

Programme
Understanding writing systems:
From core issues to implications for written language acquisition
Tenth International Workshop on Writing Systems and Literacy
Sponsored by Radboud University



May 12th-13th 2016

Huize Heyendael, Geert Grootplein-Noord 9, 6525 EZ Nijmegen

Scientific committee:

Lynne Cahill, Terry Joyce, Martin Neef, Anneke Neijt, Mijntje Peters

Local organizers:

Anneke Neijt (a.neijt@let.ru.nl),
Mijntje Peters (m.peters@let.ru.nl)
Merijn Beeksma (merijn.beeksma@student.ru.nl)

(Although not included in the workshop registration fee, for those arriving on May 11, please consider joining us for dinner or drinks in Hotel Atlanta's Grand Café from 19.30 (plate service 10 euro on Wednesdays). We will make reservations, if you advise your intention to join. Grand Café Hotel Atlanta, Grote Markt 38, 6511 Nijmegen)

Thursday, May 12

9.00 Registration

9.30 Welcome and introduction

9.45 Invited: **Learning to spell in Dutch: A case in point**

Anna M.T. Bosman, Behavioral Science Institute, Radboud University Nijmegen

10.15 **Foundations of a universal theory of learning to read**

David L. Share, University of Haifa

10.45 Coffee

11.15 **Bi-alphabetism: challenges imposed by visual, phonological and executive processing load and the implications for models of reading**

Dušica Filipović Đurđević^{1,2}, Petar Milin^{1,3}, Isidora Gatarić¹, Laure Beth Feldman⁴; ¹Department of Psychology or Laboratory for Experimental Psychology, Faculty of Philosophy, University of Novi Sad, ²Laboratory for Experimental Psychology, Faculty of Philosophy, University of Belgrade; ³Department of Computational Linguistics, Eberhard Karls Universität Tübingen; ⁴Psychology Department, University at Albany, State University of New York & Haskins Laboratories, New Haven, Connecticut

11.45 The effects of writing system and language experience on the acquisition of literacy: A 2-year study of monolinguals and early sequential bilinguals learning to read Hebrew and Arabic

Zohar Eviatar¹, Mila Schwartz², Haitham Taha³, ¹Psychology Dept. & IIPDM, ²IIPDM, University of Haifa, & Academic College Oranim, ³Sakhnin College of Education, Israel

12.15 Oral poster presentations

Tendencies in parafoveal alphabetic letter identification in Russian

S. Alexeeva & A. Konina, Saint Petersburg State University, St. Petersburg, Russia

Learning to read a Semitic Abjad: The Triplex model of Hebrew reading development

Amalia Bar-On¹ & David L. Share², ¹University of Tel-Aviv, ²University of Haifa

Assessing spelling skills related to morphosyntax of German-French biliterate pupils in a multilingual educational context

Natalia Bilici¹, Sonja Ugen¹, Reinold Funke², Michel Fayol³ & Constanze Weth¹, ¹Université du Luxembourg (LU), ²Pädagogische Hochschule Heidelberg (D), ³Université Blaise Pascal, Clermont-Ferrand (F)

How does the complexity of the French writing system impact its learning in first-grade?

Catherine Brissaud¹, Corinne Totereau¹ & Laurence Pasa², ¹University Grenoble Alpes, ²University Toulouse Jean Jaurès

The rise and rise of the orthographic kiss (in UK CMC)

Lynne Cahill^{1,2} & Annaliese Bagley¹, ¹University of Sussex, UK, ²now at University of Oxford

A longitudinal study of verbal and visuospatial working memory and emergent writing in English

Kathleen Carroll & Fiona Lyddy, Maynooth University, Ireland

Rethinking logography in Japanese kanji writing

Keisuke Honda, Imperial College London

Revealing the transposed-letter effects in monolingual and bilingual reading minds using mouse tracking

Yu-Cheng Lin¹ & Pei-Ying Lin², ¹Department of Psychological Science, University of Texas Rio Grande Valley, USA, ²Department of Educational Psychology & Special Education, University of Saskatchewan, Canada

Comparative Studies of Orthography Development in Indonesia Alphabet Design Workshop vs Korean Hangeul Development

N. Nazarudin, Leiden University

Digitally created script in the Japanese writing system: Observations of computer-mediated discourse from sociolinguistics perspectives

Yukiko Nishimura, Faculty of Global Communications, Toyo Gakuen University, Tokyo, Japan

Challenges in the development of written corpus of adult speakers

Marina Olujić, Jelena Kuvac Kraljević, Gordana Hrzica, University of Zagreb, Laboratory for Psycholinguistic Research

Grammar first, for children who learn to write Dutch verbs?

Mijntje Peters, Johan Zuidema, Anna Bosman & Anneke Neijt, Centre for Language Studies and BCN, Radboud University Nijmegen

Non-orthographic spaces in German and Scandinavian. An OT analysis

Vilma Symanczyk Joppe, Bergische Universität Wuppertal, Germany

Negative L1 Orthographic Processing Transfer on L2 Literacy

Emina Tuzovic & Marjorie Lorch, Applied Linguistics & Communication, Birkbeck, University of London

Principle analysis for orthography development

Mira Valkama, University of Helsinki

Double Dutch Vowel Spelling

Johan Zuidema & Anneke Neijt, Centre for Language Studies, Radboud University Nijmegen

12.30 Lunch

13.30 Invited: **From phonemic spelling to distinctive spelling**

Geoffrey Sampson, University of South Africa & emeritus Sussex University

14.30 **The pros and cons of IPA-based writing systems**

Martin Neef, TU Braunschweig

15.00 Coffee & posters

16.00 **Writing morphophonology, reading lexical tone: evidence in favor of morphographic spelling in Kabiye (Togo)**

Dave Roberts & Steve Walter, www.nyinyu.com, blog: cornishmaninafrica.wordpress.com

16.30 **Principles of English spelling: from Beowulf 2 bootylicious**

Des Ryan, Centre for Language & Communication Studies, Trinity College, Dublin

17.00 **Beyond Conventional Writing Boundaries: Non-Standard Language Practices in Dutch Youths' Computer-Mediated Communication**

Lieke Verheijen, Centre for Language Studies, Radboud University Nijmegen

17.30 AWLL business meeting

19.00 Dinner in town:

Restaurant Het Savarijn, Van der Brugghenstraat 14, 6511 SL Nijmegen

Friday, May 13

9.30 Invited: **The Digital Literacy Instructor (DigLin)**

Ineke van de Craats, Center for Language Studies, Radboud University

10.15 **Language segmentation practices result from literacy: a cross-linguistic study**

Dorina Veldhuis, Hogeschool iPabo, vakgroep Nederlands

10.45 Coffee

11.00 **Among Phoenicians: a comparison of learning and teaching an alphabetic and syllabic writing system**

Yonas Asfaha¹, Jeanne Kurvers² & Sjaak Kroon², ¹University of Asmara, ²Tilburg University

11.30 **Core syllables vs. moraic writing**

Gene Buckley, U. Pennsylvania

12.00 **Just mixed up or a pretty neat idea? Some reflections on the multi-script nature of the Japanese writing system**

Terry Joyce¹ & Hisashi Masuda², ¹Tama University, Japan, ²Hiroshima Shudo University, Japan

12.30 Lunch

13.30 **Naturalness of scripts and writing systems Prolegomena to a Natural Grapholinguistics**

Dimitrios Meletis, Department of Linguistics, University of Graz

14.00 **Orthographic and Visual Clues to Metalinguistic Knowledge, Script Acquisition, and the Cognition of Reading and Writing of Ancient Mayan Scribes**

David F. Mora-Marín, University of North Carolina at Chapel Hill

14.30 **How does writing something down change one's (mental) representation of it?**

David Olson, emeritus University of Toronto, Canada

15.00 Panel discussion: **orthographic data basis, typology of writing systems**

16.00 Conclusion & Drinks

Abstracts

Understanding writing systems: From core issues to implications for written language acquisition

Tenth International Workshop on Writing Systems and Literacy

Sponsored by Radboud University

May 12th-13th 2016

Huize Heyendael, Geert Grooteplein-Noord 9, 6525 EZ Nijmegen



Scientific committee: Lynne Cahill, Terry Joyce, Martin Neef, Anneke Neijt, Mijntje Peters

Local organizers: Anneke Neijt (a.neijt@let.ru.nl), Mijntje Peters (m.peters@let.ru.nl), Merijn Beeksmā (merijn.beeksmā@student.ru.nl)

May 12, Morning session

9.45 Invited: **Learning to spell in Dutch: A case in point**

Anna M.T. Bosman, Behavioral Science Institute, Radboud University Nijmegen

Learning to spell depends on the characteristics of the speller, the orthography, but above all on the instruction skills of the teacher. I will provide an overview of each of these elements concerning the teaching of Dutch spelling. Spellers differ in their ability to acquire the spelling of their native orthography. Alphabetic languages share the attempt to somehow provide a more or less consistent relationship between letters and sound. Each written language, however, has its specifics; some orthographies are more transparent than others. Transparent orthographies are easier to learn than opaque written language. Although both these aspects play a part in the acquisition of spelling knowledge, I will defend that teachers' skills fundamentally determine the outcome of the learning-to-spell process. Knowledge of the orthography and direct instruction skills provide the necessary components for successful didactics.

10.15 **Foundations of a universal theory of learning to read**

David L. Share, University of Haifa

Spoken language unites the human species, dividing humans from animals, but it is *written* language and the literacy it affords, that divides humanity; separating cultures, communities and

individuals from one another. The foundations of literacy are the basic reading and writing skills taught on entry to formal schooling around the world. Literacy learning is much more than the foundational word reading skills discussed here, but it is nothing without them.

The scientific study of reading has made important strides over the past several decades but is still entrenched in Anglocentric and Eurocentric/alphabeto-centric theoretical frameworks (Share, 2008; Share, 2014). This presentation offers some general guidelines for constructing a universal theory of learning to read, one that seeks universals yet embraces the enormous diversity among the world's languages and writing systems.

I claim that there exists a fundamental and universal dualism in printed word learning that applies to all words in all possible orthographies (Share, 2008). Because *every* printed word is, at one point, unfamiliar, the reader must possess some means of independently identifying units of meaning (words and morphemes) encountered for the first time. I propose that this is true for every orthography; alphabets, abjads, akshara-based scripts, syllabaries and morpho-syllabaries. In addition, the reader must eventually achieve a high degree of unitization or "chunking" either of letter strings, aksharas, stroke patterns or character compounds to enable rapid, holistic/parallel, and near-effortless recognition of familiar words and morphemes (Anderson et al. 2013, Kwok & Ellis 2015; LaBerge & Samuels 1974; Logan, 1988; Perfetti, 1985). And because separate morphemes necessarily have distinct visual forms, each must be individually learned as an unique visual configuration of the limited set of letters (in phonemic scripts), aksharas in Brahmi-derived Indic-based scripts or character combinations in morphosyllabaries (Shu et al., 2003).

In order to cater to the reader's needs, an efficient writing system must have both *decipherability/learnability* (via phonological transparency) and *unitizability/automatizability* (via morphemic transparency). A writing system, like spoken language, must therefore be "combinatorial" (Hockett, 1960), combining and recombining a limited, and hence learnable number of sub-lexical elements to generate an unlimited vocabulary. I outline how this decipherability/automatizability dualism plays out in the five main varieties of writing system drawing out its graphonomic/linguistic, psychological and pedagogical implications.

10.45 Coffee

11.15 Bi-alphabetism: challenges imposed by visual, phonological and executive processing load and the implications for models of reading

Dušica Filipović Đurđević^{1,2}, Petar Milin^{1,3}, Isidora Gatarić¹, Laure Beth Feldman⁴, ¹Department of Psychology or Laboratory for Experimental Psychology, Faculty of Philosophy, University of Novi Sad, ²Laboratory for Experimental Psychology, Faculty of Philosophy, University of Belgrade; ³Department of Computational Linguistics, Eberhard Karls Universität Tübingen; ⁴Psychology Department, University at Albany, State University of New York & Haskins Laboratories, New Haven, Connecticut

Native speakers of the Serbian language in the former Yugoslavia use two alphabets (Cyrillic and Roman) and word recognition studies based on manipulations of letters from the two alphabets have played a central role in understanding the interrelation of phonology and semantics in skilled reading. Starting more than twenty years ago (e.g., Feldman & Turvey, 1983) a literature has accrued showing that Serbian words composed only of letters that exist in both the Roman and Cyrillic alphabets are slower in a lexical decision task than are the unique alphabet transcriptions of those same words but only when some of those letters have different phonemic interpretations in each. For example latencies in a lexical decision task are slower and less

accurate to phonologically ambiguous targets (e.g., PETAK meaning “Friday” when pronounced as a Roman string /petak/ but readable without meaning when pronounced as Cyrillic as /retak/) than to the unique alphabet transcription of the same word ПЕТАК. The PETAK – ПЕТАК comparison (the ambiguity effect) is particularly compelling evidence of a phonological contribution to skilled reading because it contrasts two alphabetic transcriptions of the same word; therefore all of its lexical and semantic properties (e.g., frequency, letter length, meaning) are identical. More recent studies have considered the processing cost incurred by switching between alphabet codes (Filipović Đurđević, Milin & Feldman, 2013) and its analogs to bilingualism

We will present an overview of the work on bialphabeticism including the results of ongoing experiments that track the effects of visual, phonological and executive processing load on the detriment incurred by phonologically ambiguous words relative to their unique alphabet controls and will discuss the role of working memory components in the process of code switching while reading..

Feldman, L. B., and Turvey, M. T. (1983). Word recognition in Serbo-Croatian is phonologically analytic. *Journal of Experimental Psychology: Human Perception and Performance*, 9, 288-298.

Filipović Đurđević, D., Milin, P., and Feldman, L. B., (2013). Bi-alphabeticism: A window on phonological processing. *Psihologija*. 46(4), 421-438.

11.45 The effects of writing system and language experience on the acquisition of literacy: A 2-year study of monolinguals and early sequential bilinguals learning to read Hebrew and Arabic

Zohar Eviatar¹, Mila Schwartz², Haitham Taha³, ¹ Psychology Dept. & IIPDM, ² IIPDM, University of Haifa, & Academic College Oranim, ³ Sakhnin College of Education, Israel

We know that many factors affect the acquisition of literacy, among them the characteristics of the writing system and the linguistic experience of the learner. We examined the differences in morphological and phonological awareness in kindergarten in Hebrew and Arabic speaking monolinguals, and children enrolled in bilingual schools, who had either Hebrew or Arabic as their home language. In 1st grade, the children began to acquire literacy in their native language. Although typologically, Hebrew and Arabic are sister languages, and although both use an abjad writing system, there are important differences in the socio-linguistic and the orthographic domains: Arabic speakers are born into a diglossic situation, where the language they learn to read is different from the language they speak, and Arabic orthography is visually more complex than Hebrew orthography. We tested the effects of being bilingual on morphological and phonological awareness tests in kindergarten and 1st grade, and the relationship of these components of reading with measures of literacy: reading and writing single words. The results revealed that in both grades, the bilinguals achieved higher scores on tests of phonological and morphological awareness than the monolinguals, in both languages. In first grade, bilinguals were significantly better at spelling, but had significantly smaller vocabularies than monolinguals. The relationships between early literacy and phonological and morphological abilities in kindergarten were different in the two languages. We show that different metalinguistic abilities in kindergarten predict spelling in the two languages. In addition, we show together higher correlations between basic abilities with reading in Hebrew than with reading in Arabic. This finding converges with others that have shown that metalinguistic abilities in kindergarten are more strongly related to measures of reading in Hebrew than in Arabic. We present hypotheses about the sources and mechanisms of these differences.

12.15 Oral poster presentations

Tendencies in parafoveal alphabetic letter identification in Russian.

S. Alexeeva & A. Konina, Saint Petersburg State University, St. Petersburg, Russia

Numerous attempts have been taken in order to understand the mechanisms through which language speakers discriminate letters. However, those studies were based on Latin alphabet; when it comes to Russian, we find them practically nonexistent.

Eye-tracking presents with a valid instrument to research and measure those processes. We used Eyelink 1000+ with 1000 Hz record rate, the display refresh rate was set to 120 Hz. 24 volunteers aged from 18 to 34 participated in the pilot study.

After the standard procedures of instruction and calibration, participants see the fixation cross followed with a masked letter (*a*, 33 lower-case letters of the Russian alphabet in monospaced Courier New 18) appearing three visual angles left or right of the center of the screen. Letter delays are controlled by gaze contingent boundary paradigm: the masked letter disappears during the saccade to the stimulus. Such paradigm allows us to simulate parafoveal processing (when eyes are fixating the preceding area). The participant's oral report on which letter did they see is then recorded.

Mixed effects logistic regression showed that certain letters are identified significantly better than others ($p < 0.0001$): p, y, л, й, о, ф, е, с, ж, к, ш, и, ц, а (in decreasing order: precision for p is 92% and for a is 57%), and considerably worse than others: в, б, э, т, п, з, г, я, л, ч, ы, ъ, м (in decreasing order: precision for в is 34%, and for м is 6%). We tend to think that aside from ascenders/descenders round shape of the letter and diagonal elements contribute much to the faster letter identification. Among factors deteriorating identification speed we could list the inclusion of one letter inside the other (for instance, ъ is part of ы).

The project is supported by RSF #14-18-02135.

Learning to read a Semitic Abjad: The Triplex model of Hebrew reading development

Amalia Bar-On¹ & David L. Share², ¹University of Tel-Aviv, ²University of Haifa

In this presentation, we introduce a model of Hebrew reading development that emphasizes both the universal and script-specific aspects of learning to read a Semitic abjad. At the universal level, the study of Hebrew reading acquisition offers valuable insights into the fundamental dilemmas of all writing systems - the unfamiliar-tofamiliar/novice-to-expert dualism involves a shift from slow, effortful, piece-by-piece unskilled performance to rapid, one-step, "chunked" or large-unit skilled performance (Share, 2008). This dualism also converges with the dualistic nature of an efficient orthography (Rogers, 1995) which must serve the needs of the novice for decipherability (via phonological transparency) as well as the expert-to-be for unitizability and automatizability (via morphemic transparency). At the script-specific level, Hebrew performs this balancing act with aplomb, employing supplementary vowel signs (the pointed script) offering the beginner a consistent, phonologically well-specified script, while helping the expert-to-be unitize words and morphemes via (morpho-orthographic) spelling constancy. However, a 2-dimensional dualistic approach focused exclusively on individual words falls short of providing a complete picture of Hebrew reading acquisition owing to the fact that around one third of the words in unpointed Hebrew are homographic. This poses a second major challenge to the developing reader making

the transition from the pointed to the unpointed Hebrew. Our Triplex model, emphasizes three phases of early Hebrew reading development; a progression from lower-order, phonological (sub-lexical) sequential spelling-to-sound translation (Phase 1, Grade 1) to higher-order string-level (lexical) lexico-morpho-orthographic processing (Phase 2, Grade 2) followed by, in the upper elementary grades, a supra-lexical contextual level essential for dealing with the pervasive homography of unpointed script.

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of over-reliance on an “outlier” orthography. *Psychological Bulletin*, 134, 584-616.

Rogers, H. 1995. Optimal orthographies. In Insup Taylor & David R. Olson (eds.), *Scripts and literacy: Reading and learning to read alphabets, syllabaries, and characters*, 31–44. New York: Kluwer.

Assessing spelling skills related to morphosyntax of German-French biliterate pupils in a multilingual educational context

Natalia Bilici¹, Sonja Ugen¹, Reinold Funke², Michel Fayol³ & Constanze Weth¹, ¹Université du Luxembourg (LU), ²Pädagogische Hochschule Heidelberg (D), ³Université Blaise Pascal, Clermont-Ferrand (F)

How do multilingual learners write spellings related to morphosyntactic information in German (capitalisation of nouns) and French (plural markers of nouns, adjectives and verbs)? Our talk presents the construction and the items of a spelling test of German and French for multilingual 5th graders in Luxembourg (N=300) as well as the first results of both tests. It is the pre-test of a study that will analyse the processes of morphosyntactic agreement in spelling in the children’s first (German) and second (French) acquired language at school. It will further examine how their performances relate to their language background. The setting in Luxembourg is characterised by three school languages: While Luxembourgish is the language of pre-school, children are alphabetised in German and learn French from second grade on. German and French are second languages for most children. However, children have either a Germanic (Luxembourgish) or a Romance (Portuguese) background. The children performed spellings tests tailored to the specificities of each of the test languages but relying on comparable grammatical processes. The test framework will be presented in detail as well as first results. Based on the literature, we expect differences in performances for proper and abstract nouns as well as nominalisation with best performance for proper nouns. In French, we expect differences in performances according to the plural markers of nouns, adjectives and verbs with best performance for nouns. Additionally, we expect contextual effects of the target words within each language. Results will further be analysed according to the background language of the children.

How does the complexity of the French writing system impact its learning in first-grade?

Catherine Brissaud¹, Corinne Totereau¹ & Laurence Pasa², ¹University Grenoble Alpes, ²University Toulouse Jean Jaurès

French is a roman language but it has evolved a lot: its writing system looks quite different from the Latin one. The original Latin letters were not enough to transcribe French new vocal sounds: though diacritics and complex graphemes (for instance ai, ch, in, etc.) were used, French writing system is far from a bi-univoque grapheme-phoneme mapping relation (in, ain, ein, un can be read as /ẽ/). Furthermore, its complexity comes from final silent letters (elephant), especially the ones with a grammatical signification (-s meaning plural on nouns; nt meaning plural on verbs), several forms being pronounced in the same way though written in different ways (tu parles / il parle / elles parlent pronounced /paRl/). It is described as a mixed system with a significant semiographic principle at work (Coulmas, 2003; Fayol & Jaffré, 2008) or a high degree of

graphematic transparency but a low degree of orthographic transparency (Neef, 2011). In other words, French is easier to be read than to be written. The purpose of the communication is to shed light on this writing complexity by data from acquisition. Three words were dictated at the beginning of the school-year to 1169 students (rat, éléphant, lapin) and a sentence (Tom joue avec le rat 'Tom is playing with the rat'). They were dictated again to the same children at the end of the year with an added sentence: Les lapins courent vite ('Rabbits run fast'). These data were collected in 2013-2014 within the framework of the national research coordinated by Roland Goigoux (Goigoux et al., 2015). In the presentation of the results, the accent will be put on segmentation, recoding, and silent agreement marks. We shall highlight various levels of conceptualization (Ferreiro, 2000), which should allow to better understand the beginning of learning of the writing system.

Coulmas, F. (2003). *Writing Systems: An introduction to their Linguistic Analysis*. Cambridge, UK: Cambridge University Press. Ferreiro, E. (2000). *L'écriture avant la lettre*. Paris : Hachette Education. Goigoux R., Jarlégan A., Piquée, C. (2015). « Évaluer l'influence des pratiques d'enseignement du lire-écrire sur les apprentissages des élèves : Enjeux et choix méthodologiques ». *Recherches en didactique*, n° 19, p. 33-55. Fayol, M. et Jaffré, J.-P. (2008). *Orthographe*. Paris : Presses Universitaires de France. Neef, M. & Balestra, M. (2011). Measuring graphematic transparency. *German and Italian compared, Written language and literacy*, 14:1, 109-142.

The rise and rise of the orthographic kiss (in UK CMC)

Lynne Cahill^{1,2} & Annaliese Bagley¹, ¹University of Sussex, UK, ²now at University of Oxford

The increased use of various forms of CMC and developments in the technology used has led to rapid changes in the way language is used in CMC. One of the areas where this form of communication appears to diverge from other forms of written language is the use of a variety of devices to express non-linguistic aspects of communication such as emoticons and emojis.

In 2012, we presented the results of a study into the IM language of teenagers (15-16) and undergraduates (21-23). One of the key findings was the emergence of an apparently novel use of the orthographic kiss (x) by the younger group in places other than at the end of a complete conversation, i.e. at the end of every turn and in some instances even within a turn. We termed these "medial kisses". At the time, we speculated that this use appeared to be an alternative to the use of emoticons (this was before the widespread emergence of the emoji).

Two years later a further study was undertaken into the use of orthographic kisses. This second study was a much broader investigation with the purpose of not only discovering whether the use of medial kisses had continued, increased or indeed spread to other forms of communication, but also of gauging the attitudes to the use of orthographic kisses more generally.

An online survey was launched and a call to participate was spread via email and social media. The questionnaire asked multiple choice questions about how and when participants used kisses and every question also had a slot for further comments. A total of 567 responses were received. The age range of participants was 15-71, although many more responses were received from people in the younger age groups. Around two thirds were female. The majority of responses were from English-speaking UK residents.

The findings show that all forms of CMC, including email, SMS messaging and instant messaging, elicit the use of orthographic kisses and 95% of respondents said that they use them in some form of CMC to some people. Over two thirds of respondents said that they had sent a follow-up message with kisses that they had forgotten to include. Around a third said that they had apologised for sending a kiss, or too many kisses, in a message. Many respondents, in the further comments, referred to a sense of the etiquette of sending kisses, and a range of opinions, some

quite extreme, were revealed. The sheer quantity of additional comments, totalling over 30,000 words, is an indication of the strength of feeling about kisses.

In this paper we present the findings of the survey and discuss the implications. We also speculate on whether the recent rise in the use of emojis means that the use of orthographic kisses may have been a phase which is going to come to an end at some point in the future and to what extent the findings of our largely UK-based survey have any international validity.

A longitudinal study of verbal and visuospatial working memory and emergent writing in English

Kathleen Carroll & Fiona Lyddy, Maynooth University, Ireland

The importance of working memory for early reading development has been studied extensively. Less is known about the role of working memory in early writing. The current study examined the relationship between verbal and visuo-spatial short term and working memory and the mechanical and creative aspects of early writing skills. Thirty children (12 boys, 18 girls) participated in a three year longitudinal study, beginning during their first year of schooling (aged 4-5 years). Over four time points, verbal and visuo-spatial short term and working memory functions were measured using the Automated Working Memory Assessment (Alloway et al. 2007). Knowledge of letters, letter sounds, and morphological generation (as an index of grammar) were assessed and children provided a sample of free writing at each testing session over the three year period. Children also indicated their liking of writing at each stage. Writing samples were assessed for mechanical aspects of writing (letters, sentences, word formation, spellings, punctuation etc.), but also for creative aspects of the writing task: originality, detail, and relevance to the title of story as given.

Performance on the verbal short term and working memory tasks was associated with better letter and morphological knowledge, and with the originality of the stories produced by Time 4, but there was no correlation between verbal working memory and the number of words produced or spelling ability. Early letter and grammar knowledge predicted emergent spelling ability. There were no gender differences in the number of words or sentences produced, the proportion of correct spellings or the originality of the stories, but self-reported liking of writing increased for girls and decreased for boys over the period of the study. This pattern may have implications for later motivation to write.

Rethinking logography in Japanese kanji writing

Keisuke Honda, Imperial College London

One of the core issues in writing system research is the interpretation of logography. There are different views on what logography is, how it is related to phonography, and whether the division of logography and phonography is appropriate in the first place. As a result, there has never been a true consensus on how different types of writing systems represent language. This issue has direct implications for our understanding of Japanese kanji writing, that is, the use of kanji graphs to write words in Japanese. Although there seems to be a general agreement that kanji writing is something other than pure phonography, it is possible to distinguish three different views on how kanji graphs work: (1) Kanji are logographic or, more precisely, morphographic symbols representing individual morphemes; (2) Kanji are morphophonetic symbols representing sounds which may or may not correspond to individual words and/or morphemes; and (3) Kanji are logographic symbols, not in the sense that they directly represent words or morphemes, but in

that they represent particular sounds on a word-by-word basis. While each of these views has strengths and weaknesses, taking the third position allows us to explain facts that are otherwise difficult to account for. In this paper, I will demonstrate this point using examples of kanji writing involving diachronic changes (e.g., 不便 > 不憫 fubin ‘pitiful’), synchronic variations (e.g., 的中 ~ 適中 tekichū ‘hitting the mark’), and multi-graph representation of words lacking compositionality (e.g., 勉強 benkyō ‘study’). Based on these examples, I will sketch out a framework in which kanji are viewed as logophonetic symbols, which represent sounds and whose distribution is lexically conditioned.

Revealing the transposed-letter effects in monolingual and bilingual reading minds using mouse tracking

Yu-Cheng Lin¹ & Pei-Ying Lin², ¹Department of Psychological Science, University of Texas Rio Grande Valley, USA, ²Department of Educational Psychology & Special Education, University of Saskatchewan, Canada

Previous monolingual studies have consistently suggested that there was flexibility of letter position encoding in different alphabetic writing systems. However, to date whether the orthographic neighborhood and cross-language script similarity would modulate the magnitude of the transposed-letter effect during the second-language word recognition in bilingual minds was unknown. We address this question using a mousetracking experimental paradigm to trace the internal lexical matching processes underlying the lexical access. Our linear mixed effects models and growth curve analyses revealed that a low orthographic neighborhood can trigger a larger magnitude of the transposed-letter effect for monolinguals and bilinguals on their hand trajectories. We also found that different-script bilinguals (Chinese-English bilinguals) exhibited a larger transposed-letter effect than similar-script bilinguals (Spanish-English bilinguals) and English monolinguals. The findings offer compelling evidence that a human lexical match criterion of recognition system can be modified by neighborhood density and cross-language script similarity of readers.

Comparative Studies of Orthography Development in Indonesia Alphabet Design Workshop vs Korean Hangeul Development

N. Nazarudin, Leiden University

Developing orthography for a minority language requires a writing system that is embraced by members of the language community. This study investigates two different orthography projects that have been conducted in Indonesia: the Alphabet Design Workshop (ADW) for Woirata (Papuan, Timor-Alor-Pantar language family with >800 speaker) on Kisar Island, Southwest Maluku and the adoption of the Korean Alphabet (Hangeul) in Ciacia community (Austronesian language family with >60.000 speaker) in Baubau District on Buton Island, southeastern part of Sulawesi. ADW, which was developed by Summer Institute of Linguistics (SIL), has been used to develop more than 100 language/dialect groups in the last decade. Since, there was no orthography development program done by Indonesian authority yet, the ADW were chosen because of its methodological simplicity during Papuan language documentation project held by LIPI (Indonesian Institute of Science) in 2013 and 2014. On the contrary, the adoption of Hangeul was first initiated by the Hungmin Jeongeum Society, a scholarly association consisting of Korean linguists based in Seoul, South Korea. This project also reflects the effect of the local development policy on Baubau and also reflects an implicit criticism of the central government's

development, which was focused exclusively on Java and a desire for regional self-determination, particularly with regards to the language policy. (Song 2013) Apart from explaining the role of social and political factor behind these orthography development activities, this study also focusing on phonological comparison from these two cases of orthography development.

Berg, Rene van Den, 1991, Preliminary notes on the Cia-Cia language (South Buton) in Poeze, Harry A. and Pim Schoorl (Ed), *Excursies in Celebes*, Leiden: KITLV Uitgeverij, p. 305 – 325.

De Josselin de Jong, J.P.B. 1937. *Studies in Indonesian Culture*[.] *Woirata. A Timorese Settlement on Kisar*. Amsterdam: Uitgave van de N.V. Noord-Hollandsche Uitgevers-Maatschappij.

Seung-won Song, (2013), “Being Korean in Buton? The Cia-Cia's Adoption of the Korean Alphabet and Identity Politics in Decentralised Indonesia” *KEMANUSIAAN* Vol. 20, No. 1, University Kebangsaan Malaysia, (p. 51–80)

Soewarsono, L. Masnun, dan Nazarudin, 2013, *Revitalisasi Budaya dan Bahasa Oirata di Pulau Kisar, Maluku Barat Daya*, Maluku, Jakarta: LIPI dan PT Gading Inti Prima.

Digitally created script in the Japanese writing system: Observations of computer-mediated discourse from sociolinguistics perspectives

Yukiko Nishimura, Faculty of Global Communications, Toyo Gakuen University, Tokyo, Japan

More people are using written language to communicate than ever before (Coulmas 2013), such as in emails, blogs, and social network sites. Interactions in these informal, casual settings have been reported to exhibit features that are outside of orthographic conventions (e.g. Androutsopoulos 2000) and Japanese is no exception (Nishimura 2003, Tranter 2008). This study attempts to explore what these writing practices on the digital media can tell us about the Japanese writing system, based on blog posts linked to Japan Blog Village, Japan’s largest blog linking, ranking and aggregation site (<<http://www.blogmura.com/>>), and online data sets contained in the Balanced Corpus of Contemporary Written Japanese (BCCWJ) Version 1.1 compiled by the National Institute for Japanese Language and Linguistics (2015). Users seem to be motivated by narrowing the gap between the language they consider they employ when speaking and the written language that has been taught under the guidelines outlined by Kokugo Shingikai (the National Language Council) for the maintenance of literacy standards. At the same time they also perform their identity online by utilising linguistic resources that are beyond what the Guidelines state, such as smaller graphs for vowel kana as in *ありがとう* “aRIGATOu,” in which the graph size for “a” and “u” is unconventional. The study suggests that while the Japanese writing system with at least three different script types of morphographic kanji, syllabic katakana and hiragana, and alphabetic romaji, allow users flexibility in script choices, this flexibility is enhanced in digital platform, even to create new ones. If viewed in a longer time span, just as kana evolved out of kanji in the history of Japanese writing, to occupy an important place in the Japanese writing system, creative orthographic variation online, though peripheral now, could be a potential force in language change (Baron 1984).

Androutsopoulos, J. 2000 Non-standard spellings in media texts: the case of German fanzines. *Journal of Sociolinguistics* 4:4, 514-533.

Baron, N. S. 1984. Computer mediated communication as a force in language change. *Visible Language*. 18:2. 118-141.

Coulmas, F. 2013 *Writing and society : an introduction*. Cambridge : Cambridge University Press

Nishimura, Y. 2003. Linguistic innovations and interactional features of casual online communication in Japanese. *Journal of Computer-Mediated Communication*, 9: 0. Retrieved July 12 2015 from

<http://onlinelibrary.wiley.com/doi/10.1111/j.1083-6101.2003.tb00356.x/pdf>

Tranter, N. 2008. Non-conventional script choice in Japan. *International Journal of the Sociology of Language* 192: 133-151.

Challenges in the development of written corpus of adult speakers

Marina Olujic, Jelena Kuvac Kraljevic, Gordana Hrzica, University of Zagreb, Laboratory for Psycholinguistic Research

In two projects *Adult language processing (CSF – ALP-2421)* and *Computer assistant supporting text input for individuals with language disorders (EU – Structural fund; RC.2.2.08-050)*, being carried out by the group of experts in linguistics, language pathology and computer sciences, the final goal is to create the computer application that will support individuals with language impairments in writing process. To achieve this goal, it was necessary to collect written samples of children and adults, both with and without language impairments, and to develop the Croatian corpus of written language (CCOWL). The CCOWL presents a comprehensive database for further research and it is intended to be an online available database. The CCOWL consists of approximately 5600 samples (texts) written by 401 participants in the age range from 11 to 89+ (gender and education level controlled, as well). Moreover, the participants differ in terms of the existence of language impairment; 134 of them are healthy, 91 have aphasia (acquired impairment) and 176 have dyslexia (developmental impairment). Written samples differ in the level of text structures and the writing media (pen and keyboard).

While designing a plan on how to create a comprehensive corpus suitable for contemporary research and while collecting the written samples, a lot of questions came up: which participants to include (considering age, gender, education, etc.); how to collect written samples, by handwriting or typing; how to create tasks for different text structure levels; how to adapt these tasks to children, etc. The main aim of this paper is to present and discuss the greatest challenges in creating CCOWL since here presented challenges are prerequisites for development and analysis of written corpora in general

An orthographic system for Cypriot Greek: from graphemics to graphetics

Aspasia Papadima¹ & Stelios Kyriacou², ¹Language & Graphic Communication Research Lab, Cyprus University of Technology; ²Cyprus University of Technology

The paper presents the results of the research programme “Linguistics/Cypriot Dialect/Typographic rendering of the Cypriot dialect”, under the Language and Graphic Communication Research Lab (LGCRL), part of the Department of Multimedia and Graphic Arts of Cyprus University of Technology, aiming at devising an agreed, unified orthographic system, meeting the criteria of phonological adequacy, simplicity and learnability (Sebba, 2007). The present-day linguistic situation in the GC community involves the parallel use of two linguistic varieties: the local Greek-Cypriot dialect (GC) and the official Standard Modern Greek language (SMG). The two varieties exhibit differences on all linguistic levels with the phonological, morphological and lexical levels showing the highest degree of divergence (Goutsos and Karyolemou, 2001). The standard variety (SMG) is learnt mainly through formal education and its use pervades in formal settings such as administration, media, education and in writing since it is the only variety that is standardised and codified. On the other hand, the GC dialect is the most intimate language code for GCs and the one used for their everyday oral communication (Karyolemou, 2001). We argue that there is an increasing need for establishing a uniform way of representing CG in writing due to the emergence of new uses of its written form apart from folk literature and computer mediated communication. To cover the lack of an official

and widely accepted orthographic system for the representation of GCD distinct sounds, a series of studies conducted in our lab using online and offline questionnaires in combination with eye-tracking experiments, investigating informants' stances, preferences and attitudes towards this orthographic system (Papadima et al, 2014, Papadima, Ayiomamitou & Kyriacou, 2013). The typographic design of the new set of typographic characters aims to provide ideal reading conditions, based on visual homogeneity, text unity, rhythm and reading flow (Unger, 2007; Noordzij, 2005; Bringhurst, 2005).

Bringhurst, R. (2005). *The elements of typographic style*. Point Roberts, WA: Hartley & Marks, Publishers.

Goutsos, D. and Karyolemou, M. (2004). "Introduction." *International Journal of the Sociology of Language* 168:1-17.

Karyolemou, M. (2001). "From linguistic liberalism to legal regulation: the Greek language in Cyprus." *Language Problems & Language Planning* 25: 25-50.

Papadima, A., Ayiomamitou, I., Kyriacou, S., Parmaxis, G. (2014). Orthography development for the Greek-Cypriot dialect: language attitudes and orthographic choice. In Dyck, C., Granadillo, T., Rice, K. (ed.). *Dialogue on dialect standardization*. (pp. 63-80). Newcastle: Cambridge Scholars Publishing.

Papadima, A., Ayiomamitou, I., Kyriacou, S. (2013). "Typographic Practices and Spelling Convention for the Written Representation of a Non-Standard Dialect: the Case of the GreekCypriot Dialect". In Lachout, M. (Hrsg), *Aktuelle Tendenzen der Sprachwissenschaft*. (pp. 87-100). Hamburg: Verlag Dr. Kovac.

Sebba, M. (2007). *Spelling and society: The culture and politics of orthography around the world*. Cambridge: Cambridge University Press.

Noordzij, G. (2005). *The stroke: Theory of writing*. London: Hyphen Press.

Unger, G. (2007). *While You're Reading*. NY: Mark Batty

Grammar first, for children who learn to write Dutch verbs?

Mijntje Peters, Johan Zuidema, Anna Bosman & Anneke Neijt, Centre for Language Studies and BCN, Radboud University Nijmegen

The Dutch spelling curriculum starts in grade 1 and 2 with lessons about the influence of phonology and morphology on spelling, followed by autonomous spelling rules such as gemination and degemination of letters in grade 3. Towards the end of primary school, in grade 5, the system of how to spell Dutch verbs is introduced via an algorithm that requires knowledge of syntax. Children then learn to distinguish nouns, verbs, adjectives, finite verbs, tense, number, and past and present participles. Despite multiple attempts with different kinds of didactical approaches throughout the years, a satisfactory solution for teaching verb spelling has not been found: children still make numerous errors in writing verbs when they finish primary school. Usually, syntax and verb spelling are introduced simultaneously in grade 5. We investigated the effect of prior introduction of syntactic knowledge, in grade 4, followed by the usual approach in grade 5. Comparison of our experimental group with a control group that followed the standard procedure shows that the experimental group starts with higher syntactic proficiency in grade 5 than the control group. The difference between the two group was gone, however, at the end of grade 5. Verb spelling proficiency was not influenced by our intervention. The groups performed equally poor on verb spelling, which leaves the question of how to achieve better performance.

Non-orthographic spaces in German and Scandinavian. An OT analysis

Vilma Symaczyk Joppe, Bergische Universität Wuppertal, Germany

Most Germanic languages share the basic rule that morphosyntactic words correspond to graphemic words. In the last decades, linguists have noticed a tendency in these languages to write nominal compounds with internal spaces, in violation of this rule. Empiric studies—e. g. Hallencreutz (2001) for Swedish, Walmsness (2002) and Vollan (2007) for Norwegian,

Heidemann Andersen & Diderichsen (2011) for Danish and Symaczyk Joppe (2011) for German—could show that popular general explanations, like the influence of the English writing system or spelling reforms as a source of irritation, do not suffice to explain this trend. Instead, many of the non-orthographic spaces can be reduced to the presence of concrete, identifiable triggers, like brand names, abbreviations, names, and numbers in the respective compound, and also to the genre in which the compound appears. In nearly all of these cases, the non-orthographic spaces indicate morphological structure.

The first part of my talk will be an overview of the phenomenon of N_N compounds and the responsible factors in the languages mentioned above, with special reference to my corpus study on German compounds. A part of the compounds in this corpus could be attributed to individual writers. I will show that those writers can be roughly divided into three groups with respect to N_N compounds.

Finally, I will outline the systems for each group of writers in an OT framework. Jacobs (2005) analyses the German system of space insertion as the result of conflict between two basic constraints. I will show that—due to lack of orthographic knowledge and/or biased input—additional, overgeneralizing constraints can emerge, resulting in the non-orthographic spelling variants associated with the different types of writers.

Hallencreutz, K (2001): "Ski på längden, inte på engelskan". *Språkvård* 4/01, 4-9. ● Heidemann Andersen, M. & P. Diderichsen (2011): "Sjældne og sære sammensætninger. Om særskrivninger og sammensætninger i moderne dansk". *Nydanske Sprogstudier* 41. 40-65. ● Jacobs, J. (2005): *Spatien. Zum System der Getrennt- und Zusammenschreibung im heutigen Deutsch*. Berlin: de Gruyter. ● Symaczyk Joppe, V. (2011): "Spatien innerhalb deutscher Komposita. Eine empirische Untersuchung normwidriger Schreibungen." M. A. Thesis, University of Wuppertal. ● Vollan, M. (2007): "<Holdnings skapende handlings planer> - Særskrivning i studenttekster". *Språknytt* 4. 23-28. ● Walmsness, R. (2002): "Særskrivning av sammensatte ord". *Språknytt* 3-4. 26-29.

Negative L1 Orthographic Processing Transfer on L2 Literacy

Emina Tuzovic & Marjorie Lorch, Applied Linguistics & Communication, Birkbeck, University of London
Individuals with L1 literacy in a Semitic language, with consonantal roots and syntactically-conditioned vowels, must acquire the grapheme-phoneme correspondences and orthographic processes to read in English. Previous investigations of L2 literacy have revealed notable language-specific features in orthographic word form processing for Arabic L1/English L2 reading. Ryan & Meara (1991) and Saigh & Schmitt (2012) report that in contrast to other EFL learners, they tend to process English words by focusing on the central rather than the initial or final part of the word virtually trying to extrapolate the consonant root. The negative transfer of L1 Arabic reading processes leads to "vowel blindness" when reading English, and letter transposition recall errors are frequent in this particular group of learners (Al-Sulaimani 1990; Bowen 2011). The present study employed a letter transposition task to investigate whether more complex syllabic/phonotactic factors might be at play.

Thirty intermediate L1 Arabic/L2 English students performed a novel letter transposition task designed to determine the phonotactic/orthographic factors that influence these readers' strategies. This experiment extends the methodology of Ryan & Meara (1991) who supplied the original lexical stimuli. Ninety 10-letter words were controlled for the position of letter transposition and the type of transposition (V/V;C/C;C/V;V/C) in two conditions: identical (government-government) and transposed (government-goverment). The stimuli were presented in pairs tachistoscopically. Judgements of whether the word pair was identical were recorded for accuracy and RT.

Regarding error rate, the participants made most errors in mid and final word positions which witnesses Arabic learners processing words differently to the students with alphabetic L1 background. They were the slowest and made fewest errors in the identical condition. A large intra-group variation in the participants' scores was linked to L2 proficiency. The results indicate that language-specific aspects of L1 reading and recall strategies may be maintained in relatively high proficiency L2 learners.

Principle analysis for orthography development

Mira Valkama, University of Helsinki

Orthographies interpret several perspectives of a social and linguistic situation simultaneously: although writing is primarily realised through the representation of various levels of linguistic structure, an orthography is also an instrument in written communication and an emblem of its users. Varying goals and ideologies are reflected in the choice of graphical signs, their correspondence to linguistic units and the choice of the language variety represented by the orthography. My study explores the principles that have been suggested to guide the solutions in orthography development and the use of these principles in practice. The principles included in the study are collected from previous literature on orthographies, including both theoretical works and case studies. In order to take into account the conflicting and violable nature of orthography principles, I will analyse and reformulate them as constraints, in line with the ideas suggested in Optimality Theory (see the similar approach in e.g. Wiese 2004). The results of this theoretical analysis will be applied as a tool in describing orthography development principles used in individual orthographies. In this presentation, I will present an analysis of the orthographies devised for Northern Pwo Karen in Thailand (Cooke et al 1976; Phillips 2009). Once the tool has been initially tested, it can be used for the analysis of other orthographies. The results from a cross-orthography study may show trends on which kinds of principles are used in which contexts and if the use of certain principles correlates with the success in implementing the orthography. In practice, the principle analysis can be used in orthography development for reflecting on possible solutions and recognising strengths and weaknesses in orthography design.

Double Dutch Vowel Spelling

Johan Zuidema & Anneke Neijt, Centre for Language Studies, Radboud University Nijmegen

The BasisSpellingBank with triplets of phonemes, graphemes and spelling categories (Beeksmā, Neijt & Zuidema 2015, Zuidema & Neijt, 2012, 2015) allows for unambiguous documentation of the spelling of Dutch words and their derived or inflected forms. However, the official spelling dictionary (Nederlandse Taalunie, 2015, woordenlijst.org) neither links the spelling of words in the dictionary to their phonemic representation, nor to the rules described in the *Leidraad*. Our survey of derived and inflected forms shows that for this reason, the spelling dictionary is inadequate in this respect.

Double Dutch vowel spelling is a case in point. In Dutch, the set of vowels is larger than the set of vowel letters in the alphabet. Therefore, digraphs are used, such as *ie*, *oe* for [i, u] and *aa*, *ee*, *oo*, *uu* for [a, e, o, y]. Geminates are usually degeminated in syllables without a coda: *baal-balen-ja*, *veel-vele-mee* (not *me*, since *e* expresses schwa as well), *vroom-vrolijk-zo*, *vuur-vurig-nu*. This part of Dutch orthography is complex. It requires much time and effort in primary education.

The system continues to be complex for proficient users, since digraphs, geminates, single letters or apostrophes are used in derived or inflected forms, see the following examples from the spelling dictionary (* for forms not in the dictionary).

digraph (default)	degemination in word final position	before plural and genitive <i>-s</i> and superlative <i>-st</i>	before diminutive <i>-tje</i>
[a] - aa	opa 'grandpa'	opa's	opaatje
	la 'drawer'	la's	laatje
[e] - ee	café <i>idem</i>	*café's, cafés	cafeetje
	privé 'private'	privé's (<i>something ...</i>), *privés, privé'st	
	e 'the letter e'	e's	*eetje, e'tje
	re 'musical note'	re's	*reetje, re'tje
	pre 'preference'	pre's	*preetje, pre'tje
	ave (name of pray)	ave's	aveetje
[i] - ie	ski <i>idem</i>	ski's	skietje
	baby <i>idem</i>	baby's	*babetje, baby'tje
	sexy <i>idem</i>	sexy's (<i>something ...</i>), *sexy'st, sexyst	
	mi 'musical note'	mi's	mietje
	pi <i>idem</i>	pi's	*pietje, pi'tje
[o] - oo	cabrio 'cabriolet'	cabrio's	cabriootje
	hav _o (acronym)	hav _o 's	havootje
[y] - uu	parapl _u 'umbrella'	parapl _u 's	parapluutje
	menu <i>idem</i>	menu's	menuutje
[u] - oe	gnoe 'gnu'	*gnoe's, gnoes	gnoetje
	sudok _u <i>idem</i>	sudoku's	*sudokoetje, sudoku'tje

There are no subpatterns on the basis of which one may decide how to write inflected or derived words not listed in the dictionary. Digraphs are the default (cf. *ave* – *aveetje* and *ski* – *skietje*), though apostrophes are used as well (*pre* – *pre'tje*; *pi* – *pi'tje*). The forms *re* – *re'tje* and *mi* – *mietje* show that the spelling of the other diminutive forms of *do*, *re*, *mi*, *fa*, *sol*, *la*, *ti*, not listed in the dictionary, may be a digraph as well as an apostrophe. Lists of regular and irregular words in the dictionary need to be compatible with the description of the spelling by rules.

Beeksmā, M., A. Neijt & J. Zuidema (2015, submitted). SHOTGUN: converting words into triplets. A hybrid approach to grapheme-phoneme conversion in Dutch.

Zuidema, J. & A. Neijt (2012). Verkennend onderzoek naar de wenselijkheid en de haalbaarheid van een verrijkte Woordenlijst Nederlandse taal ten behoeve van spellingonderwijs.

<http://taalunieversum.org/sites/tuv/files/downloads/rapport%20VWS%2015022013.pdf>

Zuidema, J. & A. Neijt (2015, submitted). The BasisSpellingBank – spelling knowledge stored in a lexicon of triplets.

12.30 Lunch

May 12, Afternoon session

13.30 Invited: From phonemic spelling to distinctive spelling

Geoffrey Sampson, University of South Africa & emeritus Sussex University

The paper proposes a generalization about the history of scripts: they tend to move away from faithfully mirroring speech, towards giving the meaningful units of language written forms that are distinctive, both in the sense of ignoring morphophonemic variation, and in the sense of

maximizing the differences between writings of distinct units. This development is a desirable response to changing social conditions.

14.30 **The pros and cons of IPA-based writing systems**

Martin Neef, TU Braunschweig

A number of scholars like Sampson (1985), Coulmas (2003), and Harris (2009) claim that the International Phonetic Alphabet should be regarded as an alphabetic script. Therefore, IPA should be subject to grapholinguistic analysis. If a necessary condition for a writing system is its connection to a specific language, IPA could be regarded as a model kit that can be used to create writing systems for each single language. The IPA script supplies a large number of characters (plus additional symbols of different status). In order to deploy IPA for a specific language, only a subset of these characters is needed. Therefore, the specific IPA-based writing system for, e.g., English could be addressed as ‘English-IPAEnglish’, indicating that the writing system of English uses a modified version of the IPA-script, more precisely a specific reduced version or subset, as is the case for each IPA-based writing system. In an IPA-based writing system, the set of letters matches the set of phonological segments, all correspondence rules are simple, and neither fixed letter combinations nor graphematic constraints exist. Nevertheless, such a writing system allows writing any text that is conceivable based on the particular language system. It is, thus, a perfectly viable writing system. It is most shallow (in terms of the concept of orthographic depth, cf. Katz & Feldman 1983) and absolutely transparent (in terms of Neef & Balestra 2011). In my talk, I will discuss whether a writing system like English-IPAEnglish is optimal or defect. Such a discussion benefits from the distinction between two components of writing systems, namely graphematics and systematic orthography. In passing, I will also touch on the distinction between the terms ‘phonetic transcription’ and ‘phonological notation’, reflecting the notorious delimitation of phonetics and phonology.

15.00 **Coffee & poster session**

May 12, **Late afternoon session**

16.00 **Writing morphophonology, reading lexical tone: evidence in favor of morphographic spelling in Kabiye (Togo)**

Dave Roberts & Steve Walter, www.nyinyu.com, blog: cornishmaninafrica.wordpress.com

The shorter a word, the more likely it is to be lexically ambiguous. In the standard orthography of Kabiye, a tone language of Togo, numerous monosyllabic heterophonic homographs (tonal minimal pairs) and homophonic homographs occur in the imperative and six associated conjugations. This paper presents the complete catalogue of these verbs, and then examines them in natural contexts. It goes on to propose a morphographic spelling in which elided root-final labial consonants are written as superscript silent letters to help the reader identify the lexeme. This spelling is tested against a tonographic diacritic alternative in an oral reading experiment. Those who learned the morphographic spelling gained more in reading accuracy from the addition of superscript silent letters than those who learned the tonographic spelling did from the addition of diacritics.

KEYWORDS: orthography, morphographic spelling, homography, lexical tone, Kabiye, African languages.

16.30 Principles of English spelling: from Beowulf 2 bootylicious

Des Ryan, Centre for Language & Communication Studies, Trinity College, Dublin

This talk will present the main findings of my doctoral thesis: *Principles of English spelling*. Four major principles will be explained and exemplified and it will then be shown how different kinds of words are subject to different kinds of spelling patterns, whether they be core or peripheral spellings.

Identity Preservation Principle. The fundamental principle is that new spellings take their form from the spelling of their component parts. Amendments are only made where necessary, if possible. IPP applies to compound, derived and inflected words, formed in English (compare *signpost*, *signage*, *signs*). This is the traditional idea of ‘morphemic spelling’. The principle is generalised to include etymological spellings of borrowed words (hence *design*, *insignia*, *signet*, but not *cygnet*).

Invariance principle. English spellings are largely invariant, despite changes in phonology, over time and across varieties. Phonological rotations have little effect on the system (e.g. the Great Vowel Shift), mergers result in two spelling units for one sound (*dear*, *deer*) and splits result in one spelling unit for two sounds (*bat*, *bath*).

Distinctiveness Principle. The resulting complexities in the system can be exploited to create distinct spellings, as in *metal* – *mettle*, *bite* – *byte* and *in* – *inn*. This is particularly common in names, where spellings can be invented e.g. *Foxx*, *OutKast*, *Gorillaz*. In inflections, however, the writing system cannot facilitate a distinction between *evening* (noun) and *even-ing* (verb).

Phonographic matching principle. This is particularly important for resolving conflicts in the application of IPP: e.g. adding <e> in *buses* and *flexes*, blocking *<buss>, *<flexs>. An updated model of bi-directional correspondences will be proposed, by building of the models of Venezky, Carney, Rollings and Evertz.

The interaction of these principles will be exemplified in the following kinds of words: Inflection, derivation, clippings, blends, abbreviations, names, txt msgs, eye dialect and onomatopoeia.

17.00 Beyond Conventional Writing Boundaries: Non-Standard Language Practices in Dutch Youths’ Computer-Mediated Communication

Lieke Verheijen, Centre for Language Studies, Radboud University Nijmegen

New media play a vital role in many peoples’ lives: especially youngsters are continuously engaged in computer-mediated communication (CMC). Youths’ informal language practices in CMC often go beyond the boundaries of the standard written language, disregarding conventional spelling and grammar rules (Frehner 2008, Cougnon & Fairon 2014). Such language practices on social media are feared to have a detrimental impact upon youths’ traditional literacy skills, i.e. writing and reading (Thurlow 2006). The possible effect of computer-mediated communication on literacy is as yet unclear, since previous research has yielded mixed results (Verheijen 2013). But first, it is paramount to know what the non-standard language often used in CMC actually looks like. Therefore, two studies were carried out. First, I conducted a systematic register analysis into Dutch youths’ CMC writings, to determine how their new media texts exactly vary from Standard Dutch. CMC texts of various genres (text messages, instant messages, microblogs, social network

posts) were analyzed for features of three writing dimensions (orthography, lexis, syntax) and compared to school texts. Some data were extracted from an existing corpus; additional data were collected via voluntary donations. This study reveals that the linguistic characteristics of new media writings are significantly determined by CMC genre and age group, and that these linguistic ‘deviations’ serve specific pragmatic functions. Second, I conducted a correlational study into Dutch youths’ CMC use and their writing quality. Students of different age groups and various educational levels were tested. To gauge their use of new media, they filled in extensive questionnaires; to measure their writing proficiency, they composed school essays. The data are still under analysis, but preliminary results suggest that youths of a younger age, especially of lower educational levels, have more difficulty in keeping the informal ‘CMC language’ separate from the more formal register expected in educational contexts.

Cougnon, L.-A., & C. Fairon (eds.) (2014). *SMS Communication: A Linguistic Approach*. Amsterdam: John Benjamins.

Frehner, C. (2008). *Email - SMS - MMS: The Linguistic Creativity of Asynchronous Discourse in the New Media Age*. Bern: Peter Lang.

Thurlow, C. (2006). From statistical panic to moral panic: The metadiscursive construction and popular exaggeration of new media language in the print media. *Journal of Computer-Mediated Communication*, 11(3), 667-701.

Verheijen, L. (2013). The effects of text messaging and instant messaging on literacy. *English Studies*, 94(5), 582-602.

17.30 AWLL business meeting

19.00 Dinner in town

Restaurant Het Savarijn, Van der Brugghenstraat 14, 6511 SL Nijmegen

May 13, Morning session

9.30 Invited: **The Digital Literacy Instructor (DigLin)**

Ineke van de Craats, Center for Language Studies, Radboud University

DigLin is software for adults – immigrants and refugees – to learn to read and write for the first time in their lives in a new language and a new culture. Both the receiving countries and the learners often have the illusion that this is a fast process. The reality is different, as reported in the literature.

Funded by the Lifelong Learning Program (Grundtvig), a group of researchers, speech-technologists and practitioners took the initiative to develop basic software for this group of real beginners for four target languages: Finnish, Dutch, German and English, varying in orthography from transparent to opaque. The impact of the orthography became most noticeable in the selection of words and phonemes: for Finnish, Dutch and German almost the entire phonemic repertoire was integrated in 15 wordlists, for English not even half of it.

The aim of the program was to intensify practicing: more often, much longer and deeper. This led us to digital methods. The material stands out by using the computer right from the beginning, by its systematic approach departing rather from sounds than from letters, and its continuous and immediate feedback for every item. The most innovative feature of DigLin is that in production exercises learners can read aloud and get feedback on their speech production. This is made possible through the use of Automatic Speech Recognition (ASR).

Before demonstrating the programme I will sketch the target group: their concepts of words and script and the different role of context in spoken and written language. I will show how the above was taken into account when creating the interface: how a word sounds, how a word looks like, what it means, how it can be split up and how the segments can be glued together. All in one exercise and one screen.

You will get an impression of the multiple ways feedback is organised. Why feedback by ASR on the read-out/pronounced word is difficult to realise, for technical reasons and for this specific group of learners.

We close off by saying a few words on the pedagogical approach of high expectations we intended, although the teachers in the pilot were free to follow their own ideas. Most teachers did not use this approach, but learners managed to set their own targets, became more autonomous and said to like DigLin very much.

10.15 Language segmentation practices result from literacy: a cross-linguistic study

Dorina Veldhuis, Hogeschool iPabo, vakgroep Nederlands

Abstract Studies into the units of language have been prominent for long. The units that can be recognized are often said to be words, syllables, and phonemes. Studies of children's metalinguistic development from the 1960s onwards have most usually shown that the awareness of units in language develops gradually (cf. Clark 1978; Gombert 1992; Homer 2009; Slobin 1978; Ravid & Hora 2009). The question of whether this awareness is enhanced by literacy, or whether it is the other way round, has steadily gained attention (cf. Karmiloff-Smith et al. 1996; Kurvers & Uri 2006). In four recent studies, we investigated what units were segmented in ongoing speech by monolingual Dutch and Turkish, and by bilingual Turkish-Dutch, by mono- and bi-literate Chinese-Dutch children, by Japanese and Dutch-Japanese adults, and by multilingual Mandinka speakers in The Gambia. This latter language was included as writing in Mandinka is not commonly practiced despite adult literacy education in the language (Juffermans & Veldhuis 2012). In all of the studies, participants either indicated the 'pieces' they heard in spoken language (Dutch, Turkish, Chinese or Japanese), or indicated whether a hand-written text was spelled correctly (in Mandinka). The studies confirmed Bassetti's (2006) suggestion: segmentation of units in language is related to the writing system used, and to the experience in reading and writing a specific language (Juffermans & Veldhuis 2012; Veldhuis 2015; Veldhuis, Li & Kurvers 2012). Our study in Japanese also showed that for a language in which word boundaries are not marked in writing, segmentation either seems to depend on semantics, or on the convention in one's first writing system (Veldhuis 2011). These results imply that word awareness does not need to be trained at an early age; it develops with literacy practice.

10.45 Coffee

11.00 Among Phoenicians: a comparison of learning and teaching an alphabetic and syllabic writing system

Yonas Asfaha¹, Jeanne Kurvers² & Sjaak Kroon², ¹University of Asmara, ²Tilburg University

In their seminal article 'Teaching reading by use of a syllabary' Gleitman and Rozin (1973) advocated the use of the syllable as a unit for the initial acquisition of reading, a claim that was

polemically responded by Goodman (1973) who questioned the efficacy of all phonological principles in learning to read.

In another seminal article, Ziegler and Goswami (2006) introduced the Psychological Grain Size Theory (PGST) in which they introduced three core issues in explaining differences in beginning reading in several European orthographies: the availability or accessibility of phonological units, the consistency of the mapping between spelling and sound, and the granularity or the grain size of the scripts.

In this contribution we will revisit the debate by presenting the results of a study in Eritrea that offered a nice opportunity to compare the learning and teaching of different writing systems. Eritrea had adopted a mother tongue policy in which primary school children start learning to read in one of the nine mother tongues that use three different writing systems.

In our study we compared learning to read in two languages that use the syllabic Ge'ez script (Tigrinya and Tigre) with two languages that use the Roman alphabetical script (Kunama and Saho). The study revealed the initial advantage of learning the syllabic based Ge'ez compared to the phoneme based Roman script, despite the much bigger number of units in Ge'ez that have to be learned. Moreover, within the Roman alphabetic script, the teaching of Saho that started teaching with syllabic units produced better results than the phonemic based teaching of Kunama. The findings will be discussed in relation to Gleitman and Rozin's proposal, the grain size theory and the implications for teaching beginning reading in other languages.

11.30 Core syllables vs. moraic writing

Gene Buckley, U. Pennsylvania

The category of SYLLABARY in a traditional typology of writing systems would seem to represent syllables, but especially since Poser (1992) the view has become more common that (many) syllabaries are actually MORAIC writing systems, in which the basic sign represents the mora, a measure of syllable weight. This analysis is at least partly endorsed many more recent sources (Ratcliffe 2001, Coulmas 2003, Rogers 2005, Gnanadesikan 2012). I argue that syllabaries really do represent syllables, but typically the CORE (or minimal) syllable of a language, which is CV (and V). Although CV corresponds to a light syllable of one mora, the ways in which the basic sign is supplemented to represent more complex syllable types is not fundamentally connected to moraic structure. In particular, CV systems often write nonmoraic distinctions using the same tools that write moraic differences. This includes Japanese kana, frequently cited as moraic writing: long vowels and geminates require an additional symbol, and entail one mora, but complex onsets such as *kya* require a similar addition without a moraic difference. (Poser calls this ONSET SPLITTING, but that does not remove the phenomenon from evidence.) Even systems in which many bimoraic syllables require two symbols, such as Akkadian CV+VC, often ignore vowel length and gemination, exactly what moras represent in spoken language. In a syllabary for a language with any complexity in its spoken syllables, it is considerably more efficient to use a small set of CV symbols that can be supplemented in various ways to represent more complex configurations. Examining a range of examples, I show that the notion of mora is not crucial in such supplementations, and is often incorrect. Instead, the evidence from writing systems supports the dominant view in phonological theory that moras are units of weight, but not prosodic constituents.

12.00 Just mixed up or a pretty neat idea? Some reflections on the multi-script nature of the Japanese writing system

Terry Joyce¹ & Hisashi Masuda², ¹Tama University, Japan, ²Hiroshima Shudo University, Japan

The modern Japanese writing system is multi-script in nature—consisting of morphographic kanji, two syllabographic scripts (hiragana and katakana), and phonetic alphabet of rōmaji (Joyce, 2013; Joyce, Hodošček & Nishina, 2012)—which has prompted Kess and Miyamoto (1999: 9) to remark that the “Japanese may have the unique distinction of employing all three extant means” of written language. While the mixture is often singled out as one factor contributing to the complexity of the Japanese writing system (Joyce 2013), it is also, undeniably, the source of remarkable levels of orthographic flexibility and variation (Backhouse, 1984; Joyce, Hodošček & Nishina, 2012). Thus, studies of the Japanese writing system can potentially illuminate a number of the core issues for understanding writing systems and written language.

In that spirit, this paper tenders miscellaneous reflections on the multiple scripts of the Japanese writing system. After outlining the basic functional demarcations, both historical and contemporary, between the scripts that guide their largely complementary roles within an overall writing system, the paper also considers examples of orthographic variants, such as 玉ねぎ being more frequent than 玉葱 /tamanegi/ ‘onions’ in the BCCWJ corpus (Joyce, Hodošček & Nishina, 2012). Other aspects reflected upon include (1) the kanji-orthography representation of recent Sino-Japanese compound word formations, such as 婚活 /konkatsu/ ‘seeking marriage’ and 菌活 /kinkatsu/ ‘dieting by activating good bacteria’; (2) the katakana-orthography representation of foreign loanwords modified for Japanese phonology; (3) the rōmaji-orthography representation of foreign words and brand names particularly in advertising, where pronunciations may depend on the grapheme-phoneme correspondences of different source languages and thus require katakana glosses to convey intended pronunciations, such as in Rizap/ライザップ /raizappu/ ‘brand name for chain of fitness gyms’; and (4) some results from primed visual word recognition research with various Japanese scripts.

12.30 Lunch

May 13, Afternoon session

13.30 Naturalness of scripts and writing systems Prolegomena to a Natural Grapholinguistics

Dimitrios Meletis, Department of Linguistics, University of Graz

Natural Linguistics (with its most notable subdivisions Natural Phonology and Natural Morphology) investigates the notion of ‘naturalness’ and claims that if a linguistic element or structure can be processed with little effort both sensomotorically – articulatory and perceptually – as well as cognitively, it is more natural compared with other, more complex linguistic phenomena. Naturalness is thus understood to be a gradual notion closely related to concepts such as optimality, markedness, and complexity. Drawing on external evidence such as language change, language acquisition and language disorders, various parameters of naturalness (e.g. ease of production/perception, transparency, uniformity, iconism/iconicity) have been postulated and described, focusing on the phonological and morphological systems of language. Since a subcomponent of Natural Linguistics pertaining to scripts and writing systems is still lacking, it is

the main purpose of this paper to give an outline of how the concept of naturalness can be extended in order to account for grapholinguistic phenomena. It is shown how a method such as comparative graphematics can help reveal naturalness parameters that apply to both the material (graphetic) as well as the functional (graphematic) aspects of writing. Examples from various writing systems (mainly German, Chinese, Thai, and Arabic) such as the development of writing direction, the reduction of symmetry, the systematic visual composition of characters, and the re-orientation of writing are given in order to illustrate that there exist not only system-specific, but also underlying universal natural preferences that shape scripts and writing systems. The systematic integration of these findings into the framework of Naturalness Theory demonstrates that a so-called Natural Grapholinguistics could offer promising new insights into the nature of writing as well as a *tertium comparationis* for future comparative analyses of writing systems.

14.00 Orthographic and Visual Clues to Metalinguistic Knowledge, Script Acquisition, and the Cognition of Reading and Writing of Ancient Mayan Scribes

David F. Mora-Marín, University of North Carolina at Chapel Hill

Mayan Hieroglyphic writing was used in the Maya region (southeastern Mesoamerica) from ca. 300 BCE to ca. CE 1697. During that span of time, the script represented speech varieties of two of the six major subgroups of the Mayan language family, Greater Tzeltalan and Yucatecan. This paper aims to study the question of metalinguistic knowledge of the ancient scribes, and the related matters of script acquisition and the ease of reading and writing, by testing two hypotheses pertaining to the visual and orthographic principles of the script. The paper begins with a review of the so-called “glyph block” as a unit of visual organization of signs which is assumed by all specialists, but has received very little scholarly attention. The paper poses the hypothesis that the glyph block was employed as a punctuation system for demarcating syntactic constituents, and a random sample of inscriptions, selected using the Maya Hieroglyphic Database (Macri and Looper 1991-2014), is employed to test it. Next, the paper continues with a review of the basic types of signs (logograms, syllabograms, semantograms), followed by a discussion of the practice of phonetic complementation of logograms (Mora-Marín 2008), as a preliminary step to addressing the nature of logographic spellings. Then, the hypothesis, put forth by Mora-Marín (2010), that Mayan logography was a flexible principle (with polymorphic logography, requiring contextual disambiguation, a requisite subcategory) is then tested by means of a nonrandom sample with the goal of isolating correlating factors (e.g. text length, media type, genre type, glyph blocks). The paper concludes that scribes were trained to recognize syntactic constituents of various types (e.g. clauses, PPs, NPs, VPs), and that they were utilizing an efficiency principle based on a minimum number of visual and spelling clues that readers could employ to extract the complete morphological form of a spelled word.

Macri, Martha J. and Matthew G. Looper 1991–2008. Maya Hieroglyphic Database. Beta version available at the Native American Language Center, University of California, Davis.

Mora-Marín, David F. 2008. Full Phonetic Complementation, Semantic Classifiers, and Semantic Determinatives in Ancient Mayan Hieroglyphic Writing. *Ancient Mesoamerica* 19:195-213. 2010. Consonant Deletion, Obligatory Synharmony, Typical Suffixing: An Explanation of Spelling Practices in Mayan Writing. *Written Language and Literacy* 13: 118-179.

14.30 How does writing something down change one's (mental) representation of it?

David Olson, emeritus University of Toronto, Canada

How does writing something down change one's (mental) representation of it? I consider three theories: 1) it doesn't, 2) writing as a tool that extends the mind, 3) writing as a mode of thought in that builds on a consciousness of the implicit structures of spoken language. The structures implicit in spoken language are made explicit through a metalanguage. Some attempt will be made to distinguish the metalinguistic concepts common in ordinary speech, concepts such as "ask", "say", "tell", from those developed around written texts including such logical concepts as "implication", "definition", "semantic meaning" and "paraphrase". The goal of such analysis is to link writing and literacy to rationality, the subject of my forthcoming book *The mind on paper: Reading, consciousness and rationality*.

15.00 Panel discussion: orthographic data basis, typology of writing systems

16.00 Conclusion & Drinks