

NOUVELLES ÉTUDES SUR GLOZEL (suite)

Les tablettes incrites de Glozel

par T.D. CRAWFORD

Un corpus des inscriptions figurant sur 34 tablettes de Glozel fut soumis à l'analyse d'un ordinateur, en vue d'établir une comparaison statistique avec les corpora d'autres langues connues : basque, berbère, chinois, finnois et ibérique.

Les conclusions de cette étude sont significatives : le corpus examiné ne présente aucune des caractéristiques d'un texte établi dans une langue naturelle et les méthodes linguistiques ne permettent pas de déterminer l'authenticité du matériel.

Seules des méthodes ressortissant à d'autres disciplines pourraient établir la possibilité d'une fonction non linguistique de ces tablettes incrites. Une certitude à ce point de vue ne pourrait être obtenue que par les techniques de thermo-luminescence sur un matériel provenant de nouvelles recherches à Glozel, réalisées suivant les principes modernes de fouille.

Puisque l'authenticité des trouvailles de Glozel est revenue encore une fois en discussion, il a semblé valable d'examiner les objets inscrits provenant des fouilles, afin de déterminer s'ils portent des textes dans une langue naturelle et, dans l'affirmative, de quelle langue il pourrait s'agir.

DONNÉES.

Il parut judicieux de limiter l'étude aux tablettes incrites publiées par Morlet dans son *Glozel : Corpus des Inscriptions* (Montpellier, 1965, pp. 69-90) tout en laissant de côté, pour le moment, les autres objets inscrits de Glozel (*ibid.*, pp. 25-68) et les trouvailles de Puyravel et de Chez-Guerrier (*ibid.*, pp. 90-91).

Restaient donc à notre disposition 44 tablettes, dont 2 incrites sur les deux faces. Afin de laisser vierge un matériel qui pourrait servir à une reprise éventuelle des recherches, il fut décidé d'écarter certaines tablettes qui s'avéraient illisibles, parce que endommagées (*op. cit.*, p. 73, gauche inférieure - une face d'une tablette inscrite sur les deux faces ; p. 77, les deux tablettes ; p. 78, en haut ; p. 82, en haut ; p. 85, gauche inférieure ; p. 87, droite inférieure ; p. 89, les deux tablettes). En outre, deux tablettes furent ultérieurement rejetées ; l'une (p. 83, en haut)

était impossible à transcrire à cause de son alignement très mauvais, l'autre (p. 81, en bas) s'avéra presque identique à une autre tablette déjà incluse (p. 73, gauche supérieure), à la différence d'un seul caractère à la ligne pénultième. Ces éliminations laissèrent un *corpus définitif de 34 tablettes*, dont une inscrite sur les deux faces.

**

TRANSCRIPTION DU CORPUS.

Pour faciliter l'analyse au moyen d'un ordinateur, il était indispensable de transcrire le corpus sur des cartes mécaniquement lisibles, ce qui présentait certaines difficultés.

1) La liste des signes donnés par Morlet (*op. cit.*, pp. 20-21) n'est ni exhaustive, ni logiquement ordonnée. On trouvera un certain nombre de nouveaux signes portant le total jusqu'à 133, et sans doute pourrait-on en trouver encore plus dans le matériel non examiné. D'ailleurs, quelques signes inscrits séparément dans la liste sont pratiquement non différenciables les uns des autres, par exemple les deux premières variantes du signe 71 et celles du signe 86. C'est pourquoi, il fallait prendre certaines décisions plus ou moins arbitraires en assignant aux signes du corpus une place dans tel ou tel des groupes établis par Morlet. Cependant, ces cas difficiles ne s'avérèrent pas assez nombreux pour justifier un reclassement complet, et il n'y a aucune raison de croire qu'ils eussent influencé le résultat de l'enquête.

2) En dehors de la tablette rejetée à cause de son mauvais alignement, d'autres présentaient de légers problèmes à cet égard (*par exemple*, p. 74, en bas ; p. 75, en bas), et la transcription en est parfois incertaine. Ici également les problèmes n'étaient pas assez fréquents pour compromettre la fiabilité des résultats.

3) Si l'on suppose que les tablettes contiennent des textes dans une langue naturelle, l'orientation de l'écriture n'est pas évidente ; du reste, elle n'est pas forcément la même pour toutes les tablettes, étant donné que le matériel ne peut pas être assigné automatiquement à une et même date. Puisque la plupart des tablettes sont entièrement remplies de signes, les terminaisons de lignes ne donnent que peu d'indications concernant le sens de la lecture du texte. La tablette de la page 72, en haut, montre un vide à la gauche de la ligne inférieure, ce qui suggère une lecture de droite à gauche ; mais la tablette de la page 88 (à gauche, en haut), montre un vide comparable à la gauche de la ligne *supérieure*. Puisque nous n'avons aucune raison de croire que les tablettes aient été toutes inscrites dans le même sens (vu que tant de signes ont des contreparties en sens inverse), ce vide pourrait se trouver en réalité à la droite de la ligne inférieure, ce qui suggérerait une lecture de gauche à droite. Et qui plus est, deux tablettes (p. 69 ; p. 73, en bas) donnent une forte impression d'avoir été inscrites *boustrophédon*, en commençant respectivement à la droite supérieure et à la gauche supérieure. En raison de cette incertitude, il fut décidé de transcrire les tablettes de gauche à droite et de donner la priorité à la détermination du sens véritable par le moyen des méthodes statistiques.

**

ANALYSE PRÉLIMINAIRE.

On a noté ci-dessus que des vides ne se trouvent que sur un nombre très restreint de tablettes. Or il est tout à fait improbable qu'il en aurait

été ainsi si chaque tablette eût été utilisée pour l'inscription d'un texte indépendant ; aussi, selon Morlet (*op. cit.*, p. 89), faut-il supposer qu'une inscription se continuait habituellement d'une tablette à l'autre, en utilisant ainsi toute la surface disponible, sauf peut-être sur la tablette finale d'une série. De plus, rien ne porte à croire que ces textes hypothétiques soient de simples listes et comptes, comme, par exemple, la plupart du matériel écrit en *Linéaire B*. Cette observation dernière est significative quand on considère ce que pourrait être la nature de l'écriture. Le *Linéaire B* est notoirement mal adapté à représenter le grec ; il omet beaucoup de consonnes dans la position finale de la syllabe, il simplifie les diphtongues, il aboutit souvent à des ambiguïtés. Une telle écriture peut, au besoin, suffire à la comptabilité, mais ne permettrait pas la lecture d'Homère. Pour qu'une écriture suffise aux textes courants d'une certaine longueur, elle doit être plus étroitement adaptée aux nécessités de la langue.

Il est certain que les inscriptions glozéliennes contiennent au moins 133 signes, dont 116 apparurent dans le corpus examiné. Parmi ceux-ci, 34 ne se trouvèrent qu'une fois chacun.

On peut dire à coup sûr que *l'écriture n'est pas alphabétique*, attendu qu'il y a beaucoup trop de signes. Un alphabet efficace comporte un signe pour chaque phonème de la langue, dont le nombre ne dépasse jamais par beaucoup la cinquantaine. Même si l'on laisse de côté les 34 signes qui ne se trouvent qu'une fois, (en supposant qu'ils soient des éléments étrangers à l'écriture tels que seraient £ ou \$ dans un texte français), il en resterait toujours un excès considérable. Il est vrai que les alphabets contiennent souvent des signes superflus, mais jamais en si grand nombre.

L'écriture ne peut pas non plus être un idéogramme à la chinoise, attendu qu'en ce cas le nombre des signes serait beaucoup plus élevé.

Reste *la possibilité d'une espèce de syllabaire*. Le type le plus simple de syllabaire contient des signes représentatifs (a) les voyelles isolées, et (b) les combinaisons consonne + voyelle. Le Linéaire B constitue un exemple bien connu de ce type. Pour qu'un tel syllabaire comprenne tous les phénomènes d'une langue, si x représente le nombre des consonnes et y le nombre des voyelles, le nombre des signes du syllabaire doit être $y(x + 1)$. Par exemple une langue comportant 17 consonnes et 7 voyelles (combinaison assez ordinaire) exigerait un syllabaire de 126 signes. Il s'ensuit que, si la langue hypothétique de Glozel possédait un nombre moyen de phonèmes et une écriture efficace représentait les syllabes ainsi formées, alors le nombre des signes dans *le matériel glozélien est tout à fait compatible avec l'hypothèse d'un syllabaire de ce type simple*. En pratique, les syllabaires ne sont pas souvent parfaitement réguliers, et l'on pourrait s'attendre à trouver des distinctions non établies voire même des signes représentant une voyelle suivie d'une consonne au dehors de la série principale. Mais le nombre des signes utilisés et la nécessité apparente d'une écriture efficace concourent à suggérer un syllabaire simple du type « voyelle » ou (« consonne + voyelle ») comme la meilleure hypothèse de départ.

**

CORPORA COMPARATIFS.

Il semble utile de comparer statistiquement le corpus glozélien à des corpora de dimensions semblables tirés de diverses langues connues

et transcrites dans un syllabaire du type supposé pour la langue glozélienne. C'est pourquoi, les corpora suivants furent préparés sur des cartes mécaniquement lisibles dans une écriture syllabique : 1) un mélange de textes en *basque médiéval* ; 2) des contes populaires en *berbère marocain* ; 3) des conversations en *chinois mandarin* tirées d'un livre d'enseignement ; 4) un extrait d'un roman *finnois* ; 5) un extrait d'un drame en *espagnol* des premières années de l'Age d'Or. Les corpora ont été délibérément choisis pour que leur contenu soit aussi varié que possible. Pour autant qu'on en sache, aucune de ces langues n'a de rapport avec les autres, et elles ont été choisies afin d'y comprendre la variété la plus large possible de types structurels.

Un syllabaire ne contenant que des signes « voyelle seule » et « consonne + voyelle » présente des problèmes lorsque l'on essaie de transcrire, soit des groupes de consonnes, soit des syllabes qui se terminent avec une consonne. En principe il y a deux solutions possibles, que l'on observe l'une et l'autre dans les textes grecs du *Linéaire B* : 1) laisser tomber quelques-unes des consonnes ; 2) les transcrire comme si elles étaient suivies de la voyelle qui se trouve dans la syllabe dont il s'agit. Le *Linéaire B* cadrerait avec la première solution, mais la dernière facilite mieux la lecture et pour cela fut adoptée ici. Ainsi dans le corpus espagnol « incomparable » se transcrit comme I-NI-KO-MO-PA-RA-BE-LE, chaque groupe syllabique équivalant à un signe dans le corpus glozélien. Puisque les tablettes de Glozel n'indiquent pas les séparations entre les mots, les autres corpora furent préparés de la même façon.

**

SENS DE L'ÉCRITURE.

Il y a quatre sens possibles qu'il faut considérer : 1) unidirectionnel de gauche à droite ; 2) unidirectionnel de droite à gauche ; 3) boustrophédon, commençant à la gauche supérieure ; 4) boustrophédon, commençant à la droite supérieure. Il existe une épreuve assez simple qui permet de distinguer si un corpus (ou l'ensemble d'un corpus mélangé) suit un des sens 1 ou 2 par opposition aux sens 3 et 4. Cette épreuve exige une statistique de la longueur, du nombre et de la fréquence des séquences répétées de signes. Elle s'applique également aux textes en écriture soit alphabétique soit syllabique ; la méthode est la suivante :

Prendre n'importe quel texte imprimé en français et le lire de la façon ordinaire de gauche à droite, en notant les répétitions des mots et des groupes de mots. Normalement on trouvera que le texte contient un nombre peu élevé de mots qui apparaissent plusieurs fois dans une courte étendue. D'autres mots (et peut-être des groupes de mots) apparaîtront deux ou trois fois, et l'on aura un grand reste de mots qui n'apparaissent qu'une seule fois. Lire ensuite le même texte, commençant à la gauche de la première ligne, puis rentrant sur la seconde ligne de droite à gauche, puis sur la troisième de gauche à droite, et ainsi de suite, en notant les répétitions comme auparavant. Le nombre et la fréquence des mots et des groupes de mots répétés s'avérera de beaucoup moins élevé. Si auparavant on avait trouvé 10 occurrences du mot LA, on pourrait s'attendre qu'à peu près la moitié se transforment en AL ; de ces mots qui apparaissent deux fois au cours de la lecture unidirectionnelle, la moitié en moyenne seront rencontrés maintenant une fois dans chaque sens et seront enregistrés comme deux mots différents

survenant une fois seulement. D'où le principe général : *le nombre le plus élevé de répétitions ressort d'une lecture correcte du texte.*

On préparera donc un second corpus glozélien en renversant l'ordre des signes dans toutes les lignes paires du corpus original. D'après les tableaux ci-après, le corpus original unidirectionnel sera qualifié de « *texte-U* », et le corpus boustrophédon de « *texte-B* ». Des *textes-B* comparables furent préparés également à partir des *textes-U* des cinq langues connues.

Le tableau 1 montre les résultats d'un calcul du nombre et de la fréquence des répétitions de séquences de 3 à 10 signes, chaque syllabe dans les corpora des langues connues était considéré comme un signe.

Ce tableau inspire deux observations. Premièrement, alors que les nombres de répétitions dans les langues connues montrent la différence attendue entre les *textes-U* et les *textes-B*, pour les textes glozéliens cette différence est négligable. Deuxièmement, le nombre des répétitions dans les textes glozéliens est notablement moins élevé que dans les langues connues, qu'il s'agisse des *textes-U* ou des *textes-B*.

Selon l'hypothèse que le corpus glozélien se compose de textes dans une langue naturelle, comment expliquer la faible variation entre les répétitions du *texte-U* et ceux du *texte-B* ? Il semble qu'il y ait deux explications possibles :

1) La langue glozélienne comporterait un caractère structurel qui exige le renversement des séquences syllabiques pour des raisons grammaticales, d'où il résulterait que chaque séquence ait à peu près la même fréquence d'occurrence que la séquence inverse. Par exemple, si KA-LE-SA-TO représente le temps présent d'un verbe, alors TO-SA-LE-KA pourrait servir de temps passé. Un tel trait, s'il prédominait dans la langue, risquerait d'invalider l'épreuve.

En fait cette possibilité peut être négligée sans aucun risque. Le renversement des séquences de plus de 2 mots de longueur ne se trouve pas comme caractère des langues naturelles (bien que l'on puisse construire une langue artificielle sur cette base), et le renversement des séquences de syllabes à l'intérieur d'un mot ne se rencontre jamais. Il est tout à fait improbable que la langue glozélienne puisse constituer une exception à cette règle.

2) Ou bien le corpus glozélien comporterait des tablettes unidirectionnelles et des tablettes boustrophédon en proportions à peu près égales. Hypothèse peu probable, mais non impossible. Théoriquement on peut la vérifier. On peut établir un *texte-U* et un *texte-B* pour chaque tablette et essayer toutes les combinaisons possibles ; 34 tablettes, donc 2^{34} combinaisons. En théorie... ! En pratique, attendu que 2^{34} représente un peu plus de 17 000 000 000 combinaisons et que le calcul pour chacune nécessite au moins 2 minutes sur l'ordinateur, il faut chercher une autre solution. L'on pourrait essayer cette analyse approfondie sur une fraction seulement du corpus, par exemple sur 8 tablettes ; celle-ci permettrait $2^8 = 256$ combinaisons, nombre acceptable en fonction des exigences de l'ordinateur, et ce pourvu qu'il y ait la probabilité d'un résultat significatif. En effet, l'essai fut tenté avec les 8 tablettes des pages 69, 70, 71, 72 (les deux) et 73 (gauche supérieure, droite supérieure, droite inférieure) du livre de Morlet ; on compara la combinaison des 8 *textes-U* avec celle des 8 *textes-B* et traita également des *textes-U* et des *textes-B* (convenablement réduits) des langues connues. Les résultats se voient dans le tableau 2.

TABLE 1

		Nombre d'occurrences												
		1	2	3	4	5	6	7	8	9	10	Plus de 10		
Nombre de séquences de 3 signes	Glozel	texte-U	1299	30	2	1	0	0	0	0	0	0		
		texte-B	1300	30	3	0	0	0	0	0	0	0		
	Basque	texte-U	755	120	42	15	7	3	6	3	0	0	1x13	
		texte-B	1029	94	14	9	1	1	1	0	0	0		
	Berbère	texte-U	807	105	30	4	2	5	1	3	0	4	1x13, 1x25	
		texte-B	1027	67	13	2	6	1	1	0	1	0	1x12	
	Chinois	texte-U	1264	120	22	20	8	2	1	0	0	0		
		texte-B	1498	85	13	2	0	0	0	0	0	0		
	Finnois	texte-U	967	93	22	8	3	2	0	0	0	0		
		texte-B	1137	54	8	1	1	0	0	0	0	0		
	Espagnol	texte-U	1014	97	20	8	0	2	1	0	0	0		
		texte-B	1174	47	11	3	0	1	0	0	0	0		
	Nombre de séquences de 4 signes	Glozel	texte-U	1356	6	0	0	0	0	0	0	0	0	
			texte-B	1360	4	0	0	0	0	0	0	0	0	
Basque		texte-U	914	106	32	13	5	1	1	0	0	0		
		texte-B	1178	61	4	0	0	0	0	0	0	0		
Berbère		texte-U	1031	63	14	5	2	0	2	1	0	2		
		texte-B	1195	23	7	1	1	0	0	0	0	0		
Chinois		texte-U	1532	58	13	4	1	0	0	0	0	0		
		texte-B	1672	21	0	0	0	0	0	0	0	0		
Finnois		texte-U	1152	46	9	0	0	1	0	0	0	0		
		texte-B	1244	15	1	0	0	0	0	0	0	0		
Espagnol		texte-U	1176	56	6	3	0	0	0	0	0	0		
		texte-B	1276	18	2	0	0	0	0	0	0	0		
Nombre de séquences de 5 signes		Glozel	texte-U	1363	2	0	0	0	0	0	0	0	0	
			texte-B	1363	2	0	0	0	0	0	0	0	0	
	Basque	texte-U	1014	95	23	7	2	0	0	0	0	0		
		texte-B	1249	31	0	0	0	0	0	0	0	0		
	Berbère	texte-U	1158	34	6	2	0	0	0	1	0	1		
		texte-B	1247	10	1	0	0	0	0	0	0	0		
	Chinois	texte-U	1626	30	7	0	0	0	0	0	0	0		
		texte-B	1699	7	0	0	0	0	0	0	0	0		
	Finnois	texte-U	1215	26	3	0	0	0	0	0	0	0		
		texte-B	1272	2	0	0	0	0	0	0	0	0		
	Espagnol	texte-U	1249	31	2	0	0	0	0	0	0	0		
		texte-B	1307	5	0	0	0	0	0	0	0	0		

(à la page suivante)

TABLE 1 (suite)

			1	2	3	4	5	6	7	8	9	10	Plus de 10
Nombre de séquences de 6 signes	Glozel	texte-U	1364	1	0	0	0	0	0	0	0	0	0
		texte-B	1364	1	0	0	0	0	0	0	0	0	0
	Basque	texte-U	1079	84	14	4	1	0	0	0	0	0	0
		texte-B	1282	14	0	0	0	0	0	0	0	0	0
	Berbère	texte-U	1215	20	2	0	0	0	0	0	1	0	0
		texte-B	1265	2	0	0	0	0	0	0	0	0	0
	Chinois	texte-U	1665	16	3	0	0	0	0	0	0	0	0
		texte-B	1708	2	0	0	0	0	0	0	0	0	0
	Finnois	texte-U	1242	15	1	0	0	0	0	0	0	0	0
		texte-B	1275	0	0	0	0	0	0	0	0	0	0
	Espagnol	texte-U	1286	15	0	0	0	0	0	0	0	0	0
		texte-B	1316	0	0	0	0	0	0	0	0	0	0
Nombre de séquences de 7 signes	Glozel	texte-U	1365	0	0	0	0	0	0	0	0	0	0
		texte-B	1365	0	0	0	0	0	0	0	0	0	0
	Basque	texte-U	1123	74	10	2	0	0	0	0	0	0	0
		texte-B	1293	8	0	0	0	0	0	0	0	0	0
	Berbère	texte-U	1241	12	1	0	0	0	0	0	0	0	0
		texte-B	1268	0	0	0	0	0	0	0	0	0	0
	Chinois	texte-U	1686	8	1	0	0	0	0	0	0	0	0
		texte-B	1711	0	0	0	0	0	0	0	0	0	0
	Finnois	texte-U	1260	7	0	0	0	0	0	0	0	0	0
		texte-B	1274	0	0	0	0	0	0	0	0	0	0
	Espagnol	texte-U	1295	10	0	0	0	0	0	0	0	0	0
		texte-B	1315	0	0	0	0	0	0	0	0	0	0
Nombre de séquences de 8 signes	Glozel	texte-U	1364	0	0	0	0	0	0	0	0	0	0
		texte-B	1364	0	0	0	0	0	0	0	0	0	0
	Basque	texte-U	1161	63	7	0	0	0	0	0	0	0	0
		texte-B	1298	5	0	0	0	0	0	0	0	0	0
	Berbère	texte-U	1253	7	0	0	0	0	0	0	0	0	0
		texte-B	1267	0	0	0	0	0	0	0	0	0	0
	Chinois	texte-U	1696	4	0	0	0	0	0	0	0	0	0
		texte-B	1710	0	0	0	0	0	0	0	0	0	0
	Finnois	texte-U	1267	3	0	0	0	0	0	0	0	0	0
		texte-B	1273	0	0	0	0	0	0	0	0	0	0
	Espagnol	texte-U	1298	8	0	0	0	0	0	0	0	0	0
		texte-B	1314	0	0	0	0	0	0	0	0	0	0

(à la page suivante)

TABLE 1 (suite et fin)

		1	2	3	4	5	6	7	8	9	10	Plus de 10	
Nombre de séquences de 9 signes	Glozel	texte-U	1363	0	0	0	0	0	0	0	0	0	
		texte-B	1363	0	0	0	0	0	0	0	0	0	
	Basque	texte-U	1188	52	5	0	0	0	0	0	0	0	
		texte-B	1299	4	0	0	0	0	0	0	0	0	
	Berbère	texte-U	1260	3	0	0	0	0	0	0	0	0	
		texte-B	1266	0	0	0	0	0	0	0	0	0	
	Chinois	texte-U	1701	1	0	0	0	0	0	0	0	0	
		texte-B	1709	0	0	0	0	0	0	0	0	0	
	Finnois	texte-U	1272	0	0	0	0	0	0	0	0	0	
		texte-B	1272	0	0	0	0	0	0	0	0	0	
	Espagnol	texte-U	1299	7	0	0	0	0	0	0	0	0	
		texte-B	1313	0	0	0	0	0	0	0	0	0	
	Nombre de séquences de 10 signes	Glozel	texte-U	1362	0	0	0	0	0	0	0	0	0
			texte-B	1362	0	0	0	0	0	0	0	0	0
Basque		texte-U	1209	44	3	0	0	0	0	0	0	0	
		texte-B	1300	3	0	0	0	0	0	0	0	0	
Berbère		texte-U	1261	2	0	0	0	0	0	0	0	0	
		texte-B	1265	0	0	0	0	0	0	0	0	0	
Chinois		texte-U	1702	0	0	0	0	0	0	0	0	0	
		texte-B	1709	0	0	0	0	0	0	0	0	0	
Finnois		texte-U	1271	0	0	0	0	0	0	0	0	0	
		texte-B	1271	0	0	0	0	0	0	0	0	0	
Espagnol		texte-U	1300	6	0	0	0	0	0	0	0	0	
		texte-B	1312	0	0	0	0	0	0	0	0	0	

Notes

1) Pour démontrer la façon d'interpréter la Table 1, dans le texte-U basque il y a 755 séquences qui sont d'une longueur de 3 signes et qui n'apparaissent qu'une seule fois, 120 séquences de cette longueur qui apparaissent deux fois, 42 qui apparaissent trois fois, et ainsi de suite.

2) Par suite d'une malfonction du perçoir de cartes, une ligne de 6 signes dans le texte-B chinois fut enregistrée deux fois. Par conséquent la table montre une répétition de trop des séquences de 6 signes, deux répétitions de trop des séquences de 5 signes, trois répétitions de trop des séquences de 4 signes, et quatre répétitions de trop des séquences de 3 signes.

TABLE 2

		Nombre d'occurrences							
		1	2	3	4	5	Plus de 5		
Nombre de séquences de 3 signes	Glozel	texte-U	471	2	0	0	0		
		texte-B	473	1	0	0	0		
	Basque	texte-U	303	36	7	1	0		
		texte-B	368	13	2	0	0		
	Berbère	texte-U	296	30	4	1	1	1x10	
		texte-B	337	16	3	1	1		
	Chinois	texte-U	382	36	2	4	1		
		texte-B	444	14	3	0	0		
	Finnois	texte-U	360	7	0	0	2		
		texte-B	372	6	0	0	0		
	Espagnol	texte-U	364	16	2	1	0		
		texte-B	376	12	2	0	0		
	Nombre de séquences de 4 signes	Glozel	texte-U	474	0	0	0	0	
			texte-B	474	0	0	0	0	
Basque		texte-U	345	21	4	0	0		
		texte-B	389	5	0	0	0		
Berbère		texte-U	340	20	2	0	0		
		texte-B	369	7	1	0	0		
Chinois		texte-U	432	19	2	1	0		
		texte-B	472	4	0	0	0		
Finnois		texte-U	376	1	0	0	1		
		texte-B	383	0	0	0	0		
Espagnol		texte-U	386	8	1	0	0		
		texte-B	395	5	0	0	0		
Nombre de séquences de 5 signes		Glozel	texte-U	473	0	0	0	0	
			texte-B	473	0	0	0	0	
	Basque	texte-U	368	12	2	0	0		
		texte-B	394	2	0	0	0		
	Berbère	texte-U	360	11	1	0	0		
		texte-B	377	4	0	0	0		
	Chinois	texte-U	452	12	1	0	0		
		texte-B	475	2	0	0	0		
	Finnois	texte-U	380	1	0	0	0		
		texte-B	382	0	0	0	0		
	Espagnol	texte-U	398	3	0	0	0		
		texte-B	400	2	0	0	0		

(à la page suivante)

TABLE 2 (suite)

		1	2	3	4	5	Plus de 5	
Nombre de séquences de 6 signes	Glozel	texte-U	472	0	0	0	0	
		texte-B	472	0	0	0	0	
	Basque	texte-U	380	7	1	0	0	
		texte-B	395	1	0	0	0	
	Berbère	texte-U	372	6	0	0	0	
		texte-B	382	1	0	0	0	
	Chinois	texte-U	464	7	0	0	0	
		texte-B	478	0	0	0	0	
	Finnois	texte-U	381	0	0	0	0	
		texte-B	381	0	0	0	0	
	Espagnol	texte-U	403	0	0	0	0	
		texte-B	403	0	0	0	0	
Nombre de séquences de 7 signes	Glozel	texte-U	471	0	0	0	0	
		texte-B	471	0	0	0	0	
	Basque	texte-U	388	4	0	0	0	
		texte-B	396	0	0	0	0	
	Berbère	texte-U	379	2	0	0	0	
		texte-B	383	0	0	0	0	
	Chinois	texte-U	471	3	0	0	0	
		texte-B	477	0	0	0	0	
	Finnois	texte-U	380	0	0	0	0	
		texte-B	380	0	0	0	0	
	Espagnol	texte-U	402	0	0	0	0	
		texte-B	402	0	0	0	0	
Nombre de séquences de 8 signes	Glozel	texte-U	470	0	0	0	0	
		texte-B	470	0	0	0	0	
	Basque	texte-U	391	2	0	0	0	
		texte-B	395	0	0	0	0	
	Berbère	texte-U	380	1	0	0	0	
		texte-B	382	0	0	0	0	
	Chinois	texte-U	474	1	0	0	0	
		texte-B	476	0	0	0	0	
	Finnois	texte-U	379	0	0	0	0	
		texte-B	379	0	0	0	0	
	Espagnol	texte-U	401	0	0	0	0	
		texte-B	401	0	0	0	0	

(à la page suivante)

TABLE 2 (suite et fin)

		1	2	3	4	5	Plus de 5		
Nombre de séquences de 9 signes	Glozel	texte-U	469	0	0	0	0		
		texte-B	469	0	0	0	0		
	Basque	texte-U	392	1	0	0	0		
		texte-B	394	0	0	0	0		
	Berbère	texte-U	381	0	0	0	0		
		texte-B	381	0	0	0	0		
	Chinois	texte-U	475	0	0	0	0		
		texte-B	475	0	0	0	0		
	Finnois	texte-U	378	0	0	0	0		
		texte-B	378	0	0	0	0		
	Espagnol	texte-U	400	0	0	0	0		
		texte-B	400	0	0	0	0		
	Nombre de séquences de 10 signes	Glozel	texte-U	468	0	0	0	0	
			texte-B	468	0	0	0	0	
Basque		texte-U	393	0	0	0	0		
		texte-B	393	0	0	0	0		
Berbère		texte-U	380	0	0	0	0		
		texte-B	380	0	0	0	0		
Chinois		texte-U	474	0	0	0	0		
		texte-B	474	0	0	0	0		
Finnois		texte-U	377	0	0	0	0		
		texte-B	377	0	0	0	0		
Espagnol		texte-U	399	0	0	0	0		
		texte-B	399	0	0	0	0		

Note

Pour l'interprétation de cette table, voir la Table 1, note 1.

2v

Il était évident qu'on ne gagnerait rien en testant des combinaisons plus nombreuses, puisque l'absence presque totale de répétitions dans les sous-textes glozéliens rendrait l'essai non valable. On doit remarquer que même sur des sous-textes de si modeste étendue, l'essai est tout à fait valable pour les cinq langues connues.

**

RÉPÉTITION DES SÉQUENCES.

Mais le nombre très faible de répétitions dans les textes glozéliens exige en tout cas une explication. Quelle sorte de texte pourrait entraîner aussi peu de réitération ? Les cinq corpora auxquels le matériel glozélien fut comparé comprennent une diversité de sujets et de matières, et certainement ces différences de sujet se reflètent en partie dans la statistique. Par exemple, le nombre considérable de répétitions de longues séquences dans le corpus basque peut être attribué à certaines phrases formulaïques, ainsi qu'aux vœux fait lors d'une bénédiction nuptiale ; de son côté, le texte berbère abonde en répétitions d'un type commun aux contes populaires partout dans le monde. Mais les chiffres glozéliens sont beaucoup plus faibles que ceux des autres corpora, même ceux des *textes-B*, où la lecture est 50 % incorrecte. Si le corpus glozélien est vraiment moitié unidirectionnel et moitié boustrophédon, ses chiffres devraient se situer entre ceux des *textes-U* et des *textes-B* des langues connues.

L'on ne peut non plus expliquer cette divergence par la longueur comparée des corpora ni par leur mode de préparation. Par suite d'un faux calcul, le corpus complet chinois était plus long que celui de Glozel (1711 signes contre 1371), mais les quatre autres corpora étaient tous un peu plus courts que celui de Glozel (basque 1315, bergère 1274, finnois 1280, espagnol 1321), tout en révélant beaucoup plus de répétitions. D'ailleurs, lors de l'analyse de 8 tablettes on saisit l'occasion de réduire l'échantillon chinois à une étendue comparable à celle des autres échantillons (Glozel 477, basque 402, berbère 389, chinois 483, finnois 386, espagnol 408), mais la divergence statistique persistait. Quant à l'emploi du syllabaire, l'on pourrait raccourcir la représentation de la plupart des mots d'une façon conventionnelle procédé, utilisé dans le *Linéaire B*, qui permettrait de ne pas transcrire certaines consonnes ; cela réduirait la longueur des séquences répétées, mais on voit difficilement comment étendre ce procédé de façon à rendre les chiffres des langues connues comparables à ceux de Glozel, tout en gardant un texte lisible. Il est particulièrement frappant que *le corpus glozélien ne montre pas une seule répétition de plus de 6 signes*. C'est plutôt ce qui se produirait si les signes avaient été écrits au hasard ou selon *quelque critère non linguistique*. Certainement de telles hypothèses sont beaucoup plus confortées par la statistique que n'est celle de textes dans une langue naturelle.

**

CONCLUSION.

Le corpus examiné ne présente pas les caractéristiques que l'on pourrait attendre de textes dans une langue naturelle, et les méthodes linguistiques ne sont pas capables d'établir l'authenticité du matériel. Il appar-

tient aux représentants d'autres disciplines d'établir la possibilité d'une fonction non linguistique des tablettes, ou bien celle de leur fabrication récente. A cet égard aucune certitude ne peut être espérée sans des recherches plus approfondies avec les techniques de la datation par thermoluminescence ni sans la réouverture, sous une direction experte et impartiale, des fouilles de Glozel.

D^r T.D. CRAWFORD

*Lecturer in Linguistics Faculty of Arts,
University College, Cardiff*