

古代エジプト神官文字写本を対象とした 言語情報表示システムの試作

永井 正勝 和氣 愛仁

筑波大学 人文社会系

古代エジプト文明で使用されていた言語である古代エジプト語は、ある程度内容の理解はできるものの、言語研究の立場から見れば、文法の解明が十全に達成されていない言語である。加えて、パピルス写本に記される神官文字に至っては、今から 100 年前に出版された研究結果が現在でも基本書となるほどに、研究が停滞した状況となっている。このような課題を少しでも是正すべく、本研究では神官文字の画像を利用した公開型の WEB 連携データベースシステムの作成を試みた。システム設計上の目標として、①ヴィジュアルなシステムの実現、②適切かつ拡張可能なデータ構造の採用、③研究者同士のコミュニケーション活動の促進、④既存のシステムを利用した高速開発の 4 点を設定した。このシステムにより、資料上の文字や語の解釈に揺れが生じうる複雑な状況に柔軟に対応しつつ、一層の研究の進展を望むことができる。

Development of a Web-based Linguistic Information System for Ancient Egyptian Hieratic Manuscripts

Masakatsu Nagai, Toshihito Waki

Faculty of Humanities and Social Sciences
University of Tsukuba

The language system of Ancient Egyptian has not been completely understood yet. Especially, the study of writing system known as hieratic remains undeveloped. To improve this situation, this study aims to construct a web-based linguistic information system for Ancient Egyptian hieratic manuscripts. This system flexibly supports the complicated such situations as the interpretations of the signs, words, or grammar of a text.

1. はじめに

永井は言語学の立場から古代エジプトの神官文字資料を研究すべく、神官文字の字形データベースを作成してきた[1]。このデータベースは、文字や単語で検索するとその結果として言語情報のアノテーションと文字の画像が示されるという単純なものである。これは、Microsoft Access と FileMaker Pro を使用して作成したスタンドアロン形式のもので、その利用も作成者自身の個人的な研究に留まっていた。永井はこのシステムの限界を意識し、今まで作成してきた資料を生かしながら、WEB 公開型のデータベースを作成したいと考えていた。

そのような折、2012 年 8 月に和氣の賛同・協力を得て、WEB 公開型のデータベースを作成する運びとなった。未だ試作版ではあるが、本稿ではこのデータベースの概要を紹介する。

2. エジプト語研究の問題点

本研究の人文科学的意義は、古代エジプト語の研究現状から導かれるものである。そこで最初にこの点について述べる。

2.1. 古代エジプト語研究の問題点

古代エジプト語はエジプト文明で使用されていた言語であり、その言語的特徴と時代の違いから古エジプト語 (紀元前 3000-2000 年頃)、中エジプト語 (紀元前 2000-1300 年頃)、新エジプト語 (紀元前 1300-700 年頃)、民衆エジプト語 (紀元前 700 年頃-紀元後 500 年頃)、コプト・エジプト語 (紀元後 4-14 世紀頃) に区分される[2]。

これらのうち、コプト・エジプト語を除く段階では、原則として母音を表記しないシステムが採用されているため、書記言語から言語構造を抽出することが難しく、文法理解において学者間で見解の一致を見ていない部分が少なから

ず存在する。つまり、古エジプト語から新エジプト語までの段階については、文法解明が十全に達成されていない言語変種となっている。

2.2. 古代エジプト語の文字種とそれに対する学者の態度

文法が十全に解明されていない文献資料を扱う場合には、まずもって文字の判読が重要となるが、古代エジプト語の研究では、そもそも神官文字の判読において、学者間で解釈に揺れが生じている。

古エジプト語から新エジプト語の表記において、聖刻文字 (ヒエログリフ) と神官文字 (ヒエラティック) が使用されていた [図 1-(a),(b)].

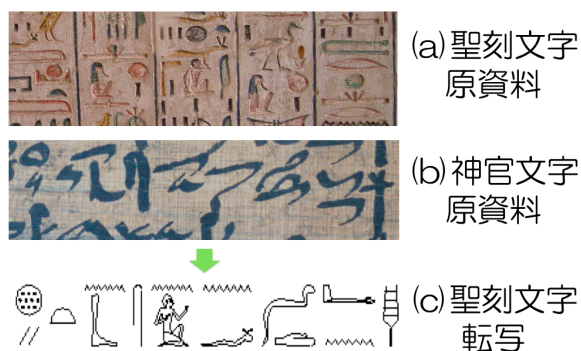


図 1 聖刻文字と神官文字の例

聖刻文字と神官文字はともに古代エジプト人の残した原資料の文字であるが、エジプト語の研究には、もう 1 つ別の種類の聖刻文字が存在する。それは、神官文字資料を学者が聖刻文字に改めた聖刻文字転写 [図 1-(c)] である。

エジプト学の研究では、神官文字資料の写真が公開されずに、神官文字資料の聖刻文字転写が「資料」と称されて刊行されることが少なくない。

2.3. 神官文字研究の現状

上で述べたように、エジプト学では神官文字資料が聖刻文字に直されて公開されるため、神官文字の字形研究が極度に遅れた分野となっている。たとえば、字形判読の基本となる字形リストは、1909-1912 年に出版された[3]が現在でも基本書となっており、その根本的な改訂がなされていない。[3]の書物は 3000 年間に及ぶ神官文字の変遷を明らかにした名著ではあるのだが、以下のような問題点を抱えている。

(1)

a: 文字の登録方法が一貫していないため、登録に過不足が認められる。

b: 掲載されている字形の例が著者自身の手書きであるために実際の写本の字形と異なる部分が見られる。

c: 字形の典型例のみが掲載されているため、個々の字形を十全に把握することができない。

d: 個々の字形がいかなる単語で使用されているのかという情報が記されていない

また、神官文字研究の問題点として、次のことも指摘しておかねばならない。

(2) 神官文字と聖刻文字を対応させる研究が十全に達成されていない。

2.4. 神官文字資料に内在する問題点

[3]に代わるデータを作るのであれば、資料を丹念に読み、字形リストを作成することになるのだが、しかしながら、そのような作業が首尾よく行われるわけではないという点に、神官文字研究の課題がある。手書きで書かれる神官文字の判読には、読み手の経験が求められ、文字の判読という基本的な作業においてすら、学者間で見解が異なることが少なくない。しかも、文字の判読の結果が、ときに形態素・単語・構文の解釈の違いにまで発展することもある[4]。

神官文字の判読に学者間で揺れが生じることの背景には、2.3.で述べたような基礎研究の不足が挙げられるが、加えて、神官文字資料に内在する次のような問題点も大きな要因となっている。

(3)

a: 字形が時代とともに変化する。

b: 書記ごとに字形が異なる。

c: 同じ文字でも、環境(単語内の位置)によって字形が異なる。

d: 文書中に文字の欠損や書き誤りが見られる。

言語そのものが十分に解明されていない現状において、特に (3c)と(3d)の状態は、学者による文字の判読に大きな影響を与えている。

3. 研究の目的

本研究の目的は、神官文字研究の問題点として指摘した(1)-(3)の課題に応えるべく、資料のデータベースを作成することにある。

エジプト語研究の立場から設定される目的は以下の通りである。

①文書中のすべての文字を対象とする (→課題 1c に対応)。

②すべての文字の画像を示す(→課題 1b に対応)。

③個々の文字に発音や文字コードなどの情報を付す。その際、文字コードについては、神官

文字と聖刻文字の文字コードの両方を付ける
(→課題 1a, 2 に対応).

④単語情報として、単語の転写や文法解説などを付す (→文法解明に寄与).

⑤文字情報と単語情報をリンクさせる (→課題 1d, 3c に対応).

⑥文字の欠損や書き誤りに対応する (→課題 3d に対応).

⑦学者間の見解の相違を示す (→文字や文法の未解決問題に対応).

目下のところ、(3a)と(3b)の課題には対応していないが、この点については資料を増加させることで達成されることになる。また、システム構築上の目的については、5で後述する。

4. 分析資料

資料は大英博物館に所蔵されている神官文字パピルス BM10221 (通称 Papyrus Abbott) を用いる。この資料は古代エジプト第 20 王朝時代に書かれた墓泥棒の裁判記録である (紀元前 1100 年頃)。この資料に記されている言語変種は新エジプト語であり、文字はすべて神官文字となる。本資料はエジプト学の研究で有名な歴史資料 (史料) であり、エジプト史の概説書で頻繁に引用される。その出版は[5]で行われているが、ここには神官文字の手書きトレースが掲載されているのみで、写真の添付はない。加えて、手書きトレースの宿命として、原資料の字形と細部において異なる部分があり、そのような違いがときに文字の解釈の違いを生じさせている。このような状況において、本資料の神官文字をデータベースとして公開すれば、言語研究者のみならず、エジプト史に関心のあるすべての人に有意義なデータとなることが期待される。

本データベースで利用する BM10221 のデジタル画像の撮影は、大英博物館の Vivian Davies 博士ならびに Richard Parkinson 博士の協力のもと、2011 年 5 月に永井が実施した。機材には中判カメラ Mamiya AFDIII, デジタルバック Phase One P40, Mamiya Sekor Macro MF 120mm f/4 を用い、それらを Mac Book Pro に連結させて撮影を行った。その結果、約 4000 万画素の高精細な画像を得ることができた。

5. システム概要

以下に本システムの URL を示す。

<https://hdb.jinsha.tsukuba.ac.jp/>

前述の研究上の問題点に対応するため、以下の 4 点をシステム構築上の目標とした。

①ヴィジュアルなシステムの実現

地理的・時間的制約を超えて、一次資料の高精細な画像を直接閲覧できることがまずなによりも重要である。そのためには、可能な限り高い解像度を持つ画像を、ネットワークを通じてストレスなく表示できることが必要である。また、画像から直接文字データ・単語データへアクセスできることが、資料そのものの理解と効率的な情報検索のために極めて重要である。

②適正かつ拡張可能なデータ構造の採用

内部的なデータ構造を、古代エジプト語の文字研究の現状に即したものとすると同時に、将来的な言語学的検証にも耐えうるような形式とする必要がある。

③研究者同士のコミュニケーション活動の促進

各国の研究者がシステムをプラットフォームとして利用することで、古代エジプト語に関する研究をより促進させることを目指す。そのために、訪問した研究者が受動的にデータを受け取るだけでなく、能動的にデータについての見解を述べることができるようにする必要がある。さらに、訪問者自ら直接的にデータを作成することに関与することができるようになればより望ましい。

④既存のシステムを利用した高速開発

どんなに高機能なシステムでも、構築にかかる手間が膨大すぎるとは導入の障壁が高くなり、利用が広がらない。今回構築したシステムは、将来的に他の人文系画像資料に対するアノテーション付与システムのプラットフォームとしても利用することを想定しており、その意味で、できるだけ構築が楽で、他の資料への応用も容易となるようなシステムにする必要がある。

以下、これらの目標を踏まえた上で、本システムの特徴について述べる。

5.1. システム構築環境

本システムは、いわゆる LAMP システム (Linux + Apache + MySQL + PHP) をベースとしつつ、いくつかの既存ソフトウェアを組み合わせることで構築した。以下、利用したソフトウェアについて述べる。

Zoomify

Zoomify[6]は、巨大な画像を自在に拡大・縮小・スクロール表示できる画像処理ライブラリである。Google Maps のようなものを想像するとわかりやすいだろう。あわせて、Zoomify では画像上にアノテーションおよびリンクを付加するための領域 (これをホットスポットと呼ぶ) を定義できる。これを利用して、画像上の文字

ごとにホットスポットを定義し、そこから文字情報を検索するためのスクリプトヘリンクを張ることで、文字データへのインターフェイスとした。また、ホットスポット定義に関する動作を変更するため、一部ソースコードに手を入れた。Zoomify の Enterprise Developer 版にはソースコードが付属しており、自由にコードを改変することができる。

なお、Zoomify には HTML5 版 (ビューアの記述言語は JavaScript) と Flash 版 (同じく ActionScript) とがあるが、今回は Flash 版を採用した。HTML5 版は後発であり、多角形のホットスポットを定義する機能がまだ実装されていない。本システムで扱う神官文字は手書きの筆記体であり、字形要素が必ずしも矩形領域に収まるものではないため、多角形領域の定義は必須である。Web 標準への準拠という点では、今後の HTML5 版の開発の進展状況によっては、HTML5 版へ移行する可能性もある。

Drupal

Drupal[7]は、GPLのもとで開発が続けられているオープンソース CMS である。基本となるコアシステムはシンプルであるものの、極めて拡張性に富んだモジュールシステムを実装しており、公式サイト上では実に 1 万を越える拡張モジュールが公開されている。これらを組み合わせることで多種多様なウェブシステムを構築することが可能であり、その拡張性の高さから、一種のアプリケーション構築フレームワークとしてとらえられることもある。本システムでは、サイトの基本構成や各ページの記述、コンテンツごとのアクセス権管理、フォーラムあるいはコンタクトフォーム等のコミュニケーション機能の実装にこれを用いた。現状では、アカウントを発行されたユーザに対してのみパピルス画像へのアクセスを許可している。

Subversion, Redmine

本システムに直接関わるものではないが、開発のサポートツールとして Subversion[8] (バージョン管理システム) および Redmine[9] (課題管理システム) を利用している。たとえ数名程度によるシステム開発であっても、これらのツールの利用は開発効率を大幅に向上させる。

開発言語

その他、サーバサイドのスクリプト記述には PHP を、クライアントサイドのスクリプト記述には JavaScript をそれぞれ使用した。Ajax を積極的に使い、極力軽快な画面レスポンスを得られるようにしてある。

以上の既存ソフトウェアの組み合わせにより、相当程度高速なシステム開発が実現できた。直

接的に開発に関与したのはほぼひとりだけであるが、設計開始から 2 ヶ月弱で暫定版の公開までこぎ着けている。

5.2. データ構造

古代エジプト語は死滅した言語であり、資料の判読そのものが問題になる。そのような状況下では、文字の解釈の揺れや、それに伴う語境の判定のずれといったことが生じる。あるいはまた、意図的であるか非意図的であるかは別として、文字が欠落する可能性もある。このように、古代エジプト語にあつては固定した読み方が決定できないため、多層的な解釈を許容するデータ構造としなければならない。そのようなことから、本システムでは、データストアシステムとして RDB を採用することとした。RDB は集合論的にデータを扱うものであり、テキストに対するマークアップと違って、言語の線条性に拘束されることがない。適正な正規化を行うことにより、ひとつの事象 (テーブル) 内で複数の解釈を保存可能である。また、テーブル同士の関連性 (リレーションシップ) を自由に表現可能であり、さらに保存すべき事象がふえた場合には自由にテーブル構造を拡張可能である。今後、古代エジプト語の文法研究の進展にともなうデータ構造を拡張する必要が生じるであろうが、その際にも比較的楽にその作業を行うことができるだろう。RDB を使用する利点は非常に大きいと言える。

5.3. テーブル定義

図 2 に本システムのデータベースの ER 図 (一部) を示す。ここに提示されていないテーブルも含め、本システムのデータベースには現状で 23 のテーブルが収められている。このうち、多層的な文字解釈・語解釈の保存に関係してくる部分は、以下の通りである。

labels : 資料上でのホットスポット定義
signs : 文字記号一覧表
signs_interpretations : ホットスポットと文字解釈との対応表
words_map : 文字解釈と語解釈との対応表
words_interpretations : 語解釈一覧表
words : 単語一覧表

また、これら以外に、語解釈と語解釈の関係性を記述するために、RDF 的な三項構造を持つ words_predicates というテーブルを用意してある。これらのデータ構造により、さまざまな文字・語解釈の可能性を排除することなく、柔軟なデータの保存が可能になる。今後は、神官文字と聖刻文字との対応関係、および、[4]に掲載されている字形をデータとして保存し、アプリ

ケーション上から参照可能なようにする予定である。

5.5. 機能

以下、本システムの基本機能のうち一部を紹介する。

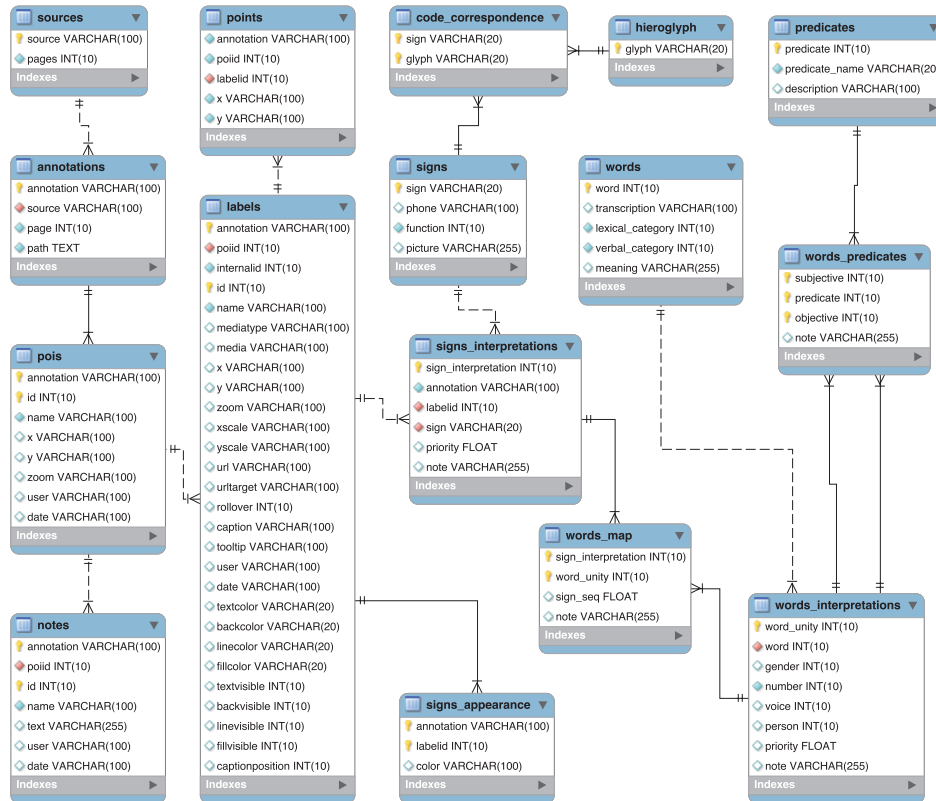


図2 ER図

5.4. データ構築法

ホットスポットの定義は、Zoomify Annotation System (以下 ZAS とする) が提供する機能を利用した。ZAS では、画像上を順にマウスクリックしていくことで多角形領域のホットスポットが定義できる (必要に応じて四角形、丸、三角形等も可能)。ただし、リンク先の定義等の細かな作業を ZAS 上で行うのは作業効率が悪いので、ホットスポットの名前および座標のみを ZAS 上で登録し、それら以外のデータについては、MySQL 上で更新することとした。ZAS のデータは XML で記述されているが、これを XSLT を用いて SQL 形式に変換し、データベースに登録した上で必要な更新作業を行った。更新後は、PHP を用いてデータベースから XML 形式でデータを出力し、この XML ファイルを再度 ZAS のデータファイルとした。また、文字解釈情報、語解釈情報等は、データベースから出力したデータをもとに Excel あるいはテキストエディタ上で編集し、これをデータベースに登録した。

5.5.1. ギャラリー

古代エジプト語の原資料画像が表示される。画像は自由に拡大・縮小・スクロールが可能である。特定の文字記号にマウスカーソルをポイントすると、多角形領域のホットスポットが示される。ホットスポットをクリックすると、Annotation info ウィンドウに、当該の文字記号および単語に関するデータが表示される。文字記号および単語の解釈が複数ある場合は、ポップアップウィンドウ上部のラジオボタンに「◎#1 ◎#2」のごとく解釈の選択肢が表示される。単

語中の文字の右にある文字 ID をクリックすると、データ検索ページに飛び、当該文字を含む単語リストが資料中から検索される。図3に画面例を示す。

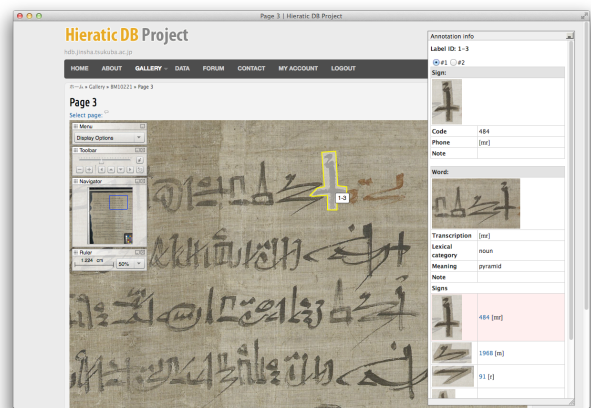


図3 ギャラリー

5.5.2. データ検索

資料名と文字 ID を選択して検索ボタンを押すと、当該文字を含む単語リストが資料中から検索される。検索結果リスト中の文字 ID をクリックすると、当該の文字を含む単語が再検索され

る。図4に画面例を示す。今後は、文字ID以外の様々な項目での検索を可能とする予定である。

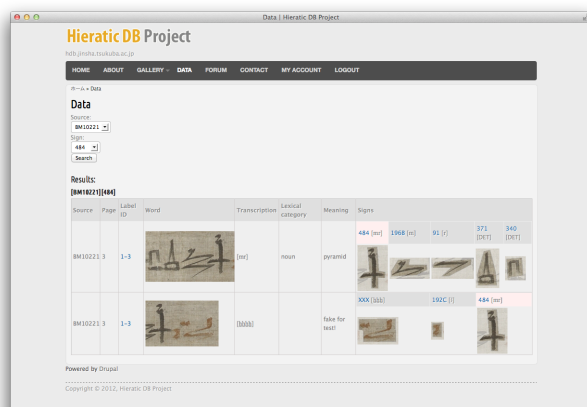


図4 検索画面

5.5.3. コミュニケーション機能

訪問者同士のコミュニケーションを促進するため、フォーラムおよびコンタクトフォームを実装した。これらについては、Drupalのユーザ役割管理機能により、非常に柔軟なアクセス権管理が可能である。また、Zoomifyのアノテーションデータにはユーザの概念があり、ユーザごとに異なるホットスポットの定義が可能である。今後、これを利用して、訪問者が自分自身のアノテーションを付すことができる機能を実装する計画である。それにより、研究者同士のコミュニケーションをより活性化させ、研究の一層の進展を望むことができる。

5.5.4. 今後の拡張計画

詳細な仕様決定はこれからであるが、外部システムとのデータ交換の仕組みとして、TEIによるXMLデータ出力機能を実装することを検討中である。この機能を実現することにより、世界的なレベルでより広範にデータを共有することが可能になると考えられる。TEIによるデータ表現は事実上人文情報学における標準的な立場を占めるに至っているが、古代エジプト語について言えば、そもそも文字や語の解釈が確定せず、場合によっては語境界が入れ子構造によって表現できない可能性もある。このようなデータをTEIでどのように扱っていくかということについて今後検討していく必要がある。

6. おわりに

本研究の発端となる目標は、古代エジプト語研究の課題を解決するために、神官文字のデータベースを作成することであった。現状では未だ試作版の作成に着手したところではあるが、目標とする古代エジプト語研究への貢献は、十分に達成が可能であると考えている。

このような当初の目的に更なる展望を付け加えるならば、古代エジプト語を対象とした本研究を足がかりとして、本システムを他言語や他文字にも対応することのできるプラットフォームへと昇華させたいと思っている。

言語資料のデータベースはすでに様々なところで作成されており、それぞれの分野で成果が上がっていると思われる。しかしながら、それらのデータベースの多くが個別システムの内部に閉じたものとなっており、プラットフォームとしての性格を帯びたものは決して多くはないのではなかろうか。

プラットフォームの構築は第一にシステムの共有によって達成されるが、さらに訪問者同士のコミュニケーション機能も重要な要素だと考えられる。

謝辞

本研究は科学研究費「基盤研究(C)：課題番号24520452」ならびに「研究活動スタート支援：課題番号22820007」の成果の一部である。なお、システムの構築については、筑波大学非常勤研究員の高橋洋成氏からアドバイスと協力を受けた。記して感謝する次第である。

参考文献

- [1] 永井正勝: 古代エジプト神官文字の画像を利用した字形データベースについて：中エジプト語の資料を中心とした事例紹介, じんもんこん 2011『人文科学とコンピュータシンポジウム論文集：「デジタル・アーカイブ」再考：いま改めて問う記録・保存／活用の技術』, 情報処理学会シンポジウムシリーズ, Vol.2011, No.8, pp.389-394, 2011.
- [2] Loprieno, A: Ancient Egyptian: A Linguistic Introduction, 2nd Edition, Cambridge, 2010.
- [3] Möller, G: Hieratisch Paläographie, 3 Vols., Leipzig, 1909-1912.
- [4] 永井正勝: 中エジプト語の進行相の否定文について：「否定辞 *nn* + 主語 + 前置詞 *hr* + 不定詞」構文の再検討, 『オリエント』, 日本オリエント学会, 第53号第2号, pp.34-54, 2010.
- [5] British Museum: Selected Papyri in the Hieratic Character from the Collections of the British Museum, Vol.2, Manchester, 1930.
- [6] <http://www.zoomify.com/>
- [7] <http://drupal.org/>
- [8] <http://subversion.apache.org/>
- [9] <http://www.redmine.org/>