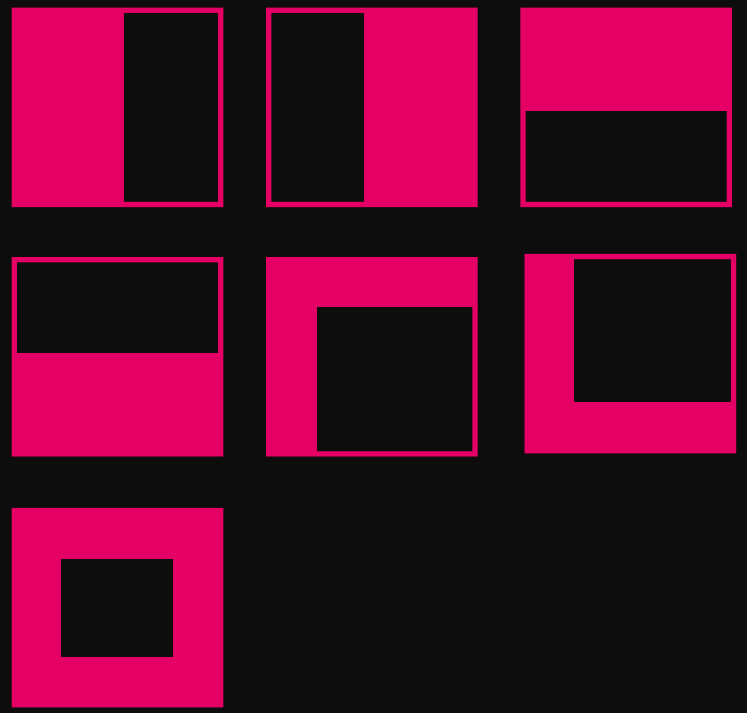
GRAPHOTACTICS,

SPATIALITY, AND WHY WRITING SHOULD BE STUDIED INDEPENDENTLY OF SPEECH

Dimitrios Meletis
University of Zurich



ON THE SYSTEMATIC NATURE OF WRITING SYSTEMS

13th Workshop of the *Association for Written Language and Literacy*

October 21st, 2021, University of North Carolina (Zoom)

Comparative concepts: Why and how?

- in the description of (even the most diverse) languages, we rely on basic and broadly defined, cross-linguistically applicable concepts such as *phoneme*, *morpheme, allophony, phonotactics*, etc. as descriptive categories
  - concepts needed to further specify the structure and (e.g., grammatical) function of these basic categories such as case or number are often fuzzy against the background of linguistic diversity; they should be defined more broadly, as comparative concepts (cf. Haspelmath 2010)
- writing may be dependent on language (and this does not equal speech!), but it is its own system that can (and should) be described
  - the dependence and different types of relations with language(s) (= types of writing systems) complicate matters
  - therefore, we should opt for comparative concepts in grapholinguistics; defining *grapheme* as such means important system-specific details need not be swept under the rug

Two identities: Dependence and independence

- writing is dependent on language: its units relate to units of language (= narrow, glottographic definition of writing)
  - and in many cases to units of sound (if indirect relations due to double articulation count, all writing is phonographic, which was postulated – more absolutely – by DeFrancis 1989)
- writing is also its own semiotic system with its own regularities both at the material (= graphetic) and functional (= graphematic) levels
- example: universals of writing
  - "no nasal vowel is represented unless some oral vowels are represented" (Justeson 1976: 69) → a reflection of writing's dependence on language
  - "[f]ew writing systems distinguish all their phonemes" (Justeson 1976: 61) → genuine feature inherent to the writing system
  - "the orientation of strokes inside written characters massively favors cardinal directions" (Morin 2018: 664) → graphetic universals are also not dependent on language

Graphotactics: Attempting a definition

- graphotactics captures “restrictions on ways in which the elements of a writing system may combine with each other” (McCawley 1994: 115)
  - which is the basis for a range of restrictions that concern, among other things, possible sequences of units, the length of well-formed units, the position of units within larger units, …; they have not been systematically organized
  - it is analogous/parallel to phonotactics and morphotactics, which does not diminish the claim that writing is its own system (it just means it works similarly to other linguistic subsystems, which is unsurprising)
  - popular in psychological research (cf. Sobaco et al. 2015; Pacton et al. 2013; Treiman & Boland 2017); linguistic literature is scarce
- it is at the core of arguing that writing is its own system
  - otherwise, writing would only follow grapheme-phoneme (syllable, morpheme)-correspondences (and irregularities)
  - example: consonant doubling in words such as <hammer>
  - this also highlights the importance of a suprasegmental (syllabic and lexical) perspective that should at least complement the segmental perspective (cf. Evertz 2018)

Types of graphotactic constraints: Graphetic

- as visual systems, scripts also have underlying systematics and regularities
- however, they are (usually) not applied productively as scripts are closed systems, many of which result from a long historical development (which was immensely slowed down or even stopped by the printing press and later the computer)
- exception: creation of new shapes, e.g., in Chinese, Japanese, or occasionally in alphabets, cf. the introduction of <ß> in German (these were all regulated orthographically and did not develop 'naturally')
  - in these cases, the systematics of a script should be considered (cf. Watt 1983)
- graphetics differs from phonetics in an important respect: phonetic units are physiologically restricted, they are a closed class; we have no comparable physiological limitations for our writing organs and the boundary seems to be cognitive instead (how much complexity we can process)

Types of graphotactic constraints: Graphematic

- positional constraints: subsegmental level
  - semantic and phonological components in Chinese predominantly appear in specific positions: semantic components on the left, phonological components on the right (cf. Taft, Zhu & Peng 1999; Ho, Ng & Ng 2003)
  - these positions determine whether non-existing graphemes are 'legal' or not, i.e., whether they are pseudographemes or nongraphemes, and influence whether they are judged as such in character decision tasks
    - even first graders have this implicit graphotactic knowledge (cf. Shu & Anderson 1999)
  - position of subsegmental graphemes in syllable blocks in Hangul is also a genuinely graphotactic feature

Types of graphotactic constraints: Graphematic

- positional constraints: segmental and suprasegmental levels
  - of course, orthographically, in phonographic writing systems, due to phoneme-grapheme (or syllable-grapheme) correspondences, the sequence is largely determined by linguistic structures
    - in processing (and graphematically?), sequence may not be as important (cf. *Cambridge University Effect*, cf. Velan & Frost 2007); cf. also typos
  - sequence of written units need not necessarily follow the sequence of the linguistic units of correspondence, leading to 'misalignment' (example: misaligned vowels in Thai, cf. Winskel 2009)
  - in fact, the position of non-linear dependent vowel 'diacritics' (as bound graphemes) in abjads and abugidas also above/below consonant graphemes is not determined by anything in phonology

Types of graphotactic constraints: Graphematic

– positional constraints (cont'd)

  – length sequencing principle (cf. Fuhrhop, Buchmann & Berg 2011), in analogy to the sonority hierarchy: the way graphematic syllables are visually structured, with 'long' shapes at edges and 'compact' shapes as nuclei, cf. <peak>

– combinatory constraints

  – in English, final <y> is replaced by <i> when a suffix is added (<glory> vs. <glorious>), except when <-ing> is added: <carrying>; this cannot be an absolute constraint, however, as there are exceptions such as <skiing> (cf. McCawley 1994)

  – combination of different types of written units: digits and 'letters' (default graphemes) may combine freely with each other: <27> or <twenty-seven>, whereas special characters such as <$> or punctuation marks do not (but cf. <?!>) (cf. Bredel 2011)

Types of graphotactic constraints: Graphematic

- length constraints
  - three-letter-rule in English: <by> vs. <buy>, <be> vs. <bee>
  - motivated by a drive for disambiguation, i.e., heterography? (this was not found for heterographic homophonous stems, cf. Berg & Aronoff 2021)
  - shows also that there can be a 'surplus' in writing (and not only in primarily morphographic systems)
  - spellings such as <have> and <give> with final <-e> that is not needed phonographically are motivated by suprasegmental written structures (cf. Evertz 2018)

Graphotactics as part of graphematics

- theoretically, graphotactics is part of the graphematic solution space (cf. Neef 2005, 2015)
  - as mentioned, correspondences between linguistic units (such as the phoneme) and graphemes would not be sufficient to account for well-formed units of writing since writing is no mere depiction of language
- in use, like everything in graphematics, graphotactics is implicit knowledge
  - is acquired very early, cf. also the judgment of doublets in Finnish (cf. Lehtonen & Bryant 2005)

Spatiality

- spatiality is constitutive of writing both as a material and a functional system
- two-dimensionality as an important feature (or affordance) that requires writing to have its 'own' rules as no (other) linguistic subsystem is two-dimensional
  - in other words: possibilities come with the necessity of constraints
- linearity (both horizontal and vertical), directionality and the complex organization of different parts of the writing surface (such as the segmental space, which is filled by much more complex units in Chinese and Korean than in writing systems using Roman script, for example)
- temporality, which is important for speech, is also important for the use of writing, which may in the long run affect the system
  - are stroke order rules in the production of sinographic graphemes graphotactic when they influence the product?

Conclusion: A methodological choice

– Eisenberg (1988: 29, my translation) captures how studying writing independently of speech does not negate that writing is subordinate in some respects:

> "The structural analysis of writing divorced from speech can make sense even when writing is functionally and genetically subordinate to speech in every respect. The postulate of a graphematics that is independent of phonology exists for the simple reason that because of it, a projection of the structure of speech to writing is avoided. This is necessary since otherwise structural features of writing could remain unseen.

– systematic features inherent to writing are no (historical) 'accidents', they reveal to us how writing functions and can only be discovered in an autonomous grapholinguistic approach (which is, however, also always just one side of the story)
– writing does depict language, but it does so in its very own way
  – if we are interested in writing, we should be interested in exactly that way

THANK YOU!

References [1/2]

Berg, Kristian & Mark Aronoff. 2021. Is the English writing system phonographic or lexical/morphological? A new look at the spelling of stems. *Morphology* 31(3). 315–328. https://doi.org/10.1007/s11525-021-09379-5

Bredel, Ursula. 2011. *Interpunktion*. (Kurze Einführungen in die germanistische Linguistik 11). Heidelberg: Winter.

DeFrancis, John. 1989. *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawaii Press.

Eisenberg, Peter. 1988. Über die Autonomie der graphematischen Analyse. In Dieter Nerius & Gerhard Augst (eds.), *Probleme der geschriebenen Sprache. Beiträge zur Schriftlinguistik auf dem XIV. internationalen Linguistenkongreß 1987 in Berlin*, 25–35. (Linguistische Studien A 173). Berlin: Akademie der Wissenschaften der DDR.

Evertz, Martin. 2018. *Visual prosody: The graphematic foot in English and German*. Berlin/Boston: De Gruyter. https://doi.org/10.1515/9783110583441

Fuhrhop, Nanna, Franziska Buchmann & Kristian Berg. 2011. The length hierarchy and the graphematic syllable: Evidence from German and English. *Written Language & Literacy* 14(2). 275–292. https://doi.org/10.1075/wll.14.2.05fuh

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687. https://doi.org/10.1353/lan.2010.0021

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. https://doi.org/10.1207/s15548430jlr3503_3

Justeson, John S. 1976. Universals of language and universals of writing. In Alphonse Juilland, Andrew M. Devine, & Laurence D. Stephens (eds.), *Linguistic studies offered to Joseph Greenberg on the occasion of his 60th birthday*, vol. 1, 57–94. (Studia linguistica et phonologia 4). Saratoga: Amna Libri.

Lehtonen, Annukka & Peter Bryant. 2005. Doublet challenge: Form comes before function in children's understanding of their orthography. *Developmental Science* 8(3). 211–217. https://doi.org/10.1111/j.1467-7687.2005.00409.x

McCawley, James David. 1994. Some graphotactic constraints. In William C. Watt (ed.), *Writing systems and cognition*, 115–127. (Neuropsychology and Cognition 6). Dordrecht: Springer. https://doi.org/10.1007/978-94-015-8285-8_7

References [2/2]

Morin, Olivier. 2018. Spontaneous emergence of legibility in writing systems: The case of orientation anisotropy. *Cognitive Science* 42(2). 664–677. https://doi.org/10.1111/cogs.12550

Neef, Martin. 2005. *Die Graphematik des Deutschen.* (Linguistische Arbeiten 500). Berlin/Boston: De Gruyter. https://doi.org/10.1515/9783110914856

Neef, Martin. 2015. Writing systems as modular objects: Proposals for theory design in grapholinguistics. *Open Linguistics* 1. 708–721. https://doi.org/10.1515/opli-2015-0026

Pacton, Sébastien, Amélie Sobaco, Michel Fayol & Rebecca Treiman. 2013. How does graphotactic knowledge influence children's learning of new spellings? *Frontiers in Psychology* 4(701). https://doi.org/10.3389/fpsyg.2013.00701

Shu, Hua & Richard C. Anderson. 1999. Learning to read Chinese: The development of metalinguistic awareness. In Jian Wang, Albrecht W. Inhoff & Hsuan-Chih Chen (eds.), *Reading Chinese script: A cognitive analysis*, 1–18. Mahwah: Lawrence Erlbaum.

Sobaco, Amélie, Rebecca Treiman, Ronald Peereman, Gaëlle Borchardt & Sébastien Pacton. 2015. The influence of graphotactic knowledge on adults' learning of spelling. *Memory & Cognition* 43. 593–604. https://doi.org/10.3758/s13421-014-0494-y

Taft, Marcus, Xiaoping Zhu & Danling Peng. 1999. Positional specificity of radicals in Chinese character recognition. *Journal of Memory and Language* 40(4). 498–519. https://doi.org/10.1006/jmla.1998.2625

Treiman, Rebecca & Kelly Boland. 2017. Graphotactics and spelling: Evidence from consonant doubling. *Memory and Language* 92. 254–264. https://doi.org/10.1016/j.jml.2016.07.001

Velan, Hadas & Ram Frost. 2007. Cambridge University versus Hebrew University: The impact of letter transposition on reading English and Hebrew. *Psychonomic Bulletin & Review* 14(5). 913–918. https://doi.org/10.3758/BF03194121

Watt, W. C. 1983. Grade der Systemhaftigkeit: Zur Homogenität der Alphabetschrift. *Zeitschrift für Semiotik* 5(4). 371–399.

Winskel, Heather. 2009. Reading in Thai: The case of misaligned vowels. *Reading and Writing* 22(1). 1–24. https://doi.org/10.1007/s11145-007-9100-z