The puzzle of ideography

Olivier Morin

Behavioral and Brain Sciences (forthcoming)

This Target Article has been accepted for publication and has not yet been copyedited and proofread. The article may be cited using its doi (About doi), but it must be made clear that it is not the final version.

The puzzle of ideography

Olivier Morin

morin@shh.mpg.de

Max Planck Institut für Menschheitsgechichte

10, Kahlaische strasse, 07745 Jena, Germany.

Institut Jean Nicod, CNRS, ENS, PSL University

29, rue d'Ulm, 75005 Paris, France.

https://www.shh.mpg.de/94549/themintgroup

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 of the European Evolution and Human Behavior Association.

Short abstract (100 words):

Writing systems do not simply encode concepts directly: their symbols encode elements of a spoken language. Generalist, self-sufficient ideographies are rare. Why? Prevailing explanations hold ideographic systems to be either inconceivable or exceedingly hard to learn. Ideographic communication, this paper argues, is conceptually and cognitively plausible, but stifled by a standardization problem. Spoken or signed languages based on cheap and transient signals are easier to standardize. Graphic codes can only be standardized for a limited number of meaning-symbol mappings. Hence, graphic notations focused on a narrow domain, like linguistic units or numerical quantities, evolve; generalist ideographies do not.

Long abstract (250 words):

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 non-linguistic 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.

Keywords: Graphic communication; Literacy; Ideography; Numerical notations; Repair; Semasiography; Semiotics; Sign language; Writing.

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, since 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 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, Maya, 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 (Rossi 2000; Lin and Biggs 2006). 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 travelled. 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 et al., 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 due to 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, etc. 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 repurposed 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 et al., 2006; Kelly et al., 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 (Section 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, due to 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), etc. For the sake of this argument, different terminological choices were made.

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

A code is a set of conventional associations between meaning and signals (Saussure, 2011; Scott-Phillips, 2014). Musical notations, road signs, writing systems, etc. are codes in this sense, and so are languages like Swahili, French, etc. 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 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 et al. 2015; Regier, Kemp, and Kay 2015; Winters and Morin 2019; Scott-Phillips 2014). 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 non-conventionalized ones.

This was shown elegantly in a series of experiments by Garrod and colleagues (Garrod et al., 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 et al., 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 et al., 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 non-conventional 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

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

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 non-codes 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 pre-existing 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).

[1] This is clearly shown by the sales figures for translated vs. untranslated manga outside Japan.

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

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 et al., 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 non-coded 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 non-codes (**Figure 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, etc., 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, etc. – can be studied with the tools of linguistic (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.

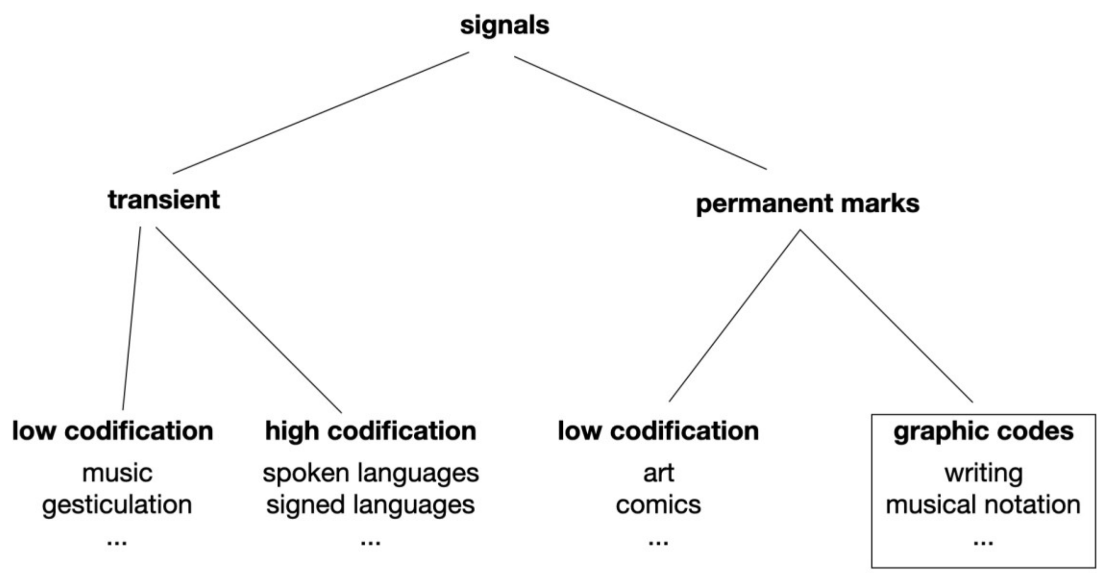

Figure 1. The place of graphic codes in human communication.

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 McCawley 1992, Casati 2017) that these codes resemble language in many ways: they carry meaning,

[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.

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

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.

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 notation 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 language 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 et al., 1991; Sevcik et al., 2018): 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.

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 1989's 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. A script can share this property, but only to the extent that it somehow transcribes its spoken-language base.
>
> (Liberman 1992:120).

[3] I am indebted to Nick Enfield for this observation.

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

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 to 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 1996:309). Examples in use in English writing include Arabic numerals (1, 2, 3, 4…), as well as conventional symbols such as £, ±, ☺, ℃, ©, &, etc. 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

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

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”, etc., 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 et al., 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 /nuεv/, 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.

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

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” – Gelb 1963; Croft 2017; Boone and Mignolo 1994) is simply a set of ideographic symbols used in conjunction with each other. Examples

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

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. du Ponceau 1838; Boltz 1993, Hill 1967), such systems are not 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 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.

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

	specialized	**generalist**
self-sufficient	**Graphic codes:** **Modern musical notation** **Modern mathematical notations (including logic)** **Coin designs** **Heraldic emblems**[1] ...	**Graphic codes:** **Few clear examples (Nsibidi symbols[3]?)**
	Other codes: *Hunting signs (e.g. Botswana[2])* *Sawmill workers' signs[2]* ...	*Other codes:* *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	**Graphic codes:** **Early musical notations**[6] **Kupesi symbols**[7] **Sand drawings (e.g. Australia[8], Vanuatu[9])** ...	**Graphic codes:** **Bliss symbolics**[10] **Pictographs (as used by, e.g., Cuna shamans[11])** **Australian message sticks**[12] **Yukaghir drawing game (the "love letter"[13])** ...
	Other codes: *Few clear examples*	*Other codes:* *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. A 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 will be answered in section 4.

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, etc. 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, etc. Others are

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

apparently more versatile — for instance the international airport signs for “toilets”, “wifi”, “luggage”, etc. — 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 et al., 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

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

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 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) are 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 and 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 connections, personal emblems, the sounds of a language, etc. 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 and 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.); 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. du Ponceau 1838; DeFrancis 1989, 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 De Francis 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

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

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: Hyman 2006; Gelb 1963) 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(:)/. 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.

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

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 counter-examples 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 (Unger and DeFrancis 1995; Sampson 2017; Handel 2015, Lurie 2006). 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: 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 2015:130 – see also DeFrancis 1989; Unger and DeFrancis 1995; Sampson 1985; Sampson 2017, Coulmas 1996). Nor is this phonetic information inert: there is massive psychological and neuropsychological evidence that Chinese readers process writing using phonological cues (Dehaene, 2010; Y. Li et al., 2014; Liu et al., 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 which 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.

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

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%) are used in a properly ideographic way, i.e., 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 on top of 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 the 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 and 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, in press; 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 in press).

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

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

(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, while 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 USA. The signed languages developed in silent monastic communities (Banham, 2015; Quay, 2015) belong in 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

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

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 across space (Morin et al., 2020). Asynchronous messages are "temporally and spatially portable" (to use the terminology of Pickering and Garrod 2021), or "location and time independent" (to use that of Lee and Karmiloff-Smith 1996). The impact of asynchronous communication on cultural evolution is twofold. 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). Secondly, it allows one single message to transmit the same piece of information multiple times, in contrast with 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.

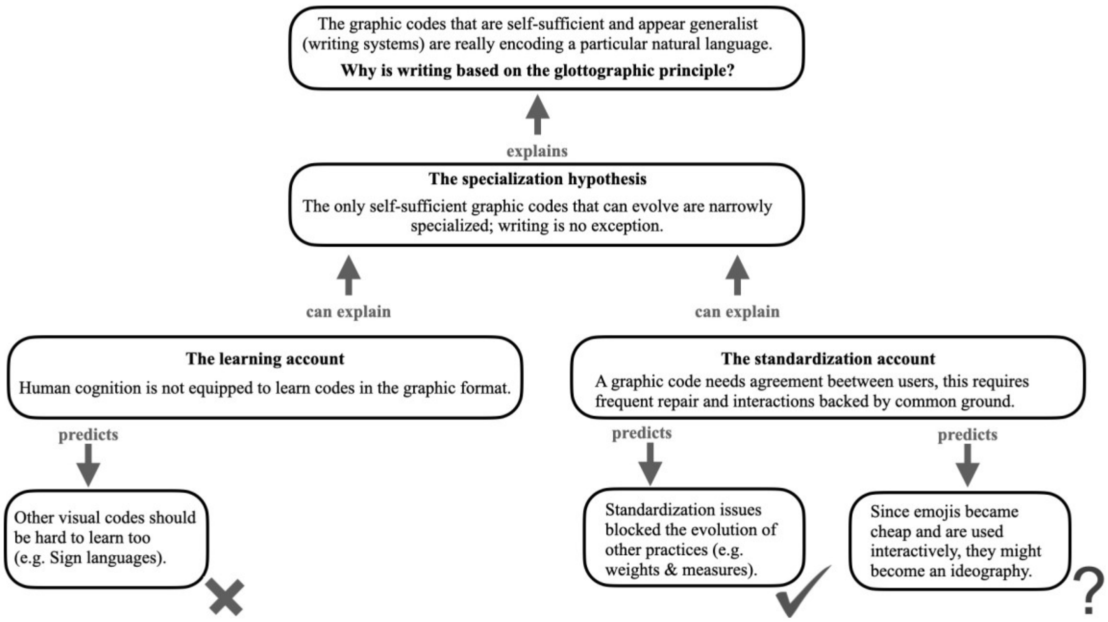

Figure 3. The argument of sections 5 to 6.

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 (compared to strings of phonemes, or bundles of gestures). On the standardization account, that is due to 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 which 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.

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

The way that the two accounts solve the puzzle of ideography is broadly the same: graphic codes are specialized because graphic codes 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 artefact, has no biological basis specifically evolved to support it (Dehaene, 2010). But stopping there would beg the question: 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.

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

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.

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

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. 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", i.e. "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.

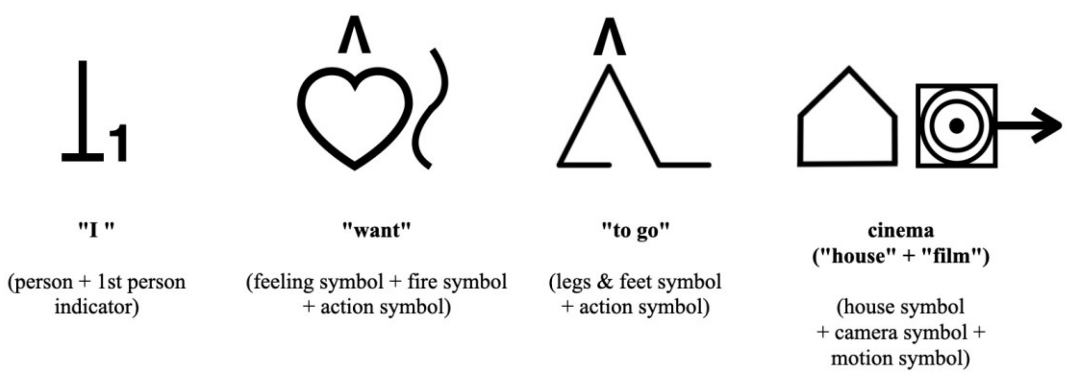

Figure 4. The phrase "I want to go to the movies" in Bliss symbolics. The Bliss symbols are reproduced by the author, copying an image from Wikimedia commons (Wikipedia 2022).

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

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

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 and 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 and Miton 2018). Combinatorial structure and compositionality are evident in writing systems, mathematical or musical notations, formal logic (Zalta, 2022), etc. 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. Since 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 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 et al., 2018; Yin & Zhang, 2021); and printed or computer-written fonts are modelled after handwritten symbols.

Thus, it seems that Liberman' 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 language 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

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

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 non-linguistic 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,

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

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 (David, 1985; David & 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 20th century (Cowan & Hultén, 1996).

A very similar problem affects the codes that we use to communicate. Since 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 modelled by assuming that each individual speaker is disproportionately more likely to use the words most frequently used by others (Pagel et al. 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 et al., 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, etc. 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, since absolute units require standardization of measuring instruments to a much more precise degree than relative units (Riggsby 2019: 86 sq.). Even for a large state with an advanced bureaucracy and a sophisticated literate culture, the challenge raised by standardization were serious enough to block the development of communication devices.

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

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 PC 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 personal computer will not allow, things which arcade or virtual reality hardware makes possible that a console cannot, etc. 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, since 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, since 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

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

transient symbols. Spoken or signed messages require little effort or time to produce and vanish soon once they are emitted (Galantucci et al., 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 et al., 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 at the level of the pair, but also at the level of entire populations, because turn-over in conversation partners leads to convergence to broader standards that everyone can share (Guilbeault et al., 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, since any signal they exchange becomes part of this common ground (Clark, 1996).

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

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 et al., 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; etc. In Clark & 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. Secondly, 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.

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

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 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, while 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); interactive alignment (repetition or imitation of speech at various levels) (Dideriksen et al., 2020; 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, since interlocutors do not inhabit the same time frame. 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 non-verbal or para-verbal channel through which to convey them, whereas conversation has eyebrow flashes, hmmming, etc. (Bavelas & Gerwing, 2011).

Access to common ground information and its exploitation through repair, backchannel communication and interactive alignment reinforce 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. Since 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 fulfils 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 fulfil 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, etc. 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 exerts over the development of signed language is suggested by a series of studies by Goldin-Meadow and collaborators (Goldin-Meadow et al., 1996, 2017; see also McNeill, 1992), who asked hearing participants to describe a scene using only gestures; compared to the co-speech 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 fulfil 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, etc. 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 ideographic communication in the digital age. Digital communication might overthrow the constraints that weigh on the evolution of

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

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 et al., 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 et al., 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 system (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, were 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 non-existence 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.

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

In the history of writing, the non-evolution 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 et al., 2016; Dennett, 1995; Mesoudi & Thornton, 2018). A design space (modelled 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, while 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.

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

Acknowledgements. 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, Elia Weber, for their helpful comments on an earlier draft. I thank the BBS reviewers, including Nick Enfield, for their input.

Funding statement. 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.

Conflicts of interest. I have no conflict of interest to declare.

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

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

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 and Schuster.

Baines, J. (2007). *Visual and Written Culture in Ancient Egypt*. OUP Oxford.

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). In *Wikipedia*. 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*. Amer 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*. 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*. The university of Chicago 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 UK. 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), E117-139. 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*, *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. Resnick, L. B, M. John, S. Teasley, & D. (Eds.), *Perspectives on Socially Shared Cognition* (pp. 13–1991). 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 *Schrift und Schriftlichkeit* (pp. 1380–1387). De Gruyter.

Coulmas, F. (2002). *Writing Systems: An Introduction to Their Linguistic Analysis*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164597

Coulmas, F. (2003). *Writing Systems: An Introduction to Their Linguistic Analysis*. Cambridge University Press.

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. https://doi.org/10.1017/9781107294264.005

Częstochowska, J., Gligorić, K., Peyrard, M., Mentha, Y., Bień, M., Grütter, A., Auer, A., Xanthos, 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. 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.

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. JSTOR. https://doi.org/10.2307/2843186

DeFrancis, J. (1989a). *Visible Speech: The Diverse Oneness of Writing Systems*. University of Hawaii Press.

DeFrancis, J. (1989b). *Visible Speech: The Diverse Oneness of Writing Systems*. University of Hawaii 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 and 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

Dideriksen, C., Christiansen, M. H., Tylén, K., Dingemanse, M., & Fusaroli, R. (2020). *Quantifying the interplay of conversational devices in building mutual understanding*. PsyArXiv. https://doi.org/10.31234/osf.io/a5r74

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.). Published for the 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. (2004a). Why is conversation so easy? *Trends in Cognitive Sciences*, *8*(1), 8–11.

Garrod, S., & Pickering, M. J. (2004b). 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. (2019a). Emoji as digital gestures. *Language@Internet*, *17*(2). https://www.languageatinternet.org/articles/2019/gawne

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

Gelb, I. J. (1963a). *A Study of Writing*. Univ of Chicago Pr.

Gelb, I. J. (1963b). *A Study of Writing*. Univ of Chicago Pr.

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*. 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*. 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). 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*, 1–46. https://doi.org/10.1017/S0140525X22000012

Hirshorn, E. A., & Harris, L. N. (in press). Culture is not destiny, for reading: Highlighting variable routes to literacy within writing systems. *Annals of the New York Academy of Sciences*, *n/a*(n/a). 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

Joyce, T. (2011). The significance of the morphographic principle for the classification of writing systems. In *Typology of writing systems*.

Kelly, P. (2019). Australian message sticks: Old questions, new directions. *Journal of Material Culture*, 1359183519858375. 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. JSTOR. 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*. 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. https://doi.org/10.1016/S0166-4115(08)62794-6

Lin, T.-J., & Biggs, M. (2006). *A preliminary study of learnable pictogram languages*. IADE. 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.

Macgregor, J. K. (1909). Some Notes on Nsibidi. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, *39*, 209–219. JSTOR. 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 GmbH & Co KG.

Mallery, G., & Ethnology, S. I. B. of A. (1886). *Pictographs of the North American Indians: A preliminary paper*. Govt. 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 *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., Cheuk, C. S.-M., & 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 Publishing Company.

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. JSTOR.

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': 10th International Conference on Web and Social Media, ICWSM 2016. *Proceedings of the 10th International Conference on Web and Social Media, ICWSM 2016*, 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*, *n/a*(n/a). 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*. 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*, *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 édition). 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. 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).

Regier, T., Kemp, C., & Kay, P. (2015). Word meanings across languages support efficient communication. In *The Handbook of Language Emergence* (pp. 237–263). John Wiley & Sons, Ltd. 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 Publishing Limited.

Sandler, W. (2009). What comes first in language emergence? In N. J. Enfield (Ed.), *Dependencies in language* (pp. 63–84). Language Science Press.

Sandler, W., Aronoff, M., Meir, I., & Padden, C. (2011). The gradual emergence of phonological form in a new language. *Natural Language & Linguistic Theory*. https://doi.org/10.1007/S11049-011-9128-2

Saussure, F. de. (2011). *Course in General Linguistics*. Columbia University 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. 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* (First Edition edition). 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. (2014). *The Psychology of Literacy* (Reprint 2014 ed. édition). Harvard University Press.

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

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

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, USA.

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* (2d edition). Blackwell.

Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J. P., Yoon, K.-E., & 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 Publishing.

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*, 859–866. 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. 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., Liu, G., & 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* (1 edition). 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*, 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. https://plato.stanford.edu/archives/fall2022/entries/frege/

Zhang, H., & Mace, R. (2021). Cultural extinction in evolutionary perspective. *Evolutionary Human Sciences*, *3*. https://doi.org/10.1017/ehs.2021.25