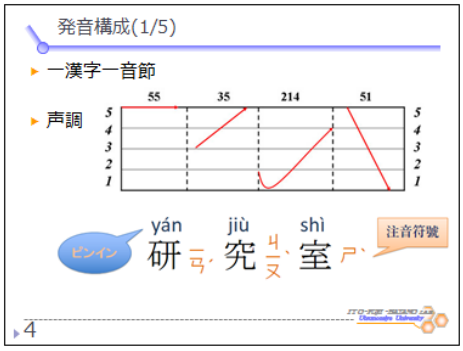
Experimental explanation slide:

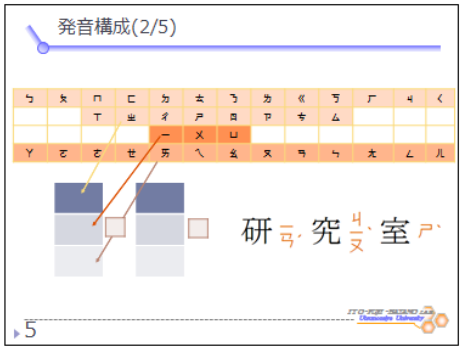


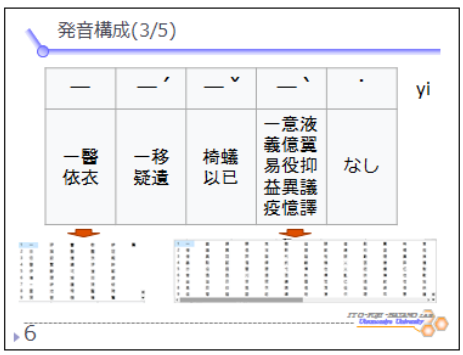












発音構成(5/5)

ㄐㄧㄡ	ㄐㄧㄡˊ	ㄐㄧㄡˇ	ㄐㄧㄡˋ	・	jiu
糾揪	なし	九酒久	就舊救究舅	なし	

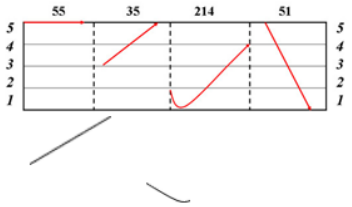
8

入力、読み

9

声調識別

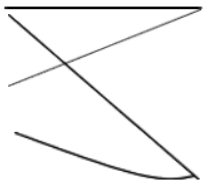
- 声調イメージ



10

一から十

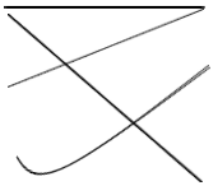
- 一, 二, 三, 四, 五, 六, 七, 八, 九, 十



11

一から十

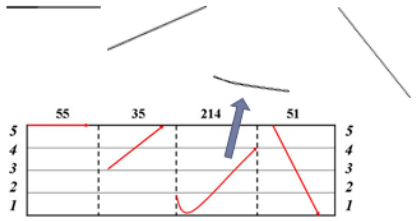
- 一, 二, 三, 四, 五, 六, 七, 八, 九, 十



12

四音節

- 一, 十, 五, 六

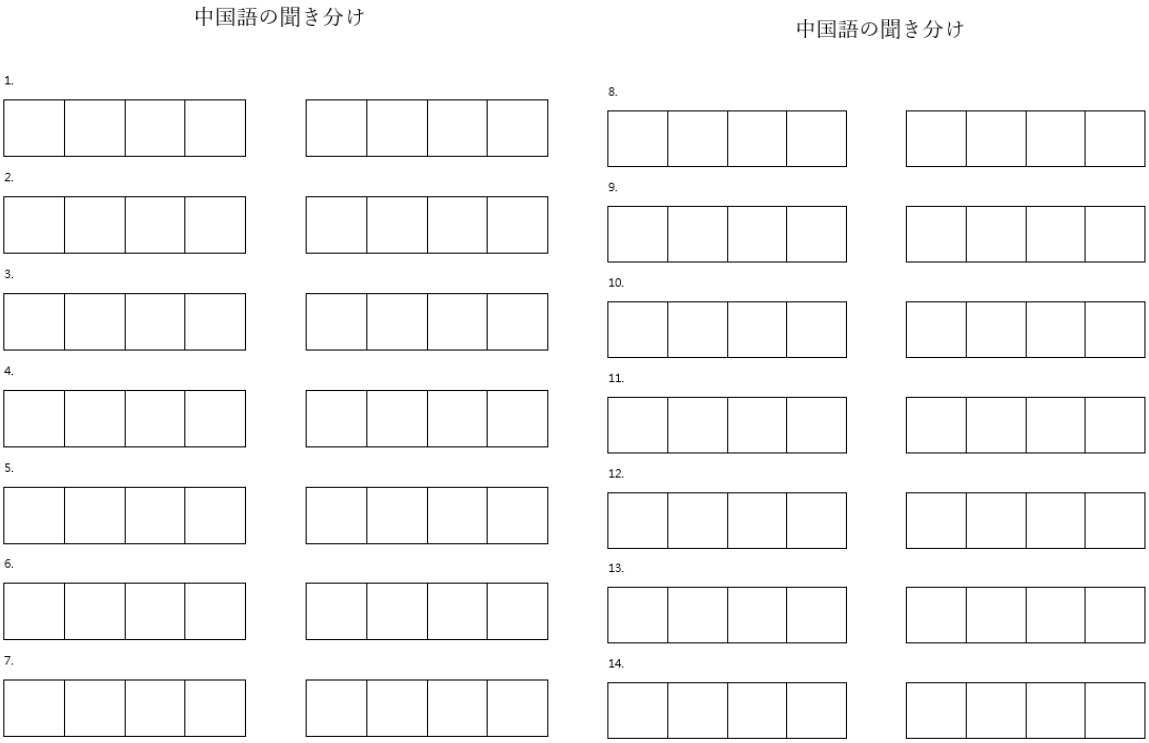


13

Four-character idioms used:

1. 弱肉強食
2. 天真爛漫
3. 森羅萬象
4. 星火燎原
5. 臨機應變
6. 哀兵必勝
7. 電光石火
8. 自由自在
9. 一刀兩斷
10. 溫故知新
11. 一字千金
12. 自問自答
13. 驚天動地
14. 自暴自棄

Answer Sheet:



Appendix 2. Contents of experiment 2

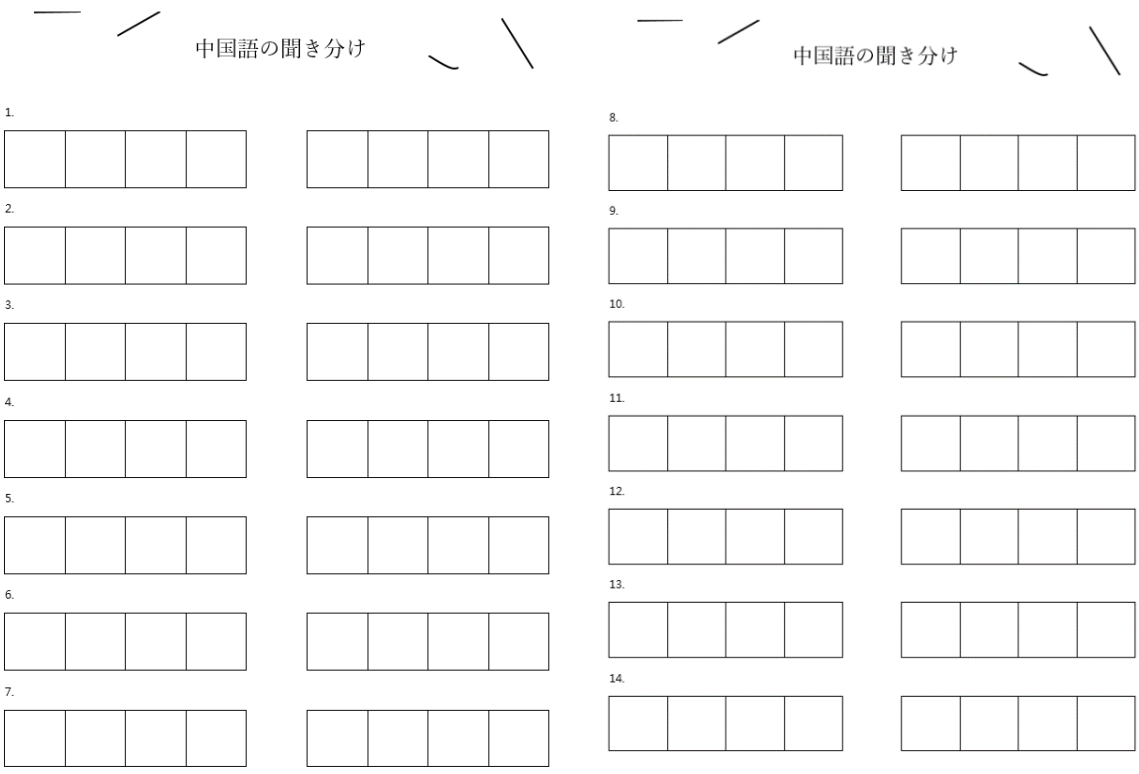
Phonemes used:

1. ㄧ
2. ㄨ
3. ㄩ
4. ㄛ

Four-character idioms:

臥薪嘗膽

Answer Sheet:



Appendix 3. Contents of experiment 3

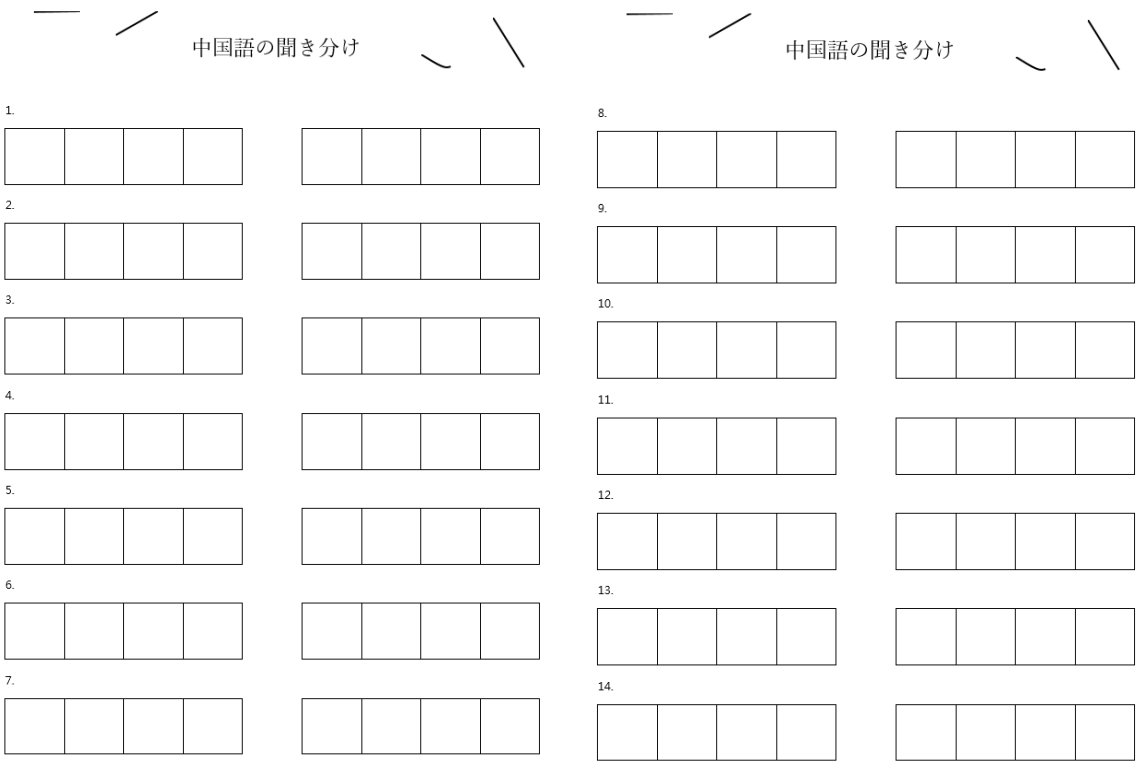
Phonemes used:

1. ㄒ
2. ㄨ
3. ㄩ
4. ㄜ

Four-character idioms:

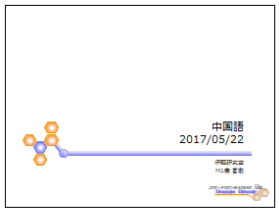
回心轉意

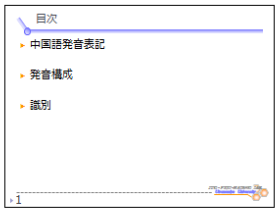
Answer Sheet:

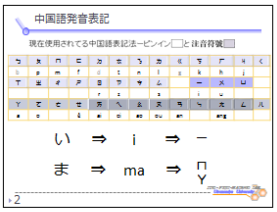


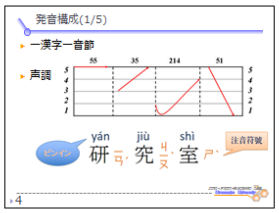
Appendix 4. Contents of experiment 4

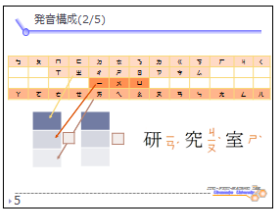
Experimental explanation slide:

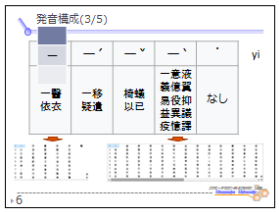


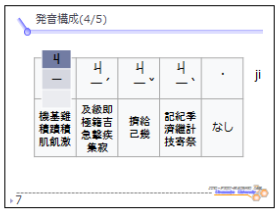


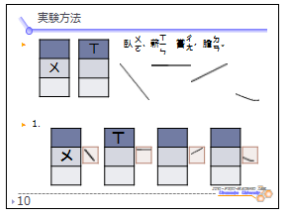


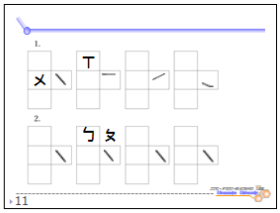


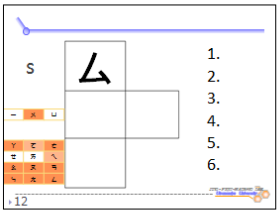


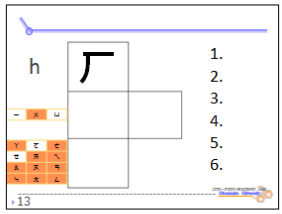


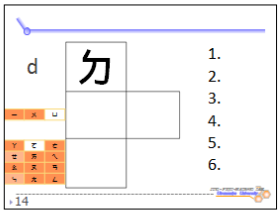


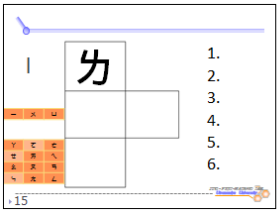


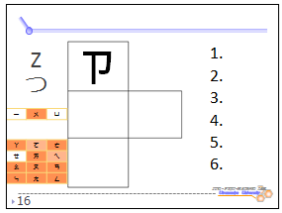


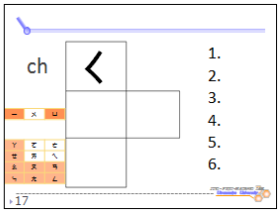


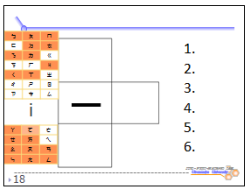


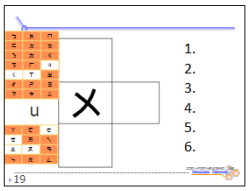


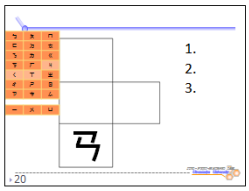


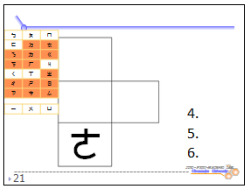


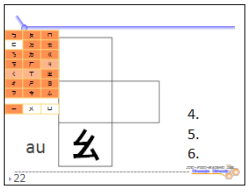


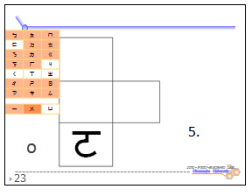


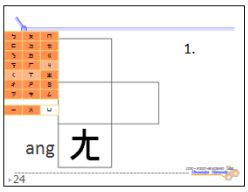


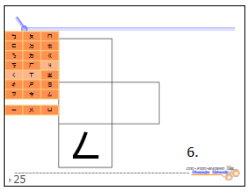


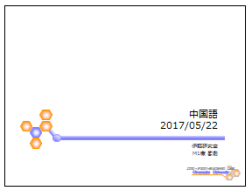












Four-character idioms :

1. 三皇五帝
2. 妻離子散
3. 三足鼎立
4. 自得其樂
5. 好騎者墮
6. 好生之德

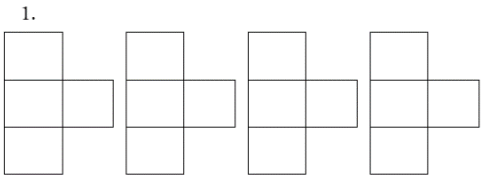
Phonemes Tested:

1. ㄙ
2. ㄏ
3. ㄉ
4. ㄌ
5. ㄗ
6. ㄑ
7. ㄧ
8. ㄨ
9. ㄢ
10. ㄜ
11. ㄠ
12. ㄛ
13. ㄤ
14. ㄥ

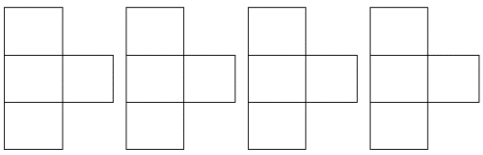
Answer Sheet:

中国語の聞き取り

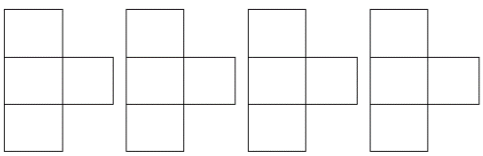
1.



2.

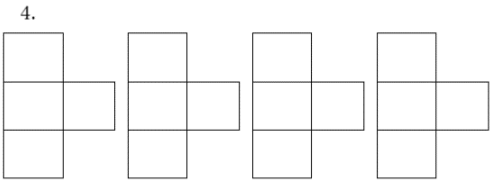


3.

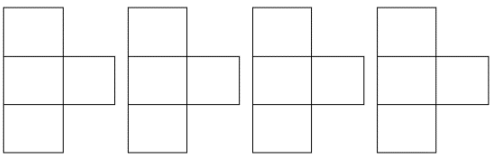


中国語の聞き取り

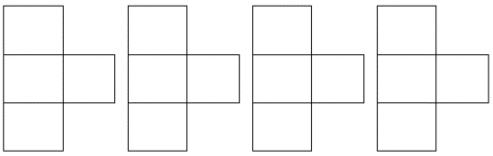
4.



5.



6.

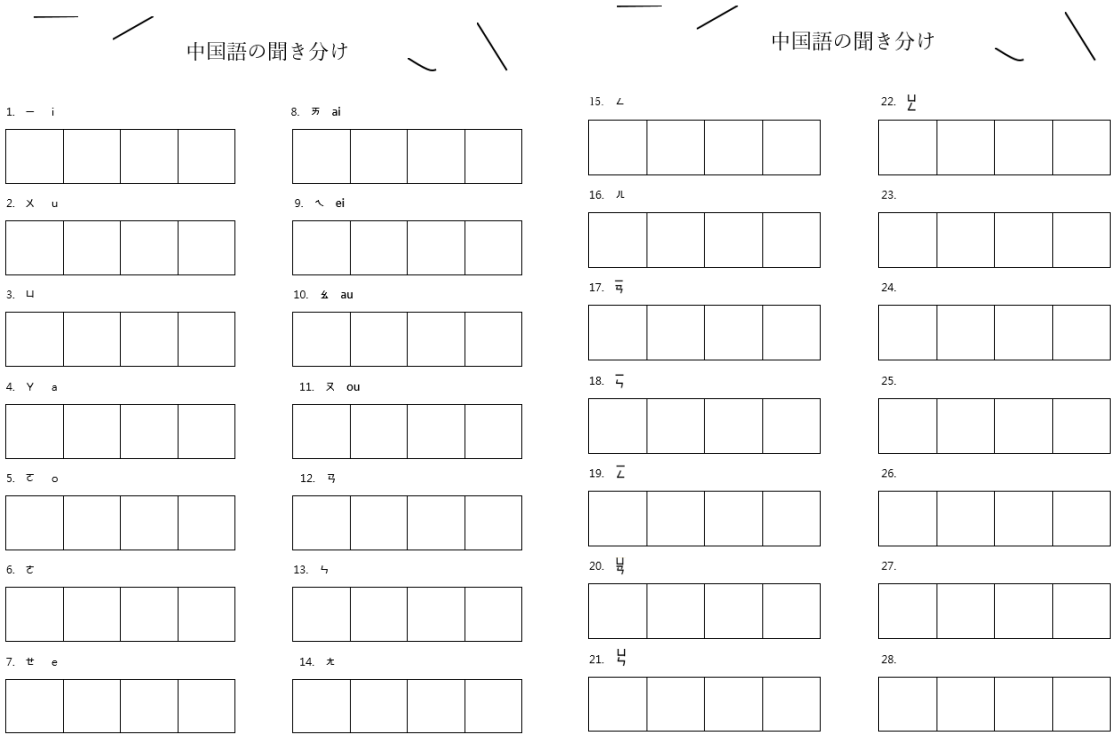


Appendix 5. Contents of experiment 5

Four-character idioms:

1. 堅忍不拔
2. 自問自答
3. 溫故知新
4. 改過自新
5. 哀兵必勝
6. 大器晚成

Phonemes Tested and Answer Sheet:



Appendix 6. Contents of experiment 6

Phonemes used:

1. ㄧ
2. ㄨ
3. ㄓ

xxxx Four-character idioms:

自暴自棄 名存實亡

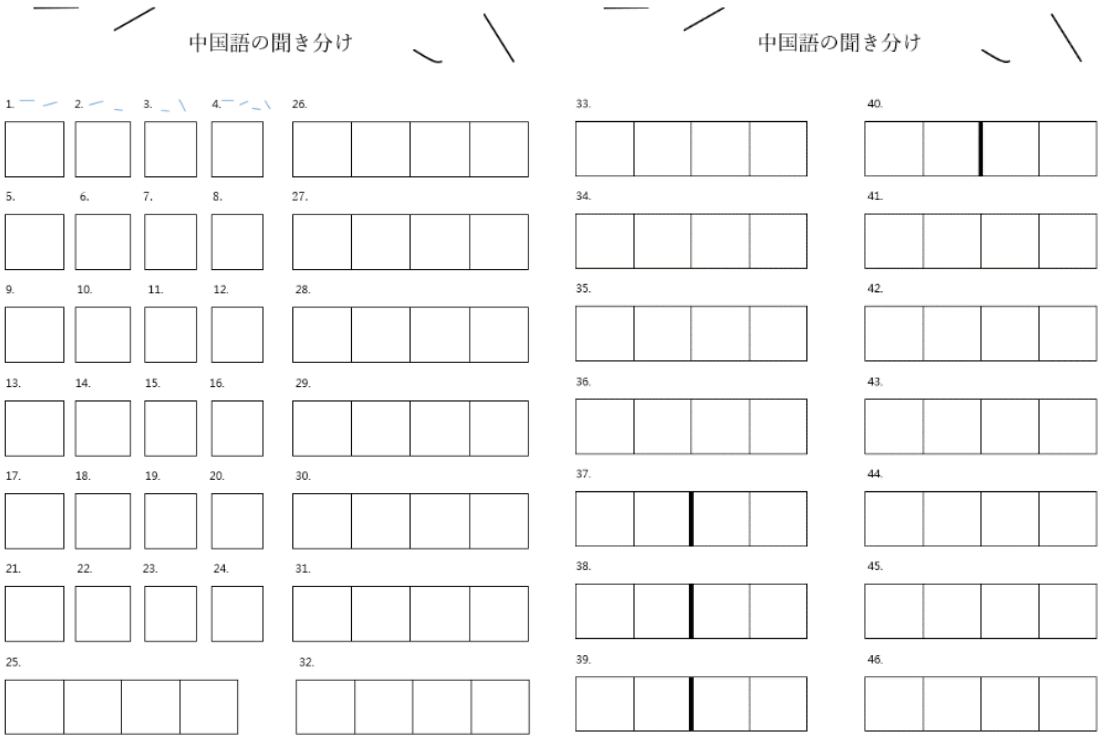
xxyy Four-character idioms:

天真爛漫 哀兵必勝 弱肉強食 驚天動地

Four-character idioms:

1. 爛醉如泥
2. 默默不語
3. 螳螂捕蟬
4. 驚慌失措
5. 事關重大
6. 心直口快

Answer Sheet:



Appendix 7. Contents of experiment 7

Phonemes used:

1. ㄏ
   ㄞ
2. ㄏ
   ㄚ

wxyz Four-character idioms:

千錘百鍊 臥薪嘗膽

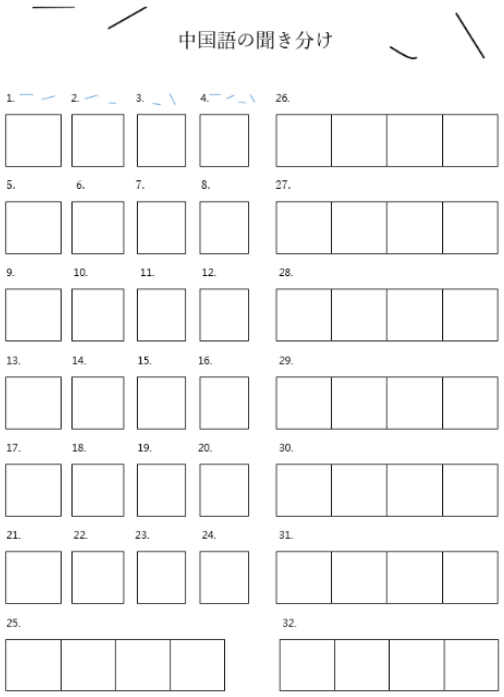
xxyy Four-character idioms x,y=1or4:

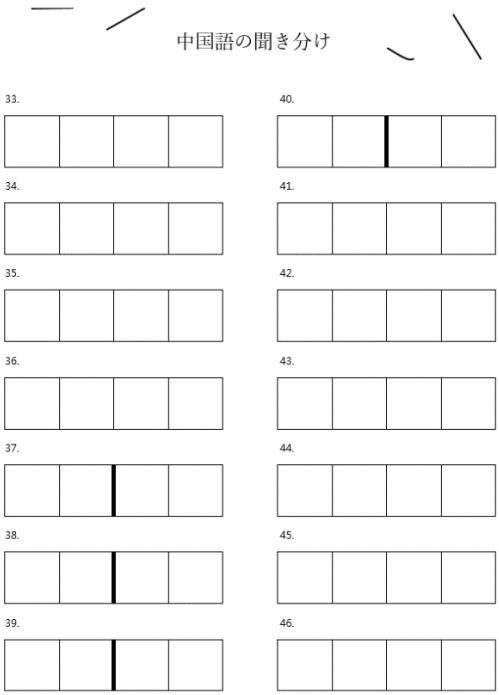
信誓旦旦 戰戰兢兢 三更半夜 四面八方

Four-character idioms:

1. 眾星捧月
2. 志同道合
3. 坐享其成
4. 直言不諱
5. 至死不渝
6. 自作自受

Answer Sheet:





Appendix 8. Contents of experiment 8

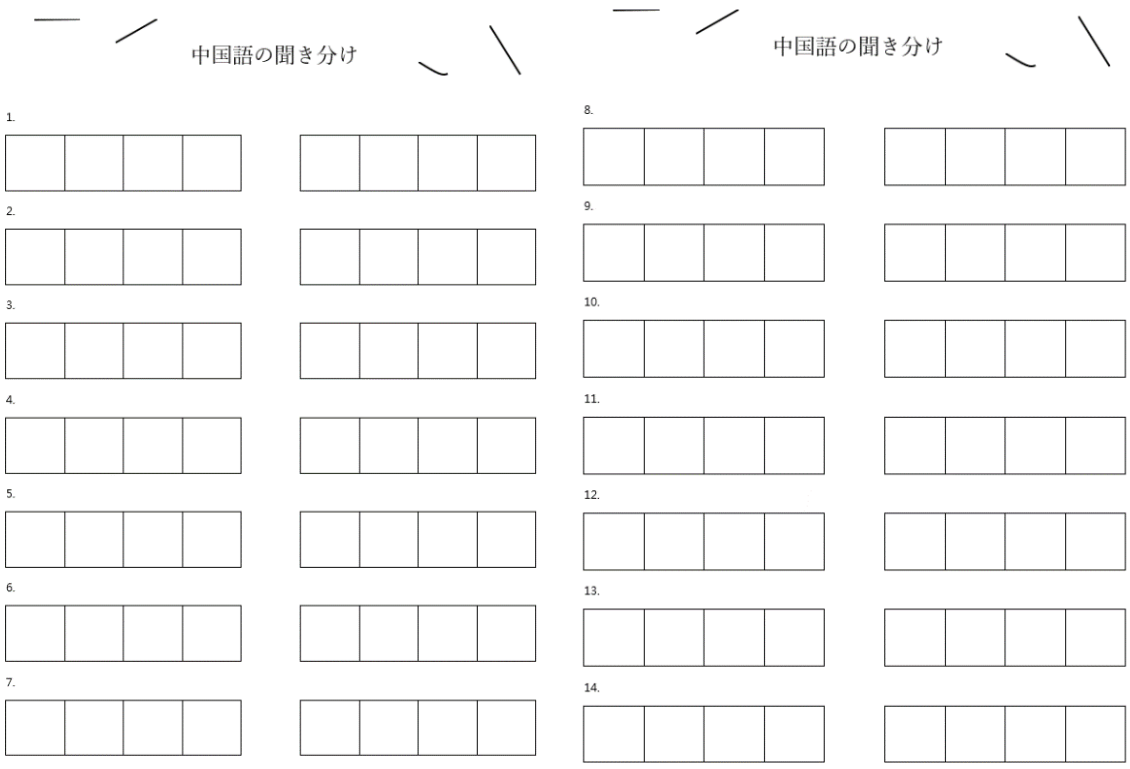
Phonemes used:

1. ㄏ
   ㄞ
2. ㄏ
   ㄚ

wxyz Four-character idioms:

1. 瞞天過海
2. 死灰復燃
3. 濃妝豔抹
4. 破釜沉舟
5. 百步穿楊
6. 妙手回春
7. 提心吊膽
8. 杏林春暖
9. 龍飛鳳舞
10. 四平八穩

Answer Sheet:



Appendix 9. Contents of experiment 9

Phonemes and character used:

1. ㄏ
   ㄞ
2. ㄏ
   ㄚ
3. ㄔ
   ㄨ
   ㄤ

Character used:

千 風 慮 謀 調 軍 雨 遠 深 順 萬 其 自 力 食 馬

wxyz Four-character idioms:

馬到成功 海枯石爛 笑裡藏刀 身不由己 同甘共苦 智勇雙全

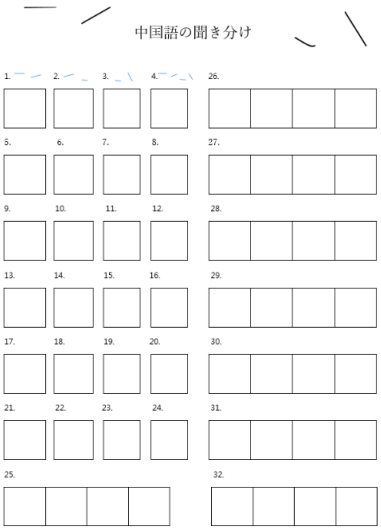
xxyy Four-character idioms （x,y = 4）:

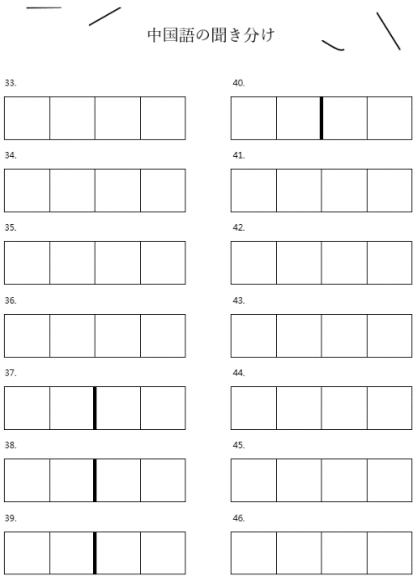
畫地自限 見利忘義 面面俱到 勝券在握

Four-character idioms:

1. 千軍萬馬
2. 風調雨順
3. 自食其力
4. 深謀遠慮
5. 恣意妄為
6. 比手畫腳

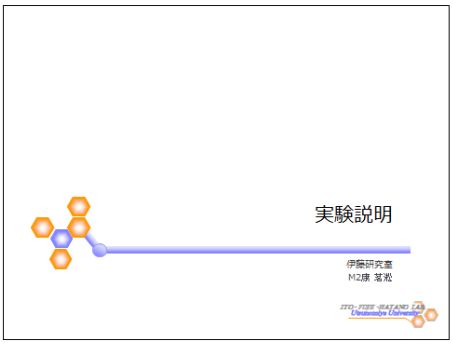
Answer Sheet:

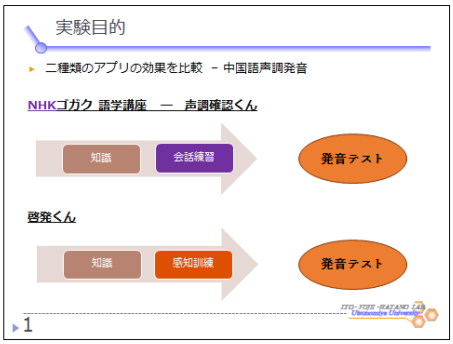


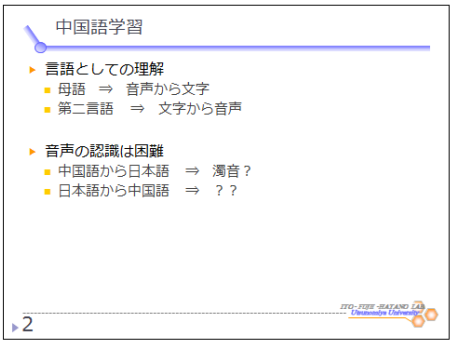


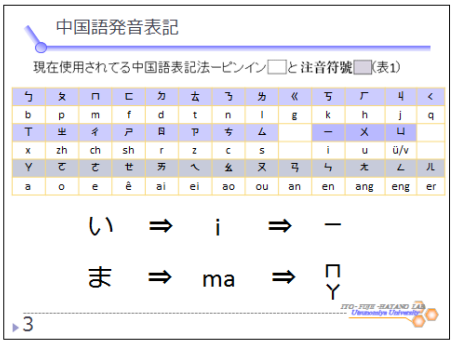
Appendix 10. Contents of evaluate experiment - comparison of effectiveness of applications to improve Chinese tone pronunciation ability

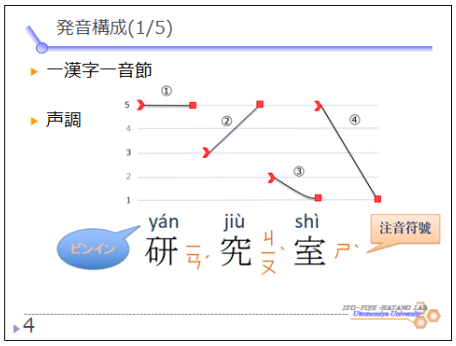
Explanatory material for the experiment:

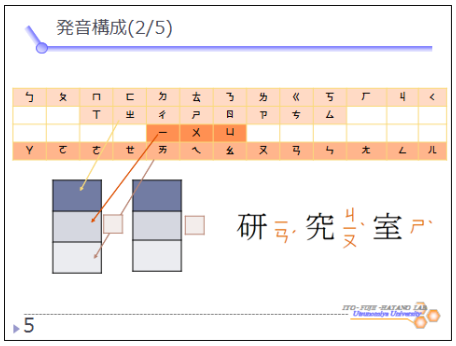


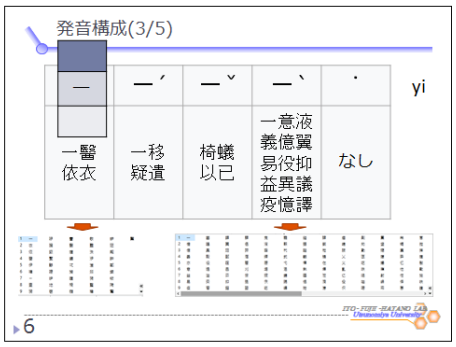


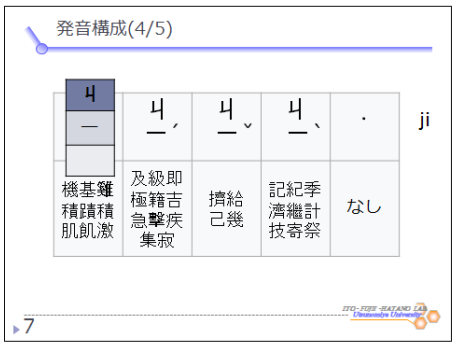












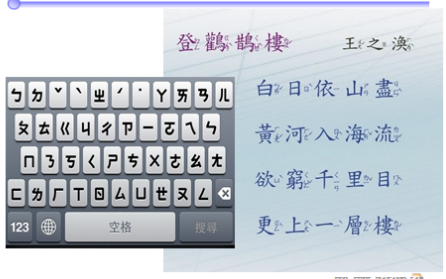
発音構成(5/5)

ㄐㄧㄡ	ㄐㄧㄡˊ	ㄐㄧㄡˇ	ㄐㄧㄡˋ	・
糾揪	なし	九酒久	就舊救究舅	なし

jiu

8

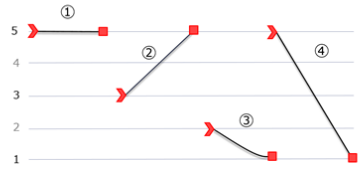
入力、読み



9

声調

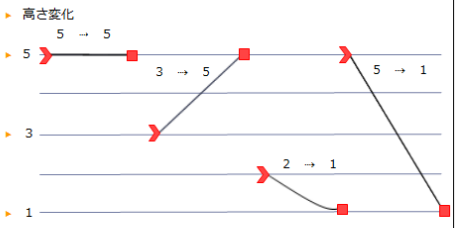
声調イメージ



10

声調の特徴

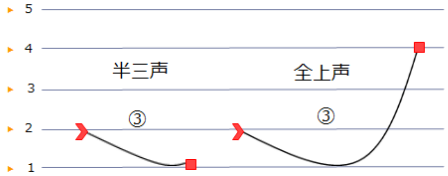
高さ変化



11

声調の特徴

第三声



12

第三声と軽声

ニ ハウ
你 好

OR

ツアウ アン
早 安

おはよう

シエ シエ
謝 謝

ありがとう

13

実験

NHKゴガク 語学講座 ― 声調確認くん

知識 → 会話練習 → 発音テスト

啓発くん

知識 → 感知能力 → 発音テスト

14

実験流れ

アプリの効果

ツールの使用 → 発音テスト → ツールの使用 → 発音テスト

三日間：
??

声調確認くん
168短文
??

啓発くん
84問+20テスト
??

15

NHKゴガク 語学講座 ― 声調確認くん

- ダウンロード数10万～50万　android
- 中国語音声サンプル再生機能
  - 一音節漢字、二音節単語、短文（男女）
- 録音+再生機能
  - 自分で音声確認
- 声の高さの軌跡を音節区切りで表示
  - イメージ確認

16

アプリ使用

- テレビで中国語（168短文）一日30分

17

第三声と軽声

- 二　ハウ
  你　好

- シエ　シエ
  謝　謝
  ありがとう

18

記録

- 一日目の30分　33．もしもし

19

アプリ紹介

啓発くん

- 以下のQRコードでユーザー登録し、氏名は名字でお願いします。

招待用QRコード

20

アプリ紹介

21

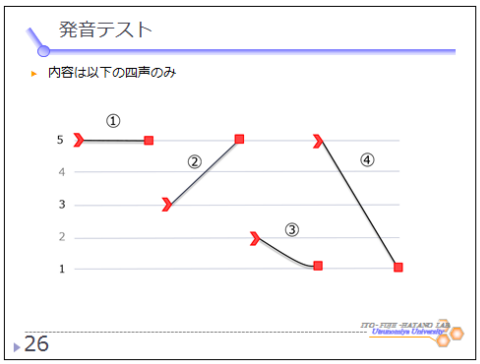
アプリ紹介

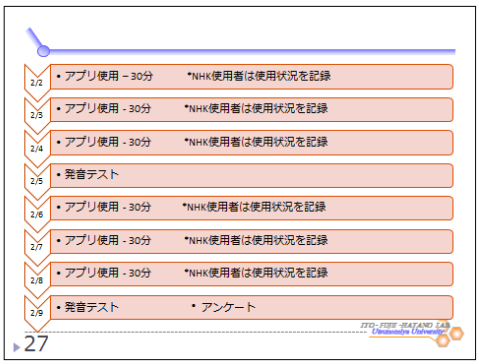
22

アプリの使用

- ログイン

23





The application usage data of two users:

活動ログ　テーマログ　テーマログ

日時	コース名		解答時間	タイプ	点数
2018-2-8 22:0	最後にやる～ 実力テスト！		11分	Study	5点
2018-2-8 21:55	女声で声調認識 - "ハ"、"イ"	（難易度：普通2）	4分	Study	50点
2018-2-8 21:49	初級 - ルール付け声調認識		5分	Study	40点
2018-2-6 21:50	女声で声調認識 - "ハイ"	（難易度：普通1）	10分	Study	20点
2018-2-4 22:56	女声で声調認識 - "ハイ"	（難易度：普通1）	3分	Study	20点
2018-2-4 22:51	ピー音声調認識 - 低音	（難易度：簡単1）	3分	Test	70点
2018-2-3 22:44	男声で声調認識 - "ハイ"	（難易度：普通1）	3分	Test	40点
2018-2-3 22:39	ピー音声調認識 - 大幅	（難易度：一番簡単）	3分	Study	90点
2018-2-3 22:34	ピー音声調認識 - 高音	（難易度：簡単3）	3分	Study	80点
2018-2-3 22:28	ピー音声調認識 - 中音	（難易度：簡単2）	4分	Study	70点
2018-2-2 8:21	ピー音声調認識 - 低音	（難易度：簡単1）	6分	Study	80点
2018-2-2 8:14	中国語声調入門 （説明）		4分	Study	50点

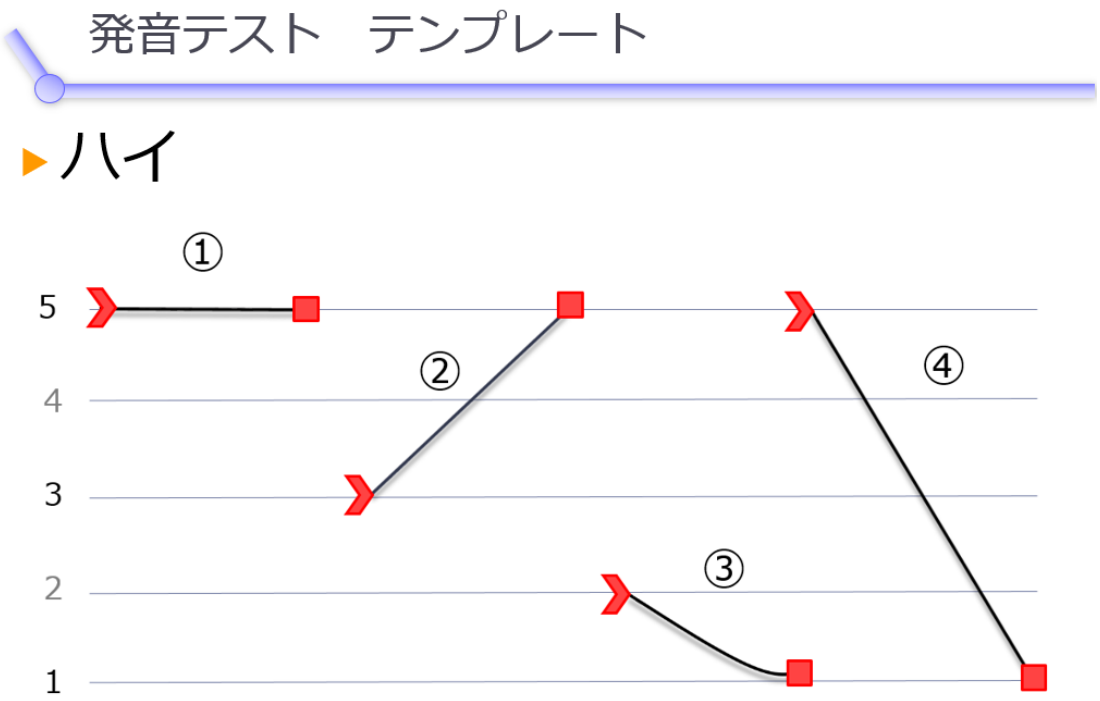
活動ログ　テーマログ　テーマログ

日時	コース名		解答時		
2018-2-4 19:14	女声で声調認識 - "ハ"、"イ"	（難易度：普通2）	3分		
2018-2-4 19:14	女声で声調認識 - "ハ"、"イ"	（難易度：普通2）	3分		
2018-2-4 19:14	女声で声調認識 - "ハ"、"イ"	（難易度：普通2）	3分		
2018-2-4 19:14	女声で声調認識 - "ハ"、"イ"	（難易度：普通2）			
2018-2-4 19:14	女声で声調認識 - "ハ"、"イ"	（難易度：普通2）	3分		
2018-2-4 19:14	女声で声調認識 - "ハ"、"イ"	（難易度：普通2）	3分	Test	100点
2018-2-4 19:14	女声で声調認識 - "ハ"、"イ"	（難易度：普通2）	3分	Test	100点
2018-2-4 19:11	女声で声調認識 - "ハイ"	（難易度：普通1）	2分	Test	60点
2018-2-4 19:6	声調認識 - 一音節漢字		4分	Study	90点
2018-2-3 16:34	男声で声調認識 - "ハ"、"イ"	（難易度：普通2）	3分	Test	91点
2018-2-3 16:30	中国語声調入門 （説明）		3分	Study	100点
2018-2-3 16:25	男声で声調認識 - "ハ"、"イ"	（難易度：普通2）	2分	Test	83点
2018-2-3 16:21	男声で声調認識 - "ハイ"	（難易度：普通1）	3分	Study	80点
2018-2-3 16:17	男声で声調認識 - "ハイ"	（難易度：普通1）	3分	Test	80点
2018-2-3 16:13	ピー音声調認識 - 高音	（難易度：簡単3）	2分	Test	100点
2018-2-3 16:10	ピー音声調認識 - 中音	（難易度：簡単2）	2分	Test	100点
2018-2-3 16:7	ピー音声調認識 - 低音	（難易度：簡単1）	2分	Test	90点
2018-2-3 0:28	ピー音声調認識 - 低音	（難易度：簡単1）	3分	Study	80点
2018-2-3 0:25	ピー音声調認識 - 大幅	（難易度：一番簡単）	2分	Test	90点
2018-2-3 0:20	ピー音声調認識 - 大幅	（難易度：一番簡単）	3分	Study	80点

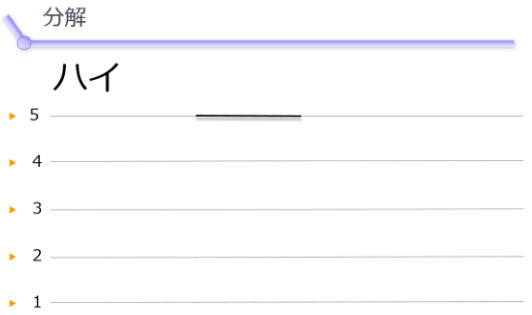
Because of data lost, the experiment result was established by asking the user

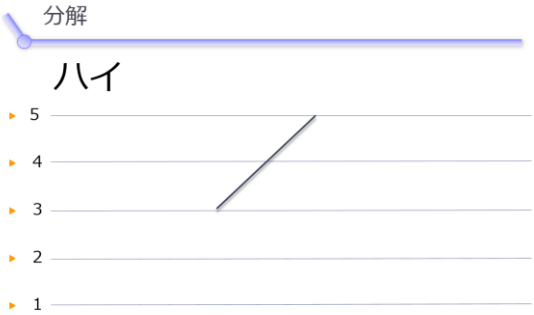
The images presented for pronunciation test:

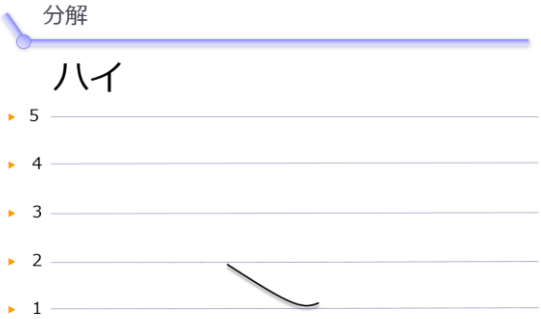
4 syllables tone template:

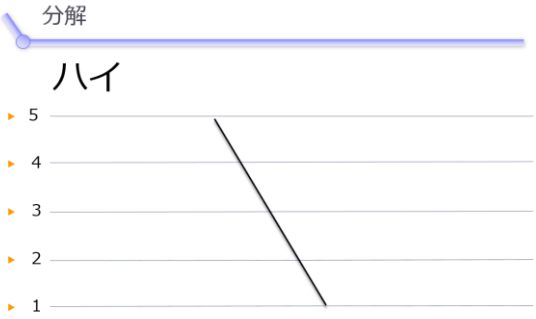


Template syllabication:

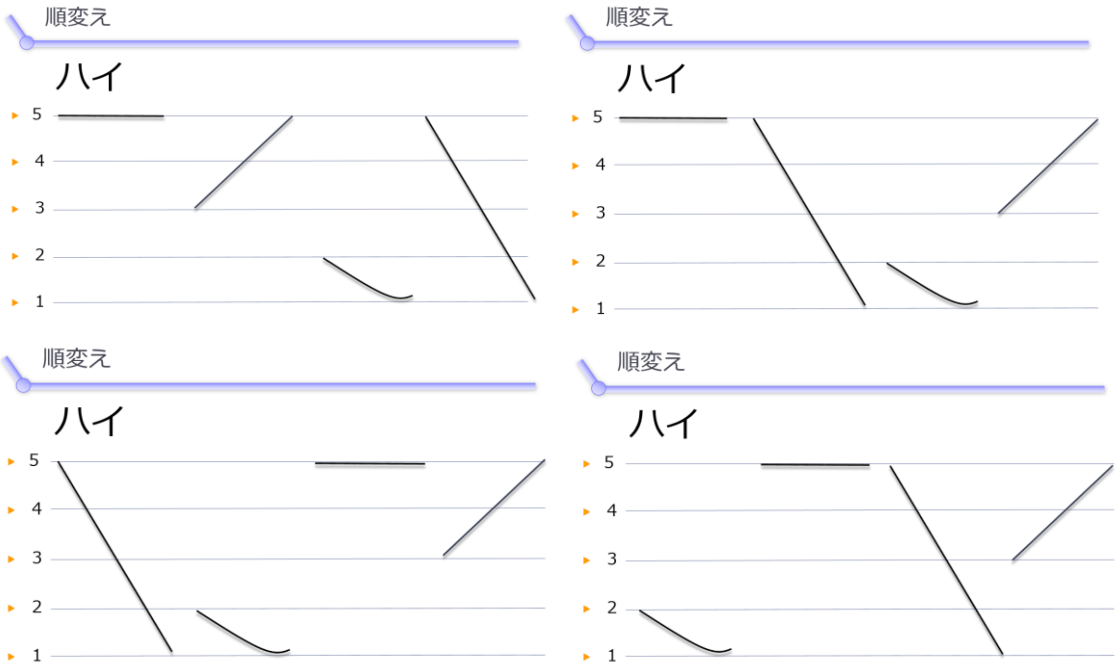






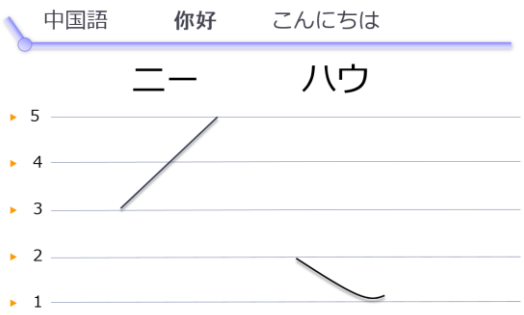


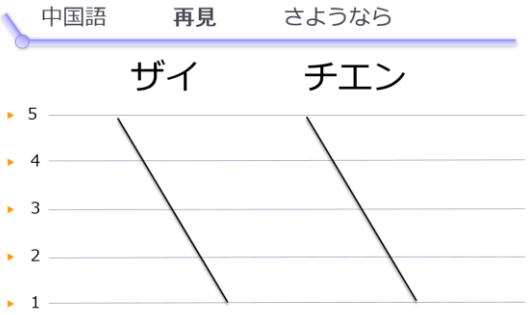
Template reordering (24 patterns):

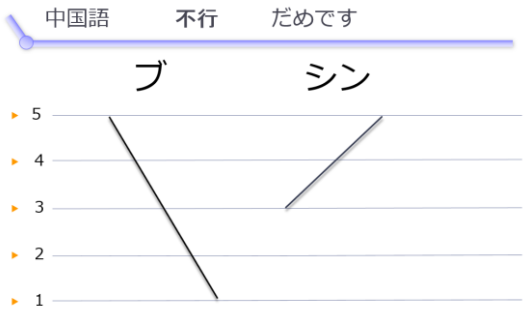


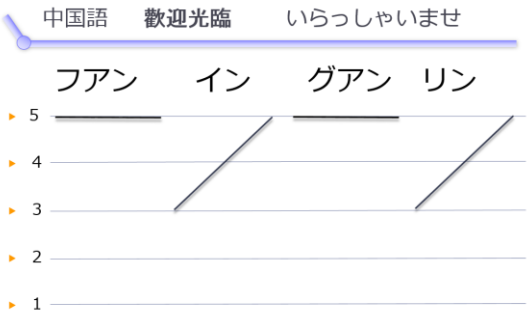
2 syllable phonemes change (15 patterns):

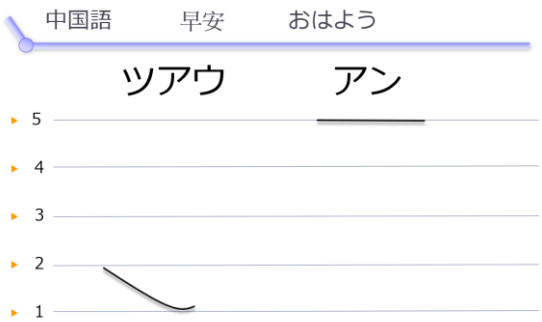
Chinese phrases:

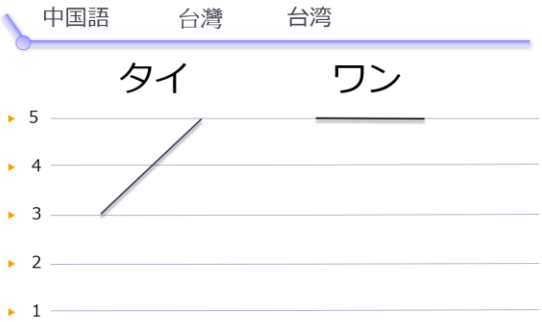


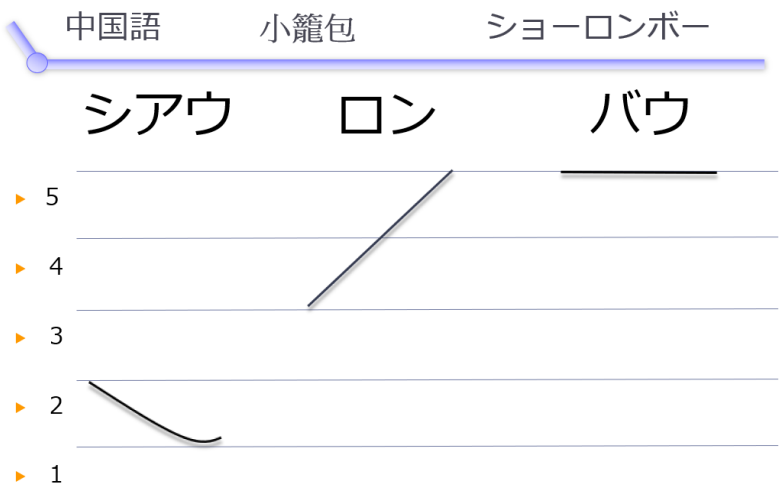












Pronunciation recording sheet:

The pre-test and post-test：

名字：

Tone Template	1234					1234				
Syllabication	1234					3214				
Reordering	1234					1243				
	1342					1324				
	1423					1432				
	4132					4123				
	4213					4231				
	4321					4312				
	3412					3421				
	3241					3214				
	3124					3142				
	2143					2134				
	2314					2341				
	2431					2413				
Phonemes “Ha Si”										
Phonemes “Ka Mi”										
Phrases imitating										
Phrases										

The final-test:

Tone Template	1234					1234				
Syllabication	1234					3214				
Reordering	1234					1243				
	1342					1324				
	1423					1432				
	4132					4123				
	4213					4231				
	4321					4312				
	3412					3421				
	3241					3214				
	3124					3142				
	2143					2134				
	2314					2341				
	2431					2413				

Phonemes “Si Ro”								
Phonemes “Yo Ru”								
Phrases imitating								
Phrases								

Questionnaire for 2 application users:

Chinese pronunciation learning tool provided by NHK's language program

アプリ(NHK)についてのアンケート

1. 使用者について以下のご質問にご回答ください。

名字		年齢		学科・学年	
性別	□ 女性 □ 男性		日本語、英語以外の言語経験		( )年
音楽経験	□ 特になし □ 半年 □ 一年 □ 二年 □ 三年 □ 三年以上				

2.啓発くんについて使用者の評価に近いものに〇をつけてください。

1 全くあてはまらない 2 あてはまらない 3 あまりあてはまらない

4 ややあてはまる 5 あてはまる 6 非常にあてはまる

項目	← あてはまらない					あてはまる →
声調の学習に専心した	1	2	3	4	5	6
(説明会) 声調の理解に役立つ	1	2	3	4	5	6
(声調確認くん) 声調の理解に役立つ	1	2	3	4	5	6
(音声サンプルを聴く) 声調の理解に役立つ	1	2	3	4	5	6
(音声サンプルの真似) 声調の理解に役立つ	1	2	3	4	5	6
(録音、再生機能) 声調の理解に役立つ	1	2	3	4	5	6
(声の高さの軌跡) 声調の理解に役立つ	1	2	3	4	5	6
(男女声) 声調の理解に役立つ	1	2	3	4	5	6
音声サンプルの真似が上手くできる	1	2	3	4	5	6
(発音テスト) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト - テンプレート) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト - 分解) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト - 順番変化) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト - 音素変化) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト - シャドイング) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト - 中国語) 声調の良い訓練になる	1	2	3	4	5	6
発音テストは面白い	1	2	3	4	5	6
中国語に興味がある	1	2	3	4	5	6

項目	← あてはまらない					あてはまる →
このアプリで声調学習は面白い	1	2	3	4	5	6
練習すると、声調の発音が上手くなると思う	1	2	3	4	5	6
中国語は聴き心地が悪い	1	2	3	4	5	6
声調の聞き分けが難しい	1	2	3	4	5	6
声調の発音が難しい	1	2	3	4	5	6
このアプリの学習方法が面倒くさい	1	2	3	4	5	6
練習しても声調の発音が上手くなる気がしない	1	2	3	4	5	6
もっと簡単な課題から練習した方がいい	1	2	3	4	5	6
(一週目) 30 分 3 回のアプリ使用をしっかり行った	1	2	3	4	5	6
(二週目) 30 分 3 回のアプリ使用をしっかり行った	1	2	3	4	5	6

The perceptual training application developed by us

アプリ(啓発くん)についてのアンケート

1. 使用者について以下のご質問にご回答ください。

名字		年齢		学科・学年	
性別	□ 女性　□ 男性		日本語、英語以外の言語経験		(　)年
音楽経験	□ 特になし　□ 半年　□ 一年　□ 二年　□ 三年　□ 三年以上				

2.啓発くんについて使用者の評価に近いものに○をつけてください。

1 全くあてはまらない　2 あてはまらない　3 あまりあてはまらない

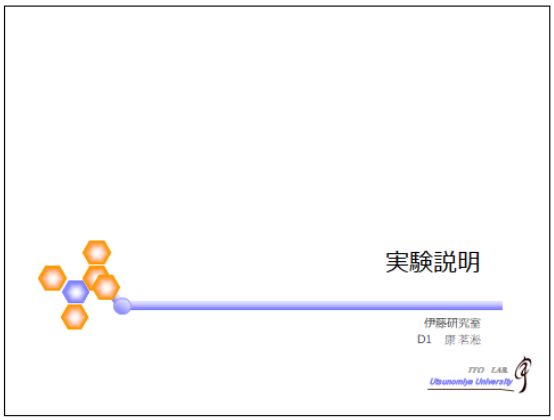
4 ややあてはまる　5 あてはまる　6 非常にあてはまる

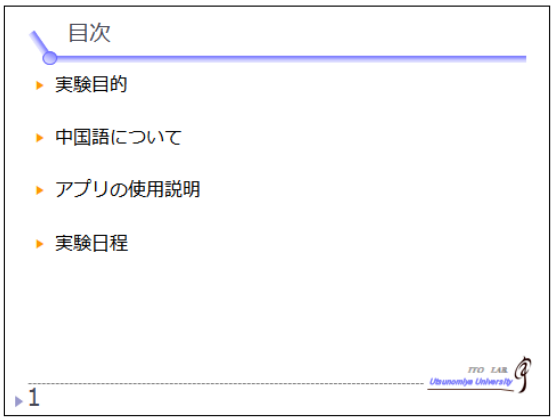
項目	← あてはまらない　あてはまる →					
声調の学習に専心した	1	2	3	4	5	6
(説明会) 声調の理解に役立つ	1	2	3	4	5	6
(アプリ全体) 声調の理解に役立つ	1	2	3	4	5	6
(文字説明) 声調の理解に役立つ	1	2	3	4	5	6
(映像説明) 声調の理解に役立つ	1	2	3	4	5	6
(ピー音) 声調の理解に役立つ	1	2	3	4	5	6
(遅いスピード) 声調の理解に役立つ	1	2	3	4	5	6
(音素を固定) 声調の理解に役立つ	1	2	3	4	5	6
(四音節に四種類声調) 声調の理解に役立つ	1	2	3	4	5	6
(ピー音の三種類高さ変化) 声調の理解に役立つ	1	2	3	4	5	6
(男女声変化) 声調の理解に役立つ	1	2	3	4	5	6
(音素変化) 声調の理解に役立つ	1	2	3	4	5	6
(実力テスト) 声調の理解に役立つ	1	2	3	4	5	6
声調 (ピー音-大幅) の聞き分けが簡単	1	2	3	4	5	6
声調 (ピー音) の聞き分けが簡単						
声調 (男声) の聞き分けが簡単						
声調 (女声) の聞き分けが簡単						
声調 (ハイ) の聞き分けが簡単						

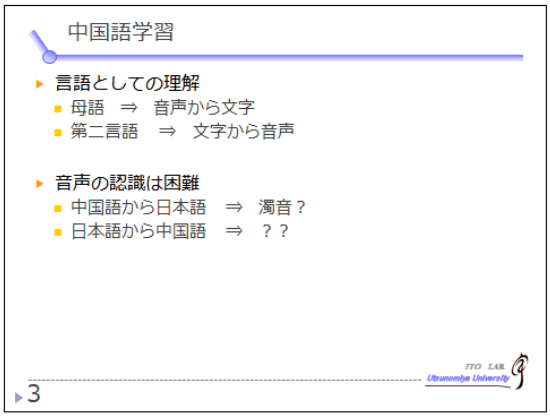
項目	← あてはまらない　あてはまる →					
声調 (ハ) の聞き分けが簡単	1	2	3	4	5	6
声調 (イ) の聞き分けが簡単	1	2	3	4	5	6
1ページ目の内容は簡単	1	2	3	4	5	6
他の声調聞き分け問題もやりたい	1	2	3	4	5	6
(発音テスト) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト – テンプレート) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト – 分解) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト – 順番変化) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト – 音素変化) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト – シャドイング) 声調の良い訓練になる	1	2	3	4	5	6
(発音テスト – 中国語) 声調の良い訓練になる	1	2	3	4	5	6
発音テストは面白い	1	2	3	4	5	6
中国語に興味がある	1	2	3	4	5	6
このアプリで声調学習は面白い	1	2	3	4	5	6
練習すると、声調の発音が上手くなると思う	1	2	3	4	5	6
中国語は聴き心地が悪い	1	2	3	4	5	6
声調の聞き分けが難しい	1	2	3	4	5	6
声調の発音が難しい	1	2	3	4	5	6
このアプリの学習方法が面倒くさい	1	2	3	4	5	6
練習しても声調の発音が上手くなる気がしない	1	2	3	4	5	6
直接単語・会話で発音練習する方がいい	1	2	3	4	5	6
(一週目) 30 分 3 回のアプリ使用をしっかり行った	1	2	3	4	5	6
(二週目) 30 分 3 回のアプリ使用をしっかり行った	1	2	3	4	5	6

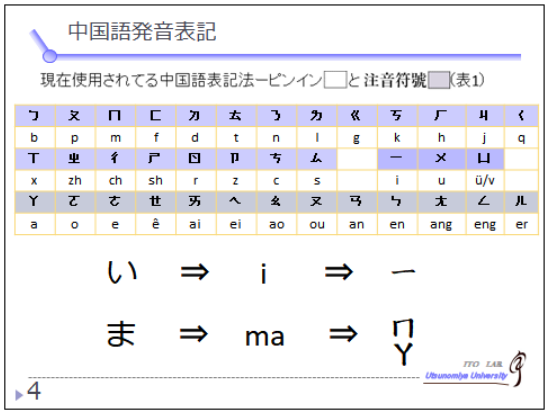
Appendix 11. Contents of evaluate experiment - comparison of effectiveness of applications to improve Chinese tone pronunciation ability

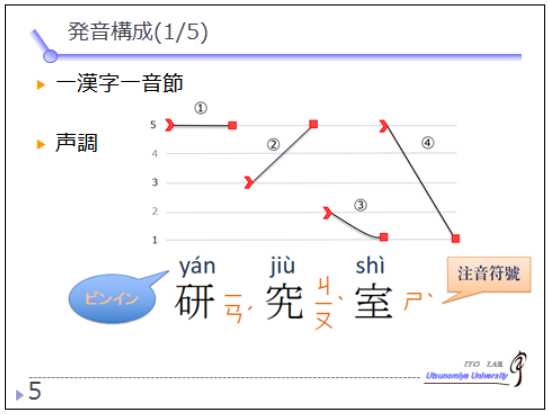
Explanatory material for the experiment:



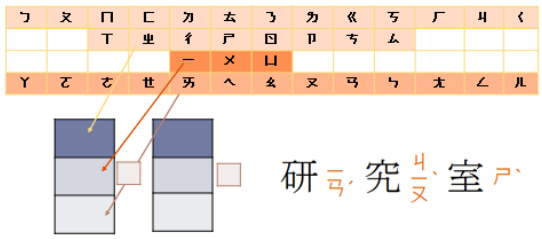








発音構成(2/5)



研 ㄧㄢˊ 究 ㄐㄧㄡˋ 室 ㄕˋ

6

発音構成(3/5)

ㄧ	ㄧˊ	ㄧˇ	ㄧˋ	˙
一醫依衣	一移疑遺	椅蟻以已	一意液義億翼易役抑益異議疫憶譯	なし

yi

7

発音構成(4/5)

ㄐㄧ	ㄐㄧˊ	ㄐㄧˇ	ㄐㄧˋ	˙
機基雞積蹟積肌飢激	及級即極籍吉急擊疾集寂	擠給己幾	記紀季濟繼計技寄祭	なし

ji

8

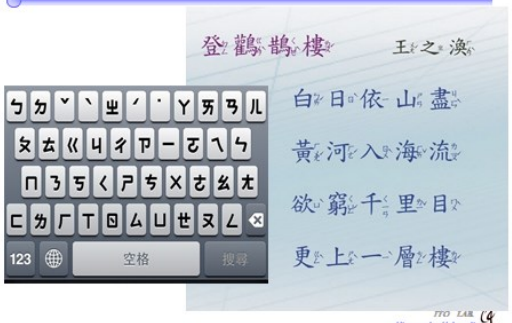
発音構成(5/5)

ㄐㄧㄡ	ㄐㄧㄡˊ	ㄐㄧㄡˇ	ㄐㄧㄡˋ	˙
糾揪	なし	九酒久	就舊救究舅	なし

jiu

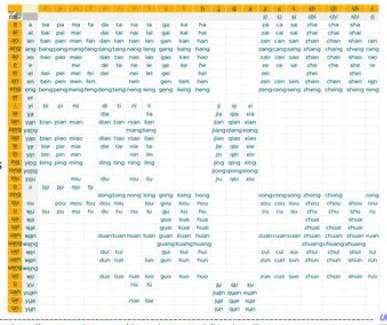
9

入力、読み



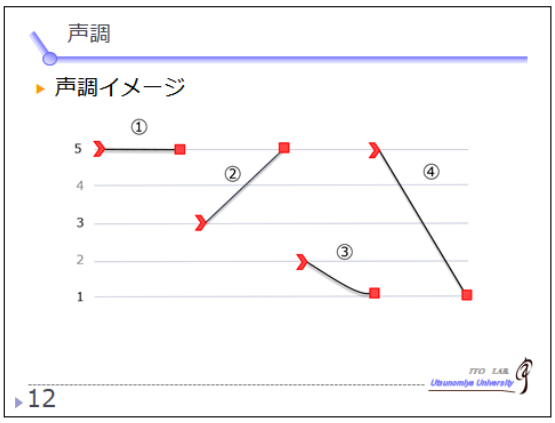
10

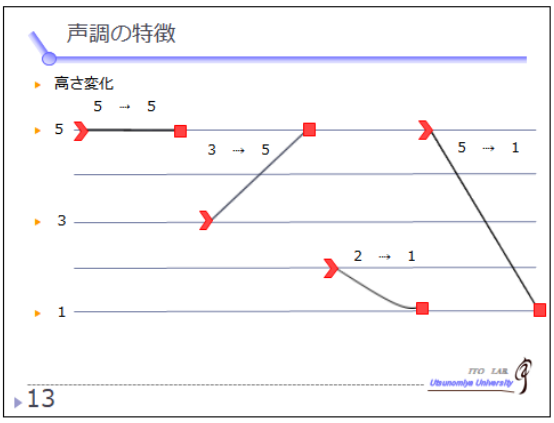
YOYOCHINESE Pinyin Chart Table

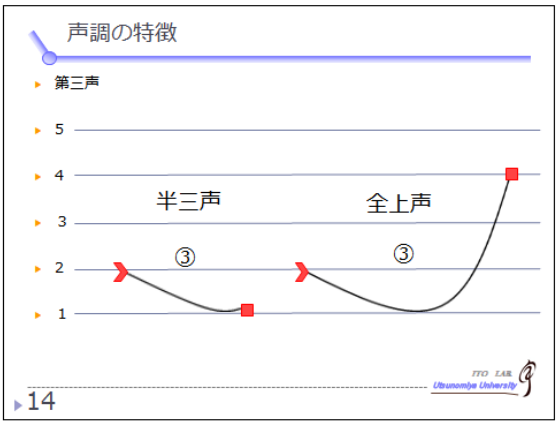


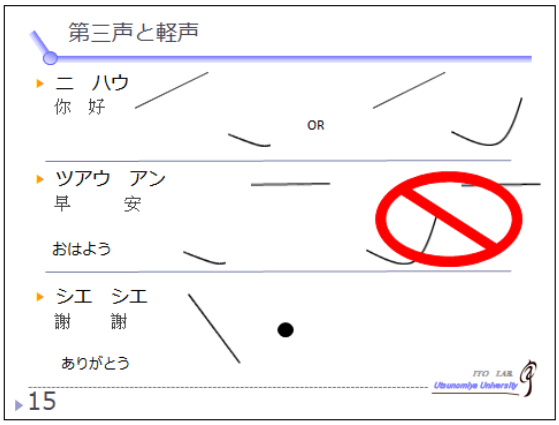
https://www.yoyochinese.com/chinese-learning-tools/Mandarin-Chinese-pronunciation-lesson/pinyin-chart-table

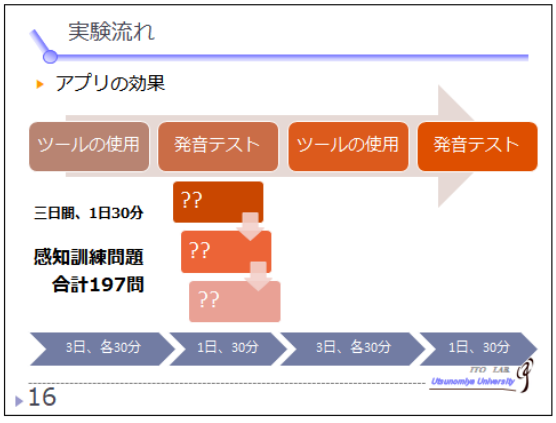
11

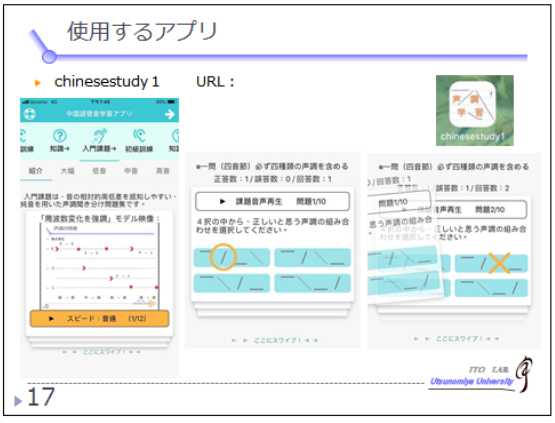

声調
声調イメージ
①
②
③
④
5
4
3
2
1
ITO LAB Utsunomiya University
12


声調の特徴
高さ変化
5 → 5
3 → 5
5 → 1
2 → 1
5
3
1
ITO LAB Utsunomiya University
13


声調の特徴
第三声
5
4
3
2
1
半三声
全上声
③
③
ITO LAB Utsunomiya University
14


第三声と軽声
ニ ハウ
你 好
OR
ツアウ アン
早 安
おはよう
シエ シエ
謝 謝
ありがとう
ITO LAB Utsunomiya University
15


実験流れ
アプリの効果
ツールの使用
発音テスト
ツールの使用
発音テスト
三日間、1日30分
感知訓練問題
合計197問
??
??
??
3日、各30分
1日、30分
3日、各30分
1日、30分
ITO LAB Utsunomiya University
16


使用するアプリ
chinesestudy 1
URL：
chinesestudy1
ITO LAB Utsunomiya University
17

使用マニュアル（エクササイズ画面ありの場合）

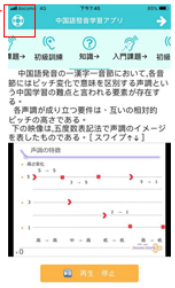
- 起動（エクササイズ画面あり）
- エクササイズ内容を確認
- スワイプ ⇅
- 映像を見ながら発音練習
  - 3つの課題を各一回以上練習する
- “練習しました！”をタップする

18

使用マニュアル

メイン画面

- 項目 - ”知識→” の内容を確認
- スワイプ ⇅
- 初めての利用は以下の順番で
  - ”知識→” ”入門課題→” “初級訓練”
- 初級訓練で上手く声調を認識できない方は画面の左上の浮き輪のアイコンをタップする
- 上手くできた方は右上の矢印タップする



19

使用マニュアル

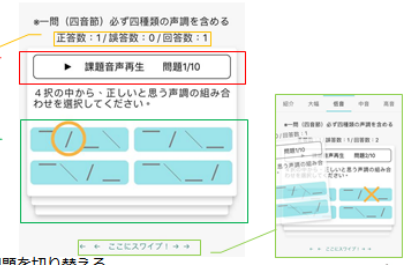
声調聞き分け訓練
（入門課題、初級訓練など）

- 項目紹介を確認
  - モデル映像
- スワイプ ⇆
- 声調問題を回答
  - モデル映像を参考に
- 正答数と回答数を記録

20

使用マニュアル

- 課題音声を再生し、声調の組み合わせを聞き分ける
- 回答状況
- 選択肢
- スワイプで問題を切り替える



21

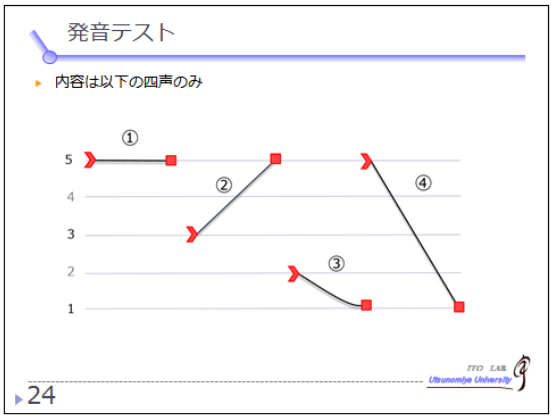
アプリ構成

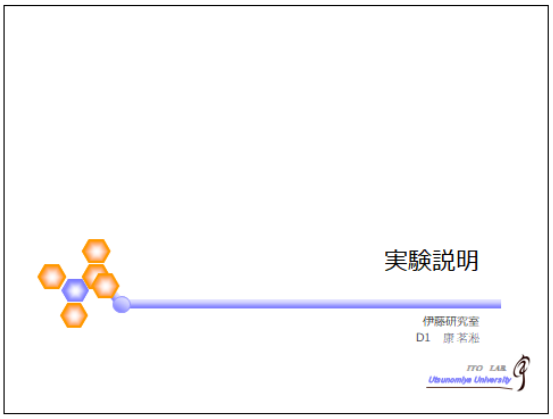
22

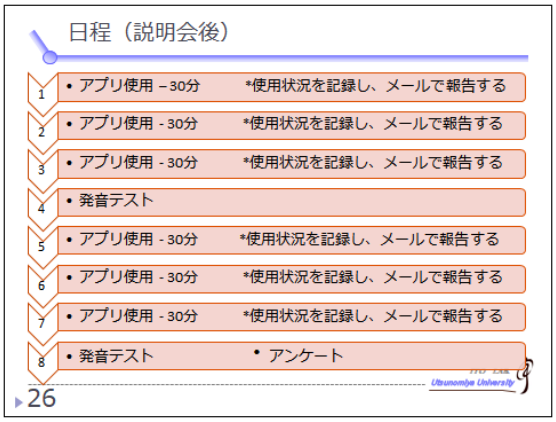
集めたいデータ

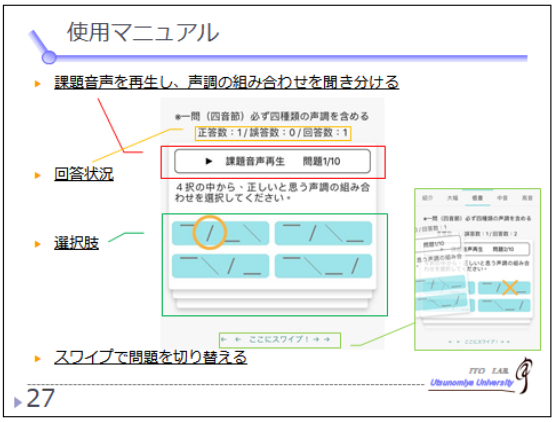
- 発音テスト結果
- アプリの各項目の成績
  - 正答数、回答数
- アプリの使用スタイル
  - 回答数、利用時間、利用項目
- 起動時に表示した発音練習映像の使用、練習状況
  - 毎回起動時に練習するか
  - 練習回数
  - 自分の発音に対する評価

23

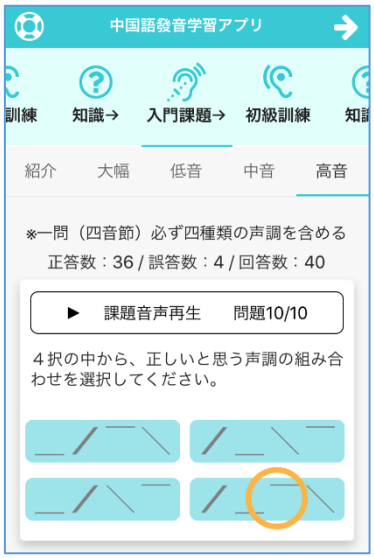


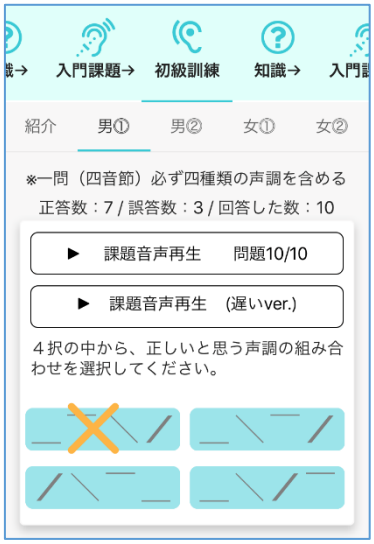


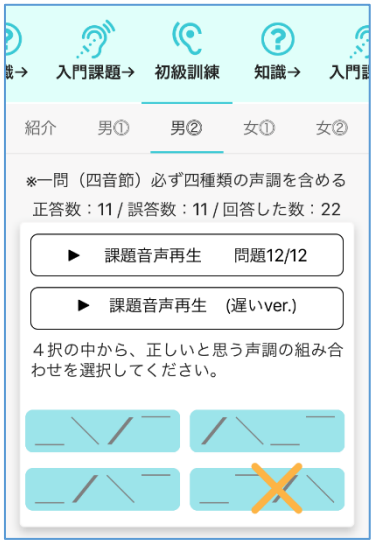


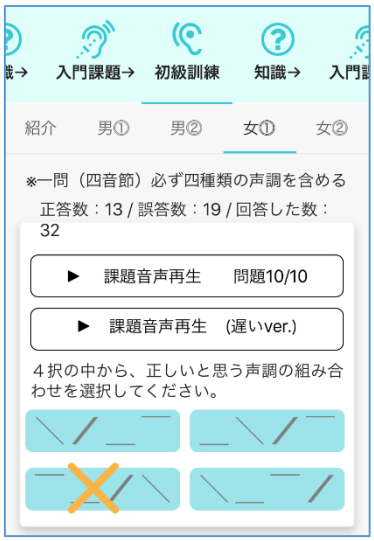


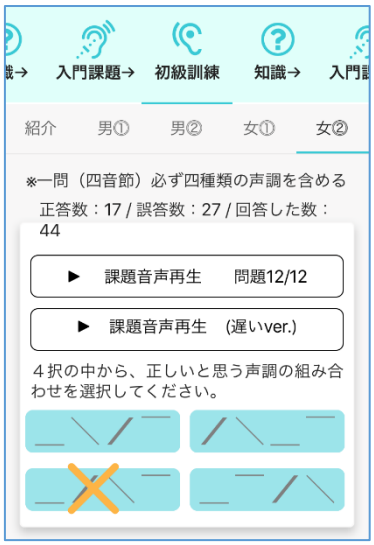
The application usage data of the user had lowest performance:

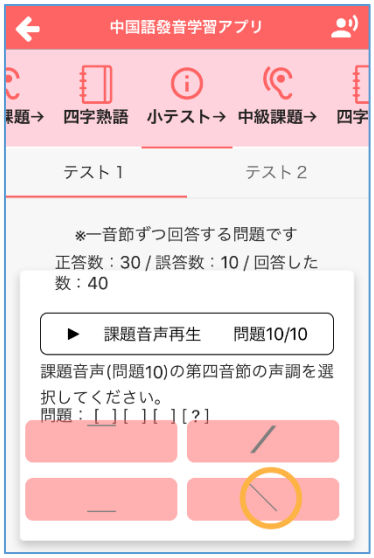

中国語發音学習アプリ
訓練
知識→
入門課題→
初級訓練
紹介
大幅
低音
中音
高音
※一問（四音節）必ず四種類の声調を含める
正答数：36 / 誤答数：4 / 回答数：40
▶ 課題音声再生 問題10/10
４択の中から、正しいと思う声調の組み合わせを選択してください。

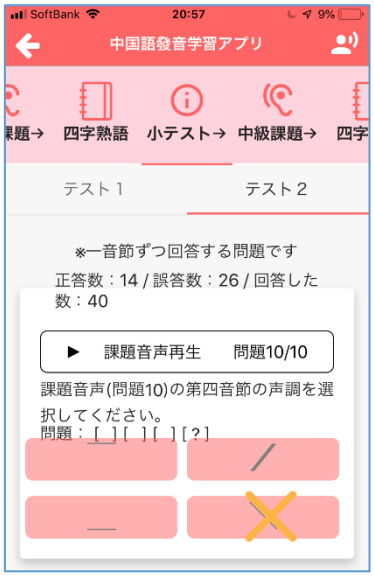

入門課題→
初級訓練
知識→
紹介
男①
男②
女①
女②
※一問（四音節）必ず四種類の声調を含める
正答数：7 / 誤答数：3 / 回答した数：10
▶ 課題音声再生 問題10/10
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。


入門課題→
初級訓練
知識→
紹介
男①
男②
女①
女②
※一問（四音節）必ず四種類の声調を含める
正答数：11 / 誤答数：11 / 回答した数：22
▶ 課題音声再生 問題12/12
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。

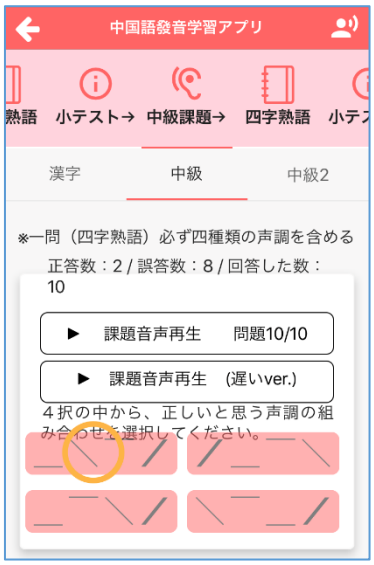

入門課題→
初級訓練
知識→
紹介
男①
男②
女①
女②
※一問（四音節）必ず四種類の声調を含める
正答数：13 / 誤答数：19 / 回答した数：32
▶ 課題音声再生 問題10/10
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。


入門課題→
初級訓練
知識→
紹介
男①
男②
女①
女②
※一問（四音節）必ず四種類の声調を含める
正答数：17 / 誤答数：27 / 回答した数：44
▶ 課題音声再生 問題12/12
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。


中国語發音学習アプリ
四字熟語
小テスト→
中級課題→
テスト１
テスト２
※一音節ずつ回答する問題です
正答数：30 / 誤答数：10 / 回答した数：40
▶ 課題音声再生 問題10/10
課題音声(問題10)の第四音節の声調を選択してください。
問題：[ ][ ][ ][?]


SoftBank
20:57
9%
中国語發音学習アプリ
四字熟語
小テスト→
中級課題→
テスト１
テスト２
※一音節ずつ回答する問題です
正答数：14 / 誤答数：26 / 回答した数：40
▶ 課題音声再生 問題10/10
課題音声(問題10)の第四音節の声調を選択してください。
問題：[ ][ ][ ][?]


SoftBank
0:41
23%
中国語發音学習アプリ
小テスト→
中級課題→
四字熟語
漢字
中級
中級2
一音節漢字声調認識
正答数：8 / 誤答数：12 / 回答した数：20
▶ 課題音声再生 問題20/20
以下の４択の中から、正しいと思う声調を選択してください。


中国語發音学習アプリ
小テスト→
中級課題→
四字熟語
漢字
中級
中級2
※一問（四字熟語）必ず四種類の声調を含める
正答数：2 / 誤答数：8 / 回答した数：10
▶ 課題音声再生 問題10/10
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。


SoftBank 19:52 40%
中国語發音学習アプリ
小テスト→ 中級課題→ 四字熟語
漢字 中級 中級2
四字熟語
正答数：2 / 誤答数：6 / 回答した数：8
▶ 課題音声再生 問題8/8
4択の中から、正しいと思う声調の組み合わせを選択してください。
SoftBank 19:57 40%
中国語發音学習アプリ
中級課題→ 四字熟語 小テスト→
四字熟語１ 四字熟語２
※一音節ずつ回答する問題です
正答数：12 / 誤答数：28 / 回答した数：40
▶ 課題音声再生 問題10/10
課題音声(問題10)の第四音節の声調を選択してください。
問題：[ ][ ][ ][ ][?]
SoftBank 20:00 40%
中国語發音学習アプリ
中級課題→ 四字熟語 小テスト→
四字熟語１ 四字熟語２
※一音節ずつ回答する問題です
正答数：18 / 誤答数：62 / 回答した数：80
▶ 課題音声再生 問題10/10
課題音声(問題10)の第四音節の声調を選択してください。
問題：[ ][ ][ ][ ][?]
SoftBank 22:46 6%
Back サポート項目
モデル映像 一音節 初級
一音節声調認識（簡単）
※モデル映像を参考に回答できます
正答数：12 / 誤答数：8 / 回答した数：20
▶ 課題音声再生 問題20/20
以下の４択の中から、正しいと思う声調を選択してください。
← ← ここにスワイプ！ → →
SoftBank 22:50 14%
Back サポート項目
モデル映像 一音節 初級
※一問（四音節）必ず四種類の声調を含める
正答数：10 / 誤答数：5 / 回答した数：15
▶ 課題音声再生 問題15/15
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。
← ← ここにスワイプ！ → →
中国語發音学習アプリ
入門課題→ 初級訓練 知識→
紹介 男① 男② 女① 女②
※一問（四音節）必ず四種類の声調を含める
正答数：6 / 誤答数：4 / 回答した数：10
▶ 課題音声再生 問題10/10
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。
中国語發音学習アプリ
入門課題→ 初級訓練 知識→
紹介 男① 男② 女① 女②
※一問（四音節）必ず四種類の声調を含める
正答数：18 / 誤答数：4 / 回答した数：22
▶ 課題音声再生 問題12/12
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。
中国語發音学習アプリ
入門課題→ 初級訓練 知識→
紹介 男① 男② 女① 女②
※一問（四音節）必ず四種類の声調を含める
正答数：31 / 誤答数：11 / 回答した数：42
▶ 課題音声再生 問題12/12
▶ 課題音声再生 (遅いver.)
４択の中から、正しいと思う声調の組み合わせを選択してください。
SoftBank 4G 13:03 65%
Back サポート項目
モデル映像 一音節 初級
一音節声調認識（簡単）
※モデル映像を参考に回答できます
正答数：7 / 誤答数：13 / 回答した数：20
▶ 課題音声再生 問題20/20
以下の４択の中から、正しいと思う声調を選択してください。
← ← ここにスワイプ！ → →