RUNESINGER:
A DEMONSTRATION OF
COUPLING KOREAN PHONICS AND SPELLING
TO A VIDEOGAME

by

Ethan Kennerly

A Thesis Presented to the
FACULTY OF THE GRADUATE SCHOOL
UNIVERSITY OF SOUTHERN CALIFORNIA
In Partial Fulfillment of the
Requirements for the Degree
MASTER OF FINE ARTS
INTERACTIVE MEDIA

May 2009

Copyright 2009 Ethan Kennerly
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of figures</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>v</td>
</tr>
<tr>
<td>Abstract</td>
<td>vii</td>
</tr>
<tr>
<td>Overview</td>
<td>1</td>
</tr>
<tr>
<td>Prior videogames decouple linguistics from play</td>
<td>3</td>
</tr>
<tr>
<td>Korean tests playful learning</td>
<td>11</td>
</tr>
<tr>
<td>Become a runesinger</td>
<td>13</td>
</tr>
<tr>
<td>Begin as a baby</td>
<td>15</td>
</tr>
<tr>
<td>Sing to learn</td>
<td>17</td>
</tr>
<tr>
<td>Sing slowly, listen joyfully</td>
<td>19</td>
</tr>
<tr>
<td>Learn to sing</td>
<td>19</td>
</tr>
<tr>
<td>Imagine the action in an environment</td>
<td>21</td>
</tr>
<tr>
<td>Bridge from shapes to glyphs</td>
<td>24</td>
</tr>
<tr>
<td>Recognize a word for food to serve that food</td>
<td>26</td>
</tr>
<tr>
<td>Spell a word to prepare food</td>
<td>30</td>
</tr>
<tr>
<td>Speak and spell a few Korean words</td>
<td>34</td>
</tr>
<tr>
<td>Assess language acquisition</td>
<td>34</td>
</tr>
<tr>
<td>Assess motivation to play</td>
<td>38</td>
</tr>
<tr>
<td>Assessment drives design</td>
<td>39</td>
</tr>
<tr>
<td>Conclusion</td>
<td>43</td>
</tr>
<tr>
<td>References</td>
<td>45</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

Figure 1. In Rosetta Stone online demo of Turkish, a user matches a pronounced word to a photograph that is out of context. 4

Figure 2. In TeLL Me More, visualization automates feedback on the user's pronunciation. 4

Figure 3. In 케비키즈 (KebiKids), "한글놀이 - 엄마가 아파요" (Alphabet Play - Mom Aches) a user selects the Korean word that matches the image in order to make mom feel better. 6

Figure 5. In My Spanish Coach (left), a user rapidly recognizes if a Spanish matches the English word to whack a mole (with the DS stylus). In My Chinese Coach (center), a user selects the English word that translates the Pinyin Mandarin word to click a button. The user may also trace the Simplified Mandarin that translates the English word (right), without context. 7

Figure 6. English of the Dead reappropriates a lightgun shooting game. Translate English into Japanese in order to shoot zombies and monsters. 8

Figure 4. In Tong Hua, a user clicks on character for Mandarin tree to hide behind a tree. 9

Figure 7. In Zengo Sayu, a user wears a head-mounted display and moves a box with a wand. 10

Figure 8. In a Mission to Iraq maze game, a user speaks an Arabic direction into a microphone to cue the avatar to jog in that direction. 10

Figure 9. A poster summarizes the activities, visual style, and composition of letters within Runesinger. 14

Figure 10. In a storyboard, mom feeds baby rice, which the baby has just created when the user sings that word. 16

Figure 11. A ball bounces on each syllable. At first, Hangul is replaced with icons, which makes the user feel more comfortable. 18

Figure 12. When listening or playing back (left), a green ball bounces and green headphones pulse. When recording, a red ball bounces and a red microphone pulses. 20
Figure 13. When Ms. Pac-man says "I eat a pellet," Pac-man can unpack the meaning through imagining his sensations from eating a pellet.

Figure 14. The user sees a simple 3D virtual environment, rendered in OGRE.

Figure 15. The author programs during runtime in Python-OGRE and CEGUI.

Figure 16. A 3D style storyboard is constructed from four primitive shapes.

Figure 17. The scene is comprised entirely of spheres (left), domes (center), cones (right), and the Korean letter ㄱ (giuk).

Figure 18. In a storyboard, mom orders food, which the child creates.

Figure 19. Cooking Dash displays heart for satisfaction of customer, food at each station, and an icon of the food the customer wants in a speech balloon.

Figure 20. The child returns with 김밥 (rice roll) and sings the word that the user recorded.

Figure 21. In storyboard, each letter compresses together inside the syllable block.

Figure 22. In the software, the user clicks on each letter to slide it into the syllable block.

Figure 23. The child returns and plays back 밥 (rice), which the user recorded. Now that the user has spelled the word, that word is written in 한글 (Hangul).

Figure 24. After the user has spelled 밥 (rice), the speech balloon writes 한글 (Hangul) for that word, while the unfamiliar word 주세요 (please give) is still three icons.

Figure 25. Standard Test of Proficiency in Korean (S-TOPIK) beginner level 1, presents one image to match to a Korean word. No English is presented.

Figure 26. HTML-embedded post-test of Korean adapted the style of S-TOPIK image recognition (left) and spelling (right). Grayscale clip art images are displayed.
Figure 27. HTML-embedded five-point Likert scale asks a user to self-report their enjoyment.

Figure 28. Models in Google SketchUp storyboard (left) were exported to Maya and then exported into an OGRE mesh (right). Thus, storyboards visually prototype the look and feel and scene composition.

Figure 29. By focus testing storyboards, it was discovered that: watercolor sketchy edges, masonite overlay and underlay made text less clear; and a frog eating a fly was humorous but irrelevant to the user's needs.

ACKNOWLEDGEMENTS

Several persons contributed to this software in diverse ways. At the time of this writing, those who constructed code or content include:

Music composed by Andrea Chang. Voice over by Allura Lee and Julian.
Recording produced by Joseph Spradley.
December 2008 version of spelling game developed by Tony Hsieh, Lin-Chun Huang, Kuang-Chou Chen, Kitae Bae, Taiyoung Ryu, Chetan Ahuja, and Joseph Spradley.

Maya dotScene exporter programmed by Balakrishnan Ranganathan.

Rbert prototype graphics programmed by Justin Lewis.

All mistakes, and everything else, are my fault.

Many also advised and provided feedback, which informed our work: Cultural advice from Scott Lee. Design advice from Chris Swain. Story advice from Paul De Meo. Animation advice from Scott Easley.

Prototype, storyboard, and software feedback from Peter Brinson, Ian Dallas, Bryan Jaycox, Brandi Wilcox, Ben Chang, Dooyul Park, Maya Churi, Benjamin Sherman, John Brennan, Mike Rossmassler, Jamie Antonisse, Robert Layton, Ala’ Diab, Diana Hughes, Perry Hoberman, Analisa, Ryan De La Torre, Max Geiger, Rhett Fahrney, Mark Bolas, Andrea Rodriguez, Peggy Weil, Jade Brewer, Veronica Paredes, and Charles Sanglimsuwan.

Software feedback from Anne Balsamo, Andreas Kratky, Diane Tucker, Afsoun Yazdian, Daniel Ponce, Elizabeth Swensen, Nahil Sharkasi, Hua Rex Chen, Richard Lemarchand, Tracy Fullerton, Kellee Santiago, Paul Belleza, Linda Brown, Anthony Ko, Mike Stein, Amanda Pope, Valentina Valik, Mindy King.
ABSTRACT

Runesinger is a demo of a PC videogame to practice speaking and spelling Korean. This computer-assisted language learning teaches a few meaningful nouns and verbs through the task of serving food to hungry villagers during the North Korean famine of 1997.

Inspired by melodic intonation therapy (MIT), the phrases are embedded in jingles. While the user listens and sings each note, a bouncing ball synchronizes the phonemes and graphemes. The user learns a subset of 한글 (Hangul), the Korean alphabet, through a spelling puzzle. Subsequently, the user sings to same language subtitling (SLS).

The three-dimensional (3D) virtual environment, user interface, and story were designed explicitly for the phonetic spelling of the Korean language. Thus, Runesinger harnesses the tropes of videogames to practice a foreign language in a meaningful context.

KEYWORDS: computer-assisted language learning (CALL), Korean, 한글 (Hangul), task-based language teaching (TBLT), secondary language (L2), same language subtitling (SLS), neural theory of language (NTL), immersive language learning, serious game, educational videogame, edutainment, restaurant script, assessment-driven design.
OVERVIEW

The process of secondary language (L2) learning generally includes tedious exercises for at least an hour a day for several years. Therefore, improving the emotional satisfaction of the language learning exercises would improve the quality of life for millions of practitioners. Furthermore, motivation to continue using computer-assisted language learning (CALL) is a potential barrier to the software’s efficacy. Apart from the educational content, such drills have little intrinsic motivation. This lack of motivation makes a case for a serious game. During the past decade, videogames have begun to supplement some secondary language learning. For the most part, the educational content has been out of context and the use of rules and play has been trivial.

Runesinger is a demo of a PC videogame to practice speaking and spelling Korean. This computer-assisted language learning teaches a few meaningful nouns and verbs through the task of serving food to hungry villagers during the North Korean famine of 1997.

Inspired by melodic intonation therapy (MIT), the phrases are embedded in jingles. While the user listens and sings each note, a bouncing ball keeps the user synchronized between the phonemes and iconic placeholders for the graphemes. The user learns a subset of 한글 (Hangul), the Korean alphabet, through a spelling puzzle. Subsequently, the user sings along accompanied by same language subtitling (SLS) in Hangul.
Embodied cognitive linguists believe that in order to learn a natural language, the learner must interact with a physical environment (Feldman 158). Although an immersive interface is beyond the scope of this project, Runesinger displays an interactive virtual environment in three dimensions (3D). On screen, iconic objects are rendered in three-dimensions, which according to a neural theory of language (NTL), enhances the neural binding of image schemas to spatial tasks and language use (Feldman 144). The abstract representation of the characters and objects makes them easy to recreate and apply to diverse experiences.

A single English word is shown to teach each Korean word, and no English is provided for the phonetics. Except for introducing the interface and story, the user is gently immersed into the target language. While playing, the user is presented with minimal English, which would interfere with language acquisition. Animation, sound, and graphic design provide context to the learning. In small increments, new letters, syllables, and words are introduced and practiced.

The virtual environment, user interface, and story were designed explicitly for the spelling and phonetic syntax of the Korean language. Thus, a user can mentally simulate physical objects behaving under physical mechanisms that embed mnemonics of the rules for spelling and pronouncing Korean. This suggests rich possibilities for designing for other secondary languages, in which the virtual environment and user interface are tailored to that language's unique rules for syntax and grammar.
This project began with the question: How can a student design a videogame such that, while playing, a user learns a skill transferable to an academic, business, or artistic setting? In the course of a few sessions of fifteen minutes each, Runesinger practices listening, speaking, and spelling of a few Korean words. Assessment-driven design was instrumental in evolving the software to satisfy its dual criteria of language acquisition and motivation to learn.

The innovation of Runesinger is to harness the tropes of videogames for practicing a foreign language in a meaningful context. While language learning software, such as Rosetta Stone or Declan, exists, in Runesinger, the user learns through meaningful play.

PRIOR VIDEOGAMES DECOUPLE LINGUISTICS FROM PLAY

Popular software for computer-assisted language learning (CALL) has relied on a drill and repeat cycle of instruction. Runesinger teaches Korean, so for the Korean language, this has included translation drills of vocabulary and grammar, such as Declan's ReadWrite Korean (Declan Software 1999). It has also included audio-visual matching drills, while being immersed in the Korean language, such as Rosetta Stone: Korean Explorer (Fairfield Language Technologies 2001). It has also included spaced repetition of translating vocabulary on virtual flash cards in Before You Know It: Korean (military version "Rapid Rote Korean") (Transparent Language, Inc. 2006).
Automated pronunciation feedback has notably been missing for the Korean language. Whereas some languages, like German, have detailed phonetic feedback, such as TeLL me More German Premium - Complete Beginner, Beginner, Intermediate & Advanced (Auralog 2004).

Figure 1. In Rosetta Stone online demo of Turkish, a user matches a pronounced word to a photograph that is out of context.

Figure 2. In TeLL Me More, visualization automates feedback on the user's pronunciation.
Motivation to continue using these forms of computer-assisted language learning is a potential barrier to their efficacy. Devoid of educational content, such drills have little intrinsic motivation. This lack of motivation makes a case for a serious game. Although a definition has not been codified, one popular definition of a serious game is one with an agenda besides entertainment, oftentimes to persuade or train. By this definition, all language learning games fall in the domain of serious games.

Until this decade, few attempts at entertaining software have been made. In Korean, matching images to the alphabet is a shallow and out of context activity, such as 한글교실 (Alphabet Classroom) (Arisu Edu, Inc. 1998). Likewise, matching words to images is a shallow and out of context activity, such as 글쓰기 교실 (Writing Classroom) (Arisu Media, Inc. 2000). Users may match and drill Korean, such as in "Korean Activities" in Digital Dialects (Gibson 2007). Matching an arbitrary image to an arbitrary word is not especially meaningful or relevant to the user's life. Repeatedly clicking to match with identical format and challenge under utilizes the user's intelligence.

The recent extension into videogames started through play with a few foreign words in a bright, cartoon environment with a simple story. In Korea, bright evocative images are combined with simple challenges to match images and Korean words, such as 케비키즈 (KebiKids) (Familyschool co. Ltd 2000). These correspond to other language learning activities such as embedding
English alphabet into a story in Starfall (Starfall Education 2007), and matching Spanish words to cartoon images in Yeebee (GV Enterprise, Inc. 2008).

Figure 3. In 케비키즈 (KebiKids), "한글놀이 - 엄마가 아파요" (Alphabet Play - Mom Aches) a user selects the Korean word that matches the image in order to make mom feel better.

The evolution has continued to offer simple games whose winning requires translation. On the PC, users may write Japanese characters to swordfight in Kana no Senshi (Kana Warrior). On the Nintendo DS, besides translating conversational excerpts, users may stay aloft over the water by translating words in 英語が苦手な大人の DS トレーニング えいご漬け (English Training: Have Fun Improving Your Skills!) (Plato 2006). In some Japanese junior high schools, students have improved English vocabulary through drills with the Nintendo DS software, 中学英単語ターゲット 1800DS (Chugaku Eitango Target 1800 DS) (IE Software 2006).
The trend of language learning videogames has started to spread to the United States and Europe. Users on the Nintendo DS may find words, flip cards, and whack moles, all of which require translating between Spanish and English in My Spanish Coach (Ubisoft 2007), or French and English in My French Coach (Ubisoft 2007). Chinese and Japanese were also developed (Ubisoft 2008).

Figure 4. In My Spanish Coach (left), a user rapidly recognizes if a Spanish matches the English word to whack a mole (with the DS stylus). In My Chinese Coach (center), a user selects the English word that translates the Pinyin Mandarin word to click a button. The user may also trace the Simplified Mandarin that translates the English word (right), without context.

Linguists have challenged the efficacy of out-of-context and superficial translation exercises, since goal-oriented activities cue the learner (Hinkel 275). Some games have provided a meaningful context, but the language activities have been a poor fit to the scene. One example that inspired this project begins ten years ago. Typing of the Dead (Sega 1999) reappropriated the animation, audio, and events of House of the Dead (Sega 1998). Instead of shooting with a lightgun, the user fires one shot for each letter typed. This is now being
developed as English of the Dead (Sega in-development), which appears to have translation exercises for the narrative purpose of shooting zombies (using the same graphics as House of the Dead). Write a letter or match a translated phrase from English to Japanese in order to shoot a monster.

Figure 5. English of the Dead reappropriates a lightgun shooting game. Translate English into Japanese in order to shoot zombies and monsters.

Common tropes of most videogames include points, rules, and evocative real-time audio-visual feedback. Even though some applications have incorporated these tropes, the applications have coupled the game mechanisms
very loosely to the rules governing the phonics, syntax, and grammar of the secondary language. Therefore, existing applications tend to have superficial relationships between their form and content. An attempt at more meaningful mnemonics is made in Tong Hua. Users may point and click Mandarin characters through a brief adventure in Tong Hua: Adventures in the Middle Kingdom (Bryan Jaycox and Ben Chang 2008).

Figure 6. In Tong Hua, a user clicks on character for Mandarin tree to hide behind a tree.

Neurolinguists believe that language use is tightly coupled with our brain's ability to navigate and manipulate the physical environment (Feldman 69). Therefore, goals in a three-dimensional environment might, in theory, deepen memory. This hypothesis has only recently been explored. A user may pickup and move boxes while listening to Japanese in Zengo Sayu (Rose 3). A user may adventure in a mythic Japanese virtual environment while being introduced
to spoken Japanese in Kotodama: The Power of Words (Entertainment Technology Center, Carnegie Mellon University 2005). A user may command their avatar to navigate a maze rendered in three dimensions in one of the games of Mission to Iraq (Alelo 2007).

Figure 7. In Zengo Sayu, a user wears a head-mounted display and moves a box with a wand.

Figure 8. In a Mission to Iraq maze game, a user speaks an Arabic direction into a microphone to cue the avatar to jog in that direction.
Based on this analysis of language learning games, this project explored the possibility of developing an application where rules of play are tightly coupled to the rules of phonics, spelling, and syntax in Korean.

KOREAN TESTS PLAYFUL LEARNING

Runesinger is a serious game. It is a demo of a PC videogame to practice speaking and spelling Korean. This computer-assisted language learning (CALL) teaches a few meaningful nouns and verbs through the task of serving food to hungry villagers during the North Korean famine of 1997.

This project began with the question: How can a student design software such that, while playing, a user learns a skill transferable to an academic, business, or artistic setting? To explore this question, language learning is an apt test case. Language acquisition is a relatively objective and arbitrary subject matter for learning; therefore, it lends itself to evaluation. A user either learns to recognize, recall, and reproduce words in a language or does not. Such criteria can inform an evaluation of the software’s learning impact.

The process of secondary language (L2) learning generally includes tedious exercises for at least an hour a day for several years. Therefore, improving the emotional satisfaction of the language learning exercises would improve the quality of life for millions of practitioners. While the advantages of entertainment are too narrow to supplant traditional curriculum, the tedious
memorization and rehearsal portion of language learning appears to be a domain that videogames are suited to supplement.

This project selected the Korean language for biographical and pedagogical reasons. The author worked as a game designer in Korea. Due to endearing treatment by coworkers, the author was motivated to learn, yet he had a great deal of trouble learning Korean through traditional methods (college courses, immersion, textbooks, audio books, flash cards, and computer-based drills and references). Due to insufficient practice, the author has not automated comprehension and production of an intermediate level of vocabulary and grammar.

Korean has pedagogical advantages for language acquisition. The Korean alphabet is the only alphabet designed for learning. 한글 (Hangul), the Korean alphabet, is phonetic and its consonants were intended to diagram positions of the tongue and mouth. Some linguists consider it to be one of the most consistent alphabets in use. In 1443, the administration of King Sejong explicitly designed this alphabet for literacy (Wikipedia).

Korean also has practical advantages for evaluating language acquisition. Korean is not widely known as a second language. Thus, there is an abundant supply of users who do not know Korean. This makes it easier to evaluate the
effectiveness of the application, by ruling out exposure to Korean as an alternative explanation for the user's performance.\footnote{Incidentally, there is a vibrant Korean curriculum at the University of Southern California and Koreatown is two miles away, which provides practical feedback to users on the real-world efficacy of such software.}

In 2002, after localizing typing text for "Survival Typing," a typing game in the Korean casual game QuizQuiz (Nexon 1999), the author began to consider the efficacy of a language learning game. In 2006, a collection of casual games, analogous to My Spanish Coach (UbiSoft 2007), was proposed.

In 2007, after researching a theory of how the brain parses language (Feldman), an approach that embeds the rules of language into the game mechanics was conceived. As an experiment in the creation of Korean language learning software, Runesinger seeks to demonstrate the improved integration of the curriculum and gameplay, accentuating the context of the exercises, and deepening the exercises from recognition and translation into comprehension and production. The application is designed to enhance the mnemonics of the language, by investing in the rhythm and animation of the language.

BECOME A RUNESINGER

Runesinger is a PC videogame using microphone and mouse as input devices, and monitor and headphones as output devices. Through playful activities in a three-dimensional virtual environment, the software introduces a few words and letters of the alphabet in the Korean language.
The user plays as a magical child in North Korea during the famine of 1997. The user sings Korean words to create food and feed the village. To prepare the food, the user spells the words by clicking on each letter in order.

The title Runesinger is derived from ancient Finnish shaman-poets who could read letters (called runes) and recite them in song. The title has no connection to Korean culture. Users preferred and identified with this title more than a more Korean-correct title, such as 한글 히어로 ("Hangul Hero" or "Alphabet Hero") or 한글 지기 ("Hangul Guardian" or "Alphabet Guardian"). The title encompasses the notion of shamanism and magic, a callback to times when
language itself was considered magical. Magical rituals have long been used to change the habits and beliefs of its performers, so invoking such a mystical framework may be conducive to developing the new habit of speaking and spelling what appear to be strange sounds and weird symbols. Since the user does not recognize letters of the Korean alphabet yet, the title is written in the English alphabet, yet is composed into Korean syllable blocks, which incidentally has a pleasing composition. Because Hangul is comprised of square syllables, it is equally legible when written horizontally or vertically, and so the title Runesinger appears horizontally or vertically.

BEGIN AS A BABY

At the start of use, the user wears the headphone and poises an attached microphone near their mouth. The user is informed that they may listen to Korean and speak Korean.

Linguists and educators believe that meaningful contexts enhance interest and retention (Hinkel 275, Schank 50). So the user is presented with a problem, which is inspired by the famine that afflicted remote villages in North Korea from 1997 to 1999 (Noland, Robinson, and Wang 741). The universal need for food spans cultures (indeed it spans species!) and needs no translation.

In the software, the user begins as an infant. This supports three pedagogical constraints placed on the story: immersion, phonics, and playfulness. Very little English is presented to the user, which may interfere with
language acquisition (Pica 263). Simple words of the target language are introduced. Simple content and pace is appropriate for a baby (Mithen 6). The user vicariously identifies with the protagonist (O’Connor), so in order to partially overcome the psychological barriers of post-toddler language acquisition, the protagonist is a child. Of course, the physiological state of childhood cannot be adopted, yet many linguists believe that most adults, while incapable of native performance, are physiologically capable of rapid and deep language acquisition (Hinkel 503).

Figure 10. In a storyboard, mom feeds baby rice, which the baby has just created when the user sings that word.

The first word the user learns is 밥 (rice). This is ideal for teaching because it is composed of only two unique letters, ㅂ (p/b) ㅏ (a), each of whose sounds is familiar in Korean, English, Mandarin, and other languages. 밥 is also central to the story of feeding the village, as rice is the staple of Korean cuisine.
SING TO LEARN

Mother-to-infant speech emphasizes aspects of the phonemes by hyperarticulating the vowels. Such speech holds a baby's attention better than tones used in adult-to-adult speech (Mithen 13). This emphasis is common to many cultures and may accelerate the child's acquisition of their first language. Therefore, the voice actors of Runesinger were directed to speak and sing their all their lines, as if to a baby.

In the software, the mother sings a Korean phrase as a jingle. Jingles are well-known mnemonic devices exploited in advertising, as well as in songs to teach children in many cultures, such as the ABC song in English, or (Three Bears) song in Korean. On first exposure, the tempo of the jingle has been slowed down from 120 beats per minute to 60 beats per minute, in order to facilitate recognition and repetition of each syllable. Runesinger adapts a therapeutic technique to aid pronunciation. In some types of aphasia, a patient cannot distinguish the pronunciation of a word (Aetna). Some speech therapists have reported success at teaching a patient with aphasia how to pronounce a word by humming the word to a melody, which the patient repeats. Then the therapist replaces the hum with the syllable. The patient follows the melody and learns to pronounce the word (Aetna). Speech therapists call this melodic intonation therapy (MIT). Runesinger employs this same principle in each jingle, by masking words and syllables with musical notes. Only gradually is a note replaced with its Korean syllable.
Figure 11. A ball bounces on each syllable. At first, Hangul is replaced with icons, which makes the user feel more comfortable.

While each syllable is being pronounced, a synchronized ball bounces on a block. In 한글 (Hangul), the alphabet and writing system of Korean, each syllable is composed into a square block. Because Korean is a phonetic language that composes the letters of each syllable into a square block, the ball is literally bouncing on the syllable block. Due to this bouncing ball and synchronized pronunciation of each syllable block, users have quickly learned the phonological boundaries of each syllable, which aids their pronunciation and comprehension.
SING SLOWLY, LISTEN JOYFULLY

On the first exposure to a new word, the word is introduced without singing, so that the pronunciation of each phoneme is slow and clear. Then the word is slowly sung (at 60 beats per minute) in a jingle that is unique to ordering that food. In order to keep the user interested, and stimulate proficiency, when a subsequent customer orders that food, the singing speed increases to 90 beats per minute.

At 120 beats per minute, the user may recognize that the jingles are being repeated in the background music (which plays at 120 beats per minute). The same six musical phrases of the ordering and returning food are key melodies in the background music. The melody of the background music, which loops every minute, may subconsciously cue practice of speech. Andrea Chang composed the original melody to have a quick pace, a positive affect, to induce the user to feel active and upbeat. Listening to music that puts a student in a good mood increases their spatial performance (Thompson, Schellenberg, and Husain 248), so this music is intended to enhance learning through inducing a playful and happy mood.

LEARN TO SING

The above method for practicing pronunciation is novel. Therefore, before the user can start to learn to pronounce Korean, the user has to learn the
conventions of the pronunciation interface. The bouncing ball is green when
listening and red while recording. In order to clarify when the user should speak,
recording also has a pink thought bubble instead of a white balloon, a red
microphone instead of green headphones, and an audio chime of three beats,
which is rhythmically similar to a 3-2-1 countdown popular in training games
since Brain Age. The microphone model pulses on each cue. While focusing on
the bouncing ball, the user may detect this pulse in peripheral vision.

![Diagram](image)

Figure 12. When listening or playing back (left), a green ball bounces and green headphones
pulse. When recording, a red ball bounces and a red microphone pulses.

After the user pronounces a novel word, the user manually evaluates their
own phonics, by immediately listening to a playback of their pronunciation, which
should be synchronized to the bouncing ball. Then, according to conventions of
language audio practice such as Pimsleur’s Korean, following the user’s
pronunciation, the expert (the mother) pronounces the target syllables.
Unfortunately the software does not automate feedback about the user’s
pronunciation. While phonetic recognition technology exists, as demonstrated by
TeLL Me More (Auralog) and Mission to Iraq (Alelo), the complexity of phoneme
matching is beyond the expertise of the developers, and outside of the focus of the author's innovation. Even in Mission to Iraq, or in the automated customer service of AT&T, errors in current phonetic recognition can be frustrating, which casts doubts on the feasibility of programming or licensing a pleasurable phoneme matching feature. This is unfortunate, because users expected automated feedback and some users relied on feedback to improve their pronunciation.

**IMAGINE THE ACTION IN AN ENVIRONMENT**

Some cognitive neurolinguists, such as Jerome Feldman, believe that meanings of nouns are bound in the brain to physical activities of the person using those nouns. In such a neural theory of language (NTL), comprehension and production of speech or writing consists of unpacking the symbols into physical simulations involving oneself as the actor (Feldman 215). For example, a person eats; call him Pac-man (adapted from the Japanese for "eating"). Pac-man imagines himself eating. Pac-man sees Ms. Pac-man eating, and is able to understand this by imagining himself eating. When Ms. Pac-man says, "I eat a pellet," Pac-man can understand this through a chain of subconscious imagination: Something eats; Ms. Pac-man eats; Pac-man eats. 1) Pellet is an instance of a trajector; Ms. Pac-man is an instance of a landmark. 2) Pac-man sees Ms. Pac-man eating. 3) Pac-man has a motor schema for eating. 4) Pac-
man eats a pellet. Therefore, to comprehend "I eat a pellet," Pac-man imagines eating a pellet.

Figure 13. When Ms. Pac-man says "I eat a pellet," Pac-man can unpack the meaning through imagining his sensations from eating a pellet.

Human (and animal) brains are adept at performing physical simulations. Therefore, to become adept at a new word, especially in a new language, rather than translate the word into a native language, a stronger bond will be formed by binding it to a physical experience (Feldman 178). Because humans simulate physical experiences in three dimensions, the graphics engine renders a three dimensional virtual environment. Ideally, the user would manipulate and navigate a virtual or augmented environment with input devices and feedback. However, such hardware and interface design is beyond the scope of this thesis.
Therefore, a three-dimensional environment might, in theory, deepen memory. To render a three dimensional virtual environment, the author selected the Object Oriented Graphics Rendering Engine (OGRE), because OGRE is a powerful, mature, open-source, and well-maintained engine with add-ons for physics, audio, and graphical user interface. However, the feedback cycle of C++ programming is prohibitively slow, and management of memory is too tedious, for this experimental software, which has undergone many refactorings. Therefore, the project was authored in the Python-language wrapper to OGRE, called Python-OGRE, which bundles together Newton Dynamics Engine, Open Audio Library, and Crazy Eddy’s Graphical User Interface.
Figure 15. The author programs during runtime in Python-OGRE and CEGUI.

BRIDGE FROM SHAPES TO GLYPHS

The art direction is minimalist collection of mostly primitive geometric forms for a few reasons. Simpler shapes are easier to reproduce, so it is easier for the user to sketch or imagine the shapes seen on screen. Paul Cezanne composed a scene's perspective from cubes, spheres and cones; Biederman has believed that human recognition occurs by parsing visual stimuli into cubes, spheres, and cones (126). Such abstract shapes are easier to map to one's own
identity and abstract meaning onto, making the art direction a bridge between sensation and language (McCloud 52, 142).

Figure 16. A 3D style storyboard is constructed from four primitive shapes.

Figure 17. The scene is comprised entirely of spheres (left), domes (center), cones (right), and the Korean letter ↵ (giuk).

The geometric shapes are also a subtle analogy to language. English, Korean, and most written languages are constructed from atomic elements: letters that comprise words, words that comprise sentences, and so on.
Therefore, the vast majority of the virtual environment is constructed from cubes, spheres, cones, prisms, torus sections, and arches. These were the visual elements to suggest that the visual landscape itself is composed of iconic elements. Even the user interface is three-dimensional and iconic. The mouse cursor, the speech balloon, the text and letters, are constructed in the virtual environment. All models lack texture, have similar shininess, and do not cast shadows. There is also no visible ground. The ground geometry is simply not there. No user has ever commented on this, as they infer ground and accept the monotone ground as part of the art style. The lack of ground is to highlight and bring attention to the objects that are there. Each object that is there stands out, almost like a word on a blank page. Almost every object is a meaningful part of the learning: the food, the characters, and the barest furniture and architecture to establish the setting.

This abstract minimalist art style has pedagogical problems. Although all users identified creatures and simple settings, most users could not easily identify complex and unfamiliar items, such as prepared Korean food or the setting of a restaurant.

**RECOGNIZE A WORD FOR FOOD TO SERVE THAT FOOD**

After the mother teaches the baby the first word, and the user understands how to listen and sing this Korean word, the baby quickly grows up and begins feeding villagers in a restaurant. Many cognitive linguists believe that we
understand language by associating the words with a scenario that relates these words together (Petruck 3), and uses the language to accomplish a task, which is called task-based language teaching (TBLT) (Hinkel 275). Ordering food in a restaurant is one of the best understood scenarios among laypersons and cognitive linguists. Roger Schank and Robert Abelson explicated the restaurant script. In the software, the script has been reduced down to:

Entry conditions: Customer is hungry
Customer sits.
Customer asks for food.
Cook prepares food.
Cook transfers food to customer.
Customer verifies food.
(paraphrase Schank and Abelson qtd Wong 216)

So that more mental resources are available for the novel words, other aspects and characters (such as a waiter) have been merged together.

Figure 18. In a storyboard, mom orders food, which the child creates.

Although the user is more likely to be a customer than a cook at a Korean restaurant, the user is more likely to play as a cook than a customer in a videogame. In Diner Dash (gameLab 2002), the user is a waitress that seats
customers, takes their orders, delivers the food, and busses the tables. The similarity of this gameplay to Schank and Abelson's codification of the common restaurant script is uncanny. Diner Dash established a genre of service games, most of which are irrelevant to this thesis. However, the selection of food by service stations in Cooking Dash, the acceleration of pace in Turbo Fiesta 2, and the music, story, and speech balloons in Miriel the Magic Merchant inspired corresponding aspects of Runesinger.

Figure 19. Cooking Dash displays heart for satisfaction of customer, food at each station, and an icon of the food the customer wants in a speech balloon.

At first, the customer sings the food she desires. The user clicks on the food station (as in Turbo Fiesta 2, Cooking Dash, and Miriel the Magic Merchant). There is no need for a timer to prepare the food, because the user practices singing, which causes the food to be magically prepared. This singing interrupts
the multi-tasking that is a hallmark of the service genre. However, the user's
mind is already burdened with a foreign language that they are learning to sing.

Figure 20. The child returns with 김밥 (rice roll) and sings the word that the user recorded.

When the user returns with the correct food, the customer eats and gives the avatar a heart. As in Diner Dash, this heart represents good feelings. However, instead of money, hearts are accumulated. For an introduction, this simplifies the concepts and emotions. The user is directly rewarded with gratitude and does not have to consider money or how to say money in Korean.

After about three iterations, the user runs out of that food, such as rice. So the avatar goes to the kitchen to cook. In another 3D food service game,
Order Up (2008), the player gestures with a Wii remote to prepare food. Instead, for Runesinger’s pedagogical goal, cooking involves reassembling the letters that spell the word.

**SPELL A WORD TO PREPARE FOOD**

At first, the word is decomposed into its letters and reassembly is a simple short-term memory task. The user recalls which letters and in which order to spell the word. Because 한글 (Hangul) is phonetic, the user may listen to each letter and syllable, which provides hints on how to spell.

The composition of 자모 (jamo or letters) into a syllable block follows half a dozen rules (Wikipedia). Many people new to 한글 (Hangul) mistake the characters for Chinese ideograms, which have no relationship between their radicals and phonemes. To prevent this confusion, a number of steps are taken in introducing 한글. The 자모 (jamo or letters) are color-coded into 자음 (consonants) and 모음 (vowels). The rules for composing 자모 are not intuitive, so they are physically depicted as walls to a box, of whose sides may open to admit the next valid 자모.
The first letter must be a consonant (e.g., ㅂ). The second letter must be a vowel, either that compresses from the right (e.g., 바) or from the bottom (e.g., 보). Each additional letter is optional. If the second vowel is wide (e.g., 보), then the next letter may be a consonant that compresses from the bottom (e.g., 놀) or a vowel that compresses from the right (e.g., 봐). Once a tall vowel has compressed from the right, the next letter must be a consonant from the bottom (e.g., 밥, 빴). There are few more special cases, but these rules are sufficient to read and write most Korean syllables.

Human brains are adapted to thinking about objects in physical space. Because the order of spelling is strict, the letters are represented as sliding into a box and compressing the previous letters in that box.
The 3D font had originally been 궁서 (palace) font, which looks more exotic (e.g., 밥). But 굴림 (gouge) font is a san serif font, analogous to the Latin font, Arial. San serif fonts are easier to distinguish at smaller sizes (e.g. 밥) (Galitz 152). 굴림 makes the salient features of each stroke and the rules for compressing each 자모 into a square block more obvious and regular (e.g. 냥, 바, 밥).

Figure 23. The child returns and plays back 밥 (rice), which the user recorded. Now that the user has spelled the word, that word is written in 한글 (Hangul).

Once the user has spelled the word for the food (such as 밥 for rice), the food is prepared. The user delivers this food. And now, instead of showing an icon in the speech balloon, the Korean is displayed (such as 밥). At this stage,
the user is able to associate the visible grapheme with the audible phoneme. The bouncing ball was first introduced to synchronize subtitles with music since 1925 (Grant 82). More recently, same language subtitles, which are synchronized with the lyrics of music, have been effective at promoting literacy in Chitrageet, telecast in Gujarat, India (Kothari 23). Those viewers, wanting to sing those songs, learned the words, which became visual mnemonic devices to help them remember the words. The same-language subtitling in Runesinger is imperfect. As we now know, each syllable is actually composed of two to six letters, which are highlighted all at once, instead of one by one (as in Chitrageet).

Figure 24. After the user has spelled 밥 (rice), the speech balloon writes 한글 (Hangul) for that word, while the unfamiliar word 주세요 (please give) is still three icons.
SPEAK AND SPELL A FEW KOREAN WORDS

Some users did not realize until being assessed how much they had learned, as each new word is gradually introduced. The user may learn to speak Korean and spell 한글 (Hangul) for three nouns: 밥 (rice), 김밥 (rice roll), 김치 (kimchi). To prevent the user from being overwhelmed by novel phonemes and graphemes, at each new word, the user only needs to recognize one new syllable. The second noun (김밥) has the same syllable as the first noun: 밥. Therefore, the user is only exposed to one new syllable: 김. The third word (김치) shares the second word’s (김밥) syllable 김, so the user is only exposed to one new syllable: 치.

The blue note blocks are placeholders for Korean verbs. The nouns are combined with two verbs: 여기요 (here is) and 주세요 (please give). Through such introduction, in fifteen to thirty minutes, the user easily speaks and spells three nouns and two verbs.

ASSESS LANGUAGE ACQUISITION

There is some debate over what defines a serious game. The methodology clarifies the definition. A serious game is one whose assessment includes not only entertainment (or motivation) but also a change in beliefs or abilities. So Runesinger was assessed by two distinct post-test surveys: motivation to use the software and recognition of Korean.
One of the benefits of Korean as a subject for a serious game, is that language is an arbitrary association of phonemes and graphemes; assessing language recognition has been codified (Korean Institute for Curriculum and Evaluation), and most of the users have never tried to listen to or read Korean.

Assessment for high-functioning language comprehension is contentious among assessors. Automatically graded tests bear a low correlation to practical language comprehension and production, compared to task-based assessment (Wei, Mislevy, and Kanal 10). However, the automatically graded tests have amassed a corpus of data from which to draw conclusions about beginner-levels of proficiency, and accordance with such standardized tests builds confidence that the assessment has some validity by association. The test used by the Korean Studies Institute at the University of Southern California is the Standard Test of Proficiency in Korean (S-TOPIK). The beginner level S-TOPIK includes a vocabulary of 800 hundred words (Korean Institute for Curriculum and Evaluation), for which a student may spend a few hours a week for over thirty weeks to develop, representing about a year of study. Therefore, the scope of the S-TOPIK is too broad for assessing the educational value of a 15-minute lesson.
1. 가: 저것이 뭐예요?
나: ( )이예요.

Figure 25. Standard Test of Proficiency in Korean (S-TOPIK) beginner level 1, presents one image to match to a Korean word. No English is presented.

Figure 26. HTML-embedded post-test of Korean adapted the style of S-TOPIK image recognition (left) and spelling (right). Grayscale clip art images are displayed.

Furthermore, the examples and instructions to the S-TOPIK are in Korean. Many of the S-TOPIK exercises require application and comprehension that is beyond a 15-minute introduction. Instead, two of the simplest formats for questions in the S-TOPIK were adapted with English instructions. The first is recognition, in which a Korean word is matched to one of four images. The second is spelling, in which an image is matched to one of four variant spellings.
Besides the words in the scenario, a couple of other Korean words are included, which enables comparison of how well the user can figure out Korean without having been exposed to it in software.

This assessment is hampered by a few serious flaws in methodology. First, an eidetic user may notice that the words being matched in the recognition phase are the correct spellings, which may then be recalled to answer the spelling questions. Second, there is no control group. Third, the pretest in Korean is brief. Fourth, the assessment is immediate, only a few minutes after using the software, so no claims of long-term memory are possible. Fifth, the author of the software conducted most of the tests, who (despite his best attempts to divorce his deep investment from communication) may elicit favorable responses through subconscious body language.

The assessment argument also suffers from bias. The test population is not random or unbiased. Of approximately three dozen users that have been tested, most are affiliated with the Interactive Media program. Most also have repeated personal contact with the author. Therefore, the average tester has a vested interest in a favorable conclusion.

Despite its flaws, after a 15-minute session, most users recognized the Korean words: 밥, 김밥, and 김치, and some were able to distinguish their spelling.
ASSESS MOTIVATION TO PLAY

For most videogames, motivation is the only criteria. For Runesinger, motivation was a dual criterion, along with recognition of Korean. Motivation is self-evaluated on a five-point Likert scale relative to the last videogame that the user has played. Most users rated Runesinger as equally or more fun than the last published videogame they had played. Follow-up questions helped isolate confusing and frustrating elements, as well as explain their motivation. Surveys were typed by the user at the computer. Many users reported a motivation to feed the people and a feeling of accomplishment. Although few words are presented in the 15-minute demo, a few users felt their hearts were warmed and expressed a new interest in learning Korean, even through other software.

<table>
<thead>
<tr>
<th>Did you have FUN?</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>not at all</td>
</tr>
</tbody>
</table>

Figure 27. HTML-embedded five-point Likert scale asks a user to self-report their enjoyment.

Because emotions are one of the primary criteria for any game, the user's facial expressions, body language, and vocal outbursts were observed. Regardless of reason, events that correlated to chuckles and smiles were storyboarded in further detail. Contrariwise, moments which elicited blank faces, shoulder or neck rubbing, brow furrowing, were noted. These spontaneous and largely subconscious expressions were worth more than the surveys in terms of
empathizing with the user. Notes were compiled immediately after each session, but there was no videotape or other recording. By this method, while programming and animating in the subsequent week, the author empathically guessed how that user might respond to one of the many alternatives for improving the software. When a tester asked for whom this software was designed, the reply was, "It is designed for you." This often freed up the tester to be frank and share their personal reactions, which is key to communicating the feelings that motivate learning.

In manners similar to Jesse Schell's, when users made suggestions, empathic interpretation of the user's feelings was favored over literal interpretation of their speech (389, 417). One of the blessings and curses of this biased user population was their design education. They were articulate and forgiving of incomplete media. By the same virtue, users sometimes rationalized and invoked theories of media. Some of the author's own pontification misled the project from engaging a user's emotions, so instead of transcribing the user's conjectures, the user's feelings were introspected.

ASSESSMENT DRIVES DESIGN

In test-driven software development, a test for the software's correct functioning is programmed before the software (Beck xix). After initial applications eliminated most of the technical debt, the author adhered to this software methodology.
Since tests drove software development toward higher productivity and reliability, assessment drove preproduction, at the stage of storyboards, before specifying the software. Usability engineers refer to this methodology as assessment-driven design (Computer Science and Telecommunications Board). The author walked users through the storyboards with minimal narration. Then the motivation and learning posttest was proctored. This method identified much confusion and interest in the scenario, and the appropriate level of language proficiency. For example, some viewers of the storyboards felt apprehensive upon seeing foreign 자모 (letters) that they did not yet recognize. So, at the beginning, icons are presented as placeholders for those syllable blocks. The icons correspond to three-dimensional objects in the virtual environment.

Assessment criteria helped resolve conflicting requests. For example, some users have requested seeing English transliteration of the Korean. For Korean, transliteration is especially challenging, because most of the phonemes of Korean do not correspond to the phonemes of English. For example, ㅂ does not sound like "p" or "b," but rather somewhere in between. Teaching with English letters impedes learning the phonics by promoting false correlations that engender an unintelligible English accent. Instead, without any visual guide to pronunciation, most of the users’ initial attempts approximate the Korean syllables.
Figure 28. Models in Google SketchUp storyboard (left) were exported to Maya and then exported into an OGRE mesh (right). Thus, storyboards visually prototype the look and feel and scene composition.

The storyboards also served as visual prototypes, because the same SketchUp models of the three-dimensional storyboard geometry were imported into the graphics engine. The storyboard became the visual target for the art style. This enabled early assessment of the graphics. According to user comments on the legibility of text and appeal of color schemes and shaders, the art style evolved. Incidentally, complaints that the models are too crude were expected. Although users had many complaints about the interface, pedagogy and story, most user comments on the imagery were positive, citing the style as charming or cute. Due to lack of experience in graphics programming and more important pedagogical priorities, the exact SketchUp materials and compositing were not replicated.
Figure 29. By focus testing storyboards, it was discovered that: watercolor sketchy edges, masonite overlay and underlay made text less clear; and a frog eating a fly was humorous but irrelevant to the user’s needs.

As is often surmised from the art style, the author originally intended this software for children in the United States. However, because testers available on a regular basis to comment on unfinished software were colleagues, the intended audience evolved into college students in the United States, most of who study design. One of the implications of an educated audience was the appeal among some users for a historical and tragic inciting incident of a famine, which afflicted North Korea from 1997 to 1999 (Noland, Robinson, and Wang 741). A dozen scenarios were storyboarded, with several alternate settings. Of them, most users preferred a more factual, practical and familiar scenario. Most users found this premise more motivating and coherent than whimsical scenarios such as a tadpole that turns into a frog, a magic child at a monastery, a child that rides a mythic Korean phoenix, a rabbit that loves a carrot, and a frog that eats many things.

Testing with previous users became problematic. Anyone who had tested before had been trained by exposure. Detection of some confusions in the
interface or story would yield a false negative because the user had already learned to compensate for them. By extrapolation, advisors, developers, and especially the author, subconsciously compensated for many problems. Therefore, each week, new users were sought.

CONCLUSION

The virtual environment, user interface, and story were designed explicitly for the spelling and phonetic syntax of the Korean language. Thus, a user can mentally simulate physical objects behaving under physical mechanisms that embed mnemonics of the rules for spelling and pronouncing Korean. This suggests rich possibilities for designing for other secondary languages, in which the virtual environment and user interface are tailored to that language's unique rules for syntax and grammar.

In the course of a 15-minute session, Runesinger practices listening, speaking, and spelling of a few Korean words.

That most users were able to approximate pronunciation and recognize spelling suggests potential for many variant designs tailored to other semantic fields, such as playing with a pet, dressing and decorating the avatar, higher-functioning embellishments of the dining scenario.

The methods of assessment were instrumental in improving the software to meet its dual criteria of language acquisition and motivation to learn.
While playing, the user is presented with minimal English, which would interfere with language acquisition. Instead, the user is gently immersed in Korean. Animation, sound, and graphic design provide context to the learning. In small increments, new letters, syllables, and words are introduced and practiced.

Embodied cognitive linguists believe that in order to learn a natural language, the learner must interact with a physical environment. Although an immersive interface is beyond the scope of this project, Runesinger displays an interactive virtual environment in three dimensions (3D).

The innovation of Runesinger is to harness the tropes of videogames to motivate practice of a foreign language in a meaningful context. While Korean language learning software, such as Rosetta Stone or Declan, exists, in Runesinger, the user learns through meaningful play.
REFERENCES


