

Association for Written Language and Literacy
Ассоциация письменного языка и письменности
Ενωση για τη γραπτη γλωσσα και τη βασικη εκπαιδευση
书面语言和识字协会
Association for Written Language and Literacy

AWLL Newsletter: Number 15: 15 February 2024

[Click on topic to navigate to relevant section; click on text of any section banner to return to content listing]

- On distribution of 15th AWLL newsletter [Terry Joyce (newsletter editor)]
- Conference report: AWLL14 [Organizers: Terry Joyce, Dorit Ravid, Lynne Cahill]
- Currency signs: In-plain-sight instances of phonosemantic compounds, principled allography, and graph-grammar analogy [Short essay: Daniel Harbour]
- Introducing writing systems: Japanese [13] [Terry Joyce & Keisuke Honda]
- Introducing writing systems: Arabic [1] [Elinor Saiegh-Haddad]
- Thought-provoking quotations and observations [15]
- Miscellaneous matters [including recent publications by core community members]

On distribution of 15th AWLL newsletter [Terry Joyce (newsletter editor)]

In introducing this 15th AWLL newsletter (NL15), my first task is to explain about its delayed distribution, asking for your understanding. As the AWLL14 conference convened over 10-12 November 2023 (see organizers' report below)—just prior to the usual mid-November newsletter schedule—I initially thought to aim for a year-end distribution, but, due to a family bereavement at that time and recent teaching commitments, an even longer delay became unavoidable.

Looking back, I cannot help but feel that 2023 was something of a roller-coaster year for AWLL. As noted in the previous newsletter, my attempts in early 2023 to find a local organizer at a European location for AWLL14 kept running aground. But, thanks to an utterly unexpected sponsorship via Temple University Japan, 2023 ended on quite a high, because, as the first NL15 item reports, it was finally possible to gather again for AWLL14 in Rome in November 2023.

NL15 also features a short essay from Daniel Harbour about currency signs as interesting examples for writing systems studies (editor note: Daniel kindly submitted it for the *Thought-provoking ...* section, but it undoubtedly merits inclusion as a welcome short essay piece). The *Japanese* series under the *Introducing writing systems* section continues its focus on rōmaji and in lieu of the *Brahmic* series, which takes a break from NL15, I am delighted to announce the launch of a new *Arabic* series by Elinor Saiegh-Haddad. As usual, NL15 includes a *Thought-provoking quotations and observations* section and concludes with the regular *Miscellaneous matters* section with details of recent publications by AWLL community members.

As always, I very much hope that AWLL NL15 might be of interest to you. Any comments, ideas,

or items for future newsletters are always most welcome; just email them to terry@tama.ac.jp. Past newsletters are available at <http://faculty-sgs.tama.ac.jp/terry/awll/newsletters.html>

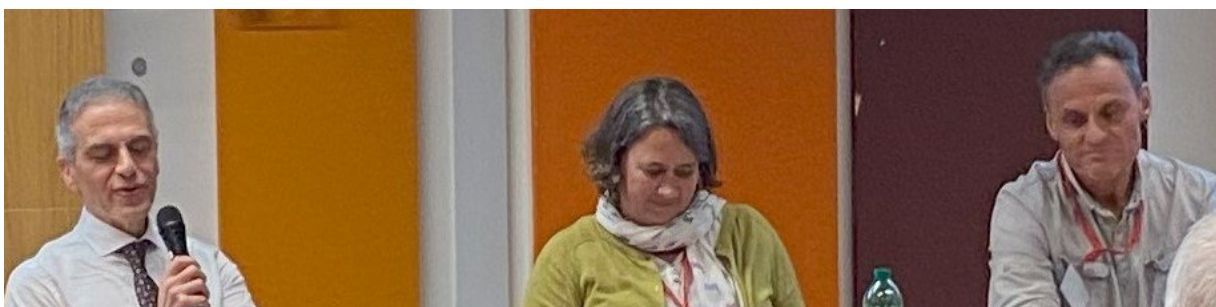
Conference report: AWLL14 [Organizers: Terry Joyce, Dorit Ravid, Lynne Cahill]



AWLL14 participants [onsite after closing remarks]

It is our immense pleasure to report that the 14th meeting of the Association of Written Language and Literacy (AWLL14), entitled the *writing/reading interface*, convened as a hybrid conference over 10-12 November 2023 at the campus of Temple University Rome (TUR), Italy. Although initially planned as an onsite meeting with only audience participation online, after the outbreak of the Israeli-Gaza conflict in October, we decided to switch to a hybrid format to enable several Israeli participants to share their accepted oral papers as online presentations.

Greatly benefitting from a unique case of funding support via Temple University Japan (TUJ) (see minutes of AWLL14's business meeting for further explanation), AWLL14 was the first AWLL conference to feature three invited keynote presentations. The first keynote talk, given by Vito Pirrelli (National Research Council, Italy), asked *Is the finger-voice span an indicator of reading proficiency?* The second keynote talk, given by Kathy Rastle (Royal Holloway, University of London), focused on *Understanding Reading, Understanding Writing*, and the third keynote talk, given by David L. Share (University of Haifa, Israel), was about *Extricating Reading Science from entrenched Anglocentrism, Eurocentrism, and Alphabetism and Embracing Global Diversity: A Personal Journey*. Moreover, the three keynote speakers graciously led the concluding panel discussion session of observations and participant interaction.



AWLL14 Panel discussion (3 keynote speakers: Vito, Kathy and David)

In addition to the three keynote talks, the three-days of the AWLL14 programme consisted of nine oral sessions, comprised of 26 oral presentations (of which five were presented online), and two poster sessions, with four presentations each. As such, the AWLL14 programme featured 37 presentations in total. In selecting the *writing/reading interface* as the theme for AWLL14, we hoped to attract researchers from various backgrounds to reflect together on how the complex interface between writing and reading plays out across the diversity of writing systems. Reflecting the theme's significance for the science of reading and writing systems research, AWLL14 presentations addressed a wide range of topics—such as diglossia, features, handwriting, heterography, homography, language-universality, literacy development, morphology, non-linearity, normativity, and spelling—across a number of writing systems—including Ancient Mayan, Arabic, Chinese, Cretan hieroglyphics, Croatian, English, French, German, Japanese, Korean, and Swedish. Moreover, counting both onsite and online registrations, up to 98 researchers participated from 27 countries, including Australia, Austria, Canada, Croatia, Finland, France, Germany, Hong Kong, India, Ireland, Israel, Italy, Japan, Korea, Luxemburg, Malta, Poland, Singapore, South Africa, Sweden, Switzerland, Thailand, The Netherlands, Togo, UK, Ukraine, and USA.

Finally, we wish to acknowledge our sincere gratitude to everyone who contributed to realizing a wonderful conference: from Chris Barnett (CEO of ICBD Holdings LLC) for his generous sponsorship; TUJ, especially Dean Matthew Wilson and Lydia Hon, for their organizational support; TUR, particularly Dean Emilia Zankina, Natasha Vlady, and Luke Shelley, for the splendid venue and to all participants for creating such a convivial atmosphere of respectful mutual curiosity.

A selection of papers will be published as a two-part special issue of *Written Language and Literacy*; conference participants will receive a call for manuscripts in due course.

Please visit the AWLL14 website to view recordings of the oral presentations and to access both abstract and presentation PDFs for all oral and poster presentations. In addition to PDFs of the conference programme with abstracts and the minutes of the business meeting, the AWLL14 website has a selection of photographs from the conference.

<http://faculty-sgs.tama.ac.jp/terry/awll/WS/14/2023-HP.html>

Currency signs: In-plain-sight instances of phonosemantic compounds, principled allography, and graph–grammar analogy [Short essay: Daniel Harbour]

Good analogies can make light work of hard concepts. For students raised with alphabets and their kin, phonosemantic compounds can be hard to grasp. That Sumerians represented *buru* ‘summer’ via a sign 𒍪 that combined the meaning of 𒍪 ‘shine’ with the sound of 𒍪 *uru*—while ignoring that 𒍪 was pronounced *kar* and that 𒍪 meant ‘clever’—can take time to sink in. Pointing out that Chinese does the same thing on a massive scale—for example, that 𣎵 *p^hrak* ‘cypress’ combines the meaning of 木 ‘tree’ with the sound of 𣎵 *b^hrak* while ignoring the sound *C.m^hok* of 𣎵 and the meaning ‘white’ of 𣎵—can easily amplify the confusion instead of lessening it. Phonosemantic compounds are, however, close to hand in day-to-day life even in alphabetic

parts of the world.

Currency signs are often built on a shared pattern. Consider $\text{A } \text{B } \text{C } \text{C } \text{C } \text{F } \text{G } \text{K } \text{L } \text{N } \text{P } \text{S } \text{T } \text{T } \text{W } \text{Y}$. All compromise a letter of the Latin alphabet crosscut by one or two horizontal, diagonal, or vertical strokes. Though the details vary, the visual of a crosscut letter has a clear semantics, ‘unit of currency’. The choice of letter in each symbol is phonetic, based (except in two cases, discussed at the end) on the first letter of the unit’s name: A austral (Argentina), B baht (Thailand) or bitcoin, C cedi (Ghana), C cent (many countries), C colón (Costa Rican), F franc (France), G guaraní (Paraguay), K kip (Lao), L manat (Azerbaijan), N naira (Nigeria), P peso (Philippines), T tenge (Kazakhstan), T tögrög (Mongolia), W wan (Korea), Y yen (Japan) or yuan (China). The same principles apply in several non-Latin-derived currency signs: Դ dram (Armenia, based on Դ capital d), € euro (European Union, based on variant Greek € e), ₴ hryvnia (Ukraine, based on italic z h), ლარი (Georgia, based on a ‘dimple-free’ allograph of ლ l), ₽ ruble (Russia, based on P capital r), ₹ rupee (India, based on ₹ r). These symbols are all phonosemantic compounds. The underlying letter choice partially encodes sound. The crosscutting strokes are a conventionalised representation of meaning.

The utility of currency signs as teaching aids for writing systems study extends beyond this simple classification. That the semantic crosscuts can vary in number and orientation is an instance of allography. Interestingly, choice of allograph is tied to orientation and properties of the underlying letter. Vertical strokes tend to be single ($\text{B } \text{C } \text{C } \text{S } \text{G } \text{L}$) whereas horizontal and diagonal strokes tend to be double ($\text{C } \text{€ } \text{N } \text{P } \text{T } \text{W } \text{Y } \text{₴}$). Single horizontal strokes mostly occur where the letter itself supplies half of the pair ($\text{A } \text{€ } \text{F } \text{T } \text{Դ } \text{P } \text{₹}$; with readjustment of the preexisting stroke in the stylised E of € , Դ of Դ , and P of ₽ ; analogous coalescence of stroke and letter part is seen, vertically, in L , based on m).

Allograph choice conforms to principles familiar from allomorphy in spoken language. Stroke allography is not wholly deterministic. S has a vertical double-stroke allograph, C a diagonal single-stroke allograph, and L a horizontal double-stroke allograph. Despite such free variation, there is an obvious principle at play in most choices of stroke allograph. Symbols avoid overcrowding. Latin capital letters and others above are taller than they are wide. So, there is more space for horizontal lines than there is for vertical ones. Hence, the dominance of single vertical strokes and double horizontal ones. A linguistic parallel for this principle is the common phenomenon of allomorph choice governed by phonological context. For instance, the accusative in Korean is -ul after a consonant but -lul after a vowel. Hence, 김을 *kim-ul* ‘Kim.ACC’ but 조를 *jo-lul* ‘Cho.ACC’. The results are uniform CVCVC sequences that avoid ‘overcrowded’ CC or VV clusters where morphemes meet.

Single horizontal lines are a second instance of overcrowding avoidance. They are used when a letter already presupplies part of the strokes. Two is fine, three is a crowd. This is a graphic counterpart of haplogly, the avoidance of sequences identical of morphemes (as in the possessive plural of *cat*, which is *cats*, not *cats’s*). Further exceptions to allograph choice avoid overcrowding. If the horizontal stroke of K were double, the three-line meeting point of K would be sandwiched between two close-set lines, each with two intersection points. Use of a single stroke for K thus avoids another form of overcrowding. Likewise, a fourth stroke allograph uses just the tips of lines, as in bitcoin ₿ . Similar variants are common for S , especially the double-

stroke allograph. Foreshortening of the double vertical stroke in € also avoids overcrowding.

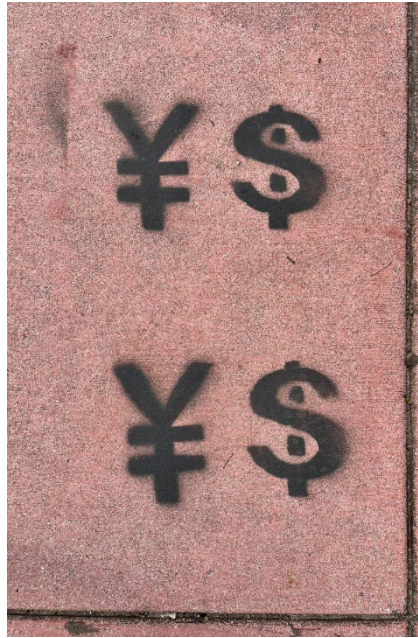
In a further linguistic parallel, allographs can acquire specialised meanings, just as *transmittal* and *transmission* are synonyms in some contexts but only *transmission* is used for ‘gearbox’. For instance, \$ has been used for a variety of currencies including the Brazilian cruzado, the Spanish American peso, and the US dollar. In contexts where multiple \$ currencies need to be disambiguated, the double-stroke variant has been used for the nondominant currency, the simpler symbol for the dominant, default option. Another disambiguation strategy parallels writing systems. Where multiple countries use the same currency name, a further phonetic complement is often added to the currency sign, as in £, £E, and £L for the British, Egyptian, and Lebanese pound. In both the dollar and pound examples, the arrangement of choices parallels spoken language. Marked (more contentful) forms are used in marked (less frequent) circumstances.

Occasional opacity provides a final parallel between currency signs and the concepts that underpin the description of writing systems (and natural language). Consider £. It is not immediately obvious to users that the underlying letter is a calligraphic L and the sound *l* bears no relation to the currency name, *pound*. The mismatch arises because *pound* derives from Latin *libra pondo*, the basic unit of weight of the Roman Empire. English has lost the first part of the compound. (The initial part lives on in French *livre* ‘pound’ and in *lira*, former currency of Italy and Turkey amongst other countries.) An abbreviation that is phonetically transparent to Latin speakers is phonetically opaque to English speakers. Nonetheless, the semantic element of £, the horizontal stroke, still marks it out as a currency symbol. Such partial analysability is typical of morphologically complex words when they travel. To English speakers, *transmit* contains a semantically recognisable prefix *trans-* alongside a bound root *mit*, the meaning of which is harder to pin down (as *admit*, *commit*, and *permit* illustrate).

Phonosemantic compounds face the same difficulties when they travel. Japanese borrowed both the characters 柏 ‘cypress’ and 白 ‘white’ from Chinese. However, the phonetic relationship between them did not travel well. Japanese *shiro* ‘white’ is a poor match for *kashiwa* ‘cypress’. That *shi* furnishes a one-syllable overlap is coincidence. Most borrowings fare worse. Chinese phonosemantic compounds thus frequently transform, in Japanese, into a semantic element and an arbitrary symbol without connection to sound or meaning of the whole character. Precisely this unanalysability affects the most ubiquitous currency sign \$. There is no *s* in *dollar*. The sign’s semantic element, the vertical stroke denoting currency unit, is clear but the underlying consonant is, for contemporary users, resolutely arbitrary (creating a gap for folk ‘etymologies’ to fill, such as deriving the double-stroke variant from a superposition of the initials of the United States).

Currency signs are readily accessible examples of phonosemantic compounds. Transferred crosslinguistically, they illustrate the same opacities as such compounds and at times undergo phonetic complementation. They also exhibit principled allography and markedness constraints, paralleling natural language allomorphy. Currency signs are, therefore, useful examples that any teacher of writing systems studies might wish to use. If tempted, you could assign this brief note as reading. Alternatively, you could send students to the online references below and give them the questions to let them come to the realisations above themselves.

Thanks to Amalia Gnanadesikan and Dimitrios Meletis. Given the informal setting of this remark, I refer readers to https://en.wikipedia.org/wiki/Currency_symbol and https://en.wikipedia.org/wiki/Dollar_sign for further information. For a general overview of linguistic allomorphy, see Eulàlia Bonet and Daniel Harbour, 2012, “Contextual allomorphy” in Jochen Trommer, ed., *The morphology and phonology of exponence*, Oxford: Oxford University Press, 195–235. Old Chinese reconstructions following William Baxter and Laurent Sagart, 2014, *Old Chinese: A new reconstruction*, Oxford: Oxford University Press.



Street art allography, Miami FL (author's photo)

Introducing writing systems: Japanese [13] [Terry Joyce & Keisuke Honda]


Continuing our brief outline of rōmaji (i.e., the Latin script), this instalment seeks to illustrate how the script is a far more integral component of the Japanese writing system (JWS) than the term romanization might generally infer; a transcription device for merely rendering parallel phonological glosses. In the previous instalment, we sketched out the development of the three romanization systems that employ rōmaji—namely, へボン式 /hebon-shiki/ ‘Hepburn system’, 日本式 /nihon-shiki/ ‘Japan system’ and 訓令式 /kunrei-shiki/ ‘Government-ordnance system’—and, although we noted their differences, we also briefly described how the systems each map to Japanese syllables. Transcribing (1)*, for example, all three systems yield the same rōmaji representation as a phonological gloss.

*Most examples in this instalment follow the pattern of primary graphematic representation, alternative representation within <>, Japanese pronunciation gloss within //, translations in italics, and explanatory comments within [].

(1) 日本語 <nihongo> /nihongo/ Japanese language

However, reflecting the extent to which rōmaji has become an indispensable element of the contemporary JWS, many instances of rōmaji graphematic representations are arguably the primary representations rather than the secondary glosses of kanji or kana representations.



Admittedly, the most frequent examples appear as company logos and product names, such as (2~5), but, naturally, by that very virtue, they are highly conspicuous aspects of the Japanese linguistic landscape.

- (2) **HITACHI** <日立> /hitachi/ Hitachi Corporation [manufacturing company]
- (3)  <セブン-イレブン> /seibun-irebun/ Seven-Eleven [convenience stores]
- (4) **Zwei** <ツヴァイ> /tsuvai/ Zwei [marriage partner introduction service]
- (5) **Châteraisé** <シャトレーズ> /shatorēze/ Châteraisé [patisserie shops chain]

Examples (2~5) illustrate the clear trend for using rōmaji for corporate logos among Japanese companies, such as Hitachi in (2), and international companies operating in Japan, such as the 7-ELEVEN logo in (3), which undoubtedly has far greater visual impact than <セブン-イレブン>. Such applications of rōmaji seek to leverage the generally positive associations of the Latin script with notions of modernity, internationalization and cosmopolitanism. Many instances are derived from English-language words, such as (3), but not all; (4) is the German word for two (an apt wordplay for a marriage-partner introduction service) and (5) is a created company name designed to deliberately elicit images of French patisseries by imitating its orthographic conventions. Such examples underscore the immense difficulties of determining the target pronunciations of rōmaji graphematic representations when they might be conforming to the graphematic conventions of any language.


- (6) **KOBAN** <交番> /kōban/ police box
- (7) **LIVE** <ライブ> /raibu/ live concert


In contrast to the corporate names of (2~5), another common example within the linguistic landscape is (6). Despite the lack of a standard materialization (both stone inscriptions and painted signs are frequently attested), it identifies police boxes. Example (7) is a particularly frequent element of advertisements for live music performances, often as the only word graphematically represented in rōmaji.

- (8)  <エヌエッチケー> /enuetchikē/ NHK
- (9)  <ジェイアール> /jeiāru/ JR

Consistent with clipping being a major principle of Japanese word-formation, there is also a marked prevalence for rōmaji abbreviations. Two prominent examples of rōmaji abbreviations

used as corporate logos are (8~9). In contrast to the abbreviation of the Japanese three-word corporation name <日本放送協会> /nihon hōsō kyōkai/ Japan Broadcasting Corporation as <NHK> in (8), the stylized logo combination of <JR> in (9) abbreviates the English of Japan Rail; the green logo being specifically for <JR 東日本> /jeiāru higashi nihon/ JR East.

- (10) **3LDK** /san erudikē/ 3 bedrooms, living-dining room, kitchen [real estate code]
- (11) **NG** /enujī/ (TV and movie) outtakes, bloopers
- (12) **KY** /kēwai/ out of touch with reality
- (13) **ggr** /gugure/ google it!
- (14) **WC** <  > /toire/ toilet

Beyond the frequent uses in corporate logos, the Japanese penchant for rōmaji abbreviations is highly conspicuous in many domains. For instance, within the world of real estate, (10) is an example of a generic code for apartment dimensions; the initial digit indicates the number of rooms suitable for use as bedrooms, <LK> stands for a combined living-dining room space and <K> for a separate kitchen. Reflecting the popularity of TV or movie outtakes or bloopers, (11) is frequently encountered within entertainment circles, where <NG> is an abbreviation of the Japanese-coined phrase No Good. Not surprisingly, rōmaji abbreviations, such as (12-13), are highly prevalent in the domains of computer-mediated communication and texting. Example (12) is now somewhat dated, but <KY> is an abbreviation for <空気読まない> /kūki yomenai/ unable to read the atmosphere. <ggr> is a particularly interesting example with its layers of lexical creativity. It is an abbreviation for <ググレ> /gugure/ (with common vowel-omitting); it ultimately stems from <グーグル> /gūguru/ Google via both initial vowel-length reduction to become <ググる>, which is then playfully interpreted as a <る> /ru/ ending verb, albeit it with an imperative <レ> /re/ inflection here. Although originally an abbreviation of water closet, (14) now effectively functions as a semasiograph, like < >, which may, depending on the context, be 'read' as /toire/ [<トイレ> toilet], /oteari/ [<お手洗い> toilet], /kyūkeishitsu/ [<休憩室> restroom], or /keshōshitsu/ [<化粧室> powder room] among other euphemisms.

- (15) **U字型** /yūjigata/ U-shape
- (16) **Yシャツ** /waishatsu/ (white) dress shirt; business shirt

As with English U-shape curve and T-shirt, some rōmaji graphematically represent morphemes related to their shape, such as (15), which parallels to the English. In contrast, the <Y> /wai/ of (15) is derived from <ホワイト> /howaito/ white.

While highly selective in range, the rōmaji examples of this instalment seek to demonstrate that contemporary rōmaji usage is not limited to merely applying an external script as phonemic

glosses of Japanese words. Rōmaji is unquestionably an integral and indispensable element of a multi-script writing system. With this second instalment on rōmaji, we conclude our brief outlines of the separate component scripts; from the next newsletter piece, we turn to explain how the multi-scripts are used together as parts of a coherent single writing system for graphematically representing the Japanese language.

Introducing writing systems: Arabic [1] [Elinor Saiegh-Haddad]

Launching a new series of *Introducing writing systems* about the Arabic writing system, this initial instalment mainly provides an overview introduction to the structures of the Arabic language followed by a briefer description of its orthography.

The term *اللغة العربية* /al-luḡa l-ʿarabiyya/ *The Arabic language* is used to refer to (Modern) Standard Arabic (MSA, a modern descendant of Classical Arabic)—a rather uniform variety and the only language of conventional reading and writing—and to the large (in fact infinite) number of spoken Arabic varieties/vernaculars/dialects (collectively referred to as either Colloquial Arabic or Spoken Arabic) that are used by all speakers: young and old, educated and uneducated, literate and illiterate, for everyday speech, everywhere. Spoken Arabic varieties do not have official orthographies and, although they are used by Arabic speakers for writing on various social media platforms, such as WhatsApp, MSA remains the only language of official writing (articles, books, newspapers, and literature including for children) and of education (at least in aspiration) as early as the first grade.

MSA and all Spoken Arabic varieties are characterized by a predominantly non-linear or non-concatenative morphological structure; the hallmark being a *root*, typically trilateral (and less frequently quadrilateral), that denotes an abstract semantic core and a *pattern*, a prosodic template with associated morpho-semantic roles. The *root-pattern* structure captures the entire Arabic lexicon and determines how speakers store and access words in both MSA and Spoken Arabic. For example, the following words all share a common root of KBR differing only in their patterns.

| | | |
|---------------------------|-----------------------------|-------------------------------|
| /kibir-kabur/ <i>grow</i> | /kabbar/ <i>raise</i> | /takabbar/ <i>be arrogant</i> |
| /kabi:r/ <i>big</i> | /mukabbir/ <i>amplifier</i> | /kubra/ <i>arrogance</i> |

There are two kinds of Arabic word patterns: verbal patterns that combine with roots to derive verbs and nominal patterns that combine with roots to derive nouns. There are 15 distinct trilateral verbal patterns in MSA, but not all are productive in the spoken dialects. For instance, only ten verbal patterns are used in Palestinian Arabic (Laks, Hamad & Saiegh-Haddad, 2019). The verbal patterns are primarily vowel templates, but some involve gemination of root consonants or vowel lengthening and others are augmented with certain consonants (/ʔ s t n/) that are prefixed to the root. The first pattern is CaCaC—and its MSA variants of CaCiC and CaCuC and dialectal CiCiC—where C stands for a consonant and the typical set of three consonants form a pattern of fixed slots, with the root consonants conforming to a specified order. Referred to as *faʿal* (from the root *fʿl* ‘do’) in traditional Arabic grammar, this first pattern is semantically basic,

syntactically neutral and constitutes the most frequent pattern in both MSA and the dialects. The nominal patterns form a very large set; indeed, Wright's (1975) grammar of Classical Arabic lists 44 nominal patterns derived from CaCaC alone and Aralex reports about 500 noun patterns in MSA (Boudelaa & Marslen-Wilson, 2010). Moreover, nouns can be primitive and not encode a root and pattern structure (e.g., جبل /jabal/ *mountain*).

Verbs are always derived non-linearly from a root and a verbal pattern. Root-based verbs with different verbal patterns constitute derivational families with the members expressing different semantic values including transitivity. For example, the root KTB *write* combines with verbal patterns to create a family of five different verbs:

| | |
|--------------------------------------|---|
| /katab/ <i>write</i> basic | /ka:tab/ <i>write to one another</i> reciprocal (tr.) |
| /kattab/ <i>make write</i> causative | /taka:tab/ <i>write to one another</i> reciprocal (intr.) |
| /inkatab/ <i>was written</i> passive | |

The relations between verbs of different patterns are derivational, being primarily manifested as transitivity alternations and other kinds of semantic relations. For example, CaCCaC verbs are usually active transitive verbs (e.g., /yassal/ *wash*) with the intransitive alternative being tCaCCaC (i.e., the reflexive verb /tyassal/ *wash oneself*).

The relations between different nouns derived from a shared root and different nominal patterns are generally less transparent. Derived nouns are often semantically close to the root meanings, which indicates that, in addition to the root meaning, general meanings are associated with certain patterns; at least, in the case of the following productive nominal patterns:

| |
|---|
| Ca:CeC (agent noun); /la:ʔeb/ <i>player</i> |
| maCCu:C (patient noun); /maktu:b/ <i>written</i> |
| miCCala-CaCCa:Ca (instrument noun); /mibraʃa-barraʃa/ <i>grater</i> |

All consonants, including glides, can function as root-radicals. However, roots that contain a glide are regarded as 'weak', in the sense that they are prone to morpho-phonological changes and are, hence, irregular. In contrast, the radicals of 'sound' roots are phonologically stable.

Arabic features a predominant system of clitics that linearly attach to word stems. Clitics may attach to a word as an unstressed prefix (proclitic) or suffix (enclitic), which can co-occur within the same word, as in /bi-bayt-i:/ *in my house* and /wa-ʔakalna:-ha/ *and we ate it*. Pronominal clitics are suffixed to verbs as direct objects (e.g., /ʔakaltu-ha:/ *I ate it*), to nouns as possessives (e.g., /bayt-i:/ *my house*), and to prepositions (e.g., /min-ha:/ *from it/her*). In contrast, clitics that are prefixed to words include several prepositions (e.g., /bi-qissa/ *in a story*), conjunctions (e.g., /wa-raqasna:/ *and we danced*) as well as other particles, such as the definite article (e.g., /(a)l-bayt/ *the house*).

This initial instalment concludes with a brief description of the Arabic script, as used in orthographically materializing the lexicon, phonology, and grammar of MSA. Arabic is written from right to left with a cursive script. All 28 letters of the alphabet represent consonants, apart from aleph which may function as a 'bearer' for an additional sign الهمزة /hamza/, representing the 28th glottal stop consonant. Thus, the Arabic alphabet is generally classified as a consonantal

alphabet or an abjad—a type of writing system where each sign always (or usually) stands for a consonant requiring readers to supply the appropriate vowels. This system is well suited to the root-pattern morphological structures of Arabic, because, as outlined above, the core semantic meaning is inherent within a consonantal root and vowel information can be recovered from the vocalic-word pattern. While each of the 28 letters of the Arabic alphabet (except aleph) represents a consonant, the three letters *ا و ي* are called *حروف العلة* /*huru:f al-Silla/ defective letters*, which act as the *matres lectionis* ‘mothers of reading’, used to represent the three Standard Arabic long vowels: high front /i:/, high back /u:/, and low /a:/, respectively. These three letters are also called *حروف اللين والمدّ* /*huru:f al-li:n wal-madd/ letters of softness and elongation* because, according to traditional views, they indicate the elongation of a preceding short vowel sound, which is orthographically represented by a diacritic-vowel mark.

In the next installment, we will elaborate on the orthographic features of the Arabic letters and on the different kinds of diacritic signs that are used in the Arabic orthography to map phonemic information, such as vowels and consonant gemination, and morpho-syntactic information, such as case and mood.

Boudelaa, S., & Marslen-Wilson, W. D. (2010). Aralex: A lexical database for modern standard Arabic. *Behaviour Research Methods*, 42, 481–487.

Laks, L., Hamad, I., & Saiegh-Haddad, E. (2019). Verbal Patterns in Palestinian Arabic. *The Mental Lexicon*, 14(2), 209-236. <https://doi.org/10.1075/ml.00005.lak>

Wright, W. (1975). *A grammar of the Arabic language* (3rd ed.). Cambridge: Cambridge University Press.

Thought-provoking quotations and observations [13]

How to transform written language into music - shared by Martin Neef

We all know what writing systems are: They are the subject of our scientific work. Here is a completely different use of the term writing system. This year, guitarist Ron Jarzombek released a CD entitled “Apps, Writing Systems and Remakes”. Three of the 21 songs belong to the writing systems section. The first two present apps that Ron Jarzombek has developed in recent years, and the third illustrates what can be produced with such apps. The basic idea is to distribute the 26 letters of the Latin alphabet over a continuous musical scale. Jarzombek calls this the diatonic alphabet. The “Create a riff” app turns a short text of up to 16 letters into a guitar riff, while the “Play the name” app converts short texts into melodies, chords and drum tracks. The song *In the name of Ron Jarzombek* shows what the app does with this name. YouTube videos demonstrate how the apps work. Now we can all turn our names into music and compare what sounds best. Abraham Lincoln comes out as an amazingly melodious name!

A tiny anecdote about typefaces/fonts - shared by Terry Joyce

This little scribble, which is a rather random add-on to a couple of group conversations at AWLL14, is proffered in the hopes of encouraging more knowledgeable others to contribute to a wider dialogue about typefaces/fonts. During the first AWLL14 conversation at one of the

delicious lunch breaks (Day 2, I think), the small-talk had turned to tofu – not the soybean curd but the blank rectangles (i.e., □□□□ resembling tofu blocks) that appear when no appropriate font is available – and I recollect a smidgen of surprise that not everyone at the small dinner table was not already aware of Google’s Noto fonts (more below). However, in the second conversation that evening over drinks, as one person recounted their efforts to ensure all the various fonts within their presentation were properly embedded within a PDF version so it would open correctly on a different computer, I also recall feeling a substantial degree of awe. My sincere admiration was less for the attention to detail (which I already strongly associate with the individual) but for their obvious familiarity with a wide variety of different fonts, which are, after all, quite a fiddly bit of working with writing systems.

As it engendered a leap in my own limited understanding of such matters, my little story, from some years ago now, concerns the proof-correction stage of a particular paper that took a while – something that is likely to resonate with many readers – (until proofs 4) to resolve the typeface issue for the Japanese language examples! Now, while I would immediately concede some responsibility for the slow emergence of a solution (at the time, I didn’t know about finding embedded fonts in PDFs via the properties menu), unfortunately, it has been a recurring issue with different publishers, so perhaps bears sharing to highlight the importance of typeface differences.

As a key background point, formal Japanese texts (i.e., academic articles) are typically rendered in a 明朝 /minchō/ (serif) font variant. Given that MS Minchō has been the de facto default with Microsoft Windows for some time (although 游明朝 /yūminchō/ has also been available since Windows 8.1; Microsoft typography), the Japanese language examples within the manuscript sent to the typesetters were in that typeface, such as the top row below. However, when the first proofs arrived, the examples were in the SimSun typeface, as on the second row. As *spot-the-differences* quizzes go, comparing the characters in these two typefaces is hardly a challenge. Yet, on asking for the typeface to be changed, the initial response was that SimSun is the standard CKJ typeface (a common abbreviation in this domain for Chinese, Korean and Japanese). According to Microsoft’s overview description, “SimSun & NSimSun is a Simplified Chinese font features mincho (serif) stroke style”. Although “Simplified Chinese” there should have been dead give-away, it did take another round of proofs before everyone acknowledged the differences between the MS Minchō and SimSun typefaces and to agree on a solution; namely, to use Google’s Noto Serif JP, as on the third row.

| | | | | | |
|----------------|---------|--------|----------|--------|---------|
| MS Minchō | 漢字 | 仮名 | 言語 | 紙 | 麵 |
| SimSun | 漢字 | 仮名 | 言語 | 紙 | 麵 |
| Noto Serif JP* | 漢字 | 仮名 | 言語 | 紙 | 麵 |
| gloss | /kanji/ | /kana/ | /gengo/ | /kami/ | /men/ |
| English | kanji | Kana | language | paper | noodles |

* The three rows of typefaces are all set at 16-point fonts, for visibility, but Noto typeface is at an extra light weighing.

On the off chance that a Google font could be the solution to a reader's typeface problem in the future, let me conclude this tiny rant with a couple of introductory remarks about Noto fonts. Starting with the name, the website lays claim to two meanings. The more erudite is that, in Latin, Noto means "I write, I mark, I note" while the slightly geeky explanation is that Noto is short for "no tofu" (again, ☐☐☐☐, not the soybean blocks), a slogan for a project that seeks to eliminate ☐☐☐☐ from all text processing. Currently, the Fonts section offers 1,595 families. Under a Fonts Knowledge link, this is also wide range of typography lessons arranged under 23 topic headings, together with a glossary link. Clearly, a lot to talk when it comes to typefaces/fonts.

<https://learn.microsoft.com/en-us/typography/>

<https://fonts.google.com/noto>

PPS: By the way, Daniel has used Noto San Cuneiform for the Sumerian examples in his essay.

Miscellaneous matters

Information about ongoing projects, upcoming conferences, events, special issues

Second Workshop on Computation and Written Language (CAWL 2024)

Torino, Italy; 21 May 2024

<https://sigwrit.org/workshops/cawl2024/>

10th Cambridge Conference on Language Endangerment

University of Cambridge, United Kingdom; 12 July 2024

<https://www.mml.cam.ac.uk/centres/celc/conference-series>

Writing and Cognition in Interdisciplinary Perspective

Monte Verità, Ascona, Switzerland; 25-30 August 2024

Traditions in the study of writing worldwide [Thematic workshop @ ICHoLS XVI]

Tbilisi Ivane Javakhsvili State University, Tbilisi, Georgia; 26-30 August 2024

<https://meletis.at/ichols16.pdf>

INSCRIBE at the Roots of Writing

Inscribe Project, University of Bologna, Italy; 11-13 September 2024

<https://site.unibo.it/inscribe/en/events/inscribe-at-the-roots-of-writing-final-conference>

Writing as Visual Experience

Views Project, University of Cambridge, United Kingdom; 20-22 September 2024

<https://viewsproject.wordpress.com/writing-as-visual-experience/>

Grapholinguistics in the 21st Century, 2024

Università Ca' Foscari, Venice, Italy; 23-25 October 2024

<https://grafematik2024.sciencesconf.org/>

~~~~~



### **Recent publications by AWLL community members**

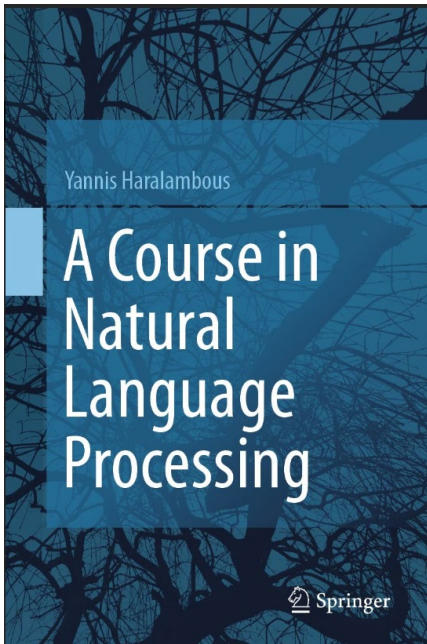
The *Miscellaneous matters* section concludes with a list of recent publications (i.e., since the last newsletter) by AWLL community members (followed by ‘mini-flyers’ for a book publication).

AWLL mailing list is open to anyone interested in receiving occasional information emails, but the core community is based primarily on participation at AWLL workshops. All who participated at recent workshops are eligible to have a brief member profile at the community page and to include recent publications under this section of future newsletters.

For further information, go to <http://faculty-sgs.tama.ac.jp/terry/awll/community.html>

- Gnanadesikan, Amalia E. (2023). Classifying and comparing early writing systems. In Marco Condorelli & Hanna Rutkowska (Eds.). *The Cambridge handbook of historical orthography* (pp. 29–49). Cambridge University Press. <https://doi.org/10.1017/9781108766463.002>
- Haralambous, Yannis. (2024). *A course in natural language processing*. Cham: Springer. <https://doi.org/10.1007/978-3-031-27226-4>
- Joyce, Terry. (2023). Typologies of writing systems. In Marco Condorelli & Hanna Rutkowska (Eds.). *The Cambridge handbook of historical orthography* (pp. 138–159). Cambridge University Press. <https://doi.org/10.1017/9781108766463.007>
- Judson, Anna P. (2023). The tablet-makers of Pylos: An experimental investigation into the production of Linear B tablets. *Annual of the British School at Athens*, 118, 147–170. <https://doi.org/10.1017/S0068245423000059> [open access]
- Meletis, Dimitrios. (2024). Schriftlinguistik interdisziplinär, multiperspektivisch, komparativ: Die Erarbeitung struktureller, psycholinguistischer und soziolinguistischer Typologien. In Sabine Krome, Mechthild Habermann, Henning Lobin & Angelika Wöllstein (Eds.), *Orthographie in Wissenschaft und Gesellschaft: Schriftsystem – Norm – Schreibgebrauch* (Jahrbuch 2023 des Leibniz-Instituts für Deutsche Sprache) (pp. 399–410). Berlin; Boston: De Gruyter. <https://doi.org/10.1515/9783111389219-022>.
- Roberts, David. (2023). Roman script orthography development in Africa: Historical and contemporary perspectives. In R. Malatesha Joshi, Catherine A. McBride, Bestern Kaani, & Gad Elbeheri (Eds.), *Handbook of literacy in Africa* (pp. 59–78). New York: Springer Nature.
- Saiegh-Haddad, Elinor. (2023). Embracing diglossia in early literacy education in Arabic: A pilot intervention study with kindergarten children. *Oxford Review of Education*, 49(1), 48–68. <https://doi.org/10.1080/03054985.2022.2090324>
- Steele, Philippa M. (2023). *Exploring writing systems and practices in the Bronze Age Aegean* (CREWS 7). Oxbow Books.

~~~~~



Haralambous, Yannis. (2024). *A course in natural language processing*. Cham: Springer. <https://doi.org/10.1007/978-3-031-27226-4>

Publisher's website:

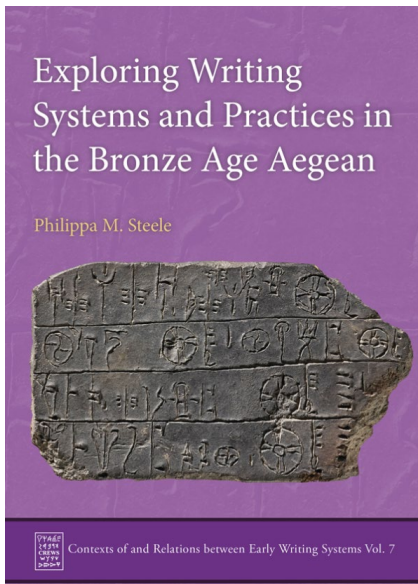
<https://link.springer.com/book/10.1007/978-3-031-27226-4#about-this-book>

Natural Language Processing is the branch of Artificial Intelligence involving language, be it in spoken or written modality. Teaching Natural Language Processing (NLP) is difficult because of its inherent connections with other disciplines, such as Linguistics, Cognitive Science, Knowledge Representation, Machine Learning, Data Science, and its latest avatar: Deep Learning. Most introductory NLP books favor one of these disciplines at the expense of others.

Based on a course on Natural Language Processing taught by the author at IMT Atlantique for over a decade, this textbook considers three points of view corresponding to three different disciplines, while granting equal importance to each of them. As such, the book provides a thorough introduction to the topic following three main threads: the fundamental notions of Linguistics, symbolic Artificial Intelligence methods (based on knowledge representation languages), and statistical methods (involving both legacy machine learning and deep learning tools).

Complementary to this introductory text is teaching material, such as exercises and labs with hints and expected results. Complete solutions with Python code are provided for educators on the SpringerLink webpage of the book. This material can serve for classes given to undergraduate and graduate students, or for researchers, instructors, and professionals in computer science or linguistics who wish to acquire or improve their knowledge in the field. The book is suitable and warmly recommended for self-study.

~~~~~



Steele, Philippa M. (2023). *Exploring writing systems and practices in the Bronze Age Aegean* (CREWS 7). Oxbow Books.

It can be downloaded at:

<https://crewsproject.files.wordpress.com/2023/10/steele-2023.pdf>

Publisher's website:

<https://www.oxbowbooks.com/9781789259018/exploring-writing-systems-and-practices-in-the-bronze-age-aegean/>

Writing does not begin and end with the encoding of an idea into a group of symbols. It is practiced by people who have learnt its principles and acquired the tools and skills for doing it, in a particular context that affects what they do and how they do it. Nor are these practices static, as those involved exploit opportunities to adapt old features and develop new ones. The act of writing then has tangible and visible consequences not only for the writers but also for those encountering what has been produced, whether they can read its content or not – with potential for a wider social visibility that can in turn affect the success and longevity of the writing system itself.

With a focus on the syllabic systems of the Bronze Age Aegean, this book attempts to bring together different perspectives to create an innovative interdisciplinary outlook on what is involved in writing: from structuralist views of writing as systems of signs with their linguistic values, to archaeological and anthropological approaches to writing as a socially grounded practice. The main chapters focus on the concepts of script adoption and adaptation; different methods of logographic writing; and the vitality of writing traditions, with repercussions for the modern world.

---

#### AWLL board

President: Terry Joyce; vice-presidents: Lynne Cahill & Dorit Ravid

AWLL website: <http://faculty-sgs.tama.ac.jp/terry/awll/index.html>;

On Facebook and Twitter [[@awll](#)]

© 2024