Understanding the “Galakei” Appeal; A Study of Language and Interaction

Lee Beckwith
Indiana University
901 E. 10th street, Bloomington IN
leebeckwith.design@gmail.com

Qualitative research methodologies were used in order to better understand how well currently available mobile phone interfaces are designed to match the way Japanese is thought of by native speakers. A conceptual design is presented at the end as a proposal for how to better design a mobile interface for writing Japanese.

Japanese, linguistics, eye-tracking, mobile phones, keitai, language input, kana, mora, design.

1. INTRODUCTION

Today we see in Japan a changing of the guard in mobile devices. The mobile phones developed in Japan, the so-called “Galapagos-island keitai” phones, or “galakei” for short, are competing with touch screen smartphones like the iPhone and Android. While the touchscreen phone offers flexible interfaces with many possibilities for use and interaction, there remains an attachment to the older galakei with simple hardware buttons among Japanese people. Often one can observe Japanese people using multiple devices, to suit the needs they have for both touch screen and keitai. This study seeks to present a model of the Japanese speaker in relation to writing messages on mobile phones, and thereby explain and predict why some mobile interfaces will see more success than others. This research focused on the entry of phonetic kana (かな), not kanji selection via input method editors.

Three perspectives of research were done at the National Institute of Informatics in Tokyo Japan, in order to understand this topic. First, a historical and artifact perspective is taken. The Japanese language and its artifacts are examined, followed by a case study showing how Japanese users have adapted technologies to write in Japanese. Using this evidence we present how Japanese people are taught the Japanese language, and how they understand and think about it. Next, ethnographic studies were done of mobile users to understand the real-life circumstances that factor into their mobile phone choices. The final perspective is an eye-tracking experiment was designed and tested on 10 participants, who wrote messages on both hardware button phones and touch-screen phones. In our experiment our users typed messages on their mobile phone for 10 minutes, and were asked to use both a smart phone and a “galakei” phone. This experiment was done to test our hypothesis that “galakei” should be easiest to write messages, and was analyzed using ELAN linguistic analysis software. Finally, we derived insights from the research done and make suggestions for further research or design. A concept is presented at the end for a hyper-ergonomic solution to text-entry that attempts to both accurately represent the way Japanese is taught and understood, as well as be faster and more complete in its method of text entry. The design concept is presented as an example of a design direction for future development, as well as an example of how to design a mobile UI with the Japanese speaker in mind.

2. HISTORICAL AND ARTIFACT EVIDENCE

When Japanese children first learn to speak, the first thing they are taught about Japanese is kana, the Japanese version of an alphabet. Children will repeat speaking the sounds, “あいうえおかきくけこさしすせそ…etc.” (pronounced like: a i u e o ka ki ku ke ko sa shi su se so…”) routinely as soon as they are able to speak, in order to learn the basic building blocks of speaking Japanese with correct pronunciation.

From here on this paper will use the following kinds of scripts. Kana are the phonetic written “alphabet” of Japanese. Of these, there are two kinds: hiragana, and katakana. Hiragana, are the most

fundamental Japanese written "letters" and is the first written Japanese taught to children. The other kana is "katakana", which is generally reserved for foreign words or occasionally brand names. In Japanese, there is also kanji, which are old Chinese characters that are still used in Japanese written language for most words. Adult-level reading usually consists of mostly kanji with hiragana for Japanese words that do not have a Chinese character, prefixes and suffixes, particles, and modifiers. I will only use one word in kanji. Finally, there is "romaji", which is the Romanized alphabet version of Japanese kana sounds. Romaji most accurately reflects the way Japanese sounds using alphabet.

2.1 On Directionality and Significance

Observe the following gojuon chart. The characters are meant to be read from top to bottom, right to left. In the modern era, Japanese people are familiar with the following ways of writing and reading the Japanese language, in the following contexts:

- Vertical columns, read top to bottom, right to left: This is the most traditional. Students learn this in their very first years. Books are still widely written in this way, and most literature consumed is written in this fashion, including novels, non-fiction, and manga (comics).
- Horizontal rows, read left to right, top to bottom (like English): This is usually for short messages. Examples include restaurant menus, text messages, websites, in-game dialogue, etc. Currently websites are mostly not equipped to write in right-to-left, although there is a CSS class to do so recently, it has not seen widespread adoption yet.

The following image is the table by which the kana are taught from the first years of life for modern Japanese people. It is called ごじゅうおん in hiragana, written 五十音 in kanji (old Chinese characters used in Japanese) meaning "fifty sounds", and pronounced like "gojuon" in it's Romanized spelling (romaji).

This chart is not the only way in which the Japanese language is conceived of. The way Japanese people would speak their "alphabet" the way English speakers speak "a b c d e f …" etc., Japanese people would say, "a i u e o ka ki ku ke ko sa shi su se so" and so on, through the gojuon chart. The spoken language and written language are reflected in the gojuon. This way of ordering sounds by the vowel sound "a i u e o" (pronounced "ahh ee ū eyy ohh") is done so in languages that use mora.

ん	わ	ら	や	ま	は	な	た	さ	か	あ
	(ゐ)	り	(い)	み	ひ	に	ち	し	き	い
	(う)	る	ゆ	む	ふ	ぬ	つ	す	く	う
	(ゑ)	れ	(え)	め	へ	ね	て	せ	け	え
	を	ろ	よ	も	ほ	の	と	そ	こ	お

Figure 1: *Gojuon chart*

2.2 Mora

The Japanese language uses "kana", which are the fundamental unit of the spoken and written language. In linguistics, Japanese spoken sounds are called "mora(s)", and very few languages share this way of speaking. A mora is briefly a sound that takes one unit of time, and is most similar to what we would think of as a consonant and vowel together. Since every written character, or kana, is one mora, every character occupies equal length of time while being spoken. To give an example, the word Japan in Japanese is pronounced "nihon", and written in kana as にほん. In English, it would appear that "nihon" is two syllables, "ni" and "hon". However, if you observe the kana spelling there are three characters, に、ほ、ん, so it would sound like "ni-ho-n". Therefore in Japanese this would be a 3-mora word, but 2 syllables if it were interpreted as an English word. In addition to the time element, the moras determine the order of the vowels. In the example of the child repeating the Japanese "alphabet", the vowels are in this order: あいうえお ("a i u e o" in romaji), and every consonant afterwards follows the same vowel order.

The reason this is significant is because interfaces developed for English speaking users are inappropriate for Japanese users because the languages are so distinctly different, in how it is written, spoken, and thought of. Any language that uses the alphabet does not share the same challenge that speakers of languages that do not. Later, we will see how the iPhone interface conflicts with the ways Japanese is written and taught.

In the case of Japan and Japanese device designers, we saw the design of devices with text input interfaces that were suitable specifically for Japanese people, and because Japanese language is only spoken in Japan, and because of restrictions that mobile carriers put on their services, these devices were restricted to the Japanese domestic market. The mobile phones that seemed to evolve in their own independent direction in Japan were dubbed "Galapagos keitai"

to describe uniqueness of Japanese phones, and they are called “ガラケイ” (galakei) for short.

2.3 Pokebell: A Case Study

The instance of the pager use in Japan is one case where users have used the 5x10 kana grid in mobile telephony. The pager in Japan was initially used for sales people and technicians who are away from the office and need to be contacted. Like normal pagers, the device only transmits the number to be dialed on a nearby pay phone. This was the case for most of the history of the device, until Japanese youth began using pagers in the early-to-mid 90’s to send messages to their friends. They sent messages by assigning numbers to the 5x10 kana chart. Aside from regular usage of transcribe kana messages, they also used many puns by combining the number meaning with the kana meaning.

漢字変換モード（漢あP）

先に押すキー ＼ 後に押すキー	1	2	3	4	5	6	7	8	9	0
1	あ	い	う	え	お	a	b	c	d	e
2	か	き	く	け	こ	f	g	h	i	j
3	さ	し	す	せ	そ	k	l	m	n	o
4	た	ち	つ	て	と	p	q	r	s	t
5	な	に	ぬ	ね	の	u	v	w	x	y
6	は	ひ	ふ	へ	ほ	z	?	!	ー	／
7	ま	み	む	め	も	¥	&	☃	☎	@
8	や	（	ゆ	）	よ	¥	#	↵	♥	スペース
9	ら	り	る	れ	ろ	1	2	3	4	5
0	わ	を	ん	゛	゜	6	7	8	9	0

Figure 2: *Pokebell mapped to Gojuuon*

Later, companies began to release pagers that were capable of displaying Japanese kana, but not long before mobile phones (at that time they were “personal handy phone” or “PHS”) took over in use and popularity. The significance of the ポケベル example is that this is a case where users, young Japanese people, have ingeniously created a way of communicating to each other through a medium that it wasn’t intended for, and of course they used 五十音.

Electronics companies took notice, and the obvious choice of design on mobile phones, “ケータイ”, employed superimposing the kana on the 10-button grid. The buttons 1-10 are the vertical columns of the 五十音, and pressing the button repeatedly selects the row. As an example, to select the fifth character of the first row on the

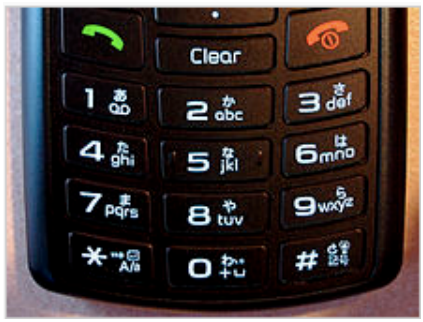

Figure 3: *Standard 10-button keitai layout*

gojuon grid, お, you must press “1” five times. The user would see an あ appear where he/she is writing, and watch as あいうえお are cycled through as they press “1” repeatedly.

2.4 Smartphones and the Japanese Language

The origin of this study is the observation I made of the many ways that smart-phones are attempting to improve upon Japanese language input. I will show some examples and give critique.

The first and most obvious candidate for critique is the iPhone and iPod Touch interface. Many smartphones mimic the design that Apple produced in the organization of kana. Apple used the same 10-button grid in the layout as the older keitai in visual organization, but added a new way of interaction. In the hardware button keitai paradigm, the user would press a physical button over and over, but with a touch-screen this is not possible. Apple mimicked this interaction by allowing the user to tap the glass screen repeatedly in the same manner, to the same effect. But they also included a way of swiping the thumb or finger in four directions away from the central button. The kana are arranged in a circle around the あ-row kana. So, since every button is the top-row on the gojuon, tapping the button will select the あ-row, and then to select one of the four other vowels, the user would slide their finger in one of four directions. If the user holds down a button, the four options around the central button will appear as in the next image. The skilled user will simply swipe in one direction. In this case, the four options do not appear, just the kana selected.

Applying Fitt’s law, the user simply has to slide their finger in one of four directions, making for a lower rate of user error than if the characters were arranged in a straight line as in the gojuon. The gojuon is good for reading and writing, but not a flat piece of glass and a thumb. The problem with this model of interaction, however, is that the kana are not arranged in a way that makes sense other than within the system they created. In Figure 4 な is in the center, and the circle begins on the left, with に.

Figure 4: *Numbers in brackets correspond to Gojuuon ordering.*

Why does the circle begin on the left, and then goes around the center clockwise? A more sensible approach would be to mimic the analog clock and put the next character at 12 o'clock, instead of putting the next character immediately at 9 o'clock. Observe in Figure 4 how the order of gojuon, numbered using brackets, creates a clockwise circle that starts at the left.

Not only are circles not found in the writing or reading of Japanese, but starting from the left makes just as little sense. This is clearly a case of speed and usability over designing with a deep understanding of Japanese language. With the competitors, like the one above, similar UI's have been developed that employ some kind of circular design around a central button.

The design in figure 5 makes the most sense in terms of reflecting Japanese out of the given examples. In the above examples. In the this example, a fan shape is employed, which instead of having the first kana in the center, the first kana appears on the left. So あいうえお are ordered from left to right.

3. MOBILE PHONES IN REAL LIFE

In an effort to understand the real-life issues that determine why some people choose to adopt a smartphone, remain loyal to keitai, or have both devices, I studied two people via ethnographic observation. My method was to follow the participant for one day with a camera, recording phone usage, and to inquire about the purpose of

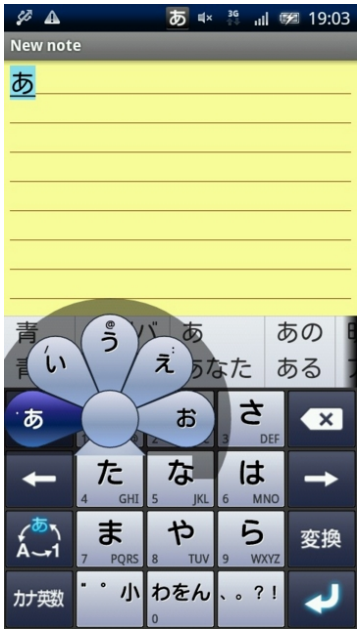

Figure 5: *Fan-shaped kana ordering*

use. Subject 1 was a 20-year old female college student, and I spent an afternoon shadowing her and inquiring about how she used her iPhone. She did not have a keitai, but she had admitted that keitai is easier to write messages on. During this afternoon, Subject1 was in-between classes and was trying to arrange a meeting place with her dance group to rehearse a dance performance. The interesting content I gathered from this was that they had simply decided to meet somewhere in-between 4th and 5th period, but had not decided a particular time or place, but they spontaneously arranged in more specificity as the meeting time approached.

Subject 2 is a 30-year old female educator, and she uses two devices: an iPhone4 and her au "galakei". I selected her because of the fact that she uses two devices, sometimes simultaneously. I followed Subject 2 for an evening, recorded her usage with a video camera, and occasionally inquired about what she was doing.

I asked in the beginning of our excursion why she is a "double-user", as we had come to call it. She explained that she uses the iPhone for work-related communication and her keitai for personal communication. The reality was that she used both devices for various uses. During our study, she frequently looked up maps on her iPhone, and emailed the details to her husband via her keitai. She explained that she felt writing messages is much easier for her on her keitai than on her iPhone. She also explained that it was free to send SMS via her keitai, but not with her iPhone, so the recipient user factored into which device she chose to send a message, and using which protocol.

Figure 6: *Subject 1 arranging a meeting place with her friend.*

3.1 Microanalysis: Eye Tracking

After observing some possible reasons why people chose to use one device or another in their real life, I wanted to see if there were reasons that were beyond what people could describe. I sought to observe the use of these devices in direct comparison with each other. My hypothesis was that the older keitai should be faster than the iPhone in writing, and with fewer mistakes.

3.1.1 Experiment Design

The experiment took two groups of participants, one identified as primarily “keitai” users, and the other as “smartphone” users. There were 5 of each, making 10 total. Each participant would be given two tasks, both identical in instruction but with a different device from the one they identified as their “main” device. This was done so that there would be no disadvantage for either user group in the second task, since both user groups would be writing a message on their less-preferred device for the second half of the experiment, so there the difference in an individual making more mistakes or slowing down as a result of tiredness will be cancelled out. The tasks were observed via a head-mounted camera and an eye-tracking camera attached to a baseball cap.

3.1.2 Data Analysis

The data collected from the EMR-9 eye tracking equipment was converted into movie files and analyzed in ELAN linguistic annotation software. In this software, I made annotations for every hand-movement and eye-movement. My hypothesis was that the visual layout of the iPhone would be confusing for the Japanese user,

Figure 7*: EMR-9 headset and monitor*

so that I might see excessive searching on the interface with the eye instead of looking at the message being written.

The keyboard equivalent would be searching for the right letter on the keyboard with your face down instead of looking at the screen and typing without looking.
What I saw was unexpected. When the subjects used their personal iPhone, they didn’t look at the buttons very often, but they didn’t look at the buttons on their keitai either. However, the eyes spent more time looking for buttons on the iPhone than on the keitai. Most of the time, the eyes were looking at the message, or at the suggested kanji (Chinese character that the operating system offers choice of based on your kana spelling. It also offers choices for completing entire phrases based on previously entered text).

Figure 8: *Subject writing a message on a iida PLY “galakei” type phone.*

Figure 9: *She writes out "だいが", she is in the middle of writing "だいがくせい"(college student)*

Figure 10: *Her eye settles on the suggestion of kanji in the IME for college student: 大学*

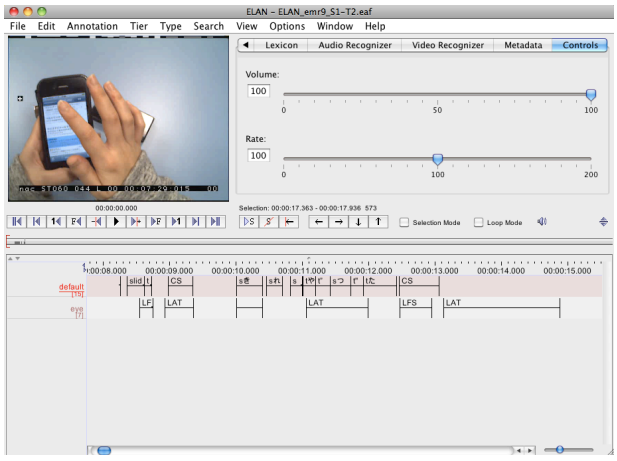

Figure 11: *Annotations for physical actions and eye placement in separate tiers of annotation using ELAN*

I annotated portions of each of the subjects using ELAN linguistic annotation software. The result is that one can map out exactly how many actions the user takes to write one character or one word. The skilled iPhone users spent less actions writing their messages, but the speed was nearly equivalent.

3.1.3 Insights

The result appears to be that among this group of users, all of whom were familiar with smart-phones, is that entering text on a smart-phone is at least as fast as a keitai. The experiment can measure how the users physically enter buttons and where their eye looks, but doesn't give much indication into the internal aspects of the user experience. Whether or not the iPhone interface takes more cognitive power to operate as a result of the mismatch of the way Japanese is taught remains unanswered. Further experimentation along this line of questioning should use a different experiment design, perhaps comparing novice users with expert users across devices.

3.2 Summary of Research Insights

Japanese people learn and memorize the kana using the 五十音 ("gojuon". The characters mean "five", "ten" and "sound", and means "fifty sounds"). In the history of mobile devices, the interface has always mapped the 10 columns of kana with the 10 buttons, and the recent iPhone Japanese interface is no exception. The problem we researched arose from this issue; why is it that many users still prefer to write messages on keitai? We developed a hypothesis that it is because the iPhone uses a system of arranging the kana in an order that does not reflect the 5x10 kana grid, while repeatedly pressing a button on the keitai more accurately resembles selecting a row on the grid. This makes sense theoretically, but our young iPhone users in our eye-tracking experiment showed that they are quite nimble and skilled on their touch-screen interfaces. The reason for this is that while the iPhone interface departs from the way Japanese is understood, it capitalizes on the ease of use of sliding the finger in one of four directions vs. tapping a button repeatedly on a keitai.

The resulting insight is that an ideal mobile interface for writing Japanese has yet to be designed. While cognitively, something resembling a 5x10 kana grid makes the most sense, it does not have to do so visually. The iPhone suffers in it's arbitrary visual arrangement, but capitalizes upon the usability of the touch-gestures via Fitt's Law.

The keitai suffers in that a button must be physically pressed several times in order to arrive upon the desired button, and if one misses the button, they must cycle through all the kana over again. For instance, if you want to type "ke", you hit "2" four times, but if you hit it five times on accident, landing on "ko", you have to hit it another four times, for a total of nine button presses. This is a clear disadvantage in usability from the iPhone interface, where a user simply has to press delete

once and try the gesture again, for a total of three actions: the first try (makes a mistake), delete, try again. Comparing the two interfaces one could say: one more closely mimics the mental model of Japanese while lacking in usability (keitai), and the other departs from the way Japanese is taught but is more usable (iPhone and most smartphone UI's).

4. DESIGN CONCEPT

The ideal human-computer interface for writing Japanese on a small, portable, hand-sized device should have two main considerations in order to be successful: 1) appropriate facilitation of the mechanics of the hand and 2) Accurately reflects the Japanese language in the way it is taught, written, and conceived of in the Japanese user's mind, using あかさたな and gojuon as a model.

The ideal interface for writing in Japanese should also involve no necessary eye-gaze towards the buttons the fingers are using for writing, meaning they should be recognizable by touch or placeholders for each finger, and a blind person could operate it equally well. It should be able to facilitate all fingers. There should be a viewer to view the message being written and the choices of kanji and sentence completion offered by the OS. This could be separate from the input interface. The method of writing in Japanese should facilitate the ergonomic workings of the hand ideally. For example, it should not require that only the thumb press buttons as in the keitai or that the hand float above the interface as in the iPhone while a single finger gestures.

4.1 Designing For the Hand

I started with the predisposition that a completely relaxed hand position is the best way to begin interacting with an object from the standpoint of design. One would not enjoy an object that is designed to be used beginning from an unnatural physical posture. This is what a hand looks like when it is completely relaxed. In order to see this on your own, individual hand, close your hand tightly into a fist and then relax and allow it to expand. The resulting form is your most natural, most relaxed hand position.
Matching Japanese with the Hand

For inputting Japanese, I determined that the hand should be facilitated more-so than it is currently on smartphones. On glass interfaces like the iPhone, a user can only use a few fingers at a time. Generally they will use only one or two fingers, as shown in my experiment data. The ideal interface for writing should use all fingers. The result is that I mapped each finger to a column of the gojuon. The gojuon is convenient because it has 10 columns and 5 rows. We have five fingers on each hand, so one can easily think of each of the fingers as representing the kana.

This is the system by which each finger will enter in Japanese kana. At this time the system is designed to be one-handed. Each finger is represented thusly: thumb = "p", index = "i", middle = "m", ring = "a", pinky = "c".

This system appears to be complex, but from the perspective of the hand it is quite easy. The columns あかさたな are each finger of the hand, and then the rows are chosen by pressing the finger corresponding 1-5. So if I write "か", I press "i" which represents 2, and "p", which represents 1. Then for the second group, はまやらわ, you have to press combinations of two fingers to select column, and then a single finger for row. So め is selected by pressing "i" index and "m" middle together, making 7, the ま column, followed by "a", 4, selecting the fourth row, め. The end-consonant ん is an exception, [pia] simultaneously because it would be the next combination in sequence if there were three simultaneous button presses, and because it is easier to do physically than [pmc] except for pianists, perhaps.

は:[pi]p	な:cp	た:ap	さ:mp	か:ip	あ:pp
ひ:[pi]i	に:ci	ち:ai	し:mi	き:ii	い:pi
ふ:[pi]m	ぬ:cm	つ:am	す:mm	く:im	う:pm
へ:[pi]a	ね:ca	て:aa	せ:ma	け:ia	え:pa
ほ:[pi]c	の:cc	と:ac	そ:mc	こ:ic	お:pc
	ん:[pia]				

わ:[ic]p	ら:[ac]p	や:[ma]p	ま:[im]p
	り:[ac]i		み:[im]i
	る:[ac]m	ゆ:[ma]m	む:[im]m
	れ:[ac]a		め:[im]a
を:[ic]c	ろ:[ac]c	よ:[ma]c	も:[im]c
濁点:[pim] 半濁点:[mac] 。:[imac]			

Figure 12: *The chart of mapping the Gojuuon onto the hand, broken into two parts for formatting.*

To switch between kana entry and kanji selection via IME (input method editor), the user can press

all fingers at once. We also included three-finger combinations for the dakuten and maru, which are marks that are added to kana. Spaces are not generally used in Japanese, and a period is simply four fingers without the thumb. The IME can be responsible for selecting more complex punctuation marks (Unicode), as it is currently done in interactive mobile systems.

This system design shows an approach towards designing for the Japanese user by focusing on mapping the hand and the language together. In this design, there is no separation between the way Japanese is thought of and the way it is written on a mobile interface.

5. CONCLUSION

No matter the design, in the end the responsibility will fall on the user to learn the new interface. This is the case with language in general, as humans have to learn how to speak, read, and write in complex systems of language, and not all humans are able to do even that. This project sought to understand how Japanese people think about language, and how that affects their experience using mobile phones, the most popular computing device in Japan. This paper presented the historical and linguistic dimension, the real-world situational dimension, and the micro-analytical usability dimension of this research project. We analyzed the results and produced a design concept and prototype that exemplifies the issues in consideration when designing for Japanese language users. Each of the dimensions covered in this paper could represent a whole body of work to be pursued that would benefit the knowledge of this subject and the efficacy of mobile user interfaces for Japanese.

As it stands, mobile UI's for writing Japanese are emergent and very clumsy. Between writing numbers in a pager to be translated in the user's head as letters, and the ultimate goal of writing as fast as the user can think, we are much closer to the pager. There is much work to be done in understanding how people think about language, and how a human can transmit that understanding through their body into a device. This paper represents an effort in the direction of making a more ideal future for Japanese mobile UI, through research and design.

6. REFERENCES

BANKS, M. 2001. *Visual Methods in Social Research*. Sage Publications Ltd, London.

GERNSBACHER, M.A., AND KASCHAK, M.P. 2003. *Language Comprehension*. Wiley Online Library.

HAMZAH., TANO., IWATA., HASHIMAYA. 2006, Effectiveness of Annotating by Hand for non-Alphabetical Languages. In *CHI2006 Proceedings*. Quebec, Canada. 22-27 April 2006.

HEATH, C., HINDMARSH, J., LUFF, P. 2010 *Video in Qualitative Research: Analysing Social Interaction in Everyday Life*. Sage, London.

IKEDA, NUMATA, KANEKO, MACHIDA. Proposal for a Multilanguage Text Input Support System that is Easy for Beginner Language Learners. In *Proceedings of the 3rd International Universal Comunication Symposium*. 109-114.

ITO, M., OKABE, D., AND MATSUDA, M. 2005 *Personal, Portable, Pedestrian: Mobile Phones in Japanese Life*. The MIT Press, Boston.

JOKINEN., NISHIDA., YAMAMOTO. 2009. Eye-gaze Experiments for Conversation monitoring. *Proceedings of the 3rd International Universal Communication Symposium.* 303-308.

SHRUM, W., DUQUE, R., AND BROWN, T. 2005 Digital Video as Research Practice: Methodology for the Millennium. *Journal of Research Practice 1*(1): Article M4, pp. 1-19.