

Ambiguity and the creation and evolution of writing systems

Jordi Fortuny

Department of Catalan Philology and Linguistics. University of
Barcelona

AWLL15. Pisa, October 29-31, 2025

Project PID2023-146903NB-I00 supported by:



On linguistic ambiguity

(1) Homonyms

- a. A Coke **can**
- b. A He **can** swim

(2) Polysemy

- a. He spent one **day** in Paris (\approx 24 hours)
- b. He only leaves his house during the **day** (vs. night)
- c. His **day** begins at six (\approx awake)
- d. I work a ten-hour **day** (\approx time at work)

(3) Syntactic ambiguity

He likes boys and girls **from Paris**

On linguistic ambiguity

Pessimistic view

“If the language assigns a unique denotation to each signal, then the interpreter’s task is relatively simple. **If signals may have multiple denotations**, then interpreters may try to disambiguate, so as to infer the producer’s intention. This entails **extra effort** on the part of the interpreter. Moreover, if the methods of disambiguation are fallible, **then ambiguity increases the chance of miscommunication.**” (Wasow et al. 2005)

On linguistic ambiguity

Optimistic view

- There is no estimate of the amount of ambiguity that triggers misunderstandings.
- Hearers are unaware of the “extra effort” that linguists attribute to ambiguity.
- Lexical ambiguity reduces the size of the lexicon (Brochhagen & Boleda 2022) and syntactic ambiguity the length of utterances (Fortuny 2024).
- Ambiguity is an unavoidable property if coding and decoding complexities are balanced (Fortuny & Corominas 2013: 264).
- Ambiguity is “a central property if one wants to understand how language is designed” (Fortuny 2024).

On the ambiguity of writing systems

Thesis 1. The awareness of linguistic ambiguity was opportunistically used to achieve full writing (cf. Mattingly 1992).

Thesis 2. The **graphic multivalence** of a script is not isomorphic to the ambiguity of its underlying language.

Thesis 3. All scripts display **graphic multivalence** (Appendix)

On graphic multivalence

- Homophony (rebus writing)
- Initial homophony (acrophony)
- Polyphony
- Graphic multivalence reduction/generation
 - 4.1 Determinatives
 - 4.2 Phonetic complements
- 5. Multiplicity of patterning
- 6. Orthographical disharmonies (appendix)

Thesis 1

The awareness of linguistic ambiguity was opportunistically used to achieve full writing

Homophony (rebus writing): a graph for a word *a* is also used to represent a semantically unrelated homophonic linguistic form *b*

Sumerian; Woods (2010, 2020)

- 𐎶𐎵 *ti* ‘arrow’ > *ti* ‘life’
- 𐎮𐎠𐎺 *gu* ‘ox’ > syllable *gu*
- 𐎠𐎺𐎠 *gi* ‘reed’ > *gi* ‘to render’
- 𐎠𐎺𐎠𐎶 *ak* ‘do/make’ > *ak* (GEN suffix)
- 𐎶𐎠 *e* ‘house’ > *e* (ERG suffix)
- 𐎶𐎠𐎶 *a* ‘wàter’ > *a* (LOC suffix)

Chinese (Boltz 1994, 1996, 1999, 2000)

- 之 *zhī* ‘to go’ > *zhī* GEN marker
- 王 *wǎng* ‘king’ > *wáng* ‘go toward’
- 來 *lái* ‘wheat’ > *lái* ‘come’
- 自 *zì* ‘nouse’ > *zì* ‘self’
- 其 *qí* ‘basket’ > *qí* ‘his’, ‘her’, ‘its’
- 者 *zhě* ‘to roast’ > *zhě* nominalizer

Homophony (rebus writing): a graph for a word *a* is also used to represent a semantically unrelated homophonic linguistic form *b*

Egyptian (Gardiner 1957)

graph	logogram	phonogram
	<i>pr</i> 'house'	<i>pr</i>
	<i>ḥr</i> 'head'	<i>ḥr</i>
	 <i>jrt</i> 'eye'	<i>jr</i>

Initial homophony (acrophony): a pictogram for a word is used to express its initial segment

Maya syllabograms; proto-Ch'olan forms (Macri & Loper 2003, Valério & Ferrara 2019)

 **kay* 'fish' > *ka* (syllabogram)

 **k'uʔ* 'nest' > *k'u*

 **naʔ* 'mother' > *na*

 *nok* 'cloth(es)' > *no*

Egyptian phonograms (Gardiner 1957, Hoch 1998, Valério & Ferrara 2019)

 ʕ(V) 'vulture' (Old Egyptian) > ?

 *raR* 'mouth' (Old Egyptian) > *r*

 *bw* 'place' (Old Egyptian) > *b*

 *nt* 'waves' (Middle Egyptian) *n*

Initial homophony (acrophony): a pictogram for a word is used to express its initial segment

Egyptian	Protosinaitic	Reconstructed word	Acroph. value	Phoenician	Greek
		* <i>alp</i> , 'ox'	ʔ		Α
		* <i>bayt</i> 'house'	b		Β
		* <i>gaml</i> 'throw-stick'	g		Γ
		* <i>dalt</i> 'door'	d		Δ
		* <i>haw</i> 'man calling'	h		Ε
		* <i>waw</i> 'hook'	w		Υ
		* <i>kap</i> 'palm'	k		Κ
		* <i>lamd</i> 'goad'	l		Λ

Conclusion (I): on ambiguity as the key to full writing

1. “Rebus tends to appear as the *first choice* for phonetization, as it is direct, intuitive, and economical. [...] Acrophony is the alternative to rebus, as it is indirect and requires some linguistic/phonological awareness.”
2. “Rebus is used to the highest extent possible, depending on the amount of homophony in the underlying language. When rebus is no longer possible, acrophony becomes necessary.”
3. “Scripts invented for languages with high monosyllabicity (predominantly monosyllabic **morphemes**) use rebus exclusively (Old Chinese, Sumerian).”
4. “Most other cases use rebus supplemented by acrophony (Egyptian, Anatolian Hieroglyphic, Nahuatl) and acrophony only (Early Maya, Proto-Sinaitic).”

Valério & Ferrara (2019)

Polyphony: a logogram for a word *a* is also used to represent a word *b* when *b* is semantically related to *a*

✳ *an* ‘sky’/‘heaven’ > *dingir* ‘god’ (Sumerian; Coulmas 2003)

𐎧𐎠𐎫 *ka* ‘mouth’ > *inim* ‘speech, word’ > *dug* ‘to speak’

𠂔 *kǒu* ‘mouth’ > *míng* ‘call out’ (Shang script; Boltz 1996)

𠂔 *mù* ‘eye’ > *jiàn* ‘to see’

☉ *r* ‘sun’ > *hrw* ‘day’ (Egyptian; Gardiner 1957)

👂 *msḏr* ‘ear’ > *sḏm* ‘to hear’

† *b* ‘mace’ > *ḥrp* ‘govern’

Thesis 2

The graphic multivalence of a script
is not isomorphic to the ambiguity of
its underlying language

Graphic multivalence reduction: determinatives in Sumerian (Coulmas 2003; Woods 2010, 2020)

✳ (*dingir*, ‘god’) > DET before the name of a god

^dInanna = the goddess Inanna

🏰 (*Ki*) > DET after a toponym

Ur^ki = the city of Ur

🎵 (*lu*, ‘profession’) > DET before the name of a profession

^lúkurgarra = a temple musician

🏠 (*é*, ‘house’, ‘building’) > DET before the name of a building

^ékur = the temple of Enlil (at Nippur)

Graphic multivalence reduction

Phono-semantic compounds in Chinese

1. 木 *mù* 'tree' >_{rebus} *mù* 'to wash (oneself)'

2. 木 *mù* 'wash (oneself)' + 氵 (det. 'water') = 沐 *mù* 'to wash (oneself)'

1. 马 *mǎ* 'horse' >_{rebus} *mā* 'mother'

2. 马 *mā* 'mother' + 女 det. 'female' = 妈 *mā* 'mother'



Determinatives in Egyptian

det. for masculine names

det. for names of buildings

det. for activities involving the mouth, head or ideas

det. for god's names

Multiplicity of patterning

Sumerian

- 𒄠: -logogram *mātu* 'land'
 - phonogram *mat*
 - DET for names of mountains
- 𒀭: -logogram *an* 'sky'
 - phonogram *an*
 - DET for names of deities
- 𒄣: -logogram *gis* 'wood'
 - phonogram *iz*
 - DET for names of wooden objects

Egyptian

- 𓂏: -logogram *pr* 'house'
 - phonogram *pr*
 - DET for names of buildings
- 𓂐: -logogram *nmtt* stride; *jw* come
 - phonogram *jw*
 - DET for words involving motion

Multiplicity of patterning and graphic composition in Chinese

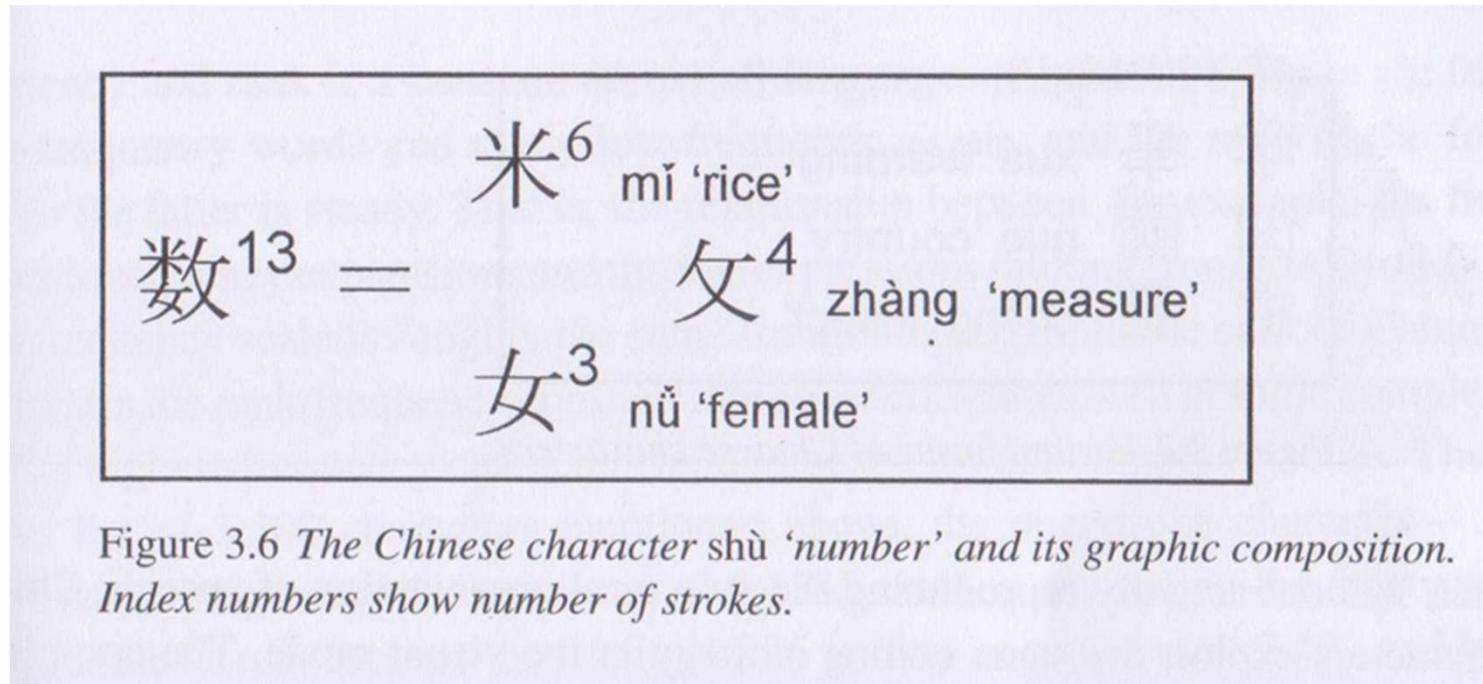
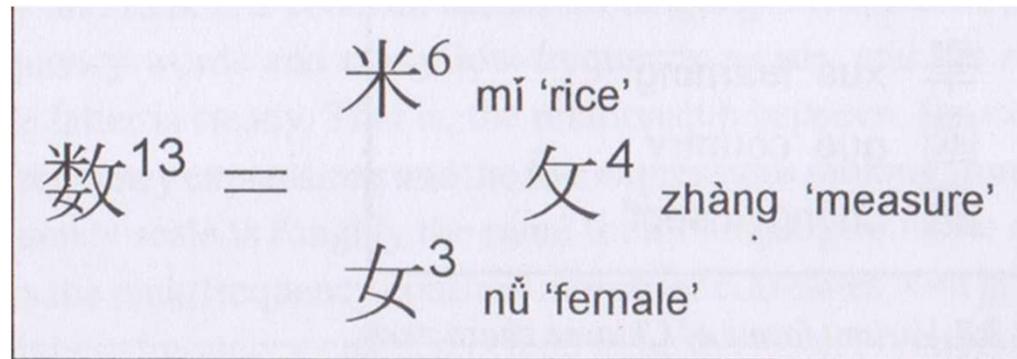


Figure 3.6 *The Chinese character shù 'number' and its graphic composition. Index numbers show number of strokes.*

Coulmas (2003)

Conclusion (II): on non-isomorphism

- Polyphonic characters ≠ polysemic words: ☉ *r*‘sun’ > *hrw* ‘day’
- A logogram and a homographic phonogram may not be strict homophones: ✧ *mātu* (‘land’) ≠ *mat* (syllable)
- Determinatives and phono-semantic compounds have no correspondence in speech (speech presents no discrete units to avoid ambiguity –Fortuny (2024))
- Determinatives create a graphic multivalence with no correlate in linguistic ambiguity
- Graphic composition ≠ layered organization of speech units (*shù* ‘number’)



A remark on theses 1 and 2

The awareness of linguistic ambiguity but also **the generation of graphic multivalence** (determinatives, graphs belonging to multiple patterning) are key factors in the development of logo-syllabic scripts

Selected references

- Boltz, W. (1994). *The origin and early development of the Chinese writing system*. American Oriental Series 78. American Oriental Society.
- (1996). Early Chinese writing. In P. T. Daniels & W. Bright (eds.), *The world's writing systems*. Oxford University Press.
- (1999). Language and writing. In M. Loewe & E. L. Shaughnessy (eds.), *The Cambridge History of Ancient China*. Cambridge University Press.
- (2000). *Monosyllabicity and the origin of the Chinese script*. Berlin: Max Planck Institute.
- Brochhagen, T. & G. Boleda (2022). When do languages use the same word for different meanings? The Goldilocks principle in colexification. *Cognition* 226.
- Cooper, J. (1996). Sumerian and Akkadian. In P. T. Daniels & W. Bright (eds.), *The world's writing systems*. Oxford University Press.
- Coulmas, F. (2003). *Writing systems*. Cambridge University Press.
- Fortuny, J. (2024). Structural ambiguity and the architecture of language. *Studia Linguistica* 1: 35-63.
- Fortuny, J. & B. Corominas (2013). On the Origin of Ambiguity in Efficient Communication. *Journal of Logic, Language and Information*: 22(3): 249-267.
- Fortuny, J. & Ll. Payrató (2024). Ambiguity in linguistics. *Studia Linguistica* 1: 1-7.
- Gardiner, A. (1957). *Egyptian Grammar: Being an introduction to the study of hieroglyphs* (3rd edition). Griffith Institute, Ashmolean Museum.
- Hoch, J. E. (1998). *Middle Egyptian Grammar: sign list*. Society for the study of Study of Egyptian Antiquities. Benben Publications.
- Macri, M. & M. G. Looper (2003). *The new catalog of Maya hieroglyphs. Volume 1 The Classic period inscriptions*. University of Oklahoma Press.
- Mattingly, I. (1992). Linguistic awareness and orthographic form. In R. Frost & L. Katz (eds.), *Orthography, phonology, morphology, and meaning* (p. 11-26). North-Holland.
- Suomi, K. J Toivanen & R. Ylitalo (2008). *Finnish Sound Structure: Phonetics, Phonology, Phonotactics and Prosody*. *Studia humaniora ouluensia* 9. Oulu: Oulun.
- Valério, M. & Ferrara (2019). Rebus and acrophony in invented writing. *Writing Systems Research* (11, 1): 66-93.
- Wasow, T. A. Perfors & D. Beaver (2005). The Puzzle of Ambiguity. In O. Orgun & P. Sells (eds.), *Morphology and the Web of Grammar*. CSLI Publications.
- Woods, C. (2010). *Visible Language: Inventions of Writing in the Ancient Middle East and Beyond*. Oriental Institute Museum Publications.
- (2020). The Emergence of Cuneiform Writing. In R. Hasselbach-Andee (ed.), *A Companion to Ancient Near Eastern Languages*. John Wiley & Sons, Inc.

Appendix. Thesis 3

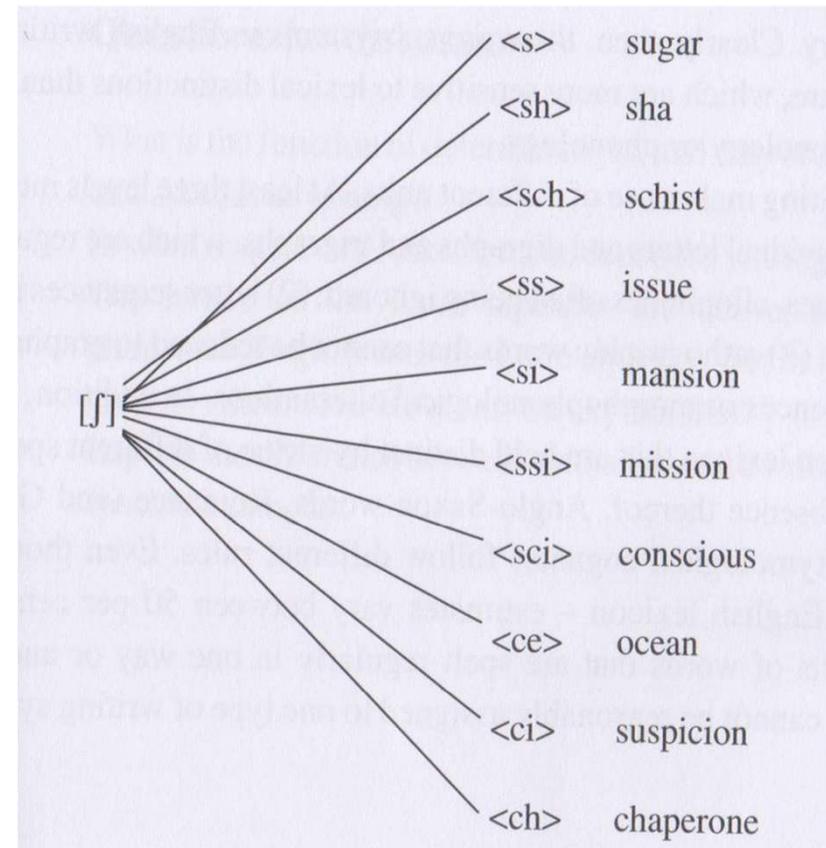
All scripts display graphic
multivalence

Orthographical disharmonies: English

Ways of spelling /u:/ in English

<u>	truly
<o>	do
<oe>	shoe
<oo>	soon
<ue>	true
<ui>	lawsuit
<ou>	routine
<wo>	two
<ew>	screwed
<ewe>	jewel
<oeu>	manoeuvre
<ous>	rendezvous
<ough>	throughout
<oups>	coups

English /ʃ/ and its graphemes



Orthographical disharmonies: Finnish 1:1 'sound'-letter correspondence (Suomi, et al. 2008)

1 sound : 2 letters

tuli 'fire' /'tuli/

tulli 'customs' /'tuɫi:/

<ll>: [l:] (<ll>: *[ll])

kengät 'shoes'

/keŋ:æt/

<ng>: [ŋ:] (<ng>: *[ng])

1 letter : two sounds (allophony)

onpa se hyvä ('That's good')

['ompɑ se 'hyvæ]

<n>: [n], [m], ...

But: <n>: /n/

zero letters : 1 sound

Spoken: *Mä en [ʔ] tiedä* ('I don't know')

Standard written form: *Mä en tiedä*

Conclusion (III): on the universality and variability of graphic multivalence

Script type		Graphic multivalence manifestations
Logosyllabaries	Sumerian	Polyphony, homophony, multiplicity of patterning
	Chinese	
Alphabets	English	orthographic disharmonies
	Finnish	
Logosyllabary + segmental component	Egyptian	Polyphony, homophony, acrophony, multiplicity of patterning; orthographic disharmonies