

## GARSHUNOGRAPHY: TERMINOLOGY AND SOME FORMAL PROPERTIES OF WRITING ONE LANGUAGE IN THE SCRIPT OF ANOTHER

George A. KIRAZ

The sociolinguistic phenomenon of using the script of one language to write utterances of another language which already has a script of its own lacks a linguistic technical term. (This is not to be confused with the act of writing a text in one language and reading it in another for which the term *alloglottography*, coined by Ilya Gershevitch,<sup>1</sup> already exists.) While the paper does not assume an intrinsic formal correlation between language and script, it does assume that there is a sociolinguistic association between them; e.g., people associate the Arabic script with the Arabic language, and the Hebrew script with the Hebrew language. The phenomenon at hand occurs when a community makes a deliberate choice, for whatever reason, to write a language that it perceives not to be its own (although it might technically be its native tongue) in a script that it perceives to be its own; e.g., writing Arabic in the Hebrew script, traditionally called Judeo-Arabic.<sup>2</sup> It is this phenomenon for which this paper seeks to find a technical term. As a point of departure, three plausible terms are considered but are deemed inappropriate as they have their own technical sense in literature on writing systems that differs from that required here. These terms are allography, transliteration, and heterography.

### 1. *Allography*

The term *allography* is tempting and indeed has been considered for this purpose.<sup>3</sup> After all, the prefix *allo-* gives the meaning of *variation* or

<sup>1</sup> I. GERSHEVITCH, *The alloglottography of Old Persian*, *Transactions of the Philological Society*, 1979, p. 114-190.

<sup>2</sup> For Judeo-Arabic, see B. HARY, *The importance of the orthography in Judeo-Arabic texts*, in *Proceedings of the Tenth World Congress of Jewish Studies*, division D, vol. 1, Jerusalem, 1990, p. 77-84; G. KHAN, *Judeo-Arabic*, in K. VERSTEEGH (ed.), *Encyclopedia of Arabic Language and Linguistics*, vol. 2, Leiden, 2007, p. 526-536.

<sup>3</sup> See note 1 above.

*departure from the normal* which may be suitable in this case. Indeed, writing Arabic in the Syriac script rather than in the Arabic script<sup>4</sup> can be seen as a variation from the norm. The prefix *allo-* has been used in this sense in a number of disciplines, especially in the sciences: *allopatric* in biology describes species that are originally similar but have become quite different due to external factors, and *allotropes* in chemistry are the same elements that have different bonds to hold their atoms together. Closer to home, the prefix *allo-* has been used extensively in linguistics: *allophones* in phonology are realized variants of the same phoneme, and *allomorphs* in morphology are realized variants of the same morpheme. Indeed, scholars of writing systems built on these concepts and coined the term *allograph*<sup>5</sup> in an analogous manner. Coulmas<sup>6</sup> gives two technical senses for the term *allograph*: Firstly, to denote a diachronic graphical variant of writing; e.g., the interchangeable use of <u> and <v> in the Latin script, or the writing of Dutch <ȳ> for <ij>. Secondly, to denote realized variants of a grapheme such as Greek <σ> and <ς> or Syriac <ⲁ>, <Ⲓ>, and <ⲓ>. In typography, the different appearance of letters from one font to another can also be considered *allographic*;<sup>7</sup> e.g., <a a a> (Arial, Garamond, Lucidia) are all *allographs*. In the field of computer processing of handwriting, computer scientists talk of *allograph recognition*;<sup>8</sup> i.e., recognizing variations of the same grapheme for which they build an *allograph lexicon*. As tempting as it may be to use the term *allography* to describe writing a language in the script of another, the fact that it has already been assigned a technical sense in the field of linguistics, particularly in writing systems, makes it

<sup>4</sup> On Arabic written in Syriac, see J. ASSFALG, *Arabische Handschriften in Syrischer Schrift (Karšūnī)*, in W. FISCHER (ed.), *Grundriß der Arabischen Philologie*, I, Wiesbaden, 1992, p. 291-303; A. MENGOZZI, *The History of Garshuni as a Writing System: Evidence from the Rabbula Codex*, in F.M. FALES – G.F. GRASSI (ed.), *CAMSEMUD 2007: Proceedings of the 13<sup>th</sup> Italian Meeting of Afro-Asiatic Linguistics*, Padova, 2010, p. 297-304 (= MENGOZZI, *The History of Garshuni*).

<sup>5</sup> Daniels objects to the term on theoretical grounds; see P. DANIELS, *Is a structural graphemics possible?*, in *LACUS Forum*, 18 (1991), p. 528-537; P. DANIELS, *Reply to HARRICK*, in *LACUS Forum*, 21 (1991), p. 425-431. Others have objected to Daniels' objections; see E.M. HARRICK, *Of course a structural graphemics is possible*, in *LACUS Forum*, 21 (1994), p. 413-424; E.M. HARRICK, *Reply to Daniels reply*, in *LACUS Forum*, 21 (1994), p. 432-440; H. ROGERS, *Writing Systems. A Linguistic Approach*, Oxford, 2005, p. 11.

<sup>6</sup> F. COULMAS, *The Blackwell Encyclopedia of Writing Systems*, Oxford (UK) – Cambridge (Mass., USA), 1999, p. 9 (= COULMAS, *Writing Systems*).

<sup>7</sup> G. SAMPSON, *Writing Systems*, London, 1985, p. 25.

<sup>8</sup> H. TEULINGS, L. SCHOMAKER, J. GERRITSEN, H. DREXLER and M. ALBERTS, *An on-line handwriting-recognition system based on unreliable modules*, in R. PLAMONDON – C.G. LEEDHAM (ed.), *Computer Processing of Handwriting. 4th Conference: Selected papers*, 1990, p. 167-85.

undesirable. Additionally, the sense that we are aiming at here differs radically from the existing technical sense of the term *allograph*.

## 2. Transliteration

Another tempting term is *transliteration*. The term, however, is problematic for not a few reasons. Firstly, the writing of one language in the script of another does not always follow a transliteration methodology. For instance, while writing Arabic in the Syriac script mostly relies on transliteration, writing Armenian,<sup>9</sup> Malayalam,<sup>10</sup> and other languages<sup>11</sup> in the Syriac script is more of a transcription rather than transliteration. Secondly, the term *transliteration* is usually used for another sense in writing systems. Transliteration usually occurs in the following two scenarios:<sup>12</sup> To embed a term (or terms) from one language into another; e.g., where one uses a transliteration scheme to render Chinese names in an English text. Or to embed a term for linguistic or scientific purposes; e.g., when one writes a Hebrew sentence in roman transcription in a linguistics or scholarly article.

## 3. Heterography

Yet another term that has already been used for this purpose is *heterography*.<sup>13</sup> This term too is spoken for in writing systems. Coulmas gives the following definition: “a differentiation in spelling which distinguishes different meanings of homophonous words or phrase.”<sup>14</sup> An example of

<sup>9</sup> On Armenian written in the Syriac script, see D.S. MARGOLIOUTH, *The Syro-Armenian Dialect*, in *Journal of the Royal Asiatic Society*, (1898), p. 839-861; S. BROCK, *Armenian in Syriac Script*, in D. KOUYMIAN (ed.), *Armenian Studies. In Memoriam Haïg Berbérian*, Lisbon, 1986, p. 75-80; H. TAKAHASHI – J.J.S. WEITENBERG, *The Shorter Syriac-Armenian Glossary in Ms. Yale Syriac 9*, in *Journal of the Canadian Society for Syriac Studies*, 10 (2010), p. 68-83; and *The Shorter Syriac-Armenian Glossary in MS. Yale Syriac 9. Part 2: Glossary in Transcription/Translation*, in *Hugoye: Journal of Syriac Studies*, 14.1 (2011), p. 87-144; H. TAKAHASHI, *Armenisch-Garschuni (Armenisch in syrischer Schrift)*, in the current volume, p.??-??.

<sup>10</sup> On Malayalam written in Syriac, see T. KOONAMMAKKAL, *An Introduction to Malayalam Karshon*, in *The Harp*, 15 (2002), p. 99-106.

<sup>11</sup> For a survey of languages written in the Syriac script, see G.A. KIRAZ, *Tūrāṣ Mamllā. A Grammar of the Syriac Language*, vol. 1: *Syriac Orthography* (forthcoming).

<sup>12</sup> R. Sproat: personal communication.

<sup>13</sup> MENGOZZI, *The History of Garshuni*, p. 297; A. MENGOZZI, art. *Garshuni*, in S. BROCK, A. BUTTS, G. KIRAZ, and L. VAN ROMPAY (ed.), *Gorgias Encyclopedic Dictionary of the Syriac Heritage*, Piscataway, 2011, p. 172 (= MENGOZZI, *Garshuni*).

<sup>14</sup> COULMAS, *Writing Systems*, p. 202.

English heterography is *right*, *rite*, *write*, and *wright*. Using it for our purposes would entail a radical departure from its current sense. Additionally, the term's prefix, *hetero-* is associated with a phenomenon not too far from our topic; viz., *heterograms*. A heterogram is a word (or morpheme) that is spelled exactly as it would be spelled in its source language, but is intended to be read in the target language.<sup>15</sup> For instance, an Iranian text written in the Aramaic script<sup>16</sup> might include Aramaic <mlk?> /malkā/ 'king' in an otherwise Iranian sentence. The heterogram <mlk?> is read *šāh* in Middle Persian, but *ʔxšēwanē* in Sogdian.<sup>17</sup> (This is not to be confused with alloglottography where the entire text is written in one language but read in another.) The close proximity of the term *heterography* to *heterogram*, but more seriously its existing technical sense in writing systems, makes it inadequate for our purposes. Hence, there remains a need to coin an unambiguous term for the phenomenon at hand.

#### 4. Garsunography

Rather than coining a term from scratch, one ought to attempt to search for a gloss that already exhibits the desired semantic features. Indeed, such a gloss already exists, albeit with limited circulation: Garšūnī (sometimes Karšūnī, also spelled Caršūnī).<sup>18</sup> While the term originally referred to writing Arabic in the Syriac script, its semantic scope became wider over time. It was later extended in speech and writing to denote the writing of a number of languages using the Syriac script. In recent times, the term was further expanded, in speech but not writing this time (at least I have not seen it in writing), to denote the writing of Syriac in scripts other than the Syriac script, something of a reverse Garšūnī.<sup>19</sup> Yet another recent expansion of the term, also in speech but not writing, is to denote the same relation between script and

<sup>15</sup> R. SPROAT, *A Computational Theory of Writing Systems*, Cambridge, 2000, p. 187-188.

<sup>16</sup> For a survey of Iranian written in the Aramaic script, see P. SKJÆRVØ, *Aramaic scripts for Iranian languages*, in P. DANIELS – W. BRIGHT (ed.), *The World's Writing Systems*, New York, 1996, p. 515-535 (= DANIELS – BRIGHT, *The World's Writing Systems*.

<sup>17</sup> *Ibidem*, p. 520.

<sup>18</sup> On the term Garshuni, see MENGOZZI, *Garshuni*, p. 172.

<sup>19</sup> I have personally heard the term used in this sense frequently. It seems that I am not alone here as someone recorded the same on the Wikipedia article on Garshuni referring to "spoken Assyrian [sic] written using Latin script" (checked Oct. 25, 2011). Similar statements were made on my Facebook page after I made an inquiry into the matter (checked Oct. 27, 2011).

language in cases when Syriac is *not* involved at all; e.g., writing Arabic in the Latin script or vice versa. As the term Garšūnī already exhibits the semantic features we are looking for, I propose to use the term *garšūnography* to refer to the writing of a particular language in the script of another whether in transcription or transliteration form, and regardless of the use of Syriac either as the source language or the target script.

As the gloss is not rooted in the Latin-Greek tradition, an objection might be raised that it is ‘foreign’ and may sound unfamiliar outside of Syriac studies. This should not preclude us from using it. The use of foreign terms in linguistic terminology is not uncommon. The term *schwa*, from the Semitic grammatical tradition, comes to mind. It too was known locally to a specific community and its circulation was then broadened first by philologists and then by linguists. Another term is *sandhi*, from Sanskrit (meaning *joining*), which refers to phonological processes that occur at the morpheme or word boundary. In writing systems Daniels<sup>20</sup> borrowed the term *abjad*, an Arabic mnemonic that refers to the Arabic ‘alphabet’, to denote a consonantary rather than a full fledged alphabet. In typography, the Persian term *kashida*, which refers to the elongation line between two graphs at the baseline, is now ubiquitous, even used to denote the same act in Arabic which has a term of its own for this purpose, viz. *taṭwīl*. All of these terms went through an internationalization process. In our case, we are fortunate enough that the internationalization process for the term Garšūnī has already begun independently. It was mentioned above that the sense of the term has already expanded in speech more than in writing. Admittedly, the term Garšūnī is some sort of an ethnolexeme: all speakers who use it in its expanded meaning are somehow sociolinguistically connected with Syriac. However, they utter the word /garšūnī/ when speaking various languages, mostly Arabic, Turkish, and Malayalam, but also a variety of European languages, due to the diaspora, particularly English, German, French, Dutch and Swedish. Arguably, the term Garšūnī, with its expanded meaning, already exists in the mental lexica of these languages but with limited circulation. Indeed, the word *Garšūnī* is already recorded in the *Oxford English Dictionary* with a citation from as early as 1856,<sup>21</sup> and in *Webster’s Third*.<sup>22</sup> The term has the additional advantage that its ety-

<sup>20</sup> P. DANIELS, *The Study of Writing Systems*, in DANIELS – BRIGHT, *The World’s Writing Systems*, p. 4.

<sup>21</sup> J. SIMPSON – E. WEINER (ed.), *Oxford English Dictionary*, 2<sup>nd</sup> ed., 20 vols., 2009.

<sup>22</sup> Art. *Karshuni*, in *Webster’s Third New International Dictionary, Unabridged*, 2002, <http://unabridged.merriam-webster.com> (25 Oct. 2011).

mology, origin, and basic meaning are unknown. It has been suggested that it might be related to Gershon of the Old Testament (Ex 2.22), and in Levantine colloquial Arabic, “speaking *Garšūnī*” means “speaking gibberish” but nothing else is known about it. That its origin is unknown gives it additional neutrality and possible universality. That its etymological meaning is unknown saves us from ambiguities: it only means writing a language in the script of another, and nothing more. In writing systems too, *garšūnography* would be unambiguous, unlike the competing terms *allography*, *transliteration*, and *heterography*.

Building on the term *garšūnography*, one can then speak of *garšūnographic* writings. It would then follow that the writing of the same text of a particular language in two or more scripts of other languages is *allogaršūnography*; i.e., different *garšūnographic* realizations of the same text. While historically it might be difficult to find such *allogaršūnographic* texts, they are common today among diaspora communities; e.g., a Syriac liturgical text is given in the Syriac script, the Latin script, and the Arabic (or Malayalam) script. In this example, the Latin and Arabic texts are said to be *allogaršūnographic*; so are the Latin and Malayalam texts. Here, our use of the prefix *allo-* agrees with all the above mentioned terms in 0 where it refers to realized variants of the same entity but in a different shape or form. In all of these terms, one can opt for using <sh> instead of <š> for simplicity the same way *šwā* is written *schwa* or *shwa*; e.g., the title and abstract of this paper. Hence, one can speak of *garshunography* and *allogarshunography*, and of *gars-hunographic* and *allogarshunographic* writing.

In the traditional usage of the term *Garšūnī*, it is implied that the source language is Arabic and the target script is Syriac. To indicate another source language, the name of that language is used as a modifier; e.g., Armenian *Garšūnī* is Armenian written in the Syriac script. My usage of the term *garšūnography* is rather neutral and uncapitalized. Hence, *garšūnography* is a system of writing regardless of the source language or the target script. In cases where Syriac is not involved, a scholarly tradition of compounding the target script with the source language already exists; e.g., Judeo-Arabic, Armeno-Turkish, and very recently Syro-Ottoman (over Ottoman *Garšūnī*).<sup>23</sup> These terms can still be preserved and used, when a need arises, as a modifier to the term *garšūnography*; hence, one can speak of *Syro-Arabic garšūnography* (i.e., traditional *Garšūnī*), *Greco-Ottoman garšūnography* (Ottoman

<sup>23</sup> B. TRIGONA-HARANY, *A Description of Syro-Ottoman*, in E. BALTA – M. ÖLMEZ (ed.), *Between Religion and Language: Turkish-Speaking Christians, Jews and Greek-Speaking Muslims and Catholics in the Ottoman Empire*, Istanbul, 2011, p. 15-41.

Turkish written in the Greek script), *Armeno-Syriac garšūnography* (Syriac written in the Armenian script),<sup>24</sup> etc.

One can extend this terminology to the graphemic level. In all garšūnographic systems, the set of graphemes of the target script is not sufficient to represent the graphemes (in the case of transliterations) or the phonemes (in the case of transcriptions) of the source language. One is forced to introduce new graphemic entities which we may call *garšūnographemes*. Their realized forms, when applicable, are then *allogaršūnographs*. Syro-Arabic garšūnography, for instance, adapts the Syriac grapheme < ܢ > into the garšūnographeme < ܢ > for Arabic < ن >. The allographs < ܢ ܢ ܢ ܢ > are then allogaršūnographs. There are instances when a grapheme is borrowed from the script of the source language; e.g., there are a few Malayalam graphemes that appear in Syro-Malayalam garšūnography. We can call such a grapheme a *hetero-garšūnographeme* (analogous to *heterogram* above, p. ??).

Defining an upper ceiling for the use of the term garšūnography requires serious consideration. I claimed above that the term Garšūnī “only means writing a language in the script of another, and nothing more.” Having said that, the term, even in its current expanded usage in modern speech, requires at least two conditions: Firstly, that the source language is associated with a script that is perceived to be its own. Secondly, that there exists readership which is either unfamiliar with the script of the source language or prefers, for whatever reason, to use the target script over the script of the source language. In other words, there exists a dichotomy between language and script, and there is intent to deviate from the norm. More formally, let  $l$  be a given language and  $\text{SCRIPT}(l)$  be a function that returns the sociolinguistically ‘normative’ script  $s$  of  $l$ , then classic cases of garšūnography require a pair of languages,  $l_1$  and  $l_2$ , and a pair of scripts,

$$\begin{aligned} s_1 &= \text{SCRIPT}(l_1) \\ s_2 &= \text{SCRIPT}(l_2) \end{aligned}$$

Writing  $l_1$  in  $s_2$  and writing  $l_2$  in  $s_1$  is said to be garšūnographic writing. This would cover all traditional Garšūnī systems that involve Syriac in either direction, as well as Judeo-Arabic, Ladino, Greco-Ottoman, Armeno-Turkish, Syro-Turkic,<sup>25</sup> etc. Recall that the current meaning of

<sup>24</sup> For Syriac written in the Armenian script, see A. SCHMIDT, *Arménien et syriaque*, in C. MUTAFIAN (ed.), *Arménie: la magie de l’écrit [exposition, Marseille, Centre de la vieille charité, 27 avril-22 juillet 2007]*, Paris, 2007, p. 345–348.

<sup>25</sup> On Syro-Turkic including further references, see W. KLEIN, *Syriac Writings and Turkic Language according to Central Asian Tombstone Inscriptions*, in *Hugoye: Journal of Syriac Studies*, 5 (2002), p. 213-224.

Garšūnī does not require Syriac to be involved which permits us to classify all these systems as garšūnographic. Next, we consider cases when the script of one language is adapted to be used as the script of another language that does not have a script of its own; e.g., the adaptation of the Arabic script to write Ottoman Turkish, Urdu, Persian, etc. Are these systems cases of garšūnography? I am personally inclined to consider these systems *nongaršūnographic* on sociolinguistic grounds. Both of the conditions just mentioned are not met: the source language has no ‘normative’ script of its own which then makes the second condition mute. In order to widen the scope of the term garšūnography to include cases like Ottoman Turkish, the term must first be freed from its sociolinguistic attachments. Doing so, however, would make defining an upper bound to garšūnography impossible. Almost all writings, at some point or another, used a script that was not normative for some initial period of time. For this reason, I consider Ottoman Turkish a case of nongaršūnography. Obviously, writing Dutch in the Latin script is also nongaršūnographic. All of this does not preclude the transformation of writing from being garšūnographic to nongaršūnographic. We have seen that Ottoman Turkish is nongaršūnographic. When writing Turkish in the Latin script was first introduced, one can argue that this writing was garšūnographic for some period of time, probably between 1929 to the mid 1930s. Once writing Turkish in the Latin script became dominant, this writing became nongaršūnographic. There remains one tempting case of writing to consider under the umbrella of garšūnography: writing one’s own dialect in the script of one’s own formal language; e.g., writing Egyptian colloquial Arabic in the Arabic script or Neo-Aramaic in the Syriac script. Here, people do not perceive themselves as writing in a script of the other.<sup>26</sup> Nor do they conceive a dichotomy between language and script. Additionally, the formal constraints set above do not apply. As such, these cases ought to be considered nongaršūnographic as well.

## 5. Glossary

It may be helpful to conclude with a glossary of all the terms that have been introduced in alphabetical order:

*allogaršūnographs* Allographs of a garšūnographeme.

<sup>26</sup> On the Neo-Aramaic case, see H.L. MURRE-VAN DEN BERG, *From a Spoken to a Written Language. The Introduction and Development of Literary Urmiah Aramaic in the Nineteenth Century*, Leiden, 1999.

*allogaršūnography* (adj. *allogaršūnographic*) The realization of variants of a *garšūnographic* text; e.g., Syriac text written in both the Arabic and Latin scripts.

*garšūnographeme* (adj. *garšūnographemic*) An adaptation of a grapheme from the target script to represent a grapheme or phoneme from the source language; e.g., <ﺀ> from Syriac <ﺀ> for Arabic <ح>.

*garšūnography* (adj. *garšūnographic*) the writing of one language in the script of another in a specific sociolinguistic setting.

*hetero-garšūnographeme* (adj. *hetero-garšūnographemic*) A borrowing of a grapheme from the script of the source language to be used in the target script; e.g., Malayalam graphemes used in Syro-Malayalam *garšūnography*.

Only time will determine if the terminology presented here will be accepted by linguists, philologists, and *garšūnographists*. If nothing else, it is hoped that this paper may help in formalizing some of the properties of what I opt to call *garshunography*.

#### **Abstract**<sup>27</sup>

This paper proposes terminology for a sociolinguistic phenomenon of writing a language in the script of another. Initially, three plausible terms are examined: allography, transliteration, and heterography. The paper will show that these terms are already spoken for in the field of writing systems with technical meanings that differ radically from the desired terminology. New terminology based on the existing term *Garšūnī* is proposed; viz., *garshunography* with various derivatives. The paper also examines some of the sociolinguistic properties of *garshunography*.

<sup>27</sup> I am grateful to Richard Sproat, my former mentor and colleague at Bell Labs, for reading and commenting on an earlier draft. My search for a term coincided with a request into the same from J. den Heijer, A. Schmidt, and T. Pataridze, editors of the current volume *Scripts beyond Borders. A Survey of Allographic Traditions in the Euro-Mediterranean World*. This paper benefited from discussions with Sebastian Brock, Daniel King, Aaron Butts, Hidemi Takahashi, and Benjamin Trigona-Harany.

