An On-Line Japanese Handwriting Recognition System integrated into an E-Learning Environment for Kanji

Steven B. Poggel
steven.poggel@gmail.com

Monday 22nd February, 2010

Contents

Chapter 1

E-Learning

1.1 Introduction to E-Learning

The term *e-learning* refers to a number of different methods, concepts and techniques. It is therefore difficult to confine the term sharply. Thus, in literature, there are different definitions of what e-learning is and what it is supposed to be. Rosenberg (2006) defines e-learning as follows:

> **E-learning** is the use of Internet technologies to create and deliver a rich learning environment that includes a broad array of instruction and information resources and solutions, the goal of which is to enhance individual and organisational performance.

It can be noted that Rosenberg (2006) defines e-learning purely by terms of instruction and information resources. Further, *the use of Internet technologies* is seen as a necessary condition for e-learning. The definition does not take into account educational software. Seel and Ifenthaler (2009) claim that the terms *e-learning* and *learning on-line* are synonymous.

Richert (2007) criticises the definition of Rosenberg because she sees no reason for such equality of terms. She constitutes her view with the fact that electronic (learning) applications are not limited to the Internet. Richert (2007) defines e-learning as:

> Unter E-learning wird das computergestützte Lernen (vorwiegend von Einzelpersonen) mit hypertextbasierten, multimedialen, interaktiven Systemen verstanden, das zeit- und ortsunabhängig sowohl online als auch offline erfolgen kann.

in English:

> E-learning is defined as computer-aided learning (mainly by individuals) with hypertext- and multimedia-based interactive systems. The learning process can take place independent of time and location both on-line and offline.

It is important to note that the term is broader than the definition of Rosenberg, but is restricted to *learning systems*. That means concretely that electronic media like dictionaries may be included in e-learning systems as a tool, however, they can only form a part of a more general e-learning environment. Electronic media itself is not necessarily understood as e-learning system.

1.2 Classification of E-Learning Systems

E-learning systems can be classified by their their degree of freedom for user interaction. On one end of the scale there are *Drill-and-Practise* programs that do not allow for freedom of interaction. On the other end there are interactive programs allowing the user to interact and control the application. Judged by the definition of Richert this classification does not seem very suitable (Richert 2007).

Another possibility to classify e-learning systems is the the kind of storage media used. This classification allows for a distinction between *on-line* and *offline* e-learning systems. *Offline systems* are those systems that are offered on passive storage media like floppy disk, CD-ROM. Offline systems are usually called *Computer Based Training* (CBT) systems. *On-line systems* on the other hand are web server based systems that fall under the category of *Web Based Training* (WBT) systems (Richert 2007).

Additionally, Richert (2007) defines *hybrid systems* that are CBT systems but use the Internet as a means of communication with other learners. Table 1.1 shows the classification of e-learning systems after (Richert 2007).

		Using the WWW for storage	
		Yes	No
Using the WWW for communication	No	WBT	CBT
	Yes	Learning platforms	Hybrid CBT

Table 1.1: Classification of e-learning systems

1.3 Technical Context of E-Learning

1.3.1 Multimedia Systems

The term *Multimedia* has several definitions. Simple versions of multimedia definitions state that multimedia refers to a combination of different forms of information from several sources. Those forms can contain textual information, graphic, video and audio. With a broad definition of that kind any television news report could be regarded as multimedia. Richert (2007) understands *multimedia* more holistically than that. She sees multimedia as a technological concept that allows for the interaction of a user and a multiple media system. More than one sensorial modality should be should be presented by the system.

1.3.2 Classification of Interactivity

Interactivity can be defined in several steps. The concept of *interaction* serves as a basis for the classification, because in a sociological sense there can, by definition, be no mutual interference between man and machine. Interactivity in the sense of interaction comprises the ability to access and control different functionalities of a software system (Richert 2007).

Six classes of interactivity can be described. They differ by their degree of interaction between the user and a software system. The gamut of interactivity is used to evaluate e-learning applications:

1. **View and absorb objects**
   The multimedia components can be viewed and played by the user. The user can not further influence the components in any way.

2. **View and absorb multiple displays**
   Program components offer more than one display. For instance, a user could click on a picture and be shown a different one. No modification of components is possible.

3. **Varying the form of representation**
   On this level, users can gain the feeling they could actively influence the multimedia components. They can scale objects or view them from different perspectives. Users can influence the form of representation but not the content.

4. **Changing the content of a component - parameter or data variation**
   Contents of a multimedia component are generated by the user. Users can input data or text. They can not change films or pictures. A usage example of that type could be the selection methods of statistics programs. Users can modify objects and the program yields different results.

5. **Generating objects or the content of a representation**
   This mode of interaction is reached by applications that offer tools to create and change content. For example visualise thoughts with mind maps, or render new forms and models.

6. **Constructive and manipulative actions through situation-dependent feedback**
   On this level of interaction symbols can be manipulated and the result of the interpretation can be interpreted by the program. That allows for the generation of useful and context-sensitive feedback. User input can be evaluated by the application.

The gamut is described after (Richert 2007).

1.4 Pedagogical Context of E-Learning

The pedagogical context of e-learning is a crucial part of any e-learning environment. The learning targets need to be defined and a conceptual design of a software needs to be based on those.

1.4.1 Learning

The term *learning* is of a complex nature. A definition of learning is therefore never sharply confined. The definition of *learning* by Lefrancois (1994) shows how broad the term can be perceived:

> *Lernen umfasst alle Verhaltensänderungen, die aufgrund von Erfahrungen zustandekommen.*

In English:

> Learning compasses all changes in behaviour that are based on experience.

The changes in behaviour include those processes that do not aim at acquiring information, but also those changes in behaviour of an unknown cause (Lefrancois 1994). According to (Richert 2007), this means the acquisition of competences of different kinds.

1.4.1.1 Educational Objectives

Cognitive learning targets comprise all targets that include acquisition of knowledge. Knowledge can refer to both reproduction of content, but also acquiring the ability to solve problems. The area of cognitive learning targets can be differentiated further. Richert (2007) distinguishes the acquisition of:

- **Declarative knowledge** or **factual knowledge**. Knowledge that can be categorised as *knowing that* as opposed to *knowing how*.
- **Procedural knowledge** or **dynamic knowledge**. Knowledge that contains approaches to problems and their resolution procedures. Procedural knowledge can be seen as a series of declarative inventory of knowledge, nevertheless it can be categorised as *knowing how*. The distinction to the regular ability of a human to solve problems lies in the fact that declarative knowledge is needed in order to solve specific types of problems. For instance, in order to be able to successfully use a map for navigating, a human needs to know that a certain object is a map, what the symbols on the map mean and where or what the four cardinal points are.
- **Contextual knowledge**. Knowledge that contains application situations. This category is centred around *when and where* to apply knowledge. What abilities and which factual knowledge can be used in which situations?

Affective learning targets are educational objectives that aim at changing behaviour. It is difficult to actualise affective learning targets by cognitive learning only. In order to achieve affective learning targets, feelings, evaluations and attitudes of humans need to be taken into account. In learning situations at school often personal enthusiasm, credibility and charisma of the teacher play a role (Richert 2007).

Psychomotor learning targets is the class of learning targets that aim at the acquisition of manual abilities and motion sequences. That includes playing of musical instruments or using tools. Analogue to the procedural knowledge learning, theoretical knowledge about the objects involved is necessary in order to achieve the psychomotor abilities. That theoretical knowledge is a necessary condition for the psychomotor learning process, but it is not sufficient, psychomotor learning involves practising motor sequences (Richert 2007).

1.4.1.2 Self-Driven Learning

Learning is often seen as a behaviourist stimulus-reaction process. Different views observe an active and constructive process. In the constructionist view on learning there is a continuum from *self-learning* to *autonomous learning*. *Self-driven learning* can be classified by that continuum.

Self-learning defines a type of learning with a focus on the self-initiative and self-responsibility of a learner. Richert (2007) reports of the opinion that it is impossible to not self-learn, because learning always assumes the intention of the learner. This view conjectures that each learner has to accomplish the task of self-construction of knowledge. However, self-learning defines solely the self-initiative of the learner, the learning material is provided by an external source.

Autonomous learning is distinct from self-learning in the way, that it focuses on teacher-independent organisation of learning. While in self-learning, the learner can decide to learn, independently of a teacher, in autonomous learning the learner is self-responsible for defining learning targets and has an analytical view on the learning process (Richert 2007).

Self-driven learning is a type of learning that can be seen somewhere on the continuum from self-learning to autonomous learning. In self-driven learning the learner is given all instructions and decisions concerning the learning process in the learning materials. The self-direction is therefore limited to the location and the time of learning (Richert 2007).

1.4.2 Intelligent Tutorial Systems

Intelligent tutorial systems fall under the paradigm of cognitive learning. The general scheme of such a system is depicted after (Richert 2007) in figure 1.1. The declarative knowledge of a system is stored in the expert module. The student module holds information about the learning progress and the course module holds the lessons of the application. The communication module interacts with the learner. A problem of these systems is that they cannot distinguish between small oversights and serious

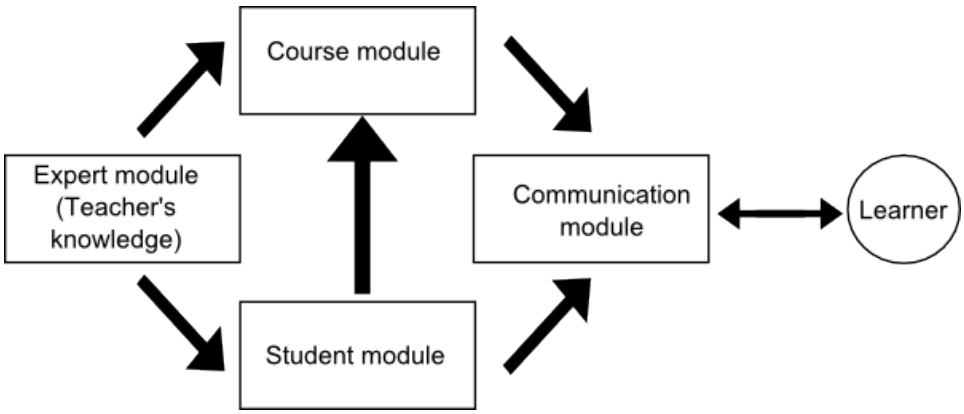

Figure 1.1: General scheme of intelligent tutorial systems

1.5 E-Learning of Languages

In *computer assisted language learning* (CALL) most the systems are not necessarily bound to the key competences for language learning. Richert (2007) defines the following key competences as possible learning foci for the e-learning of languages:

- linguistic (grammatical) competence
- socio-linguistic competence
- discourse competence
- strategic competence
- social competence
- sociocultural competence

E-learning systems can offer specific parts of the language as a learning focus. The learner decides what to learn when. Available systems, however, do not always account for this educational ideal (Richert 2007). Bailey and Meurers (2008) report about an intelligent CALL system that aims at diagnosing meaning errors. Generally, most CALL systems are based on structural analysis of language. Nevertheless, the systems often focus on fixed patterns and rely on repetition. The drill of a certain pattern is meant to automate the reproduction, however, allowing no room for creativity. The problem is that the repetition of stereotype structures is that they are just a form without a function (Richert 2007). Some e-learning systems add little or no value to the learning process, because they focused solely on the management and measurement of the training process(Ismail 2002). However, due to the impersonal nature of e-learning systems compared to face-to-face classroom work, the pedagogical acts can be objectified, free from personal animosities that might occur between a teacher and a learner (Zimmer 2009).

1.5.1 E-Learning of Japanese

1.5.1.1 Japanese CALL Systems

Some efforts for e-learning of Japanese have been made. Nagata (2002) report about the BANZAI system, that is a classical CALL application. It uses syntactical parsing in order to teach grammatical constructs. It does not aim at teaching Japanese characters, but focuses on the structure of the language. There are systems that focus on the e-learning of Japanese Kanji. These systems, by their very nature, do not comply with the key competences of language learning defined by (Richert 2007). The definition of key competences focuses mainly on spoken language, it does not account for the difficulties a learner may have in writing a language. Many Japanese e-learning systems are in fact not systems aiming at teaching the Japanese language as such, but rather only the Kanji.

1.5.1.2 Key Competences for Studying Japanese Script

The key competences[1] defined by Richert (2007) are orthogonal to those types of e-learning system focusing on the script. In addition to the key competences, other competences are needed for writing any language, Japanese in particular[2]. More concretely, the *linguistic competence*, needs to be extended from a purely *grammatical competence* to a more general linguistic competence, including the fields *orthographic competence* and *graphemic competence*. Both are necessary conditions for a successful written communication in a language. Despite the availablility of modern technologies like video conference software, and online telephony, the most popular form of both immidiate and delayed communication on the internet is still written language. That means, in order for a learner to be able to participate in world-wide communication in a specific language, it is essential to be able to read and write.

[1] Presented in section 1.5

[2] For the complexity of the Japanese writing system, see chapter **??**)

Chapter 2

Conceptual Design of Kanji-Coach

2.1 Requirements of a Kanji Teaching E-Learning Application

2.1.1 General Considerations

In order to create a concept for a Kanji teaching application, a number of different aspects need to be taken into consideration. These aspects emerge from the academic background concerning the Japanese script, pedagocical and didactic knowledge about teaching languages and general conceptions of e-learning applications.

Many efforts in designing e-learning applications are focused around the teacher's view on learning. For designing an e-learning application that is useful to students, the students view needs be taken into account (Alexander and Golja 2007). Ivašin (2009) critisises the technical dominance in e-learning and e-teaching processes, as the conceptual software designs are not always supporting the didactic purpose of the software. Therefore, the user view should be taken into account when conceptually designing an e-learning application. The requirement of a *user-focused design* follows directly from this view.

For online e-learning, it is a known that readers only scan the textual information displayed. Therefore it is not useful to provide a user with large blocks of text, but rather with smaller chunks that encourage skimming over (Hamid 2001). It can be expected that the fact that an e-learning process happens online does not greatly affect the user behaviour. Therefore, the observations made for online e-learning can probably be applied to offline desktop application based e-learning. The requirement of *keeping textual information short and concise* derives from the observation.

If e-learning is considered as a learning method in higher education, *blended learning* seems to be the most suitable form of e-learning. That means, combining classroom activities with e-learning methods (Hettinger 2008; Kahiigi et al. 2008). Language learning is not necessarily considered as higher education. In case of studying Japanese, with its specific difficulties in language and the script, language learning is taken to an intellectual level that is at least close to higher education. Therefore, an e-learning application for any aspect of the Japanese language should not have the pretensions of posessing the capability to *teach Japanese*. Japanese is a complex language with a complex script. Therefore, e-learning applications should aim at supporting a learner's classroom efforts of studying the language. The requirement of *focussing on a specific language aspect* can be drawn from this reasoning. The prototype system designed in this work does not aim at being a complete system, but rather offers individual learning components from which a user can choose what type of learning and which component best supports his study.

2.1.2 Classification of a Kanji Teaching Application

In section 1.2 different types of e-learning systems have been discussed. In the course of designing a prototype system, design choices need to be made. The design choice for the prototype will be an offline e-learning system, that runs on a desktop PC. This design choice does not follow a conceptual requirement, in fact it ignores Ivašin's (2009) criticism of technical dominance in e-learning systems. The choice is a purely technical choice, yet, it is driven by a conceptual requirement. The purpose of the e-learning environment is to test to what extend a handwriting recognition can help study the Kanji. In order to study that research question, the handwriting recognition needs to be implemented and integrated with the e-learning system. Thus, the design choice for an offline system was inevitable in the sense that the technical limitations of on-line applications form an obstacle for pen input of characters and fast recognition procedures.

In the definition given by (Richert 2007) the prototype is a *computer based training* (CBT) system, as it does not use the Internet for communication or a webserver for storage. Another criterion for an offline systems is that they are offered on CD-ROM or floppy disk. That criterion can be regarded completly obsolete, as it refers to specific storage media. Even if a higher level of abstraction is used to describe the criterion, it is

still obsolete, since *a passive storage medium* is not necessary to describe what the criterion actually tries to define. The criteria concerning communication and data storage are useful to confine different types of e-learning applications. Additionally, *installability* can be used as a criterion for offline e-learning systems. *Installability* here refers to *the possibility to install a software on a computer system*, not the *ease of installation*, which is defined in ISO9126 as *installability* as well (Chua and Dyson 2004). The ISO9126 type of installability will be taken into account during the software evaluation, which is reported in chapter **??**.

Concerning the level of interactivity described in section 1.3.2 the prototype designed in this work is aimed at a level higher than level *(4) Changing the content of a component*. It is targeted between the levels *(5) Generating objects or the content of a representation* and *(6) Constructive and manipulative actions through situation-dependent feedback*. Concretely, a user can:

- Change the ideal shape of a character by storing a new gold standard.
- Create new characters and their descriptions
- Receive situation-dependent feedback even on the newly created characters, due to the nature of the error recognition algorithm that evaluates mathematically the distance between a gold standard character and a an input. Additionally, characters are analysed structurally, therefore new characters added by the user will automatically be classified and arranged among the other characters in the database of the system.

Thus, based on the levels of interactivity (Richert 2007), it can be concluded that the prototype provides a very high level of interaction. The levels serve as an evaluation measurement for the quality of e-learning applications.

2.1.3 Conceptual Issues for E-Learning of Kanji

(Stahlmann 2004) spezielle aspekte bezueglich han-trainer pro

2.2 Approaching the Specific Difficulties of the Japanese Script

2.2.1 Character Learning Aspects

- greife typische probleme der lerner auf (siehe japanischkapite **??**) s. 11 hinten kurze auflistung. geschichten?

2.2.1.1 Character Repetition

In section 1.5 the pure repetition of grammatical structures as a learning method has been critisised. The system should account for that by not just forcing the user to reproduce fixed structures. In fact, it should leave room for creativity. Zum Beispiel - radikale vorgeben und zeichen schreiben lassen. und ganz generell: toleranzgrenzen erlauben kreativitaet allein schon deswegen, weil selbst der zeichenstift benutzt wird.

2.3 Integration of HWR Into the Learning Process

Wichtige Fragestellung: Wie sieht eine HWR in Lernumgebung aus? Mappe S. 19 genaue anforderungen s. 19! waehle bestimmte architekture unter moeglichen ansaetzen.

welche art von character recognition muss geleistet werden?

was sind die moeglichkeiten (im vergleich zu anderen produkten), die sich durch eine HWR ergeben? wie kann man die ausschoepfen? s. 16 unten und s. 15

Error Recognition what type of errors? semantical errors? cow vs sheep vs pig phonological errors (readings) kanji that sound the same. theoretically: compounds - for the kanji readings. heft: s. 52

- compare with normal paper-based learning of kanji - compare with other kanji-learning systems klare abgrenzung von skritter. s. 51 unten im heft.

2.4 Handling Errors

2.4.1 Motivation for Error Recognition

2.4.2 Possible Sources of Error When Writing Japanese Characters

error handling, see page 58.

See notes on paper, seite 58 - for example stroke number and stroke sequence - length of strokes

- stroke velocity

2.5 Use Cases

siehe 'screenshot' - grafiken von s. 2 - 9 auch: was kann man aus e-learning machen? welche (technischen) moeglichkeiten sind eroeffnet, insbesondere auch durch handschriftenerkennung?

idee: schoenschreibekurs, bei dem einzelne striche gesondert geuebt werden.

List of Figures

Listings

List of Tables

References

Alexander, S. and T. Golja (2007). Using Students' Experiences to Derive Quality in an E-Learning System: An Institution's Perspective. *Educational Technology and Society 10*(2), 17--33.

Bailey, S. and D. Meurers (2008). Diagnosing Meaning Errors in Short Answers to Reading Comprehension Questions. In *EANL '08: Proceedings of the Third Workshop on Innovative Use of NLP for Building Educational Applications*, Morristown, NJ, USA, pp. 107--115. Association for Computational Linguistics.

Chua, B. B. and L. E. Dyson (2004, Dec). Applying the ISO 9126 Model to the Evaluation of an E-Learning System. In *Proc. of ASCILITE 2004*, Perth, Australia, pp. 184--190.

Hamid, A. A. (2001). e-Learning: Is it the "e" or the Learning That Matters? *The Internet and Higher Education 4*(3-4), 311--316.

Hettinger, J. (2008). *E-Learning in der Schule (in German)*. Munich, Germany: kopaed Verlag.

Ismail, J. (2002). The Design of an E-Learning System: Beyond the Hype. *The Internet and Higher Education 4*(3-4), 329--336.

Ivašin, M. (2009). Lernen und Technologie (in German). In B. Mikuszeit and U. Szudra (Eds.), *Multimedia und ethische Bildung*, Chapter 5.5, pp. 635--648. Frankfurt am Main, Germany: Peter Lang.

Kahiigi, E. K., L. Ekenberg, H. Hansson, F. Tusubira, and M. Danielson (2008). Exploring the E-Learning State of Art. *The Electronic Journal of E-Learning 6*(2), 77--88.

Lefrancois, G. R. (1994). *Psychologie des Lernens (in German)*. Heidelberg, Germany: Springer.

Nagata, N. (2002). BANZAI: An Application of Natural Language Processing to Web-Based Language Learning. *CALICO Journal 19*(3), 583--599.

Richert, A. S. (2007, Dec). *Einfluss von Lernbiografien und subjektiven Theorien auf selbst gesteuertes Einzellernen mittels E-Learning am Beispiel Fremdsprachenlernen (in German)*. Ph. D. thesis, RWTH Aachen, Aachen. Manuscript committee: Prof. Dr. phil. Rudolf Beier, Prof. Dr. phil. Uwe Michelsen.

Rosenberg, M. J. (2006). *Beyond E-Learning. Approaches and Technologies to Enhance Organizational Knowledge, Learning and Performance*. San Francisco, USA: Pfeiffer.

Seel, N. M. and D. Ifenthaler (2009). *Online lernen und lehren (in German)*. Munich, Germany: Reinhardt UTB.

Stahlmann, R. (2004). Didaktische, inhaltliche und funktionelle Optimierung einer selbst entwickelten Chinesischlernsoftware (in German). Master's thesis, Offenburg University of Applied Sciences, Offenburg, Germany. Manuscript committee: Prof. Dr. Roland Riempp (supervisor), Prof. Dr. Thomas Breyer-Mayländer.

Zimmer, G. (2009). Bildung mit E-Learning (in German). In B. Mikuszeit and U. Szudra (Eds.), *Multimedia und ethische Bildung*, Chapter 1.3, pp. 61--92. Frankfurt am Main, Germany: Peter Lang.

Document created on Monday 22nd February, 2010 at 15:38