On the systematic nature of writing systems

13th International Workshop on Writing Systems and Literacy

Online (University of North Carolina, USA)

21-23 October 2021

Local organizer: David Mora-Marín (University of North Carolina, USA)

Programme committee: Lynne Cahill (University of Sussex, UK), Terry Joyce (Tama University, Japan); David Mora-Marín (University of North Carolina, USA); Dorit Ravid (University of Tel Aviv, Israel)

AWLL13 website: http://faculty-sgs.tama.ac.jp/terry/awll/WS/13/2021-HP.html

Zoom link (short URL): https://go.unc.edu/AWLL13   **QR code:**

AWLL13-Online is open to all, but please note that registration is required.

Participants are recommended to register in advance to avoid delays to joining the zoom meetings.

AWLL gratefully acknowledges the following academic units at the University of North Carolina for their funding support of AWLL13-Online. Thanks to their various donations, it has been possible to cover the conference's sundry expenses, without the need to levy registration fees.

Archaeology Curriculum	Department of Classics
Carolina Asia Center	Department of Linguistics
Department of Anthropology	Department of Romance Languages
Department of Asian and Middle Eastern Studies	Institute for the Study of the Americas
Department of Asian Studies	

Although AWLL13 is being convened online (as an unavoidable consequence of the ongoing corona virus pandemic) and is somewhat scaled-down in nature (in consideration of time-zones and zoom fatigue factors), we genuinely hope that AWLL13-Online's programme of presentations will be a positive experience in the tradition of AWLL gatherings.

AWLL13-Online: Programme

Zoom etiquette guidelines:

- Consistent with standard settings, please note that all participants will be muted on entering the zoom meetings, but participants are also asked to ensure that they stay muted during the presentation sessions.
- Naturally, presenters are expected to unmute themselves and have camera on during presentations (please note we will be taking screen-captures of title slides and presenter in "Presentation-Speaker" views).
- Participants are asked connect from quite locations, if possible, to help avoid unwanted 'sound intrusions' during the presentations and discussions.
- In addition to screen-captures for each presentation, we also planning to sample screen-captures of participants in gallery view (probably at the start and end of each day to later make available at AWLL13 website, as a memento of the conference).
- Please note that a support person (albeit not an IT expert) will be monitoring the chat function (in addition to session chairs) and may be able to help with some technical issues. If not, it might be necessary to leave and re-join the meeting.
- Rather than leaving AWLL13's zoom meetings during the breaks (with the need to re-join), participants are recommended to stay logged in, but please remember to check that you are muted with your camera off, if you take a break.

Oral and poster session protocols:

- Please note that session chairs will also mediate the Q&As for oral presentations, by fielding and selection audience questions via the chat window.
- To facilitate the tracking of questions in chat, we would encourage participants to prepare any questions ready to enter to the chat window once a presentation ends.
- The poster session will begin as a normal zoom session, where all presenters are to briefly (3 mins max) introduce their posters.
- After the initial introductions, participants will be free to select which breakout room/poster to visit and move between rooms, via the main room, as often as they wish.

Between-session breaks + open-exchange session:

- In order to facilitate discussions amongst participants, a number of breakout rooms will be available during the 30-min breaks between the sessions, so participants can easily break off into smaller groups (please note that these will only be open during the breaks and will be closed just prior to the next session).
- Moreover, in addition to the separate breakout rooms assigned to each poster presenter during the poster session, a number of other breakout rooms will also be available during the final session of the conference for open exchange discussions.

Programme schedule: Please note that all times are according to Eastern Daylight Time (EDT), which is four hours behind Coordinated Universal Time (UTC-4)

Day 1: Thursday, 21st October 2021

08:55 - 09:00 Opening remarks

09:00 – 11:00 Oral session 1 (chair: Terry Joyce)

09:00-09:30 **Sven Osterkamp; Gordian Schreiber:** A proposal for a formalized, expandable approach to the taxonomy of writing systems

09:30-10:00 **Arvind Iyengar:** Somewhere I belong: Non-abugidic alphasyllabaries, non-alphasyllabic abugidas and their typological classification

10:00-10:30 **Dimitrios Meletis:** Graphotactics, spatiality, and why writing should be studied independently from speech

10:30 – 11:00 Break

11:00 – 13:00 Oral session 2 (chair: Lynne Cahill)

11:00-11:30 **Mike Cahill:** Usability is not enough: The criticality of sociolinguistic factors in the establishment of new orthographies

11:30-12:00 **Leila Schroeder:** Bridging an orthography chasm

12:00-12:30 **Amalia Gnanadesikan:** Segments and Syllables in Thaana and Hangul: Comparing literate native-speaker inventions

12:30-13:00 **Robert J. Fouser:** Searching for the perfect writing system: 20th century Hangeul reform proposals

13:00 – 13:30 Break

13:30 – 14:30 Keynote presentation (chair: Dorit Ravid)

Peter T. Daniels: When is non-writing writing? or, when is writing non-writing?

Day 2: Friday, 22nd October 2021

09:00 – 10:30 Symposium: Writing systems of the Americas (chair: David Mora-Marín)

09:00-09:20 **Michael Carrasco; Joshua Englehardt:** Enduring literacies: The emergence and maturation of writing systems in Ancient Mesoamerica

09:20-09:40 **Tomi S. Melka; Robert M. Schoch:** *T'oqapu* patterns on Inqa textiles and other media: Do they constitute a writing system?

09:40-10:00 **John W. Adams†:** Notes on Gitksan crests, potlatching, and law

10:00-10:20 **David Mora-Marín:** A survey of systematic spelling practices and conventions in Mayan writing

10:20-10:30 **Discussion**

10:30 – 11:00 Break

11:00 – 12:30 Oral session 3 (chair: Dimitrios Meletis)

11:00-11:30 **Dorit Ravid; Rachel Schiff:** Hebrew affix spelling in children with developmental dyslexia

11:30-12:00 **Martin Neef:** On hidden reforms of the German official orthography

12:00-12:30 **Lynne Cahill:** A corpus analysis of the spread of standardised spelling in Middle English

12:30 – 13:00 Break

13:00 – 14:30 **Oral session 4** (chair: Philippa Steele)

  13:00-13:30 **Martin Uildriks:** A world-altering technology: Script development in Predynastic Upper Egypt

  13:30-14:00 **Rosso Manuel Senesi:** Latin <X>: seeing double

  14:00-14:30 **Corinna Salomon:** Alphabet systemisation in North Italic writing cultures

Day 3: Saturday, 23rd October 2021

9:00 – 10:00 **Keynote presentation:** (chair: Martin Neef)

  Min Wang: Phonology beyond phoneme: Contribution of suprasegmental information to reading

10:00 – 10:30 Break

10:30 – 12:30 **Oral session 5** (chair: Amalia Gnanadesikan)

  10:30-11:00 **Hana Jee; Monica Tamariz; Richard Shillcock:** Quantifying sound-graphic systematicity and application on multiple phonographs

  11:00-11:30 **Keisuke Honda:** Types of 'logography' in the Japanese writing system

  11:30-12:00 **Terry Joyce; Hisashi Masuda:** Systematic chaos or chaotic systems? Some ponderings on the complexity of the Japanese writing system

12:00 - 12:30 Break

12:30 – 14:00 **Posters and open exchange session** (poster intros chair: David Mora-Marín)

  Brian P. Bennett: The impact of Unicode: The case of Church Slavonic

  John A. Bundschuh: Predicate marking strategies in early Japanese vernacular glossing: A preliminary study

  Kevin Heffernan; Yo Sato: Predictors of script choice in Japanese: A data-driven study

  Eleonora Selvi: A story without an end. The never reached systematization of the Pamphylian alphabet

  Hao Sun; Yanwei Jin: Probabilistic grammars in open-class writing systems: A case study of Chinese characters

  Constanze Weth: Syntactic markers - a systematic category in writing systems

14:00 - 14:10 **Closing remarks**

Sven Osterkamp; Gordian Schreiber: A proposal for a formalized, expandable approach to the taxonomy of writing systems

The taxonomy of writing systems is still marred by a number of imperfections (cf. Joyce 2016): The current categories are not labeled in a uniform fashion and have originally been posited on the basis of only a fraction of the known writing systems. Some categories also conflate features belonging to different and in principle independent levels. In consequence, they often fail to accommodate cases situated in between. While the 'messy' nature of writing systems is widely acknowledged, even planar taxonomies (Sproat 2000, Rogers 2005) only take the mixture of morphography and phonography into account while different kinds of phonography are implied to be mutually exclusive within a single writing system. This leads to numerous difficulties in classification, from Persian cuneiform to Phagspa, from medieval Latin to modern English.

A viable solution suggesting itself is a formalized, expandable approach to taxonomy, focusing on the classification of typologically homogeneous *subsystems* coexisting in writing systems (Osterkamp & Schreiber 2021), thus also increasing the visibility of non-dominant subsystems. The basic distinction into segmental, syllabic and morphemic systems is complemented by an *open* number of subtypes, building and expanding on Gnanadesikan (2017) and Poser (1992, 2004). Rather than coining new labels for each subtype as *bundles* of originally independent features, the distinctive characteristics of each system in terms of their patterns of mapping and the boundedness of graphs are indicated via a closed set of symbols.

basic types	selection of subtypes		
segmental (unit: /CV/)	C[V]		Phoenician
	C.V̄ / C[V̌]		Arabic
	Cv̌. V̄ / Cv̌		vocalized Arabic
	Cv		Thāna
	C.V		English
	C(V)v	C[V] / Cv	Devanāgarī
	C(V).V	C[V] / C.V	Phagspa
syllabic (unit: /CVC/)	CV.C(V)		Devanāgarī / Phagspa
	CV.C(V) / CV[C]		early Hiragana
	CV.C		modern Hirgana
	C.VC		Bopomofo
	CVC / CV.VC		Akkadian
syllabic (unit: word)	M		Chinese
	M / M.p		Japanese
	M / P.M / M.P / P.M.P		Mayan

Arvind Iyengar: Somewhere I belong: Non-abugidic alphasyllabaries, non-alphasyllabic abugidas and their typological classification

Since their coinage a quarter-century ago, the terms *abugida* and *alphasyllabary* (Daniels & Bright, 1996; Bright, 1999) have revolutionised our conceptualisation of writing systems. Together with *alphabet*, these terms have proven invaluable in understanding and classifying subtypes of segmentaries— writing systems whose grain size is the phonological segment (Gnanadesikan, 2017). Nevertheless, there remain areas of ambiguity. While segmentaries that are neither abugidas nor alphasyllabaries are usually classified as alphabets, those that are either abugidas or alphasyllabaries—but not both—may be classified under various labels, and inconsistently so. Adding to the mix is the concept of *akshara*, which, depending on interpretation, may form a distinct typological category or a subset of an existing one. While Gnanadesikan (2017) has made significant strides in refining and constraining the above typological categories and associated terminology, she also highlights the need for additional research, for instance, on the appropriate classification of Canadian Syllabics-based writing systems. Moreover, certain writing systems such as Braille-based ones are only rarely covered from the perspective of writing system typology (Sproat, 2010), despite the potential they carry to inform our understanding of the field.

Using Gnanadesikan's (2017) typological classification of segmentaries as the point of departure, this talk proposes an augmented classification of vowelled segmentaries based on various graphe(ma)tic criteria. The classification is illustrated by applying it to various writing systems, including homoscriptal ones that are typologically distinct, as well as lesser studied ones. Particular attention will be paid to the classification of writing systems of unsettled typological status, and the justification for their classification. The talk concludes by identifying potential avenues for further refinement and inquiry.

Daniels, Peter T. & William Bright. (1996). *The World's Writing Systems*. New York: Oxford University Press.

Bright, William. (1999). A Matter of Typology: Alphasyllabaries and Abugidas. *Written Language and Literacy*, 2(1), 45–55. doi:10.1075/wll.2.1.03bri

Gnanadesikan, Amalia E. (2017). Towards a typology of phonemic scripts, *Writing Systems Research*, 9(1), 14–35.doi: 10.1080/17586801.2017.1308239

Sproat, Richard W. (2010). *Language, Technology, and Society*. New York: Oxford University Press.

Dimitrios Meletis: Graphotactics, spatiality, and why writing should be studied independently from speech

Graphotactics captures the "restrictions on ways in which" elements of writing "may combine with each other" (McCawley 1994: 115) to form larger written units. These restrictions take on many forms in the world's diverse writing systems. For instance, users of the English alphabet likely know that <v> almost never occurs word-finally as <e> is commonly added after it, cf. <have>, <give> (cf. Berg 2016: 2). Phonographically, these words (/hæv/ and /ɡɪv/) would not require a written word-final <e>, underlining that notwithstanding important parallels, graphotactics is not always dependent on phonotactics. This is a vital point in arguing that writing systems not only represent various levels of language but are also their own systems exhibiting distinct features—and it is paramount that they be studied (also) as such.

Despite graphotactics being a core feature of writing systems, research on it—especially comparative—is scarce. This talk aims to systematize various types of restrictions, including, at the subsegmental level, favored positions of semantic and phonological components inside Chinese characters that "determine whether the character is legal or not" (Ho/Ng/Ng 2003: 853). As for larger written units, in many abugidas such as Devanāgarī or Thai, secondary vowel graphemes occur in specific positions around consonants, with some being misaligned in appearing before consonants despite following them in speech (cf. Winskel 2009). A well-known length-related example is the 'three-letter-rule' in English: content words must consist of at least three letters (distinguishing <buy> from <by>).

Crucially, graphotactic regularities are independent of external orthographic regulation and thus not explicit, codified rules but intrinsic to writing systems as self-regulating systems and thus part of users' implicit knowledge. They rely fundamentally on a two-dimensional spatiality that departs from the linear temporality of speech. This makes graphotactics central in arguing that writing systems are indeed systematic in nature.

Berg, Kristian (2016): Graphemic analysis and the spoken language bias. *Frontiers in Psychology* 7: 388. DOI: 10.3389/fpsyg.2016.00388

Ho, Connie S.-H., Ting-Ting Ng & Wing-Kin Ng (2003): A 'radical' approach to reading development in Chinese: the role of semantic radicals and phonetic radicals. *Journal of Literacy Research* 35.3: 849-878. DOI: 10.1207/s15548430jlr3503_3

McCawley, James D. (1994): Some graphotactic constraints. In W. C. Watt (ed.), *Writing systems and cognition* (= Neuropsychology and Cognition, 6), 115-127. Dordrecht: Springer. DOI: 10.1007/978-94-015-8285-8_7

Winskel, Heather (2009): Reading in Thai: the case of misaligned vowels. Reading and *Writing* 22.1: 1-24. DOI: 10.1007/s11145-007-9100-z

Mike Cahill: Usability is not enough: The criticality of sociolinguistic factors in the establishment of new orthographies

There is a stark difference between the relative stability of orthographies which have been standardized for some time, and the fluidity of those which are newly proposed for previously-unwritten languages. The viability of the latter depends on many elements, including usability factors such as the bidirectional match of phonemes to graphemes, visual complexity of graphemes, and graphemes' reproducibility in electronic devices. But at least as crucial as these "objective" factors are the more human-related sociolinguistic factors, common globally, which can either aid an orthography's acceptance, or in extreme cases, prevent its use. These include:

1. Which dialect/s should be represented? Cases from Ghana, Papua New Guinea, Bangladesh, and Indonesia (Clifton 2013, Ring 1989) exhibit different solutions.
2. Should the orthography resemble the national or neighboring languages, or be distinct? Cases from Guatemala, Africa, and Mexico show different local preferences (England 2003, Hollenbach 1969).
3. Are specific orthographies associated with certain religious groups? The Americas and Asia indicate that this can potentially disenfranchise some of the population.
4. In an orthography reform effort, who developed the previous orthography? A respected community leader? An independent outsider? These can promote or block reform.
5. What government policies exist? Even in one continent, Cameroon, Ethiopia, Ghana, and Chad all differ.
6. Script choice: besides Roman, Arabic, and Brahmi-based scripts, recent proposals include the innovative Adlami, Uniskript, Musa, and Hawaiian Creole scripts. Changing the script for an orthography is unusual, but possible (e.g. Turkish, Kazakh, Konso). All the above factors affect this decision (Unseth 2008, Thomas 1989).

These factors, accessible in current situations, undoubtedly existed in the development of orthographies hundreds and thousands of years ago, but obviously are much harder to document. Still, in accounting for the systemization of those orthographies, an attempt should be made to discover what part these forces played.

Clifton, John. 2013. Dialects, Orthography, and Society. *SIL-UND Workpapers*, Vol 53.

England, Nora. 2003. Mayan language revival and revitalization politics: Linguists and linguistic ideologies. *American Anthropologist* 105.4. 733-742.

Hollenbach, Barbara E. 1969. A Note on Concepts of Political Geography. International Journal of American Linguistics, 35(3)

Ring, Andrew J. 1989. Three Case Studies Involving Dialect Standardization Strategies in Northern Ghana. In: Gloria E. Kindell (ed.). 1991. *Proceedings of the Summer Institute of Linguistics International Language Assessment Conference,* Horsleys Green, 23-31 May 1989. Dallas: SIL.

Thomas, Dorothy M. 1989. "Changing the Northern Khmer orthography." Notes on Literacy 57:47–59.

Unseth, Peter. 2008. The sociolinguistics of script choice: an introduction. International Journal of the Sociology of Language 192:1-4.

Leila Schroeder: Bridging an orthography chasm

Orthographic processing has been shown to contribute to reading fluency for students of either high or low ability, and for reading accuracy in good readers (Rakhlin, Mourges, Cardoso-Martins, Kornev & Grigorenko 2019; Probert & de Vos 2016). For most African orthographies, there is a fairly close matching of symbols and sounds. However it is common, across Africa, where the colonial languages are highly prized, for children to be confronted with English reading as well, as early as grade 2. For readers of African languages to transfer to a very deep and inconsistently spelled orthography with completely different sound-mapping associations for its twelve vowel phonemes, necessitates major changes in their decoding and word recognition strategies. Interventions in English reading pedagogy are urgently called-for (Rakhlin, Mourges, Cardoso-Martins, Kornev & Grigorenko 2019), continent-wide. I describe the most common transfer challenges for African learners/readers of English. I suggest strategies to ease the transfer by treating the English orthography as if it were regular and shallow, while developing auditory awareness of English's 12 vowel phonemes, recognition of the larger vowel grain size of English within monosyllabic words, syllable-initial and syllable-final consonant clusters, and of course recognition of "sight words". Any successful readers of English must balance their L1 decoding prowess with additional reading and spelling skills: attention to context clues and memorized spellings for both meaning and pronunciation, using a sub-lexical route to reading (Probert and de Vos 2016; Randall, in Cook and Bassetti, 2005).

Amalia Gnanadesikan: Segments and Syllables in Thaana and Hangul: Comparing literate native-speaker inventions

Korean Hangul and Maldivian Thaana form an illuminating comparison class, being deliberate inventions by literate native speakers. Hangul is well described and widely praised, but Thaana has received little attention. Both scripts represent a sharp break from earlier systems, not only in sign shape but in structural design. Both combine segmental and syllabic information in ways that are rare but similar, and the treatment of syllables shows systematic sensitivity to phonological constraints in typologically rare ways.

Both scripts represent all segmental phonemes but employ two visually distinct classes of signs: consonants and vowels. The scripts use this distinction to encode syllabic information.

Phonologically unmarked CV syllables are preferred in syllabaries (Buckley 2018). In Optimality Theory, the preference for CV syllables derives from the interplay of three universal constraints: ONSET (a syllable must have a consonantal onset), NOCODA (syllables must not end in a consonant), and *COMPLEX (no tautosyllabic consonant clusters). Using CV syllabograms to write more complex syllables can be analyzed as ranking syllabification constraints more highly in the written than the spoken language.

Hangul and Thaana also follow these syllabification constraints, somewhat differently than their spoken languages do. By using a "consonant" letter to indicate "no initial consonant," they obey ONSET strictly, even when the spoken languages do not. Thaana also obeys NOCODA strictly: one of the "vowel" letters is used when there is no vowel, converting a spoken CVC syllable into a <CV-CV> written sequence. Thaana strictly obeys *COMPLEX, but Hangul may violate it under the higher-ranked obligation to maintain a consistent spelling for each morpheme.

The syllable as a phonological unit is fundamental to the history of writing. These two redesigns of writing are fully alphabetic but without abandoning the organizing structure of the syllable. These scripts have produced high literacy rates, even under adverse economic circumstances.

Robert J. Fouser: Searching for the perfect writing system: 20th century Hangeul reform proposals

In this paper, I will investigate 20th-century attempts to reform Hangeul, the indigenous Korean script developed in 1443. Reforms reflect three areas interest: linear writing, revised graphemes, and new typography design. During the first half of the 20th century, advocates of linear writing proposals argued that they fit the "modern standard" set by dominant alphabetic Western languages. Noted advocates included Ju Si-gyeong (1876-1914), the founder of the pro-Hangeul movement and his student Choi Hyeon-bae (1894-1970) who became a leader in South Korean orthographic policy after liberation from Japan in 1945. Advocates of revised graphemes argued that changes in shape would make Hangeul more scientific, more readable, more attractive, or more adaptable to technology, such as typewriters and later computers. Choi Hyeon-bae drew on Ju's linear writing to develop his own system that included revised graphemes that looked more like the Roman alphabet. A North Korean proposal that was adopted as official orthography from 1948 to 1954 added six new graphemes to make Hangeul more consistent with Korean morphology. A proposal by Seo Yun-yeong in 1987 created new Hangeul-inspired graphemes. Interest in linear writing largely faded by the 1960s, but the spread of computers beginning in the 1980s simulated the development of new typography design, such as Kim Jeong-su's "leaning Hangeul" (1982) and fonts developed by Ahn Sang-soo beginning in the 1980s.

The paper will be based on a review of secondary literature and a close examination of noteworthy original proposals. I will show how reform proposals, in their search for perfection, drew heavily on various, often conflicting, perceptions of Hangeul's position among writing systems of the world.

Keynote presentation

Peter T. Daniels: When is non-writing writing? or, when is writing non-writing?

It was easy to say that writing was invented out of nothing three times (that we can be sure of), in Sumer, China, and Mesoamerica. That syllables were important in those inventions emerged from attention to modern inventions of writing. But in recent years, specialists in Sumerian and Mayan texts have been uncovering details about the development of cuneiform and glyphs that, perhaps surprisingly, prove to be comparable and mutually illuminating: in both cases, it seems legitimate to say that the earliest forms did not yet represent the actual writing of specific languages (Writing: A system of more or less permanent marks used to represent an utterance in such a way that it can be recovered more or less exactly without the intervention of the utterer [Daniels, passim].). Chinese writing, even though we cannot observe any essential changes in the system from its earliest known examples to the present day, also proves to have something to contribute, in the wake of a comparative study of how Chinese writing was adapted for writing neighboring (and unrelated) languages. And even half a century of experience with a semiotic system designed to be alinguistic, "Blissymbolics," has something to tell us about non-writing turning into writing.

Symposium: Writing systems of the Americas

Michael Carrasco; Joshua Englehardt: Enduring literacies: The emergence and maturation of writing systems in Ancient Mesoamerica

During the Middle (900‑400 BCE) to Late Formative (400 BCE‑250 CE) periods in Mesoamerica, an eclectic array of scribal traditions emerged across the region. These writing systems seemingly developed from initial contact with an ancestral "Olmec" iconographic and script system, whose literary forms, representational conventions, themes, and narrative genres persisted within the visual rhetoric and poetics of later iconographic and writing traditions. In this paper, we critically explore a series of specific signs, conventions, and tropes related to ritual, sacrifice, fertility, and kingship that are alluded to in the earliest Middle Formative period inscriptions, such as that on the Cascajal Block. Through tracing the lineages of these visual elements and their associated themes across regional writing systems, such as those of the Zapotec, Maya, and Postclassic Central Mexican cultures, we attempt to better contextualize the enduring scribal, artistic, and literary genres that formed a fertile ground for the emergence of regional script traditions, thereby gaining a more robust understanding of the earliest manifestations of Mesoamerican writing and the paths that led the peoples of ancient Mesoamerica to adopt and adapt shared ideas to their own ends in a diversity of sophisticated writing systems.

Symposium: Writing systems of the Americas

Tomi S. Melka; Robert M. Schoch: *T'oqapu* patterns on Inqa textiles and other media: Do they constitute a writing system?

The classic and best-known *t'oqapu* (= *tocapu*) consist of small multi-colored square units, set in a band-like or grid-like structure. The most common artifacts on which they occur are the Inqa- made or Inqa-inspired fine tapestry tunics of the apogee of Inqa (= Inca, Inka) civilization (= Late Horizon period of the Andean cultures), *circa* 1476–1532/34 CE. *T'oqapu* motifs are also found on other types of textile products, such as woven bags and pouches, as well as on some types of non-textile artifacts (e.g., on pottery or masonry designs).

The study of Inqa *t'oqapu* motifs has witnessed a steady increase over the last several decades, producing a good number of articles that can hardly go unnoticed. However, there is as yet no agreement on how to understand (and thus "interpret") this apparent semiotic system. Reviewing the current status of the field, we ask in what respects the underlying mechanisms of the *t'oqapu* system are similar to and different from those of other communicative systems. In addition, we review various plausible ideas regarding the inception and perfection of the *t'oqapu* geometric layout, and their practical implementation in pre-European times. For instance, *t'oqapu* motifs may be compared to heraldic devices of Medieval Europe. A modern analogy to *t'oqapu* may be the use of symbols such as the "peace sign" (a 1958 symbol of protest against nuclear weapons that gained popularity in the 1960s) and electronic *emoji* symbols that arguably are forming an international visual language system.

Making sense of *t'oqapu*, apart from determining their specific nature, assists in understanding better an important aspect of Tawantisuyu (= Inqa Empire), that of textile production, fashion design, and the concomitant social stratification. In a broader context, it may shed light on other unidentified symbols and aid in recognizing the fine lines and intersections that exist between phonetic and non-phonetic sign systems.

Symposium: Writing systems of the Americas

John W. Adams†: Some notes on Gitksan crests, potlatching, and law

† Presented by Alice B. Kasakoff; Bill Proudfit; John Justeson

John W Adams, 1929-2013, made several presentations on Gitksan totem poles as a system of visual symbols which he learned to interpret during 18 months of fieldwork in the mid-60s. They are erected by indigenous inhabitants of the northwest coast of North America. Their stated purpose is to commemorate recent chiefly and noble dead; they are erected by the heirs to the deceased's prerogatives upon a new chief's succession. In this version, he focused on their relationship to Gitksan law. He never published a paper.

The selection and placement of motifs on Gitksan totem poles is systematic; this paper suggests that (and how) the crests they carve on their poles, the feast at which the crests are validated, and the law binding man and nature are related. Also systematic is the relationship among the visual symbols displayed on totem poles to other visual symbols, and to both public and more private events.

The symbol system on totem poles is part of an artistic complex that is not self-contained but used and interpreted pervasively within Gitksan culture. The symbols that appear are applied, widely, to other objects, and are used in other performances. Language is a part of this system, but it is not language that is represented by its visual symbols. Totem poles and the symbols they visually present therefore participate in a broad system of visual communication that is akin to writing systems, in particular to those characterized in the literature as "semasiographic", while also serving as part of a broader representational system that operates across other situations and modalities.

Symposium: Writing systems of the Americas

David Mora-Marín: A survey of systematic spelling practices and conventions in Mayan writing

This paper presents a classification of signs and spelling practices employed by Mayan scribes during the Late Preclassic (300 BCE-CE 200) and Classic (CE 200-900) periods. Conventionally, Mayan writing has been characterized as a logosyllabic script, utilizing mainly two basic types of signs: logograms and syllabograms, the latter primarily of the consonant-vowel type (CV). This paper (see **Table 1** below) proposes that scribes also employed semantic determiners more widely and systematically than previously thought, reviews evidence for other types of signs that were contrastive and optional (e.g. duplication dots) as well as non-contrastive and optional (e.g. numeral "fillers"), and redefines the so-called semantic classifiers proposed by Hopkins and Josserand (1999) and Mora-Marin (2008) as iconographic classifiers.

The second part of the paper focuses on systematic spellings and orthographic conventions. The present classification (see **Table 2** overleaf) follows the approach to spelling conventions outlined in Mora-Marin (2001, 2005, 2010) and Kaufman with Justeson (2003), themselves influenced by authors conceiving of spelling practices as guided by usage-based pressures (Bricker 1985, 1989), as well as phonological and morphological constraints (Justeson 1989; Hopkins 1997). The present classification represents a synthesis of these approaches, with some condition-generalizing additions: seven basic spelling practices, some of which can be further subcategorized. In the process, two previously understudied conventions are defined and analyzed: consonant-insertion ligatures and vowel-insertion ligatures. The classification proposed here makes the so-called "Disharmony Hypotheses" unnecessary (e.g. Houston et al. 1998, 2004; Lacadena and Wichmann 2004): such hypotheses suggest that "fictitious" vowels of word-closing (in fact, syllable-closing) spellings functioned as diacritics, indicating the vowel complexity of a preceding syllable.

The present approach shows that these hypotheses are highly inconsistent, that they often contradict the comparative historical linguistic evidence, and that spelling practices can be better explained in terms of linguistic (phonological and morphological conditioning) and literacy (ease of writing vs. ease of writing economization) factors.

Table 1. Classification of major signs

1. Logograms
2. Syllabograms
3. Semantic determiners
4. Iconographic classifiers
5. Reduplication diacritic
6. Bar-and-dot numeral fillers
7. Allograms and allographs

Table 2. Classification of major spelling conventions

1. Polymorphemic logography
    a. Nouns
    b. Verbs
2. Phonetic complementation
3. Consonant deletion
4. CV deletion
5. Fictitious vowel insertion
    a. Obligatory synharmony due to phonological context
        1. C_1VC_1 roots and sequences
        2. CV7 roots and sequences
        3. CVCC... sequences
        4. Exceptions: Vowel-insertion ligatures (following proclitic or prefix)
    b. Obligatory synharmony at morpheme boundary
        1. ...CVC-V...
    c. Vowel conditioned by vowel of typical -VC... suffix
6. Consonant insertion
    a. Phonetic complementation
    b. Word-boundary consonant-insertion ligature (preceding suffix or enclitic)
7. CVC syllabograms

Dorit Ravid; Rachel Schiff: Hebrew affix spelling in children with developmental dyslexia

The current study explored the effect of developmental dyslexia on Hebrew affix spelling. Modern Hebrew homophony results from obsolete phonological distinctions reflected in different graphemes - e.g., spelling *t* by both ת and ט. Homophonous affix spelling may be aided by the fact that some graphemes do not have affix roles. For example, ת, but not ט, serves as an affix letter. Five factors can affect acquisition: (1) Affix type frequency; (2) Root-envelope transparency; (3) Grapheme morpho-orthographic prevalence; (4) morphological competitors of grapheme; (5) Phonological-orthographic consistency. Lower values on these criteria were expected to challenge younger and non-typically developing participants.

Performance on affix spelling was tested in a group of 21 6th graders with developmental dyslexia (DD) compared with two groups of typically developing (TD) children – 23 2nd graders with matched reading age (RA) and 24 6th graders with matched chronological age (CA). The spelling task had 244 target words with affix letters, presented in the context of short sentences.

Findings showed lower spelling scores in the DD group compared to the two TD groups. The 6th grade DD group was able to overcome lower type frequency, consistency, and morphological competition to the same extent as the 2nd grade RA group; but they were less able to demarcate affixes from roots, and also relied more on letter prevalence than both TD 6th grade RA and 2nd grade CA groups. These findings indicate that knowledge of how phonological and morphological roles are deployed and interact is critical in learning to spell homophonous affix letters. We show that in Hebrew, 6th graders with DD cannot make adequate use of morpho-phono-orthographic cues in learning to spell homophonous affix letters – unlike TD age-matched peers, and much like TD 2nd graders. These results join the general literature in pointing at DD as a linguistic impairment.

Martin Neef: On hidden reforms of the German official orthography

In this presentation, the concept of a spelling reform will be discussed using German as an example, combined with reflections on the general conditions for a successful implementation of reforms. The spelling of the German language has the peculiarity that it has an official, state-regulated status. This official orthography has existed since 1901 and applies to the entire German-speaking community, at least in written communication in the context of public institutions such as schools and the legal system. The entire 20th century was marked by attempts to reform this original standard orthography. According to official announcements, the first reform was enacted in 1998 (followed by a second one in 2006). However, two events in the middle of the 20th century can be interpreted as hidden reforms, being of quite different nature.

During the National Socialist era, Minister of Education Bernhard Rust made three attempts at reform. The first two were blocked by the Ministry of the Interior, while the third attempt from 1944 was enacted, at least briefly, before apparently being withdrawn. After the end of the rule of National Socialism, the actual situation of official spelling was somewhat confused. During this period of ambiguity, the private Duden publishing house succeeded in gaining a government-approved position. It was decided that in cases of doubt, the spellings and rules formulated in the Duden dictionary would be binding. Since central aspects of German spelling, such as separate and compound spelling and punctuation, were not included in the official rules of 1901, but were in the Duden, this move represents a significant reform of German spelling, although it was never marked as such. The Duden's official status ended in 1998, but Duden still enjoys a privileged reputation in public opinion.

Lynne Cahill: A corpus analysis of the spread of standardised spelling in Middle English

The standardisation of English spelling is known to have happened around the fifteenth and sixteenth centuries, led by the rise of Chancery English, starting in the London/Cambridge/Oxford triangle. The evidence for this comes from a wide range of documents of different types, including literature, personal correspondence and legal documents. The regional variation found across England has previously been analysed in the Linguistic Atlas of Medieval English by looking at Middle English literary manuscripts (McIntosh et al 1986). One of the problems with this approach is that it is hard to localise these texts with confidence.

The emergence of large corpora of definitively localised legal documents opens up the possibility for much more reliable and fine grained analysis of how the standardised forms spread. In particular, the MELD (Middle English Local Documents) corpus, being developed at the University of Stavanger (Stenroos et al 2017) provides transcriptions of large numbers of documents ranging in time from 1406 to 1525.

This paper examines the variation in spelling of the word *said* (including forms such as *aforesaid*). Previous work has examined the move to the use of digraphs in the spelling of English vowels, but the MELD corpus permits much more detailed examination of the development of the variant digraphs used over space and time. Specifically I look at the four main variants found in the Eastern counties: <said>, <sayd>, <seid> and <seyd>. The results of the analysis show that the two parts of the digraph, <a> vs <e> and <i> vs <y> emerge at different rates, <i> becoming fixed earlier than <a>, and that Cambridgeshire, Hertfordshire and Middlesex move to the standardised <ai> form earlier than Suffolk, Norfolk and Berkshire.

McIntosh, Angus, M. L. Samuels and Michael Benskin (1986) *A Linguistic Atlas of Late Medieval English* Aberdeen: Aberdeen University Press.

Merja Stenroos, Kjetil V. Thengs and Geir Bergstrom (2017), A Corpus of Middle English Local Documents (MELD), version 2017.1. University of Stavanger. http://www.uis.no/meld (Accessed: 10/08/2021)

Martin Uildriks: A world-altering technology: Script development in Predynastic Upper Egypt

Traditional scholarship traces Egypt's earliest language to two script systems found on funerary goods from the tail end of the Predynastic period, ca. 5,200 years ago. These two script systems, hieroglyphic and a cursive related form (hieratic), remain marginally understood and are usually analyzed in binary opposition, following Saussure's semiotic model of image (graphic) and word (concept). However, the Predynastic period shows evidence for approximately 700 years of graphic development prior and traces of incipient script development begin at 3,900 BCE. Rock drawings as well as painted decorations and markings on ceramics in particular suggest a variety of increasingly complex Predynastic linguistic profiles that may have encoded different regional dialects or languages.

I begin this paper with a consideration of Predynastic evidence from one specific region in Egypt to argue for script development on regional and local levels, rooted in exchange and connectivity of highly mobile communities, earmarked by local knowledge and complex social practices. I will then compare this evidence with contemporary evidence from other regions to suggest emerging shared and standardized visual vocabularies, adapted to various media and transfigured in social arenas to reflect particular Predynastic identities.

In doing so, I aim to 1) show that Predynastic script emergence was an internal and cross-regional phenomenon that extended across a much larger area than hitherto supposed; 2) demonstrate that the seemingly homogeneous late Predynastic imagery found in funerary contexts may have encoded linguistic variety imperceptible to standard linguistic analysis; and 3) dislodge a fundamental principle of Egyptian grammar by proposing an inversion that does not separate graphic from concept or indeed the Predynastic period from what comes after. Instead, my model develops a different logic that allows for a fluent transition from Predynastic 'images' to genuine recordings of speech.

Rosso Manuel Senesi: Latin <X>: seeing double

This talk examines a specific case of orthographic variation throughout the long-lasting period of standardization of the Latin language: the use of alternative spellings to <X> for the /ks/ cluster in epigraphic documents. The underlying reasons for such variation will be examined through a quantitative and qualitative analysis of the attested allographs: <XS>; <CX>; <CS>; <XX> and <CXS>, found in the corpora CLaSSES (http://classes-latin-linguistics.fileli.unipi.it/) and EDR (http://www.edr-edr.it/edr_programmi/res_complex_comune.php) from III BC to IV AD. The oldest and most frequent variant is <XS>. Hence, this digraph is the focal center of what Pandolfini & Prosdocimi (1990: 159) call "adjustments of the usage rules" of a writing system, which is tantamount to being the hub of an orthography reform. Adopting Pandolfini & Prosdocimi's (1990:159) terminology, the diastratic distribution of <XS> will show how subtle is the line differentiating "dirigiste" reforms by central authority stakeholders from reforms stemming from the slow and multicentric evolution of internal and external forces. Therefore, a multiple causation hypothesis will be proposed. As first suggested by Mancini (2019: 47), the use of <XS> could be related to the phenomenon of *geminatio consonantium*, which is considered a top-down modification of the orthographic system propelled by the philhellenic educated class willing to emulate Greek orthography. However, such centralized reform, involving <XS> and orthographic geminates, has probably grafted itself onto two multicentric forces already active at lower diaphasic and diastratic levels. First, the incipient prosodic drift of Latin altering the interface rules between graphematic and phonological structures could have been an internal pressure to change orthographic representations in the speaker's mind. On the other hand, it is possible to assume that language contact with Greek acted as an external force on different diastratic levels. Indeed, digraphs for /ks/ are widespread in non-canonical Greek spellings, with coeval attestations in the Italian peninsula.

Corinna Salomon: Alphabet systemisation in North Italic writing cultures

Around the late 7th century BC, the Etruscan alphabet found its way to the inhabitants of Transpadanian Italy and the Alpine region; it emerges in the course of the 6th century in the form of alphabet variants which denote the different languages of the local speaker communities – the Italic Venetic, the Celtic Lepontic, the Tyrsenian Raetic, and possibly also Camunic of as yet uncertain affiliation. These "North Italic" alphabets are very similar to each other, indeed widely indistinguishable in the absence of certain shibboleth letters or letter forms, and it is generally assumed that not all com-munities acquired writing directly from the Etruscans, but to some extent from each other, and that there was a good deal of secondary interference between the variants also across language borders.

Considering this, it is surprising how differently the evidence we have of North Italic literacy presents itself to us in terms of writing culture. We have clear evidence for highly institutionalised use of alphabets, each with their own specific uniform orthographies, at a handful of sites, while other parts of the North Italic area display a high level of local and chronological variation with no indica-tion of targeted standardisation.

The talk will give an overview of the North Italic alphabet variants and their (putative) relations to each other, and discuss possible reasons for the differences we see in terms of systemisation. What bodies are responsible for standardisation, how do they operate, and why are they only active in certain places? Consideration will also be given to questions of methodology – how does institutionalised standardisation manifest in the documents in opposition to gradual, bottom-up developments toward more systematic orthographies, and how do we tell them apart – and terminology – what do we call a discreet script/alphabet/orthographic system in a scriptscape as fragmented as the North Italic one?

Keynote presentation

Min Wang: Phonology beyond phoneme: Contribution of suprasegmental information to reading

Phonology encompasses segmental and suprasegmental information. Although both segments and suprasegmental information are salient phonological properties of spoken languages, previous research has focused heavily on segmental processing during reading; there has been relatively less research that has addressed the role of suprasegmental information. In this talk, I will present a line of research work from my lab that seeks to address the following questions central to our understanding of the role of suprasegmental information: 1) What is the relative contribution of segmental vs. suprasegmental information to reading? 2) How do children develop their sensitivity to suprasegmental information and how does the development of such sensitivity contribute to learning to read? 3) How do second language learners represent and process novel suprasegmental information in reading? 4) Is there a cross-language contribution of suprasegmental awareness in learning to read a second language? Findings from our work on native Chinese-speaking adults suggest that both segmental and tonal information play a role in lexical access; however, segmental information carries more weight than tonal information. Among 6-year-old native English-speaking children, stress sensitivity made a unique contribution to word reading over and above oral vocabulary and phonemic awareness. Children as well as adults showed better phonemic compared to stress sensitivity. Nonnative Chinese adult learners showed a greater challenge in representing tonal information in visual character recognition compared to segmental information. Finally, native Chinese-speaking children tap into their Chinese tonal awareness in learning to read English as a second language. Taken together, our findings highlight the important role of suprasegmental information in reading among diverse learners including young and skilled readers as well as second language learners. Current reading theories ought to take into account the functions of both segmental and suprasegmental information and their relative weights.

Hana Jee; Monica Tamariz; Richard Shillcock: Quantifying sound-graphic systematicity and application on multiple phonographs

Language, as a complex system, suggests coordination between subsystems. Recent studies demonstrated the meaning-form mapping existing in various languages (Blasi et al., 2016; Dautrich et al., 2017; Monaghan et al., 2014; Tamariz, 2008). The current research, for the first time, quantified form-form mapping between orthography and phonology. The letter-sound correlation in seven phonograms was examined: Arabic, Cyrillic, English, Finnish, Greek, Hebrew, and Korean. We measured all the pairwise phonological distances between phonemes in the respective alphabet system, and the corresponding pairwise orthographical distances between letters. For the phonological distance, phonemes were encoded into vectors according to the articulatory features of International Phonetic Alphabet and the distance between the vectors were calculated in various ways. Meanwhile, each letter was saved as an image file (PNG) and the distance between two letter images was measured by Hausdorff distance (Huttenlocher et al., 1993). The significance of the correlation between these two lists of pairwise distances was tested with a Monte-Carlo permutation test. The conventional orthographies generally returned small but significant positive letter-sound correlations in certain fonts: the similar letters tend to have similar sounds. Korean returned the highest correlation as expected ($r = 0.37$, $p < 0.00$) — it was artificially invented in the 15$^{\text{th}}$ century. The fine phoneme-grapheme match seems to be the reason for the correlation, which corresponds to Pitman's shorthand ($r = 0.35$, $p < 0.00$) and Shavian alphabet ($r = 0.20$, $p < 0.00$). No correlation was found in fictitious orthographies such as Aurebesh (Star Wars) and Klingon (Star Trek). By using different visual distance metrics we distinguish between symbolbased (Korean, shorthand) and effort-based (English) graphophonemic systematicity. We discuss the implications for education, and for understanding typical and atypical reading.

Keisuke Honda: Types of 'logography' in the Japanese writing system

This presentation aims to highlight the need to distinguish different types of what is traditionally labelled as 'logography' in the present-day Japanese writing system.

Logography is widely considered a key principle of writing that is categorically different from phonography or phonology-based writing. However, views vary on what counts as logography and what it means for the linguistic study of writing systems. Besides, some researchers question the appropriateness of this term and use alternative designations based on different sets of assumptions. These points have crucial implications for our understanding of the Japanese writing system: While many studies agree on the importance of logography in this system, the lack of conceptual and terminological clarity makes it difficult to examine the validity and significance of such a characterisation.

This talk will propose to tackle the above issue by identifying distinct types of writing in Japanese that are commonly subsumed under the term 'logography'. The argument is fourfold. Firstly, logography includes two broad types of lexically specified writing, one pertaining to individual written signs (愛 *ai* 'love'; を *o* 'accusative'; W *daburu* 'double') and the other to written lexical items (e.g. 多摩川 *tamagawa* 'river name' versus 玉川 *tamagawa* 'district name'; バレー *barē* 'volleyball' versus バレエ *barē* 'ballet'). Secondly, the first type is further divided into such subtypes as monomorphemic (e.g. 芋 *imo* 'potato'), polymorphemic (e.g. 杯 *saka+zuki* 'drinking cup'), and syntactic (e.g. 掌 *te=no hira* 'palm'). Thirdly, a qualitative and quantitative analysis of these types and subtypes is required to provide an accurate account of the logographic aspects of the Japanese writing system. Finally, logography should be redefined as an umbrella term covering all these aspects unless a more appropriate term can be found.

Terry Joyce; Hisashi Masuda: Systematic chaos or chaotic systems? Some ponderings on the complexity of the Japanese writing system

The Japanese writing system (JWS) has an unrivaled reputation of being the most complex in modern use and even stands as a contender for the most complicated ever devised (Joyce, 2013). Yet, as Yamada (1967) contends, by virtue of its complexity, it unquestionably warrants special attention within writing systems research. Some of its complexity comes from the fact that the JWS is an intricate amalgamation of component systems—kanji, hiragana, katakana, and rōmaji—which, although employed in complementary ways, simultaneously afford considerable levels of graphematic variation (Joyce & Masuda, 2018; 2019).

Another source of the JWS's complexity also involves a merging of two distinct systems; namely, the Sino-Japanese (SJ) and Native-Japanese (NJ) readings that are both associated with morphographic kanji. This gives rise to the JWS's interesting forms of allomorphy, such as where 水 is pronounced as /mizu/ when representing the simplex NJ morpheme 'water', but as /sui/ when representing its SJ allomorph in the compound word 防水 /bō-sui/ 'waterproof'. As this is also of special relevance for understanding writing systems, this paper discusses some of the morphological processes underlying three-kanji (防水布 /bō-sui-fu/ 'waterproof cloth') and four-kanji SJ compounds words (防水加工 /bō-sui-ka-kō/ 'waterproof finish'). More specifically, this paper argues for greater recognition being accorded to the morphographic nature of kanji and the sophisticated forms of allomorphic representation that it entails, so that systematic nature of the JWS may be more objectively assessed. This argument is motivated from three closely related notions: (1) kanji are employed in graphematically representing a considerable proportion of the Japanese lexicon (Joyce, Masuda & Ogawa, 2014); (2) kanji aid readers in capturing lexical meanings (Kobayashi, Yamashita & Kageyama, 2016); and (3) the ultimate goal of reading is to map symbols to meanings (Rastle, 2019).

Brian P. Bennett: The impact of Unicode: The case of Church Slavonic

Unicode "is designed to meet the needs of diverse user communities within each language, serving business, educational, liturgical and scientific users, and covering the needs of both modern and historical texts" (Unicode Consortium 2020: 14). There is debate about whether, on balance, the technology fosters or hinders the world's scriptal diversity (e.g., John 2012, Kamusella 2013). This paper seeks to contribute to the discussion by examining the impact of Unicode on Church Slavonic, the liturgical language of Russian Orthodoxy. Church Slavonic is historical but also modern, since new devotional texts are still composed in it (Dobrushina and Polyakov 2013). It has educational and even commercial uses (the letterforms sometimes function as a symbol or 'logo' for Russianness). Drawing on sociolinguistics and religious studies, this paper outlines the revival of the Church Slavonic tradition in post-Soviet Russia (Bennett 2011). It then summarizes some of the complexities and controversies involved in encoding Church Slavonic, a script with multiple diacritics, ligatures, *nomina sacra*, etc. (Andreev, Shardt, and Simmons 2015). Finally, it considers the reach of Unicode-conformant Slavonic in ecclesiastical, academic, and other contexts. The case study suggests that assessing the impact of Unicode is by no means straightforward.

Andreev, Aleksandr, Yuri Shardt, and Nikita Simmons. 2015. "Church Slavonic Typography in Unicode." https://www.unicode.org/notes/tn41/tn41-1.pdf

Bennett, Brian P. 2011. *Religion and Language in Post-Soviet Russia*. London: Routledge.

Dobrushina, E. R. and A. E. Poliakov. 2012. "Korpus tserkovnoslavianskogo iazyka: vozmozhnosti, metody cozdaniia, perspektivy." *Vestnik PSTGU*, III: Filologiia. Vyp. 1 (31): 32-44.

John, Nicholas A. 2013. "The Construction of the Multilingual Internet: Unicode, Hebrew, and Globalization." *Journal of Computer-Mediated Communication* 18: 321-38.

Kamusella, Tomasz. 2012. "The global regime of language recognition." *International Journal for the Sociology of Language* 218: 59-86.

Unicode Consortium. 2020. http://www.unicode.org/versions/Unicode13.0.0/ch02.pdf

John A. Bundschuh: Predicate marking strategies in early Japanese vernacular glossing: A preliminary study

Until the 8th century C.E., the Japanese language was written using only Chinese characters. However, by the mid-9th century, systems of vernacular glossing, known as *kunten*, came to be used in Buddhist temples across the country when rendering Chinese sutras and other Buddhist texts into Japanese. Although the marking schemas varied by sect, temple, and lineage, there was uniformity in the ways in which predicate morphology was represented. For example, a glosser could use a *wokototen* , or morpho-syntactic gloss, to denote tense, aspect, modality, and other grammatical categories, phonetic glosses to essentially spell out the predicate, or leave the underlying Chinese text unmarked when the proper Japanese reading is clear from the context, such as when a negation marker is present in the Chinese.

It is well known among scholars that this glossing practice was an important step in the development of the Japanese writing system we find used today (Tsukishima 1982 (1969), Frellesvig 2010, Whitman 2011, Alberizzi 2014), yet research on the orthographical choices made during its inception remains limited. Presenting numerous examples of the above predicate marking strategies from Chinese Buddhist texts glossed in 9th-century Japan, this preliminary study finds morpho-syntactic glossing to be the most prevalent, followed by unmarked and phonological glossing strategies. It argues this proportion is due to the sociolinguistic milieu of glossing in the early Japanese Buddhist context—that of marking a Chinese text as it is recited vernacularly by a lecturer—in which the speed of inscription was imperative.

Alberizzi, Valerio Luigi. 2014. "An Introduction to *kunten* Glossed Texts and Their Study in Japan." *Les dossiers de HEL: Société d'Histoire et d'Epistémologie des Sciences du Langage* 7: 1–9.

Frellesvig, Bjarke. 2010. *A History of the Japanese Language* . Cambridge: Cambridge University Press.

Tsukishima Hiroshi. 1982 (1969). *Heian jidaigo shinron* [A New Treatise on the Language of the Heian Period]. Tōkyō: Tōkyō daigaku shuppan kai.

Whitman, John. 2011. "The Ubiquity of the Gloss." *Scripta* 3: 1–36.

Kevin Heffernan; Yo Sato: Predictors of script choice in Japanese: A data-driven study

Japanese is written using three scripts: a logographic script known as kanji, and two syllabaries known as hiragana and katakana (collectively referred t o as kana). Japanese writers are hence given a choice, and tendencies vary from word to word, as shown in the examples below.

(1) ✔ 雨 / ✗ あめ / ✗ アメ (mostly kanji)
(2) ✗ 林檎 / ✔ りんご / ✔ リンゴ (mostly kana)
(3) ✔ 癖 / ✔ くせ / ✔ クセ (both kanji and kana)

Seeley (2000) points out a general trend away from kanji and towards kana but notes that this so-called "kana-ization" may depend on the genre. Shibatani and Kageyama (2016) state that kanji tends to be used for content words and kana for grammatical words. The stroke count of the kanji and the frequency of the constituent kanjis also influence the choice (Kaiho and Nomura, 1983).

To give concrete grounding to these claims, we conducted a logistic regression analysis on the genre/year-classified, PoS-annotated data in the *Balanced Corpus of Contemporary Written Japanese* (Maekawa et al., 2014).

We found the following:

1. Genre showed significant differences: the greatest tendency towards kanji is in social science, the least in magazine articles.
2. Lexical words are written in kanji more often than grammatical words.
3. Stroke count and frequency of kanji significantly correlated with script choice.
4. Author birth year did not significantly correlate with script choice.

Our results suggest that the choice reflects a complex system that balances effort against communicative efficiency, since kanji requires more cognitive effort to write but is easier to parse, while the opposite characteristics apply f or kana. Furthermore, this balance seems to be at an equilibrium and not shifting over time, as claimed by Seeley.

Itō, J., & Mester, A. (1999). The phonological lexicon. I n N. Tsujimura (Ed.), *The handbook of Japanese linguistics* (pp. 62-100). Malden, MA: Blackwell.

Kaiho, H., and Nomura, Y. (1983). *Kanji jouhou shori no shinrigaku* [The psychology of information processing for kanji]. Kyoiku Shuppan: Tokyo

Maekawa, K., Yamazaki, M., Ogiso, T., Maruyama, T., Ogura, H., Kashino, W., ... Den, Y. (2014). Balanced corpus of contemporary written Japanese. *Language Resources and Evaluation, 48*(2), 345-371. doi:10.1007/s10579-013-9261-0

Seeley, C. (2000). *A history of writing in Japan.* Honolulu: University of Hawai'i Press.

Shibatani, M., & Kageyama, T. (2016). Introduction to the handbooks of Japanese language and linguistics. I n T. Kageyama & H. Kishimoto (Eds.), *Handbook of Japanese lexicon and word formation* (pp. vii-xxxiii). Berlin: de Gruyter Mouton

Eleonora Selvi: A story without an end. The never reached systematization of the Pamphylian alphabet

The Pamphylian alphabet is a peculiar Greek alphabet attested in Pamphylia between 6th and 2nd cent. BCE by brief inscriptions. It had its own peculiar writing habits and two additional letters not known to any other Greek alphabet. Due to the interplay of different forces, both phonetic and socio-cultural, it never reached an orthographic systematization.

This research aims to understand how the unfinished process of standardization and orthographic systematization of the Pamphylian alphabet unfolded.

The Pamphylian Greek dialect was heavily influenced by Anatolian (Luwic) adstrata and substrata, especially in syntax and phonetics (cf. Skelton 2017). The Greek alphabet had to be adapted to write Anatolian personal names (cf. Brixhe 1999, Dardano 2012). Moreover, Pamphylian alphabet was surrounded by Anatolian alphabets like Sidetic, Lycian, Lydian etc., resulting in a pervasive orthographic uncertainty, as different solutions to accommodate Anatolian phonetics and Anatolian writing habits into the Greek template were tried. From the 4th cent. BCE onwards, the Pamphylian alphabet, while being a marker of Pamphylian identity, was also subjected to the pressure of homogenizing with the Greek *koiné* alphabet.

To understand how these different forces interacted to ultimately prevent a standardization of the Pamphylian alphabet and keep open the possibility for the writer to choose between different solutions, an anthropological approach to writing will be adopted, based on Prosdocimi (1990) and Rizza's (2014) works on selection and re-functionalization of the alphabetic '*corpora doctrinae*' in specific communities.

Two case-studies of non-standardized elements will be analysed more specifically: 1) the letter И, whose phonetic value was always ambiguous and made it interchangeable with other signs, being likely less important than its specific 'Pamphylian' character as an identity marker; 2) the representation of the *glide*, which is almost absent in other Greek alphabets, but common in Pamphylian, due to the proximity of Anatolian alphabets.

Hao Sun; Yanwei Jin: Probabilistic grammars in open-class writing systems: A case study of Chinese characters

Writing systems play a very important role in human languages, but little is known as to the mathematical systematicity of writing systems. Here, we propose a novel dichotomy of open-class and closed-class writing systems, capturing distinct mental representations of human language writing systems. An *open-class writing system*, for example Chinese characters, consists of a set of expandable basic units while the units in a *closed-class system* are relatively fixed, like Latin alphabets.

We aim to reveal the mathematical nature of open-class writing systems, specifically Chinese as a case study. We argue that probabilistic grammars underlie the representation of Chinese writing, by generating Chinese characters under a *probabilistic context-free grammar* (PCFG) framework. We expand and formalize traditional classifications of Chinese characters as a grammar of character *shapes* and *components*. Each character is expanded as a tree structure with shape paradigms (e.g. horizontal partition or triangle) as *nonterminal nodes* and components conditioned on each shape type as *terminal nodes*. Probabilities of grammar rules are estimated using MLE with smoothing from a character treebank that we manually constructed, covering the most frequent 3,500 Chinese characters. Exploratory analysis reveals skewed distributions of rule classes where a few prototypical shapes or components take the lion's share of probability distributions, reminiscent of information-theoretic efficiency in language encoding (Zipf, 1935). Then, 200 characters are randomly generated following rule probabilities of characters, among which 29% are real characters, even though not attested in the treebank. 78% of the noncharacters in samples are rated as acceptable, with an inter-annotator correlation of 0.47 ($p<0.05$).

Chinese writing system seems to show generative powers similar to PCFG, suggesting fundamental differences between open-class and closed-class writing systems. Given that PCFG is a common computational model of syntax, our success in applying PCFG to writing systems uncovers correspondences between graphemic and linguistic representations.

Constanze Weth: Syntactic markers - a systematic category in writing systems

The presented paper summarizes findings about a systematic category in writing systems, "syntactic markers" (Weth, 2020). Syntactic markers are graphemic elements in alphabetic writing systems that systematically indicate (morpho-)syntactic information that phonology does not provide. Examples are given for English, French, Dutch, and German.

The paper argues that syntactic markers are a specific category in writing systems as they are part of the orthographic word but also systematically tied to the presence of syntactic features above the word level. Thus, syntactic markers should be distinguished from "morphological markers", comprising inflection and derivational morphemes.

Syntactic spelling does not only refer to the correct spelling of a word but to its agreement within a given syntactical context. In syntactic reading, readers must notice the marker and interpret it correctly to understand the sentence (Funke & Sieger, 2012). Syntactic spelling and reading have been found to be, on the one hand, easy to grasp in the early years of literacy education and, on the other hand, cause difficulties even for literate adults (Fayol, Hupet, & Largy, 1999). In fact, syntactic information is often not decisive for sentence understanding, since the information can be deduced from the context. The more the syntactic structure is decisive for spelling or reading comprehension, the more the markers seem to present difficulties.

Fayol, M., Hupet, M., & Largy, P. (1999). The acquisition of subject-verb agreement in written French: From novices to experts' errors. *Reading and Writing*, 11(2), 153–174.

Funke, R., & Sieger, J. (2012). Continued access to syntactic information in reading. *Reading and Writing*, 25(7), 1769–1794.

Weth, C. (2020). Distinguishing syntactic markers from morphological markers. A cross-linguistic comparison. *Frontiers in Psychology*, 11, 1–11.