UNITS OF LANGUAGE – UNITS OF WRITING

7th International Workshop on Writing Systems and Literacy

September 30th – October 1st 2010

Paris, France

Association for Written Language and Literacy /
Laboratory Language-Music-Society FRE 3324 CNRS

PROGRAMME

UNITES DE LANGUE – UNITES D’ÉCRITURE

7ème atelier international sur les systèmes scripturaux et la littéracie

30 septembre 2010 – 1er octobre 2010

Paris, France

Association pour la langue écrite et la littéracie /
Laboratoire Langues-Musiques-Sociétés FRE 3324 CNRS
INFORMATION

Lieu / Location
Université Paris Descartes - Sorbonne - 1, rue Victor Cousin 75005 Paris
Amphithéâtre Emile Durkheim (Escalier I, 1er étage à droite / Stairs I, 1st floor)

Plan de la Sorbonne

Inscriptions / registration

<table>
<thead>
<tr>
<th></th>
<th>Before 09-28 avant le 28-09</th>
<th>Cash at the site</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chercheurs / Scholars</td>
<td>50 €</td>
<td>80 €</td>
</tr>
<tr>
<td>Etudiants / Students</td>
<td>15 €</td>
<td>30 €</td>
</tr>
<tr>
<td>Déjeuners / Lunches (Thursday &amp; Friday)</td>
<td>80 €</td>
<td>80 €</td>
</tr>
<tr>
<td>Dîner informel du mercredi / Wednesday informal dinner</td>
<td>40 €</td>
<td>40 €</td>
</tr>
<tr>
<td>Dîner du jeudi / Thursday conference dinner</td>
<td>50 €</td>
<td>70 €</td>
</tr>
</tbody>
</table>

Comité d’organisation / organization committee

Amandine Bergère (Laboratory Language-Music-Society UMR 8099 CNRS-Paris Descartes)
Terry Joyce (School of Global Studies Tama University)
Anneke Neijt (Dutch Department, University of Nijmegen)
Beatrice Primus (German Department, University of Cologne)
Restaurants

Dîner informel du mercredi 19h30
Wednesday informal dinner 7:30 pm

Name: L’autre café
Address: 62, rue Jean-Pierre Timbaud
75011 Paris
Métro: Parmentier (line 3), Goncourt (line 11) or Couronnes (line 2)

Déjeuners / Lunches (Thursday & Friday)

Name: Tugalik
Address: 4, rue Toullier, 75005 Paris
2 minutes walk away from la Sorbonne (South, uphill)

Dîner du jeudi 20h00/
Thursday conference dinner 8:00 pm

Name : Le Procope
Address: 13, rue de l'Ancienne Comédie,
75006 Paris
Métro : Odéon (line 4, line 10)
10 minutes walk away from la Sorbonne
<table>
<thead>
<tr>
<th>30th</th>
<th>SYSTEMES SCRIPTURAUX / WRITING SYSTEMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:30</td>
<td>Introduction</td>
</tr>
<tr>
<td>09:45</td>
<td>1 David Olson</td>
</tr>
<tr>
<td></td>
<td>Writing and rationality</td>
</tr>
<tr>
<td>10:30</td>
<td>2 Nanna Fuhrhop &amp; Kristian Berg</td>
</tr>
<tr>
<td></td>
<td>Vowel graphemes: in search of a definition</td>
</tr>
<tr>
<td>11:00</td>
<td>Pause / Break</td>
</tr>
<tr>
<td>11:15</td>
<td>3 Terry Joyce, Bor Hodoscek &amp; Kikuko Nishina</td>
</tr>
<tr>
<td></td>
<td>Orthographic representation within the Japanese writing system</td>
</tr>
<tr>
<td>11:45</td>
<td>4 David Roberts &amp; Stephen L. Walter</td>
</tr>
<tr>
<td></td>
<td>Representing meaning rather than sound in the orthographies of tone languages</td>
</tr>
<tr>
<td>12:15</td>
<td>Présentation orale des posters / Posters’ oral presentation</td>
</tr>
<tr>
<td>12:30</td>
<td>Déjeuner / Lunch</td>
</tr>
<tr>
<td>13:30</td>
<td>Présentation des posters / Posters’ presentations</td>
</tr>
<tr>
<td>14:30</td>
<td>5 Michel Fayol</td>
</tr>
<tr>
<td></td>
<td>Units of processing in writing</td>
</tr>
<tr>
<td>15:15</td>
<td>6 Martin Evertz, Timo B. Röttger &amp; Beatrice Primus</td>
</tr>
<tr>
<td></td>
<td>Prosodic Units in Language and Writing: Mora, Syllable, and Foot</td>
</tr>
<tr>
<td>15:45</td>
<td>Pause / Break</td>
</tr>
<tr>
<td>16:00</td>
<td>7 Dorina Veldhuis &amp; Jeanne Kurvers</td>
</tr>
<tr>
<td></td>
<td>Online language processing units and offline segmentation units</td>
</tr>
<tr>
<td>16:30</td>
<td>8 Kasper Juffermans &amp; Dorina Veldhuis</td>
</tr>
<tr>
<td></td>
<td>Learning histories and word segmentation in Mandinka textual practices</td>
</tr>
<tr>
<td>17:00</td>
<td>9 Martin Neef</td>
</tr>
<tr>
<td></td>
<td>The Potential Beginning of a Word as the crucial boundary in written representations</td>
</tr>
<tr>
<td>17:30</td>
<td>AWLL meeting / Réunion de l’association pour la langue écrite et la littéracie</td>
</tr>
<tr>
<td>20:00</td>
<td>Dîner / Dinner</td>
</tr>
</tbody>
</table>
# POSTERS

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>La translittération entre alphabet thâï et alphabet latin. Difficultés d'une normalisation et de son enseignement</td>
<td>Frédéric Carral</td>
</tr>
<tr>
<td>How the Korean alphabet Affects Acquisition of English by Korean Learners</td>
<td>Seoyoung Chae &amp; Youngjun Jang</td>
</tr>
<tr>
<td>An Optimality Theory Based Model of Intrasentential Capitalization in German</td>
<td>Annika Hübl</td>
</tr>
<tr>
<td>Syntactic process in written and oral narrative texts</td>
<td>M. Luisa Silva, Verónica Sánchez Abchi &amp; A. María Borzone</td>
</tr>
<tr>
<td>Students’ Achievement on Impromptu Writing Test versus Interactive Writing Assessment through Feedback Exchange</td>
<td>Fakhrisadat Shariati</td>
</tr>
<tr>
<td>Portuguese Language Orthographic Agreement of 1990: homographs, prescription, opacity</td>
<td>Francisco Miguel Valada</td>
</tr>
<tr>
<td>Time</td>
<td>Session</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>09:30</td>
<td>Vivian Cook</td>
</tr>
<tr>
<td></td>
<td><em>Questions for L2 WS Research</em></td>
</tr>
<tr>
<td>10:15</td>
<td>Amalia Bar-On &amp; Dorit Ravid</td>
</tr>
<tr>
<td></td>
<td><em>Syntactic and semantic processes in mapping written onto spoken Hebrew units</em></td>
</tr>
<tr>
<td>10:45</td>
<td>Pause / Break</td>
</tr>
<tr>
<td>11:00</td>
<td>Christiane Soum Favaro, Cecilia Gunnarsson &amp; Pierre Largy</td>
</tr>
<tr>
<td></td>
<td><em>Le traitement de la liaison en production écrite</em></td>
</tr>
<tr>
<td>11:30</td>
<td>Arina Banga, Esther Hanssen, Robert Schreuder &amp; Anneke Neijt</td>
</tr>
<tr>
<td></td>
<td><em>How subtle differences in writing may influence interpretation</em></td>
</tr>
<tr>
<td>12:00</td>
<td>Nadya Dich</td>
</tr>
<tr>
<td></td>
<td><em>The role of native language orthography in learning to spell in ESL</em></td>
</tr>
<tr>
<td>12:30</td>
<td>Déjeuner / Lunch</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>14:00</td>
<td>Hans Basbøll &amp; Claus Lambertsen</td>
</tr>
<tr>
<td></td>
<td><em>Matching units of language and units of writing: Danish as a test case</em></td>
</tr>
<tr>
<td>14:30</td>
<td>Susanne Borgwaldt, Emőke Jakab &amp; Patrick Bolger</td>
</tr>
<tr>
<td></td>
<td><em>Letter detection in L1 and L2 the role of Dutch orthography in reading ESL</em></td>
</tr>
<tr>
<td>15:00</td>
<td>Amandine Bergère</td>
</tr>
<tr>
<td></td>
<td><em>Littéracie en langue seconde chez des adultes chinois: un modèle interlinguistique</em></td>
</tr>
<tr>
<td>15:30</td>
<td>Pause / Break</td>
</tr>
<tr>
<td>15:45</td>
<td>Dominiek Sandra &amp; Lien Van Abbenyen</td>
</tr>
</tbody>
</table>
|       | *Twice trapped by our memory for the spelling pattern of high-frequency verb homophones:*
|       | *once during spelling and once during (re-)reading.*                                      |
| 16:15 | Esther Hanssen, Arina Banga, Robert Schreuder & Anneke Neijt                             |
|       | *The interpretation of linking elements in Dutch compounds*                              |
| 16:45 | Sandra Beyermann                                                                        |
|       | *Phonographic processing units in German*                                                 |
| 17:15 | Conclusion                                                                               |

Alternate:

21 Anat Hora & Dorit Ravid *Children's knowledge of the Hebrew spelling system and the drive to explain it: A developmental study*
Vowel graphemes:  
In search of a definition  
Nanna Fuhrhop & Kristian Berg  
Universität Oldenburg, Germany

If the unit „graphematic syllable“ is accepted, the logical question to follow is whether we can identify its core and periphery similarly to the constituents in the phonological syllable. Basically, the problem can be paraphrased in the following way: Are there two classes of graphemes, i.e. vowel and consonant (or „core“ and „periphery“) graphemes?

When it comes to defining vowel graphemes, existing approaches stress the correspondences to phonology. For them, a vowel grapheme is a grapheme that corresponds to a vowel phoneme. Contrary to this, we would like to present two analyses that are not dependent on phonology.

The formal analysis is based on the letter analysis of the roman minuscles (Primus 2004). Problems arise with the encoding of diphthongs, hiatus, and the marking of tenseness. Each of these will be addressed by using a number of different languages (English, German, Dutch, French).

The distributional analysis is based on “multidimensional scaling“ (Kruskal & Wish 1978). It reveals – quite surprisingly – differences between the aforementioned languages: English and Dutch both show two distinct sets of items (vowels and consonants in traditional approaches). In German, however, there seems to be no clear-cut boundary.

Both perspectives will finally be brought together. The conclusions that are drawn will be valid not only for the languages that were investigated.

References
Orthographic representation within the Japanese writing system

Terry Joyce¹, Bor Hodoscek² & Kikuko Nishina²
¹Tama University, Japan
²Tokyo Institute of Technology, Japan

The Japanese writing system consists of morphographic kanji, two syllabographic sets of hiragana and katakana, alphabetic romaji and Arabic numerals. Depending on one’s perspective, this unique mixture of graphic elements can be seen either as the worst possible solution to written language or as potentially supporting great orthographic flexibility. Whichever position one favors, unquestionably, the Japanese writing system provides an exceptional case study for investigating the complex relationships between units of language and units of writing.

This paper discusses some orthographic issues for the Japanese writing system from the perspective of creating some word lists based on UniDic, a morphological analyzer dictionary (Den, et al, 2007), developed as part of National Institute for Japanese Language and Literature’s ongoing Balanced Corpus of Contemporary Written Japanese (BCCWJ) Project (Maekawa, 2008). Utilizing UniDic, we have compiled from the BCCWJ corpus about 30 lists of short-word units (based on POS categories) covering approximately 93.3 million tokens. More relevant to issues of orthographic representation, the lists presently consist of approximately 126,000 lemmas that are represented by about 178,000 orthographic forms. For instance, Ogura, et al. (2010) recently discuss UniDic’s treatment of 合う /au/ and オサマル/osamaru/. Although UniDic distinguishes between two lemmas—合う /au/ “merge, unit; fit, suit” represented orthographically as あう and 合う, and 会う /au/ “meet” represented orthographically as あう and 会う, 逢う, 遭う, and 遇う—in the former case, it regards オサマル as just one lemma, even though it can be represented orthographically as おさまる, 修まる, 収まる, 収る, 治まる, 納まる, 納る, and 蔵まる. While such treatments are arguably consistent with UniDic’s objectives, as discussed in the paper, they also highlight some difficult problems surrounding the notion of orthographic representation within the Japanese writing system.
Some orthographies represent tone phonemically by means of diacritics; others favour zero marking. Neither solution is entirely satisfactory. The former leads to graphic overload; the latter to a profusion of homographs. Both may reduce fluency.

But there is a "third way" which does not align itself with either extreme: highlighting meaning rather than sound. To test this approach, we developed two experimental strategies for writing Kabiye: a grammar (meaning-based) orthography and a tone (sound-based) orthography. Both are modifications of the standard orthography which does not mark tone. We tested these in a quantitative experiment involving literate native speakers which included dictation, spontaneous writing and oral reading.

The fact that writers of the experimental grammar orthography perform faster and more accurately than writers of the experimental tone orthography suggests that they have an innate consciousness of Kabiye morphological structure, and that this awareness exceeds that of its phonology. Moreover, frequency of exposure to a particular grammatical construction in natural contexts proves to be a strong predictor of performance for those writing the experimental grammar orthography, but it confers no advantage for those writing the experimental tone orthography. This provides evidence that representing meaning instead of sound may be an effective orthographic strategy in some tone languages.

**Keywords:** Tone, Grammar, Orthography, African languages, Quantitative experiment
Prosodic units such as mora, syllable, and foot are well established within spoken language phonology. They serve as a domain for phonological rules including stress assignment. In written language research, the syllable – taken as a graphematic unit – has deserved some attention (cf. Roubah & Taft 2001, Domahs et al. 2001, Primus 2003, Weingarten et al. 2004), and there is evidence that graphematic syllabification influences phonological syllabification (cf. Treiman et al. 2002). The graphematic weight (mora) and the graphematic foot are largely neglected in written language research (but see Röttger et al., under review).

This talk will focus on the impact of graphematic weight, i.e., the complexity of the syllable rhyme, on foot structure and stress assignment in German. In order to isolate graphematic weight from phonological weight, we will focus on silent letters such as mute <h>, graphematic gemination, and complex graphemes, as in the pseudowords <Ha.bo.lu> vs. <Ha.bo.luh> = [ha.bo.lu], <Baruntap> vs. <Baruntapp> = [ba.rʊn.tap], and <Do-san.rax> vs. <Do-san.racks> = [do.zan.raks]. We will present both structural and experimental evidence based on pseudoword production studies showing that graphematic weight has an impact on foot structure and stress assignment in German.

Our results show the relevance of a graphematic hierarchy of units above the level of the grapheme. We also offer evidence in favour of the claim that graphematic structure may affect phonological structure.

References


Röttger, Timo B. et al. (under review.). Structural factors affecting the assignment of word-stress in German.


Online language processing units and offline segmentation units: the influence of literacy and language

Dorina Veldhuis & Jeanne Kurvers
Department of Language and Culture Studies, Tilburg University, The Netherlands

Studies with young, pre-literate children and illiterate adults have shown that word-awareness and literacy are related to each other (cf. Kurvers & Uri 2006; Ramachandra & Karanth 2007). Moreover, cross-linguistic studies have shown that speakers of languages that do not mark word boundaries in writing do not always easily recognize words in their language (Bassetti 2005; Hoosain 1992). Therefore, the idea that words are basic units of language is at stake, and this might be reinforced by the fact that languages differ, also in what counts as word boundaries in writing. In agglutinative languages like Turkish ‘at school’ counts, for instance, as one word (okulda), with a bound morpheme, whereas in Dutch, the same meaningful construction should be written as two separate words (op school). Accordingly, literacy in languages with alphabetic writing systems will not always lead to awareness of similar units.

To see how literacy and knowledge of specific languages affect children’s metalinguistic awareness, we tested in several offline experiments the units which monolingual Dutch and bilingual Turkish-Dutch pre-literate and literate children segment. This study showed how experience with writing and the structure of a specific language affect the units which are most easily recognized from a language.

Moreover, we investigated (in a pilot-study) whether the units which people recognize from language in offline tasks correspond to the units which people process. For this, we developed a click-task, a self-paced listening task, a production task, and an eye-tracking task, which we conducted among the same groups of monolingual and bilingual readers and pre-readers.

In our talk, findings from our offline experiments and the first results from our online pilot experiments will be discussed. With this, we will show how experience with spoken and written language relates to the units which monolingual Dutch and bilingual Turkish-Dutch pre-readers and readers recognize and process.

References


Learning histories and word segmentation in Mandinka textual practices

Kasper Juffermans & Dorina Veldhuis
Department of Language and Culture Studies, Tilburg University, The Netherlands

This paper combines two projects: an ethnographic sociolinguistic study of literacy in The Gambia (West Africa), and an experimental linguistic study on the influence of literacy on the basic units (‘building blocks’) of language.

Throughout history, linguists have addressed the question what the basic units of human languages are. Some of these have suggested that utterances are based on ‘words’. Studies into children’s metalinguistic awareness, however, have shown that young, pre-literate children do not segment language into words as adults do, but that children only become aware of words as linguistic units once they learn to read and write. What is more, word awareness depends on the writing system in which one becomes literate.

In our paper, we deploy naturalistic data from the ethnographic study to investigate how literacy in a specific language affects unit awareness. Samples were collected of grassroots writings by individuals with various learning histories and literacy-backgrounds in the multilingual ecology of Gambian society. Children in formal education are taught to read and write in the official, post-colonial language English. Gambian local languages such as Mandinka, however, are not used for literacy learning in mainstream (formal) education, but only in the (non-formal) adult literacy classes. Children and adults are thus subjected to two radically different regimes of literacy learning.

Our paper compares two versions of a Mandinka text: one written by a non-formally (Mandinka) educated man and one by a formally (English) educated younger urbanite. We address the question of word segmentation in the two versions, arguing that being literate in either English or Mandinka affects the way people process language. Different paths to literacy lead to different word boundaries in the text, and thus a different interpretation of the basic units of language.
The Potential Beginning of a Word as the crucial boundary in written representations

Martin Neef
Braunschweig Institute of Technology, Germany

In the analysis of writing systems the existence of equivalents to well-established grammatical units as phonological syllables or morphemes is often disputed. A phenomenon like hyphenation at the end of a line of German words shows a relation to these grammatical units, but neither the phonological syllable nor the morpheme alone, nor the two in combination suffice to explain the pertinent data, at least if the structure of the putative written units are derived straightforwardly from the structure of the respective grammatical units. For example, written words may be hyphenated between the prefix and the stem but not between the stem and the suffix, as demonstrated by the example Enteisung ‘defrosting’:

<table>
<thead>
<tr>
<th>(1)</th>
<th>morphological structure</th>
<th>Ent-</th>
<th>eis-</th>
<th>ung</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>prefix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>stem</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>suffix</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hyphenation of written word</td>
<td>Ent-</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>eis-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sungs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>*Ent-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>eis-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>*-ung</td>
<td></td>
</tr>
</tbody>
</table>

Thus, the position of a hyphen does not always coincide with a morpheme boundary. At the same time, examples abound where hyphens are in positions not expected if hyphenation is conceived of as being derived from syllabification:

<table>
<thead>
<tr>
<th>(2)</th>
<th>phonology</th>
<th>spelling</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>widrig ‘adverse’</td>
<td>wid-rig</td>
</tr>
<tr>
<td></td>
<td>Städte ‘cities’</td>
<td>Städ-te</td>
</tr>
<tr>
<td></td>
<td>Hexe ‘witch’</td>
<td>He-xe</td>
</tr>
</tbody>
</table>

In my talk, I will discuss a boundary called the ‘Potential Beginning of a Word’ or PW-boundary (cf. Neef 2008). The beginning of each stem and each prefix constitutes such a boundary. Hyphenation in German operates in the first place on this boundary, while other constraints apply in domains not interrupted by PW-boundaries. Other phenomena that are sensitive to PW-boundaries include specific correspondence rules and word internal punctuation marks. PW-boundaries in spelling correspond to the same kind of boundary in grammar.

Reference:
Questions for L2 WS Research
Vivian Cook, Newcastle University, United Kingdom

Second language writing systems research draws on second language acquisition research and writing system research. As SLA research, it is still in its early stages, concerned with transfer from L1WS to L2WS and with Error Analysis but has started to develop interesting ideas of its own: it is not concerned with cross-linguistic comparison but with people who learn or use more than one language. It is concerned at every level of writing from physical to direction of writing to punctuation to spelling. Looked at from a multi-competence perspective, its core is the unique ways in which L2 users learn and use writing systems, both in the L2 and the L1, and the effects of knowing two writing systems on the other aspects of cognition. L2WS research can test issues about writing system development by disengaging acquisition from maturation and by allowing experimental variation that would be unethical with children acquiring the L1WS: in other words it is not just a side issue but can cast light on the mainstream of research into writing systems and its implications for the human mind.
A prominent property of the nonvoweled Hebrew orthography is homography. For example, the homographic string מְדַבֵּר MDBR stands, inter alia, for midbar 'desert', medaber '(is) speaking', and mi-davar 'from-thing'. Homography does not hinder mapping spoken into written units (i.e. writing), but it does constitute a problem in mapping written into spoken units – that is, reading. Subsequently, accurate identification of words in a nonvoweled Hebrew text requires semantic and syntactic contextualization.

The aim of the current study was to examine how and to what extent novice, experienced and proficient readers rely on contextual linguistic processes while decoding nonvoweled Hebrew words. To this effect, 130 Hebrew-speaking pupils in six grade levels and 20 adults were administered the Homographic Garden Path Reading Task consisting of 16 homographic words, each embedded in two sentential contexts - Garden Path and Non-Garden Path. Each homographic word targeted a single phonological option with a single meaning. For example, MDBR מְדַבֵּר targeted only one option - midbar 'desert', to serve in both sentential contexts. The GP context provided an initially misleading context, resulting in reading the target word as a sententially inappropriate spoken option, thus entailing error detection and repair. In contrast, the NGP sentence supported the accurate decoding of the target option.

Results indicate that successful reading of target words in the facilitating NGP context already occurs at the beginning of 2nd grade, while erroneous reading in the misleading GP context occurred in all age groups. Novice readers made, in addition, numerous reading errors independently of the context. This study demonstrates the developmental path to skilled Hebrew reading, which requires making full usage of the semantic / syntactic context in order to map written onto spoken units. We will elaborate on how efficient error detection and repair mechanisms enabling monitoring and self-repair develop with age and schooling.
Le traitement de la liaison en production écrite

Christiane Soum Favaro¹, Cecilia Gunnarsson, Pierre Largy

¹Octogone – Laboratoire J. Lordat, R 22, Université de Toulouse 2, France

Le développement et les processus de l’acquisition de la liaison à l’oral ont été étudiés entre autres dans le but d’établir un ordre dans l’acquisition de ce phénomène typiquement français (voir par exemple Chevrot & Fayol, 2000 ; Chevrot et al., 2007, 2008 ; Wauquier-Gravelines & Braud, 2005). D’après ces études, la liaison obligatoire semble acquise à l’oral vers 6 ans.

Nous proposons d’élargir le champ de recherche en voulant étudier la liaison et son développement dans le contexte de la production écrite. Pour ce faire, nous comparons trois cas d’apprentissage de la langue écrite en français : celui des enfants, scolarisés en France dont la langue maternelle est le français, celui des enfants kanaks de la Nouvelle Calédonie ayant le drehu comme L1 et qui sont scolarisés en français L2 dès la maternelle et enfin celui d’apprenants adolescents du français langue étrangère (FLE suédophones). Dans le premier cas, l’acquisition de la langue orale a précédé l’acquisition de la langue écrite. Dans le deuxième cas, l’acquisition du français s’effectue parallèlement à leur alphabétisation. Dans le troisième cas l’acquisition de la langue orale et de la langue écrite se fait simultanément, mais les sujets maîtrisent déjà le code écrit dans leur langue maternelle.

Dans une étude préalable (Gunnarsson, Soum, 2009), nous avons pu voir que les enfants français L1 et les apprenants suédophones de français LE semblaient traiter la liaison des consonnes /z/, /n/ et /t/ à l’écrit de manière différente. Les données de cette étude indiquent que les sujets FL1 répondent aux paramètres acoustico-phonologiques tandis que les FLE sont sensibles à la fréquence. Il s’agit maintenant de comparer ces données avec celles des jeunes kanaks, qui représentent une autre situation d’apprentissage et d’acquisition du français L2, i.e. celle d’un apprentissage précoce qui démarre dès que les enfants commencent l’école où ils sont scolarisés en français.

Nous rapportons les données d’une étude auprès de 86 enfants français L1 du CE1 au CM2 (de 7 à 10 ans) ; 80 enfants kanaks français L2 du CE1 au CM2 (de 7 à 10 ans) ; 28 lycéens suédois apprenant le français LE (de 16 à 19 ans, de la 4ᵉ à la 6ᵉ année d’apprentissage). Les données sont traitées dans une perspective développementale où nous tentons d’établir un itinéraire développemental de l’acquisition de la liaison à l’écrit, chez ces trois populations.

**Références**


Part of one’s general knowledge of the world is that there are more ants in an ant hill then there are dogs in a dog house. Taking this as point of departure, the present study investigates the possible influence of one’s writing system on such numerosity judgements.

In Dutch, noun-noun compounds with the linking element en between the two constituents may occur, for example in slangenbeet ("snakebite"). This linking element en is homophonous with the regular plural affix for nouns, and previous research has shown that, for native speakers of Dutch, linking en triggers the activation of plural semantics (Schreuder et al., 1998). For Afrikaans and Frisian, the spelling of the linking element and/or the plural affix is different.

<table>
<thead>
<tr>
<th>Language</th>
<th>Plural affix</th>
<th>Linking element</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch</td>
<td>en</td>
<td>en</td>
</tr>
<tr>
<td>Frisian</td>
<td>en</td>
<td>e</td>
</tr>
<tr>
<td>Afrikaans</td>
<td>e</td>
<td>e</td>
</tr>
</tbody>
</table>

This raises the question whether speakers of Frisian or Afrikaans with Dutch as a second language have the same plurality interpretations for the left constituent in Dutch noun-noun compounds as native speakers of Dutch. The participants were exposed to two types of compounds, namely one for which the first constituent is considered as singular, as in slangenbeet ("snakebite"), and one for which the first constituent is considered as plural, as in boekenkast ("bookcase"). They indicated the degree of plurality of the first constituents on a seven-point scale.

We found differences between native speakers of Frisian and native speakers of Dutch, and also between native speakers of Afrikaans and native speakers of Dutch. We discuss these interpretation differences in the light of the spelling systems of the three languages.
The role of native language orthography in learning to spell in English as a second language

Nadya Dich
Department of Human Development, Cornell University, USA

The role of native language orthography in learning to spell in English as a second language

The presentation discusses the results of a cross-linguistic study in which native speakers of Danish, Italian, and Russian were compared on their English-as-a-second-language (ESL) spelling skills. The three orthographies – Danish, Italian, and Russian – differ significantly in their depth: Danish is a deep orthography where one letter can correspond to many sounds and vice versa. Italian is a shallow orthography with mostly unambiguous letter-sound correspondences. Finally, in Russian, converting letters into sounds is usually straightforward, whereas converting sounds to letters is often not. However, many ambiguities in spelling can be resolved by applying morphology-based rules in Russian. Thus, the question investigated in the study was how the depth and the regularity of a native language orthography influence learning to write in the opaque English orthography in ESL learners.

The study consisted of several tasks measuring different aspects of English spelling proficiency: irregular word spelling, using morphological cues to determine the correct spelling, and pseudoword spelling. A questionnaire was also administered, aimed at controlling for factors potentially influencing ESL spelling.

Participants were 100 Danish, 91 Italian and 105 Russian adult ESL learners (mean ages 23, 24, and 22 respectively) with intermediate to advanced English proficiency.

While no significant between-group differences was observed for irregular word spelling, Italians made the most errors in the task requiring the use of morphological spelling cues.

These results are discussed with respect to how first language literacy skills may be transferred to aid second language literacy acquisition and how the characteristics of first language orthography may shape the cognitive mechanisms of spelling in ESL.
Matching units of language and units of writing:
Danish as a test case

Hans Basbøll & Claus Lambertsen
Center for Child Language, Institute for Language and Communication, University of Southern Denmark

We present analyses from a project testing a model on phonology-morphology-orthography interaction (Basbøll The Phonology of Danish Oxford UP 2005) in our tool Olam:

1) Levels of phonological units: sound units, phonemes, morphophonemes. We test the relation between all three units and letters, evaluating transparency of each relation, distinguishing between the directions from sound to writing and inversely, but hereby using the same principles.

2) Domains and boundaries: syllables, morphemes, words. The phonological syllable is a crucial unit for orthography in Danish and is incorporated in the system, and we shall also show the effect of morphological boundaries.

3) Correspondences between letters and (morpho)phonemes: [+Principle] or [–Principle], and [+Sound] or [–Sound]. The mapping of the vowel of the word gæst, which is pronounced with the vowel of the letter name æ, is both +P and +S; in the rhyme word hest, the mapping of the vowel is +P but –S (also –P +S and –P –S are possible). This 4-way classification even applies to spelling errors.

4) Distinguishing different parts of the lexicon: native, “English”, etc. When the system (1-2-3) is implemented, some of the scripts which match units of language and writing are specific for e.g. English, French or German. Thus the vocabulary analysed is classified as to donator languages as far as present-day orthography is concerned, and this information can be compared to historical evidence.

5) Can the model presented be used for other languages? French as an example. We have used the principles of our Olam-system to analyse a large French lexical corpus: units (e.g. of sound) and some other linguistic parameters are redefined, and the system calculates ”pretending it is Danish”. This works surprisingly well, illustrating the potentials of our model.
Letter detection in L1 and L2: the role of Dutch orthography in reading English as second language.

Susanne Borgwaldt¹, Emőke Jakab², and Patrick Bolger³

¹University of Braunschweig, Germany
²University of Amsterdam, The Netherlands
³University of Alberta, Canada

This study investigates digraph effects during letter detection. Rey et al. (2000) demonstrated in two experiments with English and French participants that a vowel letter was detected slower if it was embedded in a digraph such as a in boat than when it appeared as a single vowel such as a in brat.

To test the time-course of this effect in English and Dutch, Bolger et al. (2009) combined the letter-detection task with the masked-priming paradigm, and embedded the grapheme manipulation in pseudoword primes that were presented at various prime durations. They found that Dutch readers displayed the digraph effect with 50 ms primes, whereas English readers did not. The authors concluded that the masked-priming digraph effect might depend on a language’s orthographic transparency, and might therefore be easier to observe in transparent orthographies like Dutch than in opaque ones like English.

The current study focuses on investigating the transfer of L1 reading strategies to the L2, using the masked-priming letter-detection paradigm with English word primes with two participants groups: L1 speakers of English, and Dutch students with English as L2. Comparing the results of L1 and L2 readers allows to investigate possible transfer of L1 (Dutch) processing strategies to the processing of L2 (English) word primes. Whereas the L1 speakers of English showed merely a weak digraph trend at 50 ms for low-frequency words only, the L2 speakers of English again showed stronger effects for both high and low frequency words at this prime duration.

Together with the results of Bolger et al., it seems that digraph effects in letter detection reflect bottom-up letter activation, and that L2 reading is influenced by either L1 reading strategies or L1 perceptual characteristics.

References


Littératie en langue seconde chez des adultes chinois :
un modèle interlinguistique

Amandine Bergère
Laboratoire Langue-Musique-Société FRE 3324 CNRS, France


L’étude porte sur un corpus de dictées de syllabes produites par des apprenants adultes chinois scolarisés en moyenne 8 ans et débutant leur apprentissage du français. La dictée était assortie d’une consigne de comptage de sons (combien de sons avez-vous entendu ?) qui permet de comprendre comment les apprenants se représentent les relations entre les unités phonologiques et les unités graphiques du français.

En conclusion, la façon dont les apprenants conceptualisent les relations entre les phonoèmes et les graphèmes du français est peu corrélée à la durée de leur scolarisation ou leur maîtrise de l’écriture chinoise, mais fortement corrélée à leur maîtrise du pinyin, la transcription alphabétique du chinois standard.

Références


Twice trapped by our memory for the spelling pattern of high-frequency verb homophones: once during spelling and once during (re-)reading.

Dominiek Sandra and Lien Van Abbenyen
University of Antwerp, Belgium

Our previous research explained why spelling errors on Dutch homophonous verb forms are so persistent, despite the descriptive simplicity of the underlying rules: the high-frequency spelling pattern tends to cause intrusion errors when spelling the low-frequency form (homophone dominance). We report two experiments, showing that the same phenomenon manifests itself in the reading process. In a phonological decision task, which required participants to decide whether two-word combinations sound correctly in Dutch, irrespective of their spelling, we found that yes-responses to incorrectly spelled verb homophones became faster as their spelling pattern became more frequent: incorrect high-frequency verb forms are apparently not easily detected. This was demonstrated in an analysis of the response times to two homophone types: homophones whose suffix is inaudible, involving a homophonic relationship between the first and third person singular of the present tense, and homophones whose suffix is audible, involving a homophonic relationship between the third person singular present tense and the past participle. Our results indicate that spellers are twice trapped by their memory for the spelling of homophonous verb forms: during the spelling process itself and again when re-reading their text. This is part of the explanation why errors on these verb forms are so persistent, even though their spelling is governed by descriptively quite simple rules.
Dutch compounds may display form variation in the choice of linking en (e.g. koekbakker and koekenbakker for 'pastrycook'). Two experimental studies with written input and one with spoken input investigated the contribution of rhythm on the plural semantics of this linking en, which is phonologically and orthographically similar to the plural suffix -en.

The first study showed an effect of rhythm on plurality in written pseudo-compounds. Pseudo-compounds that contained linking en to prevent a stress clash (burbentijpis) were considered to bear less plural meaning than compounds with en in a neutral stress context (burbenonttijp). The second study with written, existing compounds revealed no effect of rhythm, whereas the third study with spoken, existing compounds did: Spoken compounds that contained linking en to prevent a stress clash (bloemenbak) were considered to bear less plural meaning than compounds with en in a neutral stress context (zwaluwenei). One possible explanation for the presence of a rhythmic effect in written pseudo-compounds but not in existing compounds could be in terms of processing. The route from form to meaning may run directly from orthography to meaning in existing compounds, because their constituents and their combinations are stored in the lexicon. In pseudowords, however, there is no direct route from orthography to meaning because they lack a form-meaning relationship. On the way from orthography to meaning, the phonological route cannot be skipped here. In addition, the use of prosody is inevitable in speech, which could explain the presence of a rhythmic effect in spoken, but not in written, existing compounds. Overall, we conclude that prosody steers someone's interpretation of the extent to which the linking element in a compound bears plural meaning. Our studies show that the two ways of reading, with and without phonology, influence interpretation.
Phonographic processing units in German
Sandra Beyermann
Universiteit Gent, Belgium

1. Main idea
The research objective is to shed light on the interactions of phonology and orthography during reading and reading acquisition in German. The main idea is to contrast relevant phonographic processing units that serve to recognize words at different stages of reading proficiency.

2. Reaction-Time Experiments
Two visual lexical decision experiments were designed to investigate whether levels of phonological representations (i.e. feet, syllable components) play a role in visual word recognition in German. Subject groups are beginners (1st to 2nd grade), advanced (3rd to 5th grade), and skilled readers (students).

(i) Stress Pattern:
One experiment was designed in order to find out whether printed words with canonical structures are easier to process than those with non-canonical structures, a factor that is known to affect spoken word processing (e.g. in English, cf. Kelly et al. 1998).
Stimuli examples (48 bisyllabic words):

a) canonical pattern (Samen "seed", Efeu "ivy")

b) non-canonical pattern (Rabatt "discount", Kamel "camel")

(ii) Syllable Position:
A second experiment investigates whether there are syllable components in the visual word (onset, nucleus, coda) that play a major/minor role in word recognition. Since there is a difference in auditory word recognition (a different saliency of syllable components) there may also be a difference on the orthographic level. A set of 48 German words was collected, with each word containing a manipulation of onset, nucleus or coda in the second syllable.
Stimuli examples (48 trisyllabic words):

a) 16 x wrong medial onsets (Franvose > Franzose),

b) 16 x wrong medial nuclei (Pulluver > Pullover),

c) 16 x wrong medial codas (Objebte > Objekte)

3. Expected Outcome
The expected outcome is that, the more skilled a reader is, the more he/she may be affected by phonological features (stress pattern, syllable components) of printed words.

References (deleted)
Theories are not the domain of experts alone. In much the same way as expert linguists, children are moved by the drive to explore and explain their own language abilities. Relying on insights from the different domains of Theory of Mind, metalanguage, and linguistic literacy, the current study investigates Hebrew-speaking children's evolving naïve theories about the linguistic nature of spelling. Spelling is not a mere scholastic skill, but rather an object of knowledge that must be conceptualized in developing linguistic literacy.

The current study focuses on violation of the role of morpho(phono)logy in Hebrew orthography, specifically on the distinction between root letters and function letters. The design is based on the universal human 'drive for explanation' which arises, in particular, when faced by unexpected events or apparently insoluble problems. 140 Hebrew-speaking pupils in six grade levels and 20 adults were presented with a crossword puzzle which can be easily solved by young children. However, two definitions resulted in a spelling error in function letters due to homophonous grapheme e.g., the word XUT ‘thread' lead to T ט rather than the required function letter T ת indicating adjective plural feminine for YAFOT ‘pretty'.

Participants carried out the task in dyads, to evoke more complex discussions towards problem solving than single-participant situations usually yield. We examined subjects' reactions to violations, and their search for an explanation of task failure.

Results indicate that participants of all ages refused to accept a spelling error in function letters. Spelling violations provoked reactions of discomfort, and the 'drive to explain' the linguistic problem even in first graders. Age and schooling, though, brought about an increase in the degree of explicitness of linguistic explanations and in children's ability to identify, discuss, and suggest a variety of creative solutions.

This study provides new insights about relationships between linguistic knowledge, writing systems and cognition from a previously unexplored perspective.
En Thaïlande, un seul alphabet a été imposé dans tout le pays au milieu du XIXe siècle. Il est l’aboutissement d’une longue tradition historique liée à l’enseignement du pâli-sanskrit dans les écoles bouddhistes. Malgré ou à cause de nombreux aménagements cumulatifs, il est actuellement particulièrement complexe. Par exemple, il compte 44 lettres consonnes pour noter seulement 21 consonnes phonologiques. Cette complexité rend cet alphabet particulièrement difficile à enseigner tant à un public de langue maternelle que de langue seconde. Beaucoup de Thaïlandais mémorisent les formes orthographiques plus qu’ils ne maîtrisent des règles d’écriture. Par exemple, la règle de la notation de la tonalité, particulièrement alambiquée, n’est connue que de quelques enseignants experts.

Dans notre communication, nous donnerons rapidement quelques exemples des principales difficultés de l’alphabet thaï mais l’essentiel de notre communication portera sur les tentatives actuelles de normaliser les procédures de translittération entre l’alphabet thaï et l’alphabet latin. Cela concerne la « romanisation », de l’alphabet thaï vers l’alphabet latin, nécessaire en particulier pour rendre lisible toute l’onomastique aux étrangers de passage dans le pays. Cela concerne aussi la notation et l’orthographe en thaï des emprunts faits aux langues européennes. L’Institut Royal de Thaïlande optant actuellement pour des procédures sur des principes phonétiques, les deux translittérations sont distinctes et non réfléctives ; il n’est pas possible de retrouver avec certitude le mot à l’origine d’un mot translittéré.

Les solutions officielles retenues sont très insatisfaisantes et posent de nombreux problèmes dans la vie quotidienne mais aussi dans l’enseignement du thaï aux étrangers et celui des langues européennes aux Thaïlandais. Toute proposition d’amélioration se heurte à des difficultés liées tant aux écarts entre les systèmes phonologiques et graphiques qu’aux représentations sociales sur les langues.
Cross-orthographic research shows that writing systems affect the ability to identify and manipulate linguistic units in learning of second/foreign languages. Thus, features of the foreign language (FL) learners’ first language influence their FL learning (see Cook, 2005).

In this paper, we show that the writing system of Korean violently influences the learning of English as a foreign language (EFL) in that it affects the recognition of syllabic structures of English words by the Korean learners of EFL.

Korean syllable is composed of a consonant (C), a vowel (V), and a consonant at most. This can be represented as CVC (such as kan ‘liver’), in comparison with English CCCVCCCC (e.g., sprinkles) and Japanese CV (e.g., to-kyo). Any extra consonant is phonologically inactivated in Korean. Thus kaps ‘price’ is pronounced as [kap], inactivating the final consonant s. If it is followed by a vowel as in kaps+i, then it is pronounced as [kap + si] instead of [kap i]. In the writing system of the Korean language, i.e., Hangeul, this syllable structure is robustly represented. That is, any extra consonant is not written except in the coda position that we have just seen above (kaps+i ‘the price is’). How does Hangeul represent English words that contain more than one consonants in both onset and coda positions? There is no way of representing this kind of multiple consonant clusters in the Korean writing system. Thus, English word strike is not correctly transcribed in Hangeul (*ㅅㅌ라이ㅋ).

Foreign words are adjusted to the Korean syllabic structure, by adding the neutral vowel /i/, so that the extra syllabic consonant can be incorporated into a (newly created) syllable. Thus strike is transcribed as 스트라이크 (“—” is a neutral vowel). Beginning learners of English generally use English vocabulary books that provide pronunciation information written in Hangeul (not in IPA). For these learners of English, the English word strike tends to be wrongly recognized as [sɔtʰɾaikə]. These learners may retain such wrong syllable structure quite long and apply the wrong syllabic structure to every possible case.
It has long been taken for granted that a unique characteristic of German orthography is the marking of nouns with an initial capital letter, hence, of a certain part of speech, cf. the German term *Substantivgroßschreibung*. However, in recent years the focus on this phenomenon has shifted and currently, the function of capitalization is rather considered as marking a syntactic position, namely the head of the NP, an analysis that appears evident in view of (1). At present, this syntactic approach seems to be consensus, at least from a strictly linguistic perspective.

(1) Nach längerem Überlegen über Für und Wider entschied sich ihr Ich trotz der […] Ah’s und Oh’s der sie Anhimmelnden gegen das kleine Schwarze, weil ihr nach diesem schrecklichen Gestern heute an gewaltigem Theaterdonnern gelegen war (Günther 2007:161)

Nevertheless, there is evidence that the question if a lexical or a syntactic approach is more appropriate to this phenomenon, can only be answered by demonstrating that both are relevant. In accordance with GALLMANN (1997), they shall be regarded as two concepts of nominality which operate cumulatively in a core area of capitalization, but differ in their prediction in contexts where they are not compatible with one another.

I will demonstrate that Optimality Theory (OT) offers an elegant solution to analyze this interaction in order to explain and predict the use of intrasentential capital letters. In doing so, it will become obvious that the syntactic constraint is the highest one in the ranking, which supports the currently dominating syntactic approach. I will also present how to come to terms with problematical cases like nominalised adjectives. Furthermore, I will discuss if the lexical influence on capitalization should be incorporated as a markedness or a faithfulness constraint. Concluding, I will show that there necessarily is a significant interaction between the use of space characters (cf. JACOBS 2005) and intrasentential capitalization.
The present study aims to explore differences between subordinated reformulation and subordinated creation process in oral and written story retellings produced by Spanish speaking children at the end of 2nd grade of primary education.

It is assumed that differences between oral and written modalities can be found due, in part, to the cognitive demands of low level writing skills. Indeed, it has been observed that written texts produced by children are shorter and of lower quality than oral ones (Berninger, Yates, Cartwright, Rutberg, Remy & Abbott, 1992; Berninger & Swanson 1994). However, it is not well established how the transcription skills could constrain the syntactic complexity of children written texts (Silva, Sánchez Abchi & Borzone, 2010).

One hundred and sixty-three children (Mean age: 7:11; range: 7:6 -8:11) participated in this study. The children attended three different schools from the province of Córdoba, Argentina. The relation between subordinated structures in the original text and those constructions written by children was considered in order to analyse the creation and reformulation processes.

The analysis showed different textual reformulation procedures: element’s omission, substitution or addition. These procedures affect diverse simple lexical units - article, prepositions-, and complex units like nouns, verbs and complements.

It was observed that the creation of subordinated process involves the articulation between two adjacent or distant statements. The semantic nature of this articulation (type and subtype of the subordinated construction) appears in the syntactic position of the prototypical connector and the position of the articulated units. In the subordinated creation, reformulation processes were observed as well.

In both, oral and written productions, children reformulate mostly sentences with noun dependent clauses from the original text and create adverbial and adjective embedded clauses. Moreover, differences between oral and written productions in relation with the kind of elements modified and the creation processes were observed.
Students’ Achievement on Impromptu Writing Test versus Interactive Writing Assessment through Feedback Exchange
Fakhrisadat Shariati
Islamic Azad University, Najafabad Branch, Iran

The ability to write effectively is becoming increasingly important in our global community, and instruction in writing is thus assuming an increasing role in both second and foreign language education and whenever the acquisition of a specific language skill is seen as important, it becomes equally important to test that skill and writing is no exception. There has been controversy as to whether the achievement of EFL learners would be better on interactive writing assessment or an impromptu writing test and relatively few studies have compared these two types of assessment. There is not much information about the attitudes of learners toward exchange of feedback when testing their writings as well. The current study intended to compare achievement of EFL learners on a traditional approach of writing test (impromptu) with a new assessment technique (interactive writing) focusing on two types of peer and teacher feedback aiming at searching for a more efficient assessment procedure with a high level of validity and reliability. It also considers the attitudes of learners to provide them with tasks and methods of assessment which are supposed to be more interesting. To this end, forty Iranian EFL learners who were all English teaching majors were randomly divided into two groups: group A sitting for an impromptu writing test and group B first provided with a reading task and then during writing their essays with teacher and peer feedback and finally, they received a questionnaire to gather data about their attitudes toward feedback exchange when assessing writing. The results of the statistical analysis (comparing two sets of essays using t-test) indicated that there was no significant difference between the achievement of students on impromptu writing test and interactive writing assessment through feedback exchange. However, learners had a positive attitude towards interactive writing assessment through peer/teacher feedback in comparison with writing test. The pedagogical implications of the study were also discussed.

Keywords: assessment, feedback, interactive writing assessment, timed impromptu writing test, writing assessment
The occurrence of homographs in alphabetic orthographies falls within the scope of description, and the decision to delete or keep homographs depends on whether they are considered harmful or harmless to the act of communication.

However, the Portuguese Language Orthographic Agreement of 1990 (under IV and IX, 9), establishing a spelling reform of the Portuguese language, increases the number of homographs, by deleting consonants with diacritical function and diacritical marks, such as acute accents, and decreases not only the degree of certainty in terms of pronunciation, but also semantic clarity and orthographic transparency. Two examples are the nouns *coacção* (*coercion*), from the verb "coagir" ("to coerce"), and *coação* (*filtering*), from the verb "coar" ("to filter"), which are rendered *coação* and *coação* by the terms of the Agreement, and the verbal flexion *pára* (*stops/stop*), 3rd person singular present indicative and 2nd person singular imperative of the verb "parar" ("to stop"), that now become homographs of the preposition *para* (*to*). In the latter, there is also an issue of grammatical equivalence.

Considering that these homographs are laid out in a legally binding resolution, they have become part of a process of conception and prescription of homographic forms. Since the orthographic depth of alphabetic orthographies refers to consistency of grapheme-phoneme correspondences, the degree of transparency in Portuguese orthography decreases, considering that one-to-one correspondences are obliterated, as exemplified by the aforementioned cases.

**Key words:** ambiguity, homographs, opacity, orthography, transparency.