On the Writing System of Arabic: The Semiographic Principle as Reflected in Nasḫī Letter Shapes

Joseph Dichy

Abstract. Arabic writing as we know it was codified between the 1st/7th and 3rd/9th centuries. It is much more ancient and had appeared with another type of letter shapes akin to South Arabian writings around 800 years BC (Robin, 1991). The script borrowed the basis of its letter shapes from Nabataean writing (Baʿalbakī, 1981; Healey and Smith, 2009). Arabic script became more rational and regular due to what S. Auroux called, with reference to the French language, *grammatization processes*, which can be illustrated with Arabic, due to the very rich medieval sciences literature that was contemporary or immediately subsequent to its codification.

This paper is concerned with one aspect of the grammatisation process, that of letter shapes in the *Nasḫī* style of calligraphy, which was codified by Ibn Muqla (4th/10th cent.) and his followers. The paper presents the basic drawings included in the building of letters as designed by Ibn Muqla. It also highlights the fact that the relatively small number of shapes in Arabic is due to the cursive line inherited from Nabatean script. The graphic word-form with its complex set of morphological structures (extensively described in Dichy 1990a; 1997a), was another inheritance of earlier Semitic writing tradition (which isolated word-forms since Phoenician writing). Word-form structure resulted in a special style of final letters shapes, which exist in Hebrew with five letters, but were systematized in Arabic in a way that produced a basic opposition between initial and medial letter-shapes on the one hand, and final (end-of-word) shapes on the other. The result, including a few letters that escaped the opposition, is also presented.

The overall view highlights, in the author's theory of writing (Dichy, 2019), the way in which the *Semiographic Principle* parallels the *Phonographic Principle* in the writing of the Arabic language. This approach considers writing systems as analytic (Dichy, 2017). This means that the writing system can be considered as a collective cognitive analysis of the oral language. Analytic results are then projected on a writing support according to the semiographic structures of the script. The orthography of the language features the complex relations between the two principles. One point underlying the present paper is that the collective cognitive developments associated with a writing system are related to gramma-

Joseph Dichy
Professor of Arabic Linguistics Lyon - France
joseph.dichy@yahoo.fr

Y. Haralambous (Ed.), *Graphemics in the 21st Century. Brest, June 13-15, 2018. Proceedings*
Grapholinguistics and Its Applications (ISSN: 2534-5192), Vol. 1.
Fluxus Editions, Brest, 2019, p. 257–268. https://doi.org/10.36824/2018-graf-dich
ISBN: 978-2-9570549-0-9, e-ISBN: 978-2-9570549-1-6

tisation processes, which occur in a given historic period (there can therefore be successive grammatisation periods in a given language and culture).

1. Introduction: The Analytic Approach to the Emergence of a Writing System

This contribution follows previous works on the writing system of Arabic (based on Dichy 1990a,b) and the development of a cognitive approach to the emergence of a given script, which gives birth to a writing system though a metalinguistic process, the first insights of which can be found in Condillac's works (1746; 1775), and in the 20th century, in those of Vygotsky (1934). This overall view has been recently presented in Dichy (2017; 2019) in a more elaborate version, and can only be hinted at here. In very short words, the metalinguistic process of scriptural analytics operates along two interwoven lines:

- *Analytic operations*—the object of which is the oral language at phonetic, morphological and morphotactic levels—, result in *segmentation* and *identification* processes based on what I have described as the *Inventory Principle (IP)*. (In alphabetic writings, the Inventory Principle is cenemic, with a number of pleremic adjustments.)
- *Synthetic operations* project the result of analytic operations on writing supports (stone, parchment, paper, and nowadays screen) according to the conventions proper to the *semiographic characterisation* of a given language/script, i.e., according to the *Semiographic Principle (SP).*

This paper is concerned with one of the aspects of the metalinguistic processes that lead to the emergence of a writing system, namely, the shapes of letters considered as the result of synthetic operations leading to the semiographic characterisation of Arabic. This is related to the visual characterisation of both letters and words.

1.1. Short Recall of a Few Historical Facts about the Arabic Writing System

We leave it, with great relief and pleasure, to our Colleague Dr. Kamal Mansour to present, in this conference (Mansour, in this volume), the origins of Arabic Script, which is the latest of the Semitic scripts.

Written Arabic is in fact much more ancient than the writing we currently know and had appeared with another type of letter shapes

around the early 1st millennium BC[1]. The older script was borrowed from South Arabian writings. The writing in use to-day, the earliest form of which seems to go back to the 4th century CE for the least—i.e., to some 200 years before Islam—borrowed the basis of its letter shapes from Nabatean writing (Baʿalbakī 1981; Healey and Smith 2009). This entailed a good deal of adaptation, because the Syriac writing from which Arabic script was borrowed only had 22 graphemes (Robin 1991; Baʿalbakī 1981; Dichy 1990b).

The emergence of the latter into a codified writing system is related to considerable historiographic and linguistic developments that give a significant amount of information on its early traditions and progresses (Dichy 1990a,b; Abbott 1939b). On the other hand, the art of writing Arabic—which parallels that of good saying—has known remarkable developments since the beginning of Islam. The 4th/10th century[2] vizier Ibn Muqla (d. in 940) has authored, on the basis of his own experience as a calligrapher, a famous epistle on writing, which has been a basis for later Arabic calligraphy[3], and will be used extensively in the main sections of this contribution.

The emergence of Islamic culture was related to heavy needs in written documents. In addition to the Qurʾan and the Ḥadīṯ (the teachings of the Prophet of Islam), one witnessed great development in medieval sciences (mathematics, astronomy, medicine and surgery, logics, grammar and lexicography, religious sciences..., etc.), as well as many translations (mainly from Syriac and Greek). On the other hand, the emerging Islamic culture needed to ensure knowledge and transmission of the language of what we would call to-day the linguistic community of the Prophet of Islam. Basra and Kufa scholars, living in Iraq, endeavoured to compile the corpus of the poetry of the ancient Arabs in *dīwān-s* (poetic records). The compilation ensured contextualisation of the words found in the Koran and Ḥadīṯ, and a far-reaching knowledge of the language. In addition, words were extensively recorded in Arabic dictionaries from al-Ḫalīl (d. ar. 175 H/792 G) onwards. It is to be noted that, in addition to his role in the devising of the diacritic representation of vowels (Abbott, 1939a), al-Ḫalīl elaborated a method for the generation of

1. Ch. Robin has highlighted the fact, based on recently discovered inscriptions going back as far as the 8th century BC, that Arabic had been written using a South-Arabian alphabet, which was more adapted to its structure than the Syriac script, which was later borrowed (Robin, 1991, pp. 127–129).

2. Traditionally, the first date is that of the Islamic calendar ('H' for Hegira) and the second is that of common Gregorian years (symbol 'G').

3. On Ibn Muqla, see Abbott (1939a), and Osborn (2017, pp. 15–16). The first eight pages of Ibn Muqla's Epistle are reproduced in *fac simile* (after the 1663 manuscript of the Cairo National Library) with a commentary, in Massoudy (1981, pp. 40–41). One finds many quotations of Ibn Muqla in al-Qalqašandī (vol. 3, p. 23 onwards, chapters on writing, letter shapes and calligraphy).

virtual 'constructs' (in our terms, 'roots') that allowed building exhaustive dictionaries of the language of the Ancient Arabs, and gave birth to the first dictionary aiming at including the whole lexicon of a given language in the history of mankind (Dichy, 2014). An elaborate and reliable script was obviously the basis of the emerging written transmission of knowledge, which paralleled, during the first centuries oral transmission (Schoeler, 2002).

An essential development complemented the emergence of the Arabic script: conventions were needed in order to stabilise the reading of the Koranic text. In the 2nd/8th century, al-Ḫalīl devised a system of very comprehensive secondary diacritics (Abbott 1939b; Dichy 1990b), which is still in use to-day.

1.2. Illustrating the Processes That Lead to One Crucial Aspect of the Writing System

How did the writing system of Arabic emerge? To this question, one can only give conjectural answers. The fact is, writing became more rational and regular after its first period, due to what Auroux (1998; 2010) called, with reference to the French language, grammatization processes. This hypothesis can be illustrated with Arabic, on the basis of descriptive texts, due to the later appearance of its writing.

In this paper we will consider, from a descriptive standpoint, the shapes of Arabic letters and their development into a graphic system that differentiated analytically between graphemes representing phonemes, in accordance to what I described as the *Phonographic Principle*. This was not achieved immediately, as will be illustrated. Visual shapes of letters are differentiated according to the *Semiographic Principle.* In the 4th/10th century, the *Nasḫī* style of letters was systematically organized by the vizier and great calligrapher Ibn Muqla.

The graphic word-form is a complementary aspect to the shape of letters (Dichy, 1990b; 1997a). Its complex set of morphological structures occurred at a very early stage, because of long time Semitic writing traditions, which isolated word-forms in various ways since the Phoenician writing (dots separating words and special end-of-word shapes of a few letters, which appeared in early Hebraic writing...). Nevertheless, this second aspect cannot be fully developed here for lack of time and space. The figure below hints at the question of the structure of the Arabic word-form. Another aspect will be presented in some detail in relation with the shapes of letters: the word-form structure resulted in a special style of final letters shapes, which was systematized in Arabic.

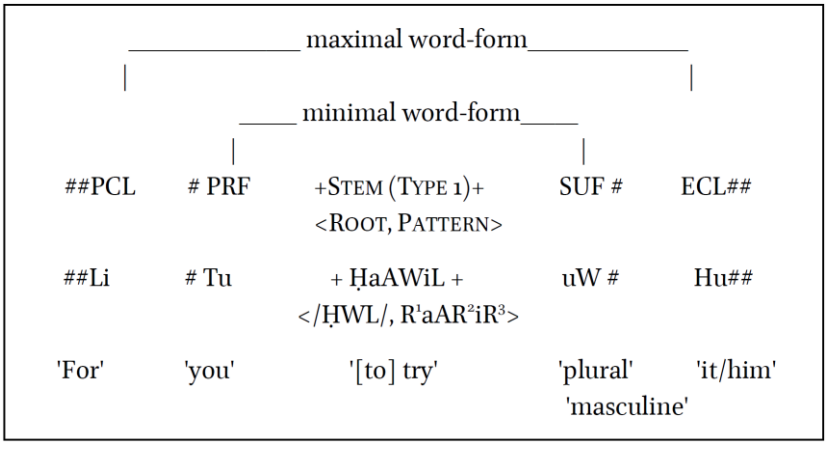

FIGURE 1. Recall of the structure of the word-form (from Dichy 1997a)

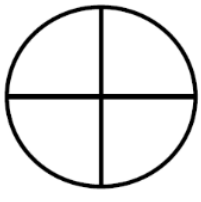

FIGURE 2. Ibn Muqla's virtual circle

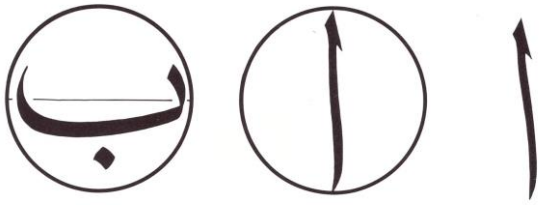

FIGURE 3. An intuitively drawn alif determines the height of the vertical diameter, which in turn determines the horizontal dimensions of a final bāʾ (from Massoudy 1981, p. 38)

2. Ibn Muqla's Epistle and the Coding of Writing on the Basis of the Nasḫī Style

Ibn Muqla is considered by medieval biographical sources as the first calligrapher who put down a codified set of regulations for letter shapes in Arabic. In this code, the dimensions of letters are reckoned after the writing of a letter *ʾalif* included in a circle featuring two perpendicular diameters.

The length of the *ʾalif* is that of the vertical diameter. The letter *bāʾ* covers the horizontal one accordingly, as shown in the figure below.

Later calligraphers have added a convention according to which the dimensions of letters are measured by dots. Ibn Muqla also broke down the movements of the pen (*qalam*, borrowed from the Greek *kalamos*) into five fundamental strokes, which H. Massoudy[4] takes up in the following explanatory drawings:

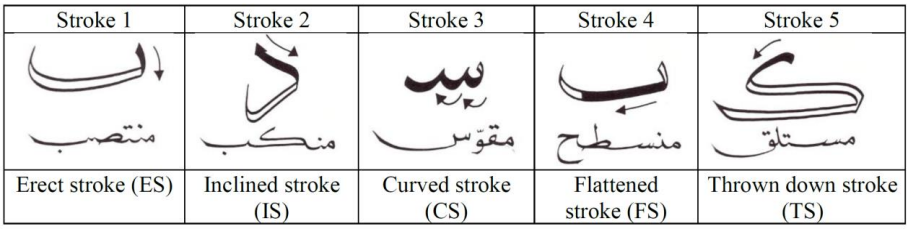

FIGURE 4. Ibn Muqla's five basic strokes (reproduced from Massoudy 1981, p. 39)

These basic movements of the pen are then developed though a description of the drawing of every letter. These can be found with more details than in Ibn Muqla's epistle in the 3rd volume of al-Qalqašandī's encyclopaedic work *Ṣubḥ al-aʿšā fī ṣināʿati l-inšā*, in the chapters on writing, letter shapes and calligraphy (vol. 3: 23 onwards). In the coding of the shapes of letters, the Nasḫī style of writing thus became a reference style, which remains globally true to-day. The very first printed edition of the Qurʾān, due to the Būlāq Printing press in 1923, was realised in Nasḫī characters.

The reference status of this style, and the coding of Arabic calligraphy by Ibn Muqla and the following traditions allows considering Nasḫī letter shapes as the basis of what I have called *semiographic characterisation,* which I will now endeavour to describe with regards to letters.

4. Hassan Massoudy is a well-known iraqi calligrapher living in Paris. See `http://hassan.massoudy.pagesperso-orange.fr/`.

The choice of the Nasḫī writing style as a basis for the study of the semiographic characterisation of the Arabic script could of course be discussed: other candidates could be proposed, such as the Kūfī style, which remained in use for Qurʾan manuscripts for a few centuries even after Ibn Muqla's referential coding (Deroche, 2004); or the Maġribī style, which has been analysed from a visual standpoint by Lakhdar Ghazal, who proposed reading betterments. These of course should be the subject of further studies, which should, in my opinion, benefit by the analyses presented here.

3. The Constrained System of Letter Shapes in Arabic and Word-Form Visual Identification

The letter shapes of Arabic are constrained by a cursive line, which means that *the number of drawings is limited*, and cannot exceed a few figures. The same is true of Syriac writings from which Arabic script, as we know, has been borrowed. Syriac had 22 letters, whereas Arabic includes 29 letters or 28, depending on alphabetic conventions[5]. The 'ancestor' of Arabic script therefore required a number of adaptations. This situation provides an explanation for two general features of Arabic writing (Dichy, 1990b):

- I have described diacritical dots, the width of which is the same as that of the cursive line, as *primary diacritics.* These actually belong to the letter—as opposed to *secondary diacritics* used for short vowels, undetermined case endings and the doubling of a consonant, in addition to diacritic *ḥamza*s. The width of secondary diacritics is much thinner, because they are drawn with the tip of the pen.
- In order to complement primary diacritical dots, the script resorts to *breaking the cursive line* to identify six letters. These breaks used to be called mnemotechnically, in bygone days, by people working in printing presses, "the *dwār* letters": دوار. One needs to add to these four letters ذ and ز, which feature an additional diacritical dot.

5. The alphabet, as it is nowadays taught in schools, includes 28 letters. From medieval times until the 1950s, it was considered as comprehending 29 letters, the first letter *ʾalif* being a *ḥamza*, in accordance with the acrophony principle. In order to represent the long vowel *ā*, a *lam-alif* was traditionally added at the end of the alphabet between *wāw* and *yāʾ*. The double letter *lam-alif* was used because the long vowel *ā* could not be started with. (These point were outlined in the 4th/10th century by Ibn Jinnī, *Sirr Ṣināʿat al-ʾiʿrāb,* vol. I: 41). The omission of the *lam-alif* in the current teaching of the alphabet in schools does not only reduce the number of letters from 29 to 28, it also affects the previous series of *muʿtall* letters (i.e., "weak" or subject to transformation phonemes) positioned at the end of the alphabet. We sincerely hope one of the Ministers of Education will read Ibn Jinnī and restore the age-old traditional Arabic alphabet.

The counterpart of letters in most writings is the word-form. I have recalled above the fact that Semitic writings—among other scriptural traditions—indicated *word boundaries* by a dot, a stroke or the end of a line. Hebraic writing did in addition, as is still the case in Modern Hebrew, have special end-of-the-word shapes for five letters. It is also the case of eight letters in Syriac. This phenomenon has become a mainstream feature in Arabic, where the style of letter has come to oppose two basic letter shapes:

initial or mid-word shape *vs.* end-of-the-word shape.

Nevertheless, end-of-word letter shapes cannot be considered as a strictly enforced principle, because it encounters a few practical impediments, most of which are related to the breaking of the cursive line by some letters. Let us now consider the basic forms of letters.

4. Basic and Initial Letter Shapes

End-of-word forms are analysed in §4.3. Under the constraint of cursive writing, initial letter shapes are limited to two basic models: *the stroke* and *the rounded form*. The first group of letters can be described on the whole with the five strokes identified by the vizier Ibn Muqla. The second one is introduced in this presentation:

- The first group develops the Erect stroke (ES), in combination, when needed, with
  - the Leaning stroke (LS),
  - the Curved stroke (CS),
  - the Thrown down stroke (TS),
  - a combination of two or more of these strokes,
  - the breaking of the cursive line.
- The second group introduces round shapes (RS), in combination, in one case with a breaking of the cursive line.

These two groups allow the building of a taxonomy, which covers all the letters of the alphabet, every item being strictly included in a single class. This can be seen in the tables below. Final shapes of letters are also divided into two parts (§4.3).

In the figures 5 to 7 below, letters are presented between two small lines, showing whether it goes below the cursive line, or above.

5. End-of-Word Shapes

Letters are presented in Fig. 7 both attached to the cursive line and in 'isolated' form, the latter being liable to be drawn or printed slightly

	Single stroke letters			Two-strokes letters		Three-strokes letters		
	Simple Erect stroke	Triple Curved stroke	Vertically prolonged stroke	Inclined + Flatte-ned strokes	Inverted inclined + Flattened strokes	Inclined + Thrown down + Flattened strokes	Inclined (horizontal) + Flattened + Simple Erect strokes	Inclined (horizontal) + Flattened + Vertical prolonged strokes
Letters in initial and middle-of-word form	ـبـ ـتـ ـثـ ـنـ ـيـ	ـسـ ـشـ	ـلـ	ـجـ ـحـ ـخـ	ـعـ ـغـ	ـكـ	ـصـ ـضـ	ـطـ ـظـ
Letters with cursive line breaking	ـرـ ـزـ		ـاـ	ـدـ ـذـ				
Letters with a different mid-word shape					ـغـ			

FIGURE 5. The basic model is that of the different types of strokes. It covers 23 letters

	Single rounded form letters		Curved inclined stroke + rounded form letters
	Simple rounded form	Simple rounded form, inverted drawing	A curved inclined stroke is continued into a rounded form
Letters in initial and middle-of-word form	ـفـ ـقـ		ـهـ
Letters with cursive line breaking	ـوـ	ـمـ	
Letters with a different mid-word shape			ـھـ

FIGURE 6. The rounded form model only concerns 5 letters

higher (ex.: ـنـ ـ ن) and/or in a somewhat different way than the former; examples of small differences : ـكـ ـ ك / ـيـ ـ ي; bigger differences can be found in the forms of initial, mid-word and end-of-word shapes of the letters ع and ه.

The above shapes visually present readers with the left-hand word boundary, i.e., with the end of the word-form.

This is not the case with the دوار group of letters (د، ذ، ر، ز، و، ا), which break the cursive line, for obvious reasons (the final and the mid-word or initial form of the letter remains the same). The existence of these letters is partly responsible for the fact that letter-shapes indicate word boundaries most of the time, but not always.

		Final shapes in contrast (in Nasḫī writing)
The ب model	ـب ـ ب / ـت ـ ت / ـث ـ ث	ـف ـ ف /
The ن model	ـن ـ ن / ـل ـ ل ـس ـ س / ـش ـ ش / ـص ـ ص / ـض ـ ض	ـق ـ ق /
The ي model	ـي ـ ي / ـى ـ ى	
The ك model	ـك ـ ك	
The ع model	ـع ـ ع / ـغ ـ غ / ـج ـ ج / ـح ـ ح / ـخ ـ خ	
The ط model	ـظ ـ ظ	
The م model	ـم ـ م	
Th ه model	ـه ـ ه / ـة ـ ة	

FIGURE 7. Final shape of letters in letters featuring no cursive line breaking

6. Conclusion

The *Semiographic Principle* illustrated above is related to a theoretical approach which considers writing systems as *analytic* (Dichy 2017; 2019). This means that the writing system can be considered as a collective cognitive analysis of the oral language. We have endeavoured here to illustrate the way in which the result of this analysis is projected on paper semiographically, i.e., in a systematic way in which letter shapes are clearly contrasted. The above elaboration owes a lot to the vizier Ibn Muqla's systematic encoding. It emphasises in addition the role of final letter shapes in the identification of word-forms and in the reading process.

References

Abbott, Nabia (1939a). "The Contribution of Ibn Muklah to the North Arabic Script". In: *American Journal of Semitic Languages and Literatures* 56, pp. 70–83.

———— (1939b). *The Rise of the North Arabic Script and Its Kur'anic Development, with a Full Description of the Qur'ān Manuscripts in the Oriental Institute*. Chicago: The University of Chicago Press.

Atanasiu, Vlad (1999). *De la fréquence des lettres et de son influence en calligraphie arabe*. Paris, Montréal: l'Harmattan.

Auroux, Sylvain (1998). *Le langage, la raison et les normes*. Paris: Presses Universitaires de France.

———— (2010). "Sur le mythe de la langue". In: *Pour la (socio)linguistique. Pour Louis-Jean Calvet*. Ed. by Médéric Gasquet-Cyrus et al. Paris: L'Harmattan.

Baʿalbakī, Ramzī [بعلبكي، رمزي] (1981). *[الكتابة العربية والسامية Arabic and Semitic Writing]*. Beirut: [دار العلم للملايين Dār al-ʿilm li-l-malāyīn].

Condillac, E. Bonnot de (1746). "Essai sur l'origine des connaissances humaines". In: *Œuvres philosophiques de Condillac*. Ed. by G. Le Roy. Paris: Presses Universitaires de France.

———— (1775). "Grammaire". In: *Œuvres philosophiques de Condillac*. Ed. by G. Le Roy. Paris: Presses Universitaires de France.

Deroche, François (2004). *Le livre manuscrit arabe. Préludes à une histoire*. Paris: Bnf.

Dichy, Joseph (1990a). "Grammatologie de l'arabe I. Les sens du mot ḥarf, ou le labyrinthe d'une évidence". In: *Studies in the History of Arabic Grammar II*. Ed. by K. Versteegh and M.G. Carter. Amsterdam, Philadelphia: Benjamins, pp. 111–128.

———— (1990b). "L'écriture dans la représentation de la langue: la lettre et le mot en arabe". Doctorat d'État. Université Lumière Lyon 2.

———— (1997a). "Deux grands "mythes scientifiques" relatifs au système d'écriture de l'arabe". In: *L'Arabisant* 32–33, pp. 67–86.

———— (1997b). "Pour une lexicomatique de l'arabe: l'unité lexicale simple et l'inventaire fini des spécificateurs du domaine du mot". In: *Meta* 42, pp. 291–306.

———— (2014). "Al-Ḫalīl's Conjecture: How the First Comprehensive Dictionary in History Was Invented". In: *Arab and Arabic Linguistics: Traditional and New Theoretical Approaches*. Ed. by Manuella Giolfo. Oxford: Oxford University Press, pp. 39–64.

———— (2017). "The Analytics of Writing, Exemplified by Arabic, the Youngest of the Semitic Scripts". In: *Approaches to the History and Dialectology of Arabic, in Honour of Pierre Larcher*. Ed. by M. Sartori, M.E. Giolofo, and Ph. Cassuto. Leiden, Boston: Brill, pp. 29–56.

———— (2019). *Études sur le système d'écriture et la linguistique de l'écrit en arabe*. Villeurbanne (Lyon): Aradic-Monde arabe.

Driver, Godfrey R. (1976). *Semitic Writing: From Pictograph to Alphabet*. 3rd ed. Oxford: Oxford University Press.

Grainger, Jonathan, Joseph Dichy, et al. (2003). "Approche expérimentale de la reconnaissance du mot écrit en arabe". In: *Faits de langue* 22, pp. 77–86.

Healey, John F. and G. Rex Smith (2009). *A Brief Introduction to the Arabic Alphabet: Its Origins and Various Forms*. London, San Francisco, Beirut: Saqi.

Ibn al-Ḥāǧib [ابن الحاجب] (1975). "الشافية في علمي التصريف والخطّ [Commentary on the Satisfying [Book] on the Two Sciences of Morphology and Writ-

ing]". In: الأستراباذي، شرح الشافية *[Commentary on the Shāfiya]*. Ed. by Nūr al-Ḥasan [الحسن, نور]. Beirut: دار الكتب العلمية [Dār al-kutub al-ʿilmiyya].

Jaffré, Jean-Pierre (1988). "Graphèmes et idéographie. Approche psycholinguistique de la notion de graphème". In: *Pour une théorie de la langue écrite*. Ed. by Nina Catach. Paris: Éditions du CNRS, pp. 93–102.

Joly, André (1982). "De la théorie du langage à l'analyse d'une langue. Remarques autour de la Grammaire de Condillac". In: *Condillac et les problèmes du langage*. Ed. by J. Sgard. Genève, Paris: Slatkine, pp. 243–256.

Mansour, Kamal (in this volume). "On the Origin of Arabic Script".

Massoudy, Hassan (1981). *Calligraphie arabe vivante*. Paris: Flammarion.

Osborn, J.R. (2017). *Letters of Light. Arabic Script in Calligraphy, Print, and Digital Design*. Cambridge, MA, London: Harvard University Press.

al-Qalqašandī, Abū l-ʿAbbās ʾAḥmad ibn ʿAlī [القلقشندي أبو العبّاس أحمد بن علي بن أحمد] (1985). صبح الأعشى في صناعة الإنشاء *[The Morning of the Dim-Sighted [Treatise] on the Art of Composition]*. Cairo: الهية العامة للكتّاب [al-Hayʾa l-ʿāmma li-l-Kitāb].

Robin, Christian (1991). "Les Écritures de l'Arabie avant l'Islam". In: *Revue du Monde Musulman et de la Méditerranée* 61, pp. 127–137.

Schoeler, Gregor (2002). *Écrire et transmettre dans les débuts de l'Islam*. Paris: Presses Universitaires de France.

Vygotsky, Lev Semenovich [Выготский, Лев Семёнович] (1934). *Мышление и речь [Thought and Language]*. Moscow: Государственное Социально-Экономическое Издательство [Gosudarstvennoe Socialno-Ekonomicheskoye Izdatelstvo].

———— (1935). "Проблема обучения и умственного развития в школьном возрасте [The Problem of Teaching and Mental Development at School Age]". In: *Психологическая наука и образование [Psychological Science and Education]*, pp. 3–19.