

Contribution informatique à l'analyse phonologique de *u* en latin

Christian TOURATIER

Etant en train de finir un programme en basic pour TRS 80 qui donne automatiquement n'importe quelle forme de n'importe quel verbe latin en affichant successivement la structure morphologique, la séquence phonologique, la réalisation phonétique et la transcription graphique de ladite forme, nous avons bien dû régler le délicat problème du statut phonologique des semi-voyelles.

Le latin note par une lettre *u* (ou *i*) deux sons reconnus comme différents par les grammairiens latins eux-mêmes, à savoir la consonne [w] (ou [j]) et la voyelle [u] (ou [i]). Et tout le problème consiste à savoir si [w] et [u] (ou [j] et [i]) fonctionnent comme deux phonèmes différents ou doivent être considérés comme deux réalisations différentes d'un seul et même phonème. A notre avis, ces deux sons différents ne fondent jamais une opposition distinctive. Il est parfaitement possible de prédire en fonction du contexte leur caractère consonantique ou vocalique, ce qui implique que le locuteur ne choisit pas entre [u] et [w] et que par conséquent la différence entre le caractère consonantique de [w] et le caractère vocalique de [u] n'est pas distinctive en latin.

Avant d'établir la distribution complémentaire entre [w] et [u], il faut récuser la fausse paire minimale

uōlūit "il a voulu"/*uōluīt* "il a fait tourner",

qui correspond à une prononciation [woluwit] en face de [wolwit] et peut s'analyser comme une opposition entre /uolu+ui+t/, parfait en *u* de *uelle*, et /uolu+i+t/, parfait sans *u*

de *uoluere* (cf. Touratier, 1973, 154-155). Il faut donc faire comme si le parfait du verbe *uelle* était écrit avec deux *u*. Ce point précisé, on distinguera les trois positions que la lettre *u* peut avoir dans un mot latin : position initiale (B=1, si B représente la position de *u*), position finale [B=LEN(A\$), c'est-à-dire la longueur du mot, si celui-ci est représenté par A\$] et position intérieure [1<B<LEN(A\$)].

A l'initiale d'un mot, la lettre *u* représente une voyelle, quand elle est suivie d'au moins une consonne (*ubique, ululāre, uncus, unde, uterque...*) ou quand elle est suivie de la marque de quantité, à savoir : dans les transcriptions phoniques, mais * dans notre programme (*ūber, ūnus, ūtor...*); mais elle note toujours un son consonantique, quand elle est suivie d'une voyelle ou d'un second *u* (*uacuus, uēlum, uigilia, uocāre, uulgu...*). Il est intéressant d'observer que tous les mots qui commencent par deux *u* ont une syllabe initiale fermée en [wu] + consonne (*uulnus, uulpes, uulua*) : il n'existe pas de mots commençant par [uw], si [u:w] initial est bien attesté par exemple dans *ūuidus, ūua* et *ūuor*. Dans un programme un test suffira à rendre compte de tout cela : si *u* initial est suivi en B+1 d'une voyelle graphique, il se réalise [w].

A la finale d'un mot, *u* a un traitement encore plus simple : il note toujours une voyelle (*tamdiu, gelu, cornu, ueru...*), sauf si la lettre qui précède (c'est-à-dire celle qui est en B-1) est un *e*, comme dans *seu*, auquel cas *u* correspond au son consonantique [w].

En position intérieure, la situation est plus complexe. Il faut distinguer quatre grands cas différents suivant que premièrement *u* est une voyelle longue (c'est-à-dire si B+1=*), que deuxièmement *u* est suivi d'un autre *u* (c'est-à-dire B+1=*u*), que troisièmement *u* est suivi d'une autre voyelle et enfin que quatrièmement *u* est suivi d'une consonne. Il est nécessaire d'envisager à part le cas de *u* intérieur suivi de *u* et celui de *u* intérieur suivi d'une autre voyelle, puisque la voyelle graphique *u* ne note pas forcément un son vocalique.

Quand *u* intérieur est long, il ne se pose que le problème de son éventuel abrégement devant une autre voyelle. Il est tentant de décrire cet abrégement, au moins quand il s'agit des phonèmes fermés /u:/ et /i:/, comme une réalisation consonantique du segment final de la voyelle longue (cf. Zirin, 1970, pp.74-75), *audio* correspondant à la réalisation [awdijo:] de la

séquence phonématique /audi:o:/. Rien n'empêche alors d'expliquer par cette même variation du phonème /u:/ devant voyelle *arguo* et *arguunt* à partir de /argu:o:/ et /argu:unt/, en postulant un radical verbal avec voyelle longue comme dans le participe *argūtus*, ce qui situerait au niveau phonologique et non pas au niveau morphologique la différence apparente de quantité entre *argūō* et *argūtus*. On peut même penser que *iuvo*, prononcé [juwo:], est la réalisation de /iu:o:/ avec un radical verbal *iu-* comme dans le parfait *iūuī* et le participe *iūtus*, ce qui amènerait à préciser que devant un *u* l'abrégement de /u:/ n'a lieu que si le second *u* est suivi d'une consonne comme dans *arguunt* de /argu:unt/ en face de *iūuī* de /iu:ui:/. La particularité de verbe *iuvo* serait alors purement orthographique : la séquence phonique [uwo] y était transcrite *uuo* et non pas seulement *uo* (comme dans *abluwo*: écrit normalement *abluo*), parce qu'elle était précédée d'un *i* : une suite graphématique *uo* ne pouvant se lire après voyelle que [wo] (cf. *auo* de *auus* ou *lauo*), une graphie **iuo* aurait correspondu à une prononciation [iwo:], comme dans *niuusus* ou *nominatiuo* de *nominatiuus*, et non pas à [juwo:].

Quand *u* intérieur est suivi d'une voyelle graphique autre que *u*, il faut regarder la nature du phonème qui précède le *u* intérieur (c'est-à-dire ce qu'il y a dans la position B-1); et il peut y avoir cinq cas de figure différents. Le premier est celui où B-1 est une sifflante. Deux éventualités peuvent alors se présenter : soit B-1 est aussi première lettre du mot, auquel cas *u* se réalisera obligatoirement sous forme consonantique (cf. *suadeo*, *suavis*, *suauium*, *suesco*, *suetus* où le *u* après *s* initial n'est pas, comme dans *sub*, *summus*, un noyau vocalique); soit B-1 n'est pas la première lettre, mais B-2 est une frontière morphologique que notre programme note par / et qui indique que B-1 est le début du second élément d'un mot composé; le traitement est alors le même qu'après un *s* initial et *u* se réalise *w* (cf. *con/suesco*, *in/suetus*, etc.). Il y a cependant un mot à sifflante initiale dont cette règle de variation ne rend pas compte; c'est l'adjectif possessif *suus*. Il est d'ailleurs le seul, car ni le verbe *suo* ni le nom au génitif *suis* ne font problème, puisqu'ils correspondent de toute évidence à /su:o:/ et /su:is/, ainsi que le montrent nettement le participe *sūtus* et le nominatif *sūs*. Nous proposerons d'analyser de la même façon l'adjectif possessif, en disant que son radical se termine par une voyelle longue et qu'il représente donc une séquence phonématique /su:us/.

Le deuxième cas de figure est celui où *u* intérieur est précédé d'une liquide, c'est-à-dire où B-1 est un *r* ou un *l*. Deux éventualités peuvent alors se présenter : ou la liquide est elle-même précédée d'une voyelle, et alors *u* reçoit forcément une réalisation consonantique (cf. *aluus, seruus, uoluo, soluo, eruum, seruitium*); ou la liquide est initiale ou bien n'est pas précédée d'une voyelle, alors *u* note un phonème vocalique /u/ qui avait toute chance de présenter une variante [uw], *ruo* et *luo* devant très certainement se prononcer [ruwo:] et [luwo:] plutôt que [ruo:] et [luo:] avec hiatus. Cette règle de variation s'applique aussi à *deruo, diruo* ou *eruo*, qui sont des mots composés et où, dans notre programme, la liquide *r* est précédée d'un / pour indiquer précisément la composition. Dans *polluo* par contre, la latérale n'est pas précédée d'une voyelle; *u* note alors un phonème vocalique, et l'on a une prononciation [polluwo:].

Troisième cas de figure : si B-1 est une voyelle, ou la marque * de la quantité, ou la marque / de la composition ou la lettre *q* du premier segment d'une labiovélaire sourde, alors *u* correspond à un son consonantique : par exemple *auus, auarus, auidus, soniuus, quadriuium; uua, primitiuus, ciuis, flauus; conuiuium, aduocare, inuidere; sequor, quattuor*.

Quatrième cas de figure : si B-1 est un *g*, alors *u* correspond à [w], c'est-à-dire au second segment de la labiovélaire sonore, lorsque B-2 est un *n*; sinon *u* note la voyelle /u/, qui bien sûr se réalise [uw] devant une autre voyelle : *lingua, distingo* en face de *indigūi, prodigus*. Le verbe *arguo* n'entre pas dans cette catégorie, la lettre *u* y notant une voyelle longue, comme le montre le participe *argūtus*; et cette voyelle longue présente la variante attendue devant voyelle, à savoir le segment [uw] (cf. ce qui a été dit plus haut sur le *u* intérieur long).

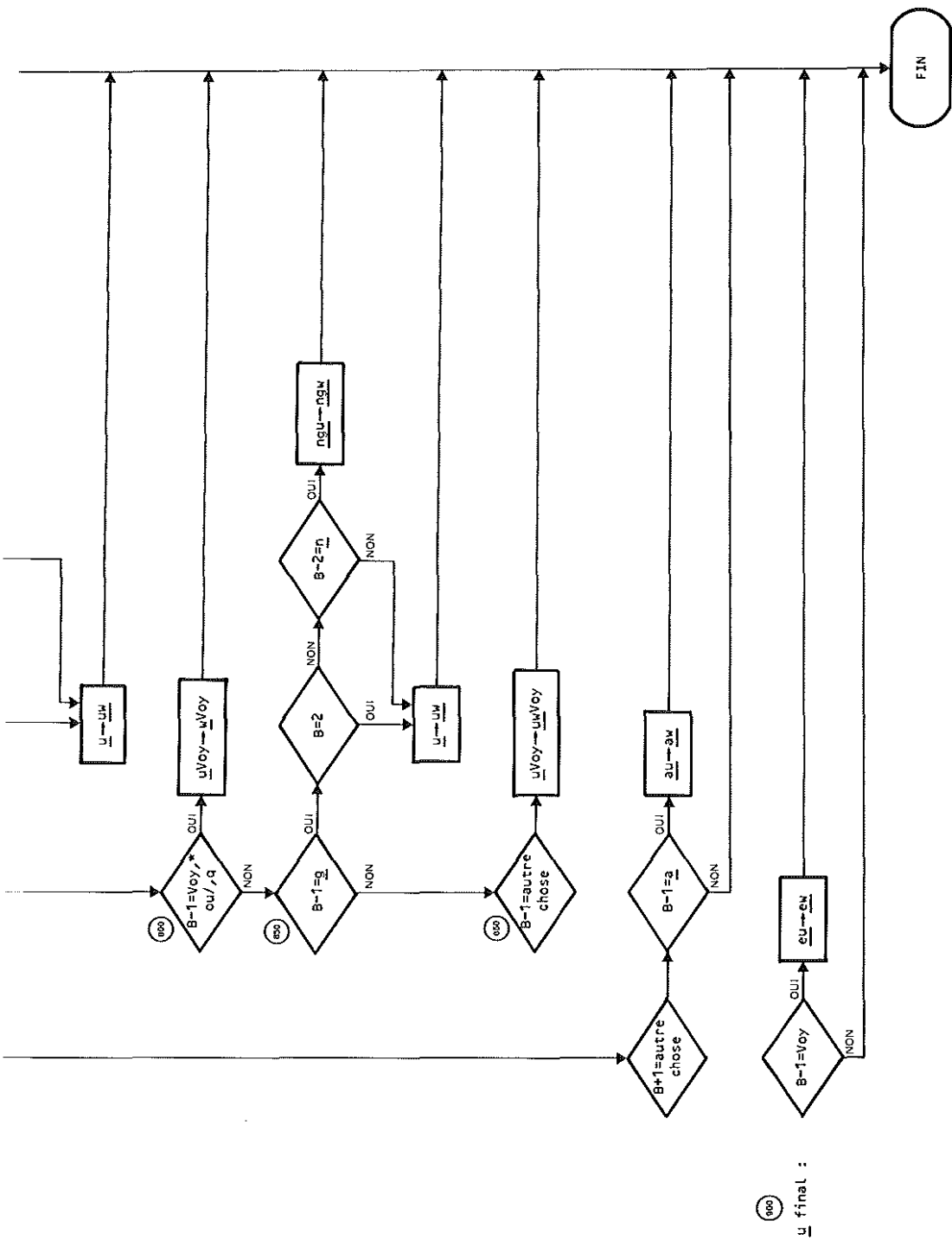
Le dernier cas de figure est celui où *u* suivi d'une voyelle n'est précédé ni d'une sifflante initiale ou pseudo-initiale, ni d'une liquide, ni d'une voyelle, ni des marques * ou /, ni de *q* ou *g*; bref il s'agit de tous les autres contextes antérieurs de *u* suivi d'une voyelle. La lettre *u* note alors un phonème vocalique qui se réalise [uw], puisqu'il est devant voyelle : par exemple *monui, debui, timui, miscui, censui*, dont la finale est traditionnellement reconnue comme se prononçant [uwi:] (cf. Niedermann 1953, p. 23).

Quand *u* intérieur est suivi d'un second *u*, il reçoit le même traitement que devant une autre voyelle, si toutefois le second *u* n'est pas suivi d'une voyelle; car c'est seulement devant une consonne que le second *u* note une voyelle.

Quand enfin *u* intérieur n'est suivi ni d'un *u*, ni d'une autre voyelle, ni d'une marque * de longueur, il note toujours une voyelle, sauf quand il est précédé d'un *a*, car il correspond alors à l'élément consonantique de la diphtongue *au*. Dans *laudo* ou *augeo*, la lettre *u* note un son consonantique; dans *deus*, *pius*, *idoneus* par contre elle note un son vocalique. Le mot *eurus* est une exception qui n'infirmes pas la règle postulée : étant un emprunt au grec, il n'est pas surprenant que ce mot ne se conforme pas entièrement aux règles phonologiques du latin.

Toutes les réalisations phoniques de *u* sont organisées dans l'organigramme présenté ci-après. Nous allons maintenant traduire cet organigramme dans le Basic du micro-ordinateur très grand public qu'est le ZX 81 de Sinclair. Mais nous y ajouterons un programme introducteur qui fera entrer une chaîne de caractères, examinera chaque caractère de cette chaîne, et enverra à l'instruction 100, c'est-à-dire au programme de variation de *u*, qui sera donc traité comme un sous-programme, chaque fois qu'il identifiera un *u* dans cette chaîne, et finalement affichera la réalisation phonétique de cette chaîne, du moins en ce qui concerne les sons notés par *u*. Nous ajouterons également à partir de l'instruction 1000 un sous-programme destiné à faire la même chose que l'instruction INSTR de certains Basic, c'est-à-dire qui permettra de trouver dans une chaîne la présence d'une sous-chaîne donnée et de calculer la position dans la chaîne du début de la sous-chaîne en question. Ce sous-programme sera principalement utilisé pour identifier la nature du contexte antérieur ou postérieur de *u*.

La lecture de l'ensemble du programme sera facilitée par les observations suivantes. Au début du programme l'instruction 10 met dans la variable V\$ la transcription graphique des cinq timbres vocaliques du latin; et les instructions 12 à 18 permettent d'économiser de la mémoire en remplaçant dans la suite du programme les constantes numériques par des variables numériques, ce qui nécessite moins d'octets, mais rend la lecture un peu plus difficile. Au début du sous-programme intitulé VAR(IATION) DE U, l'ordinateur calcule une variable T qui va prendre la valeur 200 si *u* est initial, la valeur 300 s'il



```

10 LET V$="AEIOU"
12 LET Z=1
13 LET Y=2
14 LET X=3
15 LET W=4
17 LET U=100
18 LET V=1000
20 INPUT A$
30 PRINT A$
40 FOR B=Z TO LEN A$
50 IF A$(B)="U" THEN GOSUB U
60 IF B=LEN A$ THEN GOTO 80
70 NEXT B
80 PRINT A$
85 GOTO 20
90 STOP
100 REM VAR DE U*****
110 LET T=Z*(B=Z)+Y*(B>Z AND B<LEN A$)+W*Y*(B=LEN A$)
120 GOTO U+U*T
200 REM U INITIAL.....
202 LET C$=V$
204 LET D$=A$(B+Z)
206 GOSUB V
210 IF D<>0 THEN LET A$(Z)="W"
220 RETURN
300 REM U INTERIEUR.....
302 LET C$="X"+V$
304 LET D$=A$(B+Z)
310 GOSUB V
320 LET T=Z*(D=Z)+Y*(D=W+Y)+X*(D>Z AND D<W+Y)
330 GOTO X*U+U*T+40*(D=0)
340 REM DIPHT AU, EU
350 IF A$(B-Z)="A" THEN LET A$(B)="W"
360 IF B=Y AND A$(B-Z)="E" THEN LET A$(B)="W"
370 RETURN
400 REM U*V
401 IF B=LEN A$-Z THEN RETURN
402 LET C$=V$
404 LET D$=A$(B+Y)
410 GOSUB V
420 LET T=Z*(D=W+Z)+Y*(D>0 AND D<W+Z)
430 GOTO 430+20*T+10*(D=0)
440 RETURN
450 IF B=LEN A$-Y THEN GOTO 470
452 C$=V$
454 LET D$=A$(B+X)
458 GOSUB V
460 IF D<>0 THEN RETURN
470 LET A$(B+Z)="W"
480 RETURN
500 REM UU
510 IF B=LEN A$-Z THEN RETURN
512 LET C$=V$
514 LET D$=A$(B+Y)
518 GOSUB V
520 IF D<>0 THEN RETURN
600 REM UV
602 LET C$="SRL"+V$+"*/QG"
604 LET D$=A$(B-Z)
610 GOSUB V

```



```

620 LET T=Z*(D=Z)+Y*(D=Y OR D=X)+X*(D>X AND D<W*X)+W*(D=W*X)
630 GOTO 650+U/Y*T
650 LET A$=A$(TO B)+"W"+A$(B+Z TO)
660 LET B=B-Z
670 RETURN
700 REM SUV
710 IF B=Z THEN LET A$(B)="W"
715 IF B=Y THEN RETURN
720 IF A$(B-Y)="/" THEN LET A$(B)="W"
730 RETURN
750 REM R,LUV
760 IF B=Y THEN GOTO 650
762 LET C$=V$
764 LET D$=A$(B-Y)
768 GOSUB V
770 IF D=0 THEN GOTO 650
800 REM V,/ ,QUV
810 LET A$(B)="W"
820 RETURN
850 REM GUV
860 IF B=Y THEN GOTO 650
870 IF A$(B-Y)="N" THEN LET A$(B)="W"
880 IF A$(B-Y)<>"N" THEN GOTO 650
890 RETURN
900 REM U FINAL.....
902 LET C$=V$
904 LET D$=A$(B-Z)
908 GOSUB V
910 IF D<>0 THEN GOTO 810
920 RETURN
1000 REM INSTR(C$,D$)*****
1005 FOR D=Z TO LEN C$
1010 IF D$=C$(D) THEN RETURN
1020 NEXT D
1030 LET D=0
1040 RETURN

```

est intérieur ou la valeur 900 s'il est final, et par un GOTO calculé passe à l'instruction qui correspond à la valeur de T, laquelle représente le début de l'une des trois principales parties du traitement de la variation de u. En 200 et en 900 l'ordinateur utilise le sous-programme INSTR pour savoir si la lettre qui suit u initial ou celle qui précède u final se trouve dans la chaîne V\$ et est donc une voyelle. A partir de l'instruction 300, un appel au sous-programme INSTR et un GOTO calculé envoient l'ordinateur en 400 si u intérieur est suivi de la marque de longueur, en 500 s'il est suivi d'un second u, en 600 s'il est suivi d'une autre voyelle, ou en 340 dans les autres contextes. A partir de l'instruction 600 l'ordinateur traite u intérieur suivi d'une voyelle après avoir calculé l'adresse dans le programme d'un des cinq cas de figure possibles, à savoir 700 si u est précédé d'une sifflante, 750 s'il est précédé d'une

liquide, 800 s'il est précédé d'une voyelle, de la marque de longueur, de la marque de composition ou d'un *q*, 850 s'il est précédé d'un *g* et 650 dans tous les autres cas.

Si l'on n'estime pas intéressant en soi de traiter un problème de linguistique, et qui plus est de linguistique latine, à l'aide d'un outil technologique qui révolutionne la société contemporaine et marque en quelque sorte l'entrée dans une ère nouvelle, on se demandera quel est pour le linguiste l'intérêt d'un programme comme celui qui vient d'être esquissé. Il est possible de répondre que l'ordinateur permet de tester une partie de la description phonologique du latin, en prouvant pour ainsi dire empiriquement que le trait consonantique de [w] est mécaniquement prédictible et donc que les latins ne choisissaient pas entre la consonne [w] et la voyelle [u]. Cela fait avancer la théorie phonologique, qui, après cela, ne peut plus nier que [w] et [u] sont en latin des variantes d'un même phonème. L'ordinateur permet aussi de gérer sans erreur les données phonologiques quand celles-ci sont trop nombreuses et complexes : la complémentarité entre les différentes distributions de [w] et [u] par exemple est loin d'être d'emblée évidente; et quand on a réussi à répertorier la plupart ou la totalité de ces distributions, il est assez difficile, voire impossible de savoir s'il y a ou non compatibilité entre les différentes règles de variation qui semblent rendre compte de ces distributions. Avec l'ordinateur il devient presque facile de dominer l'enchevêtrement de toutes ces règles, car la moindre contradiction ou omission donnera des résultats en désaccord avec les données. Et si l'on veut faire entrer ces règles dans l'ensemble des règles phonologiques du latin et gérer correctement tout le système phonologique de cette langue, il est quasiment impossible de le faire simplement sur le papier. Le recours à l'ordinateur est alors le seul salut.

L'ordinateur peut enfin faciliter la recherche de nouvelles explications. C'est ainsi qu'en ajoutant de 530 à 560 des instructions qui remplacent le groupe *uu* entre consonnes par une voyelle longue, ou de 352 à 355 des instructions qui remplacent *wu* par un *u* long, on constate que ces modifications rendraient compte de *solūtus* ou *locūtus*, mais produiraient **solūnt*, **eqūs*, **aūnculus* et **loqūntur* au lieu de [solwunt], [ekwus], [awunkulus] et [lokwuntur]. Mais ces erreurs permettent de voir que les mots où elles se produisent présentent une particularité que ne connaissent pas les mots non fautifs : soit les deux *u* qui ne doivent pas être remplacés par *u* long ne sont pas devant une

frontière de morphème comme dans *auuncul+us*, soit les deux *u* sont séparés par la dernière frontière de morphème du mot comme dans *equ+us*, *loqu+untur* ou *solu+unt*. On entrevoit alors que le paradigme des désinences casuelles pour les noms ou des désinences personnelles pour les verbes doit être ce qui empêche la modification du radical qu'entraînerait la réalisation [u:] des deux phonèmes /uu/. Et il devient possible de dire que le participe passé des lexèmes verbaux qui du fait de la règle morphologique de thématization (cf. Touratier, 1972, pp.148-154) présente une variante en *i*, connaît une thématization en *u* lorsque le radical verbal se termine par un *u* et que cette thématization n'intervient pas au voisinage de la dernière frontière de morphème de la forme verbale, d'où /lokuu+t+us/ et /soluu+t+us/ en face par exemple de /lokui+tur/, /solui+t/ ou /loku+untur/, /solu+unt/. Cela posé, il n'y a plus qu'à admettre que deux phonèmes /uu/ présentent une variante [u:] après consonne et devant une frontière de morphème, ce qui expliquera *solūtus* ou *locūtus* en face de *soluunt* ou *loquuntur*. Pour que notre programme obtienne ces résultats, il faut lui ajouter les instructions suivantes

```
525 REM (C)UU+(C)
530 IF A$(B+2)<>"+" THEN GOTO 600
540 LET A$(B+1)="*"
550 LET A$=A$(TO B+1)+A$(B+3 TO)
560 RETURN
```

L'instruction 550 n'ayant pour but que de supprimer la frontière de morphème qui suit les deux phonèmes /u/. Dans un programme plus vaste qui traiterait tout le système phonologique, il n'y aurait pas à faire une instruction *ad hoc* comme celle-ci : il faudrait soit ne supprimer les frontières de morphème qu'après le traitement de *u*, ce qui obligerait à modifier légèrement notre programme afin que le *u* de *au+is* par exemple reçoive le traitement de *u* intervocalique malgré la présence de la frontière de morphème, soit plus simplement peut-être traiter la variation de /uu/ en [u:] avant la suppression des frontières de morphème, qui dans notre programme de conjugaison latine intervient après le rhotacisme synchronique et juste avant la variation de /u/.

L'ordinateur permet de constater très rapidement que cette description phonologique de *locūtus* ou *solūtus* n'est pas entièrement satisfaisante; car un mot comme *argūtus* pourrait correspondre à deux séquences phonologiques différentes et être la réalisation phonétique soit de /argu:t+us/ avec un lexème verbal à voyelle finale longue soit de /arguu+t+us/ avec

un lexème verbal à voyelle finale brève, une thématization en *u* et la variation de /uu/ en [u:]. Pour éviter une telle ambiguïté phonématique, il suffit de limiter cette variation aux cas où les deux *u* se trouvent après un *q* ou une liquide, c'est-à-dire aux cas où le premier *u* n'est pas syllabique et devrait donc recevoir une réalisation consonantique. Si l'on ajoute l'instruction

```
535 IF A$(B-1)<>"q" AND A$(B-1)<>"l" AND A$(B-1)<>"r" THEN GOTO 600
```

notre programme ne pourra plus attribuer la réalisation phonétique [argu:tus] qu'à la séquence phonématique /argu:tt+us/.

Si un programme de phonologie comme celui que nous proposons à l'intérêt de pouvoir être testé sur n'importe quel mot latin contenant un /u/, il serait intéressant de le tester systématiquement sur tous les mots d'un ou de plusieurs textes latins suivis. Une entreprise aussi fastidieuse et démesurée devient, grâce aux progrès de l'informatique, parfaitement envisageable. Il suffit pour cela de disposer, sous forme de fichiers ou même d'une base de données, de certains des textes latins que le L.A.S.L.A. de Louis Delatte a eu l'heureuse idée de traiter informatiquement. On remplacera alors notre instruction 20 d'entrée des données par une procédure de lecture de fichiers, qui permettra de lire un à un les mots du texte latin mis en fichier, mais ne retiendra, pour les soumettre à notre programme, que les mots contenant au moins un *u*. Après avoir fait tourné un tel programme pendant tout le temps nécessaire, on obtiendra sur l'imprimante un listing qui donnera tous les mots retenus, ainsi que leur prononciation. Et il sera très facile de voir si les règles phonologiques ici proposées décrivent correctement la prononciation de toutes les occurrences du phonème /u/. Et en cas d'erreurs, il ne sera pas trop difficile de corriger notre description phonologique; car la cause de chaque erreur pourra être localisée avec précision dans l'ensemble des règles qui concernent les variations du phonème /u/, ou même dans l'ensemble des règles du système phonologique, si comme dans notre programme de conjugaison on a essayé de traiter informatiquement presque tous les problèmes phonologiques et graphématiques du latin classique.

Références

Niedermann, M., (1953), *Précis de phonétique historique du latin*, Paris, Klincksieck, 3e éd., 208 pp.

- Touratier, C., (1972), "Morphophonologie du verbe latin" in *B.S.L.*, 67.1, pp. 139-174.
- Touratier, C., (1973), Compte-rendu de Sandor Kiss, *Les transformations de la structure syllabique en latin tardif*, Debrecen, 1972, in *B.S.L.*, 68.2, pp. 151-156.
- Zirin, R.A., (1970), *The phonological basis of Latin prosody*, The Hague, Mouton (Janua linguarum, Series practica 99), 91 pp.