ܣܰܒܪܳܐ (Sabra), a Syriac TEX system

Yannis Haralambous*
187, rue Nationale,
59800 Lille, France,
Fax: +33 20.40.28.64,
haralambous@univ-lille1.fr

Abstract

In this paper we present "Sabra", a typesetting system for the Syriac script, based on TEX and METAFONT. We cover Serto, Estrangello, as well as East Syriac, both for the Syriac language as for Garshuni. Ligatures as well as stretching connections (keshideh) are automatically performed. Straight or curved isolated Serto *olaph* are chosen according to a set of rules described in the paper. This work is part of the ScholarTEX project.

ܒܡܐܡܪ ܗܢܐ ܡܬܚܘܝܐ ܬܘܪܨܬܐ ܕ «ܣܒܪܐ Sabra» ܕܐܬܘܬܗ ܬܘܪܨܬܐ ܡܛܠ ܬܘܣܦ ܐܬܘܬܐ ܠܟܬܒܐ ܣܘܪܝܝܐ ܠܦܘܬ ܬܘܪܨܬܐ ܕ «TEX» ܘܡܗ; ܕ «METAFONT». ܬܘܪܨܬܐ ܕܟܠ ܦܠܚܐ ܒܟܠ ܙܢ̈ ܐܬܘ̈ܬܐ ܣܘܪ̈ܝܝܬܐ. ܐܣܛܪܢܓܠܐ: ܐܡܪ ܐܢܐ: ܘܣܦܩܐ ܘܐܬܘ̈ܬܐ ܡܢܬܫ̈ܬܐ ܘܗܠܝܢ ܕܐܠܝ̈ܢ ܡܛܠ ܚܙܝܘܢ.

Contents

1 Introduction: Difficulties of Syriac Typesetting

1.1 Scripts, and their Names

Like Greek and Hebrew, Syriac is a living language and a language of scholarly interest. But unlike Greek and Hebrew, Syriac has not been written in a unique way. Moreover, political or religious communities have used script variants as a distinguishing feature. The result is that —at least

*The author would like to thank George Kiraz (Cambridge), Johannes de Moor (Kampen), Murielle Debré (Paris) and f. Hazaïl Soumi (Louvain-La-Neuve) for their friendly support and guidance.

in older scholar texts[1]— the different Syriac scripts are often called by names of Christian sects, which have used them in the past:

- *Nestorian*, which we call “East Syriac”;
- *Jacobite* (Syrian Monophysites), which we call “Serto” (ܣܪܛܐ);
- *Melchitic*, a variant of Estrangello, not covered by this paper.

These names will not be used in this paper (native Syriac language speakers find them offensive).

Other scripts, with less problematic names, are “Estrangello” (ܐܣܛܪܢܓܠܐ), *Peshito* (a variant of Serto) and *Mandean* (a very quadratic variant of Estrangello, close to Arabic Kufic) the later two not covered by this paper.

1.2 Technicalities

Syriac presents the same typesetting problems as Arabic[2]: right-to-left writing direction, contextual analysis, multitude of vowels and diacritics. It should be noted that Syriac needs almost the double amount of vowels/diacritics than Arabic: in fact, Syriac uses many vowel systems, sometimes mixed, and Garshuni Syriac (Arabic written in Syriac script) uses all Arabic vowels and diacritics.

1.3 Solutions: TeX, Ω

The ܣܒܪܐ (*Sabra* = “hope”, in Syriac) system is build upon TeX and hence heritates all of TeX’s power and flexibility. Nevertheless, the many problems involved in Syriac typesetting make the use of supplementary tools unavoidable: a pre-processor is used for ‘input encoding → font encoding’ translation, and contextual analysis; a multiple font loading mechanism is used to access all the signs, and TeX--XeT is necessary for switching writing directions.

This rather cumbersome way of adapting TeX to Syriac will be obsolete once the Ω extension of TeX[3] is fully operational. Ω allows the use of 16-bit fonts (up to 65,536 characters) which is more than enough for Syriac; the contextual analysis and remapping of codes is handled by internal Lex-like filters and TeX--XeT is integrated by default.

With respect to the forthcoming availability of this new TeX system, we will not enter too much into TeX- and METAFONT-specific techniques which become obsolete under Ω.

1.4 The Serto isolated *olaph*

The Serto letter *olaph*, when in isolated contextual form, can take two different shapes: “straight” ܐ and “curved” ܐ. Here are the rules for choosing between straight and curved form:

1. At word begin, the **curved** form is used;
2. Inside the word, or at word end, the **straight** form is used, except in the following cases:
3. When the letter *olaph* is preceeded by
   (a) a *dolath* ܕ, or
   (b) a *waw* ܘ, or
   (c) a *waw* followed by a *dolath* ܘܕ, or

[1] And even in more recent texts, like [8] (1980), still using only the terms ‘Jacobite’ and ‘Nestorian’.

[2] See [6] for a detailed discussion on the problems of Arabic typesetting.

[3] See [7] for a detailed discussion of Ω’s features, and follow the Ω project pointer in the `http://www.ens.fr` homepage to find out about the latest status of this very ambitious project.

(d) a double *dolath*[4] ܕܕ.

Here are some exceptions to these rules, found in [3]:

- the word ܐܐܪ (ἀὴρ = air), and it's plural form ܐܐܪ̈ܐܣ, contradict rule 2: the curved form of *olaph* is used inside the word;
- the words ܕܐܒܐ (wolf), ܕܐܪܐ (combat) contradict rule 3-a: although *olaph* is preceeded by a *dolath* it is typeset in straight form;
- the word ܘܐܘ (letter *waw*) contradicts rule 3-b: although *olaph* is preceeded by a *waw* it is typeset in straight form.

It seems that all Syriac authors/typesetters do not follow the same rules: [2] uses only the curved form of *olaph*; the text given as example in [9, p. 181] uses a curved *olaph* at word begin and a straight one otherwise, EVEN when preceeded by a *waw* or a *dolath*. In another document [5], also printed at the Imprimerie Nationale in Paris, one finds only straight *olaphs*, even at word begin...

According to f. H. Soumi[5], the rules are by far more complicated than those given above, and a good knowledge of Syriac language is necessary to determine the form of *olaph* to be used. The author would be grateful for any supplementary information.

The input methods for the different forms of *olaph* are described in section 2.4, p. 5.

1.5 The Serto initial *olaph-lomad* ligature

A similar phenomenon as for the isolated *olaph* letter, occurs for the initial *olaph-lomad* ligature: the use of this ligature obeys to the same rules as the typesetting of curved *olaph* letter: this ligature is used:

1. At word begin, except in the following cases:
2. When the letters *olaph-lomad* are preceeded by
   (a) a *dolath* ܕ, or
   (b) a *waw* ܘ, or
   (c) a *waw-dolath* pair ܘܕ, or
   (d) a double *dolath*[6] ܕܕ.
3. When the letters *olaph-lomad* are followed by *olaph*: in that case the first *olaph* is isolated curved and the letters *lomad-olaph* form a ligature. Example: ܐܠܐ (to bewail).

This rule seems to be even less followed as the one for the form of letter *olaph*; here are some exceptions, found in [3]: ܐܠܦ (letter *olaph*), ܐܠܦ (thousand) with the variant form ܐܠܦ.

The input methods for the different forms of *olaph* are described in section 2.4, p. 5.

[4] In a private communication, G. Kiraz mentions only the three first rules. The fourth one has been detected by the author in [4, p. 273].

[5] Personal communication.

[6] In the same private communication, G. Kiraz mentions only the three first rules. The fourth one has been introduced in analogy with the case of the isolated *olaph* letter.

2 Using Sabra

2.1 Requirements

To typeset in Syriac using ܣܒܪܐ, one needs a decent TeX system (in this context, this would be a TeX implementation featuring Peter Breitenlohner's TeX--XeT and an operational METAFONT implementation), a relatively powerful machine (being able to run BigTeX) and the ܣܒܪܐ (Sabra) package[7], consisting of a preprocessor (`sabra`), fonts written in METAFONT, and TeX macros. The preprocessor is written entirely in GNU Flex, without using any system-dependent subroutines; hence it can be compiled in a straightforward manner on any platform having a GNU Flex executable and an ANSI C compiler.

Once ܣܒܪܐ installed, typesetting is done in two steps: an input file is prepared using the syntax we will describe below; the preprocessor then reads this file, and produces a LaTeX 2_ε (or plain TeX) file which then can be run through TeX, in the usual way. Once Ω will be operational, there will be no pre-processor needed anymore.

2.2 Preparing the input file

If you wish to write your file in LaTeX 2_ε, you have to include the line

```
\usepackage{sabra}
```

in the preamble.

A ܣܒܪܐ input file contains text, TeX/LaTeX macros, and *preprocessor directives*. The latter concern only Syriac script. To type Syriac text you need to enter *Syriac mode*; this is done by the preprocessor directives `<E>` (Estrangello), `<S>` (Serto) or `<C>` (East Syriac). To leave one of these modes, one enters `</E>` or `</S>` or `</C>`.[8]

Once you are inside Syriac mode, you type Syriac text in Latin transcription, from left-to-right. No special indication needs to be given to TeX about font or writing direction switching, this is done automatically by TeX.[9] Section 2.4 describe the transcription you have to use as well as all other features of the preprocessor.[10]

2.3 A text in a text in a text: nesting modes

Often one needs Latin, Greek, or other left→right insertions inside Syriac text; these insertions may again contain Syriac insertions, and so forth. If it weren't for the writing direction, it would be equivalent of switching from one script to the other. Unfortunately, the situation is more complicated: every writing direction is an environment, and paragraph breaking into lines depends on the rules of this environment. The only way to deal properly with such environment is to *nest* them: if you want to switch from Syriac to Latin, you have to decide:

1. if you are finished with Syriac and you are "returning" to Latin, in which case you leave Syriac mode by typing `</E>` (Estrangello), `</S>` (Serto) or `</C>` (East Syriac);

2. if you are going to type a Latin insertion: in that case you enter (nested) Latin mode by typing `<L>`. To leave this mode, use `</L>`; this will bring you back to the script you were using before (Serto, Estrangello or East Syriac).

---

[7] ܣܒܪܐ it is part of the long awaited version 1 of the ScholarTeX package, together with *new* Greek, Arabic, Hebrew, Coptic and Akkadian cuneiform systems.

[8] A few other similar directives concern other languages of scholarly interest: `<A>` for Arabic Naskhi, `<H>` for Biblical Hebrew, `<K>` for Akkadian cuneiform. Other writing systems provided by ScholarTeX do not require a preprocessor.

[9] Nevertheless, when your text takes more than one paragraphs, you have to write the commands `\hfill\phantom{.}` at the end of the first paragraph, and `\beginR` at the begin of the next one; this is an unfortunate shortcoming of TeX--XeT, which will be fixed in forthcoming versions of Ω.

[10] Nevertheless, the document preparation syntax will remain the same so that documents prepared for the current version of ܣܒܪܐ will still be processable by Ω and produce exactly the same results.

Inside nested Latin mode the preprocessor is totally inactive: nothing is changed, up to the next `</L>`, `<S>`, `<E>`, or `<C>`. Similarly, when you are in (nested) Latin mode and enter into a (doubly nested) Syriac mode, everything is as in usual Syriac mode, with the only difference that when you leave it you are back in (nested) Latin mode, and so on…[11]

The figure below illustrates the nesting property of nesting modes:[12]

```
                          ...
                  <E>     ...    </E>
          <L>     ...     ...     ...    </L>
  <S>     ...     ...     ...     ...     ...    </S>
```

2.4 Transcription of Syriac consonants and long vowels

The Syriac transcription of letters (consonants and long vowels) is given in tables 1 (Serto), 2 (Estrangello) and 3 (East Syriac), pages 9–10 of the appendix.

Here is a simple example of code producing non-vowelized Syriac text:
`<S>tuub qTygwRys d'rysTTlys pylsuup'</S>` will produce

— ܛܘܒ ܩܛܝܓܘܪܝܣ ܕܐܪܝܣܛܛܠܝܣ ܦܝܠܣܘܦܐ, in Serto;

— ܛܘܒ ܩܛܝܓܘܪܝܣ ܕܐܪܝܣܛܛܠܝܣ ܦܝܠܣܘܦܐ, in Estrangello (`<E>` and `</E>` instead of `<S>` and `</S>`);

— ܛܘܒ ܩܛܝܓܘܪܝܣ ܕܐܪܝܣܛܛܠܝܣ ܦܝܠܣܘܦܐ, in East Syriac (`<C>` and `</C>` instead of `<S>` and `</S>`).

Some notes concerning the transcription of letters of tables 1–3:

There is no distinction between isolated, initial, medial and final forms; these are automatically applied by TEX.

Some letters have multiple transcriptions: `'` or `aa` for *olaph*, `w` or `uu` for *waw*, `h` or `ee` for *he*, `y` or `ii` for *yud*. These produce exactly the same results: choosing between them can make transcribed input more readable, especially when the text is not vowelized (for example, when used as *mater lectionis*, long vowels are more readable when transcribed as vowels, etc.).

In section 1.4, p. 2 we mentioned the two forms of Serto isolated *olaph*. If the rules described in that section are to be followed, then no extra precaution need to be taken: the transcription `'` will produce a straight or a curved *olaph* depending on the context. To disobey the rules, two extra transcriptions are given: `'|` will produce a straight isolated *olaph*, and `'@` a curved one, under `ALL` circumstances. So, for example, the exceptions seen in 1.4, p. 2 ܐܳܐܰܪ and ܕܺܐܒܳܠ are input as `<S>'@A'@ar</S>` and `<S>di'|bAl</S>` respectively.

In section 1.5, p. 3 we gaved rules about the typesetting of the *olaph-lomad* ligature. Once again, these rules are automatically applied by ܣܕܪܐ. To transgress them, one can use the same transcriptions as in the previous paragraph: `'@` for a curved *olaph* (the logical choice, since we are necessarily at word begin) or `'|` for a straight one. Here are some example: `<S>'alAhA'</S>` to obtain ܐܰܠܳܗܳܐ (god), `<S>'elA'</S>` for ܐܶܠܳܐ (to bewail) but `<S>'@Alep</S>` for ܐܳܠܶܦ (thousand).

Transcription `s|` stands for the two-form *semkath*. This letter will be typeset in isolated or final forms only, while usual `s` will have all four forms (isolated, initial, medial and final). The transcription `s+` is used for the (two-form) letter *semkath* with tail.

2.5 Transcription of Syriac vowels and diacritics

There are many systems of Syriac vowelization, and often more than one of them are used simultaneously. ܣܕܪܐ provides four types of vowels:

[11] The number of times you can enter a mode inside a mode, is given by the preprocessor internal variable `MAX_STATE_STACK_SIZE` which by default is set to 16. Besides switching from Syriac to Latin and back, you can also switch to Garshuni Syriac (see 2.6, p. 6) or from one Syriac script to another.

[12] For people with LATEX experience, it may be enough to say that every ܣܕܪܐ mode corresponds to a LATEX environment.

1. Greek vowels (see table 4, p. 11);
2. Vowel points (see table 5, p. 11);
3. East Syriac vowels (see table 6, p. 11);
4. Jacob of Edessa vowels (these are not diacritics but real letters) (see table 7, p. 12).

These vowels can be used in the same context: a word can very well contain for example Greek and East Syriac vowels: ܫܡܥܘ ܠܡ̈ܠܐ ܕܡܠܟܐ، ܕܡܠܟܗܘܢ (listen the words of their king).

In two cases (inverted Greek vowels and Jacob of Edessa vowels), there are alternate input methods. If one has to typeset texts which have more inverted Greek vowels than regular ones, or more Jacob of Edessa vowels than Greek ones, one can enter into “inverted vowel” (or “Jacob of Edessa vowel”) mode. This is done by typing `<INVVOW>` (resp. `<JACVOW>`). Once entered in one of these modes, the transcription of vowels changes, as indicated in the table. This behaviour of the pre-processor continues until the user exits the mode (by typing `</INVVOW>` or `</JACVOW>` respectively).

2.6 Transcription of Garshuni letters and diacritics

Garshuni is Arabic written in the Syriac script. Since the Arabic language has a few more sounds than Syriac, new letters had to be introduced: these are pre-existing Syriac letters with additional dots or strokes[13]. The Garshuni transcription used in ܗܒܝܒܐ is exactly the same as in the الأمل (*Al-amal* = “hope”, in Arabic) system, the Arabic part of ScholarTeX system.

This means that the transcription used to typeset Arabic *language* in Syriac script, is DIFFERENT from that used to typeset Syriac *language*, although the letters themselves are mostly the same. The transcription used for Garshuni is described in tables 11, p. 14 (Serto), 12, p. 15 (Estrangello) and 13, p. 16 (East Syriac), for the consonants and long vowels and in table 10, p. 13 for the Arabic vowels and diacritics.

This transcription is only valid in “Garshuni mode”: to enter Garshuni mode, one has to do the following steps: (a) entering into Syriac mode (by typing `<S>` for Serto, `<E>` for Estrangello or `<C>` for East Syriac), and (b) enter into Garshuni mode, by typing `<G>`. So, for example, the famous Arabic “welcome”

أَهلاً وَ سَهلاً

is transcribed `<A>'aahlAaN wa sahlAaN</A>` in الأمل; to typeset it in Syriac Serto script we just have to write `<S><G>'aahlAaN wa sahlAaN</G></S>`. Here is the result:

ܐܰܗܠܐً ܘܰ ܣܰܗܠܐً

The reader must realize that these “modes” have to be *nested*: if one enters Syriac mode and then Garshuni mode, then one must *first* leave the later and *then* the former.[14]

[13]There seem to be many ways of adapting Syriac script to the Arabic language; the author has chosen to follow the one suggested by G. Kiraz. Have there been Persian or Turkish texts written in Syriac script? The author would be grateful on any information related to this question.

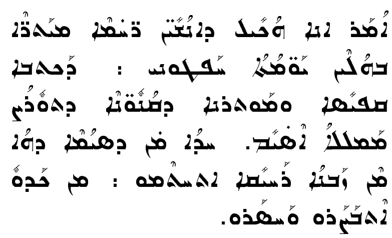
[14]This is not the case for the “inverted vowel” mode and the “Jacob of Edessa vowel” mode: these can be entered and exited at any time, without nesting with Syriac mode. To be more precise we should call these “pseudo-modes”, or “global flags”.

2.7 Stretching letter connections (کشیده)

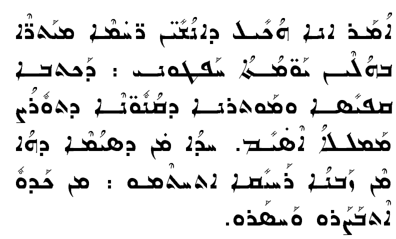
As in Arabic, Syriac words are never hyphenated. To balance white space on each line, one stretches connections between initial and medial, or medial and final forms of letters. This stretching is called کشیده "keshideh" (a Persian word, derived from the verb کشیدن, "to stretch", see [1, p. 171]).

In ܣܪܛܐ, stretching of letter connections is not activated by default. To activate it, one uses the preprocessor directive `<KESHIDEH>` (and `</KESHIDEH>` to desactivate it). This is not a "mode" in the sense of `<G>` (Garshuni mode) or `<L>` (Latin mode) and hence doesn't need to be nested with respect to Syriac mode.

Here is an example of the use of letter connection stretching:



Without letter connection stretching



With letter connection stretching

3 Notes on the design of the fonts

The Serto, Estrangello, and East Syriac fonts have been designed in the METAFONT language. Characters are described by mathematical equations (1st and 3rd degree polynomials, known as Bézier curves), using *parameters*. These parameters define the weight, scaling, optical correction and style of each character and of the entire font. The author has tried to follow traditional Syriac types as much as possible: inspirative sources have been kindly provided by G. Kiraz.

Syriac scripts carry so much beauty and offer so many challenges to the designer that giving even a half-complete description of the designing process of these fonts would be too long for this paper. The author only retained a specific phenomenon of the Serto script, which he believes will be of interest to type designers even unfamiliar with the Syriac language and script:

3.1 Anti-symmetry of Serto letters *lomad* and *'e*

One of the most interesting challenges in type design is the art of making symmetric letters look different. It is a known fact that older types (such as the Elzevirs or Venitian ones) are more readable than modern sans-serif ones; this is in part a consequence of small distinguishing features between letters which otherwise "would look too much alike". In the Helvetica type, letters 'b', 'd', 'p' and 'q' are just reflections and rotations one of each other; in Garamond, their shapes are significantly different, so that a reflected 'p' is not identical to a 'q', and so forth.

In Serto, a similar problem arises with letters *lomad* and *'e*: as can be seen in the following table, in principle *'e* can be seen as a *shortened lomad*:

Form	Isolated	Initial	Medial	Final
lomad	ܠ	ܠ	ܠ	ܠ
'e	ܥ	ܥ	ܥ	ܥ

How did Serto type designers make it easier for the eye to distinguish between these letters? By giving them **different angles**. We are talking about the angle of the main stem of these letters (or of the two parallel stems, in the case of the isolated and final forms). The following table gives the angle of these stems, when tracing them downwards and from left to right:

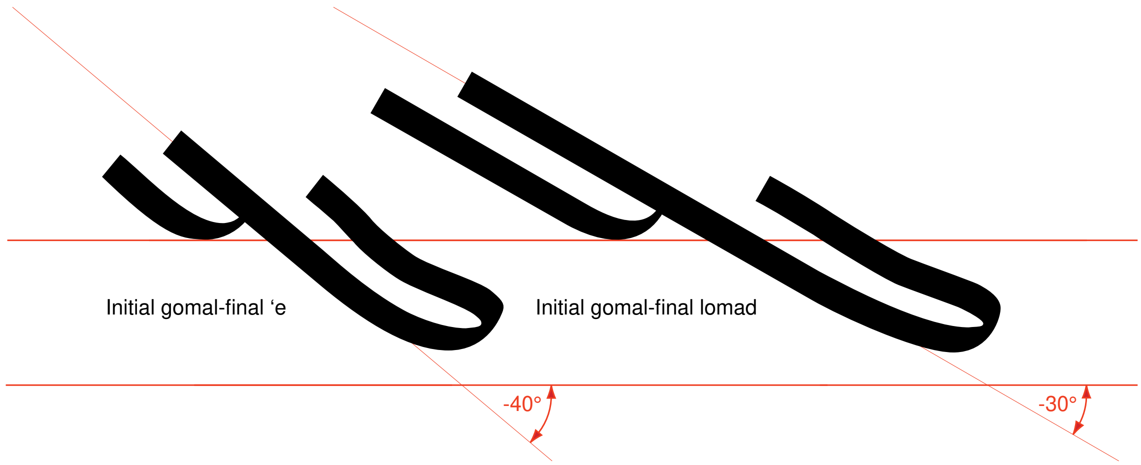


Figure 1: The Serto isolated *gomal-lomad* and *gomal-‘e* ligatures

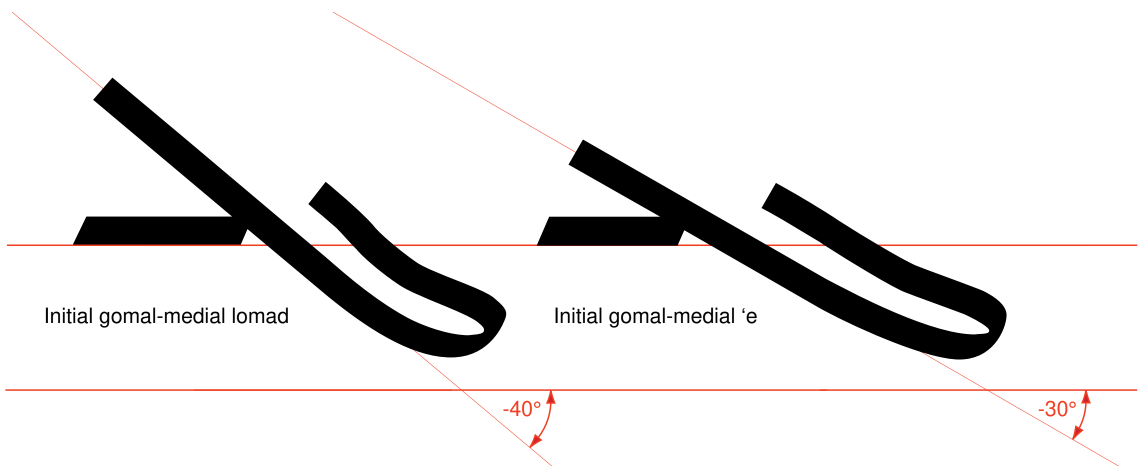


Figure 2: The Serto initial *gomal-lomad* and *gomal-‘e* ligatures

Form	Isolated	Initial	Medial	Final
lomad	-30°	-40°	-40°	-30°
‘e	-40°	-30°	-30°	-40°

This is what we call “anti-symmetry” of these two letters: whenever one of them has an angle of -30°, the other one (in the same contextual form) has an angle of -40°. Here is what it gives when these letters are close to each other:

ܠܥܠܥ ܥܠܥܠ

The story doesn’t ends here. One of the most beautiful Serto letter combinations is the *gomal-lomad* one, because that long straight *lomad* stroke is drawn as a continuation of the equally long *gomal* stroke: ܓܠ (to bubble). This means that the *gomal* stroke must have exactly the same angle as the *lomad* one, and that these two letters should be specially drawn to fit together.

What happens now when *gomal* is followed by an *‘e* rather than by a *lomad*? Once again the *gomal* must have the same angle as this letter.

This means that we need two *gomals*: one for the (final) *lomad*, with an angle of 30°, and one for the (final) *‘e*, with an angle of 40°. The amusing part is that the former will fit also exactly to the (medial) *‘e* and the later to the (medial) *lomad*. These ligatures are illustrated on figures 1 (*gomal* followed by final letters) and 2 (*gomal* followed by medial letters).

Of course ܣܒܕܐ chooses automatically the right form of letter *gomal*.

4 Tables

4.1 Syriac consonants and long vowels

ܐ	ܒ	ܓ	ܕ	ܖ	ܗ	ܘ	ܙ	ܚ
’	b	g	d	D	h, ee	w, uu	z	x
ܛ	ܝ	ܟ	ܠ	ܡ	ܢ	ܣ	ܤ	ܥ
T	y	k	l	m	n	s	s+	‘
ܦ	ܨ	ܩ	ܪ	ܪ̈	ܫ	ܬ	ܐ	ܐ
p	Y	q	r	R	S	t	’|	’@

Table 1: Syriac Serto letters and their input codes

ܐ	ܒ	ܓ	ܕ	ܖ	ܗ	ܘ	ܙ	ܚ
’	b	g	d	D	h, ee	w, uu	z	x
ܛ	ܝ	ܟ	ܠ	ܡ	ܢ	ܣ	ܤ	ܥ
T	y	k	l	m	n	s	s+	‘
ܦ	ܨ	ܩ	ܪ	ܪ̈	ܫ	ܬ	ܐ	ܐ
p	Y	q	r	R	S	t	’|	’@

Table 2: Syriac Estrangello letters and their input codes

ܐ	ܒ	ܓ	ܕ	ܖ	ܗ	ܘ	ܙ	ܚ
’	b	g	d	D	h, ee	w, uu	z	x
ܛ	ܝ	ܟ	ܠ	ܡ	ܢ	ܣ	ܤ	ܥ
T	y	k	l	m	n	s	s+	‘
ܦ	ܨ	ܩ	ܪ	ܪ̈	ܫ	ܬ	ܐ	ܐ
p	Y	q	r	R	S	t	’|	’@

Table 3: East Syriac letters and their input codes

In normal mode:	a	A	e	i	u
In inverted mode:	<INV>a	<INV>A	<INV>e	<INV>i	<INV>u
In normal mode:	<INV>a	<INV>A	<INV>e	<INV>i	<INV>u
In inverted mode:	a	A	e	i	u

Table 4: Greek vowels and their input codes (type `<INVVOW>` to enter “inverted mode”, `</INVVOW>` to exit)

<.>, .a, .o	<_.>, .e, .i

Table 5: Vowel points and their input codes

/a	/A	//A	/E
//E	|e, /e	/o	/u

Table 6: East Syriac vowels and their input codes

In default mode:	<JAC>a	<JAC>E	<JAC>e	<JAC>I
In Jacob of Edessa mode:	a	E	e	I
In default mode:	<JAC>i	<JAC>U	<JAC>u	<JAC>u/
In Jacob of Edessa mode:	i	U	u	u/

Table 7: Jacob of Edessa vowels and their input codes (type <JACVOW> to enter “Jacob of Edessa vowels mode”, </JACVOW> to exit)

<PLU>, <SEYAME>	<_PLU>, <_SEYAME>	_", <_SHWAYA>	", <SHWAYA>
<QUS>, <QUSHOYO>	<RUK>, <RUKAKHA>	<_LIN>	<^LIN>
<MEGA>	<KILO>	<TILDA>	<CUP>
	<BAREKH>	<QSO>	

Table 8: Miscellaneous Syriac diacritics

;	,	+:	\:	/:
:, |:[15]	.	::	<<	>>

Table 9: Syriac punctuation

4.2 Arabic written in Syriac script (Garshuni)

a	"a	i	u	:	*	aN	iN
uN	*a	*"a	*i	*u	*aN	*iN	*uN

Table 10: Garshuni vowels and their input codes (type <G> to enter “Garshuni mode”, </G> to exit)

[15]The transcription |: for : becomes necessary in Garshuni mode, because of a conflict with one of the *sukun* diacritic (see table 10).

ء (ء)	ܐ (ا)	ܐٓ (آ)	ܐٱ (ٱ)	ܐٔ (أ)
`||`	`A`	`’A`	`"A`	`’a`
ܐٕ (إ)	ܘٔ (ؤ)	ܝٔ (ئ)	ܒ (ب)	ܗ̈ (ة)
`’i`	`’u`	`’y`	`b`	`"t, "h`
ܬ (ت)	ܬ݂ (ث)	ܓ (ج)	ܚ (ح)	ܚ݂ (خ)
`t`	`t_`	`g`	`H`	`h_`
ܕ (د)	ܕ݂ (ذ)	ܪ (ر)	ܙ (ز)	ܣ (س)
`d`	`d_`	`r`	`z`	`s`
ܫ (ش)	ܨ (ص)	ܨ̇ (ض)	ܛ (ط)	ܛ̇ (ظ)
`S`	`Y`	`D`	`T`	`Z`
ܥ (ع)	ܓ݂ (غ)	ܦ (ف)	ܩ (ق)	ܟ (ك)
`‘`	`G`	`p, f`	`q`	`k`
ܠ (ل)	ܡ (م)	ܢ (ن)	ܗ (ه)	ܘ (و)
`l`	`m`	`n`	`h`	`w, U`
		ܝ (ي)		
		`I, y`		

Table 11: Garshuni Serto letters and their input codes (type `<G>` to enter “Garshuni mode”, `</G>` to exit)

ء (ء)	ܐ (ا)	ܐٓ (آ)	ܐٱ (ٱ)	ܐٔ (أ)
||	A	'A	"A	'a
ܐٕ (إ)	ܘٔ (ؤ)	ܝٔ (ئ)	ܒ (ب)	ܗ̈ (ة)
'i	'u	'y	b	"t, "h
ܬ (ت)	ܬ݂ (ث)	ܓ (ج)	ܚ (ح)	ܟ݂ (خ)
t	t_	g	H	h_
ܕ (د)	ܕ݂ (ذ)	ܪ (ر)	ܙ (ز)	ܣ (س)
d	d_	r	z	s
ܫ (ش)	ܨ (ص)	ܨ݂ (ض)	ܛ (ط)	ܛ݂ (ظ)
S	Y	D	T	Z
ܥ (ع)	ܓ݂ (غ)	ܦ (ف)	ܩ (ق)	ܟ (ك)
‘	G	p, f	q	k
ܠ (ل)	ܡ (م)	ܢ (ن)	ܗ (ه)	ܘ (و)
l	m	n	h	w, U
		ܝ (ي)		
		I, y		

Table 12: Garshuni Estrangello letters and their input codes (type <G> to enter “Garshuni mode”, </G> to exit)

ء (ء)	ܐ (ا)	ܐ̄ (آ)	ܐ (ٱ)	ܐ (أ)
||	A	’A	"A	’a
ܐ (إ)	ܘ (ؤ)	ܝ (ئ)	ܒ (ب)	ܗ̈ (ة)
’i	’u	’y	b	"t, "h
ܬ (ت)	ܬ̤ (ث)	ܓ (ج)	ܚ (ح)	ܟ̣ (خ)
t	t_	g	H	h_
ܕ (د)	ܕ̤ (ذ)	ܪ (ر)	ܙ (ز)	ܣ (س)
d	d_	r	z	s
ܫ (ش)	ܨ (ص)	ܨ̇ (ض)	ܛ (ط)	ܛ̣ (ظ)
S	Y	D	T	Z
ܥ (ع)	ܓ̣ (غ)	ܦ (ف)	ܩ (ق)	ܟ (ك)
‘	G	p, f	q	k
ܠ (ل)	ܡ (م)	ܢ (ن)	ܗ (ه)	ܘ (و)
l	m	n	h	w, U
		ܝ (ي)		
		I, y		

Table 13: Garshuni East Syriac letters and their input codes (type <G> to enter “Garshuni mode”, </G> to exit)

5 Examples

5.1 Serto script, Syriac language

5.2 Serto script, Arabic language (Garshuni)

ܗܘ ܒܝܬܗ ܐܘ ܡܣܟܢܗ
ܘܐܚܝܪܐ ܐܬܐ ܡܘܡܢ ܐܪܚܣ ܦܝܗ ܐܠܓܠܐܡ ܐܫܘܪ ܐܠܒܠܘܪ ܐܠܡܠܘܢ،
ܦܪܐܣܗ ܐܘ ܐܢ ܐܓܕ ܡܢ ܡܠܐܠ، ܐܘ ܗܡܠ ܦܝ ܓܠܢ ܥܪܝܢ ܡܢ ܓܢܢ
ܥܡܪ. ܢܚܒ ܐܢ ܐܘܗܒ ܦܝ ܐܠܣܠܐܡ ܠܐܚܣܡܘܡܠ.
ܓܢܒܗ ܐܠܐܡܢܙܗ ܐܠܡܚܒܒܗ ܒܡܒܗ ܒܡܒܗܘܐ، ܘܡܠܘܐ ܦܝ
ܐܠܛܢܝܗ ܐܠܒܝܡܠܘ ܦܝܗ ܐܓܕܐܗܘܐ ܡܢ ܡܚܠܠ، ܢܠܐ ܘܢܟܢ ܐܠܐ
ܐܠܡܘܒ ܐܠܒܝ ܡܥܠܡ ܦܝܗ ܐܠܢܒܠܠ ܐܠܪܐܚܫ، ܦܪܐܗ ܒܠܝܡܠ
ܐܡܠܡܗ ܦܡܠܐܟܠܗ ܡܒܝ ܐܠܟܪܐܡܐ، ܐܘܒܗ ܐܠܐܘܠܝܗ ܒܟܠܗ
ܐܠܛܢܝܗ ܐܠܒܝ ܡܘܪܠ ܐܠܐ ܐܠܢܝܡܗ ܐܠܡܫܘܪܗ.

5.3 Serto script, Jacob of Edessa vowels

ܟܬܒܐ (ܐܰܡܳܐ)، ܐܚܝܐ (ܐܰܚܝܳܐ)، ܒܝܠܢܘ (ܒܝܠܢܘ)، ܐܠܗܘܐ
(ܐܗܘܘܡ)، ܟܬܒܠܐ (ܟܬܒܘܠܐ)، ܗܦܟܠ (ܗܦܟܘܠ)، ܠܙܦܟܠ
(ܠܙܘܟܠ)، ܝܒܠܒܐ (ܝܒܡܘܐ)، ܘܗܠܒܐ (ܘܗܡܘܐ)، ܒܢܙܒܐ (ܒܢܙܘܐ)،
ܝܒܢܒܐ (ܝܒܡܘܐ)

5.4 Estrangello script, Syriac language

ܟܘܬܪܐ

ܫܘܒܚܐ ܠܗܘ ܕܐܠܒܫܐ ܗܘܪܢܐ ܒܒܘܟܪܗ ܡܕܝܢܐ ܢܗܝܪ ܘܢܗܪܗ.
ܘܡܠܦܢܘܬܗ، ܡܢܝܢܐ ܠܥܠܡܘܗ، ܠܗܕܢܐ ܗܘܪܢܝܬܐ ܐܟܚܕ ܘܐܫܠܡ.
ܐܟܘܡܬܐ ܒܢܝܢܐ ܘܡܠܦܢܐ ܡܗܝܢܐ ܕܒܕܬܗ ܠܡܟܫܦܢ ܫܦܝܪܐ
ܐܪܡܝܐܝܬ ܗܘܒ ܘܠܫܒܘ. ܘܠܟܠܐ ܚܝܠܐ ܕܟܪܘܙܘܬܗ ܦܪܘܫܘܬܐ ܟܘ
ܟܕ ܟܘ ܦܫܡܘ ܘܬܪܓܡܘ
ܐܡܪ ܐܢܐ ܡܟܝܠ ܕܐܢܫܝ ܪܘܚܐ ܡܠܬܪܐ ܒܡܠܝܢ ܢܘܟܪܝܬܐ
ܫܡܠܝܢ : ܕܟܠܟܐ ܘܡܫܝܐ ܘܡܘܬܪܐ ܕܡܢܗܪܐ ܕܬܘܪܝ ܡܚܠܠܐ
ܐܫܦܝܢ. ܚܙܐ ܟܢ ܕܗܘܐ ܕܫܒܚܐ ܟܢ ܐܟܢܐ ܐܢܫܐ ܐܬܚܫܒܘ :
ܡܢ ܒܪܗ ܐܬܒܪܝܘ ܘܚܫܢܘ. ܕܐܪܬܝܢ ܕܝܢ ܕܫܪܪܐ ܠܗ ܢܫܬܠܡܘܢ ܒܠ

ܡܪܢܐ ܕܒܠܬܢܐ ܕܒܦܠܬܢܐ ܐܬܦܪܣܘ. ܡܢܗ؛ ܐܝܟ ܕܠܒܘܠܐ
ܪܒܬܗܘܢ. ܘܠܦܘܬ ܫܐܠܬܗܘܢ : ܠܪܡܬܐܝܬ ܪܒܬܐ ܘܦܚܢܐܝܬ
ܫܢܝܬܐ ܕܒܝܬܐ ܠܡܠܦܢܐ ܦܘܩܐ ܡܢ، ܡܠܝܟܘ ܢܦܫܟ ܕܕܪܡܣܘܩ
ܕܡܫܝܚܐ ܦܩܕܬ. ܘܠܡܠܟܘܬܗ ܬܬܒܝܐ ܒܗܢܐ ܒܠܒܐ «ܕܕܪܘܡܣܩܐ
ܫܪܘܪܐ» ܒܢܫܬ ܘܡܚܫܬ. ܟܠ ܚܕ ܚܕ ܕܝܢ ܡܢ ܦܫܩܘܬܐ : ܕܘܪܫܐ
ܡܪܝܡ ܐܘܦܩܬ ܐܒܕܢ ܕܡܫܒܚܐܬ ܘܕܠܠܐܬ ܠܗܝܐ ܠܥܠܐ ܫܒܝܚܐ
ܬܠܬܐ ܢܐܠܦܘܢ. ܘܠܡܫܚܬܐ ܕܘܫܐ ܘܐܕܪܒܬܐ ܢܚܡܘܢ ܘܢܪܦܘܢ

5.5 Estrangello script, Arabic language (Garshuni)

ܐܠܦܨܠ ܐܠܟܬܐܒܗ

ܘܩܕ ܩܐܡܬ ܐܠܐܡܡ ܒܚܪ ܐܠܐܡܡ، ܘܐܠܐܩܠܝܡ ܒܚܪ ܐܠܐܩܠܝܡ،
ܘܠܡ ܝܒܩ ܐܠܐܡܢ ܐܠܐܒܕܝ ܒܝܢܗܡ، ܘܚܒܬ ܐܠܐܡܢܝܗ
ܐܠܡܣܝܚܝܗ ܡܢܝܪܗ ܡܥܓܘܠܗ ܐܠܒܐܒܠ ܒܠ ܐܒܕܝܬܐ. ܘܟܠܡܐ
ܐܫܬܕܬ ܦܝ ܐܠܝܒܐܕ ܢܬܐܝܬ ܡܫܪܩܗ ܒܠ ܓܐܬܡܗܐ ܠܬܘ،
ܠܗܢܗ ܡܢ ܗܘ ܒܪܐܡ ܐܘ ܡܚܬܡ
ܘܐܓܝܪܐ ܐܬܬ، ܢܗܡ ܐܝܢܒܪ ܦܝܗ ܐܠܓܐܬܡ ܐܫܗܕ ܐܠܠܗ
ܬܥܐܠܝܐ. ܦܝܐܢܗ ܐܢ ܐ ܐܓܘܙ ܡܢ ܚܐܬܐ. ܐܘ ܗܚܐ ܦܝ
ܓܒܪ ܥܕܝܪ ܡܢ ܓܒܪ ܥܢܝ. ܝܒܕܐ ܐܢ ܐܓܕܬ ܦܝ ܐܠܣܐܠ
ܠܐܠܣܡܗܬܐ.
ܓܪܒܬ ܐܠܐܡܢܝܗ ܐܠܡܬܕܪܒܗ ܒܫܟܠ ܒܬܬܗܐ، ܘܫܐܪܬ ܦܝ
ܐܠܒܪܝܗ ܐܠܕܫܟܐܪ ܦܝܗ ܐܓܒܐܗܐ ܡܢ ܡܟܠ، ܣܬܐ ܘܡܠܬ ܐܠܝ
ܐܠܟܘܢ ܐܠܕܝ، ܝܬܘܡ ܦܝܗ ܐܠܬܓܠ ܐܠܡܐܢܣ، ܦܪܐܬܗ ܒܐܬܠܥܐ
ܐܒܐܬܗ ܦܫܐܠܬܗ ܫܕܝ، ܐܠܒܝܢ، ܐܬܓܗ ܐܢܬܕܠܝ ܒܠܝ ܐܠܒܪܝܗ
ܐܠܕܝ، ܝܩܕܠ ܐܠܝ ܐܠܢܕܒܝܗ ܐܠܡܣܘܪܗ.

5.6 Estrangello script, Jacob of Edessa vowels

ܟܬܒ̈ (ܐܡܗ)، ܬܥܣܝܢ (ܬܣܝܢ)، ܒܢܠܣܦ (ܒܝܒܣܒ)، ܐܬܪܗܐ،

(ܐܘܪܗܝܐ)، ܠܒܝܒܐܬܐ (ܠܟܬܒܢܘܬܐ)، ܡܦܩܩܠܐ (ܡܦܩܘܠܐ)، ܦܬܩܐ
(ܦܬܩܘܬܐ)، ܥܠܡܐܬܐ (ܥܠܡܝܘܬܐ)، ܕܥܠܡܐܬܐ (ܕܥܠܡܝܘܬܐ)، ܡܠܟܘܬܐ
(ܡܠܟܘܬܐ)، ܥܠܡܬܐ (ܥܠܡܘܬܐ)

5.7 East Syriac script, Syriac language

ܢܘܗܪܐ

ܢܘܒܗܪܐ ܠܗܘ ܕܠܐܠܗܐ ܗܘܕܝܐ ܒܟܘܡܪܐ ܡܕܝܢܬܐ ܬܡܢ ܘܕܡܝܬ.
ܘܡܠܟܘܬܗ ܡܫܝܚܐ ܠܥܠܡܝܢ ܒܗܘܢܝܐ ܗܘܕܝܐ ܐܝܟܢܐ ܘܐܬܝܬ.
ܘܐܬܪܐ ܒܣܪܐ ܘܡܠܟܝܐ ܡܗܝܡܢܐ ܕܝܢܗܘܢ ܠܡܫܡܫܢܐ ܥܒܕܐ ܐܪܡܝܐ
ܗܢܘܢ ܘܡܠܟܘ. ܘܠܡܠܟܐ ܣܓܝܐ ܕܡܕܝܢܬܐ ܦܪܘܫܬܐ ܒܗ ܥܡ ܒܗ
ܦܬܚܘ ܘܐܬܪܗܘܢ

ܐܡܪ ܐܢܐ ܫܡܥܬ ܕܐܢܫܝܢ ܒܫܡܐ ܡܢܗܪܐ ܒܗܠܝܢ ܢܘܗܪܐ ܣܓܝܐܝܢ :
ܕܡܕܝܢܬܐ ܡܫܡܫܐ ܘܡܫܡܫܢܐ ܕܫܡܝܢܐ ܕܗܘܕܝܐ ܡܠܟܘܬܐ ܐܢܐ. ܗܕܐ ܡܢ
ܕܗܢܘܢ ܕܗܝ ܡܢ ܐܝܟܢܐ ܕܫܡܥܬ ܐܬܗܝܡܢܘ : ܡܢ ܥܕܡܐ ܐܬܝܬܐ

ܡܬܬܓܪܓܪܢܐ ܡܬܒܠܐ ܕܟܬܒܐ ܥܠ ܐܝܩܪܐ. ܘܟܠܗܝܢ ܕܗܝܡܢܘܬܐ ܒ
ܕܠܡܪܝܐ ܢܛܪܗ ܡܫܪܪܐ ܐܠ ܓܝܪܡܘܬܐ ܠܗܕܐ ܠܗܢܐ ܗܘ ܗܘ ܟܕܐܢܐ
ܐܘ ܡܢܗܘܢ

ܘܐܓܕܪܐ ܐܝܟ ܗܘܢ ܐܬܟܠ ܦܝܫ ܕܠܓܒܪܝܢ ܐܫܘܢ ܕܠܠܗܘܢ ܗܡܢܡܝܢ،
ܦܪܝܫܢ ܐܦ ܐܢ ܐܓܘܢ ܡܢ ܡܢܐ، ܐܘ ܗܡܢ ܒ ܓܒܪ ܥܒܝܕ ܡܢ ܓܒܪ
ܥܡ. ܒܐܬ ܐܝ ܐܕܘܢ ܒ ܕܠܫܢܐ ܠܐܠܫܡܘܫܐ.

ܓܪܡܗ ܕܠܐܡܪܬܐ ܕܠܡܫܒܚܬܐ ܒܗܟܠܗ ܢܟܗܢܐ، ܘܗܢܪܗ ܒ ܕܠܐܒܪܢܝܐ
ܕܠܒܫܢܪ ܦܝܫ ܐܓܘܢܐ ܡܢ ܡܟܠ، ܣܗܕ ܘܫܠܡ ܐܠ ܕܠܟܘܢ ܕܠܒܝ
ܟܗܕ ܦܝܫ ܕܠܕܓܠܠ ܕܠܫܢܝܐ، ܦܪܢܗܘ ܒܢܠܗܐ ܐܡܢܡܐ ܦܗܢܠܗܘ
ܗܢܒܝ ܕܠܟܘܫܐ، ܐܪܓܡܗ ܐܢܗܒܠܝܢ ܥܠܐ ܕܠܒܪܢܝܐ ܕܠܒܝ ܢܘܫܠ ܐܠ
ܕܠܫܒܝܫܬܐ ܕܠܡܫܘܚܬܐ.

References

[1] Daniel Berry and Johny Srouji. Arabic formatting with ditroff/ffortid. *Electronic Publishing—Origination, Dissemination and Design*, 5(4):163–208, 1992.

[2] C. Brockelmann. *Syrische Grammatik, mit Paradigmen, Literatur, Chrestomathie, und Glossar*. VEB Verlag Enzyklopädie, Leipzig, 1938.

[3] S.J. Costaz, L. *Syriac-English Dictionary*. Imprimerie Catholique, Beyrouth, 1963.

[4] K. Georr. *Les catégories d'Aristote dans leurs versions syro-arabes*. Institut Français de Damas, Beyrouth, 1948.

[5] A. Guillaumont. *Sémitismes dans les Logia de Jésus, retrouvés à Nag-Hamâdi*. Société Asiatique, Librairie Orientaliste Paul Geuthner, Paris, 1958.

[6] Y. Haralambous. Typesetting the holy Qur'ān with TEX. In *Proceedings of the 2nd International Conference on Multilingual Computing—Arabic and Latin script (Durham)*, 1992.

[7] Y. Haralambous and J. Plaice. First applications of Ω: Greek, Arabic, Khmer, Poetica, ISO 10646/UNICODE, etc. In *Proceedings of the 15th TEX Users Groups Annual Meeting (Santa Barbara)*, 1994.

[8] A. Nakanishi. *Writing systems of the World*. Charles E. Tuttle Company, Tokyo, 1980.

[9] Imprimerie Nationale. *Les caractères de l'Imprimerie Nationale*. Imprimerie Nationale, Éditions, Paris, France, 1990.