Behavioral and Brain Sciences

cambridge.org/bbs

The puzzle of ideography

Olivier Morin[a,b]

[a]Max Planck Institute for Geoanthropology, Minds and Traditions Research Group, Jena, Germany and
[b]Institut Jean Nicod, CNRS, ENS, PSL University, Paris, France
morin@shh.mpg.de; https://www.shh.mpg.de/94549/themintgroup

Target Article

Cite this article: Morin O. (2023) The puzzle of ideography. *Behavioral and Brain Sciences* **46**, e233: 1–66. doi:10.1017/S0140525X22002801

Target Article Accepted: 5 September 2022
Target Article Manuscript Online: 18 October 2022
Commentaries Accepted: 14 February 2023

Keywords:
graphic communication; ideography; literacy; numerical notations; repair; semasiography; semiotics; sign language; writing

What is Open Peer Commentary? What follows on these pages is known as a Treatment, in which a significant and controversial Target Article is published along with Commentaries (p. 18) and an Author's Response (p. 56). See bbsonline.org for more information.

Abstract

An ideography is a general-purpose code made of pictures that do not encode language, which can be used autonomously – not just as a mnemonic prop – to encode information on a broad range of topics. Why are viable ideographies so hard to find? I contend that self-sufficient graphic codes need to be narrowly specialized. Writing systems are only an apparent exception: At their core, they are notations of a spoken language. Even if they also encode nonlinguistic information, they are useless to someone who lacks linguistic competence in the encoded language or a related one. The versatility of writing is thus vicarious: Writing borrows it from spoken language. Why is it so difficult to build a fully generalist graphic code? The most widespread answer points to a learnability problem. We possess specialized cognitive resources for learning spoken language, but lack them for graphic codes. I argue in favor of a different account: What is difficult about graphic codes is not so much learning or teaching them as getting every user to learn and teach the same code. This standardization problem does not affect spoken or signed languages as much. Those are based on cheap and transient signals, allowing for easy online repairing of miscommunication, and require face-to-face interactions where the advantages of common ground are maximized. Graphic codes lack these advantages, which makes them smaller in size and more specialized.

1. Introduction

In the novella *Story of Your Life* (adapted into the motion picture *Arrival*) the writer Ted Chiang (2016) imagines a species of aliens with rather strange habits of communication. The aliens produce a variety of grunts and cries unintelligible to humans, but their main mode of expression seems to consist of images. They use their tentacles to produce large circular inkblots arranged into patterns. This language baffles the linguist sent to initiate contact with the aliens, because the inkblots do not resemble any known human communication system. Unlike spoken language, the aliens' inkblots engage vision, not hearing. Unlike sign language, they are static, allowing the linguist to store them as photographs. And unlike human writing, the inkblots can be deciphered on their own: They do not bear any relation with the aliens' grunts and cries. This, the linguist notices, is exceedingly rare. Permanent images can be used for communication in many human cultures, but they usually fail to reach the degree of sophistication of a full-blown language. Whenever they do reach it, that is because permanent images are being used to encode a spoken language. The way the aliens communicate visually is puzzling.

Most linguists today would agree. And yet the aliens' visual language, or at least the possibility of it, would not have seemed so odd to a linguist from a different era. The notion of a complete language consisting entirely of images referring directly to ideas without encoding words was until fairly recently a commonplace. Western philosophers such as Leibniz or Bacon were convinced that Chinese characters or Egyptian hieroglyphs were *ideographic* (Rossi, 2000). That is, the meanings they encoded were thought to be understood directly by anyone literate in these symbols, even without knowing the Egyptian or Chinese language. This misconception has long been dispelled. Egyptian, Chinese, and Mayan, among other writing systems formerly assumed to be ideographic, have been shown to encode a natural, spoken language (if only among other things). Parallel to this, numerous attempts at building a universal ideography have failed. These ideographic languages proved exceedingly difficult to use for anyone, including their makers. John Wilkins's "philosophical language," Charles Bliss's Bliss symbolic, or Otto Neurath's picture language are the most famous examples (Lin & Biggs, 2006; Rossi, 2000). These multiple failures resulted in the widespread linguistic intuition, echoed in Chiang's short story, that full-blown ideographies are impossible.

Why? There would be, after all, many benefits to mastering an ideographic language. Such a system could exploit the iconicity of pictures to make the symbols' meanings more intuitive and easier to remember. It could transmit information across timespans and across space, which neither spoken nor signed language can do (unless backed by modern technology or by writing). It could break language barriers.

This is the puzzle of ideography: A uniquely rich mode of communication that most cultures seem to avoid.

In the evolution of communication, ideography is the road not traveled. If we can understand why, we will be in a better position to understand why writing evolved in the way that it did. Literacy is widely recognized as an epochal invention – arguably the most important technological innovation since stone tools (Coulmas, 2003; Goody, 1977; Morris, 2014). With a powerful graphic code like writing, modes of communication that were hardly possible – direct communication with distant people, with entirely unknown strangers, with dead people – can become routine (Morin, Kelly, & Winters, 2020). No serious account of cultural evolution can bypass it. But the first thing we notice when studying writing is how peculiar it is. Here is a mode of communication that seems to work almost entirely by parasitizing another mode of communication – spoken language.

One might think entire fields of research would be fighting to explain the puzzle of ideography. Instead, more energy has been spent on explaining the puzzle away. The first way to do this is to trivialize the puzzle: Writing cannot be ideographic, but that is simply a matter of definition, or it is because of some basic and obvious inability to think or communicate with pictures. The second way is to deny there is a puzzle: Ideographic writing exists, in the shape of emojis, Chinese characters, Bliss symbolics, pictographic symbols, and so on. The puzzle, I will argue, will not disappear in either fashion. But progress on a number of issues will be thwarted as long as the puzzle stands in the way. Solving the puzzle can help us trace the boundaries of human communication: It is clear today that we can express ourselves in many ways that language, narrowly construed, does not capture: Gestures, art, music, and so on (Heintz & Scott-Phillips, 2022; Schlenker, 2018; Wharton, 2009). But how far can communication go without language? Studying ideography can answer this.

Ideography can also teach us about the human brain's difficulties in dealing with visual codes such as writing. Our brains' visual areas can be recycled to process letter shapes by repurposing hardwired circuits that evolved to treat other stimuli (Dehaene, 2010; Dehaene & Cohen, 2007), helped by the fact that letter shapes are optimized to fit our visual brains' native constraints (Changizi, Zhang, Ye, & Shimojo, 2006; Kelly, Winters, Miton, & Morin, 2021; Morin, 2018). In spite of the flexibility of human neural and cultural resources, learning to read never became as natural as learning to speak, and remains a tall order for around 5% of the schooled population (Ramus, 2004; Wagner et al., 2020). Even for proficient literates, spoken or signed conversation remains much easier than reading or writing (Garrod & Pickering, 2004); the failure of ideography is another aspect of this struggle to master graphic codes.

Last but not least, understanding why ideography has not worked in the past may help us understand how technology could make it work in the future.

This paper pursues two related goals. It grounds and specifies the widespread intuition that ideographies, that is to say, general-purpose codes made of symbols that do not encode words, are extremely rare; and it explains this rarity.

Section 2 defines graphic codes, which are ways of encoding information with lasting inscriptions. It describes how these codes differ from spoken or signed languages, and argues that it is not helpful to think of graphic codes as languages. The next section (sect. 3) distinguishes writing, a graphic code that primarily encodes components of spoken languages, from ideography, which does not (graphic notations of sign languages do exist, but are culturally less significant). It has often been claimed that writing simply cannot be ideographic, either by definition or as a matter of fact.

Section 4 unpacks this widespread intuition and argues in favor of the "specialization hypothesis": The view that any graphic code that can be used in a self-sufficient way (not just as a memory prop) happens to be, at heart, a narrowly specialized notation. This applies to writing itself which, under this account, rests on an encoding of language.

Why do we not find graphic codes that are both self-sufficient and generalist? Section 5 considers two answers. One, the "learning account," is based on the notion that graphic codes are much harder to learn than spoken languages, because of a cognitive specialization for language acquisition. I argue against this account and I propose another solution in section 6. The "standardization account" considers that graphic codes may not be much harder to learn than spoken languages, but are instead considerably more difficult to standardize. To standardize a code is to ensure that all its users ascribe the same meanings to the same symbols. Spoken or signed languages are self-standardizing because they are based on cheap, fast, and transient signals, which allow for easy online repairing of miscommunication, and which constrain interlocutors to communicate in face-to-face interactions where the advantages of common ground are maximized. Being easier to standardize, spoken or signed languages have a "lock-in" effect on the evolution of other codes. This, I argue, solves the puzzle of ideography.

2. Languages are codes, but not every code is a language

This paper uses the words "code" and "language" in a rather restrictive way. There is a tendency, inherited from the semiotics research tradition, to use these two terms loosely and interchangeably, as referring to any means of expression that carries information: Paintings (Panofsky, 1939), comic books (Cohn, 2013), and so on. For the sake of this argument, different terminological choices were made.

A code is a set of conventional associations between meaning and signals (de Saussure, 2011; Scott-Phillips, 2014). Musical notations, road signs, writing systems, and so on are codes in this sense, and so are languages like Swahili, French, and so on. Means of expression that do not mostly rely on conventional signals do not qualify as codes. In particular those may be pictures, schemas, maps, and other forms of graphic communication that rely chiefly on the visual resemblance between graphic shapes

Olivier Morin is a CNRS researcher at the Institut Jean Nicod (PSL University, Paris), and leads the Minds & Traditions Research group at the Max Planck Institute for Geoanthropology in Jena. His research focuses on cultural transmission and its cognitive prerequisites. His book, *How Traditions Live and Die* (2016), seeks to explain the long-term survival of culturally evolved practices. His work with the Minds & Traditions Research group focuses on the cultural evolution of graphic communication – the human ability to transmit information by means of images. His publications span anthropology, cognitive science, and behavioral ecology. In 2016, he received the Early Career researcher Award from the European Evolution and Human Behavior Association.

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

and the things they refer to. Graphic communication interests us here only insofar as it relies on a code.

What is so special about codes? They make communication more efficient (Kirby, Tamariz, Cornish, & Smith, 2015; Regier, Kemp, & Kay, 2015; Scott-Phillips, 2014; Winters & Morin, 2019). Codes allow us to compress a long or complex message into a small number of symbols. A code works by relying on memorized associations between symbols and their meanings. Once the association is in place in someone's mind, the meaning can simply be triggered with the relevant symbol. It does not have to be explained again. As a result, conventional graphic symbols can afford to be much simpler than nonconventionalized ones.

This was shown elegantly in a series of experiments by Garrod, Fay, Lee, Oberlander, and MacLeod (2007), where participants engage in a Pictionary-style task repeatedly for the same referents. Two things happened. One is partial conventionalization: Drawings for referents such as "Opera" or "Brad Pitt" become standardized inside the group and increasingly difficult for outsiders to understand. The other is a clear simplification of the pictures. Of course, this process, taking an hour or so of experimental time, only approximates the history of real-life graphic conventions, but it captures their essential features: Graphic conventions are signal–meaning pairings that need to be learnt from the group that gave them their meaning. These can be used to compress complex messages into a few simpler shapes. Codification is the standardization of pairings between meanings and signals, making a mode of communication more efficient by making messages more compressible (Winters, Kirby, & Smith, 2018; Winters & Morin, 2019).

2.1. Not all forms of expression are codified

In theory, codification is a matter of degree. In the experiments just cited, the pairings between signals and symbols progressively become simplified and conventional, so that the associated meanings become increasingly opaque to outsiders who have not had access to previous steps. This is a gradual process, and an incomplete one: Many signals retain an iconic resemblance with their referents and are sometimes still transparent enough that their meaning can be accessed directly by outsiders. Familiarity with the previous steps of the game is a facilitator, not a requirement (Caldwell & Smith, 2012; Granito, Tehrani, Kendal, & Scott-Phillips, 2019). In that sense, conventionality can be partial.

It would be tempting to go one step further and assert that conventionality is a smooth continuum, to such an extent that any sharp distinction between conventional graphic codes (like writing systems) and nonconventional graphic expression (like artistic drawings) is bound to be moot. Two main arguments support this view, neither of them as convincing as they appear to be.

The first argument rests on the view that the visual arts make use of codified cultural conventions. This is well established (Cohn, 2013; Panofsky, 1939). An excellent case for the cultural conventionality of pictures was made by Neil Cohn in his study of comic books focusing on the contrast between Euro-American comics and Japanese manga (Cohn, 2013). Cohn shows convincingly that important aspects of Japanese graphic culture are akin to codes. One must learn them in order to understand mangas. In Cohn's view, the degree of codification of comics allows us to treat the distinct conventions of particular cultures as full-blown "visual languages." The phrase makes sense in the context of Cohn's study, but there are good reasons to resist it in general. Comic book drawings are not a language in the same sense that spoken or written Japanese or English are languages. The degree of conventionality is vastly lower for comic book drawings compared to languages. To see why, consider the amount of effort and time required to learn Japanese or English compared to the amount of learning involved in understanding the visual language of manga, or simply the fact that manga drawings require no translation, while their written text does.[1]

While the first argument emphasizes the conventionality of drawings, the second argument in favor of blurring the distinction between codes and noncodes highlights the iconic aspect of graphic codes and writing in particular. Iconicity is a multifaceted notion, but for the purposes of this argument, only one sense of it really matters: The capacity to know the meaning of a sign directly, without the preexisting knowledge of a code.

Iconicity in this sense should be distinguished from three related observations.

First, the fact that characters in some writing systems are figurative shapes – pictures of recognizable objects – does not make the system iconic. Indeed, figurative shapes may be quite unrelated to their coded meaning. Egyptian or Mayan hieroglyphics are a case in point. Even though many of these symbols are figurative depictions, their meaning is often quite different from what their iconic shape suggests. Their meaning is also coded: A naïve reader cannot make sense of them based on shape alone. No real progress was made in deciphering these writing systems as long as scholars assumed they could be read iconically (Pope, 1999).

The second observation is different. In some writing systems, there are symbols that were originally iconic, lost this quality with time, but kept traces of it – traces that can be deciphered with a bit of erudition. This is, famously, the case for many Chinese characters. In most cases, however, the pictographic meaning has become unrecognizable to an untrained eye, and the sign has acquired phonetic or morphemic values that cannot be retrieved directly from a picture.

The third and last form of iconicity is suggested by recent work on sound symbolism in letters (Turoman & Styles, 2017). This work suggests that letter shapes are not unrelated to sound shapes (Jee, Tamariz, & Shillcock, 2022), so that naïve subjects are better than chance at identifying which letters stand for the sounds /i/ or /u/, in unknown writing systems. If confirmed, such findings would show that the pairings between sounds and shapes that make up writing systems are not entirely arbitrary and possess limited iconicity. Still, there is a vast difference between the kind of above-chance guessing that these studies are interested in and the near-certain decoding that a fluent reader routinely achieves.

Thus, a few interesting boundary cases notwithstanding the distinction between graphic codes and noncoded means of expression (such as pictures) remains a crucial one. Codification is, in theory, a matter of degree: Some conventions are easily learnt with just a little familiarization, others cannot be deciphered without a hard-won knowledge of precise conventional pairings between symbols and meanings. In the case of graphic codes, the difference of degree between full-blown codes, like writing systems, and vaguely conventional graphic expression, like art, is vast enough to justify placing a clear boundary between codes and noncodes (Fig. 1).

If schemas, maps, comics, or paintings are not full-blown codes in the sense used here, it makes even less sense to call them "visual languages." In fact, the term "language" will be reserved here for spoken languages like Swahili, German, and so

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

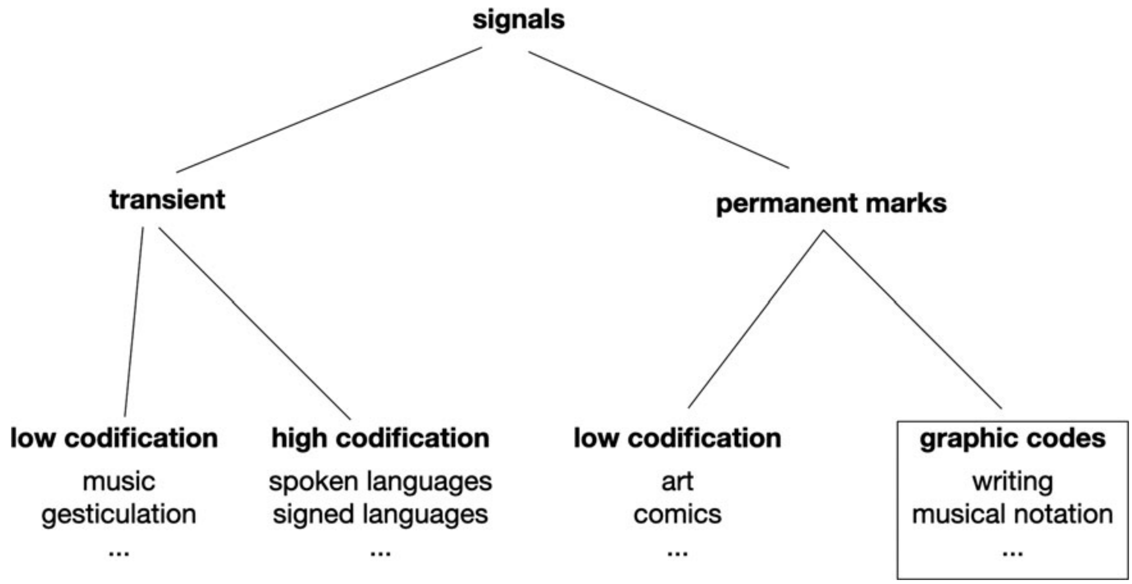

Figure 1. Place of graphic codes in human communication.

on, or signed languages like British Sign Language. Once again, this terminological choice is debatable. Some highly stimulating research is based on the premise that a wide variety of means of expression – diagrams, gestures, music, and so on – can be studied with the tools of the linguist (Schlenker, 2018). Without disputing this point, I find it convenient to reserve the term "language" for means of expression that are clearly codified, generalist, and self-sufficient.[2] This section explained what I mean by codification; the next one turns to generality and self-sufficiency.

2.2. Languages are uniquely generalist and self-sufficient

Consider musical or mathematical notations. These are fully codified graphic codes: They pair conventional graphic symbols with precise concepts (musical or mathematical). It has been noted (e.g., by Casati, 2016; McCawley, 1992) that these codes resemble language in many ways: They carry meaning, they combine a small number of symbols to generate messages, combinations of symbols obey syntactic rules, and so on. Why not count them as languages?

For the same reasons that I resist an overly extensive understanding of codes, the word "language" will only be used here to refer to spoken or signed languages in the everyday sense. This stresses the fact that languages possess unique properties. They are rich and complex codes, understood as conventional pairings of signals and meanings. Other rich and complex codes exist, but languages deserve to be singled out, being uniquely strong on two dimensions: Self-sufficiency and generality.

2.2.1. Self-sufficiency

A code is self-sufficient if we can use it to communicate with one another, without recourse to another type of code. Many codes share this property, from smoke signals to musical notations. Self-sufficiency is about usage, not acquisition: A code usually needs to be learnt linguistically before it can be used; but once fluent in the use of musical notations or smoke signals, a user may communicate without the help of auxiliary codes. This is not possible with codes lacking self-sufficiency. Few people claim to master Bliss symbolics to the point where they could produce and decipher symbols without constantly referring to a dictionary (i.e., to a linguistic gloss of Bliss's code). The fact that Bliss symbolics found its only niche as a teaching tool for children with special educational needs underscores this point (Sevcik, Barton-Hulsey, Romski, & Hyatt Fonseca, 2018; Sevcik, Romski, & Wilkinson, 1991): The ideographic code is used to make better sense of written or spoken language.

Language is self-sufficient in the sense that two people can communicate using a linguistic code and nothing else. Linguistic communication does not work in a vacuum, to be sure. It is backed by the common ground that interlocutors share; it can be complemented with gestures (some codified, some not). Some forms of linguistic exchange gain in clarity with the help of specialized codes (think of a scientific presentation in physics or mathematics). But even if we block all the side channels we use to pass information across (gestures, pauses, sighs, etc.), the linguistic code suffices to communicate a great deal. This is clearly not because all linguistic expression is immediately obvious and transparent – we regularly encounter rare or novel phrases whose opaque meaning needs unpacking, or translating. But in such cases the gloss can be given in spoken language,[3] thanks to the reflexive nature of language – its capacity to talk about itself (Taylor, 2000). In that sense, language is both self-sufficient and self-decoding.

2.2.2. Generality

A code is general if it can be used to encode information across a wide range of domains. Mathematical or musical notations are not general in this sense, but language is. "Generality" here is equivalent to what Liberman (1992) called "openness" (see also DeFrancis's, 1989, distinction between full and partial writing). I endorse his claim that languages are uniquely open:

> Spoken language has the critically important property of "openness": unlike nonhuman systems of communication, speech is capable of expressing and conveying an indefinitely numerous variety of messages.

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

> A script can share this property, but only to the extent that it somehow transcribes its spoken-language base. (Liberman, 1992, p. 120)

Liberman's second sentence considers the possibility that writing systems could also qualify as generalist (or "open") codes. His answer, which I will endorse and develop, is that they can – but only to the extent that a writing system encodes a language. The generality of writing is vicarious, derived from the language that it encodes. This classical view (Aristotle, 1962 (ed.)) considers writing to be a meta-code, a higher-level code that encodes another, more basic code. The first-level code is language, which associates meanings with spoken or signed symbols. Writing, made possible by the reflexive nature of language (Taylor, 2000), relies on our capacity to name and classify linguistic objects (and also improves upon it; Olson, 1991). It is the meta-level code that associates graphic symbols with the spoken or signed symbols of a language. It does not usually encode meanings directly, but does so by encoding spoken or signed symbols.

To say that writing encodes ideas only indirectly by encoding language is to oppose writing to ideography. But what exactly is ideography?

3. Ideography and writing, contrasted

3.1. What is ideography?

Ideographic symbols can be directly associated with a concept by a reader acquainted with the sign's meaning (following Coulmas, 1996a, p. 309). Examples in use in English writing include Arabic numerals (1, 2, 3, 4,…), as well as conventional symbols such as £, ±, ☺, ℃, ©, &, and so on. These symbols can be verbalized differently in several entirely distinct languages, and they contain no clue to their pronunciation in any of these languages. They are "translinguistic" (Chrisomalis, 2020). For this reason, such symbols can be said to encode ideas directly. Ideographic symbols are not necessarily pictographic or iconic. Some of them are (consider the Chinese numerals 一, 二, 三) but in general they need not resemble in any way the idea that they refer to, and their association with their referent can be entirely conventional (consider the signs "$" or "+").

Most ideographs can be verbalized, but not read: They do not encode linguistic information. Although ideographic symbols do not tell readers how to verbalize them in a given language, a competent speaker can, of course, match them with the corresponding words in her language. A Spanish reader can verbalize "1, 2, 3" as "uno, dos, tres," even though the ideographs do not provide her with any clue concerning the pronunciation of these words. The ♡ symbol can be verbalized as "love," "heart," "lots of love," "I love you," and so on, depending on context.

This property has been used to back the claim that ideographs indeed do encode words, but do so indirectly (Boltz, 1993; Taylor, Taylor, & Taylor, 1995). In an extreme version of this claim, Boltz argued that early Chinese pictographs (before the emergence of Chinese writing proper) were already representing the Chinese language of the time, because they could be verbalized in Chinese in various ways, even though they did not stand for specific individual words. This view should be resisted (Coulmas, 1996b; Hyman, 2006), because it blurs the distinction between those symbols that contain cues concerning the pronunciation of a word in a language and those that do not. A symbol like "9" that can be *verbalized* in a multitude of languages (as "nine," "nueve," "kilenc," or possibly "ninth" or "noveno," etc.) is not the same thing as a symbol that can be *read* (Hyman, 2006). The string of letters "nueve" can be read and it means /nweβe/, a specific spoken Spanish word. Unlike "9," it cannot encode the words "nine," "ninth," "noveno," or "kilenc." Hence, ideographs as defined here do not encode language words simply because they can be verbalized.

Use of the rebus principle is not sufficient to make ideographic symbols glottographic. Many ideographs still intuitively match one or a few spoken words in most languages in relatively straightforward ways. This property is of no small historical importance. Thanks to the rebus principle (e.g., "" = eye/can/reed = "I can read"), ideographic symbols can be made to represent phonemes. Rebus-based encodings of words are not rare, and played a key role in the emergence of glottography (Valério & Ferrara, 2019). Glottography will be defined later; it is, in short, the fact that writing, as a code, represents language. Is rebus sufficient to make a graphic code glottographic? Not if its use is neither systematic nor standardized. A series of haphazard rebus or visual puns does not amount to a systematic encoding of a language. European heraldry is a good example of a code that made frequent use of the rebus principle, but remained ideographic at heart (Pastoureau, 2007). Heraldic symbols (coats of arms), used as emblems by families or institutions, were purely ideographic most of the time, for instance symbolizing the Kingdom of Prussia with a black eagle. However, some arms ("canting arms") used the rebus principle to encode proper names (or parts thereof): The arms of Castile and León feature a castle and a lion, the arms of Berlin, a bear. But this practice was neither systematic nor standardized: Depending on the country, the arms, or the viewer, a bear could encode the corresponding sound, but it could just as well be simply a picture of a bear devoid of phonetic meaning. Neither were rebus-based encodings standardized: The sound /beə(r)/ could be encoded by one bear or by three, or by another image with the same phonetic value.

Ideographic symbols thus can be defined as symbols whose standardized and coded meaning does not include linguistic information. They may be used to convey such information indirectly (for instance, through the rebus principle), but in order to get that information, recipients need to possess linguistic knowledge (e.g., the sound of the German word *Bär*) that is not encoded by the symbol.

An ideographic graphic code (also known as "semasiography" – Boone & Mignolo, 1994; Croft, 2017; Gelb, 1963) is simply a set of ideographic symbols used in conjunction with each other. Examples include mathematical or musical notations, some shamanistic pictographic notations (e.g., Severi, 2012), heraldic emblems, commodity brands (Wengrow, 2008), formal logic, among others. These systems of symbols may be used by people who do not share a common language. Contrary to what has sometimes been claimed (e.g., Boltz, 1993; du Ponceau, 1838; Hill, 1967), such systems are neither impossible, nor are they contradictions in terms. Countless ideographic codes existed long before and after the rise of writing (Lock & Gers, 2012), ranging from tallies, property marks, and tokens to pictographic stories like Winter Counts (Mallery, 1886) or Aztec codices (Boone, 1994) that recounted sequences of events using images.

3.2. A generality/self-sufficiency trade-off

Section 2.2 argued that, of all the codes we can use for communication, languages (spoken or signed) stand out for being both self-sufficient (they can be used on their own without resorting to an

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

auxiliary code to gloss each message) and generalist (they can be used to talk about an indefinite variety of topics). I argued that this combination of self-sufficiency and generality was unique to language, putting aside writing as a possible exception that nonetheless seems intricately tied to language. This section and the next detail this claim. This section explores the trade-off between generality and self-sufficiency: Very few graphic codes seem capable of combining these two properties. The main exception seems to be writing, which the next section introduces.

Figure 2 classifies the codes we use for communication (leaving out writing for the moment) along two dimensions: Self-sufficiency and specialization. Most of the graphic codes you and I are familiar with are specialized: This includes mathematical or musical symbols, counting tools, and so on. The symbols in specialized codes may (for some codes) be combined productively according to clear and well-standardized rules to yield vast numbers of possible messages.

In spite of this, these codes remain limited in the range of topics they can tackle. Some serve but one narrow function and are strictly circumscribed to one domain: To record measurements, to encode music, to make a population census, to record a debt, to serve as emblems for families, and so on. Others are apparently more versatile – for instance the international airport signs for "toilets," "wifi," "luggage," and so on – but the number of symbols they contain is too limited to allow them to serve as a generalist code. Graphic codes are not the only kind of code to be limited by specialization: Some gestured languages are similarly restricted in their use – for instance, the specialized sign languages used by hunters (Mohr, 2015) when they must be silent, or by workers in noisy environments (Meissner, Philpott, & Philpott, 1975).

Specialist codes differ in how self-sufficient they can be. If a code is self-sufficient, a proficient user does not need an oral gloss to understand every message. An oral gloss is usually needed to learn the code, but not to use it. Mathematical or musical notations may be read in the same way that one reads print; today's most important corporate brands are recognizable without intermediates (linguistic or otherwise). Heraldic emblems could be recognized without being glossed (although large gatherings required arms to be glossed by professional heralds, often with the help of specialized directories).

A clear example of a family of specialized codes lacking self-sufficiency is provided by the early history of musical notations, as studied by Croft (2017). The earliest known inscriptions that provide instructions for the performance of a musical piece (for instance, Babylonian lyre or harp notations) are exceedingly hard to interpret, because they require implicit knowledge of the musical piece that is not encoded – the tune's rhythm, for instance. The ancestors of modern Western notations, staffless neumes, were also lacunar, leading specialists to argue that they served as mnemonic tools for melodies that were orally transmitted. These early musical notations were neither self-sufficient nor generalist. In time, they evolved to become self-sufficient, but remained narrowly specialized.

If we switch to the second column of the table, to generalist codes, we find that the vast majority of generalist graphic codes rely on an oral gloss to function. Australian message sticks, for instance, were sophisticated ideographic messages that could communicate information on a wide range of topics, as long as the messenger stood by to translate the code. Although message sticks were occasionally sent through the mail with no accompanying gloss (Kelly, 2019), this was not at all their typical use. Message sticks, in other words, are not self-sufficient codes: Most of the information they impart is not entirely encoded in the graphic message, which serves instead as a mnemonic prop. The same has been said for the pictographs used in recitation of shamanistic chants, for instance those of Cuna shamans

	specialized	**generalist**
self-sufficient	<u>**Graphic codes:**</u> **Modern musical notation** **Modern mathematical notations (including logic)** **Coin designs** **Heraldic emblems[1]** ...	<u>**Graphic codes:**</u> **Few clear examples (Nsibidi symbols[3]?)**
	<u>*Other codes:*</u> *Hunting signs (e.g. Botswana[2])* *Sawmill workers' signs[2]* ...	<u>*Other codes:*</u> *Spoken languages* *Signed languages (e.g., ASL, FSL...)* *Signed languages of hearing communities* *(e.g. Monastic sign languages[4], "Plain Indians sign language"[5])*
not self-sufficient	<u>**Graphic codes:**</u> **Early musical notations[6]** **Kupesi symbols[7]** **Sand drawings (e.g. Australia[8], Vanuatu[9])** ...	<u>**Graphic codes:**</u> **Bliss symbolics[10]** **Pictographs (as used by, e.g., Cuna shamans[11])** **Australian message sticks[12]** **Yukaghir drawing game (the "love letter"[13])** ...
	<u>*Other codes:*</u> *Few clear examples*	<u>*Other codes:*</u> *Gestured languages for hearing users, accompanying conversation — e.g., Arandic sign language[14]...*

1. (Pastoureau 2007); 2. (Mohr 2012); 2. (Meissner et al. 1975); 3. (Battestini 2006); 4. (Banham 2012, Quay 2012); 5. (Mallery 1886, Davis 2015); 6. (Croft 2017); 7. (Bell & Paegle 2021); 8. (Green 2014); 9. (Zagala 2004); 10. (Lin & Biggs 2006); 11. (Severi 2019); 12. (Kelly 2020); 13. (Sampson 1985, DeFrancis 1989, Unger 2003); 14. (Green & Wilkins 2012).

Figure 2. Typology of graphic codes, illustrating the trade-off between specialization and self-sufficiency that graphic codes face. Writing has been left out: The question whether it is specialized or generalist is answered in section 4.

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

studied by Severi (2012, 2019), or for sand drawings as used in several Pacific societies (Green, 2007; Zagala, 2004). Such codes are generalist, in the sense that the scope of all the things one may refer to using the code is rather broad; but they lack self-sufficiency.

Could we find an ideographic code that is both generalist and self-sufficient? At this stage of my argument, I have no theoretical reasons to deny this. In practice, however, examples do not come easily. One of the clearest cases (that I know of) is Nsibidi pictographs, a system of symbols in use in the Cross River region of Nigeria (Battestini, 2006; Dayrell, 1911; Griaule & Dieterlen, 1951; Macgregor, 1909). Nsibidi symbols, the preserve of a secret society (at least initially), can be gestured or inscribed; when inscribed, they can carry simple messages that recipients can understand without an oral gloss. How far the range of expression of these symbols goes is difficult to determine, given the secrecy that surrounds them. One clear limitation comes from the fact that many symbols are inaccessible below a certain level of initiation. Other examples are few and far between, with one obvious exception: Writing.

3.3. Defining writing

Writing is a versatile code, capable of encoding information on a broad range of content, and it can be used in a self-sufficient way – as you and I are using it now – to convey information across time and space without the help of an oral gloss. But how do we define writing? There are countless definitions. Some are so broad that they encompass anything that I call here a graphic code: That is what Gelb's definition does (writing is "a system of human intercommunication by means of conventional visible marks" – Gelb, 1963). But most definitions of writing oppose it to ideography (e.g., Coulmas, 2003; Daniels & Bright, 1996; DeFrancis, 1989). For those authors, writing is at heart a notation of language, even if it is only partially a notation of language.

Should we care? Definitions are cheap: Saying that writing encodes language because that is how we define it cuts little ice. Yet in this case scholarly conventions harbor an empirical truth that is anything but trivial. In the next section, I will argue that most and perhaps all self-sufficient and general-purpose graphic codes used by humans are notations of a language. This empirical claim is part of what I call the *specialization hypothesis*.

4. The specialization hypothesis

This hypothesis, in its most general form, claims that all self-sufficient and well-standardized graphic codes, including writing systems, are highly specialized notations. Unlike languages, which can encode all sorts of thoughts, self-sufficient and standardized graphic codes specialize in one or a few specific types of information: Numbers, logical connectors, personal emblems, the sounds of a language, and so on. The most important consequence of the specialization hypothesis is that writing systems, the most powerful and widespread of graphic codes, are specialized notations in spite of the wide range of uses they can be put to. A writing system, in this view, is at heart a specialized notation of a language. The generalist scope of writing systems is derivative: They inherit their versatility from the language that they encode. As a code, writing is narrowly specialized: Merely a notation of morphemes, syllables, or phonemes. One proof that writing is not actually a generalist code is given by liturgical or religious texts, which can be learnt and read by people who do not understand the target language (see, e.g., the Quranic recitations described by Scribner & Cole, 1981). These reciters know the writing system and the phonology that it encodes, but not the underlying language.

The specialization hypothesis differs from standard language-centric views of writing in a number of ways. The view that writing is mainly an encoding of spoken language is quite commonplace, dating back to Aristotle at least (Aristotle, 1962 (ed.); de Saussure, 2011), but this classical formulation was only a definition, not a strong empirical claim. Saussure's position on this matter is representative: While treating writing exclusively as a representation of spoken language, he also believed that purely ideographic forms of writing existed (e.g., Chinese characters). The specialization hypothesis is stronger. It casts doubt on the existence of any self-sufficient, generalist ideography.

In this respect, the specialization hypothesis agrees with the language-centric views of writing put forward by critiques of the ideographic interpretation of Chinese or Egyptian writing (e.g., DeFrancis, 1989; du Ponceau, 1838; Unger, 2003). This critique was spurred by three realizations. First, there was a growing awareness of the importance of morphemic and phonetic notations in scripts traditionally thought to be ideographic, like Chinese writing. The second trigger was the failure of attempts to build purely ideographic systems like Bliss symbolics. Lastly, critics like DeFrancis showed that ideographic systems used in mostly illiterate societies were not self-sufficient, but instead relied on an oral gloss. The *locus classicus* for this demonstration is the so-called "Yukaghir love letter," which Sampson (1985) presented as an ideographic message couched in a complex pictographic code. The letter was in fact no letter at all, but part of a parlor game whose participants had to guess the meaning of the cryptic message through a series of yes-or-no questions (DeFrancis, 1989; Unger, 2003). A closer look at other instances of pictographic communication, once presented as ideographies (or "semasiographies" in Gelb's terminology) in classic works (e.g., Gelb, 1963; Sampson, 1985) reveals a similar picture: Ideographic notations are heavily reliant on oral glosses, calling into question their capacity to encode a lot of information on their own (DeFrancis, 1989). This new interpretation of pictographic messages came at the same time as a series of important anthropological studies stressing the role of orality in traditional pictographic communication (Boone & Mignolo, 1994; Severi, 2012). These landmark findings transformed our understanding of pictographic communication. They also widened the gap between writing and other graphic codes.

4.1. What glottography means

The glottographic principle (also known as phonography; Gelb, 1963; Hyman, 2006) is the use of symbols to indicate linguistic information at the phonological level: Phonemes, syllables, or morphemes. Unlike ideography, the glottographic principle does not allow the direct encoding of semantic information, bypassing language. As a result, a code that makes heavy use of the glottographic principle is useless to someone who does not know the particular language that it encodes, or at least a closely related one.

How much use should a system make of the glottographic principle to count as writing? No writing system is glottographic through and through. Many systems use ideographic symbols (as in "$1"). And written representations may be richer than the spoken linguistic representations that they encode: In English, "be" and "bee" are less ambiguous than the spoken sound /bi(ː)/.

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

Because of this, writing will sometimes represent information through purely graphic cues that have no counterpart in language. Any writing system will occasionally carry information that is absent from the spoken form.

If glottography is but one aspect of writing, does this refute the specialization hypothesis? Not if the vast majority of written symbols (in contemporary systems at least) encode linguistic units (phonemes, syllables, or morphemes). Is this true?

Of the alleged counterexamples that come to mind, Chinese characters are the most famous. Chinese writing would refute the specialization hypothesis if it were true that most of them (and the most frequently used among them) primarily encode semantic information without the help of a phonetic notation. The debate on the nature of Chinese characters is not fully settled (Handel, 2015; Lurie, 2006; Sampson, 2017; Unger & DeFrancis, 1995). Yet there is a broad and robust consensus around the view that (in the words of a critic of phonocentric views), "the vast majority of Chinese characters contain phonetic elements" (Handel, 2015, pp. 117–118); indeed, "nobody is disputing the role that phonological components play in the Chinese writing system or the role that phonological recoding plays in the reading of Chinese" (Handel, 2015, p. 130 – see also Coulmas, 1996a; DeFrancis, 1989; Sampson, 1985, 2017; Unger & DeFrancis, 1995). Nor is this phonetic information inert: There is massive psychological and neuropsychological evidence that Chinese readers process writing using phonological cues (Dehaene, 2010; Li, Peng, Liu, Booth, & Ding, 2014; Liu, Vermeylen, Wisniewski, & Brysbaert, 2020).

Having said that, it is still possible that Chinese characters encode language in a way that is quite different from alphabetic or syllabic systems, that is to say, mostly at the level of morphemes instead of phonemes or syllables (Handel, 2015). If true, this would set Chinese writing apart in an interesting way, because morphemes do carry meaning, unlike syllables or phonemes that are semantically empty. Morphemic encoding makes some sense of the intuition that Chinese writing is somehow less phonetic or more ideographic than, say, an alphabet.

The morphemic encoding hypothesis is hard to evaluate, chiefly because no writing system ever sticks to one single organizing principle all the time. English orthography is occasionally logographic: Sometimes, it encodes language at the level of words (compare the written forms *write*, *right*, and *rite*), even though it is alphabetic at heart; Chinese writing presents many syllabic features (DeFrancis & Unger, 2009), even though it cannot be reduced to a syllabary.

The same reasoning applies to writing systems that adapt Chinese characters to encode another language (Coulmas, 2003). Consider the case of kanjis in Japanese writing (Matsunaga, 1996): Only a minority (around 7%) is used in a properly ideographic way, that is, to refer to a unique concept that the two languages verbalize differently. (In the same way that English and French scripts use the character "9" to mean the number verbalized as "nine" or "neuf.") Most kanjis either admit a variety of other readings in addition to their ideographic reading, or no such ideographic reading. Another possible reading is logographic. In those cases, the respective kanjis encode a word of a precise language (usually, Chinese as pronounced by the Japanese at the time and place when the character was introduced). Yet, here again, most Japanese words cannot be encoded by their own distinctive kanji. Writing them down either requires the use of a syllabary or the use of kanjis employed for their phonetic value. The same point is true of many literate cultures that adopt and adapt foreign scripts: Such adoption would not be possible without either literate bilingualism, or the use of special glosses to transcribe the new script into the vernacular (Whitman, 2011). With a few exceptions (like the numerical notations that Latin scripts borrowed from Arabic), it is rare for a script to use symbols from another script purely for their meaning, without learning the corresponding spoken form or glossing it in the local language.

Inside the Chinese language family, it is often claimed that Chinese writing enables speakers of mutually unintelligible languages to communicate, because it encodes morphemes in addition to sounds. This claim can be broken down into several notions, some true, others debatable. First, written standards are factors of linguistic unification, in China as elsewhere, because writing can be understood by speakers whose differing pronunciation would hinder mutual comprehension, and because the written standard helps in the diffusion of a unified vocabulary and grammar (Coulmas, 2003). Second, mutual comprehension can be assured by a language that is quite different from most (or any) vernacular and is only ever used in a literate context – like Latin in Europe ("diglossia"). This second factor is arguably far more important than the first in a language family as diverse as the Chinese one. Written Chinese was a literate idiom, for at least some literate Chinese, for most of its history, until it was simplified and oralized, attaining its status as *lingua franca* (Li, 2006). Lastly, the morpho-syllabic nature of written Chinese does allow its users to read some characters correctly even when they would pronounce it quite differently. However, morphemic notation only goes so far in helping this. It works to the extent that the two languages have a closely overlapping grammar (at least), and many closely related cognates (Chen & Ping, 1999; Li, 2006). Modern written Chinese cannot, for instance, encode Cantonese without modifications (Chen & Ping, 1999). Cantonese is developing a writing system of its own, with specific conventions (Bauer, 2018; Snow, 2008).

4.2. The limits of glottography

What matters, from the point of view of the glottographic principle, is the simple fact that written symbols encode linguistic information. Surprisingly, perhaps, the glottographic principle does not imply that writing systems encode sounds. Writing is not a record of speech or a phonography (*contra* DeFrancis, 1989), because writing systems encode morphemes, syllables, or phonemes, which are not sounds but contrastive categories. Of all the systems that we know, only Korean Hangul attempts to encode actual features (e.g., whether a consonant is palatal or not, etc.) (Coulmas, 2003). Other systems encode language at the phonemic level, or above it. In hearing individuals, phonemic awareness is a predictor of literacy acquisition, in keeping with the glottographic principle (Mattingly, 1972), for a broad range of scripts, including Chinese (McBride-Chang et al., 2005, 2008; Verhoeven & Perfetti, 2022). Having said that, the fact that writing encodes abstract linguistic categories as opposed to sounds opens the possibility that one could become literate in a language when one's only contact with that language is visual – through writing, fingerspelling, or lip reading (Hirshorn & Harris, 2022; Petitto et al., 2016). A close approximation of this case is provided by the minority of persons born with deep congenital deafness who nonetheless become literate (Hirshorn & Harris, 2022).

The glottographic principle, to qualify it further, is compatible with the view that reading in proficient readers rests on a broad

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

variety of mental representations, mapping written signs onto phonemes, syllables, morphemes, or (occasionally, for frequent expressions) whole words (Perfetti & Harris, 2013). Indeed, proficient readers follow two routes in accessing the meaning of a text: One that connects written words directly to their meanings through associated phonological representations, and one that connects written words to meanings without going through this phonological stage, with the two routes working in parallel, some writing systems relying more heavily on one or the other route (Harm & Seidenberg, 2004; Hirshorn & Harris, 2022; Ramus, 2004).

4.3. The case of early writing

The specialization hypothesis implies a straightforward prediction regarding the graphic codes that preceded the rise of writing: They should be lacking in self-sufficiency or generality, or both. Looking at the four civilizations that invented writing independently of one another (China, Egypt, Meso-America, Mesopotamia), it becomes clear that in three of these four cases the emergence of writing was preceded by sophisticated specialized codes. Sumer is the clearest case. Proto-cuneiform was a poorly standardized and narrowly specialized code that lacked most of the features of glottographic writing (Damerow, 2006), but was preceded and accompanied by sophisticated accounting tools (Schmandt-Besserat, 2007). Likewise, among ancient Egyptians writing was preceded by a rich system of signs, mostly used to mark goods or commodities (Baines, 2007). The Maya also had sophisticated systems of symbols encoding proper nouns before the rise of any more fully glottographic writing (Houston, 2004). As for the exception, ancient China, the lack of data before the period of oracle bone inscriptions, c. 1400 BCE, does not let us know much about the script's evolution (Wang, 2014). Thus, inventing sophisticated special-purpose graphic codes appears to be a necessary but not sufficient condition for developing writing (with one intriguing exception where the evidence is inconclusive). Numeration systems, tallying and accounting tools more generally illustrate this most clearly, because they tend to develop in state societies before the rise of writing, or in its absence (Chrisomalis, 2020).

5. One puzzle and two solutions

The specialization hypothesis implies that general-purpose ideographies are exceedingly difficult to use, and unlikely to gain currency. Because ideographies are conceptually possible, one may still invent a general-purpose ideography, just like George Bliss or Otto Neurath did. But these systems will not be used in an autonomous fashion, without the help of a written or oral gloss. In contrast, attempts to engineer a new spoken language (like Esperanto or Volapük) did not fail as languages, even though they did not become the universal languages their inventors hoped they would be (Okrent, 2010). They have (or had) communities of speakers (including native speakers) comparable in size to those of many regular languages.

Why can visual languages not be turned into self-sufficient and generalist communication devices? Graphic codes can be self-sufficient, like mathematical notations, or they can be generalist, like mnemonic pictographs. But the specialization hypothesis contends that they cannot be both at the same time: Mathematical notations are highly specialized, whereas mnemonic pictographs require an oral gloss. What the specialization hypothesis does not do is explain why this is so. This section reviews two possible reasons for the failure of general-purpose ideographies.

5.1. Unpacking the puzzle of ideography

A full ideography would combine four advantageous features: It would be generalist, language-independent, asynchronous, and visual. Each of these features is present in extant communication devices, but none combines them all.

Language-independent, visual, and generalist codes are communication devices that can be used by people having no language in common. The clearest (though poorly documented) cases are signed languages used in multilingual hearing populations, such as the "Plain Indians Signed Language" (Davis, 2015; Mallery, 1879), said to have served as a visual communication tool crossing language barriers, all over the mid-Western area of the contemporary United States. The signed languages developed in silent monastic communities (Banham, 2015; Quay, 2015) belong to this category, although they were arguably less generalist and expressive, consisting in hundreds of symbols at the most, with little in the way of syntax or morphology. The potential of visual languages (gestured or visual) to bypass the barriers of language has long been recognized: People can use these codes without sharing a spoken idiom (Knowlson, 1965; Rossi, 2000).

Such gestured codes (distinct from sign languages, among other things because their users can hear and speak) are rare in the historical record, probably because most of their functions can be filled by spoken pidgins. They also lack one feature that would make ideographies uniquely useful: Asynchronous use. The crucial advantage of graphic codes, compared to signed or spoken languages, is that, in our species' history, they were for a long time the only kind of code that allowed sending messages across time or space (Morin et al., 2020). Asynchronous messages are "temporally and spatially portable" (to use the terminology of Pickering & Garrod, 2021), or "location and time independent" (to use that of Lee & Karmiloff-Smith, 1996). The impact of asynchronous communication on cultural evolution is twofold. First, it allows information to be transmitted in one single step across potentially unlimited temporal and spatial distances, without the need for long transmission chains, which tend to lose information (Bartlett, 1932; Tamariz & Kirby, 2015). Second, it allows one single message to transmit the same piece of information multiple times, in contrast to spoken or signed messages, which do not endure and must be continuously reproduced.

So why, despite all these potential advantages, do we not communicate with ideographies? Two broad families of explanations will be reviewed. The first starts from potential cognitive difficulties raised by the learning and memorization of graphic codes (the "learning account"). Explanations of the second kind are based on the difficulty of standardizing the codes we use for communication, when communication is not face-to-face (the "standardization account"). Both explanations imply that graphic codes consisting of a small number of symbols and rules can be learnt, thus allowing for the possibility of highly specialized codes, but more generalist codes cannot (Fig. 3).

The learning account and the standardization account are both consistent with the specialization hypothesis. On both accounts, rich graphic codes using a number of symbols vast enough to rival the richness of languages cannot evolve. On the learning account, that is because graphic symbols are difficult to learn

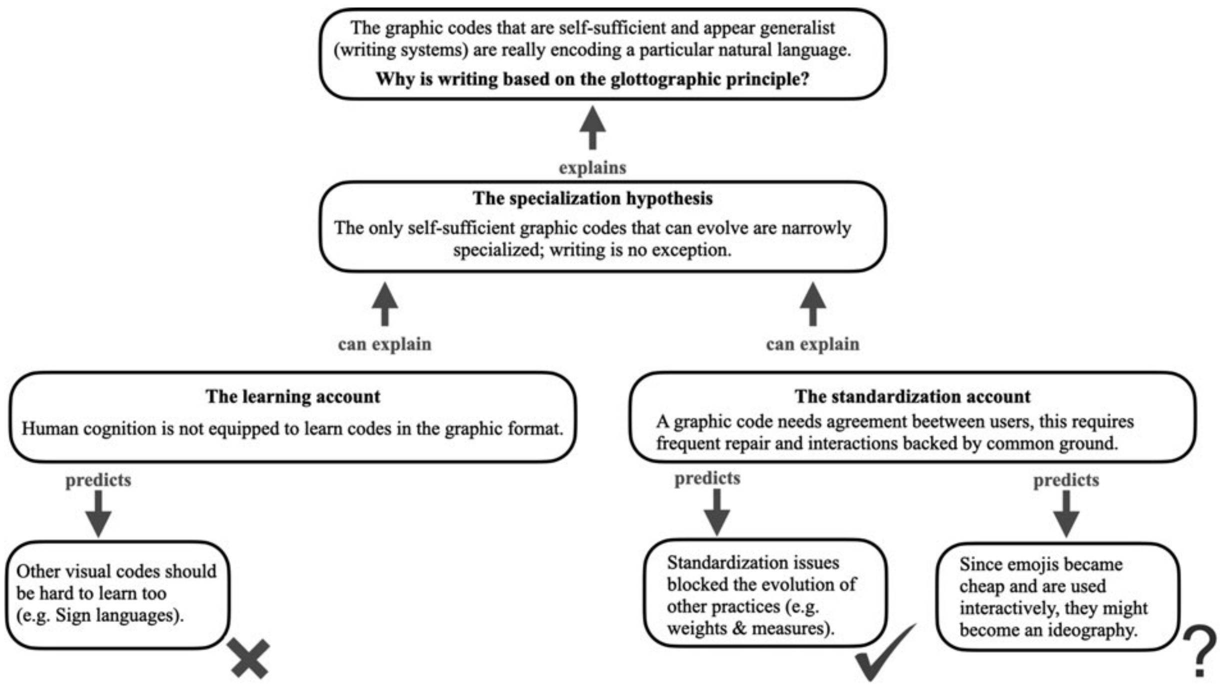

Figure 3. Argument of sections 5 and 6.

(compared to strings of phonemes, or bundles of gestures). On the standardization account, that is because of the difficulty of coordinating usage on a vast range of graphic conventions (as compared to the conventions that govern face-to-face communication: Spoken and signed languages). Under both accounts, graphic codes cannot encode a broad range of meanings, so there are only two ways for them to convey information: To rely on an oral gloss, or to be highly specialized. In other words, self-sufficient graphic codes that do not rely on oral glosses are necessarily highly specialized. Writing, an apparent exception to this rule, actually proves it. Writing is a specialized notation of language. Because languages are themselves general-purpose, writing benefits from this property of language vicariously.

The way that the two accounts solve the puzzle of ideography is broadly the same: Graphic codes are specialized because they are limited, and writing is glottographic because it is the only way for a graphic code to be both specialized and all-purpose. But the two accounts take quite different paths to reach this conclusion.

5.2. Graphic codes as a challenge for human cognition

The first family of explanations – the learning account – posits that human cognition has problems dealing with visual communication: Static images as opposed to gestures or strings of phonemes (Jakobson, 1964; Liberman, 1992). Its best proponent was the linguist Alvin Liberman (1992). Liberman was struck by the ease and naturalness with which we learn spoken languages, compared to the acquisition of reading and writing. Speech is universal; it is older than writing, phylogenetically and ontogenetically. Literacy, a localized and contingent cultural artifact, has no biological basis specifically evolved to support it (Dehaene, 2010). But stopping there would beg the questions: Why does spoken language benefit from a specific biological adaptation? What is the biological adaptation that makes spoken language, but not its graphic counterpart, so easy to acquire? Obviously, humans have been speaking for much longer than they have been writing, but then again, we need to know why writing evolved much later, and much more rarely, compared to speech.

Liberman posited that speech relied on an adaptation for phoneme perception, which worked for phonemes and only for them. Letters are not phonemes, and that is why graphic codes are difficult to learn. Can this hypothesis solve the puzzle? Before answering, I will review the things that the learning account, in my view, gets right.

5.3. Self-sufficient graphic codes must use a small number of symbols

The learning account clearly points at an important problem that graphic codes encounter, and that Liberman noted: They do not seem to possess nearly as many signs as the number of words in spoken languages, suggesting difficulties in learning a large set of graphic symbols.

The graphic codes that can be used to communicate a great deal of information without the help of an oral gloss, such as mathematical notations or writing systems, are based on a relatively small number of conventions. These conventions specify which meanings are paired with each symbol, how symbols can be combined with one another, and how to derive the meaning of a string of symbols from the meanings of the individual symbols that compose it. In the most regular graphic codes, like mathematical notations, a small number of conventions fixing the meaning of symbols is sufficient to make a great variety of messages possible to produce and to comprehend. The meaning of a mathematical expression like "2 + 2 = 4" is entirely and unambiguously given by transparent and standardized rules codified by mathematicians. Graphic codes like mathematical or musical notations possess clear syntax-like properties (Casati, 2016; Friederici, 2020; McCawley, 1992).

The orthographies of most writing systems are not as transparent and regular as this (far from it) but, as we saw, when compared to spoken languages, the number of meaning–symbol mappings that must be learnt in order to master even a complex system like Chinese characters is small relative to that of spoken language, thanks to the glottographic principle. Self-sufficient graphic codes manage to make the most of a few learnt conventions.

Attempted generalist ideographies, like Bliss symbolics, struggle to express as broad a variety of meaning as language does, in part because of the large number of conventional symbols that one would need to learn in order to make the system work, and in part because the rules that are supposed to help compose complex expressions from simpler symbols are too ambiguous. Consider the last symbol in the Bliss sentence given in Figure 4. The arrow at the end modifies the symbol that means "camera," to create a compound meaning "moving picture," that is, "film." But figuring this out requires a great deal of familiarity with Bliss. Why cannot we interpret the arrow as having a directional meaning, as in "I want to go see a picture," where the arrow would encode "to"? The grammar of Bliss is often not systematic enough to answer questions like this univocally, resulting in sentences too ambiguous to be understood without an oral gloss.

One way to understand the failure of general ideographies like Bliss symbolics is to frame it in terms of McNeill's demarcation criteria for language. According to McNeill (1992), linguistic communication differs from gestural communication in four ways. Linguistic messages have a specific hierarchical structure (they can be broken down into parts following a specific arrangement); their units can be combined productively; they can be understood out of context; they obey standards of form. In theory, Bliss has rules for ordering words, but their proper application is difficult, hence unlikely to be consistent; the production of combinatorial messages through compositionality is problematic, as just seen; decontextualized understanding (absent a written or an oral gloss) is seldom achieved or even sought. Emojis fail to function as a general visual language for the same reason (Gawne & McCulloch, 2019).

One possible explanation could be that graphic codes, in general, are simply incapable of fulfilling McNeill's criteria; but this is clearly false. Many graphic codes obey strict rules for combining the symbols they are made of. Heraldic coats of arms, for instance, must be composed in ways that forbid the juxtaposition of certain colors (Morin & Miton, 2018). Combinatorial structure and compositionality are evident in writing systems, mathematical or musical notations, formal logic (Zalta, 2022), and so on. As for context-independence and standardization, we encountered several examples of graphic codes exhibiting them. Nothing about graphic codes as such seems to prevent them from exhibiting all of McNeill's features. The problem, according to the specialization hypothesis, is that the only codes to achieve this are specialized. Enlarge the range of meanings that the code is to carry, and the system breaks down.

This suggests an obvious solution to the puzzle of ideography: The human mind cannot memorize large numbers of pairings between meanings and visual symbols. A self-sufficient code can be built on the basis of a few conventions, as long as it remains specialized and follows relatively strict rules of composition. But a more generalist code, to be usable without an oral gloss, would require users to learn an excessive number of conventions. If humans only have a large memory storing for codes and symbols when the symbols are made of phonemes, this would solve the puzzle of ideography.

5.4. Why the learning account fails

In Liberman's view, the speech faculty was specialized for processing and storing strings of phonemes. Phonemes, in his view, were quite distinct from sounds in general. His theory of the human phonetic capacity saw it essentially as a motor faculty rather than an auditory one: To represent phonemes is to represent gestures of the tongue and mouth. Because this definition excludes graphic shapes, it would explain why we can only learn a restricted number of graphic symbols.

The first issue with this account's solution to the puzzle of ideography is the fact that sign languages seem to be as easy to learn as spoken ones, even though their signs are not limited to mouth or tongue movements. A straightforward response would be to broaden the scope of Liberman's theory, so that speech includes signed speech as well as oral speech (a view considered in Lane, 1991). This move would make sense for a motor theory of language, but raises two new issues. First, it makes it harder to defend the view that speech perception is narrowly specialized. Such a view is sensible as long as speech perception is confined to the analysis of mouth and lips movements, but the perception of whole-body gestures blends into more general mechanisms of

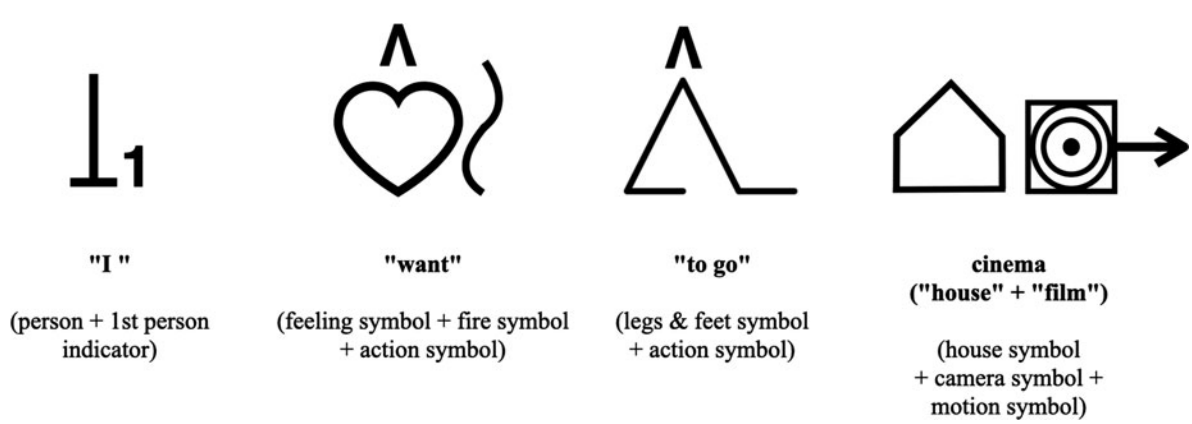

Figure 4. Phrase "I want to go to the movies" in Bliss symbolics. The Bliss symbols are reproduced by the author, copying an image from Wikimedia commons (Blissymbols, 2022).

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

action perception. The second issue is that graphic codes are gestural codes too. Graphic forms, generally, are traces of handwriting gestures. Even in the computer age, literate people learn their letters by inscribing them, affecting the way these graphic forms are represented. Even today, motor representations are involved in reading Latin-alphabet letters or Chinese characters (Schubert, Reilhac, & McCloskey, 2018; Yin & Zhang, 2021); and printed or computer-written fonts are modeled after handwritten symbols.

Thus, it seems that Liberman's motor theory of speech cannot have it both ways. If it posits a narrowly specific adaptation to process mouth and tongue gestures, it can explain why spoken language is easier to acquire compared to its graphic form, but it fails to explain why signed languages are easily and spontaneously acquired. Alternatively, it may assume that the speech faculty applies to gestures of the whole body, but in that case does not explain why traces of handwriting gestures would elude it.

5.5. The specialization constraint as a standardization problem

The learning account posits that the human mind is ill-equipped to memorize large numbers of pairings between meanings and visual symbols. This hypothesis is sufficient to derive the specialization hypothesis, but it also wrongly predicts that full-blown sign languages cannot evolve.

The standardization account focuses on the fact that any code used for communication is a *standard*: It serves its purpose only if a sufficient number of users share the same way of pairing symbols with meanings. Low standardization, I will argue, places the most serious limit on graphic codes' capacity to convey information. The *kupesi* symbols in use on the Tonga archipelago (Bell & Paegle, 2021) are ideographic symbols, often standing for clans, lineages, or mythical animals associated with them. Bell and Paegle's ethnographic work shows precisely how little shared meaning the symbols carry. Having sampled 15 *kupesi* from photographs of public spaces in Nuku'alofa, they show that none of these symbols could be named accurately by the majority of their interviewees, naming performance falling below 5% for 11 of the 15 symbols. This is what it means for a graphic code to be poorly standardized.

Building a shared standard raises a coordination problem (Lewis, 1969; Skyrms, 2010): The benefits of learning to communicate with a specific code depend on the number of others fluent in that code. This coordination problem is quite distinct from the issues that graphic codes pose for individual learning. The difficulties of standardization are surmountable for the codes used for face-to-face communication (like spoken or signed languages) because these are self-standardizing: Any occasion to use them is an occasion to learn to align with someone else's use. Mnemonic codes, being limited to private use, do not need to be standardized across several users. But to use a graphic code to communicate, and thus unlock the tremendous potential of asynchronous communication (Morin et al., 2020), a high degree of standardization is required.

6. The case for the standardization account

The standardization account implies that whether or not graphic expression develops into a full-blown code is a matter of forming and maintaining conventions between users. This view chimes in with recent claims that standardization is a key property that demarcates linguistic from nonlinguistic signs (Goldin-Meadow & Brentari, 2017). In emerging sign languages such as Al-Sayyid Bedouin sign language, the standardization of gestures into shared signs is a precondition for the emergence of phonological regularities (Sandler, 2009). More generally, standardization is an important point of demarcation between mere gestures or gesticulations, and full-blown signs (Goldin-Meadow & Brentari, 2017). I contend that difficulties linked to standardization are the reason why graphic codes remain underdeveloped compared to their spoken or signed counterparts.

6.1. Codes are standards, subject to lock-in dynamics

Many technologies benefit from the adoption of shared standards: Identical track gauges for railroads, compatible plug-in systems for electric appliances, shared coding languages for software design, and so on. The evolution and diffusion of technological standards is driven by several well-known effects: Positive feedback loops (successful standards tend to become even more successful), path-dependency, and lock-in dynamics (a small initial advantage solidifying into near-complete dominance) (Arthur, 1990; David, 1985). This last effect can lead to economic inefficiencies: Once in place, a suboptimal standard can persist indefinitely simply by virtue of being widespread. The most well-known example of such a lock-in dynamic is linked to writing: David (1985) and David and Rothwell (1996) argued that the costs of learning to type on a particular type of keyboard incentivized alignment on one standard ordering of letters (the QWERTY keyboard in many countries), to the detriment of other orderings that may have been more efficient. A similar but more consequential example of the same dynamic is the stifled development of the electric car in the twentieth century (Cowan & Hultén, 1996).

A very similar problem affects the codes that we use to communicate. Because codes are conventions, only users who have learnt the same code as other users can profit from them. If learning costs are reasonably high, this constraint can lead to a frequency-dependent advantage in favor of the codes that already have a high number of users, to the detriment of others – what Arthur calls a "positive feedback-loop" (Arthur, 1990), and Chrisomalis (2020) "networked frequency dependence." Cultural evolutionists talk of frequency-dependent cultural transmission when an agent's choice to copy a cultural trait is biased by the number of other agents having copied the trait (Boyd & Richerson, 1985). Networked frequency dependence is a special case of such dynamics, where network effects imply that it is advantageous for an agent to copy the most frequent behavior (Arthur, 1990, 2009; Chrisomalis, 2020; David, 1985; David & Rothwell, 1996).

This dynamic is evident in the case of language extinction (Zhang & Mace, 2021). The benefits of learning a language that has few speakers become less likely to outweigh its costs as the number of speakers declines, leading to an extinction spiral in which minority languages increasingly struggle to attract learners. The same type of frequency-dependent evolution asserts itself at the level of individual words: The distribution of synonym use for many meanings in English is best modeled by assuming that each individual speaker is disproportionately more likely to use the words most frequently used by others (Pagel, Beaumont, Meade, Verkerk, & Calude, 2019). Frequency-dependent advantages are not confined to human codes but are a general feature of communication signals throughout the animal world: Threat signals like warning coloration are better heeded by predators when they are common, increasing their bearer's fitness (Chouteau, Arias, & Joron, 2016). Standardization problems are, thus, not restricted to economics. They affect codes pervasively.

A code that is not standardized is useless as a communication system, although it may still be used privately as a memory prop. This issue was a roadblock in the evolution of basic information technologies. As Stephen Chrisomalis (2020) convincingly argues, standardization partly explains why it took so long for Roman numerals (I, II, III, IV, etc.) to be displaced by our current numeration system. Standardization can thus halt the displacement of a locked-in standard; it may also prevent the evolution of a useful one. This explains in part why the Romans failed to fully master tables, maps, or indexes (Riggsby, 2019). These simple and intuitive technologies seem fairly easy to invent, but such inventions are useless outside of a population of users who master the tool, having learnt to use it in the same way as others. Weights and measures are another case in point. Weights and measure are not standardized in all societies: Standards arise through commercial activity (Cooperrider & Gentner, 2019). Not all standards are equally likely to evolve: Ancient Romans possessed a refined system of *relative* weights and measures, based on fractions of an unspecified quantity: Halves, tenths, and so on. They did not manage to develop a consistent system of *absolute* weights and measures similar to the imperial or metric system. Such systems are harder to develop, because absolute units require standardization of measuring instruments to a much more precise degree than relative units (Riggsby, 2019, p. 86–88). Even for a large state with an advanced bureaucracy and a sophisticated literate culture, the challenges raised by standardization were serious enough to block the development of communication devices.

The deservedly famous example of the QWERTY keyboard highlights one key property of standardization dynamics – lock-in effects –, but arguably obscures an important fact about standardization: Standards do not always compete for the exact same niche. All English-language keyboards allow you to do the same thing: Type English words, more or less rapidly. But take video game hardware. A broad range of technologies exists to support video-gaming, from Arcade machines to personal computers (PCs) to Nintendo to virtual reality. Some of these technologies are directly competing standards offering basically the same functionalities (like Nintendo/Sony consoles); but there are things that can be done on a console that a PC will not allow, things which arcade or virtual reality hardware makes possible that a console cannot, and so on. When different standards do not compete for the exact same niche, lock-in dynamics should be less likely to arise. A competing standard can make up for an initial lack of popularity by offering services that dominant standards lack. The success of Nintendo consoles does not directly threaten the spread and development of virtual reality headsets, although it arguably holds it back. However, when a lock-in effect does occur, its consequences are more serious, because the suppressed standards do not simply offer more efficient ways of doing the same thing, but entirely new functionalities.

This, I hypothesize, is what often happened during the development of graphic codes: Spoken or signed codes, being easier to standardize, install a lock-in situation where other types of codes are less likely to evolve, even though evolving them would be beneficial, because they can do things other codes cannot do.

6.2. Cheap and transient signals are self-standardizing

All the codes that we use to communicate, including spoken and signed languages, face a standardization problem. This challenge is less daunting for languages because they are based on cheap, fast, and transient symbols. Spoken or signed messages require little effort or time to produce and vanish soon once they are emitted (Galantucci, Kroos, & Rhodes, 2010; Hockett, 1960). Being cheap and fast, they form messages that can be modified or repaired multiple times, allowing interlocutors to converge on shared meanings (Fusaroli, Rączaszek-Leonardi, & Tylén, 2014; Pickering & Garrod, 2004). Being transient (language's "rapidity of fading"; Hockett, 1960), they constrain interlocutors to face-to-face interactions, where the advantages of common ground are maximized. This leads to standardization not only at the level of the pair, but also at the level of entire populations, because turnover in conversation partners leads to convergence to broader standards that everyone can share (Guilbeault, Baronchelli, & Centola, 2021).

Transient messages have one obvious drawback: They cannot travel far in time or space. Transient symbols can only be used by interlocutors who share the same time- and spaceframe in a face-to-face setting. (I am not considering here the changes brought about by electronic recording or transmission technologies – changes that are very recent in the long-term cultural evolution of our species.) In other words, they exclude asynchronous communication. But the drawback is also an advantage: Face-to-face communication maximizes common ground, the body of information that interlocutors share by virtue of being together (Clark, 1996; Sperber & Wilson, 1995). If two people are together in a room, this includes the environment that they are both aware of, information on the identity of other interlocutors and, crucially, the knowledge that the other interlocutors know some of what I know, know that I know it, and so on. Face-to-face communication makes a lot of common ground information available without the need to infer it or to encode it explicitly (Pickering & Garrod, 2004). It also provides interlocutors with opportunities to enrich and update this common ground in real time, because any signal they exchange becomes part of this common ground (Clark, 1996).

In asynchronous communication, interlocutors also have access to common ground information, but it must be either encoded or inferred. Take the identity of a message's author. In asynchronous communication, it can be inferred from various cues (e.g., the handwriting on a note, the fact that it's pasted on the fridge door, etc.) or it can be explicitly encoded through a signature, a seal, or some other identifying mark. Both inference and explicit encoding are fallible and costly processes. They may reach the wrong conclusion and require effort in any case.

Access to common ground provides interlocutors with the means to solidify the standards and conventions of spoken language (Clark & Wilkes-Gibbs, 1986; Keysar, Barr, Balin, & Paek, 1998). Of course, interlocutors in most conversations can avail themselves of a shared language; but these conventions often need to be refined and tailored to immediate needs: Shorter, more precise ways to refer to things that matter to the conversation. A USB-A to HDMI adapter therefore becomes "the dongle" to the people frantically looking for it; the strange newcomer who just entered the bar becomes "blue hair" to the regulars; and so on. In Clark and Wilkes-Gibbs' classic experiment, ambiguous Tangram shapes get baptized in this way, long and variable descriptions swiftly morphing into short and shared conventional labels. Common ground is crucial to standardization process for two reasons. First, conventional meanings can be anchored to immediately perceptible referents: Alice points at "the thingy" on the table, and Bruno immediately knows that "the thingy" stands for his keychain. Second, the past history of a face-to-face conversation is part of the common ground that interlocutors share (Clark & Brennan, 1991). When Alice baptized "the thingy," what she did was manifest to both Alice and Bruno, and they carry this shared memory into future stages of the conversation. These cues that a message has been attended to are much more difficult to get from graphic communication.

Face-to-face communication can take full advantage of this common ground because it is based on cheap-to-produce signals. Any spoken (or signed) message can be modified or repaired at a little cost, until interlocutors align on a shared understanding (Clark & Brennan, 1991; Clark & Wilkes-Gibbs, 1986; Enfield, 2017). Cheap signals allow for a large amount of information to be exchanged, whereas face-to-face interaction prevents the flow of information from being one-sided, allowing for repair and quick turn-taking (Fusaroli et al., 2014; Levinson, 2006). High-bandwidth, two-sided communication prevents misunderstandings and ensures standardization through mechanisms such as repair (signaling misunderstandings by linguistic means); backchannel communication (subtle signals like grunts, nods, eyebrow flashes, etc.), which can function among other things to signal that repair is not needed (Schegloff, 1982); and interactive alignment (repetition or imitation of speech at various levels) (Dideriksen, Christiansen, Tylén, Dingemanse, & Fusaroli, 2023; Garrod & Pickering, 2004; Pickering & Garrod, 2004). None of those mechanisms work as efficiently in asynchronous communication as they do in face-to-face interaction – and often do not work at all. Asynchronous turn-taking (which can take place in letters exchange, for instance) is slow by definition, because interlocutors do not inhabit the same timeframe. In other cases it is simply absent: One cannot put an objection to Plato's arguments and get a reply. Compounding this problem, asynchronous repair signals need to be explicitly encoded: There is no nonverbal or paraverbal channel through which to convey them, whereas conversation has eyebrow flashes, humming, and so on (Bavelas & Gerwing, 2011).

Access to common ground information and its exploitation through repair, backchannel communication, and interactive alignment reinforces the standardization of the codes used for communication. Spoken and signed languages are "self-standardizing" because they are based on cheap and transient symbols. Transience forces interlocutors to use symbols in face-to-face settings, where common ground is rich and repair is quick. Cheap and fast production allows interlocutors to take full advantage of the possibilities offered by repair. This is how spoken and signed languages solve the challenge of standardization.

6.3. Languages have a lock-in effect on the evolution of codes

The previous section explained why languages, adapted for face-to-face communication, should evolve more readily than other codes. This section explains why, conversely, the codes fit for *asynchronous* communication, like writing systems and other graphic codes, evolve more rarely. Because they lack the self-standardizing property of languages, they are less likely to appear in the first place. And once a spoken language is in place, it fulfills most of the functions that a graphic code is useful for, making graphic codes largely redundant. As an unfortunate consequence, asynchronous communication, the one function that graphic codes are uniquely fit for, remains underdeveloped (Morin et al., 2020).

This argument assumes that, once a code exists, it inhibits the development of other codes that fulfill similar but partially different functions. This point can be made with an analogy. Most languages are based on spoken words, not gestured signs. Sign languages, however, permit some forms of communication that are not possible with spoken languages, communication with the deaf, communication in noisy environments, and so on. Yet, for most people, they remain a latent possibility, because spoken language is already the default tool for communication, and the unique benefits of signed language are not sufficient to offset the costs of developing and learning a new code. The inhibiting effects that the availability of spoken language exert over the development of signed language is suggested by a series of studies by Goldin-Meadow and collaborators (Goldin-Meadow & Brentari, 2017; Goldin-Meadow, McNeill, & Singleton, 1996; see also McNeill, 1992), who asked hearing participants to describe a scene using only gestures; compared to the cospeech gestures produced by the same participants while speaking, these "silent gestures" exhibit more discrete segmentation and a greater degree of hierarchical combination, two conditions that favor the emergence of a coherent code.

A lock-in occurs when a standard is so widespread that it hinders the rise of a different standard that would better accomplish the same task (like alternatives to the QWERTY keyboard, or the electric car), or that would fulfill different but overlapping functions (like virtual reality headsets compared to gaming consoles). The first type of lock-in can protect a deficient technology. The second type can inhibit the evolution of technologies that offer different but overlapping functionalities. Graphic codes face a lock-in of the second type. The availability of spoken or signed languages means that the benefits of developing graphic codes (including writing) do not, for most purposes, warrant the costs, because most of the things we could accomplish with them can also be performed with language.

In spite of language's inhibiting effects, graphic codes do emerge, but only for a few niche functionalities, and even then, they grow in the shadow of language. The availability of memorized oral messages means that the graphic code does not need to carry as much information as a self-standing code would; the availability of an oral gloss removes the need for spectators to understand the graphic code; misunderstandings can be repaired orally without changing or correcting the graphically encoded message. In all these ways, an oral crutch prevents graphic codes from learning to walk.

6.4. The future of ideography

This paper has so far deliberately neglected recent inventions that are too young to have left their mark on the long-term evolution of language or writing: Electronic communication, telephones, voicemail, and so on. But those have clearly started to transform human communication, graphic or oral. If the standardization account is on the right track, it opens a window for the evolution of ideographic communication in the digital age. Digital communication might overthrow the constraints that weigh on the evolution of graphic codes. Thanks to texting, graphic signals are becoming almost as fast and effortless to send as spoken words or gestures; the amount of information that participants in a digital interaction can have in common has exploded, to the point that it can rival the common ground shared by face-to-face interlocutors. The standardization account should thus predict that emojis, gifs, and other digital pictographs should become increasingly endowed with precise and shared meanings. There has been an explosion in emoji and emoticon use, concomitant with the rise of digital communication, but whether these are replacing writing or complementing it is a matter of debate. The prevalent view does not see emojis as genuine alternatives to verbal communication. Instead, they are viewed as filling some of the functions that paraverbal signals, such as intonation, facial expression, and hand gestures, would fulfill in face-to-face conversation (Derks, Bos, & von Grumbkow, 2008; Gawne & McCulloch, 2019; Vandergriff, 2013).

Why are emojis not yet ready to replace writing? As the standardization account would predict, a lack of agreement over the symbols' meanings is partly to blame. Survey studies show substantial disagreement over the meaning of the Unicode-encoded emojis (Częstochowska et al., 2022; Miller et al., 2016; Tigwell & Flatla, 2016). Agreement is weak even for very frequent symbols like 😬, 😭, or the seemingly obvious 🙂 (whose interpretation ranges from intense positivity to awkwardness). Even on basic dimensions such as valence (whether an emoji is associated with positive, negative, or neutral emotions), participants' ratings disagree 25% of the time (Miller et al., 2016; Tigwell & Flatla, 2016, find qualitatively similar results for valence and arousal). This limits the symbols' expressive power and forces users to rely on context.

The standardization account would, however, also predict that agreement over the meaning of emojis should grow over time, insofar as written digital communication approaches the conditions of synchronous face-to-face interaction. It is unclear whether digital communication can become cheap and fast enough for this; the pace of repair and turn-taking in normal conversation is so rapid (Stivers et al., 2009) that even sophisticated videoconferencing tools cannot always avoid disrupting it. Still there seems to be, for the first time ever, a possibility that digital communication breaks the chains that keep ideographic communication bound. Should this happen, the evolution of the online ideographic language should start with a consolidation of emojis that serve the restricted function of encoding paraverbal cues, such as the facial expressions of emotions. Their ambiguity should decrease over time, in keeping with experimental results showing increasing specificity for online signals in referential communication (Guilbeault et al., 2021; Morin, Müller, Morisseau, & Winters, 2022). Once this stage is passed the code could grow to include an expanding range functions, gradually becoming more generalist. This is based on the fact that complex graphic codes tend to grow on top of simpler ones, as we saw in the case of writing systems (pre-existed by numeration systems), or musical notations that grew increasingly generalist (Croft, 2017).

7. Conclusion

Everyone knows writing has a special relationship with language, but that relation has often been minimized or trivialized in the scholarly literature. Minimized by showing that writing encodes much more than sounds, by casting doubt on the importance of phonemic or syllabic notations in writing systems like Chinese or Egyptian, by giving credence to the notion that some writing systems are ideographic. Trivialized too: If writing mainly encodes language, is that not merely because we chose to define it in this way? Do we not possess many graphic notations that do not encode language, but are just as powerful as writing?

Against these tendencies, this paper sought to show that the glottographic nature of writing is neither a falsity nor a platitude, but a puzzle. The puzzle is the absolute rarity, in the current or past record, of a fully ideographic code that can be used autonomously, not just as a mnemonic prop, to encode information on a broad range of topics. I have attempted to make the case for this claim – the specialization hypothesis – reviewing two possible answers to the puzzle that it raises, as well as some implications for the evolution of writing. I argued against one plausible explanation of the puzzle. Ideographies, if they existed, would not necessarily be overly difficult to learn: Sign languages demonstrate that visual codes can be just as easy to acquire as spoken ones. Ideographies are not hard to learn; they are hard to standardize. One can build an ideographic code and learn how to use it, but getting a sufficient number of people to go along with it is the real challenge. This problem is specific to graphic codes. It does not apply to the same degree to spoken or signed codes. Languages, spoken or signed, have been used in an exclusively synchronous fashion until very recently, and face-to-face interaction makes it easier for interlocutors to resolve any salient discrepancy between my code and your code. Words and gestures are quick and effortless to produce (compared to graphic symbols), making it easier to change codes and converge on shared symbols. Spoken or signed codes get a first-mover advantage from this. They are likely to be in place before graphic codes can evolve.

This answer to the puzzle requires testing, and the puzzle itself is an empirical claim. The nonexistence of generalist and self-sufficient ideographies is not a simple consequence of how we choose to define the word "writing." The graphic codes in use before the rise of the glottographic principle, and during the long periods where the principle was known but scarcely used, should lack either generality or self-sufficiency, or both. This claim is exposed to empirical refutation – I contend it has not occurred yet.

In the history of writing, the nonevolution of ideography is the dog that did not bark: Its absence tells us much about the nature of graphic codes, their power, and their limitations. In evolutionary terms, a complete ideography could be seen as a peak in the design space of graphic codes (Acerbi, Tennie, & Mesoudi, 2016; Dennett, 1995; Mesoudi & Thornton, 2018). A design space (modeled on Sewall-Wright's fitness landscapes) is a representation of the quality of different solutions to a given problem, plotted against their similarity. A set of good solutions is represented as a peak in the design space, whereas a set of bad solutions will appear as a valley or a plain. Some peaks are easily accessible through a gradual evolutionary progress, because they are surrounded by similar solutions that are also relatively good. But some peaks are located in an area of the design space that can hardly be reached, because there is no smooth evolutionary path leading to them. In this view, writing is an isolated peak in the design landscape of graphic codes. The availability of spoken language kept most of human communication away from it. Future work could uncover interesting parallels with other domains of technology where the constraints of standardization kept good ideas from being realized.

Acknowledgments. I wish to thank Nicolas Baumard, Pascal Boyer, Roberto Casati, Paul Égré, Silvia Ferrara, David Gil, Yoolim Kim, Helena Miton, Thom Scott-Phillips, Oleg Sobchuk, Kristian Tylén, Mason Youngblood, and Elia Weber, for their helpful comments on an earlier draft. I thank the BBS reviewers, including Nick Enfield, for their input.

Financial support. The work leading to this paper has received funding from the "Frontiers in Cognition" EUR grant, ANR-17-EURE-0017 EUR and the Max Planck Society.

Competing interest. None.

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

Notes

1. This is clearly shown by the sales figures for translated vs. untranslated manga outside Japan.
2. Programming languages such as C+ or Java will not be considered here either, because this paper only concerns itself with means of communication between humans.
3. I am indebted to Nick Enfield for this observation.

References

Acerbi, A., Tennie, C., & Mesoudi, A. (2016). Social learning solves the problem of narrow-peaked search landscapes: Experimental evidence in humans. *Royal Society Open Science*, *3*(9), 160215. https://doi.org/10.1098/rsos.160215

Arthur, W. B. (1990). Positive feedbacks in the economy. *Scientific American*, *262*(2), 92–99.

Arthur, W. B. (2009). *The nature of technology: What it is and how it evolves*. Simon & Schuster.

Baines, J. (2007). *Visual and written culture in ancient Egypt*. Oxford University Press.

Banham, D. (2015). The old English monastic sign language. In J. Bakken Jepsen, S. Lutalo-Kingii, G. De Clerck, & W. McGregor (Eds.), *Sign languages of the world a comparative handbook* (pp. 901–910). De Gruyter Mouton.

Bartlett, F. (1932). *Remembering: A study in experimental and social psychology*. Cambridge University Press.

Battestini, S. (2006). L'écriture nsibidi de la Cross River Region (Nigeria). In S. Battestini (Ed.), *De l'écrit africain à l'oral: Le phénomène graphique africain* (pp. 245–262). L'Harmattan.

Bauer, R. S. (2018). Cantonese as written language in Hong Kong. *Global Chinese*, *4*(1), 103–142. https://doi.org/10.1515/glochi-2018-0006

Bavelas, J. B., & Gerwing, J. (2011). The listener as addressee in face-to-face dialogue. *International Journal of Listening*, *25*(3), 178–198. https://doi.org/10.1080/10904018.2010.508675

Bell, A. V., & Paegle, A. (2021). Ethnic markers and how to find them: An ethnographic investigation of marker presence, recognition, and social information. *Human Nature*, *32*(2), 470–481. https://doi.org/10.1007/s12110-021-09401-z

Blissymbols. (2022). *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Blissymbols&oldid=1086221021

Boltz, W. G. (1993). *The origin and early development of the Chinese writing system*. American Oriental Society.

Boone, E. H. (1994). Aztec pictorial histories: Records without words. In E. H. Boone & W. Mignolo (Eds.), *Writing without words: Alternative literacies in Mesoamerica and the Andes* (pp. 50–76). Duke University Press.

Boone, E. H., & Mignolo, W. (1994). *Writing without words: Alternative literacies in Mesoamerica and the Andes*. Duke University Press.

Boyd, R., & Richerson, P. (1985). *Culture and the evolutionary process*. University of Chicago Press.

Aristotle. (1962). *On interpretation* (T. de V. Cajetan & J. T. Oesterle, Eds.). Marquette University Press.

Caldwell, C. A., & Smith, K. (2012). Cultural evolution and perpetuation of arbitrary communicative conventions in experimental microsocieties. *PLoS ONE*, *7*(8), e43807. https://doi.org/10.1371/journal.pone.0043807

Casati, R. (2016). The structure of standard music notation. In L. Zaibert (Ed.), *The theory and practice of ontology* (pp. 187–201). Palgrave Macmillan. Retrieved from https://doi.org/10.1057/978-1-137-55278-5_10

Changizi, M., Zhang, Q., Ye, H., & Shimojo, S. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *The American Naturalist*, *167*(5), E117–E139. https://doi.org/10.1086/502806

Chen, P., & Ping, C. (1999). *Modern Chinese: History and sociolinguistics*. Cambridge University Press.

Chiang, T. (2016). *Stories of your life and others*. Vintage Books.

Chouteau, M., Arias, M., & Joron, M. (2016). Warning signals are under positive frequency-dependent selection in nature. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(8), 2164–2169. https://doi.org/10.1073/pnas.1519216113

Chrisomalis, S. (2020). *Reckonings: Numerals, cognition, and history*. MIT Press.

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). American Psychological Association.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*(1), 1–39. https://doi.org/10.1016/0010-0277(86)90010-7

Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. A&C Black.

Cooperrider, K., & Gentner, D. (2019). The career of measurement. *Cognition*, *191*, 103942. https://doi.org/10.1016/j.cognition.2019.04.011

Coulmas, F. (1996a). *Blackwell encyclopedia of writing systems*. Wiley-Blackwell.

Coulmas, F. (1996b). Typology of writing systems. In H. Gunther & O. Ludwig (Eds.), *Schrift und Schriftlichkeit* (pp. 1380–1387). De Gruyter.

Coulmas, F. (2003). *Writing systems: An introduction to their linguistic analysis*. Cambridge University Press. Retrieved from https://doi.org/10.1017/CBO9781139164597

Cowan, R., & Hultén, S. (1996). Escaping lock-in: The case of the electric vehicle. *Technological Forecasting and Social Change*, *53*(1), 61–79. https://doi.org/10.1016/0040-1625(96)00059-5

Croft, W. (2017). Evolutionary complexity of social cognition, semasiographic systems, and language. In C. Coupé, F. Pellegrino, & S. S. Mufwene (Eds.), *Complexity in language: Developmental and evolutionary perspectives* (pp. 101–134). Cambridge University Press. Retrieved from https://doi.org/10.1017/9781107294264.005

Częstochowska, J., Gligorić, K., Peyrard, M., Mentha, Y., Bień, M., Grütter, A., … West, R. (2022). On the context-free ambiguity of emoji. *Proceedings of the International AAAI Conference on Web and Social Media*, *16*, 1388–1392.

Damerow, P. (2006). The origins of writing as a problem of historical epistemology. *Cuneiform Digital Library Journal*, *1*, 1–10.

Daniels, P. T., & Bright, W. (1996). *The world's writing systems*. Oxford University Press.

David, P., & Rothwell, G. (1996). Standardization, diversity and learning: Strategies for the coevolution of technology and industrial capacity. *International Journal of Industrial Organization*, *14*(2), 181–201.

David, P. A. (1985). Clio and the economics of QWERTY. *The American Economic Review*, *75*(2), 332–337.

Davis, J. (2015). Plain Indians sign language. In J. Bakken Jepsen, G. De Clerck, S. Lutalo-Kingii, & W. McGregor (Eds.), *Sign languages of the world: A comparative handbook* (pp. 911–931). De Gruyter Mouton.

Dayrell, E. (1911). Further notes on "Nsibidi signs with their meanings from the Ikom district, southern Nigeria." *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, *41*, 521–540. https://doi.org/10.2307/2843186

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. University of Hawai'i Press.

DeFrancis, J., & Unger, J. M. (2009). Rejoinder to Geoffrey Sampson, "Chinese script and the diversity of writing systems." *Linguistics*, *32*(3), 549–554. https://doi.org/10.1515/ling.1994.32.3.549

Dehaene, S. (2010). *Reading in the brain: The new science of how we read*. Viking.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*(2), 384–398. https://doi.org/10.1016/j.neuron.2007.10.004

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meaning of life*. Simon & Schuster.

Derks, D., Bos, A. E. R., & von Grumbkow, J. (2008). Emoticons and online message interpretation. *Social Science Computer Review*, *26*(3), 379–388. https://doi.org/10.1177/0894439307311611

de Saussure, F. (2011). *Course in general linguistics*. Columbia University Press.

Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2023). Quantifying the interplay of conversational devices in building mutual understanding. *Journal of Experimental Psychology: General*, *152*(3), 864–889. https://doi.org/10.1037/xge0001301

du Ponceau, P. S. (1838). *A dissertation on the nature and character of the Chinese system of writing. To which are subjoined a vocabulary of the Cochin Chinese language by J. Morrone [&c.]*. (J. M. Morrone, Ed.). American Philosophical Society.

Enfield, N. J. (2017). *How we talk: The inner workings of conversation*. Basic Books.

Friederici, A. D. (2020). Hierarchy processing in human neurobiology: How specific is it? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1789), 20180391. https://doi.org/10.1098/rstb.2018.0391

Fusaroli, R., Rączaszek-Leonardi, J., & Tylén, K. (2014). Dialog as interpersonal synergy. *New Ideas in Psychology*, *32*, 147–157. https://doi.org/10.1016/j.newideapsych.2013.03.005

Galantucci, B., Kroos, C., & Rhodes, T. (2010). The effects of rapidity of fading on communication systems. *Interaction Studies*, *11*(1), 100–111.

Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, *31*(6), 961–987. https://doi.org/10.1080/03640210701703659

Garrod, S., & Pickering, M. J. (2004). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11. https://doi.org/10.1016/j.tics.2003.10.016

Gawne, L., & McCulloch, G. (2019). Emoji as digital gestures. *Language@Internet*, *17*(2), 1–10. https://www.languageatinternet.org/articles/2019/gawne

Gelb, I. J. (1963). *A study of writing*. University of Chicago Press.

Goldin-Meadow, S., & Brentari, D. (2017). Gesture, sign, and language: The coming of age of sign language and gesture studies. *Behavioral and Brain Sciences*, *40*, E46. https://doi.org/10.1017/S0140525X15001247

Goldin-Meadow, S., McNeill, D., & Singleton, J. (1996). Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. *Psychological Review*, *103*(1), 34–55. https://doi.org/10.1037/0033-295x.103.1.34

Goody, J. (1977). *The domestication of the savage mind*. Cambridge University Press.

Granito, C., Tehrani, J., Kendal, J., & Scott-Phillips, T. (2019). Style of pictorial representation is shaped by intergroup contact. *Evolutionary Human Sciences*, *1*, E8. https://doi.org/10.1017/ehs.2019.8

Green, J. (2007). *Drawn from the ground: Sound, sign and inscription in central Australian sand stories*. Cambridge University Press.

Griaule, M., & Dieterlen, G. (1951). *Signes graphiques soudanais*. Hermann.

Guilbeault, D., Baronchelli, A., & Centola, D. (2021). Experimental evidence for scale-induced category convergence across populations. *Nature Communications*, *12* (1), 327. https://doi.org/10.1038/s41467-020-20037-y

Handel, Z. (2015). Logography and the classification of writing systems: A response to Unger. *Scripta*, *7*, 109–150.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*(3), 662–720. https://doi.org/10.1037/0033-295X.111.3.662

Heintz, C., & Scott-Phillips, T. (2022). Expression unleashed: The evolutionary & cognitive foundations of human communication: Version December 2021. *Behavioral and Brain Sciences*, *46*, E1. https://doi.org/10.1017/S0140525X22000012

Hill, A A. (1967). The typology of writing systems. In W. M. Austin (Ed.), *Papers in linguistics in honor of Léon Dostert* (pp. 92–99). Mouton.

Hirshorn, E. A., & Harris, L. N. (2022). Culture is not destiny, for reading: Highlighting variable routes to literacy within writing systems. *Annals of the New York Academy of Sciences*, *1513*(1), 31–47. https://doi.org/10.1111/nyas.14768

Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*, 4–12.

Houston, S. D. (2004). The archaeology of communication technologies. *Annual Review of Anthropology*, *33*, 223–250.

Hyman, M. D. (2006). Of glyphs and glottography. *Language & Communication*, *26*(3), 231–249. https://doi.org/10.1016/j.langcom.2006.03.001

Jakobson, R. (1964). On visual and auditory signs. *Phonetica*, *11*(3–4), 216–220. https://doi.org/10.1159/000258396

Jee, H., Tamariz, M., & Shillcock, R. (2022). Systematicity in language and the fast and slow creation of writing systems: Understanding two types of non-arbitrary relations between orthographic characters and their canonical pronunciation. *Cognition*, *226*, 105197. https://doi.org/10.1016/j.cognition.2022.105197

Kelly, P. (2019). Australian message sticks: Old questions, new directions. *Journal of Material Culture*, *25*(2), 133–152. https://doi.org/10.1177/1359183519858375

Kelly, P., Winters, J., Miton, H., & Morin, O. (2021). The predictable evolution of letter shapes: An emergent script of West Africa recapitulates historical change in writing systems. *Current Anthropology*, *62*(6), 669–691. https://doi.org/10.1086/717779

Keysar, B., Barr, D. J., Balin, J. A., & Paek, T. S. (1998). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language*, *39*(1), 1–20. https://doi.org/10.1006/jmla.1998.2563

Kirby, S., Tamariz, M., Cornish, H., & Smith, K. (2015). Compression and communication in the cultural evolution of linguistic structure. *Cognition*, *141*, 87–102. https://doi.org/10.1016/j.cognition.2015.03.016

Knowlson, J. R. (1965). The idea of gesture as a universal language in the XVIIth and XVIIIth centuries. *Journal of the History of Ideas*, *26*(4), 495–508. https://doi.org/10.2307/2708496

Lane, H. (1991). *Comment: Dr. Harlan and Mr. Lane*. In *Modularity and the motor theory of speech perception: Proceedings of a conference to Honor Alvin M. Liberman* (pp. 171–174). Psychology Press.

Lee, K., & Karmiloff-Smith, A. (1996). The development of cognitive constraints on notations. *Archives de Psychologie*, *64*(248), 3–26.

Levinson, S. C. (2006). On the human "interaction engine." In S. C. Levinson & N. J. Enfield (Eds.), *Roots of human sociality* (pp. 39–69). Routledge.

Lewis, D. (1969). *Convention: A philosophical study*. Wiley-Blackwell.

Li, D. C. S. (2006). Chinese as lingua franca in greater China. *Annual Review of Applied Linguistics*, *26*, 149–176. https://doi.org/10.1017/S0267190506000080

Li, Y., Peng, D., Liu, L., Booth, J. R., & Ding, G. (2014). Brain activation during phonological and semantic processing of Chinese characters in deaf signers. *Frontiers in Human Neuroscience*, *8*, 1–11. https://www.frontiersin.org/article/10.3389/fnhum.2014.00211

Liberman, A. M. (1992). The relation of speech to reading and writing. In R. Frost & L. Katz (Eds.), *Advances in psychology* (Vol. 94, pp. 167–178). North-Holland. Retrieved from https://doi.org/10.1016/S0166-4115(08)62794-6

Lin, T.-J., & Biggs, M. (2006). *A preliminary study of learnable pictogram languages*. IADE. Retrieved from http://uhra.herts.ac.uk/handle/2299/7435

Liu, X., Vermeylen, L., Wisniewski, D., & Brysbaert, M. (2020). The contribution of phonological information to visual word recognition: Evidence from Chinese phonetic radicals. *Cortex*, *133*, 48–64. https://doi.org/10.1016/j.cortex.2020.09.010

Lock, A., & Gers, M. (2012). The cultural evolution of written language and its effects a Darwinian process from prehistory to the modern day. In E. Grigorenko, E. Mambrino, & D. Preiss (Eds.), *Writing: A mosaic of new perspectives* (pp. 11–36). Psychology Press.

Lurie, D B. (2006). Language, writing, and disciplinarity in the Critique of the "Ideographic Myth": Some proleptical remarks. *Language & Communication*, *26*(3), 250–269. https://doi.org/10.1016/j.langcom.2006.02.015

Macgregor, J. K. (1909). Some notes on Nsibidi. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, *39*, 209–219. https://doi.org/10.2307/2843292

Mallery, G. (1879). *Sign language among North American Indians compared with that among other peoples and deaf-mutes*. Walter de Gruyter.

Mallery, G., & Ethnology, S. I. B. of A. (1886). Pictographs of the North American Indians: A preliminary paper. Government Printing Office.

Matsunaga, S. (1996). The linguistic nature of kanji reexamined: Do kanji represent only meanings? *The Journal of the Association of Teachers of Japanese*, *30*(2), 1–22. https://doi.org/10.2307/489563

Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by hand and by ear: The relationship between speech and reading* (pp. 133–148). MIT Press.

McBride-Chang, C., Cho, J.-R., Liu, H., Wagner, R. K., Shu, H., Zhou, A., … Muse, A. (2005). Changing models across cultures: Associations of phonological awareness and morphological structure awareness with vocabulary and word recognition in second graders from Beijing, Hong Kong, Korea, and the United States. *Journal of Experimental Child Psychology*, *92*(2), 140–160. https://doi.org/10.1016/j.jecp.2005.03.009

McBride-Chang, C., Tong, X., Shu, H., Wong, A. M.-Y., Leung, K., & Tardif, T. (2008). Syllable, phoneme, and tone: Psycholinguistic units in early Chinese and English word recognition. *Scientific Studies of Reading*, *12*(2), 171–194. https://doi.org/10.1080/10888430801917290

McCawley, J. D. (1992). Linguistic aspects of musical and mathematical notation. In P. Downing, S. D. Lima, & M. Noonan (Eds.), *The linguistics of literacy* (pp. 169–190). John Benjamins.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Meissner, M., Philpott, S. B., & Philpott, D. (1975). The sign language of sawmill workers in British Columbia. *Sign Language Studies*, *9*, 291–308.

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B: Biological Sciences*, *285*(1880), 20180712. https://doi.org/10.1098/rspb.2018.0712

Miller, H., Thebault-Spieker, J., Chang, S., Johnson, I., Terveen, L., & Hecht, B. (2016). "Blissfully happy" or "ready to fight." *Proceedings of the 10th international conference on web and social media, ICWSM 2016*, pp. 259–268.

Mohr, S. (2015). Tshaukak'ui – hunting signs of the Ts'ixa in northern Botswana. In J. Bakken Jepsen, G. De Clerck, S. Lutalo-Kingii, & W. McGregor (Eds.), *Sign languages of the world* (pp. 933–953). De Gruyter Mouton.

Morin, O. (2018). Spontaneous emergence of legibility in writing systems: The case of orientation anisotropy. *Cognitive Science*, *42*(3), 664–677. https://doi.org/10.1111/cogs.12550

Morin, O., Kelly, P., & Winters, J. (2020). Writing, graphic codes, and asynchronous communication. *Topics in Cognitive Science*, *12*, 727–743. https://doi.org/10.1111/tops.12386

Morin, O., & Miton, H. (2018). Detecting wholesale copying in cultural evolution. *Evolution & Human Behavior*, *39*(4), 392–401. https://doi.org/10.1016/j.evolhumbehav.2018.03.004

Morin, O., Müller, T. F., Morisseau, T., & Winters, J. (2022). Cultural evolution of precise and agreed-upon semantic conventions in a multiplayer gaming app. *Cognitive Science*, *46*(2), e13113.

Morris, I. (2014). *The measure of civilization: How social development decides the fate of nations* (new edition). Princeton University Press.

Okrent, A. (2010). *In the land of invented languages: Adventures in linguistic creativity, madness, and genius*. Random House.

Olson, D. (1991). Literacy as metalinguistic activity. In D. Olson & N. Torrance (Eds.), *Literacy and orality* (pp. 251–269). Cambridge University Press.

Pagel, M., Beaumont, M., Meade, A., Verkerk, A., & Calude, A. (2019). Dominant words rise to the top by positive frequency-dependent selection. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(15), 7397–7402. https://doi.org/10.1073/pnas.1816994116

Panofsky, E. (1939). *Studies in iconology: Humanistic themes in the art of the renaissance*. Oxford University Press.

Pastoureau, M. (2007). *Traité d'héraldique* (5e edn.). Editions A & J Picard.

Perfetti, C. A., & Harris, L. N. (2013). Universal reading processes are modulated by language and writing system. *Language Learning and Development*, *9*(4), 296–316. https://doi.org/10.1080/15475441.2013.813828

Petitto, L. A., Langdon, C., Stone, A., Andriola, D., Kartheiser, G., & Cochran, C. (2016). Visual sign phonology: Insights into human reading and language from a natural soundless phonology. *Wiley Interdisciplinary Reviews. Cognitive Science*, *7*(6), 366–381. https://doi.org/10.1002/wcs.1404

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*(2), 169–190. https://doi.org/10.1017/S0140525X04000056

Pickering, M. J., & Garrod, S. (2021). *Understanding dialogue: Language use and social interaction*. Cambridge University Press. Retrieved from https://doi.org/10.1017/9781108610728

Pope, M. (1999). *The story of decipherment: From Egyptian hieroglyphs to Maya script*. Thames and Hudson.
Quay, S. (2015). Monastic sign language from medieval to modern times. In J. Bakken Jepsen, S. Lutalo-Kingii, G. De Clerck, & W. McGregor (Eds.), *Sign languages of the world a comparative handbook* (pp. 871–900). De Gruyter Mouton.
Ramus, F. (2004). The neural basis of reading acquisition. In S. Gazzania (Ed.), *The cognitive neurosciences* (3rd ed., pp. 815–824). The MIT Press.
Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. In B. MacWhinney, & W. O'Grady (Eds.), *The handbook of language emergence* (pp. 237–263). John Wiley. Retrieved from https://doi.org/10.1002/9781118346136.ch11
Riggsby, A. (2019). *Mosaics of knowledge: Representing information in the Roman world*. Oxford University Press.
Rossi, P. (2000). *Logic and the art of memory: The quest for a universal language*. University of Chicago Press.
Sampson, G. (1985). *Writing systems: A linguistic Introduction*. Stanford University Press.
Sampson, G. (2017). *The linguistics delusion*. Equinox.
Sandler, W. (2009). What comes first in language emergence? In N. J. Enfield (Ed.), *Dependencies in language* (pp. 63–84). Language Science Press.
Schegloff, E. A. (1982). Discourse as an interactional achievement: Some uses of "uh huh" and other things that come between sentences. In *Analyzing discourse: Text and talk* (pp. 71–93). Georgetown University Press. Retrieved from https://repository.library.georgetown.edu/bitstream/handle/10822/555474/GURT_1981.pdf
Schlenker, P. (2018). What is super semantics? *Philosophical Perspectives*, *32*(1), 365–453. https://doi.org/10.1111/phpe.12122
Schmandt-Besserat, D. (2007). *When writing met art: From symbol to story* (1st edn.). University of Texas Press.
Schubert, T., Reilhac, C., & McCloskey, M. (2018). Knowledge about writing influences reading: Dynamic visual information about letter production facilitates letter identification. *Cortex*, *103*, 302–315. https://doi.org/10.1016/j.cortex.2018.03.020
Scott-Phillips, T. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Palgrave Macmillan.
Scribner, S., & Cole, M. (1981). *The psychology of literacy*. Harvard University Press.
Sevcik, R., Romski, M. A., & Wilkinson, K. (1991). Roles of graphic symbols in the language acquisition process for persons with severe cognitive disabilities. *Augmentative and Alternative Communication*, *7*(3), 161–170. https://doi.org/10.1080/07434619112331275873
Sevcik, R. A., Barton-Hulsey, A., Romski, M., & Hyatt Fonseca, A. (2018). Visual-graphic symbol acquisition in school age children with developmental and language delays. *Augmentative and Alternative Communication*, *34*(4), 265–275. https://doi.org/10.1080/07434618.2018.1522547
Severi, C. (2012). The arts of memory: Comparative perspectives on a mental artifact. *HAU: Journal of Ethnographic Theory*, *2*(2), 451–485. https://doi.org/10.14318/hau2.2.025
Severi, C. (2019). Their way of memorizing: Mesoamerican writings and native American picture-writings. *Res: Anthropology and Aesthetics*, *71–72*, 312–324. https://doi.org/10.1086/706117
Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford University Press.
Snow, D. (2008). Cantonese as written standard? *Journal of Asian Pacific Communication*, *18*(2), 190–208. https://doi.org/10.1075/japc.18.2.05sno
Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition* (2nd ed.). Blackwell.
Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., … Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(26), 10587–10592. https://doi.org/10.1073/pnas.0903616106
Tamariz, M., & Kirby, S. (2015). Culture: Copying, compression, and conventionality. *Cognitive Science*, *39*(1), 171–183. https://doi.org/10.1111/cogs.12144
Taylor, I., Taylor, M. M., & Taylor, M. M. (1995). *Writing and literacy in Chinese, Korean and Japanese*. John Benjamins.
Taylor, T. J. (2000). Language constructing language: The implications of reflexivity for linguistic theory. *Language Sciences*, *22*(4), 483–499. https://doi.org/10.1016/S0388-0001(00)00016-4
Tigwell, G. W., & Flatla, D. R. (2016). *Oh that's what you meant! MobileHCI '16 Proceedings of the 18th international conference on human–computer interaction with mobile devices and services adjunct, MobileHCI 2016*, pp. 859–866. Retrieved from https://doi.org/10.1145/2957265.2961844
Turoman, N., & Styles, S. J. (2017). Glyph guessing for "oo" and "ee": Spatial frequency information in sound symbolic matching for ancient and unfamiliar scripts. *Royal Society Open Science*, *4*(9), 170882. https://doi.org/10.1098/rsos.170882
Unger, J. M. (2003). *Ideogram: Chinese characters and the myth of disembodied meaning*. University of Hawai'i Press.
Unger, J. M., & DeFrancis, J. (1995). Logographic and semasiographic writing systems: A critique of Sampson's classification. In I. Taylor & D. R. Olson (Eds.), *Scripts and literacy: Reading and learning to read alphabets, syllabaries and characters* (pp. 45–58). Springer Netherlands. Retrieved from https://doi.org/10.1007/978-94-011-1162-1_4
Valério, M., & Ferrara, S. (2019). Rebus and acrophony in invented writing. *Writing Systems Research*, *11*(1), 66–93. https://doi.org/10.1080/17586801.2020.1724239
Vandergriff, I. (2013). Emotive communication online: A contextual analysis of computer-mediated communication (CMC) cues. *Journal of Pragmatics*, *51*, 1–12. https://doi.org/10.1016/j.pragma.2013.02.008
Verhoeven, L., & Perfetti, C. (2022). Universals in learning to read across languages and writing systems. *Scientific Studies of Reading*, *26*(2), 150–164. https://doi.org/10.1080/10888438.2021.1938575
Wagner, R. K., Zirps, F. A., Edwards, A. A., Wood, S. G., Joyner, R. E., Becker, B. J., … Beal, B. (2020). The prevalence of dyslexia: A new approach to its estimation. *Journal of Learning Disabilities*, *53*(5), 354–365. https://doi.org/10.1177/0022219420920377
Wang, H. (2014). *Writing and the ancient state: Early China in comparative perspective*. Cambridge University Press.
Wengrow, D. (2008). Prehistories of commodity branding. *Current Anthropology*, *49*(1), 7–34. https://doi.org/10.1086/523676
Wharton, T. (2009). *Pragmatics and non-verbal communication* (1st ed.). Cambridge University Press.
Whitman, J. (2011). The ubiquity of the gloss. *Scripta*, *3*, 95–121.
Winters, J., Kirby, S., & Smith, K. (2018). Contextual predictability shapes signal autonomy. *Cognition*, *176*, 15–30. https://doi.org/10.1016/j.cognition.2018.03.002
Winters, J., & Morin, O. (2019). From context to code: Information transfer constrains the emergence of graphic codes. *Cognitive Science*, *43*(3), e12722. https://doi.org/10.1111/cogs.12722
Yin, Y., & Zhang, Q. (2021). Chinese characters are read using not only visual but also writing motor information. *Psychophysiology*, *58*(1), e13696. https://doi.org/10.1111/psyp.13696
Zagala, S. (2004). Vanuatu sand drawing. *Museum International*, *56*(1–2), 32–35. https://doi.org/10.1111/j.1350-0775.2004.00455.x
Zalta, E. N. (2022). Gottlob Frege. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2022). Metaphysics Research Lab, Stanford University. Retrieved from https://plato.stanford.edu/archives/fall2022/entries/frege/
Zhang, H., & Mace, R. (2021). Cultural extinction in evolutionary perspective. *Evolutionary Human Sciences*, *3*, E30. https://doi.org/10.1017/ehs.2021.25

Open Peer Commentary

The disadvantage of ideography

Ignasi-Xavier Adiego[a] and Miguel Valério[b]

[a]Institut Universitari del Pròxim Orient Antic (IPOA), Universitat de Barcelona, Barcelona, Spain and [b]Department of Prehistory, Universitat Autònoma de Barcelona, Bellaterra, Spain
ignasi.adiego@ub.edu
miguel.valerio@uab.cat

doi:10.1017/S0140525X23000584, e234

Abstract

Morin argues that fully fledged ideography is hindered by a problem of standardization. In general, the argument makes a great deal of sense, but we find that it is easier to grasp if, unlike the author, we consider ideography a language – one whose symbols (graphs) are not as effortless to produce as those of spoken and sign languages (phones, gestures).

Morin's article poses a puzzle and a solution. The puzzle is that ideography is "conceptually and cognitively plausible" (target article, short abstract), but rare. Perhaps nonexisting is more accurate, as there is no demonstrable real-life example of a functioning, full-blown ideography: Nsibidi is reportedly a Nigerian secretive code whose precise nature cannot be scrutinized by outsiders, whereas the designed ideographies of Wilkins, Bliss, and Neurath did not gain a community of users. By ideography

Morin means a graphic code (a code made of pictures, and not necessarily iconic ones) that is both *self-sufficient* and *generalist*. The former implies that it encodes meaning directly and does not need aid or glossing from other codes (including spoken language) to communicate properly; the latter means that the topics on which it communicates are broad (potentially infinite, one might say). Writing does both those things, and Morin does well to stress that it does so because it borrows such traits from spoken language, which is what it encodes (it is a glottography), always through a mix of semantic and phonetic signs. Hence, writing does not count.

The solution Morin offers for the puzzle is that the development of fully fledged ideographies is deterred by a problem of standardization: It is difficult to get all users of the code to "ascribe the same meanings to the same symbols" (target article, sect. 1, para. 12). In that regard, he well notes that spoken and signed languages have an advantage over ideography because they produce "cheap," "fast," and "transient" (target article, sect. 6.2) symbols. Phonemes and gestures "require little effort or time to produce" (target article, sect. 6.2) and so they can easily be corrected through replacement. These traits make spoken and signed languages self-sufficient as well as "self-standardizing". With them, communication can be maintained and repaired without the aid (or intromission) of another code.

The argument makes sense, but we find that this is easy to grasp only after we solve a paradox created by Morin. He contends that "it is not helpful to think of graphic codes as languages" (target article, sect. 1, para. 9) and sees convenient to reserve the latter term "for means of expression that are clearly codified, generalist, and self-sufficient" (target article, sect. 2.1, para. 10), by which he means preferably spoken languages. This is the paradox. It is hard to think of a code that encodes *meaning* directly (rather than a spoken idiom, like writing), is self-sufficient (it needs no glossing), and is used for communication in potentially limitless domains – three things spoken and signed languages do – *but is not* a language. From that perspective, the kind of ideography Morin has in mind – generalist, self-sufficient, and encoding meaning – *is* as much *a language* as spoken idioms and signed languages. The only difference is that, rather than *gestures* or *phones*, its symbols are made up of images or, if we will, *graphs*. To prove this point: Not even Morin can escape terming "visual languages" (target article, sect. 2.1, para. 3) the kind of graphic codes he envisages.

Once we accept that ideography is as much a language as spoken and signed ones, we can compare them. It then emerges that what makes ideographies less practicable is the way the code's symbols are *produced*. To make graphic signs, humans need to grab a pencil, a pen, a brush, or chalk and apply ink or a mineral on some surface; to carve stone or metal with some tool; use a finger or even a stick to draw shapes in the sand; make incisions or impressions on clay; press the keys of a device; and so on. In other words, even when our hands and fingers can go without using an instrument, we will still need to resort to an *external* medium to produce images. We need to manipulate something from outside our bodies. Conversely, spoken and signed languages can do only with human anatomy, so that the production of their symbols will always require less time and effort than the creation of graphs. Ultimately, what is difficult is to imagine (for now) a scenario in which a graphic language could gain independence from spoken language and evolve in parallel to it – not even speech impairment, which signed languages address, seems to do. (Unless, of course, we enter the terrain of speculative fiction and imagine a future in which human anatomy has evolved to the point where it can produce fast and transient pictures on its own.)

We could also be suspicious about the idea that any new fully fledged language, with its own syntax and rules, might develop bypassing spoken language. Language mediates in all human action and some form of verbalization would surely be involved in the creation and first acquisition of a new code. Yet, suitably, Morin seems to have in mind a process comparable to the historical development of pidgins (among speakers of different languages) and their evolution (on occasion) into full-blown creole languages, and the creation of signed languages and their subsequent transmission: "Self-sufficiency is about usage, not acquisition" (target article, sect. 2.2, para. 3).

If despite its disadvantage of production a viable ideography were still to emerge and spread in the future, unlike Morin we do not think it could "break language barriers" (target article, sect. 1, para. 3). At least no more than any spoken or sign language, however large its number of users. First, users of other languages would have to learn it. Second, it would hardly ever be the only language around. Once set in motion, we can suspect it would undergo variation and change, eventually splitting into varieties of codes that would ultimately become mutually unintelligible – like any language. Its potential asynchronous nature would then emerge as its main advantage. Yet spoken and signed languages are now just as asynchronous, thanks to audio and video recorders, which Morin is perhaps too quick to dismiss as very recent changes.

Financial support. This research received support from a María Zambrano Postdoctoral research grant (M.V.), funded by the European Union-Next GenerationEU.

Competing interest. None.

A cognitive account of the puzzle of ideography

Xerxes D. Arsiwalla[a,b]

[a]Universitat Pompeu Fabra, Barcelona, Spain and [b]Wolfram Research, Champaign, IL, USA
x.d.arsiwalla@gmail.com

doi:10.1017/S0140525X23000730, e235

Abstract

We posit a cognitive account of the puzzle of ideography, which complements the standardization account of Morin. Efficient standardization of spoken language is phenomenologically attributed to a modality effect coupled with chunking of cognitive representations, further aided by multisensory integration and the serialized nature of attention. These mechanisms explain why languages dominate graphic codes for general-purpose communication.

The puzzle of ideography can be broadly described as the near absence or rarity of self-sufficient generalist graphic codes for use in human communication. Morin proposes a resolution to this puzzle in terms of a "specialization hypothesis" (target article, sect. 4), which is subsequently explained in terms of a "standardization account" (target article, sect. 6). The latter does well in predicting empirical features in communication practice. Morin also pays due diligence to the alternative "learning account" (target article, sect. 5.4) explanation, which helps the discussion. However, both these explanations have primarily been cast in behavioral terms. *We argue that there also exists a "cognitive account" which complements the behavioral explanation*.

The cognitive account we propose is anchored on two well-studied cognitive phenomenon: (i) *Modality effect*, and (ii) *Chunking*. We show that (i) and (ii) offer a complementary explanation to the specialization hypothesis. Furthermore, combining the above two phenomena with two design features pertaining to cognitive processing: *Multisensory integration*, and the *serialized nature of attention*, explains what enables human languages to become self-standardizing, and also the precise role of learning.

The modality effect has extensively been studied in experimental psychology (Murdock, 1968; Penney, 1989). It refers to improved recall of items on a list when presented verbally in contrast to a purely visual representation. Subsequent studies, investigating this advantage of auditory over visual text modality during learning, have attributed the effect to factors including reduction in extraneous cognitive load, effective expansion of working memory, and early-sensory processes (Rummer, Schweppe, Fürstenberg, Scheiter, & Zindler, 2011; Tabbers, Martens, & Van Merrienboer, 2001). Additionally, the instructional benefits of presenting information across modalities found support in a meta-analysis involving 43 independent studies (Ginns, 2005).

Prima facie, the modality effect concerns learning and memory. However, one could argue that what can be learned or recalled better is also easier to repair during the process of turn-taking, and hence easier to standardize within a population. With respect to the specialization hypothesis, what this means is that at the population level, spoken forms of information exchange will dominate those that are purely visual. This comprises one part of our account.

The other part relevant to the cognitive account is chunking. In cognitive psychology, this refers to the process by which individual pieces of an information set are bound together into a meaningful whole (de Groot, 1968; Miller, 1956). Evidence for perceptual chunking is found in how primitive stimuli are grouped into larger conceptual groups, such as the manner by which letters are grouped into words, sentences, or paragraphs (Johnson, 1970). Chunking is an effective strategy for overcoming capacity limits of working and long-term memory via coherent grouping of information (Laird, Rosenbloom, & Newell, 1984). It has been observed in paradigms involving verbal learning, learning of perceptual-motor skills, expert memory, language acquisition, and learning multiple representations (Gobet et al., 2001).

What is relevant here is that chunking facilitates the creation of higher-order cognitive representations specific to the individual's perceptions and past experiences. Thus, allowing for greater abstraction and generality of cognitive representations. One could argue that this is not only about learning or memory, but underlies the core cognitive mechanism for generating flexible higher-order conceptualizations.

Why is this so important vis-à-vis the specialization hypothesis? The answer is that ideographs and other graphic codes that are not languages, allow for very limited abstraction and higher-order representation via compositionality of symbols. In language, compositions from only a limited set of letters allow for creating an astoundingly large number of new concepts and higher-order representations. The same is not true for graphic codes (where base symbols already represent a specific concept). Take for example, the case of Bliss: Composing symbols does not allow for reduction in complexity of expressions or creation of higher-order chunks. In contrast, a study presenting stimuli generated by an artificial grammar showed that subjects unintentionally learned to respond efficiently to the underlying structure (Servan-Schreiber & Anderson, 1990). Moreover, it was demonstrated that the learning process was chunking and that grammatical knowledge was implicitly encoded in a hierarchical network of chunks.

The point we want to emphasize is that the mechanism of chunking is precisely the brain's operationalization of compositionality of cognitive representations. That the latter happens to be what languages are designed for, explains why the specialization hypothesis holds at all. Compositionality is the reason languages beat graphic codes for general-purpose communication. This also explains why writing is the only exception of a graphic code that is general-purpose: Because it simply tags upon language.

Finally, we discuss two cognitive design features relevant to the puzzle. First, multisensory integration. The effectiveness of a stimulus of one modality in eliciting attentive behaviors and faster reaction times was dramatically affected by the presence of a stimulus from another modality (Hershenson, 1962; Stein, Meredith, Huneycutt, & McDade, 1989). The brain reinforces saliency of representations that use congruent evidence from multiple modalities, thus, making the case for enhanced learnability of concepts involving unambiguous auditory and visual representation. This is true for words of a spoken language but not so for graphic codes that one may verbalizable in different ways and hence be difficult to standardize.

The second feature is the serialized nature of attention. This matters at the decoding/read-out stage. The brain's attention system seems designed to (consciously) process one task at a time (Baars, 1998). Studies have shown divided attention impedes performance (Pashler, Johnston, & Ruthruff, 2001). Spoken language is strictly temporal and hence serialized. Written language tags along that temporality. However, complex ideographs may have layouts as two-dimensional patterns, such that different decoders may assume a different read-out order. At the population level, this directly speaks to the difficulty in standardizing such codes.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Baars, B. J. (1998). Metaphors of consciousness and attention in the brain. *Trends in Neurosciences*, *21*(2), 58–62.

de Groot, A. D. (1968). Thought and choice in chess. In P. C. Warson & P. N. Johnson-Laird (Eds.), *Thinking and reasoning* (pp. 145–151). Penguin.

Ginns, P. (2005). Meta-analysis of the modality effect. *Learning and Instruction*, *15*(4), 313–331.

Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236–243.

Hershenson, M. (1962). Reaction time as a measure of intersensory facilitation. *Journal of Experimental Psychology*, *63*(3), 289.

Johnson, N. F. (1970). The role of chunking and organization in the process of recall. In *Psychology of learning and motivation* (Vol. 4, pp. 171–247). Academic Press.

Laird, J. E., Rosenbloom, P. S., & Newell, A. (1984). Towards chunking as a general learning mechanism. In *AAAI'84: Proceedings of the Fourth Association for the Advancement of Artificial Intelligence (AAAI) Conference on Artificial Intelligence, Austin, Texas, USA, August 1984*, pp. 188–192.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81.

Murdock Jr, B. B. (1968). Modality effects in short-term memory: Storage or retrieval?. *Journal of Experimental Psychology*, *77*(1), 79.
Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology*, *52*, 629.
Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition*, *17*(4), 398–422.
Rummer, R., Schweppe, J., Fürstenberg, A., Scheiter, K., & Zindler, A. (2011). The perceptual basis of the modality effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *17*(2), 159.
Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive chunking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(4), 592.
Stein, B. E., Meredith, M. A., Huneycutt, W. S., & McDade, L. (1989). Behavioral indices of multisensory integration: Orientation to visual cues is affected by auditory stimuli. *Journal of Cognitive Neuroscience*, *1*(1), 12–24.
Tabbers, H. K., Martens, R. L., & Van Merrienboer, J. J. (2001). The modality effect in multimedia instructions. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the Annual Meeting of the Cognitive Science Society 2001, Edinburgh, Scotland, UK*, Vol. 23, pp. 1024–1029.

Pragmatic interpretation and the production of ideographic codes

Leda Berio[a]*, Berke Can[b]*, Katharina Helming[c]*, Giulia Palazzolo[b,d]* and Richard Moore[b]

[a]Institut für Philosophie II, Ruhr-Universität Bochum, Universitätsstraße 150, Bochum, Germany; [b]Department of Philosophy, University of Warwick, Coventry, UK; [c]Department of Psychology, University of Warwick, Coventry, UK and [d]Department of Letters and Modern Cultures, University La Sapienza, Rome
Leda.Berio@rub.de; ledaberio.com
Berke.Can@warwick.ac.uk
Katharina.Helming@warwick.ac.uk; https://sites.google.com/site/katharinaahelming/home
Giulia.Palazzolo.1@warwick.ac.uk; https://warwick.ac.uk/fac/soc/philosophy/people/palazzolo
Richard.Moore@warwick.ac.uk; https://warwick.ac.uk/fac/soc/philosophy/people/summaries/moore/

doi:10.1017/S0140525X23000717, e236

Abstract

We argue that the problem of ideographic codes stems from neither learnability nor standardization, but from a general issue of pragmatic interpretation. As ideographic codes increase in expressive power, in order to reduce ambiguity, they must become more detailed – such that production becomes more cumbersome, and requires greater artistry on the part of users, limiting their capacity for growth.

The puzzle of ideography is, Morin argues, that in the history of written sign systems, there seem to be no self-sufficient, generalist ideographic codes. Ideographic codes are sign systems that represent ideas without the mediation of an auxiliary code. Most writing systems are generalist because they can be used for a wide range of communicative functions. However, they are not ideographic because they are glottographic: They represent semantic contents only by representing the sounds of words in spoken languages. Without knowledge of the relevant spoken languages, written languages remain uninterpretable. Although ideographic codes exist, their use is restricted to specific and specialized contexts. For example, emojis and musical notation are self-sufficient codes, because they encode meaningful units without the need for external mediation. However, their use is restricted to the expression of simple emotions and music – making their use of limited generality.

Morin's question is why generalist ideographic codes, despite being conceptually possible, have not developed. He rejects one answer to this question, which proposes that ideographic codes are more difficult to learn than nonideographic ones (the "learnability problem") on the grounds that it underestimates human capacity for learning. Instead he defends a "standardization problem": The problem with ideographic codes is getting everyone to agree on which code to use. In face-to-face communication, he argues, standardization issues do not arise because people can give and receive real-time feedback to clarify which sign system they are using. However, because ideographic codes are used for communicating with temporally and spatially distant interlocutors, this opportunity for such feedback is missing, and standardization remains a problem.

Morin's account gets something right – namely, the interpretability of ideographic signs. However, this problem of interpretability is not best characterized as a problem of standardization. Standardization would not be an insurmountable challenge if generalist ideographic languages were independently viable. Even if ideographic codes are largely used in distal communication, a viable, generalist ideographic code could be learned and used in face-to-face interaction, with opportunities for face-to-face repair, and in such cases a standardization challenge could be overcome. In that case, the problem with ideographic codes must be more fundamental. We think the underlying issue is not standardization of learnability, but pragmatic interpretation.

As proponents of both the learnability and standardization views agree, ideographic codes are limited. In Morin's words, both explanations agree that "graphic codes cannot encode a broad range of meanings" (target article, sect. 5, para. 1). We interpret this claim as consensus that, in contrast to glottographic codes that exploit the countless combinatorial possibilities of representing phonemes, there may be a comparatively small number of ideographic signs that can be both efficiently reproduced and used to represent ideas in a reasonably unambiguous way. This consensus supports a fairly crude and empirically untested hypothesis, which we nonetheless find intuitive.

Untested hypothesis: There is a limit to the complexity of the messages that ideographic codes can be used to express before they become either unwieldy or uninterpretable. This is a consequence of the very format of ideographic representation, which requires (something akin to) drawing skills and the production of more complex symbols. The limited use of ideographic codes therefore stems from an intrinsic feature (their ideographic format), rather than a contingent factor (the fact they are mostly used asynchronously). This is the main reason for the nonexistence of generalist ideographic codes.

If we are right, then the users of ideographic codes must at some point attempt to plot a course between Scylla and Charybdis. Scylla is the threat of unwieldiness; Charybdis the problem of ambiguity. As messages get more complex, the users of ideographic codes face a choice. Either they must produce more elaborate, and more detailed ideographic codes sufficient for the visual discrimination of similar but potentially important semantic differences; or they must accept the limited expressive power of their code. In the former case, producing the code will

*These authors contributed equally to the authorship of this commentary.

become slower and more cumbersome, especially in comparison to “cheap and fast” glottographic signs. Producing ideographic codes will also require greater artistry on the part of users, making the ideographic code a less appealing tool for communication than glottographic codes already in use. In the latter case, where the limited expressive power of the code is accepted, complex messages will remain ambiguous – and a verbal gloss will be needed for interpreting complex messages. This will undermine the self-sufficiency of the ideographic code, and make it less useful for the kinds of distal communication for which it is supposedly well suited.

This problem does not seem to be faced by glottographic codes. As pressure increases for a more expressive – and thus more general – sign system, natural languages expand. Their users innovate new words and grammatical forms, to permit the expression of a wider range of ideas (e.g., Moore, 2021; Progovac, 2015). However, these new communicative tools seemingly consist largely of the introduction of grammatical forms, and so punctuation marks, rather than on the introduction of new glottographic marks. New words can therefore be handled using existing alphabets, placing little pressure on existing glottographic codes. Even if ideographic codes recombine elements, they surely cannot do so in a manner as minimal and elegant as the couple of dozen phonemes used by individual spoken languages. This is seemingly a consequence of the format of graphic codes: Increasing expressive power means increasing the complexity of ideographic representations. This is not the case with glottographic codes.

We reiterate that our hypothesis is crude and untested. However, it generates a prediction. It suggests that code systems that start off as ideographic may, over time, become more glottographic – as pressure grows for a more expressively powerful code system. In such cases, a natural compromise is that, as ideographic capacity is reached, then the expressive power of the code can be increased by incorporating glottographic elements of extant natural languages. Meanwhile, pure ideographic codes will remain suitable only for specialized use.

Financial support. Leda Berio’s work is supported by the NRW Profillinie “Interact!” (PROFILNRW-2020-135). Berke Can is supported by a University of Warwick Chancellor’s International Scholarship. Katharina Helming and Richard Moore are supported by UKRI Future Leaders Fellowship grant MR/S033858/1: The Communicative Mind. Giulia Palazzolo is supported by a La Sapienza University International Scholarship.

Competing interest. None.

References

Moore, R. (2021). The cultural evolution of mind-modelling. *Synthese*, *199*(1), 1751–1776.
Progovac, L. (2015). *Evolutionary syntax*. Oxford University Press.

The feasibility of ideography as an empirical question for a science representational systems design

Peter C.-H. Cheng

Department of Informatics, University of Sussex, Brighton, UK
p.c.h.cheng@sussex.ac.uk; http://users.sussex.ac.uk/~peterch/

doi:10.1017/S0140525X23000766, e237

Abstract

The possibility of ideography is an empirical question. Prior examples of graphic codes do not provide compelling evidence for the infeasibility of ideography, because they fail to satisfy essential cognitive requirements that have only recently been revealed by studies of representational systems in cognitive science. Design criteria derived from cognitive principles suggest how effective graphic codes may be engineered.

As Morin states in the conclusion of the target article, whether a generalist and self-sufficient graphic code is possible is ultimately an empirical question. However, an alternative to Morin’s explanation of why no successful graphic code has been created – so far – comes from the perspective of cognitive science.

The history of the invention of airplanes provides an instructive analogy. Before 1903 independent heavier-than-air flight was essentially the domain of birds (and pterodactyls). Human flight was restricted to hot air balloons and kites. Repeated attempts by well-resourced inventors met with failure. So, the feasibility of independent human flight then seemed intuitively implausible. Bradshaw (1992) attributes the success of the Wright brothers to their systematic identification of key functions essential to the operation of airplanes. The brothers developed mechanisms to perform those functions. For example, they recognized the importance of attitudinal control and decomposed it into subfunctions to separately control pitch, roll, and yaw (i.e., ailerons, wing warping, and rudder). Only when they had mechanisms for all the necessary functions did the Wrights integrate them into a complete airplane.

It appears that the design of graphic codes is currently at a pre-Wright brother’s stage where the creation of graphic codes has so far focused on parametric permutations of features of whole ideographic systems. The failure of extant graphic codes may be attributed to missing cognitive mechanisms to localize the control of essential semantic functions. By identifying such functions we can make the prospect of designing a full graphic code more realistic and take a step towards answering the empirical question.

Recent studies in cognitive science, particularly on visuospatial representational systems, suggest key ideographic functions. These are functions that enable *producers* to encode meanings in, and *readers* to access meanings from, symbols in representational systems. Many factors are now recognized, such as: *locational indexing* of information (Larkin & Simon, 1987); *isomorphic mapping* between concepts and tokens (Gurr, 1998); levels of *specificity* (Stenning & Oberlander, 1995); *free rides* (Shimojima, 2015); the *matching of quantity scales* (Zhang, 1996); and *multilevel coherent interpretive schemes* (Cheng, 2002, 2011).

From such findings we may distil essential representational functions that a full ideography needs. Here is one attempt at such a collection:

(1) *Ground symbols* are primitive symbols that establish foundational concepts from which higher order ideas can be derived (including conceptual dimensions and situational models, see below). The notion of *base-level* categories (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) may aid the selection of ground symbols.
(2) *Conceptual dimensions* are generic concepts that are underpinning characteristics of ideas (cf., Gardenfors, 2014). Some examples include: states versus actions; spatial scale;

temporal scale; states of matter; frames of reference; psychological valence; truth value; and degrees of certainty. A graphic code should possess classes of symbols whose function is to identify values on such dimensions for the purpose of qualifying the meaning of given symbols. The selection of conceptual dimensions may be guided by ideas about *core knowledge* (Spelke & Kinzler, 2007), *geometric conceptual spaces* (Gardenfors, 2014), quantity scales (Stevens, 1946), and others.

(3) *Context modelling* builds graphical expressions to represent meaningful situations as compositional configurations of symbols and conceptual dimensions. Context models provide rich settings to guide users to intended meanings, for instance by triangulation of conceptual dimensions. Ideas about *image schemas* (Lakoff & Johnson, 2008) may provide constraints on the form of context models.

(4) *Indexicalization* takes a context model and, in a decompositional manner, identifies some subconfiguration of symbols, or a part of a single symbol, to introduce a symbol for a new specific meaning.

(5) *Iconization* converts a complex graphic symbol (context model) into a simple icon-like symbol for ease of future use, by disposing of the details that were necessary to initially establish the meaning but that are not essential to the intended concept.

(6) *Meta-semantics* is the greatest challenge for creating a full graphic code. This concerns the specification of abstract and intangible concepts, which in the case of words are often defined with reference to other abstract words. A graphic code needs mechanisms to generate symbols for abstractions and reifications, and for generalizations and specializations. Commonality and contrasts of meanings across existing symbols could be one approach to drive the coining of symbols for abstract concepts. This will require the provision of *meta-symbols* whose purpose is to instruct the user to associate a new graphical object with the implied meaning.

Such functions must operate symbiotically in a graphic code. For instance, the code can be generative, in and of itself, when it possesses mechanisms to create symbols for new meanings from existing symbols without appealing to an external natural language oracle (cf., Bliss symbolics). The code can use the full richness of spatial, geometric, topological, and mereological graphical devices to encode meanings and eschew the linear concatenation of symbols as the basic organizational format of expressions. (In terms of the airplane analogy, linear concatenation in Bliss symbolics is flapping wings.) Further, when users are familiar with the representational tools for each function, one can reasonably imagine users communicating in real time by jointly editing each other's expressions. For instance, a user who has a conventionalized symbol could rewind the iconization or indexicalization processes by drawing the context model and applying meta-semantics to unfold that symbol's meaning. Thus, contrary to Morin's claim, it is possible that a graphic code could share with spoken languages the "cheap and transient signals, allowing for easy online repairing of miscommunication … where the advantages of common ground are maximized" (target article, sect. long abstract, para. 1). In turn, this undermines one pillar of Morin's argument for the standardization problem, but the learnability problem remains as an empirical question.

The idea of deploying semantic functions in the design of a graphic code is not purely theoretical. My own efforts at designing graphical notations for conceptually challenging information intensive topics – albeit specialist – show that when such functions are satisfied, users find the new notations cognitively superior to conventional representations (e.g., Cheng, 2002, 2011, 2012, 2020).

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Bradshaw, G. (1992). The airplane and the logic of invention. In R. N. Giere (Ed.), *Cognitive models of science* (pp. 239–250). University of Minnesota Press.
Cheng, P. C.-H. (2002). Electrifying diagrams for learning: Principles for effective representational systems. *Cognitive Science*, *26*(6), 685–736.
Cheng, P. C.-H. (2011). Probably good diagrams for learning: Representational epistemic re-codification of probability theory. *Topics in Cognitive Science* *3*(3), 475–498.
Cheng, P. C.-H. (2012). Algebra diagrams: A HANDi introduction. In P. Cox, B. Plimmer, & P. Rodgers (Eds.), *Diagrammatic representation and inference: 7th International conference on diagrams 2012* (pp. 178–192). Springer Verlag.
Cheng, P. C.-H. (2020). Truth diagrams versus extant notations for propositional logic. *Journal of Language, Logic and Information*, *29*, 121–161.
Gardenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. MIT Press.
Gurr, C. A. (1998). On the isomorphism, or lack of it, of representations. In K. Marriott & B. Meyer (Eds.), *Visual language theory* (pp. 293–306). Springer Verlag.
Lakoff, G., & Johnson, M. (2008). *Metaphors we live by*. University of Chicago Press.
Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*, 65–99.
Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439.
Shimojima, A. (2015). *Semantic properties of diagrams and their cognitive potentials*. CSLI Press.
Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science*, *10*(1), 89–96.
Stenning, K., & Oberlander, J. (1995). A cognitive theory of graphical and linguistic reasoning: Logic and implementation. *Cognitive Science*, *19*(1), 97–140.
Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, *103*(2684), 677–680.
Zhang, J. (1996). A representational analysis of relational information displays. *International Journal of Human Computer Studies*, *45*, 59–74.

Functional ideographies are composite semiotic systems

Stephen Chrisomalis

Department of Anthropology, Wayne State University, Detroit, MI, USA
chrisomalis@wayne.edu; https://clasprofiles.wayne.edu/profile/dz6179

doi:10.1017/S0140525X23000675, e238

Abstract

All sufficiently large functional notations (ideographic or otherwise) are composites of discrete, structured elements (e.g., phonemes, morphemes, numerals). We must consider not only the modality but also the structure of the existing, workable ideographic/semasiographic systems we know (e.g., musical and numerical notation) to establish the cognitive limitations militating against humans memorizing and standardizing domain-general ideographies that would parallel written language.

Debates over the nature and feasibility of ideography are not going away anytime soon. Often, these have not advanced us much further than Edgerton (1941), who noted over 80 years ago the widespread presence of ideograms in English (and indeed many other writing systems), such as numerals and signs like $, which cannot be subsumed under terms like "logogram." Morin has undertaken a salutary task here – asking not only whether ideograms exist (they surely do), nor even whether fully ideographic notations can exist, but how, where, and why they exist – and conversely, why domain-general, self-sufficient notations are rare or nonexistent. By drawing our attention to how notations are shared and transmitted, Morin has laid a pathway for future empirical studies both in cognitive science and in the field of writing systems.

We should be careful not to conflate Liberman's "openness" or Morin's "generality" with the number of messages capable of being conveyed. A notation can be infinite while still being highly restricted. Many numerical notations are infinite (adding another zero to the end, or another tally-mark to the score) and musical notations are capable of encoding an infinite number of musical themes and scores. It is that they are domain-specific, rather than that they are infinite, that should attract our attention. But this raises the question: What affords a notation this infinity? For Chomsky (1980), the relevant property is what he calls "discrete infinity" – that a limited set of discrete elements can be combined to produce a potential infinity of representations of the sort necessary for language. But this is precisely what is needed for nonlinguistic codes as well. A numerical notation which required that users find a different sign for every number would require an impossibly inhuman mind, and, of course, would constantly require recourse to inventing new symbols (Chrisomalis, 2020, pp. 1–26).

So, when we are looking for notations that will work – that humans will be able to accept and share – we should be looking for systems that have a componential or composite structure (Chrisomalis, 2018). It is easy to be misled by issues of modality (visual vs. auditory) here. Sign languages are for all relevant purposes, languages, and have all the properties that languages have (including phonology – see Brentari, 2019) and are scaffolded by the same cognitive capacities that underlie spoken languages. A general-purpose system of nonlinguistic gestures would be just as nonfunctional as an emoji-only pseudo-language. As Morin points out in the case of Blissymbolics, composite structure is not a sufficient condition, but I believe it is a necessary one. We need more inquiry into how much or what kinds of componential structure make a notation shareable and learnable.

The problem with ideographies is that many ideas are not easily encoded componentially – or at least, no one has done so without also encoding a lot of linguistic information. Given that languages are universal to humans, and languages already have a componential structure, perhaps this is not so surprising. Why bother with a new system for ideas when every speaker already has one at hand? Parts of systems – such as Chinese radicals or Egyptian hieroglyphic determiners – can be described as ideographic or semasiographic, but these are far less important to these scripts than the myths tell (DeFrancis, 1986). Perhaps we have imagined that ideographies are far more desirable than they really are. The obsession of Leibniz, Wallis, and other early modern Europeans with universal language schemes (and their concomitant interest in Chinese writing) may be a solution to a problem no actual script inventor actually had (Knowlson, 1975).

Morin turns our attention, rightly, to "lock-in effects" that facilitate standardization and consistency in graphic notations as opposed to language, which develops differently because of its synchronous and face-to-face qualities. But of course notations do change, and sometimes quite rapidly. The transformation of the West African Bamum script from a logographic or ideographic notation to an alphasyllabic one took place within a generation (Kelly, 2018) under rapidly changing conditions of colonization and resistance. At a somewhat longer scale, the introduction of phonography as proto-cuneiform transformed into Sumerian and later scripts (while retaining important nonphonographic elements) shows how an essentially translinguistic notation need not be eternally so. Although the strongest teleological, unilineal evolutionary arguments are clearly wrong, there are nonetheless important cultural-evolutionary patterns on display (Trigger, 1998). If Morin is right, we should be able to find evidence that standardization and social sharing motivated these diachronic transformations.

Moreover, there are highly unusual systems that deserve more attention, precisely because their combinatorial semiotics defy simple labels such as "phonography" and "ideography." For instance, the symbols of the Western Apache shaman Silas John (Basso & Anderson, 1973) encoded semantic, kinesic, and pragmatic information in addition to recording the words associated with particular prayers. Some of the challenges faced in analyzing systems like the notations of the Iron Age Scottish Picts (Lee, Jonathan, & Ziman, 2010) or the *rongorongo* of Rapa Nui (Valério & Ferrara, 2019) may rest in relying on a too-narrow binary, where every system must be a pure record of written language, or else it loses theoretical interest. We should abandon this byproduct of a progressivist mindset in which writing is the gold standard against which other notations are evaluated.

Finally, although Morin does not make much of the distinction between "ideography" and "semasiography," his interest in standardization and sharing motivates exactly this distinction. The terminological issue may seem trivial or driven by the aim of creating new terms for their own sake, but it is neither. Ideography focuses on the idea, and thus on individual cognition, whereas semasiography draws our attention to the *sema*, the sign, and thus to how meanings are shared among interlocutors. The semiotic focus on what makes the interindividual sharing and transmission of signs and their meanings difficult, rather than what makes it conceptually hard for any individual to grasp a notation, suggests that semasiography is really what we ought to be talking about here.

Competing interest. None.

References

Basso, K. H., & Anderson, N. (1973). A Western Apache writing system: The symbols of Silas John. *Science* (*New York, NY*), *180*(4090), 1013–1022.
Brentari, D. (2019). *Sign language phonology*. Cambridge University Press.
Chomsky, N. (1980). *Rules and representations*. Columbia University Press.
Chrisomalis, S. (2018). The writing of numbers. Recounting and recomposing numerical notations. *Terrain. Anthropologie & Sciences Humaines*, *70*. https://journals.openedition.org/terrain/17506.
Chrisomalis, S. (2020). *Reckonings: Numerals, cognition, and history*. MIT Press.
DeFrancis, J. (1986). *The Chinese language: Fact and fantasy*. University of Hawai'i Press.
Edgerton, W. F. (1941). Ideograms in English writing. *Language*, *17*(2), 148–150.
Kelly, P. (2018). The invention, transmission and evolution of writing: Insights from the new scripts of West Africa. In S. Ferrara & M. Valério (Eds.), *Paths into script formation in the ancient Mediterranean* (pp. 189–209). Studi Micenei ed Egeo-Anatolici.
Knowlson, J. (1975). *Universal language schemes in England and France 1600–1800*. University of Toronto Press.
Lee, R., Jonathan, P., & Ziman, P. (2010). Pictish symbols revealed as a written language through application of Shannon entropy. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *466*(2121), 2545–2560.
Trigger, B. G. (1998). Writing systems: A case study in cultural evolution. *Norwegian Archaeological Review*, *31*(1), 39–62.
Valério, M., & Ferrara, S. (2019). Rebus and acrophony in invented writing. *Writing Systems Research*, *11*(1), 66–93.

Communication consistency, completeness, and complexity of digital ideography in trustworthy mobile extended reality

Kevin B. Clark[a–j]

[a]Cures Within Reach, Chicago, IL, USA; [b]Felidae Conservation Fund, Mill Valley, CA, USA; [c]Expert Network, Penn Center for Innovation, University of Pennsylvania, Philadelphia, PA, USA; [d]Network for Life Detection (NfoLD), NASA Astrobiology Program, NASA Ames Research Center, Mountain View, CA, USA; [e]Multi-Omics and Systems Biology & Artificial Intelligence and Machine Learning Analysis Working Groups, NASA GeneLab, NASA Ames Research Center, Mountain View, CA, USA; [f]Frontier Development Lab, NASA Ames Research Center, Mountain View, CA, USA; [g]SETI Institute, Mountain View, CA, USA; [h]Peace Innovation Institute, Stanford University, Palo Alto, CA, USA and The Hague, Netherlands; [i]Shared Interest Group for Natural and Artificial Intelligence (sigNAI), Max Planck Alumni Association, Berlin, Germany and [j]Biometrics and Nanotechnology Councils, Institute for Electrical and Electronics Engineers (IEEE), New York, NY, USA
kbclarkphd@yahoo.com
www.linkedin.com/pub/kevin-clark/58/67/19a; https://access-ci.org

doi:10.1017/S0140525X23000791, e239

Abstract

Communication barriers long-associated with ideographs, including combinatorial grapholinguistic complexity, computational encoding–decoding complexity, and technological rendering and deployment, become trivialized through advancements in interoperable smart mobile digital devices. Such technologies impart unprecedented extended-reality user hazards only mitigated by unprecedented colloquial and bureaucratic societal norms. Digital age norms thus influence natural ideographic language origins and evolution in ways novel to human history.

Morin contends that, of all possible codes humans and perhaps other intelligent agents may invent, adapt, and employ for communication, spoken or signed languages have greatest eco-evolutionary fitness for being both self-sufficient and general tools of expression, with the notable rare exception of flourishing language-linked graphic codes or writing systems. As justification for the biocultural dominance of language over ideography, Morin's assertions emphasize the uniqueness of language to optimize limiting trade-offs between self-sufficiency (i.e., use of language without need to annotate messages through auxiliary codes) and generality (i.e., use of language to produce and comprehend an indefinite variety of affective prosody and propositional linguistic content). But, as the author also recognizes, the capacity of language to optimize trade-offs may differ according to protolanguage and language sophistication, exemplified by the contrasting independent evolution of ancient languages in China, Egypt, Meso-America, and Mesopotamia. Poorly standardized and narrowly specialized codes missing prominent traits of glottographic writing, such as Mesopotamian protocuneiform (Damerow, 2006; Schmandt-Besserat, 2007), failed to inspire later languages and corresponding writing systems with robust versatility and longevity. Based on this and additional evidence, the degree of graphic code standardization and specialization seems to help reciprocally drive the effectiveness, fitness, or optimality of languages, with proliferating apex writing systems representing more than just a few dimensions of their successful languages, such as mere notations for language phonetics, affect, semantics, or structure (Clark, 2017b, 2018). Morin applies this somewhat compelling author-coined "specialization hypothesis" (target article, sect. 4, para. 1) when concluding graphic codes, particularly ideographies, can be self-sufficient or general, such as respective mathematical notations and mnemonic pictographs. Significantly, the exclusion properties of this hypothetical condition are suggestive of the Gödelian incompleteness theorems (Clark, 2018; Clark & Hassert, 2013; Gödel, 1931; Kreisel, 1967), where formal axiomatic systems, including languages and graphic codes, must exist in a universe of graded logicomathematical consistency (i.e., all theorems are true syntax-correct propositions of the system) and completeness (i.e., all true syntax-correct propositions of the system are theorems).

The rational Gödelian incompleteness theorems, similar to predictions from the specialization hypothesis, precisely and accurately prove mutual exclusiveness between strong axiomatic consistency and strong axiomatic completeness. In the case of ideographies and other axiomatic constructs, systems may be nonetheless logicomathematically identified along the continuum of consistency and completeness, allowing ideographies, for example, to be only strongly consistent and specialized, only strongly complete and general, or both weakly consistent and specialized and weakly complete and general (Gödel, 1931; Kreisel, 1967). Contradicting Morin's perspective, the Gödelian framework validates the conjecture that some ideographies may too demonstrate optimality comparable to languages and their writing systems through weaker consistency–completeness conditions and that they should predictably instantiate a bigger presence in the history of human language origins and evolution. Why this is not readily evident after millennia of human language and writing development is perhaps the truer "puzzle of ideography" and one that indeed merits focused attention. Juxtaposing Morin's minor speculations on old-to-new technology influences, the main delimiting constraints on widespread, effective ideography innovation and use arguably have been its inherent combinatorial grapholinguistic complexity, computational encoding–decoding complexity, and difficult technological rendering and deployment (Clark, 2012, 2014, 2015, 2018), each of which relate to Morin's learning and specialization accounts. Traditional communication barriers associated with complexity of pictograms and alternate ideographs now become trivialized through modern advancements in interoperable mobile digital devices, such as smart phones and tablets, smart wearables (e.g., smart glasses), and smart mirrors (De Buyser, De Coninck, Dhoedt, & Simoens, 2016; Lee et al., 2020; Miotto, Danieletto, Scelza, Kidd, & Dudley, 2018). Artificial intelligence/machine learning (AI/ML)-powered virtual technologies enable communicants to easily generate, exchange, interpret, store, and adapt ideographic messages beyond simple stylized emojis in real time, in person nearby or at-a-distance, and within and across populations, cultures, and generations of users, promoting both self-sufficient and general ideographic language emergence and transition (Clark, 2017a, 2020).

Morin's fragmentary views on technology-assisted human performance may be further examined and perfected by the study of contemporary languages and how they may originate, evolve, and devolve through varying combinations of seamless, secure digital technology integration and trustworthy digital ideographic standardization and translation (Clark, 2014, 2017a, 2017b, 2020; Roff, 2020). Digital technology standardization and trustworthiness remain important debated concerns for the discipline and industry of communications, motivating creation and assembly of joint stakeholder caucuses (e.g., government, corporate, consumer, etc.) to devise and enforce laws, policies, and practices that advance technological capabilities in accordance with fundamental human values, principles, rights, and duties (e.g., Glikson & Woolley, 2020; National Academies of Sciences, Engineering, and Medicine, 2021). On-device or more intensive data-compute- and -managed off-device AI/ML capabilities (e.g., connected wireless cloud-computing resources and services for mobile devices, etc.) now enable an enormous range of communication possibilities unattainable with past technologies, such as analog landline telephone networks, wire telegraph systems, printable sheet paper and bounded books, and inscribable clay tablets. Normative technology protections are intended to safeguard users at risk for all sorts of deliberate and incidental harm coupled with human–human, human–machine, and machine–machine interactions. User safety and wellbeing may be threatened by inadequate technology operational specifications and malfunction, user error and abuses, and self-generative AI/ML biases and exploitation, among other technological and ethical dangers. Technology-free language use, of course, has its own safety faults for communicants. Ecoevolutionary pressures that force development of honest language use for meaningful, reliable communications also conserve language vulnerabilities for eavesdropping, deceit, and propaganda within and across systematics boundaries from microbes to humans (Clark, 2014). Language manipulation nonetheless often becomes exaggerated via the computational and expressive power, flexibility, and accessibility of digital communication tools, encouraging advisable standards for disambiguating and labeling honesty and deception, including, but not limited to, exchanges associated with digital user identity filters, content authenticators, cultural translators, and user consent in extended-reality platforms. Although state-of-the-art digital technologies allocate unprecedented resources to evolve ideographies into adaptive living languages, the same technologies impart unprecedented user hazards only mitigated by unprecedented colloquial and bureaucratic societal norms. These digital-age norms may facilitate and/or impede natural language origins and evolution in ways never before observed in human history, regardless of whether ideographies serve as language bases.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Clark, K. B. (2012). A statistical mechanics definition of insight. In A. G. Floares (Ed.), *Computational intelligence* (pp. 139–162). Nova Science. ISBN 978-1-62081-901-2.

Clark, K. B. (2014). Evolution of affective and linguistic disambiguation under social eavesdropping pressures. *Behavioral and Brain Sciences*, *37*(6), 551–552.

Clark, K. B. (2015). Insight and analysis problem solving in microbes to machines. *Progress in Biophysics and Molecular Biology*, *119*, 183–193.

Clark, K. B. (2017a). The humanness of artificial nonnormative personalities. *Behavioral and Brain Sciences*, *40*, e259.

Clark, K. B. (2017b). Dialect structural priming in endangered language evolution, devolution, and protection. *Proceedings of the Royal Society B: Biological Sciences*. Retrieved from https://royalsocietypublishing.org/action/downloadSupplement?doi=10.1098%2Frspb.2014.1574&file=royprsb_el.pdf&

Clark, K. B. (2018). Natural chunk-and-pass language processing: Just another joint source-channel coding model? *Communicative & Integrative Biology*, e1445899.

Clark, K. B. (2020). Digital life, a theory of minds, and mapping human and machine cultural universals. *Behavioral and Brain Sciences*, *43*, e98.

Clark, K. B., & Hassert, D. L. (2013). Undecidability and opacity of metacognition in animals and humans. *Frontiers in Psychology*, *4*, 171.

Damerow, P. (2006). The origins of writing as a problem of historical epistemology. *Cuneiform Digital Library Journal*, *1*, 1–10.

De Buyser, E., De Coninck, E., Dhoedt, B., & Simoens, P. (2016). Exploring the potential of combining smart glasses and consumer-grade EEG/EMG headsets for controlling IoT appliances in the smart home. *2nd IET International Conference on Technologies for Active and Assisted Living*, pp. 1–6.

Glikson, E., & Woolley, A. W. (2020). Human trust in artificial intelligence: Review of empirical results. *Academy of Management Annals*, *14*(2), 627–660.

Gödel, K. (1931). Über formal unentscheidbare Säze der Principia Mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, *38*, 173–198.

Kreisel, G. (1967). Mathematical logic: What has it done for the philosophy of mathematics? In R. Schoeman (Ed.), *Bertrand Russell philosopher of the century* (pp. 201–272). George Allen & Unwin.

Lee, J. H., Kim, H., Hwang, J.-Y., Chung, J., Jang, T.-M., Seo, D. G., … Hwang, S.-W. (2020). 3D Printed, customizable and multi-functional smart electronic eyeglasses (E-glasses) for wearable healthcare systems and human-machine interfaces. *ACS Applied Materials & Interfaces*, *12*, 21424–21432.

Miotto, R., Danieletto, M., Scelza, J. R., Kidd, B. A., & Dudley, J. T. (2018). Reflecting health: Smart mirrors for personalized medicine. *NPJ Digital Medicine*, *1*, 62.

National Academies of Sciences, Engineering, and Medicine. (2021). *Human–AI teaming: State of the art and research needs*. National Academies Press.

Roff, H. (2020). AI deception: When your artificial intelligence learns to lie. *IEEE Spectrum*. Retrieved from https://spectrum.ieee.org/ai-deception-when-your-ai-learns-to-lie#toggle-gdpr

Schmandt-Besserat, D. (2007). *When writing met art: From symbol to story*. University of Texas Press.

Visual languages and the problems with ideographies: A commentary on Morin

Neil Cohn and Joost Schilperoord

Department of Communication and Cognition, Tilburg School of Humanities and Digital Sciences, Tilburg University, Tilburg, The Netherlands
neilcohn@visuallanguagelab.com; www.visuallanguagelab.com
j.schilperoord@tilburguniversity.edu

doi:10.1017/S0140525X23000778, e240

Abstract

Morin argues that ideographies are limited because graphic codes lack a capacity for proliferating standardization. However, natural graphic systems display rich standardization and can be placed in sequences using complex combinatorial structures. In contrast, ideographies are not natural, and their limitations lie in their attempts to artificially force a graphic system to behave like a writing system.

In "The puzzle of ideography" Morin argues that ideographic systems of pictorial signs are limited in their communicative capacities because "Graphic codes can only be standardized for a limited number of meaning–symbol mappings" (target article, abstract). Although we agree that ideographies are limited systems, we disagree with Morin's reasons why, especially the notion that the graphic modality itself is limited in its semiotic capacities. Specifically, ideographies attempt to artificially force a graphic system to behave like a writing system, which itself is an adaptation of the vocal modality into the graphic modality.

First, it is important to recognize that "standardization" or "conventionality" is orthogonal to symbolicity (Peirce, 1940). These notions are often conflated (de Saussure, 1972 [1916]), but standardization or conventionality is how much a signal is patterned and recognized across individuals. These signals (idiosyncratic or patterned) correspond to meanings through various interfaces characterizing their signification (i.e., iconicity, indexicality, and symbolicity). To clarify: A signal in sounds or graphics can be standardized or not, and how that signal corresponds to conceptual structures characterizes its signification(s).

Despite their stereotype as "arbitrary symbols," spoken languages display *all* types of signification, as in Figure 1a (Clark, 1996; Ferrara & Hodge, 2018), and so does the graphic modality (Fig. 1b). Because all modalities use all types of signification,

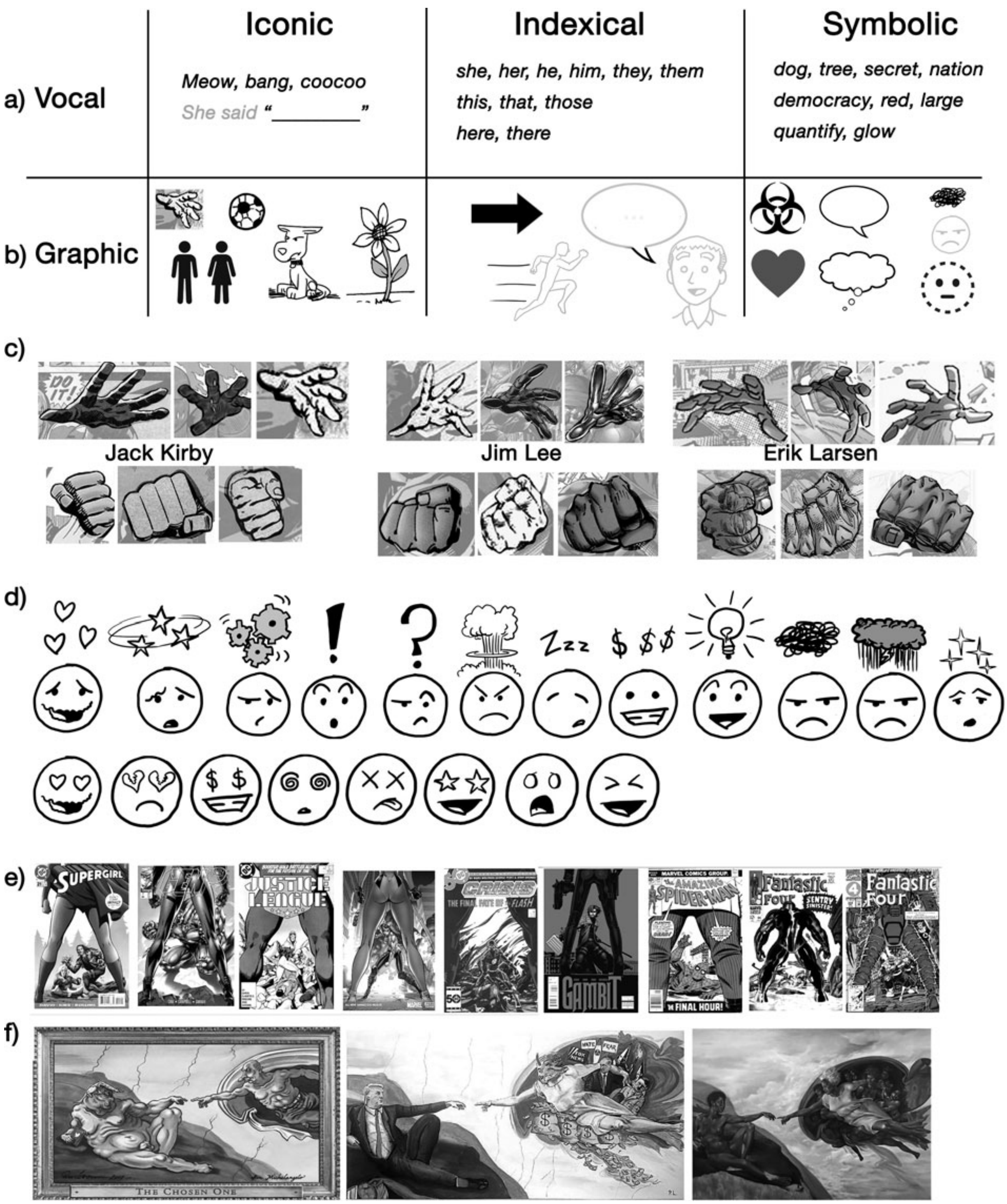

Figure 1 (Cohn and Schilperoord). (a) Vocal and (b) graphic signification, and standardized (c) components of drawings, (d) combinatorial signs, and scene templates (e) without and (f) with reference to other graphics.

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

symbolicity is not the issue, but rather the question is about standardization.

Contrary to Morin's statements, the graphic modality displays voluminous standardization at multiple levels of complexity. Although small "pictographs" like hearts, stars, or peace or radiation signs are easy to recognize as standardized (as in ideographies), all graphics use standardized building blocks (Arts & Schilperoord, 2016; Cohn, 2013; Wilson & Wilson, 1977). Drawings are constructed from low-level visual patterns, such as how people draw eyes, headshapes, houses, and flowers, as in the conventionalized hands by three comic artists in Figure 1c. Graphics also use classes of combinatorial signs, such as the inventory of elements that float above characters heads or replace eyes, which use systematic and symbolic meaning-making (Fig. 1d), in addition to visual vocabulary like motion lines, impact stars, speech balloons, and other highly standardized, culturally variable, graphic codes. Although these small visual "morphemes" can combine into novel pictures, larger units can also be systematized, whether as templates of abstract scenes (Fig. 1e) or templates in reference to other scenes, which often invoke symbolic meanings (Fig. 1f). These observations contrast with the phenomenologically based idea that drawing and graphics are about articulating one's idiosyncratic vision of what one sees, despite this notion being unsupported by cognitive, cultural, and developmental research (Wilkins, 2016; Wilson, 1988) and having erroneous origins (Willats, 2005).

Visual representations also allow a range of conventionalized sequencing. Many patterns of two-unit sequences persist to show causative before–after relations, contrasts, or analogies (Schilperoord & Cohn, 2022). Longer sequences often provide visual lists of related images, such as what is allowed in a park or on an airplane (Cohn & Schilperoord, 2022). Visual narratives also use recursive combinatorial structures for sequential images displaying structural features of linguistic grammars, but operating at a higher-level information structure than the organization of nouns and verbs (Cohn, 2013; Cohn & Schilperoord, 2022). These visual narrative sequences are natural productions of sequential images, and the specific sequencing constructions they use have been shown to vary across cultures' comics (Cohn, 2019), again indicating culturally relative standardization, not universality. In addition, their processing invokes the same neural responses as linguistic syntax and semantics, and the understanding of these visual sequences requires proficiency that is acquired through exposure to and practice with those cultural graphic systems (Cohn, 2020).

In contrast to these natural graphic systems, ideographies use a basic lexicon of simple graphics attempting to have "word" levels of information, often created top-down by individuals, rather than emerging from a language community. These graphics are intended to be sequenced through a syntax, but because graphics do not naturally afford sentence-level combinatorics, they end up parasitic to spoken languages. Writing systems themselves are adaptations of the spoken into the graphic modality, but ideographies then assume graphics should behave like writing systems – forced into sentence-level sequences – to take on "linguistic" properties. This is why ideographies are largely *invented* by specific people (e.g., Blissymbolics, as in Morin's Fig. 4), because they do not proliferate instinctively through human history or cultures.

These invented ideographies are thus systems that attempt to mimic the structures of speech/writing which serve as people's reference point for what "linguistic graphics" should be like. However, this denies the affordances and linguistic properties already displayed by natural graphics in the first place. It should be no wonder then that ideographies don't work.

There is a clear analogue to this in the bodily modality. Sign languages are natural linguistic systems that optimize the affordances of the bodily modality and thereby do things in ways that differ from the structure of speech (Liddell, 2003). Yet attempts persist to force sign languages to have the properties of spoken languages, like Manually Coded English, which maps the lexicon and grammar of spoken English onto the body (Supalla, 1991). As adaptations of one modality to another, these systems deny the bodily affordances that natural sign languages display, just like ideographies deny the affordances apparent in natural graphic systems.

To conclude, ideographies are limited because they attempt to make the natural expressive graphic modality behave like writing, itself a conversion of the spoken modality into graphics. This quality undermines Morin's claim that "understanding why ideography has not worked in the past may help us understand how technology could make it work in the future" (target article, sect. 1, para. 8), because ideographies' unnaturalness will never "make it work" (target article, sect. 1, para. 8). Rather, this discussion raises the importance of investigating the affordances of all our modalities and their meaning-making capacities, and especially a greater integration of graphics into the study of the mind and cognition.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Arts, A., & Schilperoord, J. (2016). Visual optimal innovation. In C. Fernandes (Ed.), *Multimodality and performance* (pp. 61–81). Cambridge Scholars.
Clark, H. H. (1996). *Using language*. Cambridge University Press.
Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. Bloomsbury.
Cohn, N. (2019). Structural complexity in visual narratives: Theory, brains, and cross-cultural diversity. In M. Grishakova & M. Poulaki (Eds.), *Narrative complexity: Cognition, embodiment, evolution* (pp. 174–199). University of Nebraska Press.
Cohn, N. (2020). *Who understands comics? Questioning the universality of visual language comprehension*. Bloomsbury.
Cohn, N., & Schilperoord, J. (2022). Remarks on multimodality: Grammatical interactions in the parallel architecture. *Frontiers in Artificial Intelligence*, *4*, 1–21. doi:10.3389/frai.2021.778060
de Saussure, F. (1972 [1916]). *Course in general linguistics (R. Harris, Trans.)*. Open Court Classics.
Ferrara, L., & Hodge, G. (2018). Language as description, indication, and depiction. *Frontiers in Psychology*, *9*, 1–15. doi:10.3389/fpsyg.2018.00716
Liddell, S. K. (2003). *Grammar, gesture, and meaning in American sign language*. Cambridge University Press.
Peirce, C. S. (1940). Logic as semiotic: The theory of signs. In J. Buchler (Ed.), *The philosophy of Peirce: Selected writings* (pp. 98–119). Kenga Paul, Trench, Trubner.
Schilperoord, J., & Cohn, N. (2022). Before: Unimodal linguistics, after: Multimodal linguistics: An exploration of the before–after construction. *Cognitive Semantics*, *8*(1), 109–140. doi:10.1163/23526416-bja10025
Supalla, S. (1991). Manually coded English: The modality question in signed language development. In S. D. Fischer & P. Siple (Eds.), *Theoretical issues in sign language research. Volume 2: Psychology* (pp. 85–109). University of Chicago Press.
Wilkins, D. P. (2016). Alternative representations of space: Arrernte narratives in sand. In N. Cohn (Ed.), *The visual narrative reader* (pp. 252–281). Bloomsbury.
Willats, J. (2005). *Making sense of children's drawings*. Erlbaum.
Wilson, B. (1988). The artistic tower of babel: Inextricable links between culture and graphic development. In G. W. Hardiman & T. Zernich (Eds.), *Discerning art: Concepts and issues* (pp. 488–506). Stipes.
Wilson, B., & Wilson, M. (1977). An iconoclastic view of the imagery sources in the drawings of young people. *Art Education*, *30*(1), 4–12.

Emoji use validates the potential for meaning standardization among ideographic symbols

Laurie Beth Feldman

Department of Psychology, The University at Albany, State University of New York, Albany, NY, USA
lfeldman@albany.edu

doi:10.1017/S0140525X23000651, e241

Abstract

Technological innovations for online communication reduce the impact of signal transience on meaning standardization while boosting access to reliable patterning across multiple linguistic and nonlinguistic contexts – both asynchronous and synchronous. We classify emojis as ideographic symbols, examine their interdependence with surrounding words when reading/writing, and argue that emoji use validates the potential for meaning standardization in ideographs.

Morin asserts that emojis are not yet ready to qualify as ideographic writing because, as the standardization account would predict, there is a lack of agreement over the meaning of emojis. The interpretation of isolated emojis tends to vary over participants, social groups, language/cultural as well as time and platform (Pei & Cheng, 2022). Thus, the author's cursory treatment of the category of face emojis 🙂,😭,😬 seems to support his claim that meaning standardization for emojis, as for ideographic writing systems in general, poses an insurmountable challenge. Not mentioned, however, is that the category of face emojis is notoriously ambiguous, especially when defined in terms of the variability in one-word descriptions for those emojis in the absence of a context. Face emojis may be least representative of emojis in general, whereas other categories of emojis such as objects, people, activities, and food, tend to engender greater consensus in that verbal labels are mapped more systematically to similar meanings by speakers of a language (Barach, Srinivasan, Fernandes, Feldman, & Shaikh, 2020; Częstochowska et al., 2022). Even more controversial is the author's focus on symbols in isolation when searching for standardization of meaning. When interpreting emojis, analyses often consider not only lexical but also interpersonal, social, cultural as well as legal and other technical conditions of usage (Pei & Cheng, 2022). The implication is that single-facial emojis, despite their typically high-token frequency, are not the most representative symbols on which to base an argument about the expressive power and communicative potential of all emojis and, thus, their potential as ideographic symbols.

It is possible that digital communication may break the chains that keep ideographic communication bound according to Morin. For others, emerging technologies for digital communication with menus of emojis have popularized emoji use and already attest to meaning standardization for emojis as a category of ideographs. Online communication at asynchronous as well as synchronous time scales has become relatively easy to assemble and to analyze as a cumulative pattern (Kaye, Rousaki, Joyner, Barrett, & Orchard, 2022). The author asserts that shared understanding is optimized when communication entails transient symbols and is delivered face-to-face. Neither constraint is characteristic of emoji use. Rather, new transmission technologies ease the challenge of meaning standardization because the adoption of new symbols arises in the service of shared understanding within pairs and communities of interlocutors. Emoji meaning may reveal itself incrementally, based on the volley of messages that comprise an extended online written interaction. New technologies obviate the need that they be transient and conveyed face-to-face.

Newer analytic approaches to meaning move away from the stable representation of meaning and overcome the lack of agreement about how to interpret an isolated symbol by considering usage. To detect changes of a word's senses for example, one can compare words that accompanying it across different genres of text or texts created at different points in time. These quantitative semantic analyses work from the insight that words with similar meanings tend to appear in similar contexts. In fact, not only will words that tend to occur with similar words over many instances but tend to reduce to a similar semantic vector space (Günther, Rinaldi, & Marelli, 2019), but so will emojis (Barbieri, Ronzano, & Saggion, 2016; Wijeratne, Balasuriya, Sheth, & Doran, 2017). The insight is that decontextualized interpretation is no longer the ideal problem space in which to explore meaning standardization and the communicative potential of either words or ideographic symbols –emojis included.

The author acknowledges the "explosion in emoji and emoticon use, concomitant with the rise of digital communication" (target article, sect. 6.4, para. 1) and asserts that a standardization account predicts that that "emojis… and other digital pictographs should become increasingly endowed with precise and shared meanings" (target article, sect. 6.4, para. 1). However, unlike referential communication in experimental contexts when the goal is well defined and referents repeat (Morin, Müller, Morisseau, & Winters, 2022), as emoji use propagates in natural settings, emoji meaning may diffuse from a focus on the object as a whole to its properties (Danesi, 2016). Further, emojis can take on alternative meanings contingent on real world as well as linguistic constraints. Depending on the online community that uses it, 🍆 can connote male genitalia and 🍑 can connote buttocks. Likewise, 🦋 often appears in anorexia-related contexts as well as in general contexts of biological transformation. As a general pattern, opportunities to detect alternative senses of an emoji emerge as communities and contexts of usage become more distinct. The implication is that as usage broadens, emoji meaning may become more nuanced and variable by context.

Finally, the author frames his ambitions for emojis in terms of *replacing writing or complementing it*. Although emojis seldom function as a self-sufficient ideographic system with the potential to fully replace sequences of words, the complementizing function of emojis and words already is prevalent and further, the influence is bidirectional. Experimental psycholinguistic literature demonstrates both parallels between emojis and words as well as ways in which emojis can interact with accompanying words to alter written communication. Sources of evidence are diverse and include measures of vocabulary richness in corpora of online communication with and without emojis among monolinguals as well as among bilinguals. More specifically, lexical diversity in a text tends to be more restricted when emojis are present than in comparable messages when they are absent (Feldman, Aragon, Chen, & Kroll, 2017), and the relation varies more for particular categories of emojis relative to others (Barach et al., 2020). In addition, like words, eye-fixation patterns on emojis vary according to the semantic relation between emoji and accompanying text (Barach, Feldman, & Sheridan, 2021) and recognition memory is greater for emojis that support an inference than a synonym interpretation of text (Christofalos, Feldman, & Sheridan, 2022; Feldman, Christofalos, & Sheridan, 2022). Here, emojis neither function autonomously from words nor do they serve as mnemonic props. Further, effects on written words demonstrate that emoji meaning can be systematic and their use wide-ranging rather than circumscribed to a single domain. Like words, the challenge of meaning standardization for emoji, a class of ideographic symbols, dwindles when examined systematically in the totality of their communicative function.

Acknowledgment. The author thanks Eliza Barach for our discussions.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Barach, E., Feldman, L. B., & Sheridan, H. (2021). Are emojis processed like words? Eye movements reveal the time course of semantic processing for emojified text. *Psychonomics Bulletin and Review*, *28*, 978–991. https://doi.org/10.3758/s13423-020-01864-y

Barach, E., Srinivasan, V., Fernandes, R. B., Feldman, L. B., & Shaikh, S. (2020). *It's not just what you tweet, it's how you tweet it*. Paper presented at the 7th European Conference on Social Media, Larnakes, Cyprus.

Barbieri, F., Ronzano, F., & Saggion, H. (2016). What does this emoji mean? A vector space skip-gram model for Twitter emojis. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pp. 3967–3972, Portorož, Slovenia. European Language Resources Association (ELRA).

Christofalos, A., Feldman, L. B., & Sheridan, H. (2022). Semantic congruency facilitates memory for emojis. The fifth international workshop on emoji understanding and applications in social media co-located with the 2022 Annual conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2022), 63–68. https://aclanthology.org/10.18653/v1/2022.emoji-1.7

Częstochowska, J., Gligorić, K., Peyrard, M., Mentha, Y., Bien, M., Grütter, A., … West, R. (2022). On the context-free ambiguity of emoji. arXiv:2201.06302v2. https://doi.org/10.48550/arXiv.2201.06302

Danesi, M. (2016). *The semiotics of emojis: The rise of visual language in the age of the Internet*. Bloomsbury.

Feldman, L. B., Aragon, C., Chen, N.-C., & Kroll, J. F. (2017). Emoticons in informal text communication: A new window on bilingual alignment. *Bilingualism: Language and Cognition*, *21*, 209–218. https://doi.org/10.1017/S1366728917000359

Feldman, L. B., Christofalos, A., & Sheridan, H. (2022). Memory for emojified text: A comparison of emojis with semantically redundant versus inferential functions. Paper presented at the 62 Meeting of the Psychonomic Society.

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psychological Science*, *14*, 1006–1033. https://doi.org/10.1177/1745691619861372

Kaye, L. K., Rousaki, A., Joyner, L. C., Barrett, L. A. F., & Orchard, L. J. (2022). The online behaviour taxonomy: A conceptual framework to understand behaviour in computer-mediated communication. *Computers in Human Behavior*, *37*, 107443.

Morin, O., Müller, T. F., Morisseau, T., & Winters, J. (2022). Cultural evolution of precise and agreed-upon semantic conventions in a multiplayer gaming app. *Cognitive Science*, *46*, e13113.

Pei, J., & Cheng, L. (2022). Deciphering emoji variation in courts: A social semiotic perspective. *Humanities and Social Sciences Communications*, *9*, 445. https://doi.org/10.1057/s41599-022-01453-5

Wijeratne, S., Balasuriya, L., Sheth, A., & Doran, D. (2017). EmojiNet: An open service and API for emoji sense discovery. *Proceedings of the International AAAI Conference on Web and Social Media*, *11*(1), 437–446. https://doi.org/10.1609/icwsm.v11i1.14857

Why the use of ideographic codes does not improve communicative skills in patients with severe aphasia?

Guido Gainotti

Institute of Neurology, Catholic University of the Sacred Heart, Roma, Italy
guido.gainotti@unicatt.it; WWW.Unicatt.it

doi:10.1017/S0140525X2300081X, e242

Abstract

In his target article, Morin claims that ideographic codes are exceedingly difficult to use. In my commentary I will show that the use of Bliss symbols does not improve the communicative abilities of aphasic patients with severe language disorders. This failure to remediate communication disorders may result from disruption of inner language allowing to translate ideographic codes into spoken language.

In his target article, Morin claims that we cannot find general-purpose codes entirely consisting of images that refer directly to ideas without being translated into spoken language (ideography). He maintains that these codes cannot be used to encode information on a broad range of topics, but only in highly specialized domains and claims that whenever visual codes become generalist, this is because (as in writing) visual images are translated into spoken language.

In the present commentary I will dwell on this issue, drawing on the use of ideography to improve the communicative abilities of patients with severe language disorders. The first attempt in this direction was made by Glass, Gazzaniga, and Premack (1973), who trained global aphasic patients with a pictorial system originally developed by Premack for chimpanzees, and showed that aphasic patients could learn to use this system producing simple pictorial constructions. Similar results were obtained by Johannsen-Hornbach, Cegla, Mager, and Schempp (1985) who treated global aphasics with Blissymbols and observed that some of them could benefit from therapy. Funnell and Allport (1989) showed, however, that in these patients pictorial symbols provided no communicative advantage compared with their processing of written language. Analogous conclusions were reached by McCall, Shelton, Weinrich, and Cox (2000), who showed that, after several years of repeated practice with computerized visual communication, global aphasic patients produced a restricted set of lexical items and of simple syntactic frames, but did not improve their communication.

The failure of pictorial systems to improve the communicative abilities of patients with global aphasia matches the inability shown by ideography to encode information on a broad range of topics and could be due to the fact that pictorial codes cannot replace the inner language that in severely aphasic patients is impaired as much as explicit communicative language.

Investigations exploring the nonverbal cognitive abilities of aphasic patients (recently reviewed by Gainotti, 2021) have, indeed, confirmed Head's (1926) statement that these patients are not impaired on cognitive tasks that can be performed with simple perceptual activities, but are defective when intermediate (explicit or implicit) verbal formulations are required by the task. For instance, Baldo, Paulraj, Curran, and Dronkers (2015) showed a dissociation in performance of aphasic patients between two non-verbal tasks (picture completion vs. picture arrangement tasks of the Wechsler Adult Intelligence Scale) that require differing degrees of inner verbal reasoning. Aphasic patients were, indeed, usually not impaired on the first task, which can be solved with a simple perceptual activity, whereas they were severely impaired on the second task, in which subjects must detect critical differences between pictures to form with them a reasonable and meaningful story.

On the contrary, the reliance of written language on oral language has been confirmed by cognitive neuroscience studies that have shown that structures involved in oral language processing are activated during silent reading. In particular, Perrone-Bertolotti et al. (2012) have shown that the temporal voice areas (TVAs), which are selectively activated by human voice perception, are also activated during silent reading and that this activation is increased when participants are reading attentively

Taken together, all these findings seem to show that ideographic codes cannot be used to encode information on a broad range of topics because they are not converted into spoken language, which is the communication device biologically grounded in the human species. Pictorial codes can, therefore, be used to encode general information only if they are converted into spoken language.

More in general, it could be said that ideographic codes are appealing for cognitive authors who maintain that concepts are represented in the brain in a formal, abstract manner, totally unrelated to the brain processing of sensory-motor functions. They are, on the contrary, much less appealing for authors who prefer a more embodied approach (e.g., Barsalou, 2008) assuming that the organization of semantic representations may reflect the manner in which the information most relevant for their development has been acquired

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Baldo, J. V., Paulraj, S. R., Curran, B. C., & Dronkers, N. F. (2015). Impaired reasoning and problem-solving in individuals with language impairment due to aphasia or language delay. *Frontiers in Psychology*, *6*, 1523. https://doi:10.3389/fpsyg.2015.01523

Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617–645. https://doi.org/10.1146/annurev.psych.59

Funnell, E., & Allport, A. (1989). Symbolically speaking: Communicating with Blissymbols in aphasia. *Aphasiology*, *3*, 279–300. doi:10.1080/02687038908248995

Gainotti, G. (2021). Is there a causal link between the left lateralization of language and other brain asymmetries? A review of data gathered in patients with focal brain lesions. *Brain Sciences*, *11*, 1644. doi: doi.org/10.3390/brainsci11121644

Glass, A. V., Gazzaniga, M. S., & Premack, D. (1973). Artificial language training in global aphasics. *Neuropsychologia*, *11*, 95–110. doi:10.1016/0028-3932(73)90069-9

Head, H. (1926). *Aphasia and kindred disorders of speech*. Macmillan.

Johannsen-Hornbach, H., Cegla, B., Mager, U., & Schempp, B. (1985). Treatment of chronic global aphasia with a nonverbal communication system. *Brain and Language*, *24*, 74–82. doi:10.1016/0093-934x(85)90098-7

McCall, D., Shelton, J. R., Weinrich, M., & Cox, D. (2000). The utility of computerized visual communication for improving natural language in chronic global aphasia: Implications for approaches to treatment in global aphasia, *Aphasiology*, *14*(8), 795–826. doi:10.1080/026870300412214

Perrone-Bertolotti, M., Kujala, J., Vidal, J. R., Hamame, C. M., Ossandon, T., Bertrand, O., … Lachaux, J. P. (2012). How silent is silent reading? Intracerebral evidence for top-down activation of temporal voice areas during reading. *Journal of Neuroscience*, *32*(49), 17554–17562. doi:10.1523/JNEUROSCI.2982-12.2012

Ideography in interaction

Greta Gandolfi and Martin J. Pickering

School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, UK
greta.gandolfi@ed.ac.uk; https://www.ed.ac.uk/profile/greta-gandolfi
martin.pickering@ed.ac.uk; https://www.ed.ac.uk/profile/martin-pickering

doi:10.1017/S0140525X23000572, e243

Abstract

The standardization account predicts short message service (SMS) interactions, allowed by current technology, will support the use and conventionalization of ideographs. Relying on psycholinguistic theories of dialogue, we argue that ideographs (such as emoji) can be used by interlocutors in SMS interactions, so that the main contributor can use them to accompany language and the addressee can use them as stand-alone feedback.

Dialogues are joint activities (Clark, 1996; Pickering & Garrod, 2021), where interlocutors act in coordination, constantly switching their conversational roles from speaker to addressee and providing each other with feedback, which ultimately supports mutual comprehension. In SMS dialogues, such feedback can involve ideographs – smiling faces, thumbs-up, question marks, cartoon airplanes, or actual photographs or gifs. Here we expand on the standardization account predictions for the use of ideography in the age of digital communication (target article, sect. 6.4), by focusing on the role interlocutors take when using it in dialogues.

In face-to-face dialogues, interlocutors produce cospeech gestures, including head movements and hand and facial gestures (see Bavelas, 2022). In Clark's terms (1996, p. 155; following Peirce, 1932), interlocutors produce linguistic or nonlinguistic signals (e.g., waving at someone or saying "*Hello*") or a combination of both (e.g., waving at someone, while saying "*Hello*"). Signals can include signs that relate to objects by convention (symbols, e.g., nodding to accept a request or using the word "*dog*" to refer to a specific animal), by resembling the objects they refer to (icons, e.g., drawing a notebook in the air while saying "*Can you hand me the notebook?*"), or by having a causal relationship with them (indexes, e.g., demonstrating writing while saying "*Can you hand me the notebook?*").

In SMS dialogues, similarly to iconic gestures, ideographs can be embedded in main contributors' utterances (e.g., "*John was caught cheating and looked* 😳" – i.e., meaning John was embarrassed), or they can add information to a complete utterance (e.g., "*John was caught cheating and looked embarrassed* 😒" – i.e., meaning the writer is annoyed with John's action and attitude). In both examples, the emoji do not make sense without the linguistic context – the emoji are dependent on the words in the SMS. In the first example, the text is incomplete without the emoji, whereas in the second example the text is complete, and the emoji adds an additional message (see Grosz, Kaiser, & Pierini, 2021, for analysis).

Alternatively, ideographs can function in the same way as symbolic gestures, acting as substitutes for atomic words (i.e., words that need not be embedded in syntactic structures to form full utterances, such as "*Yes*" or "*Hello*"; Clark, 1996, p. 163).

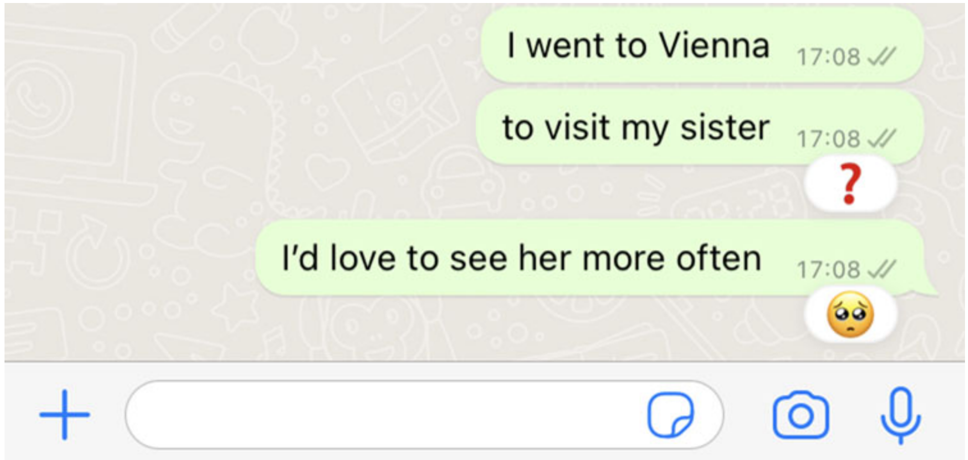

Figure 1 (Gandolfi and Pickering). Example of reactions in SMS dialogue.

So, the word "*Yes*" can be replaced by an actual thumbs-up in face-to-face communication or by a thumbs-up emoji in SMS. And like such gestures, addressees can use ideographs to provide complete feedback, indicating understanding or failure to understand, and regulating turn taking (Sacks, Schegloff, & Jefferson, 1974).

In face-to-face dialogues, addressees often indicate understanding (or alignment) by producing brief linguistic feedback (e.g., "*Yeah*") or nonlinguistic feedback (e.g., nodding) and misunderstanding by producing equally brief contributions (e.g., "*Eh?*"), specific or generic questions (e.g., "*Which one?*" or "*Who?*"), or movements (e.g., frowning). Moreover, addressees are often good at indicating where the problem is (Dingemanse et al., 2015): If a person with two sisters, one with blond and one with dark hair, says "*I visited my sister in Vienna*," their addressee might ask "*The blond one?*" rather than "*Who?*," to indicate that the locus of difficulty is the identity of the referent. Critically, addressees need not interrupt main contributors to provide feedback: They can use backchannels (Yngve, 1970), which have been shown to influence the quality, shape, and timing of speakers' utterances (Bavelas, Coates, & Johnson, 2000; Tolins & Fox Tree, 2014). Speakers use the feedback provided by their addressees to adjust their plans, to retain or vacate the floor to their interlocutor, and to repair their contributions on the fly, as the dialogue unfolds (Pickering & Garrod, 2021, Ch. 5).

Historically, the type of feedback that printed communication (e.g., letters, emails, post-its) allowed was produced slowly or limited to typewritten characters. However, current platforms for instant messaging (WhatsApp, Telegram, Slack, Teams, etc.) support flexible and highly informative forms of feedback that approximate to what happens in face-to-face interactions. Not only can addressees reply to main contributors' turns almost instantly in a subsequent turn (often called replies), but they can also comment on distinct turns by using just ideographs (often called reactions). As backchannels, reactions allow addressees to indicate the source of the problem, without interrupting the main contributor (see Fig. 1 for an example: While the first emoji indicates misunderstanding, the second might express empathy and neither of those interrupts the main contributor).

Ideographs can thus play the same role as face-to-face feedback: First, they can provide generic or specific evidence of understanding. Writers can use conventional positive or negative ideographs (e.g., 👌, ✅ or ❓, 👎) to provide generic feedback or combine ideographs (e.g., using 💁 and ❓ instead of typing "*The blond one?*") to provide specific feedback, which indicates exactly what needs to be repaired (see some examples of conventionalized pairs of emoji in Gawne & McCulloch, 2019). Second, emoji can regulate turn taking by allowing addressees to provide feedback without interrupting the main contributors. Interestingly, in group chats, attendees either can show agreement by texting individual contributions or can just wait for one of the addressees to add a reaction (e.g., 👍), and just click on it, which results in a sum of all the clicks (e.g., 👍$_5$ for five clicks).

In summary, ideographs are used flexibly in SMS: When used by the main contributors, they can be incorporated into the utterance or can serve as an additional but dependent contribution, but they cannot replace the main linguistic contribution. When used by the addressees, they can serve as a complete contribution in their own right, and function as feedback to the main contribution. Analysing SMS interactions from the perspective of dialogue, we support the standardization account in predicting an evolution of ideographic language, but only when such language comprises feedback, and not main contributions. This might explain why the ideographs that are usually associated with – more generic – feedback seem to be the ones with established conventional meaning (e.g., 👍 or 👏).

Financial support. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement (grant no. 859588).

Competing interest. None.

References

Bavelas, J. B. (2022). *Face-to-face dialogue: Theory, research, and applications*. Oxford University Press.

Bavelas, J. B., Coates, L., & Johnson, T. (2000). Listeners as co-narrators. *Journal of Personality and Social Psychology*, *79*(6), 941–952. https://doi.org/10.1037/0022-3514.79.6.941

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Dingemanse, M., Roberts, S. G., Baranova, J., Blythe, J., Drew, P., Floyd, S., … Enfield, N. J. (2015). Universal principles in the repair of communication problems. *PLoS ONE*, *10*(9), e0136100. https://doi.org/10.1371/journal.pone.0136100

Gawne, L., & McCulloch, G. (2019). Emoji as digital gestures. *Language@ Internet*, *17*, article 2 (urn:nbn:de:0009-7-48882). https://www.languageatinternet.org/articles/2019/gawne/index_html.

Grosz, P., Kaiser, E., & Pierini, F. (2021). Discourse anaphoricity and first-person indexicality in emoji resolution. *Proceedings of Sinn Und Bedeutung*, *25*, 340–357. https://doi.org/10.18148/sub/2021.v25i0.941

Peirce, C. S. (1932). *The collected papers of Charles Sanders Peirce, vol. II: Elements of logic* (C. Hartshorne & P. Weiss, Eds.). Harvard University Press.

Pickering, M. J., & Garrod, S. (2021). *Understanding dialogue: Language use and social interaction*. Cambridge University Press.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*(4), 696. https://doi.org/10.2307/412243

Tolins, J., & Fox Tree, J. E. (2014). Addressee backchannels steer narrative development. *Journal of Pragmatics*, *70*, 152–164. https://doi.org/10.1016/j.pragma.2014.06.006

Yngve, V. H. (1970). On getting a word in edgewise. *Chicago Linguistics Society, 6th meeting, 1970*, pp. 567–578.

A bigger problem for ideography: The pervasiveness of linguistic structure

Daniel Harbour

Department of Linguistics (SLLF), Queen Mary University of London, London, UK
d.harbour@qmul.ac.uk;
www.qmul.ac.uk/sllf/linguistics/people/academic/profiles/harbour.html

doi:10.1017/S0140525X23000638, e244

Abstract

Writing systems display ubiquitous linguistic structure, from the recursive syntactic properties of their glyphs to the morphology/phonology of their combinatorics. This extends to Ancient Egyptian, Chinese, and Sumerian ideograms. Pure ideography requires switching this influence off. The pervasive linguistic tinge to the fabric of writing systems suggests that the chances of breaking what Morin terms language's lock-in effect are slim.

For a self-sufficient, generalist ideography to arise, language must keep its hands to itself. The evolution of writing admits little chance of this. When writing systems diverge from the languages they are used to represent, as an ideography must, they frequently converge on properties of other natural languages. This suggests that language is not only the source of writing systems' self-sufficiency and generality, as Morin contends, but also that linguistic forces guide the development of such systems. Real though Morin's problem of standardisation is, the key challenge facing language-independent ideography – manifest in the evolution of writing and its comparative study – is the pervasiveness of language itself.

A defining property of syntax is recursion, that outputs can also be inputs (Berwick & Chomsky, 2017). Recursion is similarly key to ideograms, and other sign types. Signs that depict or suggest what they denote are a common trope in early-stage writing. These signs typically build on one another. Sumerian uses water in an enclosure to suggest "marsh"; Chinese uses the sun behind a tree to suggest "east" (Fig. 1). An ideography in which every sign differs substantively from every other is conceivable but not how ideograms in early scripts work.

Recursion in ideography (and other logograms) is reminiscent of natural language both in degree and kind. Deep recursion is an overriding characteristic of Chinese signs, revealing hallmarks of natural language. Hierarchy matters in recursion. "Unlockable"

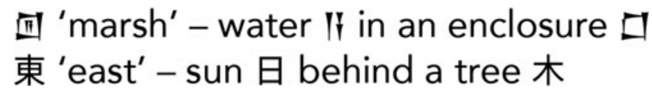

Figure 1 (Harbour). Examples of ideography (Sumerian, Chinese).

is ambiguous between *un*-[*lockable*] "not capable of being locked" and [*unlock*]-*able* "capable of not being locked." The same morphemes differently combined have different meanings. Equally, different combinations of a sign set are interpretatively distinct. "Tree," "tree," and "gate" combine into hierarchically distinct *gate*-[*tree-tree*] "learn" and *tree*-[*gate-tree*], a tree species (Fig. 2). These two signs are homophones, hence only semi-ideographic, but more purely ideographic examples exist.

The Sumerian sign for "waterskin" comprises "leather," "water," "steppe," and "hang" (Fig. 3). There is no phonetic relationship between *ummud* ("waterskin") and its constituents, *kuš*, *a*, *eden*, *la*. This is ideography: Waterskins as leather hangers for water on steppes (Selz, 2017). Rich in compounds, Sumerian had the left-headed noun–noun and right-headed agentive noun–verb compounds required for *leather*-[[*water-steppe*]-*hanger*]. Yet it used internally inflected phrases, not compounds, for items of this complexity (Jagersma, 2010). Written, "waterskin" thus diverges from spoken Sumerian both in being a compound and in its degree of recursion. Deeply recursive compounding is found in natural languages, though (famously, Sanskrit; Lowe, 2015). Ideograms like "waterskin" thus extrapolate native constructions to converge on a foreign language's grammatical norms.

Ancient Egyptian ideography strikingly illustrates how grammar fills the gap that arises when writing diverges from the speech it was devised to record. In hieroglyphs, Egyptian nouns were typically followed by "determinatives." These unpronounced glyphs categorised the preceding noun semantically. In *sꜣt* "daughter" (Fig. 4), the phonetic glyphs *sꜣ* and *t* are followed by "woman," classifying the referent as female. Goldwasser and Grinevald (2012) observe that noun phrases in the Mayan language Jacaltec have this same structure: *ix q'opoj* "a/the girl" consist of *q'opoj* "girl" and *ix* "woman," which also accompanies other human females. Written Egyptian has thus evolved to resemble a spoken language with noun classifiers though Egyptian was itself not a classifier language. (Jacaltec classifiers are prenominal and Egyptian determinatives postnominal, a point on which classifier languages vary; Aikhenvald, 2000.)

A syntax-centric view of sign formation has consequences. The dominant model of grammatical interaction in theoretical linguistics places syntax before morphology–phonology. The syntactic module builds structures absent of sound, and sound is supplied later in the morphological–phonological module. Despite their different modalities, spoken and signed languages have been shown to have analogous morphological and phonological properties (Brentari, Fenlon, & Cormier, 2018). If sign inventories are constructed by syntactic recursion and then externalised via a hand-held implement, then the externalisation process is potentially subject to similar morphology- and phonology-like forces.

The most thoroughgoing explorations of parallels between natural language grammar and writing system structure are Myers (2019) and Meletis (2020); the latter, a typological survey, the former, a study of Chinese characters. Both map extensive correlations between grammatical and graphic systems. For reasons of space, though, the next paragraphs offer briefer illustrations of, respectively, phonological and morphological parallels with writing systems.

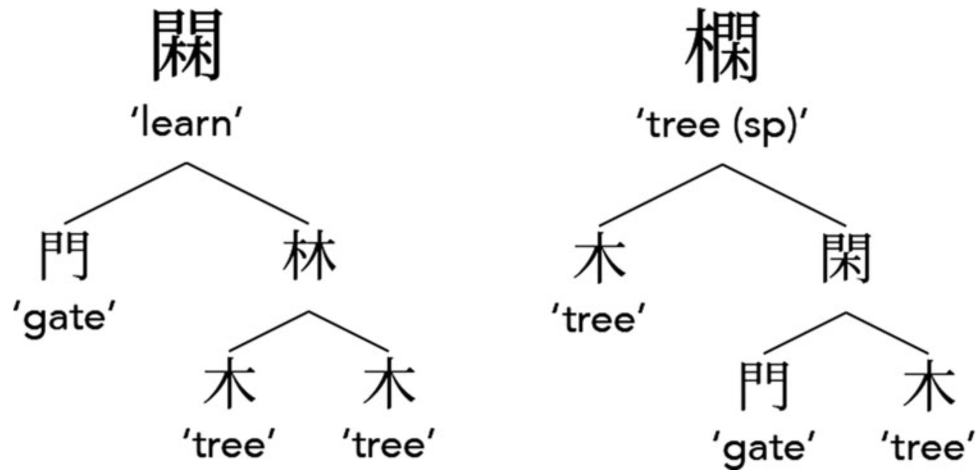

Figure 2 (Harbour). Different hierarchical arrangements of the same inputs (Chinese).

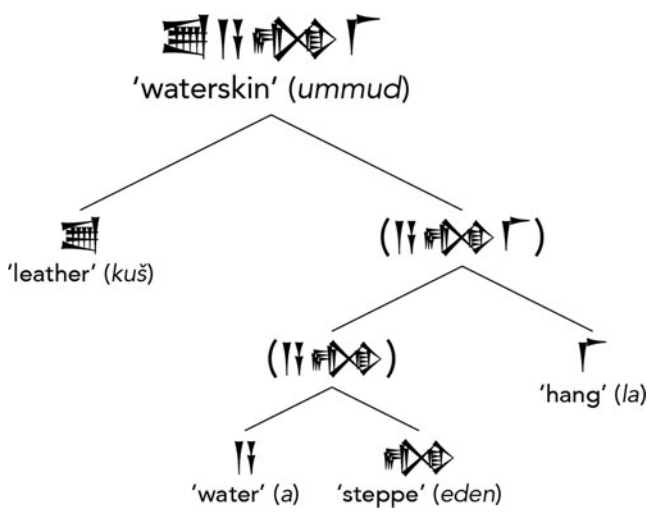

Figure 3 (Harbour). Compounding in ideography (Sumerian).

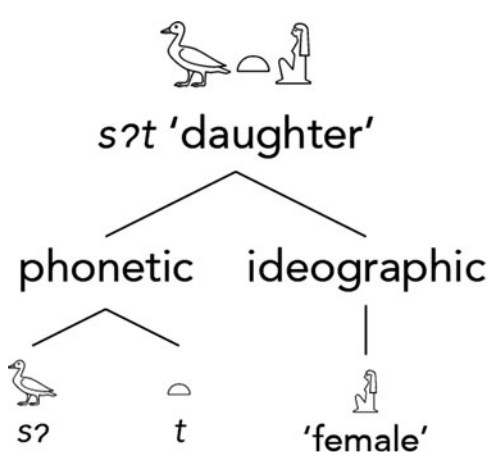

Figure 4 (Harbour). Classification by ideography (Ancient Egyptian).

Many languages have phonological minimal word constraints. These often apply to lexical but not functional vocabulary. From the irregular morass of English orthography, a minimality constraint has emerged at the graphic level (McCawley, 1994). In two- and three-letter homophone pairs, the shorter form is reserved for function words (*by*, *or*, *we*), the longer for lexical ones (*bye*, *ore*, *wee*). A different example of convergence on phonological norms is found in segmental order written in syllabic clusters, like Hangul (Korea) and Thaana (used to write Dhivehi; Maldives). These orthographic syllables obey rules familiar from syllabic structures in spoken languages but not always those of the spoken Korean and Dhivehi that they are used to write (Gnanadesikan, in press).

Afroasiatic writing systems are renowned for frequent vowel omission. Given that many Afroasiatic grammatical properties are encoded vocalically, this omission simulates morphological impoverishment. The contrasts thus neutralised are ones that other languages simply lack, so writing divergent from speech again converges with other languages (Harbour, 2021). More broadly, the complexities of segment order in Brahmic scripts display many contrasts familiar from morphological theory, such as affixality, allomorphy, headedness, and affixal levels, whereas orientation in Canadian Aboriginal Syllabics appears analogous to "ion-morphs" in American Sign Language (Gnanadesikan, 2022).

Such morphological and phonological parallels with natural languages are expected if sign inventories are constructed by the syntactic module and externalised via pathways of spoken and signed languages. Hockett (1960) famously observed that the coupling of meaning with the inherently meaningless (external gestures) is key to language. If sign systems exploit the generative infrastructure of language, for how long could a self-sufficient, generalist ideography keep both inherently meaningless, hence nonideographic, elements and the languages of its would-be users at bay?

Acknowledgments. The author thanks David Adger, Amalia Gnanadesikan, Yukiko Miller, and James Myers.

Competing interest. None.

References

Aikhenvald, A. (2000). *Classifiers: A typology of noun categorization devices*. Oxford University Press.
Berwick, R. C., & Chomsky, N. (2017). *Why only us: Language and evolution*. MIT Press.
Brentari, D., Fenlon, J., & Cormier, K. (2018). *Sign language phonology. Oxford research encyclopaedia of linguistics*. Oxford University Press. Retrieved from https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-117

Gnanadesikan, A. (2022). *Amodal morphology: Applications to Brahmic scripts and Canadian Aboriginal syllabics*. University of Maryland.
Gnanadesikan, A. (in press). Segments and syllables in Thaana and Hangeul: A comparison of literate native-speaker inventions. *Written Language and Literacy*.
Goldwasser, O., & Grinevald, C. (2012). What are determinatives good for? In E. Grossman, S. Polis, & J. Winand (Eds.), *Lexical semantics in Ancient Egyptian* (pp. 17–53). Widmaier.
Harbour, D. (2021). Grammar drives writing system evolution: Lessons from the birth of vowels. In Y. Haralambous (Ed.), *Grapholinguistics in the 21st century 2020* (pp. 201–221). Fluxus Editions.
Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*, 88–97.
Jagersma, A. H. (2010). A descriptive grammar of Sumerian (PhD thesis). Leiden University.
Lowe, J. J. (2015). The syntax of Sanskrit compounds. *Historical Syntax*, *91*, e71–e115.
McCawley, J. D. (1994). Some graphotactic constraints. In W. C. Watt (Ed.), *Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics* (pp. 115–127). Springer.
Meletis, D. (2020). *The nature of writing: A theory of grapholinguistics*. Fluxus Editions.
Myers, J. (2019). *The grammar of Chinese characters: Productive knowledge of formal patterns in an orthographic system*. Routledge.
Selz, G. J. (2017). To carry coals to Newcastle or observations concerning sign formation as an early concept in Mesopotamian "commentaries." In L. Feliu, F. Karahashi, & G. Rubio (Eds.), *The first ninety years: A Sumerian celebration in honor of Miguel Civil* (pp. 300–310). De Gruyter.

Bypass language *en route* to meaning at your peril

Lindsay N. Harris[a], Charles A. Perfetti[b] and Elizabeth A. Hirshorn[c]

[a]Department of Leadership, Educational Psychology, and Foundations, Center for the Interdisciplinary Study of Language and Literacy, Northern Illinois University, DeKalb, IL, USA; [b]Learning Research and Development Center, Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA and [c]Department of Psychology, State University of New York at New Paltz, New Paltz, NY, USA
lharris3@niu.edu
perfetti@pitt.edu
hirshore@newpaltz.edu
https://www.cedu.niu.edu/lepf/about/faculty/Lindsay-Harris.shtml
https://www.lrdc.pitt.edu/people/researcher-detail.cshtml?id=308
https://webapps.newpaltz.edu/directory/profile/elizabethhirshorn

doi:10.1017/S0140525X23000729, e245

Abstract

The learning account of the puzzle of ideography cannot be dismissed as readily as Morin maintains, and is compatible with the standardization account. The reading difficulties of deaf and dyslexic individuals, who cannot easily form connections between written letter strings and spoken words, suggest limits to our ability to bypass speech and reliably access meaning directly from graphic symbols.

Morin presents an interesting, novel argument about the lack of writing systems based on ideographs rather than language. Much of the argument, especially regarding the role of standardization in supporting systematicity, seems both correct and not incompatible with the dominance of language-based writing. Morin proposes two explanations for why there has not been a self-sufficient, systematic graphic code: The learning account and the standardization account. He dismisses the first and champions the second. But the learning account – the notion that humans are not equipped to learn large numbers of graphic symbol–meaning pairings without relying on language – cannot be dismissed so easily and is, in fact, compatible with a role for standardization.

In his target article, Morin notes that there are limits to glottography, such that the linguistic information an orthography encodes is not limited to phonological information. (It is for this reason that we refer to a "language constraint" and not "glottography" in identifying constraints on written language; Perfetti, 2003; Perfetti & Harris, 2013.) But the converse is evidently true: There are limits to our ability to bypass speech and reliably access meaning directly from graphic symbols. The strongest evidence for this learning account exists in the form of the millions of individuals worldwide who, because of biological constraints imposed by deafness or dyslexia, cannot easily form connections between written letter strings and spoken words. For these individuals, the most expeditious route to literacy, were it a viable one, would be to bypass language entirely in forming a reading network. Processing letter strings as unitary graphic symbols that map directly to meaning – that is, as ideographs – would unlock written text for clinical populations who otherwise struggle to access it.

This does not happen. Despite enormous social and economic pressures and abundant opportunity, deaf and dyslexic individuals have tremendous difficulty reading. The average deaf student finishes secondary education reading at the level of a primary-aged child (Lederberg, Schick, & Spencer, 2013), and individuals diagnosed with dyslexia in childhood continue to display poor reading and writing skills throughout life (Reis, Araújo, Morais, & Faísca, 2020). In fact, a preponderance of evidence indicates that whole-word reading instruction – a method that essentially encourages children to treat writing as an ideography – provides fertile ground for dyslexia to flourish (Perfetti & Harris, 2019).

In his description of the limits of glottography, Morin cites a recent article in which we review evidence that some percentage of deaf and dyslexic individuals become proficient readers by deemphasizing phonology in the formation of their neural reading networks (Hirshorn & Harris, 2022). Although this is true, the subset of deaf and dyslexic individuals who manage to forge these atypical pathways is distressingly small, and neuroimaging evidence has yet to reveal individuals who omit phonological areas entirely from the reading network. Moreover, we proposed in our review article that those deaf and dyslexic individuals who attain literacy by constructing phonology-deemphasized networks may represent a minority of the larger population with exceptional visual memory capacity (Hirshorn & Harris, 2022). The probability that humans would have stumbled upon or cultivated a graphic symbol–meaning network that excludes language entirely, in the absence of the current pressures that compel predisposed members of clinical populations to develop a limited version of them, therefore seems low, even if we did not have language to parasitically host writing.

What to make, then, of Morin's rejection of the learning account? Curiously, his dismissal of it is premised on his dismissal of the motor theory-of-speech perception, which, among other things, purports to account for children's learning to speak more easily than their learning to read. The reader may or may not find Morin's criticisms of motor theory persuasive – if not, there plenty of other criticisms of it on offer (see, e.g., Hickok, 2014). In any case, it is not clear how the viability of the learning account is tethered to the viability of motor theory. Rejection of motor theory's claim that humans have an innate sensitivity to speech segments, or mentally represent speech as articulatory gestures, does not entail acceptance of the claim that we can easily learn large numbers of graphic symbol–meaning pairings without parasitizing neural speech networks. Morin's argument against the learning account relies heavily on rejecting the assumption that the human mind is "ill-equipped" (target article, sect. 5.5, para. 1) to memorize large numbers of form–meaning pairings. On Morin's argument, this assumption is undermined because it "wrongly predicts that full-blown sign languages cannot evolve" (target article, sect. 5.5, para. 1). This, too, seems to be a non-sequitur. Instead, sign languages can evolve but are not able to replace language-based writing. They find their evolutionary niche without crowding out the dominant species.

We do agree that standardization is a factor in developing writing systems. Morin's argument is that "spoken or signed codes, being easier to standardize, install a lock-in situation where other types of codes are less likely to evolve" (target article, sect. 6.1, para. 6). The question is, *why* are they easier to standardize? Language-based writing, whether the graphs used map to speech segments, whole syllables, morphemes, or whole words, provides standard graph–language pairs of cognitively manageable numbers that generate the infinity of messages with relatively small means.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Hickok, G. (2014). *The myth of mirror neurons: The real neuroscience of communication and cognition*. WW Norton.

Hirshorn, E. A., & Harris, L. N. (2022). Culture is not destiny, for reading: Highlighting variable routes to literacy within writing systems. *Annals of the New York Academy of Sciences*, *1513*(1), 31–47. https://doi.org/10.1111/nyas.14768

Lederberg, A. R., Schick, B., & Spencer, P. E. (2013). Language and literacy development of deaf and hard-of-hearing children: Successes and challenges. *Developmental Psychology*, *49*(1), 15. https://doi.org/10.1037/a0029558

Perfetti, C. A. (2003). The universal grammar of reading. *Scientific Studies of Reading*, *7*(1), 3–24. https://doi.org/10.1207/S1532799XSSR0701_02

Perfetti, C. A., & Harris, L. N. (2013). Universal reading processes are modulated by language and writing system. *Language Learning and Development*, *9*(4), 296–316. https://doi.org/10.1080/15475441.2013.813828

Perfetti, C. A., & Harris, L. N. (2019). Developmental dyslexia in English. In L. Verhoeven, C. A. Perfetti, & K. Pugh (Eds.), *Developmental dyslexia across languages and writing systems* (pp. 25–49). Cambridge University Press. https://doi.org/10.1017/9781108553377.002

Reis, A., Araújo, S., Morais, I. S., & Faísca, L. (2020). Reading and reading-related skills in adults with dyslexia from different orthographic systems: A review and meta-analysis. *Annals of Dyslexia*, *70*(3), 339–368. https://doi.org/10.1007/s11881-020-00205-x

Ideography, Blissymbolics, standardization, and emergent conformity

Harry Howard

Department of Spanish & Portuguese, Tulane University, New Orleans, LA, USA
howard@tulane.edu; https://www2.tulane.edu/~h0Ward/

doi:10.1017/S0140525X23000626, e246

Abstract

There is an extensive literature on the usage of Blissymbolics in augmentative and alternative communication that contradicts Morin's contention that it fails as an ideography. Morin's notion of "standardization" (target article, sect. X, para. X) is at odds with the highly developed understanding of this notion in linguistics. What Morin seems to have in mind corresponds to the notion of emergence in iterative and multiagent models of language learning.

Morin claims that Blissymbolics are unworkable:

> Attempted generalist ideographies, like Bliss symbolics, struggle to express as wide as broad a variety of meaning as language does, in part because of the large number of conventional symbols that one would need to learn in order to make the system work, and in part because the rules that are supposed to help compose complex expressions from simpler symbols are too ambiguous. (target article, sect. 5.3, para. 4)

This claim appears to rest on the "evidence" from Morin, Kelly, and Winters (2020):

> At the opposite end of the spectrum, one finds graphic codes like Bliss symbols (Okrent, 2010), which were invented by Charles Bliss in an attempt to create an international writing system that would allow communication between different linguistic communities. The system consists of several hundred symbols for basic concepts that can be combined to create novel or more complex expressions. Although, in principle, Bliss symbols were designed to be compositional, in practice, their users would have needed to learn as many 900 individual ideographs, along with their conventional rules of combination. No one (not even the system's inventor) really became fluent in their use.

However, this single observation is contradicted in the literature on augmentative and alternative communication. Even a cursory search through Google Scholar reveals positive results from the following reports of attempts to teach Bliss symbols, or the more recent coinage, blissymbols: Beck and Fritz (1998); Burroughs, Albritton, Eaton, and Montague (1990); Clark (1981); Ecklund and Reichle (1987); Funnell and Allport (1989); Hurlbut et al. (1982); Mizuko (1987); and Poupart, Trudeau, and Sutton (2013). Muter (1986) conducted a review with many more instances of positive results. And this is surely an undercount, because many abstracts do not indicate the name of the ideography that is used in the study. Thus it seems premature to dismiss Blissymbolics at the current stage of research.

Morin refers to the back-and-forth refinement of reference of a particular ideogram as "standardization." This is not the accepted usage of the term in linguistics.

To refer to a readily available overview, Standard Language (2023) cites (Van Mol, 2003): "A standard variety can be conceptualized in two ways: (i) as the sociolect of a given socio-economic stratum or (ii) as the normative codification of a dialect, an idealized abstraction."

A popular model of the latter was set forth in Haugen (1966a, 1966b), in which standardization advances in four phases. The first is norm selection, a process of choosing between competing varieties. Once a variety is settled on, it is codified by elaborating a range of reference works, such as dictionaries, grammars, spelling manuals, and style guides. This new code is "implemented" when a society accepts and transmits among speakers. Finally, the norm can be elaborated further to extend it to new functions or changing sociopolitical conditions.

In Haugen's model, Blissymbolics is most certainly standardized. Blissymbols (2023) explains that Blissymbolics Communication International published a reference guide (Wood, Storr, & Reich, 1992), containing 2,300 vocabulary items and detailed rules for the graphic design of additional characters. Thus Morin's terminology leads to considerable confusion, if not contradiction, with the facts about an existing ideography.

To be charitable, we can imagine that Morin would maintain that the implementational phase of Haugen's standardization process has never been achieved for Blissymbolics. We concede that the studies cited above are all short term and performed on small subject populations. This is why we maintain that the purported unviability of all ideographies is still up for empirical confirmation.

To our way of thinking, what Morin seems to envision as "standardization" is a kind of bottom-up, "emergent conformity" (our coinage) that arises in iterative and multiagent models of language learning, see Kirby, Griffiths, and Smith (2014) for a recent review. The literature in this paradigm tends to use the term "emergence" to label how the final state of the algorithmic process results from its initial state. In our reading, this notion of emergence is not discussed on its own merits, because it is held to be something that dynamical systems just do. What is clear, however, is that the algorithmic process is NOT labeled "standardization."

If this terminological impasse can be resolved, we believe that Morin's analysis enters into a fruitful conversation both with iterative and multiagent modeling and with linguistic standardization. The reason is that the obstacles that Morin posits that an ideography would have to overcome to obtain wide currency (cheapness, transiency, common ground, repair, interactive alignment, lock-in, frequency-dependent advantage) are presupposed in the algorithmic architecture and deserve to be disentangled and examined on their own. In a similar vein, these factors can also be wielded to decompose Haugen's implementational phase of standardization, which, much like in the computational modeling literature, is treated as an unanalyzed black box.

Financial support. I received no financial support for the research, authorship, and/or publication of this article.

Competing interest. None.

References

Adams, F. (2020). Global aphasia and the language of thought. *Theoria. Revista de Teoría, Historia y Fundamentos de la Ciencia*, *35*(1), 9–27.

Beck, A., & Fritz, H. (1998). Can people who have aphasia learn iconic codes. *Augmentative and Alternative Communication*, *14*(3), 184–196.

Blissymbols. (2023). *Wikipedia*. https://en.wikipedia.org/wiki/Blissymbols

Burroughs, J. A., Albritton, E., Eaton, B., & Montague, J. (1990). A comparative study of language delayed preschool children's ability to recall symbols from two symbol systems. *Augmentative and Alternative Communication*, *6*(3), 202–206.

Carmeli, S., & Shen, Y. (1998). Semantic transparency and translucency in compound Blissymbols. *Augmentative and Alternative Communication*, *14*(3), 171.

Clark, C. R. (1981). Learning words using traditional orthography and the symbols of Rebus, Bliss, and Carrier. *Journal of Speech and Hearing Disorders*, *46*(2), 191–196.

Ecklund, S., & Reichle, J. (1987). A comparison of normal children's ability to recall symbols from two logographic systems. *Language, Speech, and Hearing Services in Schools*, *18*(1), 34–40.

Funnell, E., & Allport, A. (1989). Symbolically speaking: Communicating with Blissymbols in aphasia. *Aphasiology*, *3*, 279–300.

Haugen, E. (1966a). Dialect, language nation. *American Anthropologist*, *68*(6), 922–935.

Haugen, E. (1966b). *Language conflict and language planning: The case of modern Norwegian*. Harvard University Press.

Kirby, S., Griffiths, T., & Smith, K. (2014). Iterated learning and the evolution of language. *Current Opinion in Neurobiology*, *28*, 108–114.

Morin, O., Kelly, P., & Winters, J. (2020). Writing, graphic codes, and asynchronous communication. *Topics in Cognitive Science*, *12*(2), 727–743.

Muter, P. (1986). Blissymbolics, cognition, and the handicapped (9, Trans.). In E. Hjelmquist & L.-G. Nilsson (Eds.), *Communication and handicap: Aspects of psychological compensation and technical aid*.

Okrent, A. (2010). *In the land of invented languages: Esperanto rock stars, Klingon poets, Loglan lovers, and the mad dreamers who tried to build a perfect language*. Random House.

Poupart, A., Trudeau, N., & Sutton, A. (2013). Construction of graphic symbol sequences by preschool-aged children: Learning, training, and maintenance. *Applied Psycholinguistics*, *34*(1), 91–109.

Standard Languages. (2023). *Wikipedia*. https://en.wikipedia.org/wiki/Standard_language

Van Mol, M. (2003). *Variation in modern standard Arabic in radio news broadcasts: A synchronic descriptive investigation into the use of complementary particles*. Peeters.

Wood, C., Storr, J., & Reich, P. (1992). *Blissymbol reference guide*. Blissymbolics Communication International.

The centrality of practice in ideographic communication, and the perennial puzzle of positivistic thinking

Lucas B. Mazur[a,b] and Sandra Plontke[c]

[a]Jagiellonian University, Kraków, Poland; [b]Sigmund Freud University, Berlin, Germany and [c]Ruhr-Universität Bochum, Bochum, Germany
lucas.mazur@uj.edu.pl
sandra.plontke@rub.de

doi:10.1017/S0140525X23000687, e247

Abstract

To the extent that we expect ideographs to be closer to the reality they depict than spoken or written words we are succumbing to the perennial allure of positivistic thinking. Morin powerfully argues that human communication, including ideography, cannot be understood apart from practice, thus removing the positivistic assumption that made the "puzzle of ideography" puzzling in the first place.

Morin convincingly argues for the important role played by the social practice he calls "standardization" in the evolution and stabilization of language. It is not that any particular symbolic system is inherently better or worse as the foundation of human communication, but rather that certain systems, within certain contexts, may more readily cultivate online repairing of communication. This argument appears to help solve the "puzzle of ideography." However, by underscoring the centrality of practice within communication it actually removes the expectation that particular symbolic systems can be inherently better than others, thereby undercutting the foundational assumption upon which the matter was puzzling in the first place, namely, that ideographs are closer to the reality they depict than spoken or written words. To the extent that we continue to find the "puzzle of ideography" puzzling, we are in effect reminded of the continued allure of positivistic thinking regarding human psychology.

The "puzzle of ideography" flourished in the nineteenth and early twentieth centuries, a time uniquely colored by positivist and even utopian thinking. It was hoped that we would come to see and capture the world and our place therein *as it is in reality*; to the degree that we would become truly "objectively scientific." The effects of this current of thought on the development of the natural sciences, the social sciences, and even the humanities have been well-documented, and have been collectively presented as involving a set of "naturalistic" epistemological assumptions (e.g., Daston & Galison, 2010). Some hoped that the advance of science would free us from the shackles of intergroup differences and divisions, and unite us under the shared banner of a common humanity. Various attempts to develop ideographic languages arose from this spirit (e.g., Bliss language, Neurath's international picture language). To the extent that we believe ideograms to be somehow closer to the reality they symbolically depict than their linguistic counterparts, and to somehow retain that meaning apart from a community of practitioners, we are cultivating a naturalistic epistemology that assumes ideograms to be "brute facts" (to borrow the language of Taylor, 1985) that can exist outside of social interaction.

The assumption that ideographs depict things in the world somehow more directly, or immediately, than do written or spoken words is an echo of the broader, long-standing hope that – if used correctly – symbolic systems, including language, can directly capture reality. This hope blossomed in the Enlightenment and flourished in the nineteenth century. As our faith in the universal language of science increased, the use of the universal language of Latin declined and the use of local languages in science spread; a widely used language practice was replaced with the widely held expectation that we would come to know, and to represent, the world as it is in reality, unadulterated by cultural variation. This expectation of direct contact with objective, universal reality would come to take various forms over the centuries – for example, as famously seen in the hopes of the logical positivists. Each new incarnation of this expectation would be met with arguments for its folly – as in Wittgenstein's famous rejection of logical positivism. Wittgenstein's private language argument attests to the inherently social, praxiological aspect of language, and rejects the idea that symbols are connected with nonsymbolic reality, or even have symbolic meaning, outside that social practice. Thus, as naturalistic schools of thought spread, so too have emerged schools of thought that can be described as falling under the umbrella term "interpretive social science," which do not share these naturalistic assumptions. These approaches understand the symbolic dimensions of human life as requiring different methods of study than those found in the natural sciences. For example, in defining cultural psychology as a form of interpretive psychology, Bruner (1990, p. 118) argued it to be the study of "the rules that human beings bring to bear in creating meanings in cultural contexts. These contexts are always contexts of practice: it is always necessary to ask what people are doing or trying to do in that context."

The expectation that we can directly capture reality within symbolic depictions thereof is nevertheless very much alive in our current, technological age – perhaps even more so than ever before. For example, Bredekamp (2011) writes about the "principle of disjunction," whereby the more natural and "real" such scientific images appear, the more constructed and artificial they are (see also, Bredekamp, Schneider, & Dünkel, 2008). As images are increasingly created by technological devices, we have come to think of ourselves as neutral observers of reality, whose "objectivity" is assured to the degree that our tools remove our "subjectivity"; to the degree that objective facts replace subjective interpretation. "Objectivity preserves the artifact or variation that would have been erased in the name of truth… objectivity is blindsight, seeing without inference, interpretation, or intelligence" (Daston & Galison, 2010, p. 17). Modern technologies, such as neuroimaging, promise to bring the viewer closer to the phenomena they depict – they claim to present reality in an unmediated manner. As to be expected, objections have been raised to such claims. For example, Fleck (1935) argued that even the most scientific and "objective" of images is ultimately meaningless without the social practices of the communities that use them (e.g., medical doctors, research scientists). Morin's piece is a powerful argument for the centrality of practice for language, and an important reminder that our attempts to symbolically capture reality, regardless of the forms they may take, are matters of practice – even ideographs. Perhaps what remains truly puzzling, is *our expectation* that in the case of ideography it would be otherwise.

Competing interest. None.

References

Bredekamp, H. (2011). In der Tiefe die Künstlichkeit: Das Prinzip der bildaktiven Disjunktion. In H. Bredekamp & J. M. Krois (Eds.), *Sehen und Handeln* (pp. 206–224). Akademie Verlag. https://doi.org/10.1524/9783050062389.206
Bredekamp, H., Schneider, B., & Dünkel, V. (2008). *Das Technische Bild. Kompendium zu einer Stilgeschichte wissenschaftlicher Bilder*. Akademie Verlag. https://doi.org/10.1524/9783050062198
Bruner, J. (1990). *Acts of meaning*. Harvard University Press.
Daston, L. J., & Galison, P. (2010). *Objectivity*. Zone Books.
Fleck, L. (1935). *Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv*. Benno Schwabe.
Taylor, C. (1985). Peaceful coexistence in psychology. In C. Taylor (Ed.), *Human agency and language* (pp. 117–138). Cambridge University Press.

A source- and channel-coding approach to the analysis and design of languages and ideographies

Mihnea Moldoveanu

Desautels Centre for Integrative Thinking, Rotman School of Management, University of Toronto, Toronto, ON, Canada
mihnea.moldoveanu@rotman.utoronto.ca; https://www.rotman.utoronto.ca/FacultyAndResearch/Faculty/FacultyBios/Moldoveanu

doi:10.1017/S0140525X23000705, e248

Abstract

Can we explain the advantage natural languages enjoy over ideographies in a way that enables us to attempt the design of an ideography that "works"? I deploy an adapted version of Shannon's source- and channel-coding partitioning of a communication system to explain the communicative dynamics and shortfalls of ideographies, and reveal ways in which entrenchable, generalist ideographies could be designed.

Morin's *Puzzle of ideography* advances a persuasive argument for why general-purpose, entrenched ideographies are not widespread, in spite of the advantage that they can be deployed in communication among speakers competent in different languages. But it does not allow us to predict whether or not a new ideography is more or less likely to fail than were its predecessors. Can we conceptualize the difference between ideographies and natural languages that allows us to do that, and perhaps to *design* an ideography that is both generalized and entrenched? I outline such an approach, adapting Shannon's (1948; Cover and Thomas, 2006) model of a communication system as the concatenation of a source and channel encoder/decoder (Fig. 1) to represent the evolution and deployment of *both* a natural language and an ideography:

(1) The source encoder maps a plaintext message/picture into a set of codewords designed to allow the decoder to reconstruct the message from the code. The source decoder inverts the encoder's operation to recover the original message from the output of the channel decoder.
(2) The channel encoder maps the output of the source encoder onto a set of codewords that protect against distortions introduced by the channel by intelligently reintroducing redundancies in the signal. The channel decoder inverts the operations in ways informed by insight into the channel conditions – to recover the input to the source decoder.

The separation of source- and channel-coding functions arises from their different purposes:

(1) the source encoder squeezes redundancy from the original message by exploiting hidden regularities to produce the shortest code from which the message can be reconstructed by the decoder.
(2) The channel encoder adds redundancy to the output of the source encoder to generate a transmitted signal maximally intelligible to a receiver over a channel.

For a natural language: Meaning-bearers (written words, propositions) source encode underlying messages (referents: "states of affairs," "events," "objects," "relationships"). Redundancy – repetition, rephrasing, explication – and interpersonal feedback signals like queries and gestures are added to the message in ways adaptive to different channels, which could be synchronous (in person speech acts), or asynchronous (emails) and corrupted by noise that may be "white" (bad phone connection), or selectively affect different parts of the message (accented speech, ungrammatical text).

For an ideography: Meaning-bearers ("ideograms") encode underlying referents using stylized pictures that do not require knowledge of a natural language to decode. Channel encoding is more restrictive. As Morin points out, it may comprise acts of pointing, magnification, or repetition. One can argue some written languages – like English – already incorporate some redundancy into the design of the source encode: English vowels are largely redundant, and added to increase the intelligibility of text, which is often comprehensible from consonants alone ("th qck brn fx …"). They are also useful as prompts for correct pronunciation to help listeners distinguish among different words when they are spoken. Such languages lump together some source- and channel-coding *operations*, but they do not alter the parsing of communicative tropes into source- and channel-coding *operations*.

This approach to representing communication tools illuminates several aspects of the differences between languages and ideographies, such as:

(1) Morin's explanation of the difficulty of establishing a generalist ideography that is entrenched: Even though natural languages entail ambiguities mapping sign to referent that can make it difficult to uniquely recover a pointer to a specific referent (e.g., an object) from the source coder output (the word that spells the name of the object), they enable ample opportunities for error correction over both synchronous and asynchronous channels that ideographic communication lacks.
(2) The selective appeal of ideographies in specialized domains such as mathematical physics, which Morin relevantly references: Ideograms (mathematical symbols and formulas) efficiently encode their referents (operations and entities such as sets and functions defined on sets) in ways that make the source message more efficiently and less ambiguously reconstructible from received messages than alternative representations in natural language, suggesting source-coding advantages can compensate for more restricted channel-coding opportunities.

If, as Morin argues, there is no "hard-wired" advantage to the development of natural languages, in the form of neurological constraints or adaptive or exaptive developments, then one explanation for the absence of widely entrenched general-purpose ideographies is that attempts to build them have not been guided by an intelligent information-theoretic conception of a communicative device, which has led to the design of ideographies that mimic natural language sentences instead of exploiting the idiomatic source-coding advantages of ideograms and optimizing their communicability over plausible channels.

What could an ideography that did so look like? Shannon's analysis of source and channel coding offers some guidance: His *source-coding theorem* (Cover & Thomas, 2006) links the *average size of source codewords* to the *entropy of the source signal* – which measures the number of possible objects or states of affairs a codeword can refer to given the frequency with which they ecologically occur. His *channel-coding theorem* links the amount of redundancy we need to add to a signal sent over a

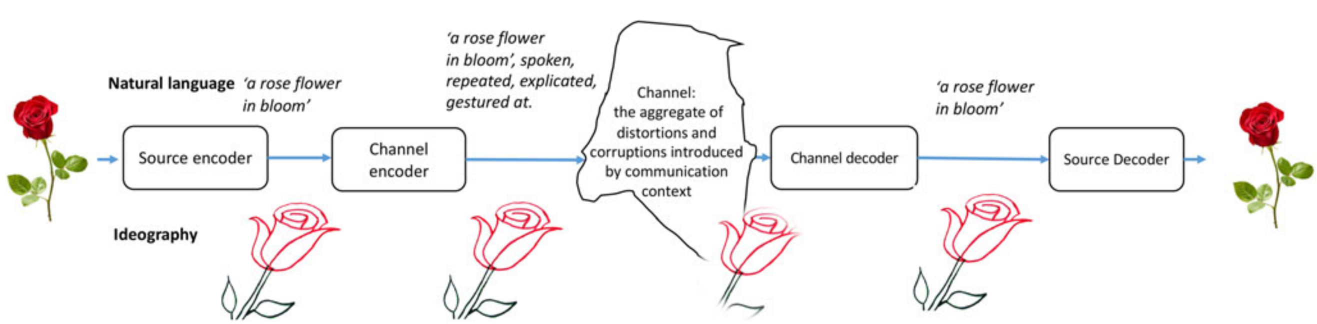

Figure 1 (Moldoveanu). Source- and channel-coding decomposition of the functions of ideographies and natural languages.

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

channel with a given capacity (i.e., the mutual information of sent and received signals) – to the probability of recovering the signal with at most a certain error. Thus:

(1) An "intelligently designed" ideographic *source encoder* can minimize the number of possible states of affairs that needs to be encoded by the source encoder by featuring a set of *objects* (ideograms) that have maximal *translational and scale* invariance in space–time, and maximize contextual invariance by encoding objects in terms of their *causal powers* as opposed to nonessential properties ("being within *Y* feet of *X*").
(2) To maximize its channel-coding *gain*, it can enable multiple levels of *resolution* and spatial *rotation/translation* (zooming and panning), decomposability (whole → parts) and compositionality (parts → whole) to make *purely graphical explanations* possible; and – a set of *ideographically idiomatic communicative acts* (zoom, decompose, pan) that are different from the queries of natural language and designed to make a purely graphical communicative act subject to pointy requests for clarification.

Financial support. This work was supported by the Desautels Centre for Integrative Thinking at the Rotman School of Management, University of Toronto.

Competing interest. None.

References

Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. Wiley.
Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423, 623–636.

Ideography insight from facial recognition and neuroimaging

Benjamin C. Nephew[a], Justin J. Polcari[a] and Dmitry Korkin[b]

[a]Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA, USA and [b]Department of Computer Science, Worcester Polytechnic Institute, Worcester, MA, USA
bcnephew@aol.com
jjpolicari@wpi.edu
dkorkin@wpi.edu

doi:10.1017/S0140525X23000602, e249

Abstract

One novel example and/or perspective in support of "Why the learning account fails" is the impressive ability of humans to recognize and memorize facial features and accurately and reliably connect those to related identities. Furthermore, neuroimaging analysis presents an example in support of the crucial role of standardization in the lack of adoption of ideography.

One novel example and/or perspective in support of, "Why the learning account fails," that the lack of viable ideographies is a learnability issue, is the impressive ability of humans to recognize and memorize facial features and accurately and reliably connect those to related identities. Faces are an aggregation of graphical features. Individuals are capable of learning and remembering hundreds of faces and facial expressions over decades and can identify extremely minor differences in facial features starting from a very young age (Nelson, Morse, & Leavitt, 1979). Human facial recognition is robust additional evidence that we are fully capable of processing and storing graphical features as well as phonemes. The challenge or failure of sophisticated artificial intelligence algorithms in correctly identifying faces compared to humans is a prime example of human ability in the graphical realm (Cavazos, Phillips, Castillo, & O'Toole, 2021; Phillips et al., 2018). This is not only an ability that humans are skilled at learning, it is also a highly evolved inherent skill crucial to maternal–infant bonding and successful early development. Brain regions critical for the identification of faces exhibit increased activity in new mothers compared to nulliparous women and these changes are closely linked to empathetic concern (Rigo et al., 2019; Zhang et al., 2020). It is possible that the failure of some ideographies is because of a lack of emotional relevance, where an increased focus on emotional impact in the development of graphic codes will result in enhanced adoption.

Although most cognitive traits possess modest heritability, twin studies of facial recognition reveal a high-genetic contribution (Wilmer et al., 2010). However, this ability does not strongly correlate with measures of visual and verbal recognition, leading to the conclusion that facial recognition is both specific and highly heritable, underscoring its importance to our survival as a mammalian species highly dependent on parental care and social interaction. Although this specificity of facial recognition could be seen as support for learnability-based theories of challenges to ideographies, it can also be interpreted as a lack of relevant and valid tools to reliably assess general visual recognition (Diaz-Orueta, Rogers, Blanco-Campal, & Burke, 2022). Along these lines, perhaps what is needed for the development of effective ideographies is a focus on less abstract, more organic shapes and features commonly found in human faces. Given the findings from studies of mothers and the heritability of facial recognition, it could indeed be that emojis are the future of effective ideography.

Neuroimaging analysis presents an example in support of the crucial role of standardization in the lack of adoption of ideography. Functional neuroimaging involves the processing and interpretation of brain activity that is often depicted solely by graphical representation. Similar to facial recognition algorithms, neuroimaging analysis often involves numerous assumptions and compromises and the never-ending methods development in this field has led to overall inconsistency (Poldrack et al., 2017). The lack of standardization in the analysis of neuroimaging data is widely acknowledged as a key reason for poor replication and repeatability in neuroimaging studies in the past (Kennedy et al., 2019), similar to the weak agreement on the meaning of various emojis. Accordingly, recent efforts to broadly increase standardization in neuroimaging (Niso et al., 2022), functionally increasing agreement on the meaning of graphical images of brain activity, may substantially enhance replication across the field. At the individual level, the standardization in functional neuroimaging that is focused on changes in brain activity has accelerated the learning process within and across trainees, resulting in the increased adoption of what could be referred to as a very complex form of ideography.

An integral aspect of neuroimaging analysis is the discrimination of patterns of functional brain activity versus patterns indicative of statistical noise/variability because of various confounds such as electrical interference or subjects moving during scans. This can occur through two methods, one based on organic human learning and training, the other based on a form of artificial intelligence, deep learning. The identification and extraction of these confound-based patterns can be accomplished through the training of individuals and/or the use of an image recognition program. Similar to facial recognition, the best performance is often obtained through the use of both image recognition programs and subsequent human confirmatory inspection. However, variability in training and human performance has adverse effects on reproducibility and current efforts are focused on the application of deep learning to neuroimaging analysis (Abrol et al., 2021). Perhaps we should consult ChatGPT on how best to design a self-sufficient graphic code?

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Abrol, A., Fu, Z., Salman, M., Silva, R., Du, Y., Plis, S., & Calhoun, V. (2021). Deep learning encodes robust discriminative neuroimaging representations to outperform standard machine learning. *Nature Communications*, *12*(1), 353. doi:10.1038/s41467-020-20655-6

Cavazos, J. G., Phillips, P. J., Castillo, C. D., & O'Toole, A. J. (2021). Accuracy comparison across face recognition algorithms: Where are we on measuring race bias? *IEEE Transactions on Biometrics, Behavior, and Identity Science*, *3*(1), 101–111. doi:10.1109/TBIOM.2020.3027269

Diaz-Orueta, U., Rogers, B. M., Blanco-Campal, A., & Burke, T. (2022). The challenge of neuropsychological assessment of visual/visuo-spatial memory: A critical, historical review, and lessons for the present and future. *Frontiers in Psychology*, *13*, 962025. doi:10.3389/fpsyg.2022.962025

Kennedy, D. N., Abraham, S. A., Bates, J. F., Crowley, A., Ghosh, S., Gillespie, T., … Travers, M. (2019). Everything matters: The ReproNim perspective on reproducible neuroimaging. *Frontiers in Neuroimaging*, *13*. doi:10.3389/fninf.2019.00001

Nelson, C. A., Morse, P. A., & Leavitt, L. A. (1979). Recognition of facial expressions by seven-month-old infants. *Child Development*, *50*(4), 1239–1242.

Niso, G., Botvinik-Nezer, R., Appelhoff, S., De La Vega, A., Esteban, O., Etzel, J. A., … Rieger, J. W. (2022). Open and reproducible neuroimaging: From study inception to publication. *NeuroImage*, *263*, 119623. https://doi.org/10.1016/j.neuroimage.2022.119623

Phillips, P. J., Yates, A. N., Hu, Y., Hahn, C. A., Noyes, E., Jackson, K., … O'Toole, A. J. (2018). Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *PNAS*, *115*(24), 6171–6176. doi:10.1073/pnas.1721355115

Poldrack, R. A., Baker, C. I., Durnez, J., Gorgolewski, K. J., Matthews, P. M., Munafò, M. R., … Yarkoni, T. (2017). Scanning the horizon: Towards transparent and reproducible neuroimaging research. *Nature Reviews Neuroscience*, *18*(2), 115–126. doi:10.1038/nrn.2016.167

Rigo, P., Kim, P., Esposito, G., Putnick, D. L., Venuti, P., & Bornstein, M. H. (2019). Specific maternal brain responses to their own child's face: An fMRI meta-analysis. *Developmental Review*, *51*, 58–69. https://doi.org/10.1016/j.dr.2018.12.001

Wilmer, J. B., Germine, L., Chabris, C. F., Chatterjee, G., Williams, M., Loken, E., … Duchaine, B. (2010). Human face recognition ability is specific and highly heritable. *PNAS*, *107*(11), 5238–5241. doi:10.1073/pnas.0913053107

Zhang, K., Rigo, P., Su, X., Wang, M., Chen, Z., Esposito, G., … Du, X. (2020). Brain responses to emotional infant faces in new mothers and nulliparous women. *Scientific Reports*, *10*(1), 9560. doi:10.1038/s41598-020-66511-x

Notational systems are distinct cognitive systems with different material prehistories

Karenleigh A. Overmann

Department of Anthropology, Center for Cognitive Archaeology, University of Colorado, Colorado Springs, CO, USA
koverman@uccs.edu; https://uccs.academia.edu/KarenleighOvermann

doi:10.1017/S0140525X2300078X, e250

Abstract

Notations are cognitive systems involving distinctive psychological functions, behaviors, and material forms. Seen through this lens, two main types – semasiography and visible language – are fundamentally differentiated by their material prehistories, emphasis on iconography, and the centrality of language's combinatorial faculty. These fundamental differences suggest that key qualities (iconicity, expressiveness, concision) are difficult to conjoin in a single system.

Notations are cognitive systems that can be analyzed through their components: Brains (psychological functions; neurological regions), bodies (distinctive behaviors; physiological capacities), and material forms (sets of visually appreciated, manually engaged elements) (Malafouris, 2013). Viewing notations through this lens reveals two main types: Semasiography and visible language. Their unique material predecessors and developmental histories potentially bear on Morin's ideographic puzzle.

Language is not central to semasiographic notations (music, numbers). In fact, adding language destroys the visual concision needed to bring large volumes of graphically represented information together for simultaneous viewing. These notations also involve a significant implicit component, knowledge the user must master to make proper sense of them.

Semasiographic notations can be rendered into language but do not express language: ♪♫♪ can be put into words, but the result is not music. Musical notations encode information about pitch, duration, and emphasis in a way that informs the motor manipulations needed to produce sounds with an instrument, including the human voice. The implicit component involves knowing how specific notations inform particular movements and acquiring proficiency in their production; the latter in particular involves distinctive neurological reorganizations (Gaser & Schlaug, 2003).

Numbers are more complex. Operational signs (+−×÷) are mechanical instructions that work equally well with written numerals (knowledge-based calculation) and abacus beads (calculation that is more manipulation based). As for the numbers they manipulate, Western numerals (0–9) descend from Mesopotamian and Egyptian notations (Chrisomalis, 2010) that meant particular quantities by instantiating them: They *meant* quantity by *being* quantity (Overmann, 2022). For example, the sign for *five* had five elements: Mesopotamian 𒐊; Egyptian 𓏾. Notations share this representational mode with their precursor technologies: Fingers ✋, tally marks |||||, tokens/pebbles •••••, and cuneiform/hieroglyph numerals all mean quantity by instantiating it, without any need to encode language. Repetition also deemphasizes the need for iconic variation.

As number systems elaborate, their representations become more concise. Conventions for large values (10; 60; or 100) are added, and instantiate notations become conventions whose value is learned, rather than depending on element countability: 7 is concise but not explicit in its value; 𒐌 and ||||||| instantiate value explicitly but are not concise (Overmann, 2023). Concision is critical to numbers: It enables them to be brought together at volumes and with a semantic succinctness that can reveal their relations and patterns. Like musical notations, operations and numerals can be spoken, but the result is ephemeral, memory dependent, and loses significant manipulability, accuracy, and complexity.

Granted, the implicit component of music and numbers is largely acquired and experienced through language. However, a substantial motor component is also involved, consistent with manipulating objects – both physical and conceptual – and differing significantly from neural activity associated with processing language (Amalric & Dehaene, 2018; Grotheer, Ambrus, & Kovács, 2016; Johansson, 2008; Perfetti & Tan, 2013; Vandervert, 2017).

In contrast, language is central to visible language. So-called glottographic scripts express particular languages with fidelity. For modern scripts like the one used here, the graphic elements are contrastive and recombinable in ways that suggest the cooption of a key language function: The ability to combinatorially recombine conceptual and phonetic units. Concision is deemphasized; signs instead become more visually complex by incorporating techniques (determinatives; phonography) for specifying the intended morphemes, syllables, or sounds. The implicit component involves knowing the language being expressed and acquiring the suite of behavioral and neurological reorganizations that recognize written characters, associate them with the meanings and sounds of language, and coordinate and guide productive movements (Dehaene & Cohen, 2007; Nakamura et al., 2012; Roux et al., 2009).

Early writing did not express language with fidelity. In Mesopotamia, it was so poor in this regard that scholars still debate the language associated with it. Early writing consisted of pictures and conventions with approximate meanings. Pictures are ambiguous in what they mean: Resemblance suggests a range of related semantic meanings, and as Morin notes, pictures may acquire cultural meanings independent of what they resemble. This ambiguity can motivate the inclusion of techniques for specifying particular words.

The process by which early writing becomes script is complex and includes morphological change in the written form. Considered across cultural spans of time – from early writing to, say, a thousand or so years later – this morphological change is consistent with the neurological reorganizations that constitute literacy (Overmann, 2016): As characters became recognized topologically – through combinations of their structural features and spatial relations – they no longer needed to maintain their original forms; this freed the set of characters to converge on features and contrasts that maximized discriminability and individualization. This in turn facilitated their recombination in ways that borrowed or mimicked the combinatorial faculty of language, and over time, sets of characters shrank (e.g., syllabaries and alphabets; Chinese is exceptional in this regard). Along the way, the neurological reorganizations needed to read and write the characters became paramount: Without them, scripts were no longer meaningful.

Like script, sign language transforms over time from emergent gestures to a complex state involving combinatorial recombination (Senghas & Coppola, 2001). Handwritten or signed, visible language involves a significant motor component like semasiographic notations do. Yet neither script nor sign has a similar material prehistory: There are no direct material precursors, no counterparts to the stalagmites, flutes, tallies, and tokens/pebbles used in prehistoric music and counting. Visually perceived forms (seals; reliefs; art) undoubtedly influenced the ideography used in visible language, but they did not constitute a prehistory of material engagements like those in music and numbers.

These material prehistories suggest there are two distinctive notational types (which, granted, interact and overlap at multiple levels in highly complex ways). In semasiography, notations follow from and concentrate a lengthy material sequence and act as precise instructions for manipulating physical or conceptual objects. In visible language, notations lack direct material precursors, initially emphasize the use of iconography to convey meanings, and can become increasingly expressive by recombining their elements like language does. Given these differences, extracting and blending iconicity, expressiveness, and concision, as Morin proposes, may well remain a challenge.

Competing interest. None.

References

Amalric, M., & Dehaene, S. (2018). Cortical circuits for mathematical knowledge: Evidence for a major subdivision within the brain's semantic networks. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *373*(1740), 20160515.

Chrisomalis, S. (2010). *Numerical notation: A comparative history*. Cambridge University Press.

Dehaene, S., & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, *56*(2), 384–398.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience*, *23*(27), 9240–9245.

Grotheer, M., Ambrus, G. G., & Kovács, G. (2016). Causal evidence of the involvement of the number form area in the visual detection of numbers and letters. *NeuroImage*, *132*, 314–319.

Johansson, B. B. (2008). Language and music: What do they have in common and how do they differ? A neuroscientific approach. *European Review*, *16*(4), 413–427.

Malafouris, L. (2013). *How things shape the mind: A theory of material engagement*. MIT Press.

Nakamura, K., Kuo, W.-J., Pegado, F., Cohen, L., Tzeng, O. J.-L., & Dehaene, S. (2012). Universal brain systems for recognizing word shapes and handwriting gestures during reading. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(50), 20762–20767.

Overmann, K. A. (2016). Beyond writing: The development of literacy in the Ancient Near East. *Cambridge Archaeological Journal*, *26*(2), 285–303.

Overmann, K. A. (2022). Early writing: A cognitive archaeological perspective on literacy and numeracy. *Visible Language*, *56*(1), 8–44.

Overmann, K. A. (2023). *The materiality of numbers: Emergence and elaboration from prehistory to present*. Cambridge University Press.

Perfetti, C. A., & Tan, L.-H. (2013). Write to read: The brain's universal reading and writing network. *Trends in Cognitive Sciences*, *17*(2), 56–57.

Roux, F., Dufor, O., Giussani, C., Wamain, Y., Draper, L., Longcamp, M., & Démonet, J. (2009). The graphemic/motor frontal area Exner's area revisited. *Annals of Neurology*, *66*(4), 537–545.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, *12*(4), 323–328.

Vandervert, L. R. (2017). The origin of mathematics and number sense in the cerebellum: With implications for finger counting and dyscalculia. *Cerebellum & Ataxias*, *4*(12), 1–16.

How standardized must a code be to be useful?

Andrew M. Riggsby

Department of Classics, University of Texas at Austin, Austin, TX, USA
ariggsby@mail.utexas.edu; https://utexas.academia.edu/AndrewRiggsby

doi:10.1017/S0140525X23000699, e251

Abstract

If, as it appears, failure of standardization blocks the rise of general-purpose ideography, then a more precise characterization of "standardization" should help illuminate aspects of the process. Comparison is made with several histories of standardization to outline relevant dimensions and thresholds. This line of inquiry is particularly important for the forward-looking question of whether such ideography can ever arise.

Morin argues that a standardization-oriented account best explains the nonexistence of "generalist, self-sufficient ideographies" (target article, sect. 1, para. 12). As against the competing learnability accounts his argument is compelling, but the space available means the notion of "standardization" is unavoidably somewhat underspecified. A more granular account will provide insight into the precise mechanisms at work. Morin's article looks directly at a variety of codes, but also argues analogically from histories of standardization in other technologies, and I will proceed similarly.

"Standardization" it is not a simple property. First, at a minimum (and as glanced at in the article), it often is a scalar rather than a binary. Metrological standardization, for instance, takes place within some range of tolerance. Second, the thresholds vary with how extensive the scope of any standard is in time and space. So, metrological standardization in the Classical Mediterranean world as a whole was quite loose (Riggsby, 2019, pp. 100–114), but for individual construction projects it could be extremely precise (Riggsby, 2019, pp. 115–116). Third, outside of pure metrology, standardization is multidimensional; ancient Mediterranean transport vessels standardized volume, shape, and perhaps weight (Olmer, 2001). Even weight systems sometimes standardize shape as well as mass, often to the detriment of precision in the latter (ancient Rome: Luciani & Lucchelli, 2016; ancient Near East: Schon, 2015; early modern Burma: Gear & Gear, 1992). Finally, multidimensional standardization often requires the choice of dimensions along which difference will be ignored. For instance, the rise of the modern commodity grain trade in the American Midwest involved the collapsing of numerous qualitative distinctions into a small number of standard categories to permit large-scale, automated handling of the product (Cronon, 1991, sects. 2485–2621).

These considerations then raise the following question in the context of the target article. How much standardization of what sort(s) across what sort of group(s) would an ideography require to succeed? In the following I consider in these terms two cases of codes that *have* been successful as a way of organizing the inquiry.

Scripts that allow the writing of languages are codes and thus they should also require standardization. (That the scripts are simpler than the languages themselves, vastly so in most cases, presumably reduces but does not eliminate the demand here.) Current research on evolution of scripts points immediately to a limit on the standardization required. For instance, Miton and Morin (2021) choose to ignore "font" or merely "typographic" variants of character shapes (e.g., sans-serif or italic forms of Latin script). This seems correct, but what are the limits?

Consider the attested variation in Sumerian cuneiform and hieroglyphic Mayan forms shown in Figures 1–3.

The reader may compare the more exhaustive listings for Sumerian (Mittermayer, 2006) and Mayan (Kettunen & Helmke, 2020; Prager, 2014) to see that this level of diversity is common and that vastly more varied cases also exist (e.g., Sumerian MAḪ or ĜIR$_2$ or Mayan /u/ or TZ'AK "whole"). The diversity cannot be significantly reduced by choosing to ignore features (as in the commodity case above). Moreover, although there is considerable evolution over time in both scripts (cf. Miton & Morin, 2021, for some general principles at work), there is also a clear local dimension. For Mayan, this can be seen directly in the site-specific syllabaries collected by Boot (2010a, 2010b, 2010c), and for cuneiform the time-limited sign lists collected by the Cuneiform Digital Library Initiative (2016). In this context I would suggest that the standardization of early-complex scripts is perhaps surprisingly low, certainly extending well beyond what is recognizable as variation in "font." This non-standardization appears to be one of the reasons it has been particularly to develop computer character recognition for cuneiform scripts (Bogacz, Gertz, & Mara, 2015; Bogacz & Mara, 2022).

We can also look a little more closely at the role of scope of standardization. That is, given some notional target level of standardization, how broadly is it achieved across spaces or populations? In the Sumerian and Mayan cases, writing and reading were done principally by a professional scribal class (Delnero, 2010; Houston, 1994), and most of our surviving texts were meant for consumption within the same circles in which they were produced. This should be a congenial environment for the development of small islands of high standardization. At the same time, the same scripts were also used for communication outside those immediate contexts (Michalowski, 2011; Tedlock, 2010, pp. 146–164), so their success was not purely an artifact of isolation.

A similar pattern arises from looking at another set of codes: Languages themselves. Morin is undoubtedly right that

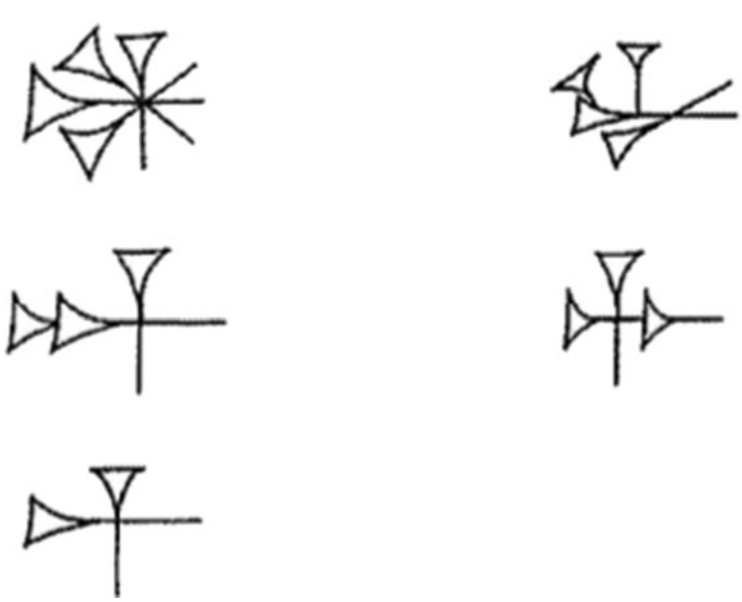

Figure 1 (Riggsby). Variant forms of AN "sky, heaven" (Mittermayer, 2006, p. 5).

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

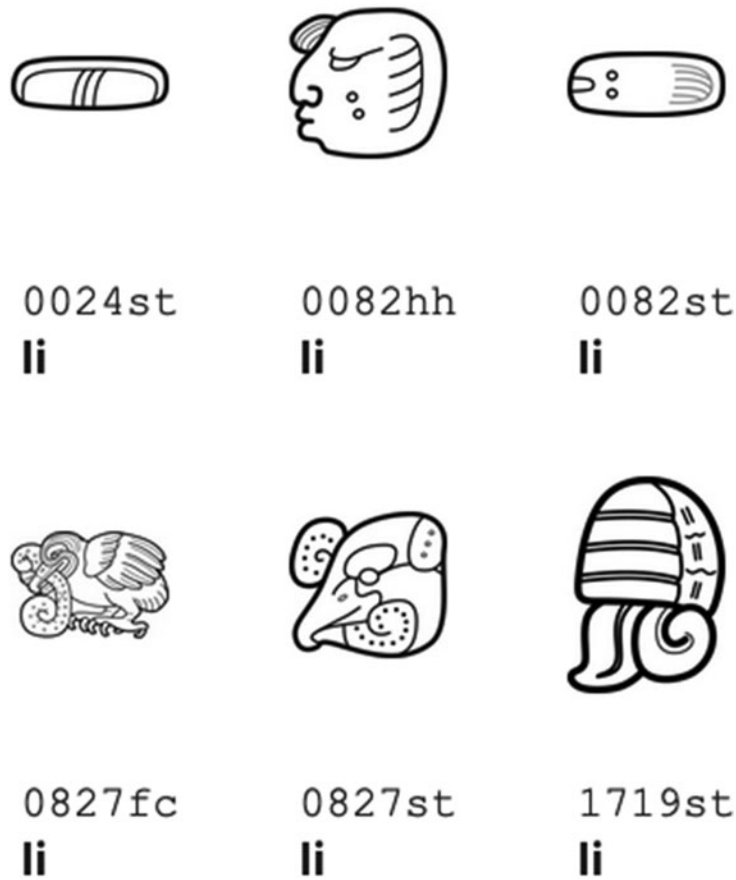

Figure 2 (Riggsby). Variant forms of phonetic /li/ (Prager, 2014).

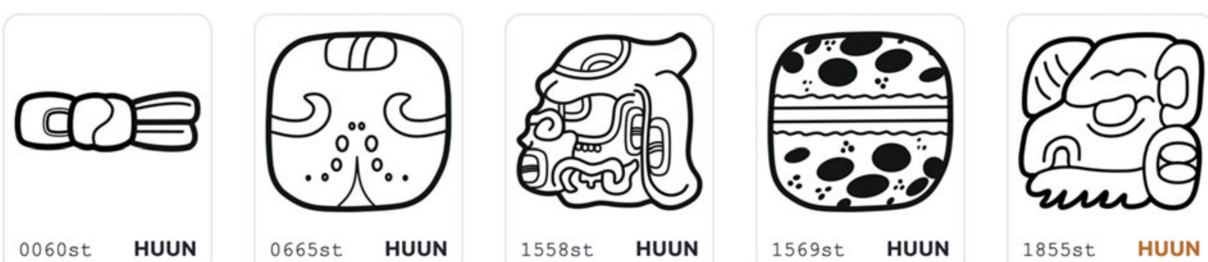

Figure 3 (Riggsby). Variant forms of logographic HUUN "writing material" (Prager, 2014).

face-to-face language (oral or signed) has self-standardization mechanisms that no graphic code does, but the leap from "the level of the pair" to "entire populations" (target article, sect. 6, para. 8) of course elides some complexity. Of particular relevance here is the phenomenon of dialect continua, in which chains of mutually comprehensible language forms cumulate large differences from end to end (Chambers & Trudgill, 1998). In both cases (scripts, spoken dialects) overall standardization is low but the codes in question remain useful. This is apparently sustainable because of the existence of pockets of relatively high standardization.

None of the above seems to me to cast any doubt on the fundamental truth of Morin's historical argument. Nonetheless, these issues are worth raising for two reasons. Historically, how much standardization has been achieved by individual codes in particular contexts? In particular, to what extent have small communities been able to serve as nuclei to produce and diffuse complex standards as an intermediate step to universalization? Prospectively, the answer to that helps us calibrate an answer to Morin's question whether emoji (or anything else) could evolve into the world's first general-purpose ideography.

Acknowledgment. I would like to thank Mallory E. Matsumoto for invaluable guidance on the Mayan material. Any errors of interpretation are my own.

Financial support. This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Competing interest. None.

References

Bogacz, B., Gertz, M., & Mara, H. (2015). Cuneiform character similarity using graph representations. In P. Wohlhart & V. Lepetit (Eds.), *20th computer vision winter workshop* (pp. 1–8). Universitätsbibliothek Heidelberg.

Bogacz, B., & Mara, H. (2022). Digital assyriology – Advances in visual cuneiform analysis. *ACM Journal on Computing and Cultural Heritage*, *15*(2), article 38.

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

Boot, E. (2010a). Maya glyph blog 2010, No. 1: Chichen Itza. Retrieved January 15, 2023, from http://maya-glyph-blog.blogspot.com/2010/05/maya-glyph-blog-2010-no.html
Boot, E. (2010b). Maya glyph blog 2010, No. 2: Xcalumkin. Retrieved January 15, 2023, from http://maya-glyph-blog.blogspot.com/2010/07/?m=0
Boot, E. (2010c). Maya glyph blog 2010, No. 3: Tikal. Retrieved January 15, 2023, from http://maya-glyph-blog.blogspot.com/2010/09/maya-glyph-blog-2010-no.html
Chambers, J. K., & Trudgill, P. (1998). *Dialectology* (2nd ed.). Cambridge University Press.
Cronon, W. (1991). *Nature's metropolis: Chicago and the great west*. W. W. Norton. Kindle edition.
Cuneiform Digital Library Initiative. (2016). Sign lists. Retrieved January 16, 2023, from https://cdli.ox.ac.uk/wiki/sign_lists
Delnero, P. (2010). Sumerian extract tables and scribal education. *Journal of Cuneiform Studies*, *62*, 53–69.
Gear, D., & Gear, J. (1992). *Earth to heaven: The royal animal-shaped weights of the Burmese empires*. Twinstar.
Houston, S. (1994). Literacy among the Precolumbian Maya: A comparative perspective. In E. Boone & W. Mignolo (Eds.), *Writing without words: Alternative literacies in Mesoamerica and the Andes* (pp. 27–49). Duke University Press.
Kettunen, H., & Helmke, C. (2020). *Introduction to Maya hieroglyphs* (17th revised ed.). Retrieved January 15, 2023, from https://wayeb.org//download/Kettunen_Helmke_2020_Introduction_to_Maya_Hieroglyphs_17th_ed.pdf
Luciani, F., & Lucchelli, T. (2016). Pondera exacta ad Castoris. *Aquileia Nostra*, *79*, 122–147.
Michalowski, P. (2011). *The correspondence of the Kings of Ur an epistolary history of an ancient Mesopotamian kingdom*. Eisenbrauns.
Miton, H., & Morin, O. (2021). Graphic complexity in writing systems. *Cognition*, *2021*, 104771.
Mittermayer, C. (2006). *Altbabylonische Zeichenliste der sumerisch-literarischen Texte*. Academic Press/Vandenhoek & Ruprecht.
Olmer, F. (2001). Le anfore e le misure di capacità. In C. Corti & N. Giordani (Eds.), *Pondera: pesi e misure nell'antichità* (pp. 227–236). Museo di Bilancia.
Prager, C. (Ed.). (2014). Text database and dictionary of Classic Mayan, a working list of Maya Hieroglyphic signs and graphs. Retrieved January 15, 2023, from https://mayawoerterbuch.de/zeichenkatalog/
Riggsby, A. M. (2019). *Mosaics of knowledge: Representing information in the Roman world*. Oxford University Press.
Schon, R. (2015). Weight sets: Identification and analysis. *Cambridge Archaeological Journal*, *25*(2), 477–494.
Tedlock, D. (2010). *2000 years of Mayan literature*. University of California Press.

Mind the gap: Why is there no general purpose ideographic system?

Kim Sterelny

School of Philosophy, Research School of the Social Sciences, Australian National University, Acton, ACT, Australia
Kim.Sterelny@anu.edu.au

doi:10.1017/S0140525X23000596, e252

Abstract

Morin has identified an intriguing puzzle about human communication systems, and one element of the solution: Inscriptional sign systems pose more coordination problems, making sender/receiver coadaptation more difficult. But I reject his view of written language, concluding that inscriptional sign systems can be generalist. The upshot is a cost-based proposal about why generalist ideographic systems are essentially unknown.

I agree with the core ideas of the target article. Although there are many inscription-based sign systems, most of these, perhaps all, are specialist. A paradigm is the Chess Informator system of the 1960s; a Yugoslavian innovation that combined an existing alphanumeric convention for recording chess moves, icons based on the most common design of chess pieces, and a small set of arbitrary symbols expressing chess-specific judgements (e.g., "++" after a series of moves meant "with decisive advantage"). This innovation made it possible for the publisher in question to produce a languageless series of volume recording and analysing the main games of the preceding half year. Until the internet took over, these were the main archive of the chess world. Chess Informator shows that language-independent ideographic communication was both possible and advantageous.

This example is not unique: Music is another. As Morin shows, ideographic systems are typically specialist, designed for communicating about a specific domain. There is a gap in the space of ideographic systems, with Morin arguing convincingly that there are no general purpose, language-independent ideographic systems. Chinese scripts have been taken to be counterexamples to this claim, in part because they have been repurposed to express Japanese, and in part because a common ideogram menu was used in China despite very considerable regional differences in spoken languages. But Morin, to my mind, is convincing in arguing that these scripts are not fully language-independent. He is much less convincing in arguing that there are no general-purpose ideographic systems at all, as this obliges him to reject the view that written English, for example, is a general purpose system. He claims instead that it is a special purpose system for representing (in this case) the phonology and morphology of spoken English, saying at one point that "writing, as a code, represents language" (target article, sect. 3, para. 4) and a little later that it is, as a code "highly specialized, merely a notation of … phonemes" (target article, sect. 4, para. 1). Morin supports this with the observation that in liturgical contexts we find individuals who can read and recite sacred texts that they do not understand. Although striking, it is far from clear that this competence is productive; that they can write as well as read, or read novel texts or novel layouts of a known text.

This move misrepresents an important insight. It is true that there is a sense in which written English is parasitic on spoken (or signed) English. No one learns written English as their first or only language: Literacy is always scaffolded by a prior linguistic competence, and this is indeed relevant to explaining the gap. But this unnecessary claim pushes Morin into a strange corner. He has to deny the existence of a phenomenon once common at universities: Individuals literate in a language they cannot speak (this is the inverse of his ecclesiastic example above). He has no natural way to describe the competence of the literate but congenitally deaf, because written English has a much less direct relationship to signed English than to spoken English. Moreover, although Chinese ideograms are not language-independent, they are not plausibly described as specifications of the voicing of their spoken equivalents. Finally, it seems that Morin has to say strange things about the semantics of written sentence: If written English represents its spoken version, presumably the written sentence "Berlin is the capital of Germany" is true, because it correctly represents the structure and sounding of the spoken twin of that sentence.

The relationship between written and spoken language is asymmetrical not because written language is about spoken language, but through constraints on the acquisition of language (and to a lesser extent, its use). As Morin points out, when agents are acquiring a rich and flexible system (especially one that can be used idiosyncratically, on the fly, to cope with unexpected communicative challenges), misfires are inevitable, and efficient repair is essential. In face-to-face interaction, misfires can usually be detected and corrected quickly, because sign production is rapid, cheap, and supported by nonlinguistic signals. That allows conversation to be multiparty, free-flowing but repairable (if Steve

Levinson is right, supported by cognitive mechanisms tuned to both the speed of conversational flow, and the sequential appearance of signs over time; see, e.g., Levinson, 2019). That is not true of ideographic systems. These allow individuals to communicate over space and time, but they do not support conversation at a place and time: They are not interactable in the way voice or sign is. All forms of spontaneous interaction are much more laborious. To that extent Morin is right: We cannot do everything in written English that we can do in spoken English (but also vice versa: Most of us would find it difficult to produce research papers without recourse to language in its written form). These constraints on interaction, Morin suggests, explain why literacy has to be scaffolded by voice or sign, and the same will be true of less specialised ideographic codes.

The upshot is that ideographic, language-independent general purpose systems are not impossible in principle. But their acquisition would have to be scaffolded, probably by both spoken and written forms of language. Acquiring the ability to read Informator analyses required very little investment: The sign menu is small, much of it is iconic, and sign strings are just specifications of chess moves. Acquiring a general purpose system would be very demanding: The elementary sign menu would have to be very large, iconicity would offer less help, and the combination rules would have to enable the system to approach the expressive lower of language. The up-front investment looks heavy. Why would anyone pay? A language-independent general purpose ideographic system allows users to communicate across space and time, but literacy in a shared language already allows this. Although such a system also allows communication without a shared language, the first mover burden is crippling. If a community of users already existed, it might be worth paying to join, but until then, it looks more rational to invest in learning an extra language: At least one knows, for example, that others can read French. Morin's insight into the economics of ideographic codes is independent of his eccentric analysis of literacy, and does the explanatory work he needs.

Financial support. This study was supported by Australian Research Grants Council DP210102513.

Competing interest. None.

Reference

Levinson, S. C. (2019). Interactional foundations of language: The interaction engine hypothesis. In P. Hagoort (Ed.), *Human language: From genes and brain to behavior* (pp. 189–200). MIT Press.

The different paths to cultural convergence

Larissa Mendoza Straffon, Aliki Papa, Heidi Øhrn and Andrea Bender

Department of Psychosocial Science, SFF Centre for Early Sapiens Behaviour (SapienCE), University of Bergen, Bergen, Norway
larissa.straffon@uib.no;
https://www.uib.no/en/persons/Larissa.Mendoza.Straffon
aliki.papa@uib.no
heidi.ohrn@uib.no
andrea.bender@uib.no

doi:10.1017/S0140525X23000560, e253

Abstract

Morin envisions the adaptive landscape of graphic codes as an unfertile valley where writing rises as an isolated peak that humans managed to reach only on four occasions throughout all of history. By exploring the different paths to cultural convergence, we suggest an alternative landscape occupied by a mountain range of visual art systems. We conclude that graphic communication through visual art worked well enough to render writing contingent but not necessary in most cases.

The very fact that we have been able to read and comment on this target article proves one of Morin's points: That writing, like no other technology before it, has allowed us to encode and transmit precise information across time and space, opening up almost endless possibilities in human communication and learning. One cannot help but wonder with the author why then, despite these enormous benefits, writing systems arose independently only four times in the history of the world. Furthermore, how come that a fully autonomous ideographic code is yet to be developed? This is what Morin calls the "puzzle of ideography" (target article, sect. 1, para. 4). We are enthusiastic about his masterly review of the topic. Nevertheless, we will attempt to recast his "puzzle" as a question of cultural evolutionary convergence, in a reformulation that leaves room to explore alternative answers to those questions.

Convergent cultural evolution may be generally defined as the independent emergence or invention of similar cultural traits in multiple societies. Design spaces, such as the adaptive landscape of peaks and valleys described in the target article, can indeed help visualize convergence (Peregrine, 2018). Because the problem of asynchronous communication was solved through writing only in four occurrences, Morin depicts writing as an isolated peak at the end of a single arduous path in the adaptive landscape of graphic codes. In this scenario, writing would be akin to other well-known cultural convergence examples, like the bow-and-arrow, which show that unrelated groups arrive at analogous solutions to common challenges (Carignani, 2016). However, more than one path can lead to evolutionary convergence (Blount, Lenski, & Losos, 2018).

One alternative is that populations may develop similar outcomes, yet as a means to solve different problems. It is possible that this was one of the paths taken toward writing. For instance, cuneiform might have emerged in Mesopotamia as a response to the administrative need of recording goods, whereas Chinese writing seemingly originated as a divinatory aid, each thus occupying a separate niche originally (Schmandt-Besserat & Erard, 2007).

Another path, which will be our main focus, is that populations may settle on a lower adaptive peak, and hence fail to reach the most optimal solution to a problem. As Morin mentions, this could be caused by the availability of "good enough" alternatives. But it could also be as a result of social or material constraints, like taboos or competition with neighboring groups, which may hold back the development of a technology. Let us look at Classic period Mesoamerica, where artistic traditions, cultural practices, and religious beliefs largely converged, but writing did not. Teotihuacan, the most powerful city of the region at the time, never adopted formal writing in spite of being in close contact with the Maya and Zapotec groups who had developed comprehensive writing systems (Langley, 1994). This apparent dismissal of writing has been attributed to religious or political motives (Pasztory, 1994). If correct, this would mean that although the Teotihuacanos were well aware of the most optimal solution for encoding and transmitting information graphically, they settled on their local optimum, dominated by a complex style of visual art and a simple notational system (Langley, 1994). In addition to problem-solving dynamics, we therefore need to consider that historical contingencies related to population size, contact, conflict, or resources also influence the presence or absence of cultural traits (Sterelny, 2016), no less so in the case of writing.

Similar to Teotihuacan, the Andean societies of the Moche and Inca were able to establish and sustain sophisticated states without developing formal writing. Instead, their intricate artistic motifs on various media (ceramics, textiles, sculpture, painting) likely filled the economic and ritual niches that writing initially occupied in Mesopotamia and China respectively (Arnold, 1997), and were seemingly effective enough to allow them to build large empires while remaining at a lower adaptive peak in the landscape of graphic communication.

The previous examples invite revising Morin's vision of the adaptive landscape of graphic codes: Not as towered by the isolated peak of writing, but occupied by multiple peaks of adequate solutions made up by a diversity of visual art and notation systems. Visual narratives were long used across the globe to encode and share complex information asynchronously prior to, or in the absence of literacy (Cohn, 2016; Straffon, 2019). Some properties of visual narratives, such as sequential comprehension, may in fact have been coopted in writing (Cohn, 2019). As the cases above suggest, even though the principles of writing might have been known and understood in more than those four populations in which it emerged, they would have remained un- or underexploited without a social incentive to assume the costs of developing a new technology (Derex, 2021). It seems to us then, that the path to writing as a permanent codified system was constrained not only by transient spoken or sign languages, but also by equally long-lasting forms of visual art.

The few instances in which convergence in writing occurred make it difficult to draw definite conclusions about the specific initial conditions that favored its evolution. In contrast, the many examples of art as an efficient system of asynchronous information exchange even cast doubts on the primacy given to glottographic codes as the supreme mode of visual communication (Brokaw & Mikulska, 2022). It is our contention that when it comes to encoding and transmitting information graphically, visual art has performed sufficiently well throughout human history to render writing highly contingent rather than most optimal in an otherwise diverse adaptive landscape.

Financial support. This work was partly supported by the Research Council of Norway, through its Centres of Excellence funding scheme, SFF Centre for Early Sapiens Behaviour (SapienCE), project number 262618.

Competing interest. None.

References

Arnold, D. Y. (1997). Using ethnography to unravel different kinds of knowledge in the Andes. *Journal of Latin American Cultural Studies*, *6*(1), 33–50.

Blount, Z. D., Lenski, R. E., & Losos, J. B. (2018). Contingency and determinism in evolution: Replaying life's tape. *Science*, *362*(6415), eaam5979.

Brokaw, G., & Mikulska, K. (2022). Homo alphabeticus, la definición de la escritura y las escrituras Mesoamericanas. *Revista Espanola de Antropologia Americana*, *52*(2), 201–224.

Carignani, G. (2016). On the origin of technologies: The invention and evolution of the bow-and-arrow. In F. Panebianco & E. Serrelli (Eds.) *Understanding cultural traits* (pp. 315–339). Springer.

Cohn, N. (2016). Interdisciplinary approaches to visual narrative. In N. Cohn (Ed.), *The visual narrative reader* (pp. 1–15). Bloomsbury.

Cohn, N. (2019). Visual narratives and the mind: Comprehension, cognition, and learning. In *Psychology of learning and motivation* (Vol. 70, pp. 97–127). Academic Press.

Derex, M. (2021). Human cumulative culture and the exploitation of natural phenomena. *Philosophical Transactions of the Royal Society B*, *377*(1843), 20200311. https://doi.org/10.1098/rstb.2020.0311

Langley, J. C. (1994). Symbols, signs, and writing systems. In K. Berrin & E. Pasztory (Eds.), *Teotihuacan: Art from the city of the gods* (pp. 128–139). Thames & Hudson, The Fine Arts Museums of San Francisco.

Pasztory, E. (1994). Teotihuacan unmasked: A view through art. In K. Berrin & E. Pasztory (Eds.), *Teotihuacan: Art from the city of the gods* (pp. 44–63). Thames & Hudson, The Fine Arts Museums of San Francisco.

Peregrine, P. (2018). Toward a theory of recurrent social formations. In J. A. Sabloff & P. L. W. Sabloff (Eds.), *The emergence of premodern states: New perspectives on the development of complex societies* (Vol. 2, pp. 271–295). Santa Fe Institute Press.

Schmandt-Besserat, D., & Erard, D. (2007). Origins and forms of writing. In C. Bazerman (Ed.), *Handbook of research on writing* (pp. 7–22). Routledge.

Sterelny, K. (2016). Contingency and history. *Philosophy of Science*, *83*(4), 521–539.

Straffon, L. M. (2019). Evolution and the origins of visual art: An archaeological perspective. In A. Prentiss (Ed.), *Handbook of evolutionary research in archaeology* (pp. 407–435). Springer.

Fractals and artificial intelligence to decrypt ideography and understand the evolution of language

Cédric Sueur[a,b] and Marie Pelé[c]

[a]Université de Strasbourg, CNRS, IPHC UMR 7178, Strasbourg, France; [b]Institut Universitaire de France, Paris, France and [c]ANTHROPO-LAB, ETHICS EA7446, Université Catholique de Lille, Lille, France
cedric.sueur@iphc.cnrs.fr; marie.pele@univ-catholille.fr
https://sites.google.com/site/cedricsueuranimalbehaviour/; www.ethobiosciences.com

doi:10.1017/S0140525X23000808, e254

Abstract

Self-sufficient ideographies are rare because they are stifled by the issue of standardization. Similar issues arise with abstract art or drawings created by young children or great apes. We propose that mathematical indices and artificial intelligence can help us decode ideography, and if not to understand its meaning, at least to know that meaning exists.

Morin stipulates that self-sufficient ideographies that are understandable to others are rare because they are stifled by the issue of standardization and need to be explained through other means. We agree with this point and argue that mathematical indices and artificial intelligence can assist us in decoding ideography, and if not to understand its meaning, at least to know that meaning exists. Issues such as a lack of representativeness (Martinet & Pelé, 2021) by an external viewer arise, of course, with abstract art and graphic productions created by young children who are not yet able to speak and explain their art. Moreover, examining the evolution of ideographies, which may be a first language before writing, is difficult because no other species, except for *Homo sapiens*, exist in the *Homo* genus. However, it is possible to mathematically study the ontogeny of ideographies, the scribblings, drawings, or sketches of children and its phylogeny through apes that have the ability to draw.

Morin refers to Ted Chiang's novel, *Story of Your Life*, in which aliens draw spherical stains or semagrams. Linguists and other scientists have analyzed this alien language using what appear to be network analyses and mathematical tools that were developed to understand how words are linked in the human language (Batagelj, Mrvar, & Zaversnik, 2002; Cong & Liu, 2014), and how animal sounds are linked in vocal sequences (Allen, Garland, Dunlop, & Noad, 2019; Weiss, Hultsch, Adam, Scharff, & Kipper, 2014). These studies show that the sequences of vocalizations are not random and that some calls are similar to logical connectors in human language because they are central to the vocalizations network. The linguistic approach adopted in human communication can be used to better understand child or animal vocalizations, writings, or drawings. Fractals in mathematics can be used to comprehend the complexity of physics, economy, and urbanism, as well as the complexity of animal behavior in terms of temporal sequences and spatial distribution as with Lévy flights (MacIntosh, 2014). We applied these indices for the first time to drawings created on touchscreens by chimpanzees, children, and adult humans (Beltzung et al., 2022a; Martinet et al., 2021). We analyzed the trajectories of finger drawings as animals moving in nature, foraging, resting, and so on. We demonstrated that chimpanzees do not draw randomly but have a lower efficiency than 3-year-old human children. The efficiency of drawings (measured with a spatial fractal index based on trajectories) increases from 3-year-old to 10-year-old children before reaching a plateau during adulthood because adults add too many details to their graphical productions. Regarding ideography, this result suggests that the drawings that are understandable at a glance are those with few details as sketches and that there may be a kind of selection of codes. The temporal complexity of drawings (based on sequences of drawings and nondrawings) follows a similar pattern.

Combining these fractal indices with other measures, such as the entropy or Gini index, we differentiated the scribbles of children from those of adults; it was not possible to discern these differences with the human eye (Sueur, Martinet, Beltzung, & Pelé, 2022b). Three interpretable dimensions were highlighted in the drawings: Efficiency, diversity, and sequentiality. This means that by observing the coordinates of graphical communication, a drawing, sketch, or semagram, on a three-dimensional graph, we could precisely determine whether the production has meaning and potentially link it to the age, or even the country, of the human creator. Efficiency can be defined as an ability to avoid wasting materials, energy, efforts, and so on in an activity. In the context of graphical communication, we suggest that efficiency is representativeness combined with few details (i.e., optimality). Diversity is almost exclusively linked to the use of colors, whereas sequentiality is linked to graphical complexity. When the complexity of the drawing is increased to make something representative, the number of sequences and the stochasticity increase. Of course there is an interpretation bias in these studies linked to anthropomorphism and the fact that the graphical productions were created by hand (Sueur, Beltzung, & Pelé, 2022a). Extending this work to orangutans (Pelé et al., 2021), we showed that a female orangutan exhibited a higher graphical complexity than her congeners, and also changed the colors and shapes of drawings (circle, triangles, or fan patterns) according to the seasons/context, which indicated that there was meaning in her drawings. For instance, the subject used red extensively after the birth of an orangutan baby and blue and yellow after the visit of schoolchildren who were wearing yellow hats and blue coats. This study's findings were verified using deep-learning models (Beltzung, Pelé, Renoult, Shimada, & Sueur, 2022b). The convolutional neural networks used in orangutan drawings showed a seasonal effect on the style and content of the productions.

These results indicate that the puzzle of ideography can be extended to the sketches or drawings of children and great apes. The premises of semagrams can already encode meaning without requiring language. Our contribution proposes to go deeper into the ontogeny and phylogeny of communication. Mathematics can decode many nonunderstandable codes for human senses and cognition, and the methodology we applied to understand representativeness in apes' and children's graphical productions can be extended to domains such as comparative psychology, anthropology, and semiotics. We believe that we can create elements and research frameworks for the future of ideography using new technologies, analytics, and support. New technologies combined with new mathematical methods appear to be extremely useful and provide new possibilities to test mental states, intentions, and emotions beyond graphical representations (Watanabe & Kuczaj, 2012). Touchscreens score the coordinates of each drawing point, eye-tracking measures anticipation, and encephalograms or magnetic resonance imaging (MRI) can detect whether the communication is indeed a form of language. Finally, our research framework can be extended to psychopathologies such as autism (Jolley, O'Kelly, Barlow, & Jarrold, 2013) or simply be used to measure learning processes and creativity in language or writing (Lee & Hobson, 2006).

Financial support. This work was supported by an interdisciplinary grant from the CNRS (French National Scientific Research Center, MITI) and an IDEX exploratory research grant from the University of Strasbourg.

Competing interest. None.

References

Allen, J. A., Garland, E. C., Dunlop, R. A., & Noad, M. J. (2019). Network analysis reveals underlying syntactic features in a vocally learnt mammalian display, humpback whale song. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1917), 20192014.

Batagelj, V., Mrvar, A., & Zaversnik, M. (2002). *Network analysis of texts*. University of Ljubljana, Institute of Mathematics, Physics and Mechanics.

Beltzung, B., Martinet, L., MacIntosh, A., Meyer, X., Hosselet, J., Pele, M., & Sueur, C. (2022a). To draw or not to draw: Understanding the temporal organization of drawing behavior using fractal analyses. *Fractals*, 2350009. doi:10.1101/2021.08.29.458053.

Beltzung, B., Pelé, M., Renoult, J. P., Shimada, M., & Sueur, C. (2022b). Using artificial intelligence to analyze non-human drawings: A first step with orangutan productions. *Animals*, *12*(20), 2761.

Cong, J., & Liu, H. (2014). Approaching human language with complex networks. *Physics of Life Reviews*, *11*(4), 598–618.

Jolley, R. P., O'Kelly, R., Barlow, C. M., & Jarrold, C. (2013). Expressive drawing ability in children with autism. *British Journal of Developmental Psychology*, *31*(1), 143–149. doi:10.1111/bjdp.12008

Lee, A., & Hobson, R. P. (2006). Drawing self and others: How do children with autism differ from those with learning difficulties? *British Journal of Developmental Psychology*, *24*(3), 547–565. doi:10.1348/026151005X49881

MacIntosh, A. (2014). The fractal primate: Interdisciplinary science and the math behind the monkey. *Primate Research*, *30*(1), 95–119. doi: 10.2354/psj.30.011.

Martinet, L., & Pelé, M. (2021). Drawing in nonhuman primates: What we know and what remains to be investigated. *Journal of Comparative Psychology*, *135*(2), 176–184. doi:10.1037/com0000251

Martinet, L., Sueur, C., Hirata, S., Hosselet, J., Matsuzawa, T., & Pelé, M. (2021). New indices to characterize drawing behavior in humans (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Scientific Reports*, *11*(1), 1–14. doi:10.1038/s41598-021-83043-0

Pelé, M., Thomas, G., Liénard, A., Eguchi, N., Shimada, M., & Sueur, C. (2021). I wanna draw like you: Inter- and intra-individual differences in orangutan drawings. *Animals*, *11*(11), 11. doi:10.3390/ani11113202

Sueur, C., Beltzung, B., & Pelé, M. (2022a). Vers une création picturale non humaine: Quel degré d'auteurisation pour les animaux et les machines? *Polygraphe(s)*, *4*, 45–60.

Sueur, C., Martinet, L., Beltzung, B., & Pelé, M. (2022b). Making drawings speak through mathematical metrics. *Human Nature*. Retrieved from http://arxiv.org/abs/2109.02276
Watanabe, S., & Kuczaj, S. (2012). *Emotions of animals and humans: Comparative perspectives*. Springer Science & Business Media.
Weiss, M., Hultsch, H., Adam, I., Scharff, C., & Kipper, S. (2014). The use of network analysis to study complex animal communication systems: A study on nightingale song. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1785), 20140460.

The design space of human communication and the nonevolution of ideography

Walter Veit[a] and Heather Browning[b]

[a]Department of Philosophy, University of Bristol, Bristol, UK and
[b]Department of Philosophy, University of Southampton, Southampton, UK
wrwveit@gmail.com; https://walterveit.com/
DrHeatherBrowning@gmail.com; https://www.heatherbrowning.net/

doi:10.1017/S0140525X23000754, e255

Abstract

Despite the once-common idea that a universal ideography would have numerous advantages, attempts to develop such ideographies have failed. Here, we make use of the biological idea of fitness landscapes to help us understand the nonevolution of such a universal ideographic code as well as how we might reach this potential global fitness peak in the design space.

Universal ideographies – graphic languages in which symbols encode conceptual rather than linguistic content – hold many apparent advantages, such as transmission of information across time and space, operating across language barriers, and the potential iconicity of symbols increasing ease of learning. Despite this, there are no successful examples of such ideographic communication systems. Morin's proposed solution to this "puzzle of ideography" is to explain their absence as resulting from a standardization problem, with such systems suffering from the inherent challenges raised through the need for everyone to use the same meaning-to-symbol mappings. This becomes ever more difficult as the number of symbols increases and thus restricts them to narrow domains. Here we aim to further advance Morin's suggestion that the nonevolution of ideography is largely a result of spoken (or signed) languages having been "locked-in" earlier because of their easier standardization, to the detriment of other codes. We do so through use of the concept of the "fitness landscape," which can be borrowed from its biological context to aid in understanding the nonevolution of "bad" solutions to cultural problems.

At the very end of his article, Morin notes that a "complete ideography could be seen as a peak in the design space of graphic codes (Acerbi, Tennie, & Mesoudi, 2016; Dennett, 1995; Mesoudi & Thornton, 2018)" (target article, sect. 7, para. 3). This type of thinking about cultural artefacts in terms of a "design space" inspired by the notion of fitness landscapes has proven highly useful in the past, and we wish to explore the suggestion further here, particularly in relation to the "lock-in dynamics" (target article, sect. 7, para. 4) Morin discusses. Wright's (1932) fitness landscapes posit that we can model the relative fitness of different phenotypes as a "landscape" across which there are fitness "peaks" where organisms are doing as well as possible within the "local" set of possible phenotypes, and "valleys" in which they would be doing very poorly. They provide a useful tool for thinking about why some species appear to be "stuck" in suboptimal solutions to their ecological problems, with the path towards a higher peak involving passing through a fitness valley, requiring the organism to become temporarily less fit than others in the population, and thus often blocking the path towards better solutions. Similarly, cultural innovations such as communication systems may be stuck at a local fitness peak in the design space with no way to move to a better system (the global optimum) because any individual shifting their strategy would be initially worse off, through the high costs of learning a new system, and inability to communicate with others in the community.

The effects of standardization that Morin describes may very well be the reasons for the existence of fitness "valleys" that prevent the development of ideographic communication. This is in line with another example Morin raises – that of the lock-in of the QWERTY keyboard which, as Morin points out, is now commonly regarded as less quick or efficient than other keyboard arrangements (David, 1985; David & Rothwell, 1996). However, its early adoption has led to it becoming a local fitness peak, where movement to another (perhaps higher) peak carries the cost of having to temporarily move across a lower space in the fitness landscape.

One common criticism of using the model of fitness landscapes is that, as they are typically presented, they are static and fixed. However, this is of course only an idealization and one that has been frequently criticized (Kaplan, 2008) – not a necessary feature of the model. It is entirely possible and now common to construct dynamic fitness landscapes that represent changing conditions. For example, as environmental conditions change, a strategy or technology that was once the most optimal might turn instead from a fitness peak into a fitness valley. The more rapid pace of cultural change makes this model even more plausible for cultural fitness landscapes.

Thinking about a dynamic design space allows us to explore the technological and societal changes that may be required to create slopes or neutral ridges that would shift agents towards the alternative peak of a universal ideography. As Morin has argued, spoken language has restricted us from exploring alternative strategies and here we may find ways to promote the advantages of ideographic communication. This requires acknowledgement of the difficulties facing such a change. As Morin proposes that the cultural "fitness" (target article, sect. 7, para. 4) of different communication systems is largely driven by standardization of conventions between users, this will thus be a key issue for improving the design of ideographic communication systems. For instance, network effects make languages more useful the more people use them and thus force standardization between users. This implies that the only way to make ideography viable is to improve it through use of new means.

Here, as Morin also suggests in the conclusion to his article, we think that use of new technologies provides an opportunity. In particular, online communication provides many of the benefits Morin attributes to face-to-face spoken and signed communication – signals are cheap, (semi-)transient, and there is opportunity to repair miscommunication. Indeed, this has already brought us quite a long way – think of the standardization of emojis across platforms. Although Morin is right to point out that there is still disagreement about the meaning of emojis, we think he underestimates how standardized their usage already is, especially among those populations that use them most frequently and have grown up with them. The differences in use occur most often between cohorts, not within them. This then suggests that we might be on our way towards the elimination of this ambiguity, or at least for it to be diminished, to the same extent as there is persisting acceptable ambiguity in spoken languages. Standardization of meaning does not have to imply universal agreement. We suggest that changes in communication technology may sufficiently alter the fitness landscape to make the peak of a general ideography accessible, but that more work would be needed to refine the model and test the predictions.

Financial support. This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number 101018533).

Competing interest. None.

References

Acerbi, A., Tennie, C., & Mesoudi, A. (2016). Social learning solves the problem of narrow-peaked search landscapes: Experimental evidence in humans. *Royal Society Open Science*, *3*(9), 160215. https://doi.org/10.1098/rsos.160215

David, P. A. (1985). Clio and the economics of QWERTY. *The American Economic Review*, *75*(2), 332–337.

David, P., & Rothwell, G. (1996). Standardization, diversity and learning: Strategies for the coevolution of technology and industrial capacity. *International Journal of Industrial Organization*, *14*(2), 181–201.

Dennett, D. C. (1995). *Darwin's dangerous idea: Evolution and the meaning of life*. Simon & Schuster.

Kaplan, J. (2008). The end of the adaptive landscape metaphor? *Biology & Philosophy*, *23*(5), 625–638.

Mesoudi, A., & Thornton, A. (2018). What is cumulative cultural evolution? *Proceedings of the Royal Society B: Biological Sciences*, *285*(1880), 20180712. https://doi.org/10.1098/rspb.2018.0712

Wright, S. (1932). The roles of mutation, inbreeding, crossbreeding, and selection in evolution. *Proceedings of the Sixth International Congress on Genetics*, *1*(8), 355–366.

Graphic codes, language, and the computational niche

James Winters

School of Collective Intelligence, Université Mohammed VI Polytechnique, Rabat, Morocco
James.Winters@um6p.ma; https://j-winters.github.io/

doi:10.1017/S0140525X23000614, e256

Abstract

Human language looms large in the emergence and evolution of graphic codes. Here, I argue that language not only acts as a strong constraint on graphic codes, but it is also a precondition for their emergence and their evolution as computational devices.

Graphic codes are ultimately a collection of human technologies that serve a computational role: To store, transmit, and process information across space and time. In this respect, the emergence and evolution of graphic codes is (partly) a story of how humans have continually optimized and expanded the computational resources at our disposal. Thinking of graphic codes as occupying a computational niche helps enrich Morin's general argument in two ways. First, because of the presence of language, which itself is a powerful computational system for thinking and communication, we should expect graphic codes to fill in functional gaps in the storage, transmission, and processing of information. Second, any expansion of graphic codes is dependent on language, which serves as a strong constraint on the emergence and evolution of such codes.

As Morin aptly put it, language acts as an "oral crutch" that "prevents graphic codes from learning to walk" (target article, sect. 6.3, para. 4). This is evident in the evolution of writing that was initially restricted to transcribing proper names (Morin, 2022). The latent potential of writing, as both a general-purpose glottography and its use as an asynchronous communication device, evolves centuries after its invention (Morin, 2022; Morin, Kelly, & Winters, 2020) and illustrates how language acts as a strong constraint: It is not immediately obvious that a general-purpose glottography is useful when oral language already exists, and it is only when this functionality is discovered that asynchronous communication is distinctly advantageous. However, although spoken and signed languages, because of the ease by which they are standardized, constrain and delay the emergence of sophisticated graphic codes, such as writing, there is a case to be made that language is also an important enabling condition.

One possibility, which was absent in Morin's target article, is that language lowers the barrier for graphic codes to emerge in the first place. A tally system, for instance, is far easier to invent and disseminate in a species where language is the basis for communication and learning. This is possible because: (1) The expressive power of language allows graphic codes to be massively underspecified and (2) language serves as the basis by which humans acquire knowledge of how to use the code. By filling in gaps in inference and interpretation, language makes it possible for simple graphic codes to exist by enriching the context in which these codes are learned and used. Moreover, the use of language as a pedagogical tool helps explain how graphic codes can rapidly spread and become standardized in a community. It is, of course, possible to envisage graphic codes that emerge and are standardized through observation and other nonlinguistic behaviours. However, it is telling that we do not observe even rudimentary graphical notation in nonhuman animals – simple graphic codes appear out of reach for the inventive capabilities of most species. In cases where we do observe the use of graphic codes in nonhuman species, such as Kanzi and his lexigrams (Rumbaugh et al., 1973), the underlying systems are invented by humans.

A similar argument can be made for the impact of writing on the emergence of subsequent graphic codes. The standardization account can point to why powerful and specialized graphic codes, such as rich systems of mathematical and musical notation, are difficult to discover without writing. A world in which writing has been invented, and serves as a coordination device in a population, makes it far easier for individuals to invent, standardize, and learn novel graphic codes. Crucially, it is the ability to communicate general-purpose information asynchronously, which lowers the barrier for our modern systems of mathematical and musical notation to exist. This leaves us with a key unanswered question: Are such systems likely to emerge in a counterfactual world where writing was never invented?

Lastly, if writing is adapted to exploit and expand the computational niche in which it is situated, then this helps explain why a fully fledged ideography is unlikely: A richly structured ideographic system is unnecessary in a world where writing exists.

Writing is eminently more learnable than an ideography, largely because of its parasitic relationship with language, and it fulfils all of the same functional roles in the computational niche as a hypothetical ideography. In fact, as Morin highlights, the key advantage of an ideography is its independence from language, which, in principle, would mean an individual who speaks Mandarin could readily communicate with an individual who spoke English or Darija – so long as populations invested time and resources in learning this system. Yet, even in this specific instance of cross-linguistic communication, it seems unlikely an ideography is particularly advantageous, especially in a world where translation technologies are at our disposal. But ¯_(ツ)_/¯ (who knows)?

Competing interest. None.

References

Morin, O. (2022). The piecemeal evolution of writing. *Lingue e Linguaggio* (2), 217–237. https://doi.org/10.31235/osf.io/a6ket

Morin, O., Kelly, P., & Winters, J. (2020). Writing, graphic codes, and asynchronous communication. *Topics in Cognitive Science*, *12*(2), 727–743. https://doi.org/10.1111/tops.12386

Rumbaugh, D. M., Gill, T. V., Brown, J. V., von Glasersfeld, E. C., Pisani, P., Warner, H., & Bell, C. L. (1973). A computer-controlled language training system for investigating the language skills of young apes. *Behavior Research Methods & Instrumentation*, *5*(5), 385–392. https://doi.org/10.3758/BF03200213

On the semiotic and material constraints of ideographies

Izzy Wisher and Kristian Tylén

Department of Linguistics, Cognitive Science and Semiotics, Aarhus University, Aarhus, Denmark
kristian@cc.au.dk
izzywisher@cc.au.dk
https://pure.au.dk/portal/en/persons/kristian-tylen(9950d8bc-e1cd-400a-a547-359cd0b07157).html
https://pure.au.dk/portal/en/persons/isobel-wisher(d0e2fb0e-cb67-4ff2-9786-3c9dff72d291).html

doi:10.1017/S0140525X23000742, e257

Abstract

Despite obvious advantages, no generalised ideographic codes have evolved through cultural evolution to rely on iconicity. Morin suggests that this is because of missing means of standardisation, which glottographic codes get from natural languages. Although we agree, we also point to the important role of the available media, which might support some forms of reference more effectively than others.

Recently, there has been renewed interest in the role of iconicity in human communication (Fay, Ellison, & Garrod, 2014; Perlman, Dale, & Lupyan, 2015). Although arbitrariness has historically been considered a central design feature of language (Hockett, 1960), sound symbolism appears more prevalent than first acknowledged (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Monaghan, Shillcock, Christiansen, & Kirby, 2014). Similarly, new studies on the possible evolutionary roots of language (whether spoken, signed, or written) suggest a stage of iconically grounded reference in, for instance, pantomime, vocalisations, or figurative depiction (Garrod, Fay, Lee, Oberlander, & MacLeod, 2007; Nölle, Staib, Fusaroli, & Tylén, 2028; Perlman & Lupyan, 2018; Zlatev, Żywiczyński, & Wacewicz, 2020).

Indeed, preceding any forms of conventionalised writing, our Palaeolithic ancestors used iconic depictions in parietal and portable art to denote animals, humans, and perhaps even narrative scenes (Aubert et al., 2019). From an ontogenetic perspective, iconicity facilitates early-visual communication as children typically learn to draw before they write, and pictorial narratives appear intuitively accessible even to preschool children (Jolley, 2009). Given these observations, it is puzzling that no generalised ideographic systems have evolved to be fully dependent on iconicity.

Morin suggests that the reason these apparently fundamental forms of human signification never evolve to become generalised ideographies is because of the lack of procedures for standardisation. We agree that any generalised form of communication will depend on the continuous contextualised negotiation of meaning (Dideriksen, Christiansen, Tylén, Dingemanse, & Fusaroli, 2022). However, we also suggest that there might be semiotic factors intrinsic to the materiality of certain media that challenge the evolution of ideographies – in particular, to the extent they depend on iconicity. Heraldic signs, coins, or commodity brands work well in their specialised systems for marking identities (families, cities, values, or products), but they have no means of representing the course of multiple events; they cannot tell a story.

In written language, meaning is built-up as new words are added in a continuous linear succession. Any new detail or event manifests as a spatial prolongation of the text. In an iconic depiction, the addition of new detail will instead correspond to a transformation of the depiction itself (Lotman, 1975). Thus, a purely iconicity-based code faces the choice of three possibilities: (1) Presenting a single-static depiction of a salient moment from unfolding events to support the reader's inferences about immediately preceding and succeeding events; (2) using the linearisation principle from glottography to present a series of depictions representing a succession of events (like a comics strip); or (3) if the medium allows, animating the image to create a continuous transformation of the icon (like a cartoon).

The first option is what we appear to find already within examples of early or contemporary rock art. However, these depictions are often limited to a single event (e.g., a hunting scene; Aubert et al., 2019) or require a preexisting knowledge of the narrative being communicated. In both Australian Aboriginal and San rock art, depicted scenes represent complex narratives related to how the world was created (e.g., the *Dreamtime*; McDonald, 2013; Tacon, 1989) or spiritual encounters and trance states. However, without an understanding of associated mythologies or certain cultural behaviours, the art cannot be decoded. This is highlighted by the misinterpretation of San rock art by Western ethnographers, where poor translations of oral traditions have confused concepts embedded in the art (Challis, Hollmann, & McGranaghan, 2013; McGranaghan & Challis, 2016). Thus, the potential for static depictions to serve as ideographies is limited, despite icons sometimes having standardised elements (e.g., X-ray depictions in Australian art represent living animals, solid-filled depictions represent dead animals; Tacon, 1989). Rather, these iconic depictions can be considered mnemonic devices that do not communicate independently.

The second option, the comics principle, has numerous historical instantiations (e.g., the famous Bayeux Tapestry; Brilliant, 1991). Although it overcomes some of the outlined limitations with respect to representing complex sequences of events and has established itself as a rich genre of graphic storytelling (Cohn, 2013; Stjernfelt & Østergaard, 2013), it has not evolved into conventionalised ideographic codes. Among the reasons, as discussed by Morin, is probably that it resists the compression and standardisation needed to become an economic medium of communication. For instance, comics do not come with a functional system of anaphora by which a character or object mentioned in an earlier scene can be referred to with a shorthand "she" or "it." Rather, reappearing characters are redrawn in new configurations in every panel. Bliss symbols could be considered a solution, combining schematic icons with some arbitrary elements, grammatical categories, and the linearisation principle from verbal language (Nawar, 2012). However, as a kind of "creole," it depends on language-like conventions for composing meaning, bringing back issues of standardisation and the need for an oral gloss.

The third option – the animated cartoon – may be the solution most true to the inherent semiotic nature of iconicity (not piggybacking on discretisation and linearisation principles from language; Lotman, 1975). A hypothetical cartoon language would have several advantages for effective communication. On the receptive side, it would be highly intuitive and accessible, thus overcoming language barriers and possibly requiring minimal formal training (Berney & Bétrancourt, 2016). Although such an ideographic language has not (yet) evolved, it is probably not because of cognitive factors pertaining, for instance, to the architecture of our visual system or working memory. It is also not because of social factors alone, because such a communication system would potentially require fewer conventions to standardise. The main reason is probably instead the lack of a suitable medium that would support such a code, which makes it an unfeasible solution for general communication. If the effortless production, transmission, and reception of cartoons could be effectively supported by available material media, this could hypothetically constitute Morin's missing ideographic code. The historical success of glottographic codes is thus not only a matter of the way it relies on standardisation processes from natural language, but also how it is supported by the available technological solutions and material media that historically has included clay tablets, pen, paper, and print, and is currently evolving with digital media.

Financial support. The authors are supported by an EU ERC consolidator grant (eSYMb, grant number: 101044626), awarded to Kristian Tylén.

Competing interest. None.

References

Aubert, M., Lebe, R., Oktaviana, A. A., Tang, M., Burhan, B., Hamrullah, … Brumm, A. (2019). Earliest hunting scene in prehistoric art. *Nature*, *576*, 442–445.
Berney, S., & Bétrancourt, M. (2016). Does animation enhance learning? A meta-analysis. *Computers & Education*, *101*, 150–167.
Brilliant, R. (1991). The Bayeux Tapestry: A stripped narrative for their eyes and ears. *Word & Image*, *7*(2), 98–126.
Challis, S., Hollmann, J., & McGranaghan, M. (2013). "Rain snakes" from the Senqu River: New light on Qing's commentary on San rock art from Sehonghong, Lesotho. *Azania: Archaeologial Research in Africa*, *48*(3), 331–354.
Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. A&C Black.
Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2022). Quantifying the interplay of conversational devices in building mutual understanding. *Journal of Experimental Psychology: General*, *152*(3), 864–889.
Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19* (10), 603–615.
Fay, N., Ellison, M., & Garrod, S. (2014). Iconicity: From sign to system in human communication and language. *Pragmatics & Cognition*, *22*(2), 244–263.
Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, *31*(6), 961–987.
Hockett, C. F. (1960). The origin of speech. *Scientific American*, *203*(3), 88–97.
Jolley, R. P. (2009). *Children and pictures: Drawing and understanding*. John Wiley.
Lotman, J. M. (1975). The discrete text and the iconic text: Remarks on the structure of narrative (F. Pfotenhauer, Trans.). *New Literary History*, *6*(2), 333–338.
McDonald, J. (2013). Contemporary meanings and the recursive nature of rock art: Dilemmas for a purely archaeological understanding of rock art. *Time and Mind: The Journal of Archaeology, Consciousness and Culture*, *6*(1), 65–72.
McGranaghan, M., & Challis, S. (2016). Reconfiguring hunting magic: Southern Bushman (San) perspectives on taming and their implications for understanding rock art. *Cambridge Archaeological Journal*, *26*(4), 579–599.
Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369* (1651), 20130299.
Nawar, H. (2012). Multicultural transposition: From alphabets to pictographs, towards semantographic communication. *Technoetic Arts*, *10*(1), 59–68.
Nölle, J., Staib, M., Fusaroli, R., & Tylén, K. (2018). The emergence of systematicity: How environmental and communicative factors shape a novel communication system. *Cognition*, *181*, 93–104.
Perlman, M., Dale, R., & Lupyan, G. (2015). Iconicity can ground the creation of vocal symbols. *Royal Society Open Science*, *2*(8), 150152.
Perlman, M., & Lupyan, G. (2018). People can create iconic vocalizations to communicate various meanings to naïve listeners. *Scientific Reports*, *8*(1), 1–14.
Stjernfelt, F., & Østergaard, S. (2013). FONK! HONK! WHAM! OOF!: Representation of events in Carl Barks – and in the aesthetics of comics in general. In N. Pedri & L. Petit (Eds.), *Picturing the language of images* (pp. 483–508). Cambridge Scholars Press.
Tacon, P. S. C. (1989). From the "dreamtime" to the present: The changing role of Aboriginal rock paintings in Western Arnhem Land, Australia. *The Canadian Journal of Native Studies*, *IX*(2), 317–339.
Zlatev, J., Żywiczyński, P., & Wacewicz, S. (2020). Pantomime as the original human-specific communicative system. *Journal of Language Evolution*, *5*(2), 156–174.

Chinese offers a test for universal cognitive processes

Ming Yan[a,b] and Reinhold Kliegl[c]

[a]Department of Psychology, University of Macau, Taipa, Macau; [b]Center for Cognitive and Brain Sciences, University of Macau, Taipa, Macau and [c]Department of Sport and Health Sciences, University of Potsdam, Potsdam, Germany
mingyan@um.edu.mo; https://www.um.edu.mo/fss/psychology/staff_yanming.html
reinhold.kliegl@uni-potsdam.de; https://www.uni-potsdam.de/en/trainingswissenschaft/staff/rkliegl

doi:10.1017/S0140525X23000663, e258

Abstract

The Chinese writing system is unique in its implementation of graphemic, phonological, morphological, and semantic features. We add nuances to its portrait in the target article and highlight research on radically different timelines of phonological and semantic activation during reading of Chinese and alphabetic script, paving the way for the identification of universal and culture-specific cognitive processing.

Because of their ideographic nature, some visual forms of Chinese characters resemble the concepts they represent for trained readers. In addition, Chinese encodes a natural spoken language just as alphabetic scripts do. We review the graphemic, phonological, morphological, and semantic properties of the Chinese writing system to add some nuance to Morin's arguments and to convey a clearer understanding of the unique features of the language. We focus on language-universal and -specific cognitive processes during Chinese reading with reference to recent psycholinguistic research and conclude that, because of its fundamental difference from alphabetic scripts with respect to the relations between orthography, phonology, semantics, and morphology, Chinese offers unique opportunities to test theoretical accounts based on cross-language comparisons of scripts.

In the Western point of view, Chinese typically uses pictures for concepts. For example, some characters are of a more "ideographic" nature, largely recognizable even by untrained eyes (e.g., 山 for mountain and 田 for farmland). These characters include the so-called pictograms, simple ideograms, and compound ideographs. Presumably, they are derived from the earliest character forms, often found on ancient artifacts such as ox bones and turtle shells. Mostly, these characters are visually simple and represent common and concrete concepts, therefore, they are optimized for fast semantic access. Indeed, the first experimental evidence for parafoveal semantic processing was based on such characters (Yan, Richter, Shu, & Kliegl, 2009). However, these characters are rare and not representative of the language. The majority (i.e., over 80%) of modern Chinese characters are compound phonograms (DeFrancis, 1989), which are comparatively less "ideographic." These compound phonograms include two or more "mini-characters," called radicals. A semantic radical often transparently indicates the meaning category of the whole character, whereas a phonetic radical merely provides an unreliable clue to its pronunciation (Shu, Chen, Anderson, Wu, & Xuan, 2003; Zhou, 1980). Thus, Chinese characters are formed according to different principles and vary in their degrees of ideography, but generally provide no clues to their pronunciations, or clues that are unreliable, indicating that the language prioritizes semantic knowledge (Hoosain, 1991).

Written Chinese was not created initially to encode a natural spoken language. The first nationally standardized Chinese character set, the small seal script, was promulgated during the Qin dynasty in about 220 BC (Diringer, 1982). By then, bamboo boards were the primary media for carrying text, implying that the actual speech flow must have been abstracted before the costly process of carving down characters. For this reason, Chinese (and Japanese kanji) characters were designed to convey clear meanings and ideas, whereas phonological distinguishability was not a major concern. Traces of such a characteristic can still be found in modern Chinese: A high degree of homophony in Chinese makes it a terrible language in which to encode pronunciation, as a syllable by itself can be morpho-semantically ambiguous. For instance, Lion-Eating Poet in the Stone Den, a famous passage composed of 94 homophonic characters all pronounced as /shi/ in Mandarin Chinese, illustrates that phonology does not necessarily mediate semantic access in Chinese; the passage is fully intelligible in print but incomprehensible when read aloud. Relatedly, pronunciations of Chinese characters are vulnerable, with large variations across time and regions. The same character can have very different pronunciations in different dialects and languages; therefore, it is not orally communicable. In contrast, the orthographic forms and their meanings are highly standardized, so that cross-dialect communication via printed media is largely feasible. To some degree, even cross-language communication using written Chinese, such as between Chinese and Japanese, is possible. Indeed, some studies have demonstrated that morphological awareness (i.e., the ability to disambiguate homographic/homophonic characters) plays a more important role than phonological awareness (i.e., the ability to be aware of and to manipulate phonological structures) for literacy development in Chinese (McBride-Chang et al., 2005).

The linguistics-related features reviewed above imply that phonological and semantic activation in Chinese reading can be rather unique, as also pointed out in the target article. In alphabetic scripts, the activation of phonological representation often temporally precedes that of semantic properties (e.g., Frost, 1998; Perfetti & Bell, 1991; Van Orden, 1987). In contrast, although there is phonological activation in Chinese, its role during lexical access is still under debate. Although some studies suggest phonology is activated earlier than semantics (e.g., Tan, Hoosain, & Peng, 1995), many favor a late phonological activation. For example, in isolated word recognition, phonological priming effects suggest a late phonological activation among Chinese adults (Chen & Shu, 2001; Zhou & Marslen-Wilson, 1999, 2000; Zhou, Marslen-Wilson, Taft, & Shu, 1999). Thus, semantic access may not depend on phonological activation in Chinese. There is also a theoretical debate about parafoveal lexical processing during natural sentence reading. Based on fixation durations measured in the gaze-contingent boundary paradigm (Rayner, 1975), several studies concluded that semantics is more important than phonology, that earlier and larger parafoveal semantic than phonological priming effects were found for Chinese adults during silent reading (e.g., Pan, Yan, & Yeh, 2022; Tsai, Kliegl, & Yan, 2012; Yan et al., 2009).

The theoretical aim of exploring lexical processing during reading of Chinese is not only the specific script, but a language-universal model of reading (Frost, 2012). For instance, for decades there was no evidence for parafoveal semantic effects for reading of English before the first report based on Chinese reading appeared. The findings of these Chinese studies then led to numerous extensions in alphabetic scripts, yielding positive evidence in German (Hohenstein & Kliegl, 2014; Hohenstein, Laubrock, & Kliegl, 2010), Korean (Kim, Radach, & Vorstius, 2012; Yan, Wang, Song, & Kliegl, 2019), and English (Schotter, 2013; Veldre & Andrews, 2016). These cross-language findings then, in turn, inspired new developments of computational models of reading to capture parafoveal effects and other individual differences such as those related to aging (e.g., E-Z Reader: McGowan & Reichle, 2018, and SWIFT: Laubrock, Kliegl, & Engbert, 2006; Risse & Kliegl, 2011). Cross-language comparative results, revealing differences in the temporal dynamics at which various types of information become available, might provide additional constraints to test theoretical distinctions such as the learning and standardization accounts contrasted in the target article.

Financial support. The work was supported by a FDCT grant from the Macao Science and Technology Development Fund (Project code: 0015/2021/ITP).

Competing interest. None.

References

Chen, H.-C., & Shu, H. (2001). Lexical activation during the recognition of Chinese characters: Evidence against early phonological activation. *Psychonomic Bulletin & Review*, *8*(3), 511–518. https://doi.org/10.3758/BF03196186

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. University of Hawai'i Press.

Diringer, D. (1982). *The book before printing: Ancient, medieval, and oriental*. Courier Dover.
Frost, R. (1998). Toward a strong phonological theory of visual word recognition: True issues and false trails. *Psychological Bulletin*, *123*, 71–99. https://doi.org/10.1037/0033-2909.123.1.71
Frost, R. (2012). Towards a universal model of reading. *Behavioral and Brain Sciences*, *35* (5), 263–279. https://doi.org/10.1017/S0140525X11001841
Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 166–190. https://doi.org/10.1037/a0033670
Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology: Learning Memory and Cognition*, *36*, 1150–1170. https://doi.org/10.1037/a0020233
Hoosain, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*. Lea.
Kim, Y.-S., Radach, R., & Vorstius, C. (2012). Eye movements and parafoveal processing during reading in Korean. *Reading and Writing: An Interdisciplinary Journal*, *25*(5), 1053–1078. https://doi.org/10.1007/s11145-011-9349-0
Laubrock, J., Kliegl, R., & Engbert, R. (2006). SWIFT explorations of age differences in eye movements during reading. *Neuroscience & Biobehavioral Reviews*, *30*(6), 872–884. https://doi.org/10.1016/j.neubiorev.2006.06.013
McBride-Chang, C., Cho, J.-R., Liu, H., Wagner, R. K., Shu, H., Zhou, A., … Muse, A. (2005). Changing models across cultures: Associations of phonological awareness and morphological structure awareness with vocabulary and word recognition in second graders from Beijing, Hong Kong, Korea, and the United States. *Journal of Experimental Child Psychology*, *92*(2), 140–160. https://doi.org/10.1016/j.jecp.2005.03.009
McGowan, V. A., & Reichle, E. D. (2018). The "risky" reading strategy revisited: New simulations using E-Z Reader. *Quarterly Journal of Experimental Psychology*, *71*, 179–189. https://doi.org/10.1080/17470218.2017.1307424
Pan, J., Yan, M., & Yeh, S.-L. (2022). Accessing semantic information from above: Parafoveal processing during the reading of vertically presented sentences in traditional Chinese. *Cognitive Science*, *46*(2), e13104. https://doi.org/10.1111/cogs.13104
Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory and Language*, *30*, 473–485. https://doi.org/10.1016/0749-596X(91)90017-E
Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*, 65–81. https://doi.org/10.1016/0010-0285(75)90005-5
Risse, S., & Kliegl, R. (2011). Adult age differences in the perceptual span during reading. *Psychology and Aging*, *26* (2), 451–460. hhtps://doi.org/10.1037/a0021616
Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language*, *69*(4), 619–633. https://doi.org/10.1016/j.jml.2013.09.002
Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development*, *74*(1), 27–47. https://doi.org/10.1111/1467-8624.00519
Tan, L., Hoosain, R., & Peng, D. (1995). Role of early presemantic phonological code in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(1), 43–54. https://doi.org/10.1037/0278-7393.21.1.43
Tsai, J.-L., Kliegl, R., & Yan, M. (2012). Parafoveal semantic information extraction in traditional Chinese reading. *Acta Psychologica*, *141*, 17–23. https://doi.org/10.1016/j.actpsy.2012.06.004
Van Orden, G. C. (1987). A rows is a rose: Spelling, sound, and reading. *Memory & Cognition*, *15*, 181–198. https://doi.org/10.3758/BF03197716
Veldre, A., & Andrews, S. (2016). Is semantic preview benefit due to relatedness or plausibility? *Journal of Experimental Psychology: Human Perception and Performance*, *42* (7), 939–952. https://doi.org/10.1037/xhp0000200
Yan, M., Richter, E. M., Shu, H., & Kliegl, R. (2009). Chinese Readers extract semantic information from parafoveal words during reading. *Psychonomic Bulletin & Review*, *16*, 561–566. https://doi.org/10.3758/PBR.16.3.561
Yan, M., Wang, A., Song, H., & Kliegl, R. (2019). Parafoveal processing of phonology and semantics during the reading of Korean sentences. *Cognition*, *193*, 104009. https://doi.org/10.1016/j.cognition.2019.104009
Zhou, X., & Marslen-Wilson, W. (1999). Phonology, orthography, and semantic activation in reading Chinese. *Journal of Memory and Language*, *41*, 579–606. https://doi.org/10.1006/jmla.1999.2663
Zhou, X., & Marslen-Wilson, W. (2000). The relative time course of semantic and phonological activation in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1245–1265. https://doi.org/10.1037/0278-7393.26.5.1245
Zhou, X., Marslen-Wilson, W., Taft, M., & Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese. *Language and Cognitive Processes*, *14*, 525–565. https://doi.org/10.1080/016909699386185
Zhou, Y. (1980). 汉字声旁读音便查 *[Pronunciations of phonetics within compound characters]*. Jilin People's Publisher. (Original work in Chinese).

The stranding of the ideography: A nonnegligible role of the spoken language

Bohua Zhang[a], Xueping Hu[b], Qing Li[c] and Antao Chen[a]

[a]School of Psychology, Shanghai University of Sport, Shanghai, China; [b]College of Education, Huaibei Normal University, Huaibei, Anhui, China and [c]Faculty of Psychology, Southwest University, Chongqing, China
2426711673@qq.com; huxpxp@163.com; fightingliqing@126.com; chenantao@sus.edu.cn

doi:10.1017/S0140525X2300064X, e259

Abstract

Morin suggested that one of the reasons for the difficulty in standardizing graphic codes is that the production of spoken language reduces the need for graphic codes. Here we try to extend their claims from a psychological perspective, which allows us to conclude that the puzzle of ideography is perhaps related to human psychological traits and psychological evolution.

Morin's standardization hypothesis is important for understanding the puzzle of ideography, but we suggest that understanding the ideographic puzzle, from psychological perspectives, has great potential value. Because people can be advantageous on learning, memory, and survival if spoken languages can be combined with graphic codes, people may prefer to combine spoken language with graphic codes and the target article should strengthen this part of the discussion. In this commentary, the effects that spoken languages have on graphic codes would be further discussed from three perspectives.

Perspective of error-driven learning

First, in terms of psychological mechanisms, the combination of spoken language and graphic codes contributes to standardization. Spoken language is the cheap and transient signal, and when people combine spoken language with graphic codes, they can verbally correct misunderstandings about the meaning of graphic codes, which is related to error-driven learning (Li et al., 2021). We have built an error-driven standardization model. Therefore, we can understand the mechanisms by which spoken language corrects misunderstandings of graphic codes at a psychological level (Fig. 1). As the figure shows, when the ideography appears, each interlocutor will have an initial expectation for it (expectation 1), followed by a first interaction with other interlocutors to exchange meanings (response 1). During the information communication, each interlocutor will receive information from other interlocutors (feedback 1), which may result in support, opposition, or supplement to the interlocutor's existing views and then give rise to prediction error (prediction error 1). When the prediction error happens (i.e., the expected view does not match the received feedback), the interlocutor will receive the error signal, first performing error monitoring to detect the mismatched information, then making targeted

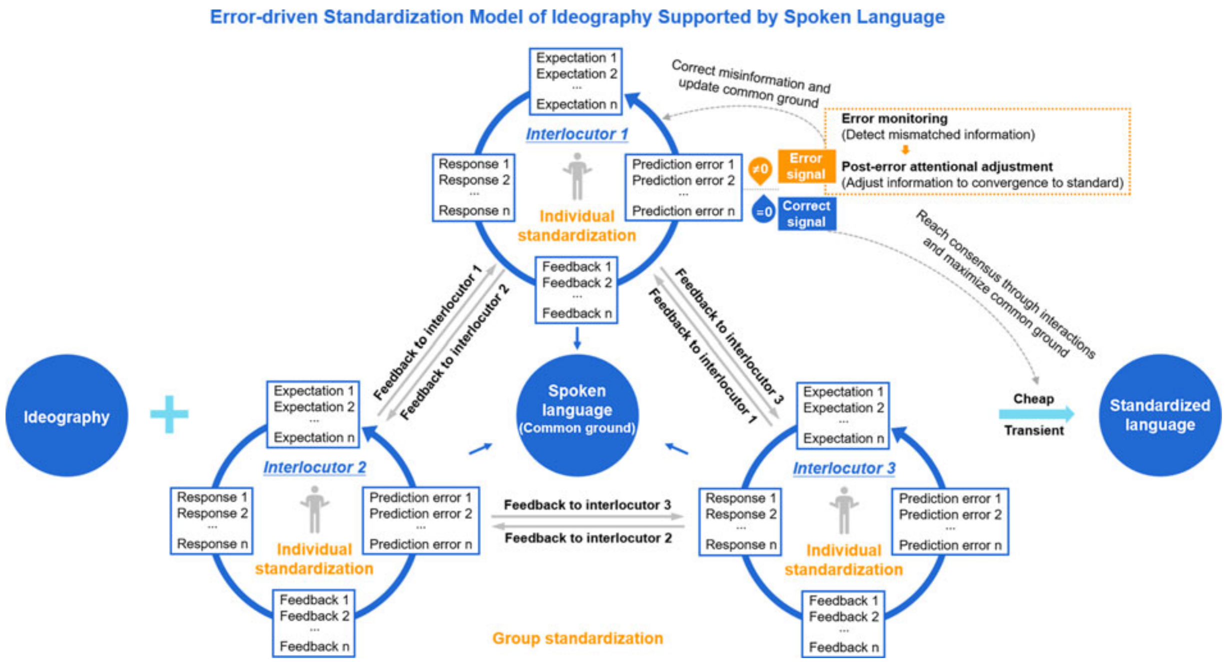

Figure 1 (Zhang et al.). Error-driven standardization model of ideography supported by spoken language.

posterror attentional adjustments to integrate their information (Li, Wang, Li, & Chen, 2022), and eventually becoming identical to the existing standard. The interlocutor generates new expectations (expectation 2) by correcting misinformation and updating common ground. This is followed by a second interaction to process individual standardization. In the interaction, the interlocutors support, debate, or correct each other, cycling through the above processes in order to revise and update the information several times. Until the end of the *n*th interaction, each interlocutor's prediction error would become zero (i.e., the expected viewpoint is the same as the received feedback). It means that the group has reached a consensus through the interaction, forming a uniform standard that can be shared and completing group standardization. In short, interlocutors standardize ideography by engaging in a cheap and transient interaction process of individual standardization combined with group standardization, which is compatible with the natural tendency to combine spoken language with graphic codes.

Perspective of memory

Second, combining spoken language with graphic codes is more beneficial to human memory. For example, both working memory and long-term memory are better with pronunciation added than without pronunciation (Hopkins & Edwards, 1972; MacLeod, Gopie, Hourihan, Neary, & Ozubko, 2010; Tan, Li, & Bai, 2022). Memory for pictures with pronunciation is also more effective than without pronunciation, when content-rich pictures are used as stimuli (Zormpa, Brehm, Hoedemaker, & Meyer, 2019). From the neuroscience perspective, vocabulary memorization with pronunciation also has an activation advantage, compared with vocabulary memorization without pronunciation (Bailey et al., 2021). The learning hypothesis in the target article claims the human mind cannot memorize large numbers of pairings between meanings and visual symbols. However, we argue that humans are capable of remembering large numbers of pairings between meanings and visual symbols; it's just that memory efficiency and performance could be largely improved by combining pronunciation with visual symbols, therefore, humans choose the discarded less-efficient way of memorizing things.

Perspective of psychological evolution

As mentioned above, it can be demonstrated that adding spoken language enables humans to achieve better performance on learning and memory. Human's learning and memory are closely related to long-term psychological evolution, who may have evolved a tendency to learn and remember out of survival interests (Nairne, Thompson, & Pandeirada, 2007). For our ancestors, in order to solve the crisis of survival and repopulation, it was necessary to use a system of information that could be passed on from generation to generation. Thereby, the crisis could be better resolved and corresponding experiences could be preserved for a long time. For survival, our ancestors would have preferred a more efficient way of learning and memory. After some practice, this operation of combining spoken language does not need to consume cognitive resources, and have become an automatic operation after psychological evolution (Hu et al., 2017; Yin, Sui, Chiu, Chen, & Egner, et al., 2019; Zhang, Ding, Li, Zhang, & Chen, 2013). Automatically combining spoken language with graphic codes may be an important reason for the historical absence of ideography.

Chinese characters

To sum up, from a psychological perspective, people will combine spoken language with graphic codes. Although Morin supposes Chinese is a code for spoken language, this opinion is absolute and controversial (Li, 1996; Zhang, 2011; Zhang et al., 2012; Zhu, 1995). If all Chinese characters are the record or code for spoken language, they cannot explain the existence of words that describe the shape of things but have no pronunciation clues in Chinese. For example, the Chinese oracle bones use such forms as , , and  to represent sun, moon, and mountains, respectively. In modern Chinese characters, sun is "日," moon is "月," and mountain is "山." Only 26% of the characters in the earliest oracle bones are associated with sound clues. Therefore, Chinese is a graphic code combined with spoken language, and cannot be interpreted absolutely as a record or encoding of spoken language (Zhang, 2011; Zhang et al., 2012; Zhu, 1995).

Conclusion

In summary, from a psychological perspective, this article suggests that because of the advantages of learning and memory, people will be inclined to choose a graphic code that incorporates spoken language as opposed to one that does not, and this operation of combining may have become automatic over time. People always automatically want to combine spoken language with graphic codes, which may be the reason why ideography has been missing since ancient times, and Chinese is a script that combines spoken language with graphic codes.

Financial support. This work was supported by the National Natural Science Foundation of China (Grant No. 32171040).

Competing interest. None.

References

Bailey, L. M., Bodner, G. E., Matheson, H. E., Stewart, B. M., Roddick, K., O'Neil, K., … Newman, A. J. (2021). Neural correlates of the production effect: An fMRI study. *Brain and Cognition*, *152*, 105757.

Hopkins, R. H., & Edwards, R. E. (1972). Pronunciation effects in recognition memory. *Journal of Verbal Learning and Verbal Behavior*, *11*(4), 534–537.

Hu, X. P., Wang, X. P., Gu, Y., Luo, P., Yin, S. H., Wang, L. J., … Chen, A. T. (2017). Phonological experience modulates voice discrimination: Evidence from functional brain networks analysis. *Brain and Language*, *173*, 67–75.

Li, G. Y. (1996). On the morphological and meaning systems of Chinese characters with morphemes. *Social Science in China*, *3*, 186–193.

Li, Q., Hu, N., Li, Y. L., Long, Q. S., Gu, Y., Tang, Y. C., & Chen, A. T. (2021). Error-induced adaptability: Behavioral and neural dynamics of response-stimulus interval modulations on posterror slowing. *Journal of Experimental Psychology: General*, *150*(5), 851–863.

Li, Q., Wang, J., Li, Z. F., & Chen, A. T. (2022). Decoding the specificity of post-error adjustments using EEG-based multivariate pattern analysis. *Journal of Neuroscience*, *42*(35), 6800–6809.

MacLeod, C. M., Gopie, N., Hourihan, K. L., Neary, K. R., & Ozubko, J. D. (2010). The production effect: Delineation of a phenomenon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(3), 671–685.

Nairne, J. S., Thompson, S. R., & Pandeirada, J. N. S. (2007). Adaptive memory: Survival processing enhances retention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(2), 263–273.

Tan, X. J., Li, X., & Bai, X. J. (2022). The influence of reading mode and learning frequency on production effect of Chinese written word. *Studies of Psychology and Behavior*, *20*(5), 600–606.

Yin, S., Sui, J., Chiu, Y. C., Chen, A., & Egner, T. (2019). Automatic prioritization of self-referential stimuli in working memory. *Psychological Science*, *30*(3), 415–423.

Zhang, L., Ding, C., Li, H., Zhang, Q., & Chen, A. (2013). The influence of attentional control on stimulus processing is category specific in Stroop tasks. *Psychological Research*, *77*(5), 599–610.

Zhang, X. X. (2011). Meaning–spelling theory of the Chinese characters: Insight into the nature of written Chinese from the perspective of cognitive psychology. *Journal of South China Normal University (Social Science Edition)*, *4*, 5–13.

Zhang, X. X., Fang, Z., Du, Y. C., Kong, L. Y., Zhang, Q., & Xing, Q. (2012). The centro-parietal N200: An event-related potential component specific to Chinese visual word recognition. *China Science Bulletin*, *57*(1), 1–16.

Zhu, B. X. (1995). On the original relationship between Chinese characters and the Chinese language. Jianghan Forum, 3, 79–82.

Zormpa, E., Brehm, L. E., Hoedemaker, R. S., & Meyer, A. S. (2019). The production effect and the generation effect improve memory in picture naming. *Memory (Hove, England)*, *27*(3), 340–352.

Author's Response

Puzzling out graphic codes

Olivier Morin[a,b]

[a]Max Planck Institute for Geoanthropology, Minds & Traditions Research Group, Jena, Germany and [b]Institut Jean Nicod, CNRS, ENS, PSL University 29, Paris, France
morin@shh.mpg.de; https://www.shh.mpg.de/94549/themintgroup

doi:10.1017/S0140525X23002418, e260

Abstract

This response takes advantage of the diverse and wide-ranging series of commentaries to clarify some aspects of the target article, and flesh out other aspects. My central point is a plea to take graphic codes seriously as codes, rather than as a kind of visual art or as a byproduct of spoken language; only in this way can the puzzle of ideography be identified and solved. In this perspective, I argue that graphic codes do not derive their expressive power from iconicity alone (unlike visual arts), and I clarify the peculiar relationship that ties writing to spoken language. I then discuss three possible solutions to the puzzle of ideography. I argue that a learning account still cannot explain why ideographies fail to evolve, even if we emancipate the learning account from the version that Liberman put forward; I develop my preferred solution, the "standardization account," and contrast it with a third solution suggested by some commentaries, which says that ideographies do not evolve because they would make communication too costly. I consider, by way of conclusion, the consequences of these views for the future evolution of ideography.

R1. Introduction

One of my goals in presenting the target article to commentators was to show that the study of graphic codes can and should be a field of study in its own right; that the conventions linking inscribed symbols to specific meanings deserve to be studied on their own terms. Disciplinary divides have tended to relegate the study of writing, emblems, pictographs, specialized notations, and so on to two approaches that do not quite give them justice. One approach lumps them together with all possible means of expressions – from visual arts to gesture and language –, to be studied by a general theory of signs. Another approach treats them as borderline cases of linguistic communication – interesting insofar as they can validate theories developed within linguistics, but treated overall as linguistics' poor relation. Solving the puzzle of ideography requires (or so I argued) an approach of graphic codes that takes them and their unique properties seriously. Most of the contributors to this rich and diverse collection of commentaries appear to share this ambition. I am grateful to all of them for showing that the study of graphic codes may interest a broad range of disciplines, from semiotics to anthropology, from generative linguistics to typology, and from archaeology to neuroscience.

This response is organized into six sections. Section R.2 addresses the commentaries that see writing and other graphic codes as continuous with visual arts, comic books, and other means of expression that rely on iconicity; it defends the view that graphic codes do not rely on iconicity alone, and possess special properties by virtue of being codes. Section R.3 clarifies the claim that writing is a specialized notation of spoken language. Section R.4 revisits the "learning account," a solution to the puzzle of ideography based on the view that a generalist ideography would be unlearnable. I address new arguments that the commentaries put forward in favor of the learning account, but also against it. Section R.5 clarifies my preferred solution to the puzzle of ideography, the "standardization account"; it holds that generalist ideographies fail to evolve not because they are unlearnable, but because it is difficult for their users to align on a shared code. Section R.6 addresses a different solution to the puzzle of ideography – a simpler one, which says the reason ideographies do not develop is because graphic symbols are too costly to produce. Section R.7 concludes with an attempt at synthesizing the many commentaries that responded to my speculative views on the future of ideography in the digital world.

R.2. Graphic codes are more than mere images

R.2.1. On distinguishing codes from noncodes

The target article was focused on the evolution of graphic codes, rather than graphic communication in general. The key property that marks out graphic codes from other forms of visual communication is the importance of conventional (or standardized) mappings between symbols and meanings. Writing systems, heraldic emblems, musical notations, and so on are highly codified in this sense. Visual art forms like paintings, graffiti, comic books, and so on, are not. This does not mean visual art does not carry information: As **Sueur & Pelé**'s commentary notices, even simple abstract doodles or abstract paintings produced by apes possess informational content, in the sense that they form shapes that are both visually complex and predictable to a degree. But graphic codes carry information in a way that is quite different, and much more powerful. Graphic codes possess two important properties: They allow their users to compress a great deal of information into a few simple signs (Garrod, Fay, Lee, Oberlander, & MacLeod, 2007; Tamariz & Kirby, 2015; Winters & Morin, 2019), but they require users to learn the code. This capacity to compress information is both a defining property and a key advantage of graphic codes, as **Moldoveanu** notices when he aptly defines graphic codes as a form of "source coding." Being compressed, codified signals tend to be simpler, cheaper to produce, and less cumbersome to store. In contrast, noncodified representations convey a much smaller quantity of information. They tend to be more complex, thus more expensive to produce; but they can be interpreted immediately without the need to master a code. How do we recognize that a mode of expression relies on codification? Two simple cues are the amount of time or effort needed to learn the code, and the need for translation. Using this yardstick, we can easily see, for instance, that comic books are clearly less codified than natural language, and that Chinese characters are as codified as the vocabulary of a spoken language.

These points did not persuade all commentators. **Zhang, Hu, Li, & Chen (Zhang et al.)** argue that Chinese writing is iconic, not symbolic: The characters' meaning is immediately accessible without knowing the underlying code. **Cohn & Schilperoord** argue that comic books are a visual language on a par with spoken languages. Other commentaries (e.g., **Straffon, Papa, Øhrn, & Bender [Straffon et al.]**; **Wisher & Tylén**) broadly agree with the target article but underscore the importance of visual arts and iconic resources in graphic communication.

R.2.2. On iconicity

Signs are iconic, according to the target article's definition, if their meaning can be transparently derived from their form alone, without being acquainted with a specific conventional code. Following a broad scholarly consensus, I claimed that writing, like language, is not iconic in this way. Not all commentators agree. **Zhang et al.** (and perhaps **Yan & Kliegl**, more cautiously) argue that Chinese writing is iconic to a certain extent, whereas **Wisher & Tylén** cast doubt on the general idea that language is an arbitrary code. Zhang et al. argue that most Chinese characters directly describe the shape of things, citing characters such as 日 (rì, sun), 月 (yuè, moon), and 山 (shān, mountain). Yan & Kliegl echo this view in a more watered-down version, arguing that "… some [Chinese] characters are…largely recognizable even by untrained eyes (e.g., 山 for mountain and 田 for farmland)." Is this true?

The studies that investigated the iconicity of Chinese characters (Koriat & Levy, 1979; Luk & Bialystok, 2005; Xiao & Treiman, 2012) generally use a two-alternative forced choice (2AFC) paradigm. Participants are shown a character (e.g., 山) and asked to choose between two associated meanings, the correct one (mountain) and an incorrect distractor (e.g., lake). A 2AFC experiment is considered conclusive when a statistically significant proportion of guesses reaches above the level of chance (50%). This sets a rather low bar for a symbol to count as iconic. True iconicity would be obtained if participants spontaneously guessed a symbol's accurate meaning on seeing the symbol, without being primed with two meanings that include the right one. (This happens less than 2% of the time in Luk & Bialystok, 2005.) Even using the 2AFC task, for cherry-picked sets of iconic characters, performances are mediocre. In the most recent study on the topic (Xiao & Treiman, 2012), the characters cited by the commentaries, 山, 田, 月, and 日 are recognized at rates of 65, 85, 55, and 35%, respectively. And are such characters the majority? No, as **Yan & Kliegl** recognize. Out of the 213 (highly common) characters studied by Xiao and Treiman (2012), only 15 are guessed above chance in the 2AFC.

The example above underlines the perils of relying on 2AFC tasks to show that symbols or wordforms are iconic as opposed to arbitrary. Such studies show that symbols provide some information about their referents by iconicity alone, and they usually show that this amount is small. They do not demonstrate that the symbols derive most of their informative power from iconicity alone. More generally, studies showing nonrandom associations between (some) wordforms and their meaning are intriguing (Dingemanse, Blasi, Lupyan, Christiansen, & Monaghan, 2015; Monaghan, Shillcock, Christiansen, & Kirby, 2014); but they do not claim to challenge the view that language mostly rests on conventional mappings between symbols and referents (Lewis, 1969; Skyrms, 2010). They actually show the opposite: For instance, Monaghan et al. (2014) estimate at 0.02% the share of variance in one-syllable English wordform properties that can be predicted by iconicity. Such differences in degree are massive enough to justify drawing a sharp boundary between spoken or signed languages, on the one hand, and on the other, means of expression that rely overwhelmingly on iconicity, such as visual arts.

R.2.3. Comics are not a visual language

In this spirit, I presented two simple arguments against the view that comic books rely on a visual language (Cohn, 2013). First, when exported to a different country, comic book drawings require no translation, whereas their written dialogues do. Second, the conventions of a particular genre (manga, for instance) can be assimilated in a few hours, compared to the years required to attain fluency in spoken language. More fine-grained arguments could be given. **Wisher & Tylén**, for instance, note that "the comics principle … has not evolved into conventionalized ideographic codes," in particular when it comes to encoding anaphoric relations. In their commentary, **Cohn & Schilperoord** address none of these objections.

Most of the disagreements between me and **Cohn & Schilperoord** spring from the way we understand conventionality, or standardization (two words that I use interchangeably). The topic of the target article is the evolution of codes, understood as standardized (or conventional) mappings between symbols and their meanings (de Saussure, 2011; Lewis, 1969; Millikan, 1998; Scott-Phillips, 2014). I took the view that the kind of standardization that matters is the standardization of the code that pairs meanings and symbols – in other words, codification. Linguistic codes, I argue, are far more standardized than many visual codes (including comics). Thanks to this, they can carry and store much more information, compared to merely iconic means of expression. Cohn & Schilperoord's reply stems from a different view of conventionality, or standardization. In their definition, contrary to mine, standardization has nothing to do with the way images map to meanings. Thus, purely iconic signs can be highly standardized. They provide a number of examples for iconic signs that are conventional, in the sense that artists in different traditions depict eyes or fists in the same way, quite different from the way that is taught in other traditions.

So far, we simply have different ways of using the same words; but the disagreement becomes substantial when **Cohn & Schilperoord** go on to claim that standardization has nothing to do with the distinction between iconic or noniconic signs. This, in my view, is tantamount to dismissing the contribution that standardized codes make to communication. Indeed, they go on to show that comic books, being standardized, are a full-blown language on a par with spoken languages. The fact that comic books overwhelmingly rely on iconicity is irrelevant in their view, because spoken languages (they claim) also rely on iconicity to some extent. (A point I addressed in the previous section.)

Suppose we agree, and the formal standardization of styles is all that matters for comic book styles to count as full-blown languages. That would compel us to put any form of graphic expression showing variations in style on the same footing as spoken language. This includes cave art (Guthrie, 2006), culturally transmitted pottery decorations (Crema, Kandler, & Shennan, 2016); beyond this, cultural patterns have been found, or claimed, for nest building in some bird species (Madden, 2008), for primate tool making, and so on. All these things may be called "visual languages" – and why not? They are, after all, visual forms of expression, forming cultural patterns. Yet this simplification comes at a cost. It renders us incapable of putting a name on something that makes spoken and signed languages uniquely powerful and informative. That something is a shared code. A shared code allows language users to convey information cheaply and efficiently, saving them the effort of depicting or explaining meanings that are already encoded. Having a shared code requires standardizing mappings between symbols and meanings – mere standardization of forms does not suffice.

R.2.4. Visual arts cannot fulfill all the functions of graphic codes

I welcome **Straffon et al.**'s eloquent plea for taking seriously the variety of visual communication systems, and I fully agree with them that writing is very far from being the only efficient way of transmitting information with images. I am less persuaded by the claim that visual arts alone can enable rich forms of communication comparable to those that graphic codes make possible. I worry that we may be falling into an old trap: The temptation to underestimate the degree to which visual communication is codified. This temptation has been a recurrent problem in the study of American graphic codes – the region on which Straffon et al. focus their commentary. Cuna shamans' pictographs were dismissed as mere drawings before anthropologists like Severi (2019) showed how they worked; the very idea of a writing system native to America was dismissed until quite late (e.g., Gelb, 1963); and it took a long time before the sophisticated encoding system of khipus – not a writing system but a complex graphic–haptic code – was taken seriously (Urton, 2017). Precedents like this make me wary of embracing the highly controversial view that Teotihuacan had no writing (Helmke & Nielsen, 2021, make a good case for a native glottographic writing system at Teotihuacan). Likewise, I doubt the Inka could have held together a vast empire using visual arts alone: They could hardly have done it without their khipus. These minor disagreements aside, I agree on Straffon et al.'s most important point: The evolution of writing was highly contingent and unpredictable, due in part to the availability of graphic codes that could fulfill some of its functions.

To summarize, the puzzle of ideography is a puzzle about the evolution of graphic *codes*. Graphic codes are mostly conventional: They cannot be read or produced fluently by someone who does not know the underlying standard for pairing symbols with meanings. As a corollary, many ideographic codes are not iconic at all (algebraic signs, logical symbols, monetary symbols, etc.). Some are only residually iconic, like musical notations. The relative height of notes on the staff score is iconically linked to their pitch, but this does not get us very far if we want to know a note's exact pitch, because the clef symbol is not encoded iconically, and neither are the alterations, and many aspects of the code have no iconic meaning – for example, the fact that white notes are longer than black ones. This is why codes like this one interest me: Because they can only evolve if their users are taught a series of conventions. Forms of expression that are not strongly codified, like visual arts, were not my primary concern, because they lack the crucial power of graphic codes: The power to compress information, thus allowing us to store and transport it.

R.3. The language–writing nexus

Relatively few commentaries took issue with the specialization hypothesis. This hypothesis claims that graphic codes necessarily

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

encode a small range of meanings, which confines them to restricted topics, such as music, mathematics, or (to use **Sterelny**'s excellent example) chess games. What makes the specialization hypothesis important is its application to writing. Writing, being a graphic code, is a specialized notation, not a general-purpose one. What is it specialized for? I defend the (rather banal) view that writing mostly encodes elements of language – be they phonemes, syllables, or morphemes. This does not mean that I consider writing a record of speech, or a phonography: Linguistic elements like phonemes are abstract, contrastive categories, not sounds. But my view does entail that the basic components of writing systems encode linguistic units and cannot be processed without an understanding of the encoded language.

R.3.1. Why literacy requires linguistic competence

Three commentaries that do not object to the specialization hypothesis in general nonetheless wish to nuance the claim that writing encodes language. Two of them (**Yan & Kliegl**; **Zhang et al.**) are specific to Chinese writing. **Sterelny**'s remarks are much more general. He makes three[1] objections against the view that writing encodes language.

First, holding that view forces me to deny the possibility that someone achieves literacy in a language they cannot speak. This happens (in **Sterelny**'s view) to academics who learnt a language through books alone. This case is very close to situations of literate diglossia (which the target article does mention). Diglossia typically occurs when a language spawns a literate variant that is overwhelmingly used by clerks for literate use. Classical Chinese, literate Arabic, or Renaissance Latin are cases in point. Such literate languages can develop on their own and become quite distinct from their spoken counterpart. The people who can read and write these languages can also read them aloud, and are trained to do so in specific settings (e.g., classical poetry recitations). This, to me, counts as a kind of linguistic competence, albeit one that lacks fluency. Today's classicists are able to speak Latin – their own literate kind of Latin. Sterelny seems to disagree, but holding this stance would force him to say that linguistic competence in Sumerian, ancient Egyptian, Hittite, and so on has entirely disappeared, turning most paleographers into strange impostors.

Sterelny turns to semantics for his last objection. If it were true that writing represents spoken language, he notes, then written sentences would be representations of spoken sentences, not of states of the world. Thus, the written sentence "Berlin is the capital of Germany" would have the corresponding spoken sentence as its truth condition: It would not be a statement of fact. This would make standard semantics inapplicable to written sentences. Sterelny reads a lot into my use of the verb "represent"; I more often wrote that writing systems *encode* spoken languages – probably a more adequate verb. Inscriptions that encode sentences should not be confounded with the truth-bearing proposition that the encoded sentence may express, as Frege (for instance) made clear:

> A sentence which an author writes down is primarily a direction for forming a spoken sentence in a language whose sequences of sounds serve as signs for expressing a sense. So at first there is only a mediated connection set up between written signs and a sense that is expressed. But once this connection is established, we may also regard the written or printed sentence as an immediate expression of a thought, and so as a sentence in the strict sense of the word. (Frege, 1920/1981, p. 260)

Written sentences, in this Fregean view, can be analyzed on two levels. On the first level, the printed inscription "Berlin is the capital of Germany" is an encoding of a possible spoken sentence, that is to say, a set of instructions or a recipe for forming a sentence (a set of possible spoken sentences, to be precise, because accentuation, prosody, etc. are not usually encoded). As an encoding, the inscribed sentence may be more or less accurate (for instance, it may contain typos); but it lacks full-blown truth-conditional meaning. On a second level, the sentence that the inscription encodes may express a proposition with truth conditions, if some additional conditions are fulfilled: The proposition must be expressed with the right assertoric force, and at the right moment (the proposition was neither true nor false in 1960 or 1750, but it is true today). **Sterelny**'s challenge dissolves when we distinguish propositions, which are truth-bearers, from inscriptions, which are not.

R.3.2. Chinese writing

There was on the whole, relatively little push-back against the glottographic view of Chinese writing advocated in the target article. **Zhang et al.** dispute it, but only concerning its earliest manifestation, the writing on Shang oracle bones. Even though the use of phono-semantic compounds seems attested on oracle bones (Boltz, 1993), and even though Old Chinese phonology has been successfully reconstructed, I agree that the writing system was highly ambiguous (Demattè, 2022). In this it is similar to other pristine inventions of writing, where the mapping between syllables and symbols took a long time to become systematic and standardized (see, e.g., Hermalin & Regier, 2019). **Yan & Kliegl** broadly agree that "Chinese encodes a natural spoken language just as alphabetic scripts do," but also point out the rather unique aspects of the relation between Chinese writing and language. Their insightful commentary contains claims that I doubt – how easy it is for Japanese readers to read Chinese, or how iconic some Chinese characters may be. There are also claims that I agree with: It is true that "phonetic radical[s] merely provides an unreliable clue to [their] pronunciation." However the same claim can be made (to varying degrees) about most other writing systems except the most regular ones (e.g., Finnish or Hungarian writing), because many writing systems are highly irregular (including English). Overall, I take Yan & Kliegl's welcome qualifications as matters of nuance.

R.3.3. Strengthening the specialization hypothesis: The centrality of language

Thus, only a minority of commentaries challenge the specialization hypothesis, or the glottographic view of language defended in the target article, and they do not provide strong arguments to reject it. Other commentaries seem generally to endorse it, and three commentaries go much further than the target article. **Overmann** sees a more watertight boundary between writing and other graphic codes than I do; **Harbour** argues that graphic codes rely on the language faculty even when they do not encode the language that their users speak; **Winters** argues for a strong role of language not just in the emergence of writing, but in the evolution of all graphic codes.

Overmann's commentary emphasizes the uniqueness of writing (which she calls "visible language") vis-à-vis other graphic codes. She argues that writing is entirely distinct from other graphic codes (e.g., musical or numerical notations), functionally,

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

psychologically, and historically. I agree with her that writing was a rather singular invention, which needs to be distinguished starkly from other graphic codes. That being said, I am not convinced that writing has "no direct material precursors" (because all pristine inventions of writing had clear precursors, I assume the truth of Overmann's claim hinges on what we mean by "direct"). Nor do I think that the use of symbol recombination is unique to writing (see sect. R.7.3). The specialization hypothesis does not need to overplay the uniqueness of writing.

Harbour's commentary highlights the many surprising ways in which graphic codes, including writing, may rely on the language faculty, or at least on cognitive mechanisms that are shared with the language faculty broadly construed (such as recursion). Harbour deftly shows that some of the linguistic structure in writing systems is not actually derived from the language that they encode. Writing systems have a degree of autonomy from their target language, and may form structures that have no counterpart in it, like the determinatives found in Egyptian hieroglyphs. These structures can be analyzed using the tools of linguistics, because they behave according to the same rules as similar structures found in other languages – but not in the encoded language. Thus, Egyptian hieroglyphs reinvented determinatives, even though the language they encode lacks them. With examples like this one, Harbour makes a convincing case for the research program that studies writing systems as linguistic objects without reducing them to mere reflections of the structure of their target language.

Harbour goes on to suggest that, because linguistic structure pervades graphic codes, ideography was always dead on arrival (so to speak). If ideography is defined as a language-independent graphic code, and if language-like structures always find ways to sneak into graphic codes, even the codes not made to encode them, how could a language-free ideography ever evolve? This argument rests on an ambiguity in how we define "language." The way the target article defines them, ideographies are graphic codes that do not encode elements of a specific spoken (or signed) language. This leaves open the possibility that ideographies may themselves possess language-like features such as recursion, duality of patterning, or compositionality (see sect. R.7.3). Put differently, what differentiates a writing system from an ideography is the fact that writing systems encode a specific natural language that preexists in them – not the fact that it possesses general language-like features like recursion, syntax, and so on. Taking Harbour's argument to its extreme limit would lead us to a familiar position in the debates over ideography: Ideography cannot logically exist, because it is conceptually impossible to assign meanings to symbols without in some way using language (Boltz, 1993; du Ponceau 1838). If this position were true, there would be no such thing as an ideograph, and we would be completely unable to describe the difference between symbols like 2, 3, ♡, and the written words "one," "two," "love" (Edgerton, 1941[2]). Ideography is simply the fact that some symbols encode ideas directly, bypassing words. These symbols may still possess rich language-like properties.

Winters's commentary further stresses the centrality of language to all graphic codes, not simply writing. He makes many important points, all of which I endorse. Most important perhaps is his remark that self-sufficient specialized graphic codes, like musical or mathematical notations, seem to evolve much more readily in literate societies. This does not mean societies without writing do not produce rich and sophisticated graphic codes – they do – but it does mean these codes are unlikely to be self-sufficient (i.e., they should rely on oral glosses). I agree with Winters that this is a natural consequence of the standardization account, because writing, itself a product of successful coordination, in turn becomes a powerful coordination device.

Winters raises a question also broached by other commentaries (**Adiego & Valério**; **Gainotti**): To what extent can a graphic code ever be independent of spoken language? **Adiego & Valério** argue that the creation and first acquisition of codes (if not their use) must involve linguistic communication. **Gainotti** shows, on the basis of clinical evidence with aphasic patients, that learning to communicate ideographically is not a domain-general process, but rather relies on language-specific cognition. The difficulties that aphasics experience in learning Bliss or other pictorial codes do suggest this, because their language faculty is damaged but their capacities for perception or memorization are preserved. For **Winters**, creating and learning a graphic code from scratch without using language is possible in theory, but unlikely in practice. One intriguing argument that he gives is the apparent lack of codified permanent visual marks in nonhuman animals, in contrast to the importance of acoustic codes in several species of birds and primates. Here again, I am tempted to agree; bower-bird nest decorations are the closest thing to a counterexample that comes to my mind, but they do not have codified meanings in the way that, for instance, vervet monkey calls do.

R.4. For and against the learning account

The learning account is a way to solve the puzzle of ideography by showing that acquiring a generalist ideography raises serious cognitive difficulties, contrary to spoken language. Alvin Liberman's theory of writing is a good example of a learning account. Besides, the puzzle of ideography is a relatively neglected problem, and Liberman is one of the few scholars who explicitly articulated it and proposed a cogent solution. Being keen not to attack straw men, I concentrated my critique on his specific argument. No commentator explicitly attempted to defend Liberman's own views, but several commentaries proposed alternative ways of showing that generalist ideographies raise a learning problem (**Arsiwalla**; **Harris, Perfetti, & Hirshorn** [**Harris et al.**]). Other commentaries sought, on the contrary, to strengthen the case against the learning account (**Howard**; **Nephew, Polcari, & Korkin** [**Nephew et al.**]).

R.4.1. Alternatives to Liberman's learning account

Harris et al. object to my critique of the learning account, because it focuses on Liberman and his motor theory (which they quickly dismiss). In their view, I should have criticized the learning account in a much more general way, going beyond Liberman's specific theory to include other possible versions of the learning account. This discussion gives me an occasion to do exactly that.

Harris et al. present an intriguing argument to defend the general idea of a learning account. Dyslexic as well as congenitally deaf individuals cannot easily link letters to sounds, but this would not be an obstacle to literacy if they simply could treat writing like an ideography. In other words, if humans had no problem learning arbitrary pairings between signs and meanings, dyslexics and deaf people could simply treat written words as ideographs – that is, take them as pointing directly to ideas, bypassing language. This, in their view, does not happen. Only a tiny minority

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

of profoundly deaf people learn to read in this way (and then again, not very successfully). As for dyslexia, attempts to cure it by training learners to process words as ideographic symbols aggravate the condition instead of improving it. Harris et al. take this as a strong indication that ideography raises an insurmountable learning problem.

The learning difficulties that **Harris et al.** highlight are real and significant. The fact that they are also found for the Chinese script is an important reason to resist nonglottographic accounts of Chinese characters. But do they show that ideographies in general are unlearnable? No. The fact that deaf people or dyslexics usually fail to recognize written words by their visual shape alone does not make this point. Writing it is not an ideography, so attempts to learn writing as if it were ideographic are bound to fail. The people Harris et al. refer to all have some mastery of the spoken language that writing encodes; they cannot simply ignore their linguistic knowledge – as Harris et al. acknowledge. Thus, I fail to see how Harris et al.'s argument could make a case against the learnability of ideographies. Writing is the very opposite of an ideography; dyslexics and deaf people struggle with writing precisely because writing encodes language.

A counterpoint to **Harris et al.**'s position is found in **Sterelny**'s commentary. Sterelny considers that a glottographic view of writing cannot account for the success of profoundly deaf people in learning to read. Harris et al. blame the target article for taking deaf literacy too seriously; Sterelny criticizes it for failing to take deaf literacy seriously enough. I attempted to explain literacy in profoundly deaf people by noting that phonemes, syllables, or morphemes are not sounds but contrastive categories, making it theoretically possible to become literate in a language when one's main contact with that language is visual, not aural. Sterelny says this account is not "natural," suggesting that he subscribes to the view that profoundly deaf people may learn a spoken language from print only, by mapping printed words directly onto sign language words (Hoffmeister & Caldwell-Harris, 2014). This view can be disputed, however, because most deaf people who learn to read know the encoded spoken language from partial hearing, lip-reading, or signed versions of the spoken language. Whether this phonemical awareness acts as a help, a hindrance, or a mere byproduct, is not entirely clear, a point that my article should have paid more attention to. In any case, Harris et al.'s commentary makes evident the drastic limitations that profoundly deaf people face in learning to read (Hirshorn & Harris, 2022).

Arsiwalla's thoughtful version of the learning account puts forward two mechanisms that enhance the learnability of spoken language as opposed to visual ones. Arsiwalla's first argument is that ideographic messages are difficult to decompose into discrete chunks, making ideographic codes harder to memorize. To make this point, Arsiwalla relies on the assumption that ideographic codes lack compositionality – an assumption that I refute in section R.7.3. His second argument is based on the view that multimodal learning, combining visual and aural information, is more efficient compared to learning through one modality alone. It rests on a literature showing, in classroom contexts, that students are better at retaining things taught using both visual material and an oral gloss, compared to things taught with exclusively visual material (usually a diagram and a written gloss). I do not quite understand how these intriguing studies would make a case against the learnability of ideography. To make such a case, Arsiwalla would have needed to show that spoken and signed languages are learnt multimodally, and that domain-specific ideographies are not. The acquisition of codes generally relies on several modalities, and this includes ideographic codes like musical notations. Arsiwalla does not explain what would make speech acquisition special in this respect. More importantly, he does not mention sign languages, which are perfectly learnable for deaf people who have at best limited access to the aural modality.

Overall, the two commentaries written in defense of the learning account do not succeed in producing a cognitive mechanism that makes it difficult to learn ideographic codes, without raising the same problem for spoken languages. If one wants to claim that a learnability problem is what prevents us from acquiring a generalist ideography, one has to explain where exactly the problem resides. And the answer cannot be modality. The reason generalist ideographies do not take off is not because they are visual – sign language linguistics has taught us that. One could perhaps imagine other versions of the learning account that do not face this problem, and that would be an exciting research program; but the commentators have not done this yet.

R.4.2. Strengthening the case against learning account

Nephew et al. provide an intriguing argument against some versions of the learning account by highlighting human proficiency with face recognition. The fact that we routinely process and correctly recognize thousands of faces shows how far human visual memory can go. Although this argument would not affect Liberman's version of the learning account (because in Liberman's view we learn codes through a motor route not a visual one), it would clearly be a problem for anyone who claims that ideographies are unlearnable because of visual memory constraints (as **Harris et al.** seem to do). I take Nephew et al.'s point, with a handful of qualifications. The typical inventory of faces a person can identify ranges in the thousands – 5,000 on average and up to 10,000 for "super-recognizers" (Jenkins, Dowsett, & Burton, 2018). This is as high as the size of the most complex-documented graphic codes.[3] It is not, however, extremely high if compared to the vocabulary size of English (to take a well-studied example), which was estimated at 42,000 lemmas on average, with other estimates ranging from 10,000 to hundreds of thousands (Brysbaert, Stevens, Mandera, & Keuleers, 2016). Another important caveat, pointed out by Nephew et al. themselves, is the domain-specific nature of face recognition, which is cognitively distinct from neighboring skills such as voice recognition (Young, Frühholz, & Schweinberger, 2020). With these limitations in mind, I agree with Nephew et al.: Our face recognition abilities suggest that visual memory per se does not stand in the way of ideography.

Another commentary that redoubles my critique of the learning account is **Howard**'s. He concentrates his commentary on Bliss symbols that, he claims, are easily taught and learnt. He blames me for not citing the abundant literature that (he thinks) makes this point. I disagree. The authors of these studies did not actually try to teach Bliss symbolics to their participants. Their goal was different: To compare Bliss with other ideographic systems like Carrier symbols, or pictographic systems like Rebus, Picsyms, or the Picture Communication System, based on transparent pictures. To perform these comparisons, they familiarized their participants with a small number of symbols, in the short term.[4] Only one of the studies cited comes close to a genuine attempt at teaching Bliss: Funnell and Allport (1989). Finding the results disappointing, they gave up on using Bliss as an

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

alternative communication tool with their two aphasic patients. **Gainotti** points to several other studies that reached the same conclusion. There is, thus, no clear-cut evidence that Bliss symbolics is readily learnable. By itself, this lack of evidence proves little – only that seriously trying to teach an ideographic language to normal adults or children (let alone patients) would be a costly and complex enterprise that researchers are yet to tackle.

R.5. Debating and clarifying the standardization account

Several commentaries contained insightful arguments against my preferred solution to the puzzle of ideography – the standardization account. The standardization account claims that graphic symbols are difficult to standardize because they lack the properties that allow users of spoken or gestured signs to align on shared meanings and repair misunderstandings in the course of conversation. If used synchronously, graphic symbols do not facilitate repair, alignment, or rephrasings, because they are too effortful to produce and too cumbersome (they do not fade rapidly), compared to words or gestures. If used asynchronously, graphic symbols cannot be interpreted in the light of a rich common ground, and messages cannot be repaired easily. Because of this, getting users of a graphic code to align on the exact same code is a challenge that can only be overcome for small numbers of symbol–meaning pairings, characteristic of specialized codes.

The general view that face-to-face communication offers unique opportunities for repair and alignment (Clark, 1996; Enfield, 2017) was not challenged by the commentaries. **Moldoveanu** found an excellent way to phrase this point when he wrote that spoken or signed language rely much more on "channel encoding," which consists in adding enough redundancy to a signal to offset noise. Repetition, repair, but also pointing and other means of obtaining alignment between interlocutors are all forms of channel coding in his view. Graphic codes, on the contrary, cannot count so much on channel coding so they must rely on precise source coding, that is to say, unambiguous mappings between symbols and their meanings.

The kind of standardization that spoken or signed conversations allow is a fine-grained and decentralized alignment, the outcome of a process that many users of the code contribute to. Getting a committee to agree on a published dictionary does not achieve standardization in this sense. Standardization, in other words, is a social fact: As **Riggsby** aptly notes, this means standardization is both a matter of degree and a matter of scale. A code may be highly standardized for some people, not for others; may be loose overall but tight in some places; different standards often compete. A code is standardized for a community of users, at a given point in time. This point is the source of a misunderstanding about Bliss symbols, between myself and **Howard**. There is, Howard remarks, an official standard for Bliss, accompanied by textbooks, dictionaries, and so on. The problem is that this standard exists on paper; few people master it well enough to communicate fluently with others. The fact that the use of Bliss is restricted to a clinical context, to help people with severe language impairments makes it difficult to give a fair estimate of the system's potential for future success: As we saw, the literature cited by Howard does not prove much one way or the other.

Berio, Can, Helming, Palazzolo, & Moore (Berio et al.) object to the standardization account, arguing that graphic codes could simply be learned and used in face-to-face interactions, where repair, alignment, and common ground are readily available. But a graphic code that would be used exclusively in face-to-face settings would lose what makes it useful compared to other codes: Its capacity to support asynchronous communication. A graphic code that is only ever used face-to-face would have no advantage compared to spoken or signed language, with the added costs of consisting of cumbersome signs, effortful to produce. The only feature that gives graphic signals a clear comparative advantage is their permanence; that feature is lost when graphic symbols are only used for face-to-face communication.

Riggsby provides an intriguing rejoinder to the standardization account when he notes how poorly standardized writing systems could be, in their early stages. In his view, this need not be an obstacle to using a code, as long as the code is tight enough for a particular group of users, in the same way that dialect continua enable communication inside overlapping pockets of linguistic unity, even though linguistic standardization is fairly low on the whole. Riggsby's argument is persuasive and backed by compelling examples. It points out a lacuna in my target article: Standardization is described as a population-level challenge, yet the mechanisms that produce it occur at the level of the dyad (alignment, repair, etc.). More work is clearly needed to bridge these two scales.

R.6. Can we solve the puzzle of ideography with production costs alone?

Three commentaries (**Adiego & Valério**; **Berio et al.**; **Tylén & Wisher**) suggest a solution to the puzzle of ideography that seems much simpler than the one I advocate. All commentaries start from the fact that graphic messages are costly to produce, compared to speech and gesture. The costs of producing graphic messages, they argue, are what really stand in the way of a generalist ideography. If true, this cost-based account could replace my standardization account. The standardization account acknowledges the production costs of graphic message, but does not see these costs as sufficient, in themselves, to explain why generalist ideographies did not arise, while specialist ideographies did.

Berio et al. offer the most challenging version of this argument. They argue that the format of graphic messages requires them to match the complexity of their content. When that content reaches a certain level of complexity, the message's inscribers must rely on ever more intricate drawing skills, until the message simply becomes too unwieldy to produce or process. The only way to escape this trade-off is for messages to be ambiguous, which means that they will require a verbal gloss for their meaning to be fully communicated. Thus, the inherent complexity and cost of graphic messages suffice to explain the puzzle of ideography. This argument can be decomposed into three premises and a conclusion; I agree with the premises but reject the conclusion.

Premise 1: Graphic messages are costly to produce (compared to spoken ones). Other commentaries insist on this point too. **Adiego & Valério**'s commentary underlines the importance of production costs in blocking the evolution of ideography. They highlight, in particular, the limitation raised by the fact that most graphic symbols require some external support to inscribe them on,[5] and a tool to inscribe them with. **Wisher & Tylén** agree. I agree too: Graphic messages are hard to produce. In the target article, cheap production is one important reason why spoken or signed languages are easy to standardize, while graphic codes are not. That, however, is only one reason among several others; it does not, by itself, solve the puzzle, as I'll explain below.

Premise 2: Informative graphic messages tend to be complex, and thus costly. The expressive power of graphic codes is limited,

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

Berio et al. argue, by the degree of complexity that graphic symbols can achieve. **Wisher & Tylén** make a similar point: Efficient graphic communication, in their view, requires material resources that, prior to the industrial revolution, were either highly expensive (like tapestries), or just nonexistent (like cartoon motion pictures). I agree that there is a correlation between the complexity (and consequent cost) of graphic messages, and the amount of information they can encode. For instance, in writing systems, characters encoding high-level linguistic units (think Chinese characters as opposed to Latin alphabet letters) are more graphically complex (Chang, Chen, & Perfetti, 2018; Miton & Morin, 2021). On a more fine-grained level, frequent letters carry less information than rare ones (by definition, because a rare letter is more unexpected than a frequent one). Our work (Koshevoy, Miton, & Morin, 2023) shows that frequent letters are graphically simpler than infrequent ones; considering a diverse sample of 27 scripts, we found that this relation obtained inside each one of them. In short, I emphatically agree with premise 2, but I would add two caveats. First, the correlation between informativeness and cost/complexity is real and robust, but not necessarily strong. Frequent letters can be complex; and it is entirely possible to encode a great deal of information using just a few symbols (consider "$E = MC^2$"). Second, and most importantly, the mechanism that Berio et al. highlight also applies to spoken language. There is evidence that the complexity of utterances is related to the amount of information they contain. At the level of words, the length of words reflects their conceptual complexity (Lewis & Frank, 2016). Words that are long or phonotactically complex are infrequent and less likely to be ambiguous (Piantadosi, Tily, & Gibson, 2012). At the level of utterances, it seems fairly straightforward that, *ceteris paribus*, encoding a lot of information is easier to do with a ten-words English sentence compared to a three-words one. Thus, the constraint that Berio et al. highlight is very general; it applies far beyond graphic codes; and it need not be very strong.

Premise 3: There is a trade-off between code and context. Graphic codes allow us to encode some information graphically but, as with any other mode of communication, not everything can be encoded (Winters, Kirby, & Smith, 2018; Winters & Morin, 2019). What the code leaves out must be inferred pragmatically, or supplied using another code – in the case of graphic messages, an oral gloss. Whenever we use a code, we face a trade-off between making our messages too ambiguous – running the risk of being misunderstood – or overly explicit – increasing complexity, burdening ourselves and our audience with unwieldy messages. On this point **Berio et al.** fully agree with the target article.

Conclusion: Self-sufficient ideographies do not evolve because informative graphic messages are too costly to produce. This is where me and **Berio et al.** disagree. Production costs play a role in my account of the puzzle of ideography – but they do not explain everything. Costs play a role in the nonevolution of generalist ideographies, because they stand in the way of the conversational interactions that allow communication to self-standardize in spoken or signed languages. But costs do not explain the puzzle of ideography on their own. Costs, after all, can be paid. They had to be, or visual arts would not exist, and neither would specialized graphic codes. The chief benefit of using graphic signals, in spite of their cost, is durability; and durability is worth paying for. Many societies incurred huge costs in maintaining the skills and materials required to produce enduring messages. Writing is indeed cumbersome, costly, intricate – but it is worth the trouble. The reason why elaborate visual arts or specialized notations evolve, but generalist graphic codes do not, is not costs alone. It is because communication with visual art or specialized codes does not require much standardization. Production costs matter, but they matter *only* insofar as they prevent the kind of quick and easy exchanges required for alignment and repair, two key ingredients of standardization. (This point is well captured by **Sterelny**'s commentary.)

Thus, I share many assumptions with **Berio et al.** and **Adiego & Valério**'s commentaries, but I do not think that the costs of producing graphic symbols solve the puzzle of ideography on their own. **Wisher & Tylén** make a slightly different point. Their commentary only considers graphic messages that are fully iconic: Pictures that resemble their referents and need no graphic code to convey their content. They argue that the technical means needed for telling complex messages in an exclusively iconic fashion only became available quite recently – and I agree. This, however, does not solve the puzzle of ideography, which is a puzzle about the origins of graphic codes. As I argued above (sect. R.2), graphic codes are usually not highly iconic, and thanks to this, they can be much simpler than their iconic equivalents – and thus, less cumbersome and costly.

R.7. The future of ideography

The one section in the target article that elicited the most commentaries is the one that speculates about the possible evolution of a generalist ideography, made possible by the resources of the digital age (sect. 6.4). A consensus emerges that graphic symbols are indeed becoming more informative and better standardized online. Critical commentaries say I underestimate this trend, or predict it for the wrong reason; but no one appears to disagree on its possibility.

R.7.1. Standardization of communication in the digital world

The target article argued that codes based on cheap, fast, and transient signals are easier to standardize, because of two independent mechanisms. The first is cheap and fast production leading to repair and alignment: Cheap and fast signals can be modified or repaired multiple times, allowing interlocutors to converge on shared meanings. The second is face-to-face communication with transient signals. Transient signals constrain interlocutors to face-to-face interactions, where the advantages of common ground are maximized. The target article argued that digital communication shared some (not all) of the characteristics of the cheap, fast, and transient signals of spoken or signed languages. This should ease the standardization problem somewhat. Most of the commentaries that broach this issue agree with the target article on these points: Digital communication is closer to spoken or signed communication than to legacy graphic communication, and this should ease the standardization problem to a degree (**Clark**; **Gandolfi & Pickering**; **Veit & Browning**).

The target article, however, was unclear on the exact nature of the mechanism that could further standardization in digital communication. I mentioned two mechanisms: Cheap and fast repair and alignment and face-to-face communication with transient signals. The two are independent; they can be dissociated, and in the case of digital communication they clearly are. The commentaries by **Feldman** and **Gandolfi & Pickering** allow me to clarify my views on this. What, in my view, makes digital communication different is cheap and fast production leading to repair

https://doi.org/10.1017/S0140525X22002801 Published online by Cambridge University Press

and alignment. It is not face-to-face communication with transient signals. Gandolfi & Pickering's commentary make this point much better than I did. Digital communication can be synchronic, and symbols can be exchanged at a fast rate, close to the pace of spoken conversation, because they require little effort to produce. Thanks to this, ideographic symbols can play the same role as the signals used for repair and feedback in face-to-face conversation.

R.7.2. How standardized are the meanings of emojis?

Commentators (**Clark**; **Feldman**; **Gandolfi & Pickering**; **Veit & Browning**) all appear to agree that digital communication is a favorable environment for solving the standardization problem. The question then becomes: Has this standardization happened yet, and what kind of symbols would it apply to? The target article took a cautious stance on the matter – too cautious for several commentators. I argued that emojis, gifs, and other digital pictographs may acquire increasingly precise and standardized meanings; but this stage may not have been reached yet, possibly because digital communication is not yet fast enough compared to the pace of speech, or because we are only seeing the beginning of its evolution. Applying this view to emojis, I made two points. First, emojis are not yet standardized enough to function as a self-sufficient ideography; second, some emojis may reach a sufficient level of standardization in the future, starting with emojis that encode paraverbal cues.

Regarding the first point, two commentaries (**Feldman**; **Veit & Browning**) argue that I underestimate the level of standardization that emojis have already achieved (a point also suggested by **Gandolfi & Pickering**'s commentary). Veit & Browning cite the successful standardization of available emojis across platforms in support of their view; I would reply that the standardization of emoji keyboards should not be confounded with the degree to which their meanings are standardized between users. Feldman criticizes my focus on face-expression emojis, which she sees as much less standardized than other emojis. I chose to focus on face-expression emojis because, as Feldman acknowledges, they are by far the most commonly used emojis (see, e.g., Daniel, 2021). Their poor standardization tells us something important about emoji users' capacities to converge on a shared meaning when that meaning is not immediately given through iconicity. In the studies cited by Feldman (Barach, Feldman, & Sheridan, 2021; Częstochowska et al., 2022), the nonface emojis that are highly standardized stand for objects, activities, food, in a straightforwardly iconic fashion, so that associating 🍺 with "beer" or 🔑 with "key" requires little in the way of a shared code. There are, of course, nonface emojis that have acquired a standardized meaning, often quite remote from their figurative referent: Feldman mentions the 🦋 or 🍆 emojis. But how standardized are these figurative meanings? Quantitative data would be needed to estimate this – an intriguing topic for future studies.

Gandolfi & Pickering broadly agree that emojis are poorly standardized, but they make an exception for the ones used in backchannel interactions such as repair or other kinds of feedback (e.g., 👍, 👌), especially in SMS interactions, which allow for quick exchanges. The point is well taken. I see their commentary as a more specific and more accurate version of the point I attempted to make in section 6.4 in the target article – namely, that digital technologies should foster the standardization of emojis if (1) they are used as paraverbal cues and (2) the interactions are rapid and synchronous, leaving a lot of room for repair and alignment. **Gandolfi & Pickering**'s views also provide an interesting contrast with **Feldman**, who explicitly ignores single emojis that stand on their own, and focuses her commentary on emojis as embedded in sentences or messages. **Feldman** makes a point that the target article acknowledges, and which does not contradict my theory: Emojis do not require much standardization as long as they are surrounded by linguistic information that helps readers narrow down an emoji's meaning. **Gandolfi & Pickering** make the same point: Emojis are often used as complements to writing rather than replacements for it, playing the same part that paraverbal cues (cospeech gestures, nods, and facial expressions) play for speech; this limits their potential to evolve into a complete ideography.

In summary, there is no unanimity on what level of standardization emojis may have reached, but whatever we take it to be, it is not yet sufficient to allow them to function as a self-sufficient ideography. Still, the evolution of emojis highlights the new opportunities for standardization that digital communication opens up.

R.7.3. What would a generalist ideography look like?

If the target article is right and ideographies are both learnable, and capable of becoming standardized in the future, thanks to digital communication, what would such a future ideography look like? The commentaries lay down prerequisites and highlight important challenges that this ideography would need to face. Historically, inventors of ideography (like inventors of universal languages in general) tend to fall into one of two camps. There are those who think their system should improve upon natural language, by being closer to the true structure of concepts. Frege's ideography is a successful example of ideography in this sense (even though it is highly specialized) (Frege, 1883). Then there are those who think they would be lucky enough if they could produce a tool for communication that simply works. **Cheng** and **Moldoveanu** keep the first tradition alive: A good ideography should be a language of ideas that carves concepts at their joints; it should avoid ambiguity; it should express only intrinsic properties of things, not contingent ones. Most other commentaries have more modest ambitions: When they consider the prospects for a functioning graphic code, they simply ask whether it will be good enough for communication. **Mazur & Plontke** go further, denouncing the quest for a language of ideas devoid of ambiguity as a positivistic myth. **Chrisomalis** agrees: It is doubtful whether ideography in that sense was ever possible or even needed. These points are well taken.[6] Another thing the target article did not consider was the use of ideography as a universal language.[7] As **Adiego & Valério** rightly note, we cannot expect a generalist ideography to escape the laws of language evolution: Languages change, diverge, fragment, to form mutually incomprehensible dialects. I wrote that a generalist ideography would cut across language barriers, allowing its users to communicate even when they have no spoken language in common; but I never said that it could stand as a universal language or, as Charles Bliss put it, "overcome Babel."

Even with these relatively modest ambitions, a generalist ideography would still have many challenges to overcome. Compositionality and the capacity to express abstract ideas are two things that natural languages excel in, but graphic codes may struggle with – according to the commentaries. **Cheng** sees abstract and intangible concepts as the greatest hurdle, whereas

Chrisomalis points at the difficulty of using graphic codes componentially – fusing several symbols to communicate one idea. In line with this, **Overmann** sees the capacity to combine and recombine elements (be they phonetic units or morphemes) as something that spoken or written langue are uniquely good at – a point echoed by **Cohn & Schilperoord** and **Arsiwalla**. The specialization hypothesis implies that ideographies can and do express abstract concepts as well as spoken languages do, but only for a restricted range of meanings. It makes the same prediction about combinatoriality: Ideographies are capable of it, as long as the number of basic symbols to be recombined is not too high and their meanings are narrow enough. Both predictions, I think, are confirmed by existing ideographies. Abstraction is no issue for formal logic or mathematical notations, whereas combinatoriality at multiple levels is an obvious feature of many graphic codes, from musical notations to cattle branding (Youngblood, Miton, & Morin, 2023). There are rules for the combination of meaningless elements (duality of patterning) in heraldry (Morin & Miton, 2018), and rules to combine meaningful elements into meaningful compounds in musical or mathematical notations, road signs, and so on. The degree of compositionality at play in many graphic codes is arguably equal or superior to that of natural language: The meaning of compound expressions depends unambiguously on the meaning of the parts and can be derived from systematic rules. This is always true, by construction, for logical notations. But we may well ask, following Chrisomalis, whether this property can scale up for a generalist ideography. If the standardization hypothesis is on the right track, there should be no special difficulty in designing a generalist graphic code capable of high levels of abstraction and obeying compositional combination rules. The problem would lie in getting most users to align on a shared standard, one that cannot be artificially decreed but has to emerge from the back-and-forth of communicative interactions.

Notes

1. Here I overlook **Sterelny**'s second objection, concerning deaf people learning a language through print – it is tackled in section R.4.1.
2. I thank **Chrisomalis** for pointing me to this paper through his comment.
3. Some dictionaries of Chinese contain tens of thousands of characters, but most of them are archaic or rare forms, and it is highly unlikely that a single person has memorized such a dictionary in its entirety.
4. Bliss symbolics consists of around 100 basic signs, which can be combined to form a number of compound signs which varies between 886 TBC (the Unicode Corporation's conservative estimate) and 2,384 (ISO-IR norm). With one exception, the studies that **Howard** cites only taught between 11 (Poupart, Trudeau, & Sutton, 2013) and 45 (Mizuko, 1987) *compound* symbols, most studies teaching 15 compound symbols (Burroughs, Albritton, Eaton, & Montague, 1990; Clark, 1981; Ecklund & Reichle, 1987; Hurlbut, Iwata, & Green, 1982). The maximum number of symbols that were actually retained does not reach above 20 (in Mizuko, 1987). Three studies found that these iconic symbols were easier to learn compared to Bliss symbols (Ecklund & Reichle, 1987; Mizuko, 1987, both of them working with preschoolers, and Hurlbut et al., 1982, working with severely handicapped adolescents). This pattern is consistent with other studies (e.g., Alant, Life, & Harty, 2005; Kozleski, 1991; Mizuko & Reichle, 1989; Sevcik, Barton-Hulsey, Romski, & Hyatt Fonseca, 2018). Lastly, two studies cited by **Howard** did not find a clear advantage for either type of system (Burroughs et al., 1990; Poupart et al., 2013, both working with preschoolers). Interestingly, Clark (1981), working with preschoolers, found that all the ideographic or pictographic symbols that they use (Bliss, Carrier, and Rebus) were easier to learn compared to written English.
5. I agree, but exceptions should be made for tattoos, scarifications, body paintings, etc.
6. **Chrisomalis** is right to argue that the word "semasiography" (also used by **Overmann**) would have been a more fitting label for what I call "ideography": It would be in line with the literature (e.g., Gelb, 1963), and it would put the focus where it ought to be – on the signs and their meanings, rather than the ideas they designate. My only excuse (if it is one) is the need to catch the eyes of a broad interdisciplinary audience to whom the word "semasiography" means little.
7. **Adiego & Valério** remark that were a generalist, self-sufficient ideography to be found, it would then fulfill all the conditions I set to be called a language. I agree with this, but the point is rather vacuous if, as the target article claims, generalist ideographies are rare or nonexistent. Adiego & Valério also note my tendency to use the phrase "visual language" rather too loosely and inconsistently in some places; they are right about this. I invite readers to mentally correct the target article, replacing "visual language" with "visual code."

References

Alant, E., Life, H., & Harty, M. (2005). Comparison of the learnability and retention between Blissymbols and CyberGlyphs. *International Journal of Language & Communication Disorders*, *40*(2), 151–169. https://doi.org/10.1080/13682820400009980

Barach, E., Feldman, L. B., & Sheridan, H. (2021). Are emojis processed like words?: Eye movements reveal the time course of semantic processing for emojified text. *Psychonomic Bulletin & Review*, *28*(3), 978–991. https://doi.org/10.3758/s13423-020-01864-y

Boltz, W. G. (1993). *The origin and early development of the Chinese writing system*. American Oriental Society.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology*, *7*, 1116. https://doi.org/10.3389/fpsyg.2016.01116

Burroughs, J. A., Albritton, E., Eaton, B., & Montague, J. (1990). A comparative study of language delayed preschool children's ability to recall symbols from two symbol systems. *Augmentative and Alternative Communication*, *6*(3), 202–206. https://doi.org/10.1080/07434619012331275464

Chang, L.-Y., Chen, Y.-C., & Perfetti, C. A. (2018). GraphCom: A multidimensional measure of graphic complexity applied to 131 written languages. *Behavior Research Methods*, *50*(1), 427–449. https://doi.org/10.3758/s13428-017-0881-y

Clark, C. R. (1981). Learning words using traditional orthography and the symbols of Rebus, Bliss, and Carrier. *Journal of Speech & Hearing Disorders*, *46*, 191–196. https://doi.org/10.1044/jshd.4602.191

Clark, H. H. (1996). *Using language*. Cambridge University Press.

Cohn, N. (2013). *The visual language of comics: Introduction to the structure and cognition of sequential images*. A&C Black.

Crema, E. R., Kandler, A., & Shennan, S. (2016). Revealing patterns of cultural transmission from frequency data: Equilibrium and non-equilibrium assumptions. *Scientific Reports*, *6*, 39122. https://doi.org/10.1038/srep39122

Częstochowska, J., Gligorić, K., Peyrard, M., Mentha, Y., Bień, M., Grütter, A., … West, R. (2022). On the context-free ambiguity of emoji. *Proceedings of the International AAAI Conference on Web and Social Media*, 16, 1388–1392.

Daniel, J. (2021). Emoji frequency. *Unicode*. https://home.unicode.org/emoji/emoji-frequency/

Demattè, P. (2022). *The origins of Chinese writing*. Oxford University Press.

de Saussure, F. (2011). *Course in general linguistics* (P. Meisel, Ed.; W. B. E. by P. Meisel & H. Saussy, Trans.). Columbia University Press.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19* (10), 603–615. https://doi.org/10.1016/j.tics.2015.07.013

du Ponceau, P. S. (1838). *A dissertation on the nature and character of the Chinese system of writing. To which are subjoined a vocabulary of the Cochin Chinese language by J. Morrone [&c.]*. Published for the American Philosophical Society.

Ecklund, S., & Reichle, J. (1987). A comparison of normal children's ability to recall symbols from two logographic systems. *Language, Speech, and Hearing Services in Schools*, *18*, 34–40. https://doi.org/10.1044/0161-1461.1801.34

Edgerton, W. F. (1941). Ideograms in English writing. *Language*, *17*(2), 148–150. https://doi.org/10.2307/409622

Enfield, N. J. (2017). *How we talk: The inner workings of conversation*. Basic Books.

Frege, G. (1883). On the scientific justification of a concept-script (J. M. Bartlett, Trans.). *Mind; A Quarterly Review of Psychology and Philosophy*, *73*(290), 155–160.

Frege, G. (1920/1981). *Posthumous writing* (Illustrated ed.). John Wiley.

Funnell, E., & Allport, A. (1989). Symbolically speaking: Communicating with Blissymbols in aphasia. *Aphasiology*, *3*(3), 279–300. https://doi.org/10.1080/02687038908248995

Garrod, S., Fay, N., Lee, J., Oberlander, J., & MacLeod, T. (2007). Foundations of representation: Where might graphical symbol systems come from? *Cognitive Science*, *31*(6), 961–987. https://doi.org/10.1080/03640210701703659

Gelb, I. J. (1963). *A study of writing*. University of Chicago Press.

Guthrie, R. D. (2006). *The nature of Paleolithic art*. University of Chicago Press.

Helmke, C., & Nielsen, J. (2021). Teotihuacan writing: Where are we now? *Visible Language*, *55*(2), Article 2. https://doi.org/10.34314/vl.v55i2.4607

Hermalin, N., & Regier, T. (2019). *Efficient use of ambiguity in an early writing system: Evidence from Sumerian cuneiform*. Annual Meeting of the Cognitive Science Society. https://www.semanticscholar.org/paper/Efficient-use-of-ambiguity-in-an-early-writing-from-Hermalin-Regier/9a6898192ea9a9ad8f9d67fe95521a8532fdd3be

Hirshorn, E. A., & Harris, L. N. (2022). Culture is not destiny, for reading: Highlighting variable routes to literacy within writing systems. *Annals of the New York Academy of Sciences*, *1513*(1), 31–47. https://doi.org/10.1111/nyas.14768

Hoffmeister, R. J., & Caldwell-Harris, C. L. (2014). Acquiring English as a second language via print: The task for deaf children. *Cognition*, *132*(2), 229–242. https://doi.org/10.1016/j.cognition.2014.03.014

Hurlbut, B. I., Iwata, B. A., & Green, J. D. (1982). Nonvocal language acquisition in adolescents with severe physical disabilities: Bliss symbol versus iconic stimulus formats. *Journal of Applied Behavior Analysis*, *15*(2), 241–258. https://doi.org/10.1901/jaba.1982.15-241

Jenkins, R., Dowsett, A. J., & Burton, A. M. (2018). How many faces do people know? *Proceedings of the Royal Society B: Biological Sciences*, *285*(1888), 20181319. https://doi.org/10.1098/rspb.2018.1319

Koriat, A., & Levy, I. (1979). Figural symbolism in Chinese ideographs. *Journal of Psycholinguistic Research*, *8*(4), 353–365. https://doi.org/10.1007/BF01067139

Koshevoy, A., Miton, H., & Morin, O. (2023). Zipf's law of abbreviation holds for individual characters across a broad range of writing systems. *Cognition*, accepted.

Kozleski, E. B. (1991). Visual symbol acquisition by students with autism. *Exceptionality*, *2*(4), 173–194. https://doi.org/10.1080/09362839109524782

Lewis, D. (1969). *Convention: A philosophical study*. Wiley Blackwell.

Lewis, M. L., & Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition*, *153*, 182–195. https://doi.org/10.1016/j.cognition.2016.04.003

Luk, G., & Bialystok, E. (2005). How iconic are Chinese characters? *Bilingualism: Language and Cognition*, *8*(1), 79–83. https://doi.org/10.1017/S1366728904002081

Madden, J. R. (2008). Do bowerbirds exhibit cultures? *Animal Cognition*, *11*(1), 1–12. https://doi.org/10.1007/s10071-007-0092-5

Millikan, R. G. (1998). Language conventions made simple. *Journal of Philosophy*, *185*(4), 161–180.

Miton, H., & Morin, O. (2021). Graphic complexity in writing systems. *Cognition*, *214*, 104771. https://doi.org/10.1016/j.cognition.2021.104771

Mizuko, M. (1987). Transparency and ease of learning of symbols represented by Blissymbols, PCS, and Picsyms. *AAC: Augmentative and Alternative Communication*, *3*, 129–136. https://doi.org/10.1080/07434618712331274409

Mizuko, M., & Reichle, J. (1989). Transparency and recall of symbols among intellectually handicapped adults. *Journal of Speech and Hearing Disorders*, *54*(4), 627–633. https://doi.org/10.1044/jshd.5404.627

Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369* (1651), 20130299. https://doi.org/10.1098/rstb.2013.0299

Morin, O., & Miton, H. (2018). Detecting wholesale copying in cultural evolution. *Evolution & Human Behavior*, *39*(4), 392–401. https://doi.org/10.1016/j.evolhumbehav.2018.03.004

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. *Cognition*, *122*(3), 280–291. https://doi.org/10.1016/j.cognition.2011.10.004

Poupart, A., Trudeau, N., & Sutton, A. (2013). Construction of graphic symbol sequences by preschool-aged children: Learning, training, and maintenance. *Applied Psycholinguistics*, *34*(1), 91–109. https://doi.org/10.1017/S0142716411000622

Scott-Phillips, T. (2014). *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Palgrave Macmillan.

Sevcik, R. A., Barton-Hulsey, A., Romski, M., & Hyatt Fonseca, A. (2018). Visual-graphic symbol acquisition in school age children with developmental and language delays. *Augmentative and Alternative Communication*, *34*(4), 265–275. https://doi.org/10.1080/07434618.2018.1522547

Severi, C. (2019). Their way of memorizing: Mesoamerican writings and native American picture-writings. *Res: Anthropology and Aesthetics*, *71–72*, 312–324. https://doi.org/10.1086/706117

Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford University Press.

Tamariz, M., & Kirby, S. (2015). Culture: Copying, compression, and conventionality. *Cognitive Science*, *39*(1), 171–183. https://doi.org/10.1111/cogs.12144

Urton, G. (2017). *Inka history in knots: Reading khipus as primary sources*. University of Texas Press.

Winters, J., Kirby, S., & Smith, K. (2018). Contextual predictability shapes signal autonomy. *Cognition*, *176*, 15–30. https://doi.org/10.1016/j.cognition.2018.03.002

Winters, J., & Morin, O. (2019). From context to code: Information transfer constrains the emergence of graphic codes. *Cognitive Science*, *43*(3), e12722. https://doi.org/10.1111/cogs.12722

Xiao, W., & Treiman, R. (2012). Iconicity of simple Chinese characters. *Behavior Research Methods*, *44*(4), 954–960. https://doi.org/10.3758/s13428-012-0191-3

Young, A. W., Frühholz, S., & Schweinberger, S. R. (2020). Face and voice perception: Understanding commonalities and differences. *Trends in Cognitive Sciences*, *24*, 398–410. https://doi.org/10.1016/j.tics.2020.02.001

Youngblood, M., Miton, H., & Morin, O. (2023). Signals of random copying are robust to time- and space-averaging. *Evolutionary Human Sciences*, *5*, E10.