

REMARQUES AU SUJET DES VOYELLES ADDITIONNELLES

MARGUERITE DURAND

Voici les observations qui sont à la base des recherches dont il est ici question: j'ai souvent été frappée et intriguée par ce fait: des voyelles qui se glissent à l'intérieur d'un mot, voyelles non perçues, qui n'apparaissent qu'à l'enregistrement phonétique, sont souvent aussi longues, quelquefois plus longues, que de vraies voyelles, bien perçues, faisant fonction syllabique. Citons quelques exemples: dans la phrase *il y a de l'eau dans la bouilloire*, nous voyons une voyelle intercalaire et inaudible, entre *l* et *j*, mesurant 5 centisecondes, puis une autre entre *j* et *w*, mesurant 9 centisecondes, alors que le premier *a*, bien audible, mesure 7 centisecondes, et que le *ä* mesure 6 centisecondes. Dans l'enregistrement des deux vers de Boileau:

*Il n'est pas de serpent ni de monstre odieux
Qui par l'art imité ne puisse plaire aux yeux*

nous trouvons, entre *r* et *l* (*par l'art*), une voyelle inaudible mesurant 6 centisecondes, alors que, dans ce même distique, nous trouvons des voyelles audibles et identifiables, et pourtant de durée moindre: le *i* de *ni* qui mesure 5 centisecondes, le *a* de *ne* qui mesure 3 centisecondes. Dans cette autre phrase *il faudra de l'argent*, entre le *r* et le *z* vient s'insérer une voyelle inaudible de 6 centisecondes, ce qui est précisément la durée du *a* de *argent*. Nous pourrions multiplier les exemples semblables de voyelles épenthétiques, inaudibles, dont l'enregistrement expérimental seul souligne la présence, et dont pourtant la durée est suffisante pour la perception.

A l'opposé, certaines voyelles épenthétiques ou prothétiques apparaissent avec un timbre assez clairement perçu pour que l'orthographe les note. Citons le cas du mot français *canif* avec son *a* épenthétique, du *e* prothétique développé en français comme en espagnol devant les mots latins commençant par *s* + *occlusive*, les voyelles prothétiques et assez diverses développées en grec¹ pour les racines à initiale *r* ou dans un groupe de consonnes comprenant une occlusive.

Quels sont les cas où se développe une voyelle additionnelle, prothétique, épenthétique ou finale? C'est là une question qui ressortit à la phonétique générale. Les cas peuvent se ramener à quatre principaux:

I. Une mauvaise constitution de la syllabe, c'est-à-dire une divergence entre la constitution syllabique du groupe d'après de Saussure (ouverture + fermeture des

¹ M. Lejeune, *Traité de Phonétique Grecque*, pp. 181-182.

organes phonateurs) et la constitution syllabique du groupe d'après Grammont-Fouché (contraction + décontraction des muscles phonateurs); dans ce cas, une nouvelle voyelle tend à se créer, changeant la constitution syllabique. Ce fait peut avoir lieu:

α) avant la première consonne, créant une voyelle prothétique, comme le *e* de *épée* et de *école*, le *α* de *ἀστήρ*, le *ι* de *ἰκτίνος*.

β) à l'intérieur du groupe, comme nous pouvons l'observer dans *Θάλαττος*² où le premier *α* coupe le groupe étymologique *dhn*.

II. Le fait de débiter un mot par une consonne faible telle que *l*, *r*, *z*. Il se produit alors une prothèse vocalique qui peut fréquemment s'observer dans des enregistrements phonétiques de phrases à initiale *l* ou *z* en français (ce qui est extrêmement fréquent à cause des morphèmes *je* et *le*, et il se produit de façon systématique en grec (*ἐλεύθερος* **leudh*, *ἀνεψιός* /lat. *nepos*)³); ce fait semble dû à une gémation de la consonne douce antérieure, donc être un cas particulier de la mauvaise constitution syllabique dont il a été question plus haut.

III. Une consonne peut être en danger de chute par dissimilation. C'est le risque que court le premier *r* dans des mots français comportant la séquence *r* + *occlusive* + *r*, dans les mots comme *arbre*. Dans de tels mots, le premier *r*, menacé de dissimilation par le second, ce qui causerait sa chute (*r* + *r* > *zéro* + *r*), se montre, dans les enregistrements phonétiques, séparé de la consonne suivante par une petite voyelle. De la sorte se crée une syllabe supplémentaire, faisant passer le *r* en danger de syncope de la position finale à la position initiale, de la tension musculaire faible à la tension forte, et ainsi réduisant le danger de dissimilation entre les deux consonnes semblables. Un tel fait ressortit encore à la constitution syllabique du groupe.

IV. Un mot peut comporter un groupe consonantique. Le groupement de deux consonnes, aussi bien dans deux syllabes successives que dans la même syllabe, se fait avec une interpénétration des temps des deux consonnes, l'implosion de la seconde venant avant l'explosion de la première. Un tel traitement est nécessairement le fait d'une articulation tendue. Il ne peut donc guère avoir lieu pour une séquence de consonnes faibles qui ne sont pas groupées à proprement parler, mais seulement juxtaposées, et même le plus souvent séparées par une voyelle épenthétique.

Ce que nous nous proposons d'étudier, ce sont ces voyelles non perçues quoique de durée suffisante, les voyelles additionnelles perçues posant d'autres problèmes.

Pour étudier ces voyelles, nous avons procédé par enregistrements spectrographiques et nous avons rapproché les bandes de résonances, ou formants de ces voyelles épenthétiques, des formants constitutifs des voyelles environnantes. Nous nous sommes basées sur les études du Laboratoire Haskins à New York,⁴ établissant: 1° qu'une

² M. Lejeune, *op. laud.*, p. 178.

³ M. Lejeune, *op. laud.*, p. 128.

⁴ F. Cooper, P. Delattre, A. Liberman, J. Borst, L. Gertsman, "Some experiments on the perception of synthetic speech sounds", *Journal of the Acoustical Society of America*, 24 (1952), pp. 597-606.

voyelle donnée a des formants à des hauteurs déterminées, 2° que, de ces formants, deux sont essentiels, pour l'identification de ces voyelles. En effet, si l'on reconstitue les phonèmes à l'aide d'un appareil synthétiseur, il suffit de reporter les deux formants inférieurs pour que l'identification de la voyelle soit possible.

Rappelons brièvement que, pour les voyelles orales françaises, la hauteur de ces deux premiers formants est⁵:

<i>i</i> 240-2.500 cycles	<i>u</i> 240-750 cycles
<i>e</i> 350-2.200 cycles	<i>o</i> 350-865 cycles
<i>ɛ</i> 510-1.950 cycles	<i>ɔ</i> 510-1.000 cycles
<i>a</i> 725-1.300 cycles	<i>ɑ</i> 650-1.200 cycles
<i>y</i> 240-1.850 cycles	<i>ø</i> 240-750 cycles
	<i>æ</i> 510-1.400 cycles.

Rappelons que, pour les voix féminines, et c'est souvent le cas de nos enregistrements, les formants sont un peu plus élevés.

En ce qui concerne l'élément vocalique prothétique, je n'ai pas eu la possibilité d'en rassembler de nombreux exemples. D'après ce que j'ai pu observer, cet élément ne semble pas être une voyelle. En effet, pour que nous puissions parler de voyelle, il faut qu'il y ait un accord entre plusieurs notes composantes, deux au moins; or, en initiale, dans les quelques exemples que nous avons pu rassembler, nous ne trouvons devant la consonne qu'une bande sonore très basse, vers 200 cycles, c'est-à-dire à la hauteur des vibrations laryngiennes. Nous ne donnerons pour exemple que ce tracé de la phrase *la famille royale est arrivée*, commençant par un *l*: devant le *l*, nous trouvons une bande vers 250 cycles; il ne semble pas qu'il y ait là matière à distinguer une voyelle, il doit plutôt s'agir de l'ouverture de la bouche, avec mise en fonctionnement du larynx, sans qu'il puisse s'agir d'un élément phonétique quelconque. Nous avons rassemblé divers enregistrements, de plusieurs locuteurs (tous français), et commençant par *l* ou *ʒ*, qui présentent tous la même bande grave.

La voyelle de détente, c'est-à-dire placée après le mot, se présente de façon moins simple. Comme nous croyons l'avoir vu précédemment,⁶ la voyelle qui suit une occlusive sonore (ex. *froide*) est la voyelle précédente, c'est un *a* dans *froide*, un *y* dans *Hugues*; cette conclusion provient de l'examen des spectrogrammes et de la reconstitution des consonnes finales à l'aide d'un synthétiseur; après une occlusive sourde, la détente sonore semble ne pas être vocalique, elle est réduite à une bande très basse ou à une bande à la hauteur propre de la détente de la consonne.

Ce à quoi nous nous sommes plus particulièrement attachée, ce sont les voyelles intérieures, épenthétiques, qui viennent disjoindre des groupes consonantiques de mauvaise constitution syllabique ou composés de consonnes d'articulation faible.

Soit la phrase *une belle bille de marbre blanc*: le *l* se trouve en mauvaise position en

⁵ P. Delattre, "Un triangle acoustique des voyelles orales du français", *The French Review*, XXI, nr 6, pp. 478-484.

⁶ M. Durand, "De la perception des consonnes occlusives. Questions de sonorité," *Word*, 12, 12.

fin de syllabe, il court le risque de syncope, ou encore, du moins en français, le risque de vocalisation; de plus, le premier *r* est en position plus défavorable encore car il est aussi en fin de syllabe et, de plus, il court le risque de dissimilation par le second *r*, $r + r > \text{zéro} + r$; le *j* de *bille*, consonne assez faible, se trouve aussi en mauvaise position, en fin de syllabe. La solution adoptée pour toutes ces consonnes en danger de chute a été d'insérer, entre la consonne en danger et la consonne suivante, une voyelle épenthétique qui fera passer la consonne mal placée de la tension décroissante à la tension croissante.

Quel est le timbre de ces voyelles additionnelles? Entre *l* et *b*, nous trouvons une bande centrée vers 1600 cycles et une bande basse, le tout étant situé pendant le temps d'explosion de *l*, explosion visible sur la ligne d'intensité. La bande centrée vers 1600 cycles correspond peut-être au second formant essentiel du *ɛ* qui précède. B. Hála⁷ donne pour *e* la résonance supérieure 1750 cycles et P. Delattre⁸ donne pour *ɛ* un second formant de hauteur 1750 cycles. Il semble plus probable que cette résonance corresponde à la hauteur propre de *b*, et nous penserions alors que cet élément épenthétique n'est pas exactement de nature vocalique puisqu'il n'est pas exactement composé par un accord musical, n'ayant qu'un son laryngien et une seule résonance.

Différent est l'élément placé entre *j* et *d*: il a un formant centré sur 2500 cycles, exactement comme le *i* dont le *j* est très proche comme timbre et aussi dans la conscience linguistique du locuteur; le premier formant est un peu plus bas que celui de *i* tel qu'on le voit dans la voyelle qui précède; cette voyelle épenthétique est un *i*, de timbre un peu plus diffus qu'un *i* normal, une telle différence entre cet élément et un *i* risque-t-elle de rendre cette voyelle non reconnaissable pour un *i*? Des travaux ultérieurs à l'aide d'un synthétiseur nous permettront de résoudre cette question mais, autant que l'on puisse juger par l'œil de ce qui est de nature auditive, il ne semble pas que cette voyelle ne puisse pas être reconnue comme *i*.

Quant à la voyelle qui disjoint le groupe *rbr*, elle fait prévoir le *ə* qui suit ces trois consonnes, elle a le même formant centré sur 2800, mais il lui manque le premier formant centré⁹ vers 1000.

En résumé, dans ce mot *marbre* où le premier *r* est menacé d'un double danger, s'il reste dans la position de fin de syllabe que lui assigne la conscience linguistique, mais le souci orthographique interdisant l'insertion d'une véritable voyelle perceptible, nous voyons que le sujet parlant a utilisé une voyelle, mais une voyelle mutilée peut-on dire: la voyelle qui suit ce très mauvais groupe consonantique, mais à laquelle manque une résonance essentielle pour être identifiable, une voyelle qui ne peut être perçue au sens étymologique du mot, choisie entre les voyelles à l'exception de toutes les autres.

Ce procédé de la voyelle non identifiable est d'observation fréquente; signalons-le

⁷ B. Hála, *Akustika podstata somohlasek* (Prague, 1941), p. 190 et suiv.

⁸ P. Delattre, "Triangle acoustique", *op. laud.*

⁹ J. Varney-Pleasant, *Etudes sur l'e muet* (Paris, Klincksieck, 1956), p. 40.

dans la syllabe *grø* où le *g* et le *r*, consonnes faibles, sont peu aptes à se grouper: une voyelle épenthétique *a*, en commun avec le *ø* de la syllabe, le formant bas centré vers 350 cycles, mais n'a pas le second qui serait centré vers 1500; cette voyelle additionnelle est un *ø*, mais un *ø* mutilé de l'un des deux éléments qui le rendraient identifiable. Dans un autre exemple, à *la Grèce assemblée*, deux voyelles épenthétiques se sont formées dans les groupes consonantiques faibles *gr* et *bl*. La première, entre *g* et *b*, a bien le formant 2000 cycles du *ε* suivant, mais n'en a pas le premier formant; la seconde, entre *b* et *l*, a bien le premier formant bas du *e* suivant, mais le second, trop bas pour *e*, rappellerait plutôt le timbre de *ø*; on pourrait même penser qu'il y a là une anticipation du *e*, mais altéré dans son timbre de telle sorte que le locuteur obtient une voyelle trop inattendue dans ce groupe pour être choquante pour l'auditeur.

Une remarque analogue peut être faite pour la syllabe *bli*, où la voyelle épenthétique entre *b* et *l* a bien le premier formant bas du *i* suivant, mais un second formant centré vers 2000 cycles, nettement trop bas pour *i*, à peu près à la hauteur de *y*.

Des remarques analogues peuvent être faites dans d'autres langues que le français; donnons-en quelques exemples: en espagnol de Colombie, nous avons les trois enregistrements *crisis*, *cruza*, *crasis*, qui présentent trois solutions du problème de la voyelle épenthétique: dans la syllabe *cri*, la voyelle additionnelle présente deux formants qui n'ont aucun rapport avec ceux de *i*; le premier formant bas du *i* manque, un autre vers 1600 cycles est trop bas pour être le second formant de *i* (qui se trouverait vers 2500 cycles). Dans la syllabe *cru*, la voyelle épenthétique a bien l'un des deux formants essentiels de *u*, le premier formant bas; un second, très faible, placé un peu en dessous de 200 cycles, donc trop haut pour caractériser un *u*. Quant à la syllabe *cra*, elle présente une voyelle épenthétique qui a bien le premier formant de *a* (vers 1500), mais non pas le premier (vers 650). Autrement dit, dans ces trois mots espagnols, nous trouvons trois voyelles épenthétiques: l'une est totalement étrangère à son groupement phonétique, les deux autres sont la voyelle de la syllabe, mais mutilée de l'un des deux éléments essentiels qui la rendraient identifiable.

Nous croyons voir un fait analogue dans un tracé en langue ukrainienne où, entre un *d* et un *r*, se trouve un élément vocalique auquel manque un formant bas pour faire un *u*.

En résumé, parmi ces voyelles épenthétiques, nous trouvons des exemples d'anticipation de la voyelle de la syllabe, ce qui n'est pas surprenant à première vue, mais aussi des exemples d'éléments qui semblent intéressants du point de vue psychologique, des éléments qui sont la voyelle de la syllabe, mais mutilée de l'un des formants essentiels qui la feraient identifier.

On pourrait penser que ce formant a pu être tronqué à cause de la brièveté de cette voyelle épenthétique, mais il ne semble pas que nous puissions voir dans un simple fait physiologique, dans le manque de temps pour mettre les organes phona-

teurs en place, la cause de l'absence d'un formant, alors que nous pouvons trouver des exemples de voyelles dépourvues d'un formant et pourtant prises entre deux phonèmes qui en sont pourvus (comme la voyelle prise entre *d* et *r* dans *l'aérodrome d'Orly*).

Il semble qu'il y ait là une entreprise de locuteurs au courant de l'orthographe de leur langue, tenant à insérer une voyelle que les conditions du mot ou du groupe rendent inévitable sous peine de syncope d'une consonne (et ce souci est d'ordre linguistique), et faisant en sorte que cette voyelle épenthétique soit impossible à reconnaître, donc non perçue.

Pour être confirmée, ou infirmée, une telle hypothèse devrait être appuyée sur des enregistrements de locuteurs illettrés, ce qui n'est pas très facile à obtenir.

Je terminerai en remarquant qu'il est un problème connexe dont la solution peut être d'une grande complexité: lorsque cette voyelle épenthétique vient à être entendue jusqu'à modifier l'orthographe du mot, pourquoi telle voyelle et non telle autre est-elle choisie? La voyelle de la syllabe, mais mutilée de l'un de ses éléments laisse la porte ouverte à la voyelle de la syllabe (le groupe *kra* de l'espagnol de Colombie laisse la porte ouverte à *kara*). Mais des voyelles additionnelles semblent être de véritables corps étrangers et donner des solutions différentes: pourquoi la même mauvaise constitution syllabique donne-t-elle un *a* en grec (*ἀστήρ*), un *e* en français et en espagnol (*épée*, *espada*), un *i* en turc et en hongrois (*Istanbul*, *Istvan* > *Stephanus*)? La langue emploie-t-elle la voyelle la plus fréquente dans son système? et cette fréquence même serait une sorte de succédané de la voyelle mutilée car un phonème très fréquent est sans doute moins remarquable qu'un autre. Cela est peut-être vrai pour le français, où *e* est bien la voyelle la plus fréquente. Est-ce un élément de dissimilation, comme dans *canif*, où le *i* originel et le *a* adventice ont une opposition complète de diffuse/compacte? Ou bien le choix de cette voyelle ressortit-il à des tendances permanentes de la langue (ouverture, fermeture)? Ce sont là des questions à étudier, en tenant compte de ces voyelles additionnelles en rapport avec les tendances de bien des langues.