

# Japanese HWR

Steven B. Poggel  
steven.poggel@gmail.com

February 15, 2010



# Contents

Summary	5
Zusammenfassung	7



# Summary

In this work, a Japanese handwriting recognition system is being developed. The system is integrated into an e-learning environment in order to provide a Kanji teaching application with automated error correction. Conceptually, the application is an e-learning environment for Japanese characters, intended for the foreign learner of the Japanese language. Most e-learning systems of Japanese Kanji provide only a multiple choice method for the learner to reproduce characters. The present prototype offers the ability to enter characters with a stylus on a touch screen system.

The study seeks to determine to what extent it is possible to use modern NLP methods for language learning. While other studies mainly focus on grammatical correction, this application is targeted on the Kanji characters. It will be examined if a handwriting recognition engine can generate informed feedback, suitable for a learner. Additionally, the study examines if that feedback helps obtaining the ability to actively reproduce the Kanji characters.

The prototype developed in this work combines e-learning methods with natural language processing applied to the Japanese script. In order to prepare the task of creating a software that spans across the aforementioned fields of study, the work reviews the structure of the Japanese script, the current state of the art in handwriting recognition methods and e-learning techniques. The recognition engine implements a structural approach to Kanji character identification. The recognition performs partial analysis of substructures and binds the recognised elements together to form a character. Because of the structural approach it becomes possible to create an informed error recognition that considers linguistic units of the Kanji characters.



# Zusammenfassung

In dieser Arbeit wird eine Handschriftenerkennung für japanische Kanji entwickelt. Der Handschriftenerkennung ist in eine E-Learning-Umgebung integriert und liefert eine automatisch generierte Fehlerkorrektur für Lernende.

Das System ist konzeptuell eine E-Learning-Anwendung für das Erlernen der japanischen Schrift, die aufgrund ihrer morphemischen Struktur einen hohen Komplexitätsgrad aufweist und daher besonderen Lernaufwand benötigt. Die meisten E-Learning-Systeme für asiatische Schriftzeichen bieten Zeichenabfrage als Multiple-Choice an, da die Eingabe der Zeichen für einen Lernenden sonst kaum zu bewerkstelligen wäre. Der in dieser Arbeit erstellte Prototyp bietet die Möglichkeit zur handschriftlichen Eingabe von Zeichen auf einer dafür geeigneten Bildschirmoberfläche.

Die Studie untersucht, inwieweit es im Bereich des Schrifterwerbs möglich ist, NLP und Lernmethoden zusammenzubringen. Dabei wird nicht mit Parsing-Methoden die grammatische Struktur der Sprache untersucht, sondern vielmehr die interne Struktur der Kanji zugrunde gelegt und durch einen Handschriftenerkennung erfasst. Dabei sollen Schreibfehler strukturell erkannt werden. Intelligentes Feedback soll dem Lernenden dabei helfen, die Fähigkeit der aktiven Reproduktion der Kanji zu erwerben.

Da der Prototyp eine disziplinübergreifende Software ist, die in den Bereichen Handschriftenerkennung und E-Learning angesiedelt ist, wird in der vorliegenden Arbeit der Forschungsstand der beiden untersucht. Weiterhin wird die Struktur der japanischen Schrift dargestellt. Die Kombination der drei Bereiche führt dazu, dass die Substrukturen der Kanji überhaupt programmatisch analysiert werden können, wodurch die Fehlererkennung ermöglicht wird.