

<b>Zeitschrift:</b>	Études de Lettres : revue de la Faculté des lettres de l'Université de Lausanne
<b>Herausgeber:</b>	Université de Lausanne, Faculté des lettres
<b>Band:</b>	- (2004)
<b>Heft:</b>	1-2
<b>Artikel:</b>	César, parle-moi! : La synthèse de la parole appliquée au latin classique
<b>Autor:</b>	Bianchi, Olivier
<b>DOI:</b>	<a href="https://doi.org/10.5169/seals-870151">https://doi.org/10.5169/seals-870151</a>

### Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

### Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

### Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

**Download PDF:** 09.02.2026

**ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>**

# CÉSAR, PARLE-MOI ! LA SYNTHÈSE DE LA PAROLE APPLIQUÉE AU LATIN CLASSIQUE

La langue latine n'est parvenue jusqu'à nous que sous sa forme écrite. Son oralité ne doit cependant pas être mésestimée : elle offre une dimension supplémentaire à l'étude des œuvres antiques. Mais si le locuteur humain ne peut s'affranchir complètement, en tentant de restituer la prononciation du latin, des habitudes phonologiques et prosodiques propres à sa langue maternelle, la synthèse de la parole (science de simulation informatique) permet un contrôle total sur chacun des éléments de la chaîne parlée. Dans cette présentation générale de la synthèse de la parole appliquée au latin classique, nous rappelons quelles sont les sources qui nous renseignent sur sa prononciation et proposons des solutions pour pallier l'absence de locuteurs natifs et les limitations technologiques actuelles. Les étapes de cette reconstruction sont décrites en accordant une importance particulière aux composantes phonologique et prosodique.

## *1. Introduction*

Le latin peut-il ressusciter ? Cette question — à priori une simple boutade faisant écho à l'expression fâcheuse « langue morte » pour désigner la langue latine — nous conduit en fait à nous demander si nos connaissances de la langue latine suffisent à rendre compte de ce qu'a été autrefois sa dimension orale. Car il convient de ne pas oublier que le latin a été une langue parlée avant que de posséder une littérature, et que la plupart des textes qui nous sont parvenus ont été un jour lus à haute voix, quand ils n'ont pas même été conçus spécialement dans cette intention. N'est-ce pas perdre un peu de la saveur d'une pièce de Molière ou d'un discours de Churchill que de les lire sans les entendre ? La

force d'un discours de Cicéron, la beauté d'une ode d'Horace, si elles s'expriment déjà dans les mots, ne seraient-elles pas exaltées par leurs sonorités ?

La prononciation est un aspect du latin que l'étudiant aborde rarement au cours de ses études. Lorsqu'il doit « dire » des textes, sa prononciation est déterminée par la tradition linguistique attachée à sa langue maternelle : francophones, italophones, germanophones ou anglophones prononceront différemment un mot comme *Cicero* par exemple. Cependant, certains professeurs de littérature classique, aux États-unis principalement<sup>1</sup>, enseignent la prononciation du latin selon les critères établis par les linguistes et les phonéticiens<sup>2</sup> qui, depuis plus d'un siècle, ont recensé, analysé, classé les témoignages, directs ou indirects, des Anciens sur la prononciation de leur langue. Plusieurs groupements et associations ont ainsi vu le jour, notamment la *Society for the Oral Reading of Greek and Latin Literature (SORGLL)*<sup>3</sup>, sous la direction de M. Dillon (Loyola Marymount University), K. Volk (Columbia University), S. G. Daitz (City College of New York) et R. P. Sonkowsky (University of Minnesota), ainsi que la *Societas Latina*<sup>4</sup>, sous la direction de S. Albert et C. Eichenseer (Universität des Saarlandes, Saarbrücken). La prononciation dite « restituée », encouragée par ces deux organismes, vise à promouvoir l'oralité des langues classiques afin de profiter pleinement de la beauté des œuvres antiques.

Pour toute langue, la prononciation restituée veut être la plus proche possible de ce qu'a pu être la prononciation d'une langue cible à une époque cible (typiquement : le latin au I<sup>er</sup> siècle avant J.-C.). L'hébreu est l'exemple le plus célèbre d'une prononciation restituée. La fin du XIX<sup>e</sup> siècle voit la renaissance du peuple juif au travers de la restitution de l'hébreu qui, après mille six cents ans d'existence seulement écrite, retrouve sa dimension orale. Avec une grammaire basée sur celle de la Bible et une prononciation sépharade (que les pères fondateurs de l'hébreu restitué

<sup>1</sup> En voici une liste non exhaustive : W. Clausen, K. Coleman, R. Tarrant, R. Thomas (Harvard); S. G. Daitz (New York); C. Higbie (Buffalo); W. Harris (Middlebury); V. Nedeljkovic (Belgrade); R. P. Sonkowsky (Minnesota); W. Stevenson (Richmond); W. Stroh (Munich).

<sup>2</sup> Parmi les plus connus, il convient de citer W. S. Allen, C. Juret, R. G. Kent, C. W. E. Miller, M. Niedermann, E. Pulgram, E. H. Sturtevant.

<sup>3</sup> [www.rhapsodoioralgreekandlatin.org](http://www.rhapsodoioralgreekandlatin.org).

<sup>4</sup> [www.uni-saarland.de/fak5/stockmann/voxlatina](http://www.uni-saarland.de/fak5/stockmann/voxlatina).

considéraient comme étant la plus proche de la prononciation originale), cette nouvelle langue, ou plutôt cette langue ancienne retrouvée, n'en est pas pour autant strictement conforme à ce que les linguistes de l'hébreu peuvent reconstruire de la prononciation originale. On déplore entre autres un appauvrissement phonologique entraînant des confusions sémantiques. Cet appauvrissement est dû notamment aux origines linguistiques très diverses des personnes souhaitant adopter cette langue dont plusieurs phonèmes leur sont étrangers, phonèmes qui sont alors déformés ou simplement omis dans la prononciation<sup>5</sup>.

Les habitudes linguistiques des locuteurs constituent un paramètre qu'il faut absolument prendre en considération lors de toute tentative de «restitution» d'une prononciation. Si la prononciation restituée du latin n'a pas — encore ?<sup>6</sup> — la même vocation que celle de l'hébreu (la communication orale en latin, telle qu'on peut l'apprécier lors des séances des organismes cités plus haut, n'a pour objectif que de faire découvrir la beauté d'une langue et d'en apprécier une facette longtemps laissée de côté), elle est soumise aux mêmes contraintes et restrictions, sur les plans phonologique et prosodique, que la restitution de l'hébreu parlé. Ainsi, toute personne tentant de prononcer le latin selon les indications dégagées par les spécialistes doit faire abstraction des habitudes linguistiques propres à sa langue maternelle. Cet exercice est rendu particulièrement difficile par le fait qu'il n'existe pas de locuteur natif du latin. En effet, les erreurs (ou les écarts par rapport à la reconstruction théorique) sont difficiles à détecter par un auditeur dont l'oreille n'est pas exercée de manière native à la réception de cette parole. À ces difficultés d'ordre phonologique s'ajoute le fait qu'un locuteur latin doit tenir compte d'un certain nombre de paramètres concernant l'accentuation, les élisions, le type accentuel, le rythme, l'intonation, etc. Intégrer l'ensemble de ces paramètres et les restituer sous la forme d'une parole fluide demande un effort particulièrement soutenu de concentration, effort qui ne peut être maintenu avec la même constance tout au

<sup>5</sup> On trouvera une information plus complète chez Yossi COHEN, «L'Hébreu parlé et l'hébreu écrit: une seule langue?», *Hamoré*, 156 (1999), p. 18-21.

<sup>6</sup> Voir à ce propos les remarques de Christophe SCHMIDT, «Le Latin : langue officielle de l'Union européenne?», *Pharos* (Journal de l'Association Antiquité Vivante, Lausanne), 16 (2003), p. 4-9.

long d'une récitation. D'autre part, la difficulté inhérente à cet exercice empêche la prise en compte des variantes théoriques qui apparaissent dans les études faites par les spécialistes de la langue latine. Car si un consensus général se dégage sur un grand nombre de points, quelques aspects de la langue sont encore sujets à la controverse, ou alors peu documentés. Relevons par exemple la question de l'ictus métrique, de sa réalisation, de sa cohabitation avec — ou de sa superposition à — l'accent verbal. Si la réponse qu'apporte Sandro Boldrini, à savoir que les vers latins se lisent comme se lit la prose<sup>7</sup>, semble aujourd'hui avoir valeur de référence, elle pose la question implicite de savoir comment se prononçait exactement la prose. Et comment, alors, les accents oratoires, d'insistance ou affectifs — qui, comme le rappelle justement C. Van den Bergh, ne relèvent pas de la structure métrique du vers, mais de son interprétation subjective<sup>8</sup> — doivent-ils être appréhendés ? La réalisation orale des vers latins demande une interprétation rythmique qui ne saurait se réduire à une étude structurelle de ces vers. Une personne désirant entreprendre la récitation de vers latins donnera sa propre interprétation rythmique, mais ne sera pas en mesure d'en changer facilement si les critères sur lesquels elle se base pour établir son interprétation doivent être modifiés.

## *2. La synthèse de la parole, un outil scientifique pour le latin*

Les spécialistes du langage et de la parole ont compris tout le parti qu'ils pouvaient tirer de l'utilisation de l'informatique pour mieux comprendre la parole humaine. Le matériel linguistique et phonétique (énoncés, mots, phonèmes, ensemble de règles) se prête particulièrement bien au traitement informatique. Non seulement il bénéficie largement de la puissance toujours accrue des

<sup>7</sup> Sandro BOLDRINI, *La Prosodia et la metrica dei Romani*, Roma : La Nuova Italia Scientifica, 1992, p. 36 : «I Latini leggevano i versi esattamente come la prosa».

<sup>8</sup> Carla VAN DEN BERGH, « Vers une théorie du vers », *Acta fabula. Revue en ligne des parutions en théorie littéraire*, site internet *Fabula. Théories de la fiction littéraire*, juillet 2001, URL (13.01.2004) : [www.fabula.org/revue/cr/116.php](http://www.fabula.org/revue/cr/116.php). C. Van den Bergh y propose une étude complète des contributions et des discussions qui ont donné lieu à la publication *Le Vers français. Histoire, théorie, esthétique. Actes du colloque tenu en 1996 à l'Université de Paris-Sorbonne (Paris-IV)*, textes réunis par M. Murat, Paris : Honoré Champion, 2000, 412 p.

méthodes statistiques (les résultats des analyses peuvent être affinés, comparés, mis en perspective de manière plus productive), mais il est de nature à être « modélisé ». La modélisation, c'est-à-dire la création de modèles basés sur des analyses et des hypothèses, est une science qui émerge dans la seconde moitié du XX<sup>e</sup> siècle. Grâce à la modélisation, il devient possible de vérifier, sur un grand nombre d'énoncés, la pertinence d'un ensemble de traits linguistiques et phonétiques<sup>9</sup>. Le développement de cette nouvelle approche scientifique, ainsi que l'évolution technologique qui l'accompagne, permet à un nouveau domaine des sciences de simulation de voir le jour : la synthèse de la parole<sup>10</sup>, qui consiste en la reproduction sonore d'un texte écrit, codé, contenant essentiellement des informations phonétiques, temporelles (débit), d'intensité et d'intonation.

La possibilité de disposer d'une « machine capable de parler » ouvre, d'un point de vue commercial, des horizons nouveaux que d'aucuns espèrent lucratifs. Cependant, si beaucoup de projets en synthèse de la parole ont vu le jour dans le but d'atteindre des objectifs avant tout commerciaux<sup>11</sup>, un certain nombre de facteurs font que la synthèse n'est plus convoitée aujourd'hui uniquement comme source potentielle de profit, mais qu'elle devient un véritable outil scientifique : on pense notamment aux difficultés financières rencontrées par les services de télécommunications, à la gratuité de systèmes de synthèse développés par des collectivités publiques largement aussi performants que les systèmes payants développés par les entreprises, et enfin à la difficulté d'obtenir des synthèses d'une qualité qui justifie, d'un point de vue commercial, l'investissement consenti.

<sup>9</sup> « La modélisation est considérée comme un outil complémentaire à l'analyse, en ce sens qu'elle vérifie la justification ainsi que la qualité de l'analyse originale. » (Eric KELLER, « Les Théories de la parole dans l'épreuve de la synthèse », in *Les Défis actuels en synthèse de la parole*, éd. E. Keller et B. Zellner, *Études de Lettres* 1997/3, p. 11).

<sup>10</sup> Pour un état des lieux complet de la question, voir *Les Défis actuels en synthèse de la parole*, et plus particulièrement pour la genèse des systèmes de synthèse actuels : E. Keller, « Les théories de la parole », p. 9-27.

<sup>11</sup> On peut s'en convaincre en songeant notamment 1° à l'implication des grandes compagnies de télécommunications (France Télécom, AT&T, British Telecom, Telia (Suède), Telefonica (Espagne), Swisscom, etc.); 2° à la création d'un nombre important de compagnies commerciales indépendantes ou semi-indépendantes (*Élan* à Toulouse, *Babel* à Mons, *SVox* à Zurich, etc.); 3° à l'institutionnalisation de cours en « sciences du langage »

En tant qu'outil scientifique destiné aux sciences de la parole, la synthèse possède des atouts qui la rendent de plus en plus indispensable : sa capacité à traiter un très grand nombre de règles autorise des expérimentations sur des systèmes complexes tout en garantissant la mise en application correcte de chacune des règles ; la flexibilité de sa conception permet de vérifier individuellement l'adéquation de chacune des règles à l'ensemble du système ; elle reste subordonnée à l'oreille humaine, juge impitoyable qui détecte toute anomalie<sup>12</sup>.

Un tel système permet non seulement de mettre en lumière, de manière empirique, les mécanismes complexes de la parole et de les confronter à une théorie, mais aussi, étant donné qu'un locuteur humain n'a pas, de manière permanente et volontaire, le contrôle total sur chacun des éléments constitutifs de la parole, de bénéficier d'un outil scientifique permettant de s'affranchir de cet aspect aléatoire en proposant un système dans lequel chaque composante peut être contrôlée et maîtrisée.

La synthèse de la parole, en permettant le contrôle permanent de tous les éléments constitutifs de la chaîne parlée, offre ainsi au chercheur un outil puissant pour l'étude de la langue latine. Alors que nous sommes contraints de manière rédhibitoire par nos habitudes linguistiques et dans l'incapacité d'avoir un contrôle complet sur tous les paramètres nécessaires à la restitution d'une langue dont il n'existe pas de locuteur natif depuis plusieurs siècles, la synthèse de la parole appliquée au latin permet de réduire de manière significative les risques d'erreur et d'approximation, et ainsi de pouvoir tirer le meilleur parti des analyses phonétiques, linguistiques et métriques effectuées depuis un siècle : grâce à la modélisation, il est possible d'intégrer dans un système de synthèse les nombreuses règles (d'articulation, d'élosion, d'accentuation, etc.) issues de l'analyse théorique, et même de confronter celles qui peuvent être concurrentielles voire contradictoires.

dans les Écoles polytechniques fédérales ; 4° aux besoins commerciaux pour la création d'annonces publiques, téléphoniques, pour l'internet, etc.

<sup>12</sup> Dans la mesure où la synthèse de la parole se veut aussi proche que possible de la parole humaine, la sensibilité extrême de notre oreille représente à priori un inconvénient : elle finira immanquablement par distinguer le locuteur humain de la machine. Pour le chercheur, cette sensibilité constitue toutefois le garant indispensable de la qualité de la synthèse : elle lui permet de déceler la moindre faille du système et de la corriger.

### *3. Quel locuteur latin ?*

Les systèmes de synthèse pour le français, l'allemand, l'anglais et les autres langues modernes ont pu voir le jour grâce à l'existence de locuteurs natifs pour ces langues. Des enregistrements ont été effectués, et leur analyse a permis d'en dégager les éléments constitutifs de la parole. Ces éléments ont été ensuite modélisés de façon à pouvoir être intégrés dans un système de synthèse. Les résultats sonores obtenus grâce à ces synthèses ont été soumis pour vérification et validation à des auditeurs dont la langue testée était la langue maternelle. Les imprécisions, voire les erreurs, ont pu être repérées et les corrections effectuées, de manière à approcher au plus près la qualité de la parole humaine, dans les limites imposées par la technologie actuelle.

Cette méthode expérimentale ne peut pas s'appliquer de manière analogue au latin, étant entendu que nous ne pouvons procéder à des enregistrements d'énoncés produits par des locuteurs natifs. Si cet état de fait semble compromettre gravement la création de la synthèse d'une langue entièrement fidèle à son modèle parlé il y a maintenant deux millénaires, il est cependant possible, pensons-nous, de nous en approcher en contournant l'obstacle, c'est-à-dire en cherchant ailleurs que dans la production sonore d'un locuteur natif les éléments indispensables à la création d'une telle synthèse.

Au reste, pouvons-nous affirmer, quand bien même nous disposerions d'un locuteur natif, que nous serions en mesure de capter l'essence même de ce que nous appelons le latin classique ? La synthèse du français, de l'allemand ou de l'anglais rend-elle compte d'un état de la langue existant ? Les locuteurs natifs de ces trois langues sont-ils réellement représentatifs de leur langue respective ? D'un Marseillais, d'un Lillois, d'un Parisien ou d'un Vaudois, lequel parle un français que l'on pourrait qualifier de « standard » ? Car la synthèse de la parole, en l'état actuel de la technologie du moins, nous force à faire des choix. Si le latin ne nous est plus connu que sous sa forme écrite, il faut se souvenir que c'est une langue qui a été parlée jadis par des millions d'êtres humains. Cette réalité prend toute son importance en simulation de la parole. Comme toute langue parlée, le latin a subi de nombreux changements et a connu de nombreuses variantes tant sur un plan synchronique que diachronique, aussi bien au niveau phonétique que sur le plan du lexique et de l'accentuation. Non seulement la langue latine s'est propagée jusque dans des territoires

très éloignés de Rome et a par ce fait subi l'influence, au moins localement, des parlers indigènes, mais dans la Ville même, les disparités au sein de la population en matière d'éducation ainsi que d'un point de vue culturel et social avaient pour conséquence une inégalité notoire dans la connaissance et la pratique de la langue.

Dans ces conditions, quel latin pouvons-nous espérer reconstruire ? Quelles possibilités nous laissent entrevoir les limitations auxquelles nous devons faire face, tant au point de vue technologique qu'au niveau de nos connaissances de la langue ?

### *3.1. Les sources de notre connaissance de la prononciation latine*

C'est à n'en pas douter pour la période de la fin de la République que nos sources sont à la fois les plus précises et les plus abondantes. Avec Cicéron, le latin atteint l'apogée de ce que nous appelons le « classicisme ». C'est ce latin « classique », celui de l'aristocratie romaine, que nous nous sommes fixé comme objectif de simuler. Examinons en premier lieu les sources de notre connaissance pour la prononciation du latin de cette époque :

a) Nous sommes renseignés pour une part par les auteurs anciens eux-mêmes : réfléchissant sur leur propre langue, ils ont pris soin d'en noter les particularités, les exceptions à partir desquelles on déduit les règles générales. À l'époque classique, c'est l'*Orateur* de Cicéron qui constitue notre principale source d'informations. Mais paradoxalement, c'est chez des auteurs plus tardifs que les renseignements sur la prononciation du latin classique sont les plus nombreux : conscient de l'évolution de la langue et soucieux de remettre à l'ordre du jour la « pureté » du style cicéronien, Quintilien, dans son *Institution oratoire*, met le doigt sur les erreurs de langage et de prononciation de ses contemporains (1.5.33) :

Donc, l'absence de tous les défauts dont nous venons de parler constituera ce qui s'appelle en grec l'όρθοέπεια (l'orthoépie), c'est-à-dire une émission exacte et agréable des sons : car c'est bien ainsi que la prononciation peut être entendue comme correcte. (Trad. J. Cousin)

De nombreux traits phonologiques peuvent être ainsi repérés à la lecture attentive de l'œuvre de Quintilien<sup>13</sup>. Plus tard (IV<sup>e</sup>-V<sup>e</sup>

---

<sup>13</sup> On citera pour exemple la remarque de Quintilien concernant la mode d'affubler d'une aspiration des occlusives sourdes qui en sont dépourvues

siècles), d'autres grammairiens s'efforceront à leur tour de dénoncer les erreurs et imprécisions qui émaillent la prononciation. Lorsque plusieurs variantes sont en concurrence, ils prennent généralement parti pour l'une d'entre elles, comme dans cet exemple tiré du traité *De l'orthographe* de Vélius Longus (*Grammatici Latini VII* 65.11-14 Keil):

Quant à la préposition *per* en composition, elle se maintient intacte dans tous les cas, sauf quand elle entre en contact avec la lettre *l*, consonne voisine, que l'on préfère géminer, dans un style soigné, plutôt que de prononcer la lettre *r*, comme quand on préfère dire *pellabor* que *perlabor*.

b) Les jeux de mots ou, plus généralement, de langage nous donnent quelques indices intéressants concernant la prononciation. Ainsi, un passage bien connu de Cicéron vient confirmer l'hypothèse selon laquelle la semi-voyelle *u* avait, à l'époque classique, une articulation vélaire (*De la divination* 2.84):

Alors que Marcus Crassus embarquait son armée à Brindes, un marchand qui vendait sur le port des figues amenées de Caunos criait: *Cauneas!* Nous dirions, si vous le voulez bien, que Crassus avait été averti par ce marchand de ne pas embarquer; il n'aurait pas péri, en effet, s'il avait obéi à cet augure.

Le jeu de mot entre *Cauneas* (« de Caunos ») et *caue ne eas* (« Prends garde ! ne pars pas ») ne peut s'expliquer que si l'on postule une homophonie presque parfaite entre les deux expressions, laquelle n'est possible que si l'on attribue une valeur de semi-voyelle (/w/) au *u* de *caue*.

Suétone, quant à lui, rapporte une anecdote qui vient accréditer la thèse selon laquelle, dans un langage soigné, la prononciation en diphongue du digramme *au* a perduré, alors que, dans la langue populaire, l'évolution de la diphongue s'est faite en *o* long (*Vespasien* 22):

étymologiquement (*Institution oratoire* 1.5.20): « Les anciens usaient très sobrement de l'aspiration [...]. Puis on a longtemps veillé à ne pas aspirer les consonnes comme dans *Gracci* et *triumpi*. Pendant un court laps de temps a surgi l'usage abusif <inverse>, si bien qu'on lit sur des inscriptions *choronae*, *chenturiones*, *praechones*, et il y a, sur ce sujet, une épigramme bien connue de Catulle » (trad. J. Cousin). Quintilien fait ici référence aux deux premiers vers de l'épigramme 84 de Catulle: *Chommoda dicebat, si quando commoda uellet/ dicere, et insidias Arrius hinsidias*, « “Havantages”, disait Arrius, quand il voulait dire “avantages”, et “hembûches”, au lieu d’“embûches”» (trad. G. Lafaye).

Le consulaire Mestrius Florus lui ayant fait observer [sc. à Vespasien] qu'il fallait prononcer *plausta* plutôt que *plostra*, le lendemain il le salua du nom de *Flaurus*. (Trad. H. Ailloud)

c) Les emprunts mutuels entre latin et grec fournissent également des indications très précieuses : le jeu des transcriptions et des translittérations vient éclairer ou confirmer la prononciation d'un certain nombre de phonèmes. Nous sommes ainsi renseignés sur la valeur sourde des *s* intervocaliques en latin à l'époque classique grâce notamment à la transcription du mot latin *Caesar* en grec Καῖσαρ; en effet, à cette époque, le grec marquait la sifflante sonore /z/ par  $\zeta$  et la sourde /s/ par  $\sigma$ ,  $\varsigma$ . Si les Grecs ont fait le choix d'un sigma pour la transcription du nom de César, c'est probablement qu'ils entendaient une sourde et non une sonore.

d) Les textes des inscriptions, et plus particulièrement les nombreuses fautes d'orthographe qu'on y trouve, se révèlent très utiles pour qui s'occupe de phonétique. Bon nombre d'inscriptions ont été réalisées par des gens de peu d'instruction ; ignorants parfois l'orthographe correcte de certains mots, ils les ont écrits selon ce que leur oreille leur dictait, c'est-à-dire phonétiquement. Ainsi *pleps* (*CIL* II 34) apparaît là où on attendrait *plebs*, *optinuit* (*CIL* IX 2845) est mis pour *obtinuit*, nous renseignant sur le dévoisement (assourdissement) que subissait l'occlusive sonore *b* devant une consonne sourde. Dans le même ordre d'idées, la valeur double du *i* intervocalique est mise en évidence par le dédoublement graphique du *i* dans certaines inscriptions : *cuius*, *maiorem* (*CIL* III 13374) pour *cuius* et *maiores*.

e) La métrique, qui repose sur des alternances réglées de syllabes longues et brèves, est une source non négligeable de renseignements. Un exemple intéressant nous est fourni par le verbe composé *reicio* (c'est-à-dire *reiicio*, de *rē-* et *iacio*). On trouve chez Lucrèce la forme *rēicīt* en première position de l'hexamètre<sup>14</sup>. Cette position devant nécessairement être occupée par une syllabe longue, il faut postuler un allongement par position de la première syllabe de *rēicīt*. Cet allongement ne peut s'expliquer que si l'on pose une transcription /rejjicit/. La syllabe initiale, se trouvant alors devant une géminée, compte comme

---

<sup>14</sup> Lucrèce, *De la nature* 1.34-35 : [...] in gremium qui saepe tuum se / *reicit*, aeterno deuictus uolnere amoris, « [Mars] souvent vient chercher asile sur tes genoux, vaincu à son tour par la blessure éternelle de l'amour » (trad. A. Ernout).

longue<sup>15</sup>. La valeur géminée que l'analyse métrique nous conduit à postuler dans le cas de la semi-voyelle /j/ en position intervocalique se trouve confirmée par un témoignage du grammairien Vélius Longus qui prend pour exemple le verbe composé *coicio* (*Grammatici Latini VII* 54.20-21 Keil):

De là l'usage de la gémination se répand, et on commence à écrire *coiiicit* avec trois *i*: une première syllabe *coi* suivie de deux syllabes *iicit*.

*f)* Finalement, il est possible de tirer un grand bénéfice de l'étude de l'évolution du latin dans les langues romanes. Particulièrement instructif est, dans ce contexte, le traitement des voyelles. Le latin, en effet, notait par le même graphème la brève et la longue. Dans beaucoup de cas, la métrique, l'analyse morphologique ou encore la méthode comparative nous renseignent sur la longueur des voyelles en latin; dans un certain nombre de cas, cependant, seule l'étude de la filiation des formes dans les langues romanes permet de trancher: prenons le mot latin *pēdem*; il a donné en français *pied*, en italien *piè* et en espagnol *pié*. D'un autre côté, le mot latin *sēta* a donné en français *soie*, en italien *seta* et en espagnol *seda*. Cette différence de traitement montre que les deux voyelles *e* en cause, même si elles ne se distinguent pas dans la graphie, ne représentent pas le même phonème: le premier *e* de *pēdem* est une voyelle brève ouverte alors que le *e* de *sēta* est une voyelle longue fermée.

### 3.2. Les limites actuelles de la synthèse de la parole

Malgré les efforts soutenus des laboratoires travaillant dans le domaine de la simulation de la parole, les synthèses, bien que d'une qualité très satisfaisante, n'en demeurent pas moins limitées sur un certain nombre de points, et force est de constater que les performances de ces systèmes ne sont pas équivalentes à celles de la parole humaine. Il n'y a pas lieu, dans le cadre de cet article, de faire un tour complet de la question de la position actuelle de la synthèse de la parole dans le domaine des sciences de la simulation: on se référera pour cela à deux articles récents

<sup>15</sup> Pour des développements complets sur le statut de la voyelle *i* en latin, voir Christian TOURATIER, «Quelques problèmes de phonologie à propos de -i-», in *Mélanges François Kerlouégan*, éd. D. Couso, N. Fick et B. Pouille, *Annales littéraires de l'Université de Besançon* 515, Paris: Les Belles Lettres, 1994, p. 623-632.

d'Eric Keller<sup>16</sup>. Nous nous contenterons ici de signaler deux limitations qu'il faut considérer pour la création d'une synthèse pour le latin.

*a)* La première limitation concerne les styles de parole. Par « style de parole », on entend les différents types d'expression qu'un locuteur peut produire, ainsi que les contextes d'énonciation dans lesquels ils s'inscrivent. On parle ainsi d'un style soigné, d'un style relâché, d'un style conversationnel, d'un style théâtral, etc. Très peu de styles de parole sont gérés par les systèmes de synthèses actuels. Cette limitation est due principalement à la façon dont les synthèses sont construites. Non seulement chaque style de parole possède des caractéristiques temporelles et intonatives propres, mais des variations apparaissent également entre différents locuteurs pour un même style de parole. Pour chacun des styles que l'on désire simuler, la somme de travail est considérable et l'investissement important : il faut compter plusieurs semaines de travail — voire des mois — pour une personne seule, même expérimentée.

*b)* La seconde limitation à laquelle on se heurte est la difficulté à simuler l'expressivité de la langue dans toute sa variété. Si, grâce à une approche psycholinguistique<sup>17</sup>, les modes (déclaratif, interrogatif, exclamatif, etc.) peuvent aujourd'hui être simulés de manière tout à fait acceptable, des nuances comme l'ironie, la fierté, la joie etc. ne peuvent pas être modélisés par les systèmes actuels à cause de la complexité extrême des références contextuelles — ainsi que des réalisations prosodique et acoustique — qui sont mises en œuvre.

Dans ce contexte, la question se pose alors de savoir quel style de parole est le plus approprié à la synthèse du latin. En ce qui concerne les langues modernes, tous les laboratoires ont fait le choix de simuler la lecture à haute voix. C'est le choix que nous avons fait nous aussi pour le latin. S'il est vrai qu'il nous permet

<sup>16</sup> Eric KELLER, « La Vérification d'hypothèses linguistiques au moyen de la synthèse de la parole », *Cahiers de l'institut de linguistique*, 28 (2003), à paraître ; Eric KELLER, « Towards Greater Naturalness : Future Directions of Research in Speech Synthesis », in *Improvements in Speech Synthesis*, éd. E. Keller, G. Bailly, A. Monaghan, J. Terken et M. Huckvale, Chichester (UK) : Wiley & Sons, 2001, p. 3-17.

<sup>17</sup> C'est l'approche retenue pour les synthèses développées dans notre laboratoire.

de bénéficier de l'expérience déjà accumulée par d'autres, il s'impose en fait de lui-même eu égard à la double difficulté (limitations technologiques, absence de locuteur natif) à laquelle nous sommes confrontés. Il était essentiel, dans ces conditions, de faire le choix d'un style de parole le plus neutre possible : la parole lue est un style qui demande moins d'expressivité marquée que ne le requiert une récitation théâtrale par exemple ; il est aussi plus codifié (et donc prédictible) du point de vue de la syntaxe qu'un style conversationnel ; la lecture à haute voix, enfin, se réalise à une vitesse d'énonciation adaptée à une prononciation et à une articulation soignées.

Nous pouvons dès lors, pour résumer, dresser le portrait du locuteur idéal que nous cherchons à simuler grâce à la synthèse de la parole : il s'agit d'un aristocrate romain de la Ville vivant à la fin du I<sup>er</sup> siècle avant J.-C., de langue maternelle latine, sans défaut de prononciation ; c'est un homme adulte (le locuteur que nous avons choisi pour les enregistrements nous l'impose) ; il doit connaître parfaitement le grec (pour pouvoir prononcer correctement les mots grecs qui figurent dans les textes) ; il doit être cultivé et posséder d'excellentes connaissances littéraires (de façon à pouvoir lire aisément tout type de texte). On soumettra à ce locuteur idéal des œuvres latines pour qu'il en fasse une lecture à voix haute<sup>18</sup>.

#### *4. Les étapes de la reconstruction*

Pour créer une synthèse de la parole pour le latin, tous les renseignements que l'on peut tirer des différentes sources que nous avons énumérées précédemment doivent dans un premier temps être analysés et classés. De cette façon, ils pourront être mis pleinement à profit lors des différentes étapes de la construction du système.

<sup>18</sup> On peut rappeler ici que la lecture à voix haute (par opposition à la lecture silencieuse, *lectio tacita*) est la seule que connaissaient les Romains à cette époque ; selon le témoignage d'Augustin, Ambroise de Milan (IV<sup>e</sup> siècle) fut le premier homme de l'Antiquité à pratiquer la lecture mentale (*Confessions* 6.3.3) : « Lisait-il, ses yeux couraient sur les pages dont son esprit perçait le sens ; mais sa voix et sa langue se reposaient. Souvent quand je me trouvais là [...], je le voyais lisant silencieusement et jamais autrement. [...] Et puis la nécessité de ménager sa voix, qui se brisait aisément, pouvait être encore une juste raison de lire silencieusement » (trad. P. de Labriolle, légèrement modifiée).

On peut schématiser les composantes majeures d'un système de synthèse de la façon suivante :

1. entrée : il s'agit du texte latin que l'on soumet au logiciel ; cela peut être un texte scanné ou tapé au clavier : il se présente sous la forme d'un document texte
2. traitement : cette partie est constituée de trois composantes :
  - a. composante phonologique
    - i. traduction graphémo-phonétique
    - ii. traitement phonologique
  - b. composante prosodique
    - i. structuration temporelle (ms)
    - ii. fréquence fondamentale (F0, Hz)
    - iii. modulation de l'amplitude (dB)
    - iv. styles de parole et modulations de la voix
  - c. composante de traitement du signal
    - i. pour les systèmes concaténatifs : sélection de diphones, élaboration de la structuration temporelle et de la modulation F0, concaténation, lissage
3. sortie : il s'agit du son que l'utilisateur entend par l'intermédiaire des enceintes de son ordinateur.

Les étapes 1 et 3 n'appellent pas de commentaire particulier : elles sont largement identiques pour toute synthèse et ne représentent que les points de départ et d'arrivée du système. C'est à l'étape 2, qui concerne le traitement des données, que nous allons nous intéresser maintenant. Cette étape comprend trois composantes : une composante phonologique, une composante prosodique et une composante de traitement du signal.

#### *4.1 La composante phonologique*

Le texte latin soumis au logiciel ne peut être traité en l'état. La première étape consiste à « traduire » les graphèmes en phonèmes. Le résultat obtenu est une chaîne phonétique simple, non annotée. Ainsi que le montrent les différentes sources de notre connaissance de la prononciation du latin énumérées plus haut, nous sommes bien renseignés sur le système phonologique de cette langue. La traduction graphémo-phonétique peut donc être qualifiée de raisonnablement fiable. Il ne reste que quelques cas litigieux, qui ont été abondamment discutés par les linguistes et les phonéticiens. On citera notamment les questions de savoir si l'on aspirait le *h* initial, comme dans *homo*, si l'on prononçait, et de quelle manière, la nasale *n* devant *s* et *f* dans un mot comme

*consul*<sup>19</sup>, ou si le *m* final, comme dans *urbem*, disparaissait en entraînant la nasalisation de la voyelle précédente.

Un point délicat de la traduction graphémo-phonétique est la transcription des voyelles. Comme nous l'avons brièvement évoqué plus haut, la graphie ne notait pas la longueur des voyelles. Or la transcription phonétique — et donc la réalisation phonique — d'un *e* bref ouvert [ɛ], par exemple, n'est pas la même que celle d'un *e* long fermé [e:]. Pour transcrire correctement la longueur des voyelles, nous devons utiliser un système complexe de règles fonctionnant sur la base d'une analyse morphologique obtenue à partir d'un dictionnaire électronique annoté manuellement.

Une fois cette première étape réalisée, un certain nombre d'améliorations et de corrections doivent être apportées à la chaîne phonétique obtenue : la réalisation des phonèmes peut être influencée par le contexte phonologique dans lequel ils s'inscrivent. Ainsi, pour reprendre un des exemples cités plus haut, on constate en latin un dévoisement (assourdissement) préocclusif : *plebs* doit être transcrit / pleps /. Ces traitements complémentaires de la chaîne phonétique sont eux aussi déterminés par un jeu de règles.

#### 4.2 *La composante prosodique*

Si, comme nous l'avons vu, nos diverses sources d'informations nous offrent un panorama très complet du système phonologique latin, elles contiennent en revanche beaucoup moins d'indices sur la prosodie. En synthèse de la parole, on entend par prosodie les phénomènes sonores en relation avec la durée (structuration temporelle), la mélodie (variation de la hauteur de la voix), l'amplitude (force de la voix) et, d'une manière plus générale, avec les caractéristiques propres à la voix (qui indiquent — ou sont des signes de — l'état psychologique et émotif du

<sup>19</sup> L'histoire de la nasale *n* dans cette position peut être reconstruite ainsi : très tôt, la nasale *n* a perdu son occlusion devant *s* et *f*. La perte de l'occlusion a sans doute eu pour conséquence, dans un premier temps, le transfert des vibrations nasales de la consonne sur la voyelle précédente, avec allongement de cette dernière par compensation. Puisque cette nasale n'existe plus en tant que consonne, elle a disparu dans l'orthographe et dans la prononciation des classes les moins éduquées. Cependant, comme souvent en latin, des raisons d'analogie ou de sentiments étymologiques ont rétabli ou maintenu le *n* dans l'orthographe officielle avec une constance absolue.

locuteur : ironie, peur, chagrin...). La technologie Mbrola<sup>20</sup>, qui repose sur un système de concaténation de diphones<sup>21</sup> et dont nous nous servons actuellement dans notre laboratoire pour la gestion des voix et de la reproduction sonore<sup>22</sup>, ne nous permet de manipuler que les deux premiers paramètres : la durée et la mélodie.

Pour les langues modernes, les valeurs de durée et les contours mélodiques sont obtenus sur la base de modèles statistiques construits à partir d'analyses d'un corpus de phrases préalablement enregistrées<sup>23</sup>, produites par un locuteur natif de la langue testée, et segmentées<sup>24</sup>. La vérification des modèles ainsi conçus se fait en grande partie de manière empirique : le résultat sonore obtenu est soumis au jugement d'un locuteur natif de la langue, grâce auquel ces modèles peuvent être, selon les cas, améliorés ou corrigés.

Ne possédant pas de locuteur natif pour le latin, l'enregistrement du corpus de phrases — étape inévitable dans le processus de création d'un système de synthèse — devait être effectué par une personne possédant les nombreuses qualités lui permettant de prononcer ces phrases dans le strict respect des critères que nous lui imposons. C'est ainsi que notre choix s'est porté sur un chercheur de nationalité tchèque qui possède, de par son parcours académique et ses expériences personnelles, toutes les qualités

<sup>20</sup> Le projet Mbrola (<http://tcts.fpms.ac.be/synthesis/mbrola.html>), qui a vu le jour au TCTS Lab de la Faculté polytechnique de Mons (Belgique), a pour objectif d'encourager et de dynamiser la recherche en synthèse de la parole en mettant gratuitement à disposition des bases de données et des systèmes de synthèse pour un grand nombre de langues.

<sup>21</sup> Un diphone représente, en synthèse de la parole, un fichier sonore qui s'étend sur deux phonèmes contigus, de la moitié du premier à la moitié du second. L'utilisation d'unités découpées de cette façon permet de concaténer deux diphones par leur partie la plus stable du point de vue du signal acoustique, limitant ainsi le lissage nécessaire aux jointures.

<sup>22</sup> Il est prévu, à terme, que la synthèse du latin dispose de son propre moteur diphonique, actuellement en cours d'élaboration dans notre laboratoire. Un tel outil permettrait de s'affranchir de certaines des limites actuelles du système Mbrola.

<sup>23</sup> Le corpus de phrases enregistrées contient toutes les transitions phonème à phonème de la langue que l'on veut simuler. Chaque transition apparaît, dans la mesure du possible, en plusieurs endroits de la phrase (début, milieu, fin de groupe temporel, début, milieu, fin d'énoncé, etc.).

<sup>24</sup> La segmentation est le découpage, automatisé dans un premier temps, puis affiné manuellement, du signal de la parole enregistré numériquement.

requises : au bénéfice d'un diplôme d'études de lettres françaises et de linguistique / phonétique générale, titulaire d'un doctorat tchèco-français en phonétique, il parle couramment, en plus du tchèque qui est sa langue maternelle, le français, l'allemand et l'anglais, et possède d'excellentes connaissances de l'italien, du croate et du russe auxquelles s'ajoutent des notions d'arabe et de latin. Ses connaissances théoriques de linguiste et de phonéticien, complétées par une expérience pratique dans plusieurs langues vivantes aux systèmes phonologiques parfois très différents, lui confèrent un excellent contrôle et une grande souplesse articulatoire qui en font un candidat idéal pour l'enregistrement d'un corpus de phrases latines.

En tchèque, comme en latin, la durée vocalique est phonologique. Elle permet parfois à elle seule de distinguer deux formes (par exemple le mot latin *lēgit*, « il lit », en face de *lēg̊it*, « il a lu »). En tchèque, comme en latin, les syllabes sur lesquelles tombe l'accent verbal peuvent contenir une voyelle longue comme une voyelle brève. Ces deux éléments communs au tchèque et au latin représentent un atout supplémentaire lors de l'enregistrement du corpus dans la mesure où le locuteur tchèque qui lit du latin est amené de manière naturelle par sa langue maternelle à reproduire les différences entre voyelles longues et voyelles brèves, accentuées et non accentuées<sup>25</sup>: les durées vocaliques que l'on mesure pour ce locuteur peuvent donc servir de base à la création d'un modèle statistique pour la structuration temporelle en latin.

La seconde composante prosodique paramétrable en synthèse de la parole est, comme nous l'avons déjà mentionné, la courbe mélodique (variations de la fréquence fondamentale). Pour beaucoup de latinistes, le terme « mélodique » évoque la question longtemps débattue — et toujours ouverte — de savoir si l'accent latin était, selon la terminologie consacrée, intensif ou mélodique (de hauteur). Sans reprendre ici les arguments des uns et des autres<sup>26</sup>, il est à noter que des études récentes montrent que les

<sup>25</sup> Un francophone aura tendance à allonger de manière artificielle et démesurée les voyelles longues pour distinguer, par exemple, *causā* (ablatif) de *causā* (nominatif); un italophone, quant à lui, sera enclin à allonger une voyelle brève qui porte l'accent verbal : il prononcera *mēdicus* en allongeant le *e*, alors que ce dernier est bref en latin (*mēdicus*).

<sup>26</sup> On trouvera une synthèse de ce débat, avec des références bibliographiques, chez Jean HELLEGOUARC'H, « Métrique et oralité », in *Les*

notions d'intensité et de mélodie ne s'excluent pas et que l'accent verbal étend son influence aussi bien au niveau de la fréquence fondamentale et de la durée segmentale que de l'amplitude<sup>27</sup>. Le locuteur tchèque a reçu comme consigne, lors des enregistrements, de produire une légère inflexion de la voix, aussi neutre et naturelle que possible, sur la syllabe qui reçoit l'accent verbal. Ce choix de marquer, même légèrement, la place de l'accent permet de garantir un certain naturel dans la production de l'ensemble de l'énoncé : une prononciation sans accentuation aurait induit, de par son caractère totalement artificiel, des phénomènes parasites invalidant bon nombre d'analyses prosodiques réalisables sur la base de ces enregistrements. Même si elle n'est marquée que légèrement, la présence de cet accent verbal et les incidences qu'a ce dernier tant sur un plan phonologique que prosodique peuvent servir de point de départ, grâce à la flexibilité offerte par la synthèse, à la création de plusieurs modèles destinés à mettre à l'épreuve les théories concurrentes brièvement évoquées plus haut.

### *Conclusion*

Tant au niveau phonétique que prosodique, rien ne permet d'affirmer avec certitude — ni non plus de nier — que nos mesures et nos restitutions correspondent à la réalité de la langue latine du I<sup>er</sup> siècle avant J.-C. Le latin, néanmoins, comme toute langue (qui a été) parlée, ne peut avoir échappé aux contraintes universelles que nous impose notre appareil phonatoire. Les structures physiologiques de ses composants (cordes vocales, palais, dents, langue, etc.) rendent certaines coarticulations sinon impossibles, du moins peu probables. Les renseignements que nous pouvons réunir sur la prononciation du latin nous permettent d'établir un certain nombre de parallèles avec d'autres systèmes phonologiques proches pour lesquels il existe des locuteurs natifs : grâce à ces derniers, nous pouvons obtenir des enregistrements et les

---

*Structures de l'oralité en latin. Colloque du Centre Alfred Ernout, Université de Paris IV, 2, 3 et 4 juin 1994*, textes réunis par J. Dangel et C. Moussy, Paris : Presses de l'Université de Paris-Sorbonne, 1996, p. 265-266.

<sup>27</sup> Voir en particulier Philippe MARTIN, « Modelling F0 in Various Romance Languages : Implementation in Some TTS Systems », in *Improvements in Speech Synthesis*, p. 104-119.

analyser. Ces parallèles peuvent nous aider, dans une certaine mesure et en prenant toutes les précautions qui s'imposent, à mieux appréhender les mécanismes subtils d'une langue phonologiquement proche, de manière à affiner et à parfaire nos modèles pour le latin.

Car la synthèse de la parole appliquée au latin classique doit être avant tout un outil scientifique susceptible d'aider le chercheur à mieux comprendre, à mieux percevoir la langue latine.

Olivier BIANCHI

Université de Lausanne

Institut d'archéologie et des sciences de l'Antiquité (IASA)

Laboratoire d'analyse informatique de la parole (LAIP)