

Proto- und Linear-Elamisch: Berechnung des Jaccard-Index zur Valorisierung der grafischen Ähnlichkeit [Version 05/2020]

Michael Mäder
michael.maeder@isw.unibe.ch
Institut für Sprachwissenschaft, Universität Bern

[A] Einleitung

Die beiden hier behandelten Schriftsysteme wurden ursprünglich als eines angesehen (vgl. Scheil 1905). Dies änderte sich erst, als die kohärente und im Vergleich reduzierte Zeichenauswahl der elamischen Strichschrift (künftig LE) erkannt wurde sowie die Tatsache, dass hier anders als im Proto-Elamischen (PE) keine numerischen Zeichen vorkommen. Später hat die Archäologie mit gesicherten Datierungen des PE (3200 – 2900 v. Chr.) und des LE (ab 2250 v. Chr.) endgültig eine Bresche zwischen die beiden Schriftsysteme des alten Irans geschlagen. Intuitiv lieferte aber die grafische Ähnlichkeit sowie die mit zunehmenden Funden immer deutlichere Überlappung der Verbreitungsgebiete den beteiligten Forschern weiterhin guten Grund, die These im Hinterkopf zu behalten, dass LE ein Abkömmling des PE darstellt. Kürzlich konnte zudem mittels syntaktischer Analysen aufgezeigt werden, dass LE fürwahr die elamische Sprache abbildet (Mäder et al. 2018).











Hinzufügen möchte ich eine schrifttheoretische Überlegung: Von den (je nach Zählweise) mehreren Hundert archaischen Schriftsystemen sind nur gerade deren vier (Woods 2010:15) bis sieben (Mäder 2019:12) endogen, d.h. wurden unabhängig von bestehenden Schriftkenntnissen erfunden. Dass eine Schrift ohne Modell erschaffen wurde, ist also in derartigem Mass eine Ausnahme, dass wir im Zweifelsfall eine Vorläuferschrift anzunehmen haben. Schon allein dahier ist eine Untersuchung der graphemischen Ähnlichkeit ist also gerechtfertigt. Kritikern, die möglicherweise den berühmten Ausspruch von Peter Daniels (1996:22) anführen – "The comparison of shapes alone without attention to sound values [is] the oldest fallacy in the study of writing systems." – möchte ich gleich von Anbeginn den Wind aus den Segeln nehmen: Daniels warnt zurecht davor, Verwandtschaftsbehauptungen zwischen Schriftsystemen sowie Entzifferungsvorschläge allein aufgrund grafischer Ähnlichkeit vorzubringen. Doch bei den altiranischen Schriften steht es andersrum: Ein Verdacht auf Verwandtschaft ist kulturgeschichtlich wie auch schriftarchäologisch gegeben, und in der vorliegenden Arbeit geht es darum, die Daten zu liefern, um diesen Verdacht zu überprüfen.

[A] Bisherige Einschätzungen, neue Datenlage


















Vorauszuschicken ist, dass kein Schriftstück bekannt ist, das archäologisch zwischen das Ende des Proto-Elamischen (2900 v. Chr. gemäss mittlerer, 2800 v. Chr. gemäss kurzer Chronologie) und das Auftauchen des Strichschrift-Elamischen (2250 v. Chr. gemäss mittlerer, 2150 v. Chr. gemäss kurzer Chronologie) datiert wird (Dahl 2009:27). Paläografisch hingegen scheinen die zwei Schriftsysteme zwar anders, aber nicht voneinander getrennt zu sein, wie die folgenden Beispiele aus der Literatur zeigen: Piero Meriggi (1971) nennt in der bisher ausführlichsten paläografischen Studie das Proto-Elamische "scrittura proto-elamica A" und das Linear-Elamische "scrittura proto-elamica B". Marie-Joseph Steve (1992:4) geht gar soweit zu behaupten, dass "[L'élamite linéaire] a dû coexister à la fois avec les pictogrammes proto-élamites les plus récents et l'écriture d'emprunt suméro-accadienne (...)." Beide Forscher schätzen die Anzahl Zeichen, die LE und PE gemein haben, auf 30-35 (Meriggi 1971:185 sowie 193-203, § 508; siehe Tab. 1). Steve reportiert (1992:4): "Déjà en 1971, P. Meriggi avait mis en évidence la relation entre ces deux écritures et inventorié une trentaine de signes communs aux deux systèmes."



Linear Elamite			Proto-Elamite		Meriggi's Beurteilung (Meriggi 1971: 193–202)
siglum (Plachtzik et al. 2017: Tab. 7)	siglum (Meriggi 1971: 193–202)	shape	shape	siglum (Meriggi 1974: 8–24)	
	1			M297 sgg.	sim.
	1 b			M298	id.
	2			M301	id.
	4			M308 a	id.
	5			M218	id.
	6			M219	id.
	7			M229g	id.
	8			M220	id.
	11			M223	id.
	13			M252g	sim. ?
	14			M248 sgg.	sim.
	15			M263	sim.
	19			M374	sim. ?
	20			M371	id.
	22			M24	id.
	26			M99	id.
	26 a			M96	id.
	27			M344	id.
	28			M58	id.
	29			M25	sim.
	34			M9	id.
	36			M4	sim.
	42			M316	sim. ?
	43			M94 l	sim. ?
	44			M347	sim.
	44 a			M347	id.
	46			M69 l	sim.
	47			M259	sim.
	48			M319	id.
	50			M341	sim. ?
	51			M57	id.
	53			M134 a	sim.
	53a			M131	id.
	57			M372	sim.
	57 a			M372	sim.
	58			M372 a	sim.

Tab. 1: Die von Meriggi (1971:193-203, § 508) als in beiden Schriftsystemen vorkommend betrachteten Zeichen. Sie sind entweder identisch (id.) oder ähnlich (sim.).

Für weitere 43 Zeichen fand er keine Entsprechung. Dies sind 1a ; 3 ; 9 ; 10 ; 12 ; 14a ; 14b ; 15 ; 15b ; 15c ; 15d ; 16 ; 16a ; 16b ; 16c ; 17 ; 18 ; 21 ; 23 ; 24 ; 25 ; 28a ; 29a ; 30 ; 31 ; 32 ; 33 ; 35 ; 37 ; 37a ; 38 ; 39a ; 40 ; 41 ; 45 ; 45a ; 46a ; 49 ; 51 ; 55 ; 56 ; 59 ; 60 . Mit leichten Abweichungen in den Details können die 34 Identifikationen aus Tab. 1 auch mit heutigem Wissen bestätigt werden; siehe Tab. 3 und Tab. 4. Dass heute mehr identische und ähnliche Zeichen ausgemacht werden können, ist darauf zurückzuführen, dass die beiden Forscher ihre Einschätzungen auf eine noch stark limitierte Datenbasis stützen mussten. Bis vor wenigen Jahren waren bloss 18 LE-Inschriften mit 553 Zeichen bekannt, während wir heute durch Neufunde und die Öffnung diverser Privatsammlungen 41 Inschriften (bei Mitzählen aller Fragmente sind es deren 50) mit insgesamt 1731 Zeichen zur Verfügung haben (OCLEI 2020).

[A] Mögliche Zwischenformen zwischen PE und LE

Die LE-Inschrift ^{Susa}M, die wohl jünger ist als die von Puzur-Inshushinak in Auftrag gegebenen Texte (Dahl 2009:29), trägt auf der ersten Zeile der Rückseite die Zeichen . Meriggi (1971:185) macht nun darauf aufmerksam, dass die ganze erste Zeile aus typischen und überaus häufigen PE-Zeichen bestehe und hält sie für proto-elamisch. Besonders häufig sei im PE die Zweiersequenz , die im PE x Mal vorkomme. Heute ist aufgrund der erweiterten Datenlage allerdings einzuwenden, dass die vier Zeichen auch im LE häufig genug sind, um zum Standard-Syllabar gezählt zu werden:  belegt von den insgesamt 99 Zeichentypen den 4. Häufigkeitsrang (mit 55 Attestationen),  den 29. Rang (mit 15 Attestationen),  den 24. Rang (21 Attestationen) und  immerhin den 32. Rang mit 14 Vorkommnissen (Plachtzik et al. 2017:Tab. 7). Als zweites Beispiel zu dieser Inschrift führt Meriggi (ebd.) die Sequenz  als typisch proto-elamisch an, doch ist auch hier auf ^{Susa}F und ^{Susa}I zu verweisen, wo diese Sequenz jeweils den Anfang des ansonsten  geschriebenen Eigennamens *Pu(?)–uz(?)–ri–shu–shi–na–k* darstellt. ^{Susa}M dürfte also nicht, wie von Meriggi angenommen, auf der Vorderseite LE und auf der Rückseite PE wiedergeben, sondern wie ursprünglich von Hinz (1969:28) postuliert ein reines linearelamisches Dokument sein. Das Beispiel –auf weitere will ich verzichten – zeigt aber umso mehr, wie verwechselbar die Zeichenformen der zwei hier verglichenen Schriftsysteme zuweilen sind.

Zwei andere Schriftstücke möchte ich hingegen als intermediär zwischen PE und LE vorschlagen. Das erste ist die von Hinz (1969: Taf. 16-17) noch als strichschriftelamisch aufgefasste, spätestens aber seit Desset (2012:93, Fn. 2) aus dem LE-Corpus ausgeschlossene Inschrift O (Abb. 1). Sie besteht, wenn man sie aus LE-Sicht betrachtet, ausschliesslich aus hapax legomena, d.h. keines seiner 41 lesbaren Zeichen kommt in einer anderen linearelamischen Inschrift vor, jedenfalls nicht als identische grafische Form. Immerhin xx Zeichen sind einer LE-Type ähnlich, wie Tab. 2 zeigt.¹ Nebst der Hypothese, dass O eine Zwischenstufe zwischen PE und LE darstellt, sind weitere Möglichkeiten in Betracht zu ziehen. Die Erkenntnis, dass sich kein einziges der 41 lesbaren Zeichen wiederholt und jede der acht Zeilen genau sechs Zeichen aufweist (Dahl 2009:29), könnte auf einen Schultext hinweisen (darauf deuten unschön ausgeführte Zeichen wie  anstelle der wohl beabsichtigen Type  hin), oder aber, wie Dahl (ebd.) vorschlägt, auf einen Traumtext, entstanden aus Fantasie oder religiöser Eingebung.

Tab. 2: Die Inschrift O und ihre grafischen Korrespondenzen mit LE bzw. PE. (Abb. aus Hinz 1969: Taf. 17).

¹ Für eine Definition der Begriffe *identical*, *similar* und *different* vgl. die methodischen Grundlagen der GEAS – Alice Kober Gesellschaft für die Entzifferung antiker Schriftsysteme, zu finden unter https://center-for-decipherment.ch/standards/GEAS_Methodology_for_Statistical_Script_Comparison.pdf.

Doch nicht nur ^{Div}O, sondern auch ^{Susa}J, ^{Susa}K und ^{Susa}L bestehen ausschliesslich aus Zeichen, die auch in PE vorkommen. Gleichzeitig sind dies jene Inschriften, die nicht ins Syntax-Muster passen, das die anderen Susa-Inschriften zeigen.² Und sie zeigen Zeichenformen wie \boxtimes (^{Susa}L II 5), die ausschliesslich im PE, nicht aber im LE vorkommen. All diese Punkte könnten darauf hinweisen, dass ^{Susa}J, ^{Susa}K und ^{Susa}L gar nicht reine LE-Inschriften sind, sondern ebenfalls zu dieser Zwischenstufe oder gar zu einer bisher unbekannten Kategorie kursiver – d.h. nicht mit einem Keilgriffel eingedrückten, sondern mit einem spitzen Gegenstand in den Ton geritzten – Inschriften zu zählen sind. Diese Hypothese muss in einer eigenen Untersuchung geprüft werden.

Für eine allfällige Zwischenstufe kommen vielleicht auch die wenigen nicht-numerischen, d.h. nicht wie administrative Dokumente wirkenden, protoelamischen Inschriften in Frage. Das längste dieser seltenen Art ist in Abb. 1 abgebildet. Meriggis (1971-1974) umfassender Studie stand sie leider noch nicht zur Verfügung. Eine Sammlung und anschliessende syntaktische und paläografische Analyse ist ein dringendes Desiderat.

Abb. 1: Eine nicht-numerische protoelamische Tafel. (Fotografie aus Amiet 1986:260. Fig. 48).

[A] Zeichenkorrespondenzen zwischen PE und LE

Als Grundlage für die vorliegende Studie werden in Tab. 3 die grafischen Entsprechungen zwischen PE und LE, d.h. jene Zeichen, die in beiden Schriften vorkommen, aufgelistet. Es wird zwischen *identischen* (= *identical*) und *ähnlichen* (= *similar*) Zeichenformen unterschieden. Zeichen, die keines der beiden Kriterienbündel erfüllen (= *different*), werden selbstredend nicht aufgeführt. Definiert sind die hervorgehobenen Begriffe in der GEAS-Terminologie³, die wie folgt lautet:

- 1) A sign form is a new sign type (e.g. L instead of K), if it has always or sometimes another shape and always another sound value than other signs of the same corpus.
- 2) A sign form has a new ordnungszahl (e.g. K₂ instead of K₁), if it has the same sound value but a different or similar shape [in the sense of 3) below] from the other ordnungszahlen of a sign type.
- 3) The shape of two signs can be "different (and not similar)", "(different, but) similar" or "the same (= identical)".
 - 3.1) A sign shape is called "different (and not similar)", if two or more of the following conditions – so-called "differences" – are fulfilled: a) The sign has at least one supplementary, intentionally drawn stroke; b) the length of a single stroke intentionally differs more than 35% relative to the corresponding stroke of the sign compared to it; c) the angle between two strokes intentionally differs by more than 30 degrees; d) a stroke is intentionally rounded (or intentionally unrounded) in contrast to the corresponding stroke of the sign compared to it; e) when a sign is intentionally mirrored or turned by more than 30 degrees.
 - 3.1.1) Definition of "intentional": An intention is given, if either a) a difference in the sense of 3.1) is repeated more than once in the same sub-corpus, or b) a difference in the sense of 3.1) is present in an inscription which i) shows a high level of calligraphic regularity, and ii) shows, as a whole, no stylistic feature (i.e. no serifs or similar style mutations) which could account for the formal difference in question.
 - 3.2) A sign shape is called "(different, but) similar", if A) exactly one of the parameters for "different" in 3.1 [a), b), c), d) or e)] is observed.
 - 3.3) A sign shape is called "the same (= identical)", if it is not different (3.1) and not similar (3.2).
- 4) A sound value can be "different (and not similar)", "(different, but) similar" or "the same (= identical)".
 - 4.1) A sound value is called "different (and not similar)", if neither place of articulation nor manner of articulation are the same.
 - 4.2) A sound value is called "(different, but) similar", if either place of articulation or manner of articulation are the same.
 - 4.3) A sound value is called "the same (= identical)", if both place of articulation and manner of articulation are the same.

Diese Methodik anwendend, werden nun in Tab. 3 die identischen aufgelistet und in Tab. 4 die ähnlichen, d.h. jene Zeichen, die sich genau durch eines der folgenden Merkmale – und zwar durch ein

² Dies zeigt sich daran, dass es weder Mäder et al. (2018:60-75) noch Desset (2012:112) gelingt, in ^{Susa}J, ^{Susa}K und ^{Susa}L signifikante Sequenzen zu finden, die in anderen LE-Texten vorkommen.

³ https://center-for-decipherment.ch/standards/GEAS_Methodology_for_Statistical_Comparison.pdf.

vom Schreiber beabsichtigtes⁴ – vom verglichenen Zeichen unterscheiden: a) durch maximal einen zusätzlichen oder fehlenden Strich, b) durch maximal eine (um mehr als 35%) abweichende Strichlänge, c) durch maximal einen (um mehr als 30°) abweichenden Winkel, d) durch Rundung oder Nichtrundung oder e) durch Spiegelung oder (um mehr als 30° betragende) Drehung. Nicht aufgelistet sind letztlich die Zeichen, die *different* sind, d.h. bei dem sich mehr als einer der unter a) bis e) genannten Unterschiede in Bezug auf eine tentativ verglichene Zeichenform festzustellen sind.

Elamische Strichschrift			Proto-Elamisch	
Siglum (Plachtzik et al. 2017: Tab. 7)	Form	Neue Type?	Form	Siglum (Meriggi 1974: 8–24)
	l	y	!	M1b
	↖	y	↖	M3c
	+++	y	+	M6; 137g
	≠	y	≠	M7; M328
	∴	n	∴	M372
		y		M11
	⊠	y	⊠	M24c; 262d'
	⊠⊠	y	⊠	M29
	⊐	n	⊐	M32
	⊐	n	⊐	M32a
	⊠	y	⊠	M29j
	*	y	*	M46
	≡	y	≡	M57
	⌘	y	⌘	M57a; M57a'
	≡≡	y	≡≡	M58
	⊖	y	⊖	M68b
	⊖	y	⊖	M68d
	▽▽▽	y	▽▽▽	M69i; M131; 160a
	⌘	y	⌘	M96; M305k
	⌘	y	⌘	M99
	⌘	y	⌘	M107
	⌘	n	⌘	M116a
	⌘	y	⌘	M130
	⌘	y	⌘	M134b
	⌘	y	⌘	M148a _{ins.} ; 154a
	⌘	y	⌘	M179; M286
	⌘	y	⌘	M203

⁴ Das Kriterium der Absichtlichkeit ist in den GEAS-Richtlinien, die ursprünglich für die Ähnlichkeits-Quantifizierung der verschiedenen norditalischen Schriften entwickelt worden sind, zentral. Die Absichtlichkeit ist dort in Punkt 3.1.1 ("Definition of 'intentional'") eingegrenzt. Spezifisch für die hier untersuchte elamische Strichschrift bzw. die Proto-Elamische Schrift werden folgende Eigenschaften nicht als intentional betrachtet, da sie nach derzeitigem Wissen weder durch die Zeichenfunktion noch durch die Paläografie (d.h. geografische oder chronologische Unterschiede) zu erklären sind, sondern fakultative, wohl rein stilistisch motivierte Eigenheiten darstellen: Es sind dies die Punkte an den Strichenden (z.B. ⌘/⌘ oder ⊐/⊐) sowie Existenz und Art der an Rhomben angebrachten "Fähnchen" (z.B. ⌘/⌘/⌘ oder ⌘/⌘/⌘).

	y		M214
	y		M218
	y		M219
	y		M220
	y		M221
	y		M222
	y		M223a
	n		M223e
	y		M225
	y		M226f
	y		M227b; M248
	y		M229g
	y		M229h; M236a; M264h
	y		M233
	n		M243c
	y		M243g; M244b'
	n		M243i
	y		M244e; M263
	y		M260h
	n		M248a
	y		M252g; M353
	n		M262a
	y		M296
	n		M297; M298
	y		M296 + M301 _{ins.}
	y		M302e
	y		M306
	y		M308
	n		M308a
	y		M315g
	y		M316f; e; M321l
	n		M319
	y		M319c
	y		M343
	n		M343a
	y		M364
	y		M374b
	y=51 n=13		

Tab. 3: Liste der identischen Zeichen in den beiden Schriftsystemen.

Elamische Strichschrift			Proto-Elamisch	
Siglum (Plachtzik et al. 2017: Tab. 7)	Form	Neue Type?	Form	Siglum (Meriggi 1974: 8–24)
		n		M318
		n		M316; M316l
		n		M302f
		y		M305n
		y		M290h
		n		M266; M295x
		y		M266h; M266g
		n		M267d
		n		M286a; M286b
		y		M288; M305r
		y		M289a; b; c; 305o
		n		M263a
		n		M252o; M292g'
		n		M258; M259a; b; M62; M257a; b
		n		M260; M305q'
		y		M248h; 256; 280
		y		M237e
		y		M243b; M246
		n		M225b
		n		M226b
		n		M226c; M226d
		y		M228a'; M228b
		y		M216; M320; j
		y		M205b; e
		n		M157a
		n		M143a
		y		M96a; M97h
		y		M71c
		n		M79
		y		M85; M185; M83

	n		M68g
	y		M59i
	n		M53a'; c; M54
	n		M41c
	y		M30e
	n		M26c
	y		M27
	n		M26h; j
	n		M28; M30a
	n		M25
	n		M24a
	y = 17 n = 24		

Tab. 4: Liste der ähnlichen Zeichen in den beiden Schriftsystemen.

Während für die elamische Strichschrift (LE) nach derzeitigem Stand xy grafische Varianten verteilt auf 99 Zeichentypen bestehen, sind es im Proto-Elamischen (PE) ca. 1500 grafische Varianten verteilt auf 393 Zeichentypen (Meriggi 1974:24). Um objektive Schlüsse aus den Tab. 3 und 4 zu gewinnen, wollen wir zuerst definieren, dass nicht grafische Varianten, sondern Zeichentypen (d.h. Syllabar-Einträge) verglichen werden. Daraus folgt zweitens, dass beim kleineren Korpus, also dem LE mit seinen 99 Typen, nur jene Varianten addiert werden, die einer neuen, d.h. bisher nicht addierten, Type zugeordnet werden. In diesem Fall wird in Spalte 3 ("New Sign Type?") ein y für "yes" notiert. Gehört die verglichene Variante hingegen zu einer Type, die vorangehend bereits addiert wurde, wird ein n für "no" notiert. Die Summe der y-Einträge ist somit die Anzahl aus den 99 EL-Zeichentypen, die ein identisches (Tab. 3) oder ein ähnliches (Tab. 4) Pendant aus dem PE hat. Die Summe aus y- und n-Einträgen hingegen ist die Anzahl aus den xy insgesamt im LE vorhandenen grafischen Varianten, die ein Pendant aus dem PE hat. Folgende Liste fasst die Summen zusammen:

- 68 der 99 EL-Zeichentypen (67,3 %) sind identisch oder ähnlich.
- 51 der 99 EL-Zeichentypen (50,5 %) sind identisch.
- 17 der 99 EL-Zeichentypen (16,8 %) sind ähnlich.
- 105 der 413 EL-Formvarianten (25,4 %) sind identisch oder ähnlich.
- 64 der 413 EL-Formvarianten (15,5 %) sind identisch.
- 41 der 413 EL-Formvarianten (9,9 %) sind ähnlich.

Dies ist der momentane Stand unter der angewandten Methode. Er kann und wird sich mit jeder weiteren gefundenen Inschrift verändern.

[A] Systematische Gemeinsamkeiten

Kleinere Auffälligkeiten innerhalb der Variantengruppen deuten ebenfalls auf eine Verwandtschaft der beiden Schriften hin: So gibt es sowohl im PE wie auch im LE verschiedene Variante von winkelförmigen Zeichen, sog. Tetrissteine. Sowohl im PE wie im LE steht jede Variante auf dem Stamm (/) oder auf der Hypothenuse (/), nie aber auf der langen Seite oder der Spitze.

Letztlich ist die doch sehr spezifische und zur fraglichen Epoche nirgendwo sonst vorgefundene Eigenschaft der eingeschriebenen Zeichen zu erwähnen. Dies sind komplexe Zeichen, die aus zwei anderen in der Schrift vorhandenen Zeichenformen zusammengefügt wurden, wobei eines als Matrix und das andere als verkleinertes Insertum fungiert:

Linear Elamite			Proto-Elamite	
Matrix Sign	Inserted Sign	Combined Sign	Combined Sign	sigl. Meriggi 1971
				M233
				M229h; M236a; M264h
				M228i
				M222
				M229g
				M226f
				M226k
				M233
				M225b
	+			
	+ +			M225c
		;		
		;		
		;		
				M195C _{ins.}
				M109
				M263
				M29j
				M68h
(vacat)				M295x; M278i
(vacat)				M267d
(vacat)	+	; ; ;	x	M277b; M279; M278a' _{ins.} ; M277e _{u.} ; M267c; M276f; M290h
				M57a; M277b; M57a'
				M348b; M352i
			+ _{ins.}	M296 + M301 _{ins.}
				M85; M185; M83

Tab. 4: Einschreibung (Insertion) im LE und im PE

Die typischen Matrix-Zeichen sind im PE wie auch für LE dieselben: Rhomben (◊) sowie Rhomben mit Fähnchen sind in beiden Schriftsystemen die mit Abstand häufigsten Matrices, und allem Anschein nach die einzigen produktiven (d.h. jene, die mit beinahe allen Formen inskribiert werden können). Weitere Zeichen werden ein oder zweimal als Matrices genutzt und andere gar nie; dies gilt für LE wie auch für PE. Auch die Inskripta scheinen nicht gleichverteilt zu sein, es gibt typische wie oder sowie etliche mehrfach vorkommende, doch die Mehrzahl der Zeichen – sowohl im PE wie im LE – wird niemals inskribiert. In beiden Schriftsystemen existieren etliche Fälle, wo das Matrixzeichen mehrere Diakritika annehmen kann, so bei xy im PE oder ; ; ; ; ; im LE. Weiter gibt es in beiden Schriftsystemen die Dichotomie zwischen punktierten und unpunkteten Linien, z. B. so / (M44a;

b) vs. ≪ (M44) oder ≪ (M59) vs. ≪ (M59a) im PE oder ≪ vs. ≪; ꝛ vs. ꝛ und viele andere. Bei den Rhomben gibt es bezüglich den "Fähnchen" eine in beiden Schriften gleich geartete Auswahl: Die Rhomben erscheinen a) ohne Fähnchen (◊), b) mit einfachem senkrechten Fähnchen, c) mit einfachem geneigten Fähnchen (◊), d) mit geknicktem geneigtem Fähnchen, e) mit zwei einfachen geneigten Fähnchen, f) mit zwei geknickten geneigten Fähnchen, g) mit doppelt geknicktem Fähnchen; h) mit drei einfachen Fähnchen, i) mit drei einfachen Fähnchen sowie einem Querbalken. Die Fähnchen sind immerzu oben und in seltenen Fällen oben und unten angebracht. Niemals kommen andere Fähnchen-Anbringungen vor wie z.B. an den seitlichen Ecken angebrachte.⁵ Solche systematischen Übereinstimmungen bezüglich den "erlaubten" bzw. "unerlaubten" Features sind meines Erachtens nicht durch Zufall zu erklären.

[A] Berechnung des Jaccard-Indexes

Die relative Ähnlichkeit bzw. Übereinstimmungsrate zwischen zwei vermutlich oder sicherlich miteinander verwandten Schriftsystemen kann durch den Concordance Index (I_c) bestimmt werden. Für den (hier vorliegenden) Fall, dass nur zwei Schriftsysteme miteinander verglichen werden, wird ein Spezialfall des I_c angewandt: der auch in anderen Forschungsgebieten verwendete Jaccard-Index. Die Formel lautet wie folgt (vgl. Wandeler & Fuls 2019:1):

$$T / (A_1 + A_2 - T) = J$$

Dabei ist T die Anzahl graphisch identischer Zeichen. Dieser Wert wird ins Verhältnis zur Gesamtanzahl von Zeichen in beiden Systemen (hier PE_t bzw. LE_t) gesetzt. Der Jaccard-Index (J) liegt immer zwischen Null und Eins. Je höher der Jaccard-Index, desto näher sind sich die verglichenen Schriftsysteme. So herrschen zwischen den Lateinischen Alphabeten, die bekanntlich den Grossteil der Zeichentypen gemeinsam haben, hohe Jaccard-Werte zwischen 0,575 und 0,897. Die grösste Ähnlichkeit haben das englische und dänische Alphabet ($J = 0,897$), die sich nur durch die drei zusätzlichen dänischen Zeichentypen Æ, Ø und Å unterscheiden. Der Tiefste gemessene Wert innerhalb der von Wandeler & Fuls (2020:3) ausgewählten Latein-Alphabete ist mit $J = 0,575$ der Jaccard-Wert für das lateinische (Spätantike) und das tschechische, da letzteres zahlreiche zusätzliche diakritisierte Buchstaben beinhaltet. Dass man, wie in obigem Beispiel, nur Zeichentypen (Alphabeteinträge) und nicht alle bekannten grafischen Varianten einbezieht, ist nur bei entzifferten Schriften möglich. Bei teil- oder unentzifferten Schriften hingegen ist die Funktion und somit die Zugehörigkeit jeder Variante noch nicht ermittelt. Folglich ist die Anzahl Zeichentypen unbekannt oder unsicher. Dies stellt aber kein Problem dar, da der Jaccard-Index auch unter Einbezug aller grafischen Varianten ermittelt werden kann. A_1 und A_2 haben dann absolut höhere Werte, doch die Relation bleibt valide, ja ist aufgrund der höheren Werte sogar aussagekräftiger, da einzelne Entscheide, ob eine Variante ins Inventar aufzunehmen ist oder nicht, weniger ins Gewicht fallen. Wenn ganze Varianten-Inventare verglichen werden, sinken die Werte tendenziell, ein Quervergleich zwischen Jaccard-Werten von kanonisierten Schriften wie den modernen Latein-Alphabeten mit teil- oder unentzifferten Schriften ist somit nicht aussagekräftig. Innerhalb von teil- oder unentzifferten Schriften sind Vergleiche hingegen valide. Ein Beispiel einer Gruppe von miteinander verwandten teilentzifferten (und somit mit all ihren grafischen Varianten erfassten) Schriften sind die norditalischen Alphabete. Wandeler & Fuls (2020:4) ermitteln Jaccard-Werte zwischen 0,140 und 0,263. So haben Lepontisch (225 grafische Varianten) und Rätisch (183 grafische Varianten) 85 identische Zeichenformen gemeinsam. Wir setzen in obige Formel somit $T = 85$, $A_1 = 225$ und $A_2 = 183$ ein und erhalten den Jaccard-Index von 0,263:

⁵ Nach derzeitigem Wissen scheinen Punkte und Fähnchen eher orthografische Varianten als funktionale (d.h. semantische oder phonologische Diakritia) zu sein, siehe Fn. 4. Dies schmälert aber nicht die Systematik, ähnlich wie wir auch in der modernen Typographie Fonts mit und ohne Serifen haben, deren Anbringung systematisch, nicht aber funktional ist.

$$85 / (225 + 183 - 85) = 0,263$$

Den kleinsten Jaccard-Wert unter den norditalischen Alphabeten hat das Schriftenpaar Etruskisch (177 grafische Varianten) und Germanische Runen (125 grafische Varianten): Sie teilen 37 identische Zeichenformen und haben somit einen Jaccard-Wert von 0,140.⁶

Nach dieser Kontextualisierung können wir uns nun der Ermittlung des Ähnlichkeitsmasses zwischen PE und LE widmen. Die Gesamtzahl grafischer Varianten des Proto-Elamischen ist dabei PE_t , jene des Linear-Elamischen LE_t :

$$T / (PE_t + LE_t - T) = J$$

[B] Der Jaccard-Index auf Basis aller grafischen Varianten

Für das PE zählt Meriggi 2276 grafische Varianten. Sie verteilen sich auf 393 Typen (Meriggi 1974:8-24). Zuerst muss der grundsätzliche Entscheid getroffen werden, ob alle Varianten oder nur die kondensierten Typen verglichen werden sollen. Dass bei noch unentzifferten Schriftsystemen die Anzahl Zeichentypen einer erheblichen Willkür unterliegt, wurde oben bereits erwähnt; die Zahl ist von der intuitiven Entscheidung des Kompilators abhängig, welche Variantengruppen einen eigenen Typeneintrag erhalten und welche nicht. Einer noch grösseren Willkür ist aber die Anzahl grafischer Varianten unterstellt, denn die Differenzen in der Entscheidung des Kompilators steigt linear mit der Anzahl Einheiten. Dabei sticht Meriggis (1974) Kompilation dadurch hervor, dass auch kleinste grafische Unterschiede, z.B. zwischen und , dazu führten, eine neue Variante aufzulisten. Dies führt – zusammen mit der Tatsache, dass Meriggi auch die aufgrund von Korrosion des Schriftträgers nur teilweise erkennbaren Zeichen auflistet (z.B.), zu der sehr hohen Zahl von 2276.⁷ Ein weiterer Effekt, der die Anzahl grafischer Varianten unzulässig erhöht, ist die im PE zahlenmässig stärker als im LE ins Gewicht fallende, beinahe vollproduktive Diakritisierung, d.h. die Einschreibung von Inskripta in ein Matrixzeichen (siehe unten Tab. 4). Beinahe die Hälfte von Meriggis Formvarianten sind solche kombinierten Grapheme. Da die vorliegende Arbeit stark tentativen Charakter hat und sich für den quantitativen Vergleich noch keine saubere Methodik entwickelt hat, schlage ich gleichwohl vor, versuchsweise mit diesen Zahlen zu arbeiten. Wir setzen also $PE_t = 2276$. Für die elamische Strichschrift – für die das eben beschriebene Problem auch gilt, zusammen mit dem gegensteuernden Effekt des Fundzufalls – sind derzeit 413 grafische Varianten des LE verzeichnet.⁸ Wir setzen somit $LE_t = 413$. Den Wert für T schliesslich können wir dem aus Tab. 3 und Tab. 4 ermittelten Resultat entnehmen wonach LE mit PE 64 identische Formvarianten gemeinsam hat. Somit gilt:

$$64 / (2276 + 413 - 64) = 0.024$$

[B] Der Jaccard-Index auf Basis der Zeichentypen

Vielsagender ist es, mit den Typen statt den Formvarianten zu arbeiten. Während die ungleichen Kompilationsmethoden verschiedener Forscher bei den Formvarianten zu schier unüberwindlichen Unstimmigkeiten führen (siehe oben "Jaccard-Index auf Basis aller grafischen Varianten"), wiegen diese Probleme bei den Zeichentypen aus folgenden Gründen weniger schwer: Erstens werden bei der Syllabarerstellung, also bei der Zuordnung der Formvarianten zu Typen, bereits intern oder extern

⁶ Für die Basisdaten der norditalischen Schriften siehe PNI.

⁷ Ein weiterer Effekt, allerdings mit entgegengesetzter Wirkung (d.h. einer Reduktion der grafischen Varianten), ist der Fundzufall. Er ist allerdings bei den ca. 1700 Inschriften des PE zu vernachlässigen.

⁸ OCLEI (2020). Dieser Wert, gilt für den Stand Mai 2020 und wird sich durch neugefundene Inschriften verändern. Er kann auf folgende Art aktualisiert werden: Man wähle im OCLEI-Entzifferungstool (<https://center-for-decipherment.ch/tool/#elam>) unter "Settings" die Syllabar-Variante "Each sign separately" und kopiere die so gewonnene Zeichenliste in ein Programm, das die Zählung der Zeilen erlaubt. Die Anzahl Zeilen entspricht der Anzahl Zeichenvarianten.

begründete methodische Regeln befolgt. Zweitens stützen sowohl Meriggi (1974) als auch Plachtzik et al. (2017) ihre Syllabare auf Erkenntnisse, die sie aus Sequenzvergleichen: So wurde fürs LE beispielsweise erkannt, dass aufgrund der Sequenzen « ⊗ !» und « ⊗ !» die Zeichen ⊗ und ⊗ Formvarianten der selben Type sind, auch wenn dies der Intuition widerspricht. Zusammengefasst heisst das, dass die Auflistung der Formvarianten ausschliesslich auf Intuition und willkürlichen Entscheiden beruht, während die Erstellung des Syllabars zumindest teilweise faktisch begründet ist. Die Berechnung des Jaccard-Indexes auf der Grundlage der Zeichentypen ist somit aussagekräftiger. Für das PE erstellt Meriggi (1974:8-24) ein Syllabar von 393 Typen, für das LE zählen Plachtzik et al. 99 Typen. Wichtig ist nun, dass wir den Resultaten aus Tab. 3 und 4 nicht die identischen Formvarianten, sondern die identischen Zeichentypen entnehmen, also 51. Somit gilt:

$$51 / (393 + 99 - 51) = 0.115$$

Es ergibt sich ein Jaccard-Index von 0.115. Es ist aufgrund der noch schlecht erprobten Schriftvergleichsmethode noch nicht möglich, aussagekräftige Quervergleiche zu anderen, d.h. nicht mit den elamischen verwandte Schriftsysteme zu vergleichen. Je öfter die Methode aber angewandt werden wird, desto besser wird dieser Wert kontextualisiert werden können.

[A] Zusammenfassung

Literatur

- Amiet, Pierre (1986):** L'âge des échanges Inter-Iraniens, 3500 - 1700 avant J.-C. Notes et documents des Musées de France 11. Paris.
- Dahl, Jacob L. (2009):** Early writing in Iran, a reappraisal. *Iran* 47, 23-31.
- Daniels, Peter T. (1996):** The First Civilizations. In: Daniels, P. & Bright, W. (Hrsg.): *The World's Writing Systems*. Oxford, 21-28.
- Desset, François (2012):** Premières écritures iraniennes – Les systèmes proto-élamite et élamite linéaire. Napoli.
- Mäder, Michael; Balmer, Stephan; Plachtzik, Simon und Rawyler, Nicolai (2018):** "Sequenzanalysen zur elamischen Strichschrift", in: B. Mofidi-Nasrabadi, D. Prechel, A. Pruß (Hrsg.): *Elam and its Neighbors. Recent Research and New Perspectives. Proceedings of the international congress held at Johannes Gutenberg University Mainz, September 21-23, 2016* (*Elamica* 8), 49-104.
- Mäder, Michael (2019):** Ist die Donauschrift Schrift? Eine systematische Untersuchung der Zeichensequenzen aus der Vinča-Kultur (5200 – 3400 v. Chr.). Budapest: Archaeolingua.
- Meriggi, Piero (1971):** "La scrittura B", in: *La scrittura proto-elamica. Parte Ia: La scrittura e il contenuto dei testi*. Roma. 184-224.
- OCLEI:** Online Corpus of Linear Elamite Inscriptions, www.elamicon.org
- Plachtzik, Simon & Rawyler, Nicolai (2017):** Das Syllabar der elamischen Strichschrift. https://center-for-decipherment.ch/pubs/plachtzik-et-al-2017__das-syllabar-der-elamischen-strichschrift/
- PNI: Mäder, Michael (forthcoming):** Paleografia norditalica. *Zixu* 14.
- Scheil, Vincent (1905):** Documents en écriture proto-élamite. MDP 6, Paris.
- Steve, Marie-Joseph (1992):** Syllabaire Elamite – Histoire et Paléographie. Recherches et Publications. Neuchâtel/Paris
- Woods, Christopher (2010):** Visible language – inventions of writing in the ancient Middle East and beyond. Chicago.