

Le statut de l'oralité dans l'enseignement-apprentissage des langues

Julio MURILLO PUYAL

Laboratori fLexSem, Universitat Autònoma de Barcelona

Résumé

Alors que nous célébrons le centenaire de la publication du *Cours de Linguistique Générale*, le principe selon lequel les langues « ont une manifestation normale et primaire qui est phonique » a acquis valeur axiomatique en Linguistique. Vu les relations traditionnellement ancillaires entre la Didactique des Langues et la Linguistique, il y aurait lieu de penser que ce même principe a été adopté dans l'enseignement-apprentissage des langues et que la didactique de l'oralité ne fait plus problème.

Le traitement réservé à la matière phonique dans l'enseignement des langues étrangères, et la notion même de langue orale sur laquelle se fondent ces approches révèlent cependant que, dans le meilleur des cas, l'oralité est réduite et circonscrite au flux phonique et/ou à l'activité des organes du tractus phonatoire.

Or, d'innombrables phénomènes de parole (variabilité phonique, rythme, métaphonie, macromotricité, modalités énonciatives, proprioception, etc.) témoignent de ce que les manifestations phoniques sont générées dans le processus du *dire* bien avant qu'elles soient produites dans le tractus phonatoire, et avant que se structure le flux phonique en phonèmes : le processus de parole s'amorce dans une phase prélinguistique et est régi par la « Logique du vivant ». C'est dans cette phase d'émergence du langage que l'intervention pédagogique a sa raison d'être. Les *valeurs de la langue parlée* permettent de caractériser ce processus en vue d'établir des procédures didactiques qui ne peuvent être efficaces qu'à cette étape.

Le statut de l'oralité étant ainsi défini, *l'écriture* – conçue comme représentation symbolique et non comme image analogique –, et bien entendu à condition qu'on n'y ait recours que pour compléter le processus de structuration et mémorisation de la matière phonique et non pour amorcer ledit processus, constitue un moyen « dont on ne peut se passer » (Saussure *dixit*) pour représenter la parole et l'engrammer d'une manière efficiente.

Mots-clé : Didactique de l'oral, iconicité parole-macromotricité, prosodie-grammaire/ lexicale, engrammation de l'oral et scription.

Abstract

On the 100th anniversary of the publication of *Course in General Linguistics*, the principle that languages “have a primary and normal manifestation which is phonic” (Alarcos) has achieved an axiomatic value in Linguistics. Given that Didactics of Languages has traditionally maintained subordinate relations with Linguistics, one might think that this same principle has been adopted in language teaching-learning, and that orality teaching does not represent an issue anymore.

However, the treatment given to the phonic material in foreign language teaching-learning and the notion itself of oral language on which these approaches are founded reveal that, at best, orality is often reduced and limited to the phonic flow and/or to the activity of the vocal tract which is supposed to produce it. Teaching intervention turns out to be based on physicalist descriptions (acoustic and/or articulatory) of these phonic speech performances which is particularly reductive and biased in didactic perspective.

A wide range of phenomena such as phonic variability, rhythm, metaphony, macromotorics, sentence modalities or proprioception among others prove however that articulatory or acoustic performances are created in the process of “dire” even before they are produced and before the phonological structuration of the phonic flow is accomplished by the audio-phonatory activity: The process of speech begins in the pre-linguistic phase and takes places in the “Logic of Life”. As pedagogical intervention can only be efficient if it is carried out at this stage, the “valeurs de la langue parlée” (spoken language values) enable us to characterize in order to establish didactic procedures.

From that perspective, and obviously provided it is used to settle and complete the process of structuration and memorization of the phonic material and not to initiate this process, *writing*, conceived as a symbolic representation of orality, not as an analogical image, constitutes a means “we cannot simply disregard” (*dixit* Saussure), to represent speech and to make more efficient didactic intervention and internalize the language efficiently.

Key words: Speech didactics, iconicity speech-macromotorics, prosody grammar / lexis, speech internalization and writing.

ALORS que nous célébrons le centenaire de la parution du *Cours de Linguistique Générale*, il peut paraître que la question de l'oralité ne fait plus problème, et que son statut en Didactique des Langues et des Cultures a été suffisamment défini et précisé. Le maître genevois soulignait, en effet, la primauté des manifestations sonores dans les processus de structuration de la langue, et l'introducteur du fonction-

nalisme linguistique en Espagne enchérissait sur ce principe en déclarant :

Les langues ont une manifestation normale et primaire qui est phonique (Emilio Alarcos Llorach, 1968, 519).

Il est notoire par ailleurs que le recours à l'oral a été posé en littérature comme une exigence aussi bien par des poètes que par des romanciers : Ronsard aimait « chanter et ouyr chanter ses vers », Flaubert allait dans son « gueuloir » prononcer sa prose à haute voix, et Daniel Pennac, après avoir indiqué que « Dostoïevski... ne se contentait pas de lire à voix haute, mais... écrivait à haute voix », s'insurgeait contre cette « étrange disparition que celle de la lecture à voix haute ».

Le principe d'oralité des langues est ainsi généralement admis aujourd'hui, et en linguistique acquiert valeur axiomatique. L'argumentaire qui le fonde est connu¹, mais les relations que l'oralité entretient avec l'écriture, les applications qui en ont été faites dans l'enseignement-apprentissage des langues, la spécificité des processus d'émergence, structuration et engrammation de la parole, et finalement les implications pédagogiques de la nature de l'oralité dans l'activité langagière demandent à être explicitées et précisées. Les développements du numérique accentuent encore cette exigence méthodologique.

1. L'oralité et le « prestige de l'écriture »

De fait, la primauté de l'oral dans l'activité langagière a été reconnue dans les études sur les langues bien avant d'être revendiquée par la linguistique contemporaine. Déjà en 1492, dans sa *Gramática sobre la lengua castellana*, première grammaire d'une langue romane, on pouvait lire :

No es otra cosa la letra, sino figura por la cual se representa la voz [sic] et pronunciación (Nebrija, [1492] 2011, 30).

La formule n'est pas seulement descriptive et a davantage d'implications qu'il n'y paraît. En effet, la « prononciation » est posée comme l'authentique objet linguistique, et l'écriture n'en est que la représentation (« no es otra cosa ») ; autrement dit, la valeur originelle des

¹ Les arguments sont à la fois d'ordre biologique, psychogénétique, ethnolinguistique, sociolinguistique, didactique ou relatifs à la pathologie de la parole et à la linguistique diachronique (cf. entre autres Murillo 2002 et 2015).

langues, leur essence, n'est pas dans l'écriture. La conception de l'oral ainsi posée renvoie au plan de l'épistémologie linguistique.

Jean-Jacques Rousseau reprenait presque mot pour mot la formule d'Antonio de Nebrija, mais constatait déjà que dans la pratique sociale l'oral était relégué au second plan, et ses réflexions ne manquent pas d'intérêt pour ce qui est des implications didactiques :

L'écriture n'est que la représentation de la parole, il est bizarre qu'on donne plus de soin à déterminer l'image que l'objet (J.-J. Rousseau, [1781] 1961, 1 252).

Ferdinand de Saussure relevait également le « prestige [*sic*] de l'écriture » dans « les opinions courantes qu'on colporte sur la langue » et s'interrogeait sur les « causes de son ascendant sur la forme parlée » et sur le fait que :

Chez la plupart des individus, [...] l'image graphique finit par s'imposer aux dépens du son (F. de Saussure, [1916] 1976, 46-47).

La question est toujours d'autant plus pertinente que, tout en présentant des arguments qui fondent la primauté de l'oral, l'auteur du *Cours de Linguistique Générale* ajoutait :

[...] il est impossible de faire abstraction d'un procédé [i.e., l'écriture] par lequel la langue est sans cesse figurée (Ibid., 44)².

Et le maître genevois avançait les « raisons » qui expliqueraient le « prestige de l'écriture », « raisons » qui sont toujours au moins latentes chez tout un chacun, donc aussi bien chez l'apprenant que... chez les enseignants, et qui ont conduit, de fait, à considérer que l'on peut caractériser les manifestations phoniques des langues au moyen de représentations écrites. Cette option méthodologique a des implications évidentes au plan didactique et requiert une analyse fine des relations entre le flux phonique et sa représentation.

² Les récentes avancées technologiques dans l'analyse du signal sonore, en pathologie de la parole et de l'audition ou en psycholinguistique et en neurolinguistique, tendent d'ailleurs, en révélant la complexité du signal phonique, d'une part à fournir des arguments aux méthodologues qui proposent de surseoir à l'introduction de l'écrit – i.e., de la *scription* (J. Peytard) – mais tendent également à conforter les démarches qui préconisent la prise en compte de l'écriture dans l'enseignement, même à des débutants. La réflexion critique sur ces procédures s'avère d'autant plus nécessaire.

1.1 Les représentations écrites des manifestations phoniques : conception et emploi

La création de l'alphabet phonétique international (A.P.I.) en 1886 et les actualisations qui ont suivi entendaient ainsi transcrire l'oral le plus fidèlement possible. Mais la représentation écrite d'un segment du flux phonique, aussi précise soit-elle, ne peut pas être prise pour l'objet phonique représenté, ni même considérée comme équivalente à cet objet : toute transcription phonétique implique une pré-catégorisation des segments du *continuum* du flux phonique³.

Envisagé comme une image plus ou moins symbolique du son, l'A.P.I., créé pour pouvoir avoir une représentation écrite de l'oral, s'est avéré un outil utile sinon nécessaire, par exemple, dans les enquêtes dialectologiques ou pour « représenter » les « prononciations déviantes », soit dans l'enseignement-apprentissage des langues soit en pathologie de la parole. Mais dans la pratique pédagogique, la plupart du temps, l'A.P.I. est utilisé pour « caractériser » l'oral – la « norme » – à partir de l'écrit. Les « dictionnaires de prononciation » sont un exemple patent de la démarche qui consiste à « partir de l'écrit », et il n'est pas exceptionnel que des matériels pédagogiques proposent d'enseigner l'oral à partir de l'écrit. Certaines approches didactiques préconisent même d'une manière explicite de retrouver la « forme orale » à partir de la « forme écrite » (cf. Wioland, 1991).

Du point de vue méthodologique, la démarche qui consiste à « partir de l'écrit pour retrouver l'oral » implique que l'on présuppose l'existence de relations bijectives entre l'oral et la représentation écrite⁴, et cette conception *nolens volens* préside déjà, dans une large mesure, à la

³ Louis Hjelmslev, sur le schéma qu'il proposait pour représenter la face *expression* du signe linguistique, situait d'ailleurs la phonétique à cheval entre les deux niveaux *forme* et *substance*. Eugenio Coseriu, pour sa part, attribuait aux variantes phonématiques résultant de phénomènes de métaphonie (assimilations, « harmonisation vocalique », etc.) un statut proche de la catégorisation phonologique, puisqu'il s'agit de « réalisations obligées » (fr. "absolu" → [apsoly] ; esp. "desde" → [dezde], etc.).

⁴ Toutes proportions gardées, la boutade de Magritte représentant une pipe et donnant pour titre au tableau « Ceci n'est pas une pipe » peut être appliquée à la graphie d'un phonème : un « a » (graphique) doit être compris comme *n'étant pas* un [a] (phonique), et même cette transcription phonétique implique, comme cela a été souligné plus haut, une précatégorisation de la matière du flux sonore.

création même de l'alphabet phonétique international⁵. Au demeurant, cette « relation bijective » entre l'écrit et l'oral est également latente dans la plupart des analyses et descriptions des langues (sauf, évidemment, pour l'analyse de la matière phonique) qui se réclament des grands courants de la linguistique tout au long du XX^e siècle et jusqu'à nos jours, et qui peuvent s'appliquer aussi bien à la langue écrite qu'à la langue orale.

Les travaux sur les « bigrammaires » en Afrique s'inscrivent également dans cette perspective (cf. Blanchet et Chardenet, 2014).

Il en est de même pour ce qui est des études comparées de la *langue parlée* et de la *langue écrite* en didactique : sont relevées et systématisées les différences pour ce qui est des marques morphologiques, syntaxiques ou textuelles, mais dans la plupart des cas ces analyses ne traduisent pas la spécificité de l'oralité et réduisent les différences à un changement de code qui d'ailleurs, le plus souvent, adopte comme référence de base la « forme écrite »⁶.

⁵ Comme il vient d'être indiqué, l'API est un outil utile et à bien des égards nécessaire, mais du point de vue didactique, il y a lieu de souligner qu'il s'agit d'une représentation doublement réductrice qui ne tient pas compte de la variabilité phonique inhérente à la parole, et ce à la fois au plan intralinguistique (alors qu'on ne réalise jamais de la même manière un son, toutes les réalisations par exemple de « ou » sont représentées par le graphème [u] en transcription phonétique du français, quel que soit l'entourage, l'accent ou la situation sur la courbe intonative), et au plan interlinguistique. Or un [a] en espagnol est différent d'un [a] français à la fois selon l'entourage consonantique, selon le degré accentuel de la syllabe qui le contient ou du type de syllabe (ouverte / fermée) : l'évolution diachronique du latin au français et à l'espagnol témoigne suffisamment de ces différences (lat. *mater* > fr. *mère* / esp. *madre* ; lat. *camisia* > fr. *chemise* / esp. *camisa*). En synchronie, l'exemple de l'arabe traduit d'une manière patente que dans la graphie (française) « Mohamed », « o » et « e » représentent des sons que les arabophones perçoivent comme « u » et « a », respectivement.

⁶ Quoique ces pratiques ne soient pas généralisées, certaines approches présentent, par exemple, la spécificité des marques grammaticales de l'oral. Ainsi relève-t-on que la marque la plus courante du pluriel de l'assiette nominale consiste à adjoindre la lettre « s » à la forme du singulier (*le cahier* → *les cahiers* ; *le petit enfant* → *les petits enfants*), alors qu'à l'oral la marque du pluriel est une liaison (lɛptitɑ̃fɑ̃ → lɛptizɑ̃fɑ̃) et/ou une variation de timbre de l'actualisateur (*le* [lə] → *les* [le]) ; pour la relation masculin-féminin, à l'écrit on partira du masculin auquel la règle dicte l'adjonction d'un « e » (*fort* → *forte*), à l'oral, au contraire, on propose de partir du féminin auquel on supprime la consonne finale (fɔ:ʀt → fɔ:ʀ). Ce changement de code ne rend pas compte, par exemple, de la redondance d'une marque de nombre ou de genre (" les enfants chantent " ☞ [lezɑ̃fɑ̃ʃɑ̃t]), ou des différents moyens de traduire la solidarité syntaxique (redondance de la marque à l'écrit, intonation à l'oral), etc.

1.2 Statut de l'oral et de l'écrit dans l'activité langagière

Que ce soit dans l'emploi le plus usuel de l'A.P.I. ou dans les pratiques qui préconisent de « partir de l'écrit pour retrouver l'oral », ces deux manifestations des langues sont censées être indépendantes l'une de l'autre, et dans le meilleur des cas elles sont mises sur un pied d'égalité pour ce qui est de leur statut dans l'activité langagière. Dans la pratique, la démarche revient à faire l'impasse sur les caractéristiques essentielles qui différencient l'écrit de l'oral et ce faisant, *a fortiori*, à ne pas prendre en compte la relation de dépendance de l'écrit par rapport à l'oral.

Or les relations qu'entretiennent la manifestation écrite des langues – leur *scription*⁷ – et leur manifestation phonique s'avèrent essentielles dans les analyses et procédures à visée didactique, dans la mesure, en particulier, où elles impliquent la temporalité et l'ordre de succession des séquences pédagogiques. À cet égard, la remarque péremptoire de Saussure ne peut pas ne pas être prise en compte :

Langue et écriture sont deux systèmes de signes distincts ; l'unique raison d'être du second est de représenter le premier ; l'objet linguistique n'est pas défini par la combinaison du mot écrit et du mot parlé ; ce dernier constitue à lui seul cet objet. Mais [...] on en vient à donner autant et plus d'importance à la représentation du signe vocal qu'à ce signe lui-même (GLG, 45 [94]).

Cette relation de dépendance de l'écrit par rapport à l'oral se vérifie dans l'argumentaire qui fonde la primauté de l'oral (cf. *supra* note 1), mais aussi en pathologie de la parole (potentiels évoqués, dyslexie, etc.), dans les lapsus orthographiques (confusions "b" ≈ "v" des hispanophones, "é" ≈ "-er" ≈ "ai" ≈ "-ez"... des francophones, etc.) ou dans les recherches qui utilisent l'imagerie cérébrale.

Même si, dans nombre d'analyses, la primauté de la langue parlée dans l'activité langagière n'est pas posée en tant que telle, le principe est aujourd'hui unanimement reconnu : à l'instar des études linguistiques *stricto sensu*, les recherches et applications pratiques concernant les différents champs afférents au langage revendiquent le principe de

⁷ Nous utilisons, à la suite de Jean Peytard, le terme de *scription* au lieu du terme *écriture* car *écriture* prend souvent le sens large d'élaboration du discours (cf. Barthes, *Le degré zéro de l'écriture*).

primauté de l'oralité, et l'appliquent dans les domaines les plus divers – administration, psychologie et psychanalyse, etc.⁸

Il n'en reste pas moins que l'écriture conserve toujours son prestige, et ce de par son importance sociétale (scolarisation, communications administratives, documents faisant foi, littérature, etc.), mais également de par son rôle au plan psychopédagogique dans l'activité langagière, en particulier dans l'enseignement-apprentissage des langues étrangères.

Étant donné les relations traditionnellement ancillaires de la didactique des langues par rapport à la linguistique⁹, il y a lieu, dès lors, de s'interroger sur la place relative de l'oral et de l'écrit dans la pratique didactique. Les matériels pédagogiques existant sur le marché et les orientations d'emploi des auteurs concernant la caractérisation de l'oralité et des manifestations phoniques constituent à cet effet des références significatives.

2. Oralité et pratiques didactiques : quelques jalons

Les premiers matériels pédagogiques conçus pour l'enseignement-apprentissage des langues qui postulent l'exigence de prendre en compte l'oral et reconnaissent l'antériorité et primauté de la langue parlée dans l'activité langagière datent des années qui ont suivi la Deuxième Guerre Mondiale. Ces applications bénéficièrent d'une conjoncture doublement favorable : d'une part, l'enseignement-apprentissage des langues est alors envisagé comme un facteur de paix et conçu en vue de contacts présents, et d'autre part, les avancées technologiques rendent

⁸ Il est par exemple symptomatique à cet égard que la cure psychanalytique soit essentiellement orale, ou que dans un tribunal l'intervention d'un témoin soit requise en langue parlée et ne soit pas admise par écrit.

Par ailleurs, les champs de recherche où se vérifie la primauté de l'oral dans l'activité langagière ont connu des avancées notoires au cours des dernières décennies : généralisation des implants cochléaires, investigations qui recourent aux potentiels évoqués ou à l'imagerie cérébrale, travaux concernant la parole pathologique, rééducation en orthophonie/logopédie, analyse du processus de structuration phonologique de la matière phonique, développements technologiques dans l'analyse du signal sonore, etc.

⁹ La création d'*exercices structuraux* à l'époque où le *structuralisme* s'impose est à cet égard plus qu'une anecdote. La dénomination même de « linguistique appliquée » utilisée dans de nombreuses publications et institutions pour désigner les recherches et procédures concernant l'enseignement-apprentissage des langues est également révélatrice.

accessibles au grand public – et aux enseignants, il va sans dire – les appareils d'enregistrement du son. Les programmes de formation élaborés à l'époque reçoivent l'appui des institutions nationales et internationales (l'UNESCO), et les matériels d'enseignement créés durant ces années incluent des textes oraux dûment enregistrés sur bande magnétique.

L'exigence de donner la priorité à la langue orale se fonde d'ailleurs non seulement sur des raisons ayant trait à la conception théorique et à la naturalité de l'activité langagière (cf. *supra* note 1), mais également sur des raisons sociopolitiques : après l'indépendance des anciennes colonies, anglaises et françaises principalement, qui se produit au milieu du XX^e siècle, la langue est utilisée par les anciennes puissances coloniales comme un moyen efficace pour manifester leur présence et exercer leur influence, et les importants mouvements migratoires de l'époque requièrent que les nouveaux arrivants puissent communiquer oralement le plus tôt possible.

Des organismes spécifiques sont créés – dont le CRÉDIF, en France, est l'exemple le plus représentatif – afin d'élaborer ces matériels d'enseignement, en particulier les méthodes désignées « audio-visuelles », et de former des enseignants aptes à en assurer l'utilisation efficace¹⁰.

2.1 Les méthodes audio-visuelles

Les maisons d'édition, après quelques hésitations sur leur viabilité et rentabilité économique, appuient ces initiatives et publient des « Cours » qui entendent répondre à ces orientations méthodologiques.

Comme indique leur dénomination, les méthodes audio-visuelles associent l'image et le son. Les images illustrent plus ou moins globalement les messages prononcés – généralement les répliques d'un dialogue – et/ou les situations de communication (décor, personnages, etc.), afin de traduire la naturalité des dialogues enregistrés (certains sont d'ailleurs « authentiques ») qui s'y seraient déroulés. Elles ont pour but d'aider l'élève à « comprendre »¹¹, c'est-à-dire à saisir le *sens* des messages

¹⁰ Dans certains cas, d'ailleurs, les dites *Méthodes Audio-Visuelles* n'étaient livrées à un établissement d'enseignement que si celui-ci avait au moins un professeur dûment formé à leur utilisation.

¹¹ « Seule, l'image rend possible la présentation et compréhension de multiples situations » *Voix et Images de France, Livre du maître* (1971, 24). À une deuxième époque on conseillera de « faire

enregistrés, et également à en faciliter la mémorisation. Les situations sont censées pouvoir être simulées en classe et les dialogues appris et reproduits. La connaissance de la langue orale est posée comme un objectif principal.

Quoique ces objectifs leur soient communs, les différents matériels pédagogiques « audiovisuels » existant sur le marché à l'époque sont loin de se fonder sur des principes identiques concernant le rôle respectif des illustrations et des enregistrements sonores. Les « méthodes » qui se réclament de ces orientations peuvent être regroupées en deux grandes catégories (cf. Porquier & Vives, 1974) selon que les images visent à représenter une *situation de discours* ou à *transcoder* (« illustrer ») les énoncés prononcés. Ces deux approches traduisent en fait deux conceptions radicalement différentes de la fonction de la « trace visuelle » des échantillons de parole présentés pour enclencher le processus d'apprentissage, et ce faisant deux conceptions très différentes des manifestations orales des langues et de leur statut dans l'activité langagière.

Les images qui visent à *transcoder l'énoncé* – à « l'illustrer » – introduisent dans le processus d'enseignement-apprentissage une phase intermédiaire entre le flux phonique et la représentation écrite. Que ce soit au moyen du tableau de feutre (*Frère Jacques*) ou par la représentation sur une vignette projetée ou non (Capelle, G. 1969, *La France en Direct* ; Mauger, Bruézière 1971, *Le français et la vie* ; Grandía, 1970, *Sonimage*, etc.), ces méthodes où l'image est une illustration de l'énoncé enregistré manifestent une conception essentiellement analytique de l'oralité qui s'avère particulièrement réductrice. L'utilisation d'éléments codés (« bulles » reliées au personnage qui produit l'énoncé, « ? » pour indiquer qu'il s'agit d'une question ou d'une interrogation, croix de St. André sur une partie ou la totalité de l'image pour symboliser une négation, etc.) ne tient pas compte de la variabilité de la langue parlée (en particulier de la modalisation et topicalisation des énoncés), de la valeur symbolique du langage, et, pour ce qui est de l'oral proprement dit, de la complexité du

repasser devant l'étudiant le film fixe seul, et à lui demander de retrouver le dialogue en présence des images. »

flux phonique que doit structurer l'apprenant¹². Le *sens* est circonscrit, réduit – et par-là même faussé – aux référents des vocables représentés, l'énoncé est segmenté en fonction des ancrages lexicaux détruisant ainsi la globalité du message et incitant à la traduction mot à mot « par image interposée » ; de plus, l'image laisse entendre que la relation spatiale traduit l'ordre de la séquence morphosyntaxique censé aller de soi (sujet + verbe + complément).

Pour ce qui est des méthodes où les images représentent la *situation* où se réalise la parole et qui parfois la motive (*Voix et Images de France, De vive Voix, Bonjour Line* pour le français, *Vida y Diálogos* pour l'espagnol, *All's Well* pour l'anglais, *In Bonn* pour l'allemand, ...), l'image contribue à l'accès au sens par les indications sur la fonction déictique, sur l'identité des personnages, les relations qu'ils entretiennent, parfois sur leur état affectif manifesté par la mimique et la gestualité¹³, mais les éléments visualisés ne conditionnent pas la saisie de la globalité du message, sa structure grammaticale et la composante prosodique : l'image « situationnelle » n'est pas une illustration du message, et la saisie que peut en faire l'apprenant en la « décodant dans sa langue maternelle » ne fausse pas la perception globale des énoncés en en indiquant *a priori* la segmentation en « mots ».

Parmi les différents matériels pédagogiques produits à l'époque, certaines méthodes ainsi élaborées et la démarche pédagogique proposée constituent de véritables jalons dans l'histoire de la didactique. Le processus d'apprentissage proposé était, en effet, analysé et ritualisé en cinq « moments »¹⁴ qui rendaient bien compte des opérations que

¹² Cette phase intermédiaire où le texte oral est transcodé constitue un véritable idéogramme. Le découpage en images de la suite des vocables de l'énoncé circonscrit la représentation aux lexèmes visualisables et omet évidemment les éléments grammaticaux, les solidarités syntagmatiques et les composantes prosodiques. Il s'agit, comme a montré Henri Besse, d'une pratique qui conduit inconsciemment l'apprenant à essayer de comprendre en traduisant dans sa langue par images interposées.

¹³ En général, une image est prévue pour chaque réplique mais ne l'illustre pas. Pierre Neveu, l'auteur des « images séquentielles » (P. Rivenc *dixit*) des cours produits par le CRÉDIF (Méthodes *Voix et Images de France, Bonjour Line...*), expliquait qu'il commençait par définir les relations spatiales, la distance entre les personnages, puis l'expression des visages en fonction du message prononcé, et les gestes ; les mains, seulement à la fin.

¹⁴ I.e., « présentation » (saisie globale de la situation : visualisation de l'ensemble de la scène + écoute synchronisée du dialogue avec un léger retard sur la présentation de l'image), « explication » (accès au sens), « travail phonétique et mémorisation » (maîtrise du signifiant),

requiert l'acquisition d'une langue seconde ou étrangère, mais ils étaient présentés d'une manière séquentielle qui, à l'évidence, est loin de correspondre au processus d'apprentissage réellement vécu par l'apprenant¹⁵.

Pour ce qui est des procédures didactiques, l'oral était l'objet premier du processus d'enseignement-apprentissage des langues : la langue orale était présentée avant l'écrit, sous forme de dialogues enregistrés à apprendre et à « exploiter » ensuite selon le protocole qui précisait les « moments de la classe de langue ». Certains des phénomènes de la langue parlée (métaphonies, intonation, expressivité, signes et mouvements à valeur sémiotique, etc.)¹⁶ étaient bien mis à profit pour faciliter la perception et structuration de la chaîne phonique mais seulement au niveau phonématique. Il n'y avait pas d'exercices de phonétique spécifiques : le travail phonétique était censé être réalisé dans les phases de répétition, de mémorisation ou de réemploi.

L'importance des éléments prosodiques était soulignée mais sans en indiquer les raisons, et sans en montrer non plus le statut dans l'activité langagière, laissant à l'enseignant l'initiative des procédures à appliquer. Ainsi le *Livre du maître* de la méthode *Voix et Images de France* ne donne comme procédures d'intervention sur la prononciation que la « répétition » et « l'imitation » :

... les élèves doivent répéter les « groupes sémantiques » [sic] en imitant parfaitement les groupes phonétiques [sic] enregistrés [...] C'est de cette imitation parfaite des groupes, au point de vue de la prononciation correcte des sons ... mais plus encore au point de vue du rythme et de l'intonation que

« systématisation » (économie grammaticale), « transposition-exploitation » (réemploi des énoncés dans des situations différentes).

¹⁵ L'ordre de présentation adopté a pu d'ailleurs faire l'objet de controverses même entre les auteurs de ces méthodes : le débat concernant l'ordre respectif des moments « explication » et « travail phonétique » est symptomatique du statut que l'on accordait à la manifestation phonique. Vu que, selon Rousselot, « dès que le sens apparaît à l'esprit, on néglige le son », certains méthodologues proposaient que le travail phonétique précède l'explication ; d'autres, pensant avec Bally que « la parole est avant tout porteuse de sens », prônaient, au contraire, que l'explication précède la correction phonétique. Il va sans dire qu'il s'agissait d'une fausse querelle : le travail sur le sens, même si celui-ci est illustré par une « image », ne peut se réaliser que par le truchement de la manifestation phonique des vocables, sur laquelle il n'est pas possible de faire l'impasse. Le seul débat possible concerne dès lors le degré d'exigence de correction dans la prononciation à chaque étape du processus d'apprentissage.

¹⁶ Ainsi, dans *Bonjour Line*, l'action métaphonique de [p] dans « le bol de Paul » facilite la réalisation occlusive de [b] dans « bol » ; dans *Voix et Images de France*, l'expressivité de « C'est haut ! » contribue à « fermer » le [o] final, etc.

dépend le succès de la méthode... (*Voix et Images de France, Livre du maître*, [1960] 1971, 21).

Les auteurs ajoutaient d'ailleurs « *Il ne faut jamais changer les intonations et les rythmes enregistrés* », indication qui n'est pas à proprement parler une procédure didactique mais une pétition de principe.

Il n'y a pas lieu, dans le cadre de cette contribution, d'évaluer l'efficacité pédagogique de ces différentes approches, mais pour nous en tenir à ce qui a trait à la mise en pratique du principe d'oralité au plan didactique, il ressort que l'enclenchement du processus d'enseignement-apprentissage se réalisait bien à partir d'échantillons oraux de parole, et l'objectif poursuivi était la maîtrise de la langue orale, mais les démarches suggérées ne rendaient pas compte de la spécificité de la langue parlée et donc des caractéristiques et procédures didactiques spécifiques qui doivent s'ensuivre¹⁷ ; elles ne résolvaient pas l'apparent dilemme entre l'exigence méthodologique d'accorder la primauté à l'oral d'une part, et d'autre part de prendre en compte la langue écrite dont « on ne peut pas se passer ».

Pour respecter la primauté de l'oral et aider à la mémorisation, certaines de ces méthodes « audio-visuelles » (cf. *infra* § 5.2) préconisaient que la langue écrite soit totalement exclue en début d'apprentissage durant environ une soixantaine d'heures, mais afin que l'apprenant ait une trace du travail réalisé en classe, il disposait de documents graphiques – parfois un « livre » – où étaient reproduites les images du document « audio-visuel » utilisé en classe. Chaque vignette représentait, pour chaque énoncé, *la situation* où la réplique était censée avoir été prononcée, mais sans texte écrit aucun¹⁸.

¹⁷ Les *Guides pédagogiques* sont à cet égard révélateurs. La gestualité, par exemple, est parfois requise, mais dans sa valeur sémantique (orientation dans l'espace, déictique, imitation d'une forme géométrique, geste de refus ou d'acquiescement, etc.), qui « double » le contenu du message, mais non dans sa valeur associée inhérente à la parole (rythme et/ou intonation). Ces « gestes sémantiques » ont fait d'ailleurs l'objet d'études systématiques (cf. Calbris et Porcher, 1989 ; Pavelin, 2002).

¹⁸ Ainsi, par exemple, la méthode *Voix et images de France (VIF)*, outre les bandes magnétiques (dialogues enregistrés), les films fixes (images situationnelles) et le *livre du professeur* (textes en orthographe d'usage et orientations pédagogiques), comprenait un *livre de l'élève* où étaient seulement reproduites les images projetées en classe sans aucune indication écrite.

Grâce à ces images, l'élève pourra faire seul des exercices de révision en s'efforçant de retrouver les éléments du dialogue qui accompagnent les films et qui ont été étudiés avec le professeur (Guberina P. et Rivenc P. (dir.), s.d., AVERTISSEMENT, *Voix et images de France – livre de l'élève*).

Ce n'est qu'au bout de cette soixantaine d'heures exclusivement consacrées à l'oral que des procédures de « passage à l'écrit » étaient proposées.

Le recours au « livre d'images » fut cependant abandonné. L'apprenant se trouvait dès lors dépourvu d'une trace quelconque susceptible de l'aider à mémoriser tant que l'écriture n'avait pas été introduite.

Surgirent alors des « cours », qualifiés pour l'heure de « communicatifs » – et plus récemment méthodes « actionnelles » ou « actionnistes » – qui proposaient des dialogues enregistrés et transcrits en orthographe d'usage, les textes étant accompagnés d'une ou plusieurs illustrations. C'est toujours d'ailleurs la maquette la plus fréquente dans les matériels didactiques actuellement sur le marché.

2.2 *Les méthodes « communicatives » et « actionnelles »*

Au vrai, les principes revendiqués par les auteurs de ces méthodes et/ou par les méthodologues qui s'en réclament (cf. entre autres Puren Ch., 2004, 2007 ; Piccardo E., 2014) – i.e., exigence de concevoir la *communication* comme une interaction, et l'enseignement-apprentissage comme une action, en primant les *actes de parole* et les « projets d'activités » – étaient déjà latents, et souvent formulés d'une manière explicite, dans les orientations *structuro-globales* (cf. Rivenc, Guberina, 1960 ; R. Renard, 1965, 1976 ; A. Rojo Sastre 1968 ; H. Besse, 1985 ; S. De Vriendt, 1997 ; P. Rivenc, 2000, 2003 ; J. Cortès, 2011...).

Ces *Cours* « communicatifs » ou « actionnels » proposent des dialogues censés être « naturels » (au demeurant comme dans les Méthodes Audio-Visuelles et comme il est préconisé dans la problématique structuro-globale), et des documents présentés comme « authentiques » (en général des textes extraits des médias) transcrits en orthographe d'usage et accompagnés généralement d'une ou deux illustrations plus ou moins globales de la situation ou de la thématique du

texte¹⁹. Les *corpus* ainsi constitués sont enregistrés comme échantillons de parole à reproduire par l'apprenant comme « enclencheurs » du processus d'apprentissage.

Indépendamment de l'importance donnée à la « dramatisation » de ces textes ou à la lecture, de la primauté donnée ou non au sens, de l'importance accordée à la gestuelle à valeur sémantique (cf. *supra* note 13), pour ce qui est de l'*oralité*, la démarche la plus généralisée dans ces matériels didactiques circonscrivait la plupart du temps l'objet d'analyse et d'enseignement-apprentissage au flux sonore émis, conçu comme une chaîne de segments sonores, comme un « collier » (Boë, 1997) d'unités discrètes.

Le travail ciblé sur l'oral s'articule généralement selon trois approches :

- écoute et répétition des enregistrements,
- présentation (plus ou moins explicite) de règles d'orthoépie,
- réalisation d'exercices de correction phonétique.

Les procédures d'écoute et de répétition sont posées comme allant de soi, et laissées à l'initiative de l'enseignant.

Les règles d'orthoépie – comme les règles de grammaire ou d'orthographe – présentent la « norme », et décrivent un état de langue le plus souvent à partir de l'écrit. Les indications sur ladite norme

¹⁹ Ces matériels d'enseignement-apprentissage privilégient les *actes de parole* (à la suite des travaux de Austin et de Searle) : *Un Niveau Seuil* (français), *The Threshold Level* (anglais) furent les deux premiers recueils (1976) produits par le Conseil de l'Europe à l'usage des enseignants et des auteurs de Méthodes qui se fondent sur ces analyses du langage. Le *Cadre Européen Commun de Référence* s'inscrit dans cette perspective en élargissant le domaine des aptitudes requises : compréhension et production orale, compréhension et production écrite, médiation.

Mais il y lieu de souligner que ces recueils auxquels renvoient les auteurs de *Cours* qui s'auto-définissent « communicatifs », ou « actionnistes », ne préjugent pas de la méthodologie à appliquer et *a fortiori* ne se veulent pas des matériaux d'enseignement-apprentissage à employer « tels quels » mais des outils pour définir des contenus et des niveaux de compétence. Ils ne spécifient donc pas non plus ce en quoi l'oral a la primauté dans l'activité de langage – y compris écrite – et *a fortiori* les procédures qui doivent s'ensuivre. Ces démarches au demeurant négligent les *fonctions* essentielles que réalisent les langues pour résoudre les « problèmes » qui se posent dans la *praxis humaine* (cf. la linguistique typologique, les travaux de Seiler, Lazard, etc.). La présentation qu'en fait E. Piccardo (2014) est significative : « conscience phonétique » et « compétence phonologique » sont mentionnées comme objectifs, mais aucune indication n'est donnée sur l'intervention didactique à réaliser à cette fin.

phonétique se réfèrent au niveau phonématique (timbre [e]/[ɛ]/[ə]..., nasalité, etc.) ou à des phénomènes phonotactiques (liaisons, [ə] caduc, assimilations, timbre en syllabe ouverte/fermée, allongements, loi des 3 consonnes, etc.). À l'évidence, cependant, les règles d'orthoépie peuvent au mieux contribuer à fixer des acquis antérieurs, mais non à structurer le flux phonique ; elles sont d'ailleurs appliquées de manière inconsciente et spontanée par les locuteurs natifs²⁰. Ce ne sont pas, à proprement parler, des « méthodes d'accès à l'oral » ou de « correction phonétique » mais, dans le meilleur des cas, des indications qui présupposent la maîtrise des sons prononcés.

L'application des règles d'orthoépie s'exemplifie en général par la présentation de « paires minimales » utilisées pour la réalisation de tâches de perception des « oppositions phonologiques » ; mais cette pratique, si elle peut aider à stimuler la *sensibilité auditive* et contribuer à fixer des acquis antérieurs, s'avère peu efficace pour structurer le flux phonique de la langue : les limites qu'impose le « crible phonologique » sont rédhibitoires à cet égard.

Pour ce qui est des exercices d'accès au signifiant et de correction phonétique, si, comme il a été rappelé plus haut, les méthodes audiovisuelles ne proposaient pas d'exercices spécifiques (ces exercices étaient censés être pratiqués lors de la répétition et mémorisation des dialogues, les procédures à appliquer étant laissées à l'initiative de l'enseignant), dans les méthodes « communicatives » et/ou « actionnelles », les exercices spécifiques de correction phonétique – lorsqu'ils sont prévus²¹ – traduisent une conception physicaliste des manifestations phoniques. Ils présentent de ce point de vue un intérêt tout particulier pour définir *a contrario* le statut de l'oralité au plan didactique.

²⁰ Dans leur langue propre, les locuteurs sans formation phonétique appliquent en effet ces « règles d'orthoépie » inconsciemment et d'une manière spontanée (il est bien connu que la règle d'orthoépie – comme toute règle, de grammaire ou autre – explicite une pratique et non l'inverse). Ainsi la « règle des trois consonnes » formulée par M. Grammont était appliquée d'une manière inconsciente par les locuteurs francophones avant que le phonéticien ne la relève : "vous demandez" → [vudmãde] ⇔ "vous le demandez" [vuldəməde].

²¹ Il est arrivé toutefois que l'engouement pour les approches communicatives, et parfois une lecture réductrice de la notion même de *communication*, ont conduit certains méthodologues et auteurs de *Cours* à ne pas prévoir un travail spécifique sur la prononciation (cf. *Escales*, Clé International).

2.3 La conception physicaliste des manifestations phoniques de la parole

La plupart du temps, dans ces *Cours* et *Méthodes communicatifs et/ou actionnels*, les manifestations orales des langues sont identifiées par leurs « caractéristiques articulatoires » exclusivement, autrement dit par la position respective des organes phonatoires (langue, lèvres, palais, dents, voile du palais, etc.) supposée être commune à tous les locuteurs et ne pas varier durant tout le « segment » du flux phonique correspondant à un phonème. Les phénomènes de compensation (par exemple, moindre ouverture pour compenser une position trop antérieure de la langue) ne sont pas pris en compte, et le processus de structuration de la matière phonique en phonèmes et intonèmes n'est pas envisagé : la dichotomie *ergon-energeia* (Humboldt), essentielle du point de vue pédagogique, est ignorée. Au vrai, la conception de l'oral qui sous-tend ces procédures rejoint la conception mécaniciste de l'être humain qui aux XVII^e et XVIII^e siècles avait conduit à concevoir *l'homme machine* (La Mettrie), et à construire des *machines parlantes* (Von Kempelen).

La caractérisation acoustique des sons de la parole n'est généralement pas abordée dans ces matériels pédagogiques (la lecture des documents requiert une formation spécialisée), et les applications dans les TICE (cf. entre autres, *Lab de Langues*) sont loin d'être satisfaisantes (problèmes concernant la reconnaissance de la voix de l'apprenant, la caractérisation acoustique des références standard, la comparaison des réalisations de l'apprenant avec les « modèles », afin d'évaluer le degré de déviance, et de définir l'importance relative des différentes « erreurs »).

L'analyse critique des procédures de « correction phonétique » basées sur ces conceptions physicalistes de l'oral (articulatoires ou acoustiques) a fait l'objet de travaux importants qui en ont montré les limites (Renard, 1979, 2002, 2015), voire les effets négatifs (blocages, non-respect des phénomènes idiosyncrasiques de phonétique syntaxique ou combinatoire, etc.). Mais la question essentielle qui se pose concernant la conception de l'oralité qui sous-tend ces exercices de correction phonétique a trait davantage encore au plan de l'épistémologie linguistique et du comportement psycholinguistique qu'aux procédures d'intervention didactique. La démarche, en effet, fait l'impasse sur des phénomènes consubstantiels à l'oralité et qui s'avèrent déterminants dans le processus audio-phonatoire : les composantes prosodiques (variation mélodique,

structure rythmique...) ne sont généralement pas prises en compte – ou partiellement « indiquées » sur des documents écrits – alors qu'elles sont inhérentes à toute activité langagière (Jakobson, 1980) et jouent un rôle essentiel dans la structuration syntaxique de la langue orale. Ainsi, dans la pratique, la structure syntaxique est circonscrite au type SVO, et des phénomènes aussi importants que la variation tonale des constructions clivées ou disloquées (pourtant très fréquentes) ne sont pas pris en compte.

Pour ce qui est des segments sonores correspondant à la réalisation d'un phonème, ils sont censés pouvoir être caractérisés par des propriétés (intensité, tonalité et/ou timbre) soi-disant constantes (ce qui équivaldrait à la « tenue du son » selon Paul Menzerath et Armando Lacerda). Or il n'en est rien : la fréquence du F_0 , la distribution de l'énergie sonore sur le spectre fréquentiel, et l'intensité globale du flux phonique varient constamment. Il s'est avéré ainsi, dans les essais de synthèse de la parole par « montage de bandes magnétiques », qu'un énoncé formé par une succession de segments sans variations est incompréhensible :

... la synthèse de la parole par montage de bandes magnétiques (C. Harris, 1953) avec des éléments de la taille d'un phonème (ce que l'auteur appelle les *building blocks*) s'était révélée complètement inintelligible (Boë, 1997, 32).

Par ailleurs, dans ces analyses et procédures qui circonscrivent les descriptions à la position des organes du tractus phonatoire – censée être fixe – sont également omis des phénomènes moins apparents mais non moins caractéristiques de la spécificité des manifestations phoniques des différentes langues, spécificité dont les études de diachronie²², de dialectologie ou en pathologie de la parole ont montré l'importance, et que la didactique des langues se doit de prendre en compte : le tonus musculaire²³, et les phénomènes de phonétique distributionnelle ou de métaphonie ne peuvent pas être ignorés.

²² Par exemple, le phénomène de *lénition celtique* ne s'applique pas pour la consonne initiale, i.e. la plus tendue, qui résiste à ce phénomène de lénition (*ripa* → *riba* / *rive* ; $p > b > v$; mais : *pauper* > *pauvre* / *pobre*), ce qui explique que le nombre des lemmes dans les dictionnaires des langues latines soit pratiquement le même, etc.

²³ Le fait que le tonus musculaire et l'activité de l'orbiculaire des lèvres révélée par les analyses électro-myographiques soient supérieurs pour la prononciation d'un son sourd dorso-dental en français (par exemple [s], pour la description duquel l'activité de l'orbiculaire n'est pas prise en compte dans les analyses articulatoires) que pour la prononciation d'un son sourd bilabial espagnol (par exemple [p] pour lequel l'activité de l'orbiculaire est censé être déterminante) montre à l'évidence que la didactique de la prononciation ne peut se satisfaire des descriptions physicalistes.

Les formes métaphoniques, et en général les « règles » de phonétique syntactique²⁴ sont particulièrement révélatrices de phénomènes que l'orthoépique décrit mais dont on ne relève généralement pas les implications psychomotrices et, en particulier, la séquenciation du processus d'émergence de la parole : à l'évidence, si un son que je n'ai pas encore prononcé modifie le timbre ou le caractère occlusif de celui qui est en train d'être produit, c'est que la séquence phonique est programmée avant d'être réalisée (cf. note 24). Au plan didactique, il ne suffit pas de décrire ces phénomènes comme des variations « en soi », ou de les caractériser comme des « réalisations obligées » (Coseriu) : la prise en compte des facteurs qui déterminent ces variations oblige à intervenir *en amont du processus articulatoire* proprement dit, et partant à adopter des procédures qui ne se circonscrivent pas à la « mécanique du tractus phonatoire ».

Les descriptions du flux phonique articulatoires ou acoustiques, qui se veulent « naturelles » et « conformes à la réalité », traduisent de fait une conception des langues qui réduit l'oralité à l'activité – consciente ou non, cette distinction ici ne s'avère pas pertinente²⁵ – des organes du tractus phonatoire (positions et mouvements articulatoires), alors que ladite activité s'amorce avant l'ébranlement des organes phonatoires qui produit chaque son. Elles limitent de ce fait les possibilités d'intervention pédagogique, laquelle doit opérer sur les facteurs qui déclenchent et

²⁴ Cf., par exemple, les phénomènes de métaphonie et d'« harmonisation vocalique » ("tête" → [tɛt] ⇨ "têtu" → [tɛty] : le [y] de la deuxième syllabe « non encore prononcée » fait que [ɛ] de [tɛt] ⇨ [e] dans [tɛty]) ; la réalisation des occlusives sonores intervocaliques comme approximantes en espagnol (« cabo » : "-b-" → [-β-] ; « hada » : "-d-" → [-ð-] ; « mago » : "-g-" → [-ɣ-]) ; les assimilations dites « régressives » (i.e., l'incidence d'un son non encore prononcé sur le son qui est produit (dans "absolu", [b] ⇨ [p], assourdi par [s] ; en espagnol "asno" prononcé [azno] : sonorisation de /s/ par l'influence de [n] non encore prononcé) ; etc.

²⁵ La distinction entre l'intervention inconsciente ou indirecte sur le tractus phonatoire et l'intervention consciente, délibérée et explicitée dépendait d'ailleurs la plupart du temps de la possibilité ou non de contrôler la motricité : une spatule aidait à relever la langue pour obtenir une palatale, un miroir à arrondir les lèvres, mais pour obtenir des voyelles nasales, la légère chiquenaude sur le menton afin de faire baisser l'uvule se pratiquait sans « explications » (procédure préconisée par Suzanne Mercier dans les années 1965-1970). Par ailleurs, même si un son peut être effectivement obtenu par ces procédés (par exemple : « pour prononcer [y], placer la langue comme pour [i] et arrondir les lèvres comme pour [u] »), dès que l'on revient à une élocution normale, spontanée, il n'est pas exceptionnel que l'erreur réapparaisse. Cette distinction entre comportement conscient et délibéré et comportement spontané se retrouve d'ailleurs en grammaire : il n'est pas exceptionnel qu'un apprenant répète correctement la règle et l'applique dans un exercice *ad hoc*, mais cesse de l'appliquer dans une élocution spontanée.

régissent le processus, y compris les phénomènes de perception qui s'avèrent déterminants (cf. *infra* § 5.2).

Les variables utilisées pour l'analyse des manifestations phoniques « standard » ou déviantes doivent donc permettre de caractériser l'oralité en tant qu'activité structurante du flux phonique aussi bien en perception qu'en production. Au demeurant ces deux pôles du circuit de la parole sont loin d'être autonomes : toute activité d'émission est à la fois perception. Le processus est d'ailleurs désigné *audio-phonatoire* (R. Renard), et le circuit de la parole *boucle audio-phonatoire* (A. Prochiantz).

[...] le cortex moteur contribue de façon spécifique à la perception de la parole (D'Ausilio, et al., 2009). Ces résultats viennent confirmer ceux obtenus en 2006 par Pulvermüller et collègues (D. Méligne, 2011, 57).

Il ne suffit pas dès lors que les matériels d'enseignement-apprentissage des langues, et les procédures didactiques à appliquer, revendiquent l'oralité et proclament la primauté des manifestations phoniques des langues. Encore faut-il ne pas réduire celles-ci à une suite de sons ou à la mécanique articulatoire censée les produire. C'est la conception même du processus audio-phonatoire qui doit être précisée de manière à prendre en compte les facteurs – y compris la perception – qui opèrent non seulement pendant la production des sons mais *en amont*. La problématique renvoie ainsi au *statut de l'oralité* dans l'activité de parole.

3 Le statut de l'oralité au plan didactique

Le terme de *statut*, emprunté au domaine du droit, a connu de nombreuses acceptions y compris en linguistique. Pour ce qui concerne l'enseignement-apprentissage des langues, on peut poser que le terme *statut* se réfère, par analogie à certaines des définitions que mentionne le *Trésor de la Langue Française* dans ce domaine, en particulier en linguistique, à la *situation de fait dans laquelle se trouve l'oralité relativement à son usage en didactique des langues*, situation sur laquelle il y a lieu de s'interroger.

3.1 Parole et logique du vivant : le système verbo-tonal

En effet, la « situation de fait » de l'oralité n'est pas définie, en didactique des langues, par le principe selon lequel les manifestations normales et primaires des langues sont sonores (cf. *supra* § 2.3, et note 1), ni par le rappel de la manière dont lesdites manifestations phoniques ont

été prises en compte dans la plupart des *Cours de Langue*. Le statut de l'oralité n'est pas précisé non plus explicitement dans la Méthode Verbo-Tonale : certes cette méthode s'inscrit dans la perspective de la *Linguistique de la Parole* et propose des procédures (cf. *infra* § 5) qui ne se circonscrivent pas au tractus phonatoire, mais il apparaît cependant que, tout en ayant fait preuve d'efficacité, ces approches ne donnent pas des indications explicites concernant le statut de l'oralité dans le langage, ou dans le meilleur des cas en indiquent quelques éléments par des références *a contrario* à l'écrit.

Saussure, dont on a souvent cité la dernière phrase du *Cours de Linguistique Générale* pour justifier que les linguistes circonscrivent leurs analyses à la « langue » – et se cantonnent la plupart du temps à des descriptions formalistes des langues – envisageait de produire un *Cours* sur la *Linguistique de la Parole*, et déclarait déjà :

Rien n'entre dans la langue sans avoir été essayé dans la parole (C.L.G., 231).

Le constat saussurien se référait à la diachronie, mais la structuration de *la langue*, c'est-à-dire du *système* ou du *système de systèmes* (Renard) s'effectue également, et à plus forte raison, à partir de la *parole* dans l'enseignement-apprentissage des langues. Le fait que le projet saussurien, dont Charles Bally et Albert Sechehaye ont souligné l'importance²⁶, n'ait pas abouti explique pour une large part que les analyses linguistiques et les procédures didactiques les plus généralisées se circonscrivent à la « langue ». Or, à l'évidence, l'acquisition de la langue maternelle, et *a fortiori* tout enseignement-apprentissage d'une langue seconde ou étrangère, consiste précisément à réaliser cet « essayage » dans la parole par « approximations successives » (R. Renard). Dès lors, la *situation de fait dans laquelle se trouve l'oralité relativement à son usage en didactique des langues* se caractérise au premier chef par le fait que cet « essayage » s'effectue sur la langue orale.

D'ailleurs ces « approximations successives » consistent également à *engrammer* la parole, à la mémoriser. L'oralité est non seulement la

²⁶ « L'absence d'une « linguistique de la parole » [guillemets des auteurs] est plus sensible. Promise aux auditeurs [...], cette étude aurait eu sans doute une place d'honneur [...] cette promesse n'a pas été tenue. Nous nous sommes bornés à recueillir et à mettre en leur place naturelle les indications fugitives de ce programme à peine esquissé ; nous ne pouvions aller plus loin » (Bally et Séchehaye, Préface du C.L.G.).

« manifestation normale et première » des langues en tant qu'objet de communication, mais également leur manifestation mentale et corporelle. Et les erreurs d'orthographe témoignent assez de ce que l'image engrammée des langues est essentiellement phonique²⁷.

Il est notoire qu'une variation morphologique ou une « structure syntaxique » ne sont vraiment acquises que lorsqu'elles ont été automatisées oralement. Ainsi envisagée et située, l'oralité ne peut pas être réduite à un attribut de la parole, elle en est partie inhérente et consubstantielle : il n'y a pas de parole sans oralité (que celle-ci soit intérieure » – Hécaen et Angelergues, 1965 – ou extériorisée).

En deuxième lieu, « l'essayage de la parole », en agissant sur l'oralité, requiert l'intervention incontournable de l'être humain. Les développements technologiques permettent aujourd'hui la production de parole synthétique et la « reconnaissance automatique » de la parole, mais toutes ces avancées font l'impasse sur la phase d'émergence de la parole. Une machine peut saisir la parole et la transcrire ; elle peut « générer » des textes écrits plus ou moins codifiés à partir de données préalablement fournies (L. Danlos), et les traduire en sons : ce n'est pas de la parole au sens plein du terme, ce n'en est au mieux qu'un ersatz. Le statut de l'oralité au plan didactique requiert que les manifestations phoniques ne soient pas réduites au flux phonique produit par l'activité audio-phonatoire comme s'il s'agissait d'un processus mécanique. C'est l'ensemble du processus audio-phonatoire qui doit être pris en compte, y compris les facteurs qui interviennent à l'étape du circuit de la parole antérieure à la production dans le tractus phonatoire.

Une machine ne peut pas produire la langue orale naturelle ni nuancer le flux phonique en vue d'en optimiser la perception selon la réaction de l'apprenant (par définition idiosyncrasique). Les conclusions des recherches sur la perception sont à cet égard significatives : la

²⁷ En effet, même les représentations écrites requièrent la *mise en sons* préalable. Les références qui fondent et appuient ce principe sont notoires : pour corriger un élève de français qui n'écrit pas le « r » final d'un verbe du premier groupe à l'infinitif, l'instituteur lui suggère de le remplacer par un verbe en « -ir » (**Il va mangé* → *il va partir* → *Il va manger*) ; la phrase de Marcel Proust a été associée à sa respiration asthmatique ; la confusion entre graphèmes qui ont la même prononciation en espagnol (/b/ écrit « b » pour « v » et vice versa, ou /x/ écrit « j » au lieu de « g » ou vice versa, sont des erreurs courantes des hispanophones, et inconnues des francophones ou des germanophones...) ; les exemples pourraient être multipliés.

« réalité perceptive » des manifestations phoniques diffère notablement de la « réalité physique » desdites manifestations²⁸, et d'autre part la perception auditive varie selon qu'il s'agisse de sons non verbaux (perception tonale) ou de sons de la parole (perception verbale)²⁹. Ces conclusions permettent de poser que pour ce qui est de la communication parlée, et *a fortiori* en didactique des langues, les manifestations phoniques de la parole entretiennent des relations spécifiques à la fois au plan verbal et au plan tonal, et structurent ainsi un *système verbo-tonal*. L'intervention de l'être communiquant s'avère incontournable, et se réalise, par définition, selon la *logique du vivant* (F. Jacob, [1970], 2002). La *langue orale* – « la seule vivante » (Lacan) – répond à cette logique dès que s'amorce l'acte langagier, et durant tout le processus audio-phonatoire.

Le statut de l'oralité au plan didactique se doit dès lors d'être caractérisé à l'aide des principes de ladite *logique du vivant* (Murillo, 2003 ; Rivenc, 2010) qui s'avèrent particulièrement porteurs dans une approche à visée didactique.

Dans cette perspective, les quatre principes suivants repris à la *logique du vivant*, et qui caractérisent l'activité audio-phonatoire, définissent par là-même le statut de l'oralité et peuvent/doivent être pris en compte pour décrire la langue orale (*i.e.*, le *Système Verbo-Tonal*), et élaborer des procédures efficaces dans l'enseignement-apprentissage des langues (*i.e.*, la *Méthode Verbo-Tonale*) :

- les manifestations phoniques résultent de la mobilisation globale de l'individu communiquant (dans sa dimension biologique et corporelle, psycho-sensorielle, psycho-cognitive, psycho-sociale) ;
- dans la structuration énonciative la primauté revient à l'intonation, qui manifeste l'affectivité et l'intention pragmatique ;
- toute activité de parole et sa manifestation phonique résulte de la relation écologique que le locuteur entretient avec la situation de communication et avec l'interlocuteur ;

²⁸ Pour ce qui est, par exemple, du spectre fréquentiel, les valeurs optimales pour la perception des réalisations d'un phonème ne correspondent pas aux pics d'énergie sonore du spectre. Ainsi la plage fréquentielle optimale pour la perception de [i] se situe au-dessus de 2.000 Hz et l'énergie du F1 n'est pas prise en compte pour l'identifier.

²⁹ Ce qui en audiologie a conduit à pratiquer l'audiométrie tonale et l'audiométrie verbale.

- la temporalité est une caractéristique *sine qua non* de la structuration phonique de la parole et en traduit la séquenciation.

Ces caractéristiques ne sont pas indépendantes les unes des autres et interviennent conjointement : ce n'est que pour en préciser l'incidence sur le statut de l'oralité, et en mentionner les implications dans la démarche didactique, que nous les présentons séparément.

3.2 *Le caractère global de l'oralité*

L'oralité en tant que manifestation *sui generis* de l'activité langagière de l'être humain se caractérise au premier chef par le principe de globalité propre à tout être vivant « de l'amibe au plus développé » (F. Jacob) – c'est-à-dire comme un ensemble dont tous les éléments constitutifs sont étroitement solidaires. Réduire la prononciation à la position et aux mouvements des organes du tractus phonatoire, dont on néglige d'ailleurs des caractéristiques qui s'avèrent déterminantes (par exemple, la référence au *tonus musculaire* de l'orbiculaire dans la prononciation des dentales en espagnol et en français, cf. *supra* § 2.3, et notes 22 et 23), signifie que l'on ne tient pas compte de cette prémisse essentielle de toute intervention didactique.

Le caractère global de la parole et *a fortiori* de son enseignement-apprentissage a été souligné par la Méthodologie Structuro-Globale, y compris par la mention explicite du terme dans sa dénomination, mais concernait la parole en situation. La prégnance du concept apparaît mieux encore appliquée à l'oralité, où se vérifient d'une manière patente, dans les manifestations phoniques de la parole, les inter-relations entre les plans biologique, affectif, psychosomatique, psycho-sensoriel, psychomoteur, psycho-cognitif et psycho-social (cf. *infra* § 4.4 et 4.5).

Le principe de globalité qui se manifeste dans ces relations constitue un trait fondamental et déterminant du statut de l'oralité. Il s'agit d'un principe opérationnel qui permet de mettre à profit lesdites relations pour concevoir des procédures efficaces qui recourent à la manifestation la mieux susceptible d'être appliquée (valeurs de la langue parlée, mouvements, signification du stimulus, etc.), et par là-même la plus efficace pour agir sur la caractéristique phonique objet de l'intervention didactique.

3.3 L'affectivité et l'intention pragmatique

L'affectivité et l'intention pragmatique constituent le deuxième volet du statut de l'oralité au plan didactique. La définition qu'en donnait Michel Henry – « Ce qui se sent sans que ce soit par l'intermédiaire d'un sens est dans son essence affectivité » – est généralement admise et sa prise en compte dans la démarche didactique s'avère ainsi problématique. Comment, en effet, intervenir sur le comportement autrement que par les sens ? De fait les analyses linguistiques formalistes ne prennent pas en compte *l'affectivité*.

Or il s'agit d'une composante de la parole que les variantes stylistiques manifestent (Ch. Bally, P. Guberina) – et dont dépend « même le choix des mots » (Guberina) – et qui est toujours présente :

L'affectivité accompagne l'individu durant toute sa vie depuis la naissance, et structure entre les vivants un véritable organe de coexistence (Boris Cyrulnik).

Mais comment appréhender « ce qui se sent sans que ce soit par l'intermédiaire d'un sens », et comment le prendre en compte ? Le problème s'avère d'ordre méthodologique. Certes, comme il arrivait dans les Méthodes audio-visuelles situationnelles, la mimique, certains comportements gestuels, plus ou moins codés selon les cultures, expriment l'état affectif du locuteur, mais c'est l'oralité qui traduit le plus directement et spontanément³⁰ ses émotions (dont on sait par ailleurs le rôle capital dans la mémoire) et la relation affective qu'il entretient avec l'interlocuteur.

Une première distinction s'impose à cet égard concernant les *actes de parole* selon qu'il s'agit de la catégorie « *acte 1* » ou « *acte 2* ». En général la question n'est pas posée dans les *Cours de langue*, or, pour ce qui est de l'affectivité et de l'intention pragmatique elle est primordiale. Si l'on convient, en effet, et l'accord est général à ce propos, que le dialogue est la forme la plus naturelle de la communication parlée, il est pédagogiquement utile de relever par exemple que *l'interrogation* (qui sollicite l'interlocuteur et ne requiert pas de réplique préalable) est un *acte 1*, et la *surprise* (réaction à une information inattendue, qui fait le plus souvent qu'on en demande confirmation) est un *acte 2* à forte charge

³⁰ Dans *Emma Bovary*, la muflerie de Rodolphe envers Emma dans la lettre de rupture ne serait pas jouable à l'oral.

affective³¹. Or ces modalités correspondent aux pôles intonologiques aigus, et constituent donc des valeurs optimales de référence pour la structuration prosodique du flux phonique.

Par ailleurs, *l'interrogation* est la modalité énonciative qui sollicite l'intervention de l'interlocuteur et dont les caractéristiques résultent de l'interprétation comme signal et marque affective d'une réaction de type biologique : la tension de désagrément chez l'enfant se traduirait par une tension et accélération des vibrations des cordes vocales, ce qui produirait la variation du F_0 vers les aigus, laquelle est perçue comme un appel (J. Kristeva, 1976). La courbe ascendante ainsi produite est à rapprocher également de la partie progrédiente de l'intonation assertive, qui justement correspond au thème de l'énoncé, et sollicite donc l'attention de l'interlocuteur.

La thématisation ou la focalisation d'un syntagme traduit toujours une charge affective. Les langues possèdent à cet effet des structures et des particules spécifiques (cf. en français : « *c'est...qui / que ...* » ; « *croire que c'est inéluctable, voilà l'erreur* » ...) mais, comme cela a été souligné par Creissels (2006, 114), « *ces procédés sont indissociables de l'intonation [et donc de l'oralité, c'est nous qui l'ajoutons] et [...] l'absence de prise en compte de l'intonation peut conduire à des erreurs grossières d'analyse. L'écrit peut donner une vision incomplète et éventuellement trompeuse de ces mécanismes* »³².

Outre la fonction mnémonique de *l'affectivité*, et son incidence au plan stylistique, il y a lieu également de souligner son rôle direct ou indirect dans la structuration grammaticale et textuelle. La structure parenthétique ou le « *je interprétant* » (cf. *infra* § 4.3 et 4.4), qui renvoient au plan psycho-cognitif et manifestent un moindre investissement affectif du locuteur, se traduisent par une bémolisation de l'intonation, phénomène spécifique à l'oralité³³. Les différences d'implication affective

³¹ Il y a lieu à ce propos de souligner la différence entre *surprise* et *étonnement* : cette dernière émotion causée par une information ou un événement inattendus qui laissent parfois se manifester par une intonation descendante et ne requiert pas confirmation.

³² Les représentations de l'écriture imaginées par les auteurs de bandes dessinées pour donner du relief au message transcrit sont particulièrement significatives de ce que l'orthographe « normalisée » ne rend pas compte de cette composante inhérente à la parole.

³³ L'écrit essaie de rendre compte de ce double discours par une virgule (pour différencier, par exemple, une subordonnée explicative, souvent causale, d'une subordonnée qualificative) ou par des parenthèses, mais c'est la prononciation, par un décrochement du F_0 généralement d'une

du locuteur se manifestent également à l'oral – et non à l'écrit – entre les éléments lexicaux *vs.* grammaticaux : les éléments grammaticaux, en particulier les clitiques, ne manifestent pas une charge affective, ils ne sont pas accentués et se réalisent toujours avec une tension moindre (cf. à cet égard la distinction établie par la grammaire traditionnelle entre le paradigme des pronoms « toniques » et les pronoms « atones », i.e. « clitiques »). Cette bémolisation des éléments grammaticaux (véritable érosion phonique, Haiman, 1985 ; Gauchola & Murillo, 2016) fait que, tant que ces emplois ne sont pas automatisés – automatisation qui ne peut se faire qu'à l'oral –, « l'erreur » morphologique ou syntaxique apparaît et/ou réapparaît.

Une fois maîtrisés, les éléments grammaticaux constituent un instant de détente dans l'échange langagier (de par l'absence d'accent, de la moindre tension, de l'absence de contenu lexical et de saillance communicative mais également du fait qu'ils ont été automatisés)³⁴.

Sans prétendre, il va sans dire à une présentation qui épuise la question, il y a lieu de souligner que *l'affectivité* constitue un élément essentiel du statut de l'oralité du fait qu'elle est inhérente à la parole, mais aussi de par sa place dans la *relation écologique* et dans la *temporalité* du processus audio-phonatoire qu'elle conditionne pour une large part.

3.4 La relation écologique

Ce principe de la *logique du vivant* que nous retenons pour définir le statut de l'oralité au plan didactique est généralement interprété comme renvoyant aux mêmes réalités qui constituent la *situation de communication*, ce qui n'est pas inexact mais demande à être précisé et complété³⁵, car l'être communiquant y joue un rôle actif que la simple référence à ladite situation n'explicite pas.

octave vers les fréquences plus graves, et par le moindre relief de l'incise, qui indique que cet énoncé ou partie d'énoncé correspond au « sujet interprétant ».

³⁴ Ce qui explique que l'écoute d'une langue étrangère que l'on ne maîtrise pas suffisamment produise une certaine fatigue : les éléments grammaticaux non automatisés qui devraient permettre une certaine détente exigent au contraire une attention supplémentaire puisque non accentués.

³⁵ C'est-à-dire le lieu et le décor dans lequel se réalise la parole, l'identité plus ou moins explicite du locuteur, les présupposés et préalables de la relation du locuteur avec l'allocutaire (affects, âge, courtoisie, actualité, milieu familial, voisinage, loisirs, travail, ...), le statut sociologique des interlocuteurs, leur culture réciproque, etc. Le locuteur est normalement présent et la communication s'établit par la langue parlée (Cl. Germain, 1973, 155)

Comme dans toute manifestation du vivant, la relation écologique situe les acteurs dans l'espace et dans le temps, mais traduit au premier chef l'interaction entre le locuteur et l'interlocuteur. Dans cette interaction interviennent l'état affectif, les capacités psycho-sensorielles, psycho-cognitives et même somato-sensorielles, ainsi que la proprioception, facteurs qui ne sont pas à proprement parler des éléments situationnels mais jouent un rôle déterminant dans l'activité langagière, et participent donc du statut de l'oralité. C'est la relation écologique d'ailleurs qui, la plupart du temps, fait émerger la parole.

Reste à définir la temporalité du processus audio-phonatoire dans sa double dimension séquentielle et durative.

3.5 *La temporalité du processus audio-phonatoire*

La temporalité de l'activité de parole, conçue comme la séquenciation des opérations de structuration des manifestations phoniques, n'est généralement pas mentionnée dans les analyses linguistiques *stricto sensu* ou à visée didactique ; or non seulement la durée mais l'ordre de succession desdites opérations constituent des caractéristiques essentielles à prendre en compte dans l'enseignement-apprentissage des langues, et sur lesquelles les approches didactiques fondées sur les caractéristiques articulatoires ou acoustiques du flux phonique – et plus encore si c'est sur la *scripture* – font l'impasse.

La temporalité apparaît déjà dans la manifestation même de l'affectivité :

Quand on étudie les émotions par tomographie à émission de positrons TEP chez des volontaires, un changement de conductivité de la peau précède toujours le moment où le sujet remue la main pour indiquer qu'il éprouve un sentiment (Damasio, 2003, 69).

Quelques faits de langue apparemment anecdotiques concernant la temporalité de l'activité de parole sont à cet égard spécialement significatifs. Les logiciels informatiques comportent aujourd'hui des correcteurs d'orthographe qui signalent, par exemple, les accords de nombre non respectés en écrivant « **le cahiers* », mais cette observation n'est signalée, et pour cause, qu'après avoir écrit le substantif ; or à l'oral le timbre de la voyelle [e] marque du pluriel, est prononcé avant de prononcer [kaje]. Les phénomènes d'assimilation régressive (le son

« assimilé » se prononce avant le son qui l'influence), ou de métaphonie (cf. *supra* note 23), qui traduisent l'incidence de « *ce qui n'est pas encore prononcé* » sur « *ce qui est en train d'être prononcé* », sont des indications patentes de ce que la structuration phonique – l'image sonore mentale – est antérieure à sa réalisation³⁶.

La modalité énonciative et l'intonation qui la réalise sont également programmées avant l'excitation des cordes vocales. Ainsi donc si la Méthode Verbo-tonale propose des procédures qui privilégient l'intonation, ce n'est pas seulement du fait de son rôle dans la structure du *dit*, et de par son incidence sur les unités phonématiques mais d'abord parce qu'elle est première et antérieure à la mise en mots dans la *temporalité* du processus de parole. A. Sechehaye ([1926], 1950) suggérait déjà d'explorer le niveau « *prégrammatical* » des langues, où il situait l'intonation³⁷ (cf. également Murillo, 1988).

La prise en compte du principe de *temporalité* dans la communication permet également de pondérer l'importance et la pertinence de la gestualité à valeur sémiotique (les signes codés plus ou moins déictiques : orientation, distance, forme, etc.) par rapport à l'expression verbale : si une personne dit « *là-bas, à gauche* » et montre du doigt une orientation « *à droite* », c'est le geste qui dit vrai... Les implications didactiques de ce constat sont évidentes.

Ces quelques références à la *logique du vivant* qui déterminent le statut de l'oralité montrent assez le caractère dynamique de l'activité de parole, et que tout acte langagier mobilise l'être communicant tout entier. Elles confortent la démarche qui consiste à définir l'intervention didactique sur la base de variables qui rendent compte de ce caractère dynamique de la parole et non à partir des descriptions physicalistes de la parole.

³⁶ Ce qui ne préjuge pas évidemment de la relation entre le *vouloir dire* – le sens – et le *dire*. Poser que le flux phonique résulte de l'être communicant envisagé comme un tout, et qu'il se structure avant la mise en mouvement des organes du tractus phonatoire ne signifie pas bien évidemment une quelconque prise de position concernant l'émergence de la parole, et en particulier sur une éventuelle conceptualisation antérieure à la formalisation en séquences phoniques. Nespoulous (2013, 210) rapporte à ce propos le témoignage de Jacques Lordat sur son propre vécu et qui caractérise l'aphasie de Broca : « J'ouvrais la bouche pour répondre [...]. Ma pensée était toute prête. Mais les sons [si] qui devaient la confier à l'intermédiaire n'étaient plus à ma disposition. » En tout état de cause, l'étape de structuration phonique, i.e. la « mise en sons », s'avère incontournable.

³⁷ Boë (1997) relevait ainsi une phase pré-articulatoire dans la séquenciation du processus de parole.

Mais outre que le statut de l'oralité dans la parole requiert que l'on prenne en considération la séquenciation des opérations d'émergence et structuration du flux phonique ainsi que des facteurs généralement négligés qui s'avèrent déterminants du point de vue pédagogique, il y a lieu de souligner également que, l'image mentale des langues étant phonique, l'oralité est non seulement inhérente à l'émergence de la parole, mais aussi la condition *sine qua non* de la mise en mémoire d'une langue (L_0 , L_2 , ou langue étrangère).

Dans cette perspective, le statut de l'oralité au plan didactique implique la nécessité de fixer les manifestations phoniques, y compris dans le processus d'apprentissage, et partant de prendre en compte leur composante temporelle non seulement dans leur dimension séquentielle mais encore durative. En effet, l'identification d'un phonème à partir du segment du flux qui le réalise requiert qu'il ait une durée minimale comprise entre 20 et 30 ms; or la durée moyenne des réalisations phonématiques, qui est déjà brève (occlusives $\approx$ 60 ms à raison de 12 phonèmes/seconde) est généralement biaisée car elle ne tient pas compte de la variation continue du flux phonique et s'avère une donnée souvent erronée au plan didactique, puisque si l'intensité maximale correspond à un segment du flux phonique d'une durée inférieure à 20 ms, l'intensité perçue tend à se situer aux 2/3 de la « montée ».

Conscients de ces impératifs psycho-sensoriels et des variations de la matière phonique, en particulier dans la transition entre deux phonèmes successifs, et afin de prendre en compte les propriétés des dites « transitions », certains chercheurs (cf. C. Harris, 1953 *apud* Boë, 1997) ont proposé de découper en deux, trois ou même quatre segments le flux correspondant à un phonème ; les unités ainsi constituées étaient désignées « diphones », « triphones », ou « quadriphones ». Mais outre que la durée des segments qui constituent ces unités est alors insuffisante pour leur identification, le nombre « d'unités phonématiques » s'avère excessif³⁸. Ces essais n'ont pas été satisfaisants car la réalité du flux phonique n'est pas une suite de segments mais un *continuum* :

³⁸ On en a ainsi établi plusieurs milliers (2.753 pour l'allemand, 2.265 pour l'anglais, 1.290 pour le français, 2.500 pour l'espagnol). Il est d'ailleurs significatif que leur nombre soit nettement supérieur en espagnol qu'en français, ce qui corrobore que la langue orale espagnole est moins tendue que le français... La démarche toutefois n'invalide pas l'hypothèse de la *tenue phonématique* (Lacerda et Menzerath) mais lui donne une valeur toute relative.

The concept of speech as a sequence of discrete units with distinct boundaries joined together as beads on a string is contrasted to the view of speech as a continuous succession of gradually varying and overlapping patterns (Fant, [1962] 2004, 149).

Il n'en reste pas moins que la parole, dont le flux phonique qui la manifeste est effectivement un *continuum*, est au plan communicatif – donc au plan didactique – une suite d'éléments discrets significatifs. Mais c'est l'être communiquant qui introduit des discontinuités dans le flux phonique en y relevant (la plupart du temps inconsciemment) les indices qu'il utilise pour le structurer en éléments discrets – les phonèmes³⁹.

Les maillons de la chaîne parlée étant ainsi constitués, il devient possible d'en faire la représentation symbolique par l'écriture. Comme relevait Georges Mounin, les inventeurs de l'écriture furent d'extraordinaires phonologues. Dès lors la *scription* de l'oral (cf. *supra* notes 2 et 7), à condition d'envisager l'écriture comme un objet plus ou moins symbolique et radicalement différent de la manifestation phonique qu'il représente, peut contribuer à fixer *l'aboutissement* du processus de structuration phonologique.

L'écriture peut/doit jouer un rôle essentiel pour engrammer l'oral ; elle constitue un facteur souvent déterminant du processus de mémorisation de la parole, et l'on comprend que l'on ne « puisse pas s'en passer » : l'écriture en arrive à être un référent capital de la parole⁴⁰, elle s'inscrit en grande partie dans la dimension écologique des langues (cf. *supra* § 3.4).

La création de *Alfonic* (« Un tremplin vers l'orthographe ») est à cet égard une tentative qui ne manque pas d'intérêt. Cette initiative rappelle la position des créateurs de l'A.P.I., car elle adopte le principe selon lequel :

³⁹ Les caractéristiques prises en compte, la plupart du temps de manière inconsciente, pour introduire les discontinuités varient selon les individus et selon les langues (Cf. *supra* la segmentation à partir de l'intonation, ou des « règles accentuelles »).

⁴⁰ Il aura fallu attendre le XX^e siècle pour concevoir des instruments qui permettent de ne pas assimiler graphème et phonème. Saussure s'étonnait encore de la confusion, relevée chez certains des « premiers linguistes » et des « humanistes », entre la représentation graphique et le son. Lorsque, à la fin du XV^e siècle, Nebrija veut rendre compte de la prononciation vélaire de la lettre "x", qui s'impose en espagnol, n'ayant pas un outil qui lui permette de caractériser ce nouveau phonème, il adopte comme référent la graphie, et donne comme description de la manifestation phonique de la lettre « la pronunciación propia de moros » (ch. III).

C'est toujours la même lettre qui sert pour un même son et toujours le même son qu'on prononce pour la même lettre (J. Martinet, 1975, 127).

La démarche de Jeanne Martinet va toutefois au-delà de l'A.P.I en ce sens qu'elle entend aider l'enfant à prendre conscience du processus de structuration de la matière phonique préalablement réalisé et à en saisir l'étape finale en en laissant une trace « permanente et solide ».

Alfonic n'a pas connu cependant le succès qu'espérait l'auteure. Mais, indépendamment des facteurs socio-administratifs et des « lois du marché » qui jouent un rôle non négligeable dans la diffusion d'une orientation pédagogique, les réserves à l'encontre de cette nouvelle proposition étaient dues aussi à des considérations d'ordre méthodologique. En effet, les saisies dans le processus d'appréhension de la matière phonique ainsi suggérées peuvent aider l'apprenant à caractériser inconsciemment le système phonologique, mais ne résolvent pas la question préalable et essentielle, à savoir, la structuration du flux phonique en unités discrètes : les phonèmes.

Au fond, la démarche pédagogique d'enseignement-apprentissage des langues concernant l'écrit renvoie toujours *nolens volens* au statut de l'oralité et à la relation de dépendance de la graphie par rapport à la manifestation phonique résultant du processus de structuration du flux sonore. La lettre n'est que la représentation de l'aboutissement de ce processus de structuration, et c'est donc ledit processus qu'il y a lieu d'appréhender et caractériser en vue de l'intervention didactique, y compris pour l'enseignement-apprentissage de la « langue écrite ». Dans cette perspective, comme il a été montré plus haut, la caractérisation des manifestations phoniques ne peut pas se circonscrire aux descriptions physicalistes.

4. Caractérisation à visée didactique des manifestations phoniques

Pour pouvoir agir sur la perception et prononciation de la parole, les paramètres acoustiques ou articulatoires ne sont pas opérationnels, car ils concernent le *résultat* de ces processus et non l'activité audiophonatoire en tant que telle. La question fondamentale qui se pose dans l'analyse à visée didactique de la parole consiste donc, dans un premier temps, à bien

distinguer les propriétés du résultat de l'activité audio-phonatoire d'une part, et d'autre part les caractéristiques du processus d'émergence et structuration des manifestations phoniques de la parole, autrement dit à ne pas confondre le « *dit* » et le « *dire* » (Ducrot, 1985).

4.1 La caractérisation du « *dire* »

La distinction *dire* ≠ *dit* est essentielle au plan didactique. Le *dit* est généralement conçu comme une chaîne d'objets sonores discrets alors que :

La parole est bel et bien un continuum sonore, résultat de la mise en œuvre parallèle des organes de l'articulation qui coproduisent une suite de sons, avec anticipation articulatoire (Boë, 1997, 33).

Ce *continuum sonore* est, certes, l'élément observable dans le processus de parole, et grâce au développement de la technologie du son il peut être enregistré et analysé pour en déterminer avec précision les propriétés physiques, mais il s'agit du résultat du *dire*, de l'*aboutissement* du processus de parole. Au demeurant, il en va de même pour les analyses dites articulatoires (palatographie, radiographie, électromyographie, cinéradiographie). Or l'intervention didactique ne peut être efficiente qu'en agissant sur ce processus, i.e. sur le *dire*, par le truchement des facteurs qui interviennent à cette étape du processus langagier.

Au risque de se cantonner à des suggestions vagues et impressives plus ou moins spéculatives sur lesdits facteurs, et fondées sur un « empirisme raisonné » qui, dans le meilleur des cas, se traduit par l'élaboration de recettes partielles (« au coup par coup ») ou éclectiques, l'élaboration de procédures didactiques concrètes et opérationnelles requiert que l'on ait caractérisé au préalable la langue parlée en tenant compte du statut de l'oralité dans l'activité langagière⁴¹.

Le *dire* se doit donc d'être envisagé comme étape essentielle et incontournable de la boucle audio-phonatoire⁴² – ce qui est un truisme –,

⁴¹ Ainsi, par exemple, vu que le système phonologique de l'espagnol n'a pas de constrictives sonores, et le français n'a pas d'approximantes, il n'est pas exceptionnel qu'on ne relève pas – et *a fortiori* qu'on ne corrige pas – chez les hispanophones qui apprennent le français, les réalisations de [-b-], [-d-], [-g-] intervocaliques (prononcées *[-β-], *[-ð-], *[-ɣ-]), or, comme relève le diagnostic verbo-tonal, il s'agit de l'effet d'un facteur déterminant du processus de structuration du flux phonique (cf. infra § 4.2).

⁴² Dans cette perspective, comme indiqué plus haut, les descriptions acoustiques (physiques) ou « articulatoires » (physiologiques) du *dit* n'ont qu'une utilité toute relative : la structuration du

et doit donc obligatoirement être pris en compte dans l'intervention pédagogique. Toute analyse à visée didactique se doit de caractériser cette étape de l'activité de parole de manière à permettre d'induire les procédures didactiques qui s'avèrent opérationnelles.

Par ailleurs, si les analyses acoustiques du signal phonique, les recherches sur la parole synthétique et sur la reconnaissance automatique de la parole ont permis de décrire avec précision le flux phonique à l'aide de paramètres acoustiques (durée, intensité, fréquence, timbre, etc.), il s'avère cependant, que « la réalité existe de trop » (Sartre *dixit*) : les résultats de ces travaux, qui ont certes donné lieu à des applications techniques importantes, ne rendent pas compte, comme rappelait Jakobson (1976, 34-37), de la structure phonologique, et *a fortiori* ne sont pas opérationnels au plan didactique.

Saussure insistait déjà sur le fait que les caractéristiques d'un élément de la chaîne parlée n'ont pas de valeur en soi mais seulement dans les rapports – syntagmatiques (relations de « contraste ») ou paradigmatiques (relations d'« opposition ») – qu'il entretient avec les autres éléments constitutifs de ladite chaîne, ou avec les éléments du système avec lesquels il peut commuter :

(...) dans la langue il n'y a que des différences. Bien plus : une différence suppose en général des termes positifs entre lesquels elle s'établit ; mais dans la langue il n'y a que des différences sans termes positifs (Saussure [1916] 1976, 166).

Dès lors, la structuration du flux phonique dépend du degré de pertinence relative des éléments qui le constituent et qui interviennent dans la construction et l'équilibre du système phonologique⁴³. Mais ce degré de pertinence relative ne peut être déterminé qu'à travers l'intervention du sujet communicant : il résulte au premier chef des

continuum du flux phonique en une suite de phonèmes consiste précisément à y introduire des discontinuités que l'on ne peut déterminer a priori : ces *discontinuités* en effet sont déterminées parfois par la prise en compte des saillies de l'énergie phonique (des pics de tonalité, d'intensité, etc.) et parfois au contraire en fonction des valeurs d'énergie phonique minimales ou même des silences. R. Jakobson (1976 : 34-37) a d'ailleurs souligné que, de ce fait, les descriptions acoustiques les plus précises ne rendent pas compte du système phonologique.

⁴³ « Du fait de la coordination du larynx avec les organes supraglottiques, où finit le [s] et où commence la voyelle, il existe de la friction dans ce qu'on est tenté d'identifier comme le début de la voyelle et il apparaît du voisement dans la fin de ce qu'on pourrait encore considérer comme la partie finale de la fricative sourde. » (Boë, 1997, 33).

compétences psycho-sensorielles individuelles (i.e., de la « nature humaine ») et de la langue de base du locuteur (i.e., culturelles).

Au plan didactique, la caractérisation physique – acoustique ou articulatoire – du flux phonique ne présente donc qu'un intérêt très relatif. C'aura été précisément un des apports essentiels de l'œuvre de P. Guberina, et des adeptes de la Méthode Verbo-Tonale, d'avoir fourni aux chercheurs, aux enseignants de langue, et aux rééducateurs des troubles de la parole et de l'audition, des outils méthodologiques qui permettent de caractériser les manifestations phoniques des langues sans recourir aux descriptions physicalistes⁴⁴, lesquelles font l'impasse sur des propriétés essentielles de la parole, et réduisent ainsi les résultats de l'analyse de la langue parlée à une simple « nomenclature » (Saussure F. [1916] 1976, 158).

4.2 La Méthode verbo-tonale : les valeurs de la langue parlée

La Méthode Verbo-Tonale s'avère opérationnelle pour caractériser le Système Verbo-Tonal et définir des procédures d'intervention didactique, du fait en particulier, que lesdites manifestations phoniques ne sont pas chosifiées dans leurs propriétés physiques intrinsèques et mesurables mais sont définies par leurs caractéristiques optimales obtenues empiriquement à partir de la perception qu'en a l'être communiquant⁴⁵. Ces caractéristiques désignées « *valeurs de la langue parlée* » (le terme de « valeur » est repris de Saussure⁴⁶) résultent à la fois de facteurs biologiques, affectifs, cognitifs, psychologiques et sociologiques que la didactique ne peut pas ne pas prendre en compte,

⁴⁴ Les recherches et applications qui se réclament de la Méthode Verbo-Tonale ont fait l'objet au cours des dernières cinquante années de plusieurs centaines de publications. Nous ne mentionnons ci-après que quelques principes et notions qui traduisent l'exigence de descriptions de l'oralité définies en vue de l'intervention pédagogique, et qui caractérisent la spécificité des manifestations phoniques dans cette perspective. La Méthode Verbo-Tonale contribue ainsi à préciser le statut de l'oralité au plan didactique.

⁴⁵ Ces procédures ne peuvent pas être des recettes plus ou moins ritualisées qui, comme préconisent certaines approches « conductivistes » de type skinnérien, pourraient être appliquées à tous les apprenants et dans toutes les situations d'apprentissage.

⁴⁶ Elle-même reprise de l'économie politique :

« ... même en dehors de la langue, toutes les valeurs sont toujours constituées

1° par une chose *dissemblable* susceptible d'être *échangée* contre celle dont la valeur est à déterminer ;

2° par des choses similaires qu'on peut comparer avec celle dont la valeur est en cause » (CLG, 159).

car c'est à partir de ces facteurs que peuvent/doivent être établies les procédures didactiques. Au demeurant, lesdits facteurs n'opèrent pas selon la logique mécanique qui préside aux conceptions physicalistes de la parole, dont le caractère réducteur s'est vérifié dans les analyses des matériels et les pratiques didactiques qui se disent communicatives, mais comme indiqué plus haut, selon la *logique du vivant* (F. Jacob, *op. cit.*).

Les *valeurs de la langue parlée* permettent de caractériser ledit processus de structuration du flux phonique et ce faisant d'en révéler la spécificité, du fait en particulier que sont pris en compte le *tempo* et la composante temporelle (la *durée*) des manifestations phoniques, mais aussi la *temporalité* du *processus de parole* (i.e., la succession des opérations qui interviennent dans ledit processus); et ce aussi bien au plan phonématique – en particulier par les valeurs *tonalité* et *tension* – qu'au plan prosodique – *intonation* et *rythme* principalement.

Les résultats obtenus au cours de ces dernières cinquante années (Berré, 2005) par l'application des procédures fondées sur la *linguistique de la parole*, et les recherches verbo-tonales qui s'en réclament, confortent cette approche et montrent, en particulier, que la structuration phonologique sur l'axe temporel peut être défini à la fois au niveau phonématique et au niveau prosodique⁴⁷ par des *valeurs* qui permettent de caractériser le *dit* et le *dire*, c'est-à-dire l'ensemble du processus audio-phonatoire.

4.2.1 Les valeurs de la langue parlée au niveau phonématique : tonalité et tension

La tonalité des unités phonématiques est définie à partir de l'analyse systématique de la perception phonique du spectre fréquentiel d'une langue donnée par les sujets possédant ladite langue. Cette analyse permet de dégager des *gammes fréquentielles optimales* propres aux réalisations de chaque phonème de cette langue, c'est-à-dire les gammes fréquentielles où la perception, l'identification et partant la structuration phonologique du segment du flux phonique correspondant s'avère optimale pour la majorité de la population possédant cette langue (Guberina, [1972] 2003 ; Asp, 1981, 2011)⁴⁸.

⁴⁷ Les recherches et applications concernant la vibrotactilité ont montré, par exemple, que la perception des sons graves, qui caractérisent en particulier l'intonation, s'effectue par l'ouïe et par la peau.

⁴⁸ Tout son de la parole filtré aux fréquences de ladite gamme optimale tend d'ailleurs à être perçu comme le phonème dont la gamme est l'optimale. Ainsi, par exemple, n'importe quel son ([i],

On peut donc adopter l'hypothèse que ces *gammes optimales* représentent des *tendances générales* – qui rejoignent ainsi la notion de *prototype* en linguistique – et correspondent à « l'équilibre des tonalités dans l'écoute » (Guberina, [1974] 2003, 247) de cette langue (entropie), étant bien entendu cependant que ces valeurs ne sont pas universelles et ne sont pas optimales pour toute la population d'un idiome donné et *a fortiori* pour tous les apprenants⁴⁹ : elles permettent d'émettre des hypothèses sur les réactions des apprenants mais non de préjuger du comportement perceptif idiosyncrasique de chacun d'eux. Il s'ensuit qu'il n'y a pas de « recettes » pédagogiques valables dans tous les cas, et qui pourraient être appliquées d'une manière générale et systématique (cf. note 45).

Par ailleurs, contrairement à l'hypothèse saussurienne rappelée plus haut selon laquelle « dans la langue il n'y que des différences sans termes positifs », ce qui impliquerait qu'aucun des éléments du système ne pourrait être pris comme référent pour identifier les unités qui constituent la chaîne parlée (dans la pratique cette hypothèse débouche sur des définitions circulaires, sur lesquelles J. P. Sartre a ironisé dans *La nausée*, et ne fournit pas à l'enseignant des instruments qui lui permettent d'élaborer des procédures efficaces), la méthode verbo-tonale part du constat que, dans le comportement perceptif de l'être humain, il se produit une première scission claire dans l'aire d'audition globale de l'oreille humaine aux limites fréquentielles de la matière phonique utilisée et perçue dans la parole (Guberina, 2003)⁵⁰. Ces limites constituent de véritables *seuils de structuration phonologique* (Murillo, 2007), la plage comprise entre ces seuils pouvant être ainsi définie comme la *palette sonore* d'une langue.

[u]...) filtré à la fréquence qui est censée être l'optimale de [a] est perçu comme étant un [a] (quoiqu'avec une moindre intensité). De là l'aphorisme des verbo-tonalistes « Tout son de la parole contient tous les autres sons de la parole »

⁴⁹ Les verbo-tonalistes n'ont cessé de souligner que les procédures didactiques suggérées ne sont pas des consignes et doivent toujours être adaptées – parfois même remplacées par des procédures que l'on considérerait au départ inefficaces – aux réactions des apprenants qui peuvent être ... inattendues.

⁵⁰ La notion de *seuil* n'est pas explicitement désignée comme telle dans la littérature verbo-tonale, mais elle est appliquée dans les pratiques préconisées avec le SUVAG : Dans un premier temps il est suggéré de filtrer les stimuli phoniques en passe haut et passe bas de la plage phonique correspondant à la parole.

La perception phonique de l'être humain dépend par conséquent de ses capacités auditives naturelles, qui sont universelles et propres à l'espèce (quoiqu'avec des variations individuelles qui peuvent aller jusqu'à l'hypoacousie), mais dépend aussi des propriétés acoustiques de la matière phonique qui, dans le processus de structuration phonologique de chaque langue, se sont avérées optimales pour une population donnée et sont donc spécifiques à chaque langue.

L'on peut ainsi déterminer une véritable *palette sonore* des différentes langues entre le *seuil de perception tonale* correspondant aux fréquences les plus aiguës de la parole, et le *seuil* correspondant aux fréquences les plus graves. Un troisième *pôle de structuration phonologique* a pu également être établi qui correspond, dans la plupart des langues, à la perception des réalisations de /a/, dont la *tonalité* optimale correspond à la zone fréquentielle de meilleure audibilité de l'oreille humaine⁵¹. Ces seuils sont d'ailleurs corrélés aux mouvements : aux fréquences les plus aiguës correspondent des mouvements centrifuges et aux fréquences les plus graves des mouvements centripètes.

La tension constitue la deuxième *valeur de la langue parlée* au niveau phonématique, autrement dit le deuxième facteur de discontinuité qui structure la matière phonique de chaque langue en système phonologique. Il s'agit d'une propriété qui présente un intérêt majeur : elle est étroitement associée aux mouvements et à la macro-motricité au point que s'agissant d'une valeur proprioceptive (Renard, 1971), c'est à travers les gestes et les mouvements qu'elle est le mieux perçue, et que la *tension* relative des différentes réalisations phonématiques peut être établie⁵². Les différents phonèmes peuvent ainsi être situés sur une échelle tensionnelle entre deux pôles : pôle de tension maximale (réalisations à consonanticité maximale : occlusives sourdes), et pôle de tension minimale (réalisations des voyelles). La corrélation entre macro-

⁵¹ Les praticiens de la Méthode Verbo-Tonale dans le domaine de la rééducation des troubles de la parole et de l'audition adoptent les réalisations de /a/ comme référent de base à partir duquel construire le système et faire percevoir et produire les autres réalisations. Il est symptomatique à cet égard que les indices de perceptibilité de ce phonème tendent à coïncider quelle que soit la langue, et ce même si ces langues ont des *palettes sonores* différentes.

⁵² La scalarité des valeurs tensionnelles des unités phonématiques est révélée par la proprioception de la tension corporelle : les sourdes sont plus tendues que les sonores, les occlusives plus tendues que leurs corrélats constrictifs, les consonnes plus tendues que les voyelles, etc. (Renard, 1976).

motricité et manifestations phoniques déjà, relevée concernant la *tonalité*, se vérifie également pour ce qui est de la *tension* : plus forte tension manifestée dans les gestes rapides et « raides » ; moindre tension dans les mouvements lents et ondoyants.

La *tension* intervient également dans la structuration rythmique du flux phonique en soulignant la discontinuité *syllabe accentuée* ≠ *syllabe non accentuée*, et comme manifestation des différences de mise en relief selon qu'il s'agit de segments phoniques grammaticaux (absence « d'accent tonique ») ou de « mots phoniques ayant une charge sémantique » (présence de l'accent à valeur culminative, démarcative ou expressive). La différence de tension manifeste ainsi également la distinction entre le discours du *sujet communiquant* et le discours du *sujet interprétant* (Charaudeau, 2009), et la distinction entre « catégorie grammaticale **vs** [...] lexicale » (Nespoulous 2013 *dixit*) : elle est un facteur déterminant dans la structuration grammaticale d'un énoncé aussi bien chez le locuteur que chez l'allocutaire. La différence de *tension* caractérise de manière semblable le *discours référentiel* du *discours commentaire* (ou parenthétique⁵³).

4.2.2 Les valeurs de la langue parlée au niveau prosodique : rythme et intonation

Plus encore que pour les unités phonématiques, c'est au niveau suprasegmental que le statut de l'oralité s'avère à la fois représentatif et déterminant au plan didactique tout particulièrement pour ce qui est du *rythme* et de *l'intonation*. Or il est notoire que ces *valeurs de la langue orale* ne peuvent pas être décrites par l'analyse articulatoire et ne sont pas transcrites à l'écrit (la ponctuation et les « blancs » entre les mots sont des procédés conventionnels à valeur très limitée).

Le rythme structure le flux phonique au niveau prosodique en manifestant la dimension temporelle des langues, or, comme cela a été indiqué plus haut (cf. § 3.5), la dimension temporelle est consubstantielle au *vivant* et ne peut dès lors être omise dans une analyse à visée didactique. Émile Benveniste (1966) a défini le concept de « rythme » (en le distinguant

⁵³ Cf. Guberina [1988], 2003. Les travaux concernant le rôle de la catégorie grammaticale vs. rôle de la fréquence lexicale sont à cet égard probants (cf. Nespoulous, 1999). Pour les caractéristiques phoniques des parenthétiques, cf. D. Bolinger (1989), Baqué (1998), Schneider (2007), Gachet et Avanzi (2008), Avanzi (2009).

de la notion de *structure*) et relevé en particulier que, quoique certaines propriétés qu'on lui attribue (tempo, débit...) aient pu prêter à confusion (par exemple, avec le débit), il est d'abord, et quel que soit le phénomène envisagé, l'expression d'une discontinuité dans le déroulement temporel.

Dans la parole, la syllabe constitue l'unité minimale de diction et une première structuration rythmique qui manifeste cette discontinuité se traduit par l'alternance plus ou moins régulière de syllabes mises en relief (syllabes accentuées ou marquées) ou non (syllabes non accentuées, non marquées). Les *pauses* intérieures au « discours » manifestent également par définition l'arrêt du flux phonique et le structurent en fonction de variables morphosyntaxiques, sémantiques et/ou communicatives.

Ces caractères et propriétés du rythme en tant que *valeur de la langue parlée* inhérente à l'oralité et incontournable (Jakobson 1980) sont notoires, et les analyses de « l'Art poétique » ont pu déterminer des figures rythmiques qui participent du statut de l'oralité et qui caractérisent la parole et même les langues⁵⁴. Au plan didactique il est particulièrement intéressant de relever à cet égard qu'il s'agit d'une composante de la parole qui ne se circonscrit pas au flux phonique. Le père de la Méthode Verbo-Tonale déclarait en effet à ce propos d'une manière péremptoire :

L'oreille ne suffit pas pour saisir le rythme (Guberina [1985] 2003, 412).

Partant de ce constat, les recherches qui se réclament de la Verbo-Tonale, en particulier dans le domaine de la rééducation des troubles de la parole et de l'audition, ont donné une nouvelle dimension pédagogique à cette composante de la parole en relevant que le rythme du flux phonique est parfaitement congruent avec le rythme corporel.

Cette *iconicité*⁵⁵ entre le rythme de la parole et le rythme corporel (Gauchola et Murillo 2015) se vérifie dans la relation entre la structure temporelle du flux phonique et les inflexions dudit flux phonique d'une part,

⁵⁴ Ainsi, le français est défini au plan rythmique comme une langue à scansion relativement régulière (tendance à l'isochronie syllabique) et rythme iambique, l'espagnol comme une langue à tendance rythmique trochaïque.

⁵⁵ Le principe d'*iconicité* a été énoncé dans un premier temps pour rendre compte des étroites corrélations (au point qu'on a pu parler d'une « covariation de forme et de signification ») entre les manifestations formelles et les composantes sémantiques des phénomènes langagiers. Il s'avère particulièrement porteur au plan didactique. Cf. Haiman (1985), Sonesson (2008), Nobile (2014).

et d'autre part dans la relation entre la structure temporelle et l'évolution dans l'espace de l'expression corporelle. Mais ces relations *d'iconicité* se vérifient également dans la relation entre d'une part la structure du flux phonique et d'autre part la structuration grammaticale et sémantique.

Dès lors, l'intervention sur le rythme de la parole, en particulier sur la structuration temporelle du flux phonique, peut être réalisée à partir d'indications ou de consignes ayant trait à l'expression corporelle⁵⁶ : la macro-motricité (« la musculature striée », dont Hjelmslev regrettait déjà qu'elle ne soit pas prise en compte dans les analyses linguistiques) est inhérente au *dire*, et constitue un moyen efficace d'intervention didactique, y compris au plan grammatical⁵⁷.

Envisagé dans cette perspective, le rythme, tout particulièrement dans sa manifestation corporelle incontournable – dont les analyses linguistiques formalistes ne rendent pourtant pas compte – traduit également la corrélation entre la structure du flux phonique et la structuration syntaxique et/ou lexicale, voire communicative ou pragmatique. Il intervient ainsi dans la structure intonative, et ne peut pas dès lors être envisagé de manière isolée : rythme et intonation sont deux *valeurs de la langue parlée* indissociables et inhérentes au *dire*.

L'intonation a été traditionnellement exclue de l'analyse linguistique comme étant « un fait marginal » (A. Martinet, 1965, 25-31 *apud* M. Rossi *et al.* 1981, 7). Dans l'enseignement-apprentissage des langues, quand l'accès au signifiant est envisagé en tant que tel, la plupart des *Cours de Langue* circonscrivent généralement la démarche au niveau phonématique, et l'intonation, la plupart du temps, ne fait pas l'objet de présentations ou d'exercices spécifiques ; il n'est d'ailleurs pas exceptionnel que même des ouvrages spécialisés consacrés à la prononciation lui aient réservé la portion congrue⁵⁸.

⁵⁶ Par exemple : à une syllabe accentuée correspond toujours une inflexion gestuelle (généralement inconsciente) aussi discrète soit-elle (mouvement des mains, des paupières, des muscles du visage, hochement de la tête, etc.). Pour la corporalité du rythme, cf. Murillo (2002), Asp (2011).

⁵⁷ Comme indiqué plus haut, les éléments grammaticaux, en particulier les clitiques, sont dépourvus de sens lexical et ne sont pas accentués ; les « mots pleins », à valeur sémantico-lexicale, ont toujours un accent et un seul ; autrement dit, les propriétés rythmiques sont corrélées à celles qui relèvent de la structuration grammaticale.

⁵⁸ Ainsi l'excellent manuel *Fonética y Fonología francesas* (Alicia Yllera, 1991, Madrid, UNED) consacre 4 pages à la présentation de l'intonation (pp. 318-391) ; mais la plupart des *Cours de*

Les études qui se sont multipliées au cours des dernières décennies concernant les composantes suprasegmentales de la parole ont montré cependant le rôle capital de l'intonation : elle est l'expression spontanée de *l'affectivité* du locuteur, elle indique la *modalité énonciative* (elle en est parfois le seul signe, comme dans l'interrogation totale sans mot interrogatif ou inversion du sujet), elle assure la *thématisation* (thème/rhème) des syntagmes d'un énoncé, et structure le flux phonique au point de constituer la véritable *syntaxe de la langue orale* (Ch. Bally *apud* R. Renard, 1979). L'intonation joue ainsi une fonction triple : stylistique, structurelle et sémantique⁵⁹. Elle s'avère une composante incontournable du *dire*, et partant du statut de l'oralité au plan didactique ; dans la temporisation préconisée par la Méthode Verbo-Tonale, elle occupe la place prioritaire (cf. *supra* la référence à *Voix et Images de France*).

Toutefois, si les tenants de cette orientation méthodologique ont toujours revendiqué la primauté de l'intonation dans l'enseignement-apprentissage des langues, il n'y a pas eu d'analyses intonologiques systématiques, ni de procédures didactiques précises et explicites sur l'intonation, comme cela a été fait au niveau phonématique (Mestreit, 2011).

Au demeurant, les difficultés concernant l'intervention sur l'intonation ne sont pas spécifiques à la didactique des langues :

Le premier problème qui se pose dans l'étude de l'intonation est de définir sa nature et le niveau d'analyse auquel elle se situe (Rossi, 1981, 7).

Langue actuellement sur le marché font l'impasse sur cette composante de la parole. Il y a lieu à cet égard de relever que dans la première moitié du XX^e siècle il n'était pas exceptionnel que des ouvrages entiers soient consacrés à l'intonation pour l'apprentissage des langues. Cf. Coustenoble et Armstrong (1934), *Studies in French Intonation*, Cambridge, W. Heffer & Sons Ltd.

⁵⁹ D'aucuns ont pu avancer que la dimension sémantique de l'intonation n'est pas évidente car, par exemple, si elle manifeste un rapport logique entre deux propositions, ledit rapport est déjà implicite dans la relation sémantique qu'entretiennent les deux propositions (« *Je ne sors pas, il pleut* »). Mais dans le cas bien connu des relatives, ce n'est que l'intonation qui permet de savoir si ladite proposition relative est *explicative* (à valeur caractérisante), et exprime, par exemple, une *cause* qui concerne tout un ensemble référentiel parmi d'autres, ou *déterminative* (à valeur référentielle) dans la mesure où elle ne porte que sur une partie de l'ensemble, la partie caractérisée par la relative. La *scripture* essaie tant bien que mal de rendre compte de cette distinction en encadrant la proposition relative explicative entre virgules, qui ne sont cependant que des pauses, et non la variation mélodique (« les grévistes (,) qui s'étaient opposés à l'accord (,) refusèrent de reprendre le travail » = tous les grévistes ? seulement ceux qui s'étaient opposé à l'accord ? C'est l'intonation qui donne l'information).

Les solutions apportées à ce problème par les recherches intonologiques en ont montré la complexité en soulignant en particulier le caractère pluri-paramétrique de l'intonation, où interviennent au moins la mélodie (la variation de la fréquence fondamentale, F_0), la durée et l'intensité⁶⁰. Mais dans une approche à visée didactique, ce genre d'analyses présente, comme au niveau phonématique, un intérêt tout relatif : même si effectivement l'analyse acoustique de l'intonation requiert que l'on prenne en compte ces trois paramètres, et même si la perception de l'intonation est globale dans l'activité audio-phonatoire et intègre les trois paramètres (parmi d'autres d'ailleurs), il s'avère que la mélodie est déterminante. Les variations du F_0 jouent un rôle prépondérant dans la structuration du flux phonique et permettent entre autres de saisir la *modalité énonciative*, les phénomènes de « *topicalisation* » (Alcoba et Murillo, 1998), et la *segmentation du flux phonique* en syntagmes (Alcoba, Le Besnerais et Murillo, 1992).

L'analyse des phénomènes de perception de la parole révèle également que les variations du F_0 entre ses valeurs extrêmes s'avèrent plus ou moins scalaires, et que ladite scalarité manifeste des discontinuités significatives⁶¹. Il apparaît que, toutes proportions gardées, de la même manière que la structuration de la matière phonique au plan phonématique (cf. *supra* § 3.3) peut être caractérisée à partir de trois *pôles de structuration phonologique*, au plan prosodique la structuration du flux phonique peut être caractérisée à partir de trois *pôles de structuration intonologique* : le pôle des intonèmes qui atteignent les fréquences du F_0 les plus aiguës, le pôle des intonèmes qui atteignent les fréquences les plus graves, et les intonèmes qui correspondent au *champ optimal d'audition*, c'est-à-dire à la zone fréquentielle de meilleure audibilité de l'oreille humaine (Gauchola et Murillo, 2016). Ces *pôles de structuration intonologique* correspondent d'ailleurs à des modalités énonciatives bien définies :

- l'intonème le plus aigu, qui atteint la valeur maximale du F_0 , correspond dans une diction expressive à la manifestation de la

⁶⁰ Il n'est pas exceptionnel cependant que les termes « intonation » et « mélodie » soient considérés équivalents dans des matériels d'enseignement.

⁶¹ Cf. en particulier le projet de recherche PROSOPTIMA HUM2007-61648 du Ministerio de Educación y Ciencia espagnol et Murillo (2013).

surprise, et dans la réalisation des « modalités grammaticales », à *l'interrogation totale* (i.e., sans particule interrogative, morphème interrogatif – « est-ce que » – ou *inversion du sujet clitique*) ;

- l'intonème le plus grave et qui atteint la fréquence minimale du F_0 , correspond à la *modalité jussive* (et parfois dans une diction expressive, à l'étonnement⁶²) ;
- les intonèmes du *champ optimal d'audition* au niveau prosodique (i.e., l'intonème « neutre ») correspondent à la *modalité suspensive* ou d'une manière plus générale aux temporisateurs discursifs dont « *eah* » est le représentant prototypique en français.

Ces intonèmes peuvent, certes, être décrits au moyen de paramètres acoustiques – c'est d'ailleurs le cas de la plupart des analyses et descriptions actuelles qui s'en tiennent à des approches physicalistes –, mais au plan didactique, au même titre que les unités phonémiques, ils ne sont plus à caractériser par leurs propriétés acoustiques en soi, mais par leur valeur relative (y compris proprioceptive) – qui est globalement perçue – et par les relations qu'ils entretiennent : du point de vue pratique ces données physiques n'auraient qu'une utilité toute relative pour l'enseignant ou l'apprenant, car la « réalité » ainsi représentée s'avère trop complexe et les données obtenues ne permettent pas de la catégoriser.

Comme au plan phonématique également, des seuils perceptifs peuvent dès lors être établis, par exemple entre l'intonème *interrogatif* et l'intonème expressif de *surprise* (Murillo, 2013) à partir des *valeurs optimales* qui s'avèrent opérationnelles et qui sont à déterminer à la fois par des épreuves de perception auditive, par la proprioception, et à partir de la manifestation de ces valeurs optimales dans la macro-motricité.

Les *valeurs de la langue parlée* « tonalité » et « tension » contribuent également à la caractérisation de l'intonation : les modalités *surprise* et *interrogation* sont perçues comme ayant un degré de tension supérieur à la modalité *suspensive*, tout comme les modalités *assertives* et *jussives*. L'intonation constitue ainsi, dans une approche à visée didactique, un

⁶² Une analyse fine des modalités énonciatives expressives au niveau intonatif peut d'ailleurs aider à bien différencier les vocables *surprise* et *étonnement*, qui dans le « langage ordinaire » sont souvent considérés comme ayant des sens équivalents (c. note 31).

vecteur de l'oralité qui intègre les autres *valeurs de la langue parlée* : la *tonalité*⁶³, la *tension* et le *rythme*.

Par ailleurs, toujours en adoptant une visée didactique, il y a lieu de souligner que les variations de la fréquence fondamentale (F_0) sur le spectre fréquentiel de la parole correspondent aux fréquences les plus graves du flux phonique. Or, le corps humain est spécialement sensible aux fréquences graves (Guberina, 2003 ; Asp et. al., 2011 ; Bensalah, 2013) et *l'iconicité* entre les variations du F_0 et la macro-motricité qui lui est ainsi associée peut/doit être mise à profit dans l'intervention pédagogique sur la prononciation, aussi bien en langue maternelle (dans la parole ordinaire et/ou pathologique) que dans l'enseignement-apprentissage des langues étrangères (Osorio, 2016).

Ces relations d'iconicité sont consubstantielles à l'activité langagière de l'être communicant. Elles sont donc universelles et caractérisent les manifestations phoniques de la parole au même titre que les *valeurs de la langue parlée*. Au plan didactique elles sont d'une importance capitale.

4.3 Les relations d'iconicité dans la parole

Comme il vient d'être indiqué, la caractérisation des manifestations phoniques au moyen des *valeurs de la langue parlée* permet de relever, pour les différentes langues, les corrélations – i.e., les relations d'iconicité – entre les différentes dimensions de l'expression : verbale (lexicale et grammaticale), phonique et gestuelle.

Pour ce qui est de la prosodie – rythme et intonation principalement –, ces corrélations montrent que la perception et production phoniques ne se circonscrivent pas à la « boucle audio-phonatoire », et encore moins au tractus phonatoire. Les réalisations des modalités énonciatives *surprise* et *interrogation totale*, par exemple, se caractérisent globalement par la variation de fréquence du F_0 (i.e., hauteur par rapport au *pôle de structuration intonative suspensif*, et courbe tonale ascendante), mais aussi par l'orientation de la macro-motricité⁶⁴,

⁶³ Y compris la *microprosodie* propre à chaque phonème (Di Cristo), et la place de l'accent par rapport aux inflexions de la fréquence fondamentale (F_0).

⁶⁴ Le glissando des deux intonèmes vers les aigus caractérise les deux modalités mais est nettement plus prononcé et abrupt dans la surprise. Au plan de l'expression corporelle, il se

et l'intensité ou la « force » du mouvement (i.e., « l'effort » dans la terminologie des travaux de Rudolf Laban⁶⁵). Ces corrélations qui traduisent l'*iconicité* parfaite entre la structure du flux phonique, le contenu verbal et les macro-motricités se vérifient quel que soit l'énoncé.

Les relations d'*iconicité* parole↔macro-motricité se manifestent également au plan phonématique : la proprioception optimale des réalisations de chaque phonème correspond en effet à une certaine partie du corps, et pour la perception et production des différents sons de la parole⁶⁶, certaines postures, gestes et mouvements s'avèrent optimaux. Cette *iconicité* a pu être mise à profit dans l'enseignement-apprentissage des langues étrangères et dans la rééducation des troubles de la parole en élaborant des procédures de stimulation par le mouvement (Roberge, 2002, Pavelin et Frankol, 2016).

Partant de ces considérations, les *valeurs de la langue parlée*, les *relations d'iconicité* qu'elles révèlent, s'avèrent des éléments définitoires du statut de l'oralité au plan didactique. La Méthode Verbo-Tonale, dont l'efficacité est notoire et généralement reconnue, se fonde précisément sur le statut de l'oralité ainsi défini. Mais l'approche conduit également à poser – quoiqu'il n'en ait pas toujours été ainsi – que les pratiques didactiques qui en découlent conçoivent les langues comme des objets d'enseignement-apprentissage en constante évolution, du fait que tout système linguistique est par essence un système dynamique au plan diachronique et synchronique (Seiler, 2002), mais aussi parce que les manifestations phoniques de la parole ne peuvent être définies au plan didactique sans la référence à l'être humain⁶⁷ et sans la référence aux

produit un mouvement centrifuge dans les deux cas, mais nettement moins ample pour la modalité interrogative

⁶⁵ Communication personnelle de Jacques Bouchez, et Bouchez (2009).

⁶⁶ Ainsi les locuteurs non spécialistes du langage indiquent la même localisation sur le corps pour la série de tension décroissante [s]⊆[ʃ]⊆[z]⊆[ʒ] : si on leur demande de prononcer ces phonèmes et de placer leur main sur l'axe du sternum, il y a pratiquement unanimité – quelle que soit la langue de base, le niveau socioculturel ou le sexe – à les placer dans l'ordre indiqué de haut en bas.

Les recherches réalisées en particulier pour la rééducation des troubles de la parole ont relevé, pour les différents allophones, les tendances à associer {geste ≈ son}, i.e., à relever l'*iconicité* entre parole et activité sensori-motrice.

⁶⁷ L'être humain « ondoyant et divers » (Montaigne), et *a fortiori* l'être communiquant, dans ses dimensions affective, psycho-sensorielle, psycho-cognitive et macro-motrices.

situations de communication, par définition multiples et toujours changeantes de par le vécu de tout être communicant.

La recherche méthodologique se fourvoierait donc à prétendre établir des procédures didactiques à partir des propriétés acoustiques ou « articulatoires » du *dit*, et qui plus est en les considérant *en soi* et invariables. Une démarche pédagogique fondée sur le statut de l'oralité implique la recherche permanente de procédures qui s'avèrent optimales pour agir sur les facteurs qui déterminent le flux phonique, et qui ne peuvent se réduire à des répertoires de recettes didactiques. Il n'y a pas lieu *a fortiori*, dans le cadre de la présente réflexion sur le statut de l'oralité dans l'enseignement-apprentissage des langues, de définir lesdites recettes, mais il va sans dire que cette approche et conception de la parole ne peut pas ne pas se traduire au plan didactique.

5. Les implications didactiques

Il est devenu un lieu commun de poser que l'intervention didactique doit être « centrée sur l'apprenant ». On entend généralement par-là que la démarche ne doit pas partir du système à enseigner-apprendre, ni imposer des procédures stéréotypées définies *a priori*, et qui seraient généralisables. Il s'agit bien évidemment d'un principe de bon sens.

Mais, comme il ressort de la caractérisation à visée didactique des manifestations phoniques de la parole présentée ci-dessus, et du statut de l'oralité dans l'activité langagière, l'apprenant n'est pas seulement à prendre en compte parce qu'il est le destinataire de l'intervention pédagogique ; il en est d'abord l'acteur principal et ce à un double titre : parce que c'est lui qui possède les mécanismes de perception et de production, et parce que c'est en lui que résident les difficultés, les sources d'éventuels blocages et les éléments facilitateurs. C'est l'apprenant et lui seul qui possède les capacités et potentialités d'apprentissage (quoiqu'il n'en ait généralement pas conscience). Il s'ensuit que l'intervention pédagogique ne peut/ne doit que contribuer à faire émerger ces potentialités et à les rendre opérationnelles. La démarche didactique optimale s'avère donc éminemment idiosyncrasique.

Le statut de l'oralité constitue ainsi un élément essentiel de cette exigence méthodologique : c'est dans le cadre dudit statut que la

référence aux facteurs qui déterminent le *dire* (cf. § 4.1 et 4.2) et que les *valeurs de la langue parlée* qui permettent d'en rendre compte s'avèrent être des moyens opérationnels pour que l'apprenant s'investisse et pour que l'enseignant puisse concevoir des procédures didactiques efficaces en établissant un protocole opérationnel qui intègre *production* et *perception*.

Les recherches et applications qui se réclament de la Méthode Verbo-Tonale, conçue comme un protocole de recherche-action dans l'enseignement-apprentissage des langues, permettent d'ores et déjà de distinguer quatre grands types de procédures selon qu'elles ont plutôt trait à :

- l'« *intention pragmatique* » et à la *globalité* de l'activité de parole,
- la *structuration mélodique et rythmique*,
- la *structuration phonématique*,
- la *mise en mémoire*, i.e., l'*engrammation* de la matière phonique.

Il n'y a pas lieu dans le cadre de la présente contribution de proposer une sorte de guide pédagogique qui, bien évidemment, ne peut être établi qu'en fonction des situations concrètes d'enseignement-apprentissage : niveau des apprenants, objectifs poursuivis, matériels employés, etc.⁶⁸ Mais dans tous les cas la prise en compte du statut de l'oralité au plan didactique implique ce quadruple parcours. Au demeurant ces approches sont complémentaires et la plupart du temps convergentes.

5.1 Les procédures d'intervention sur les manifestations phoniques

Quelle que soit la situation qui motive l'activité de parole, l'*intention pragmatique* est le facteur préalable à toute structuration phonique ; elle est une réalité pré-linguistique qui détermine la *modalité énonciative* (« grammaticale » ou « expressive »), le « choix des mots » (Guberina *dixit*), et qui d'une manière consciente ou inconsciente, traduit l'état

⁶⁸ Ainsi pour les *débutants complets*, le recours à des corpus oraux à imiter et la présentation d'échantillons de parole qui permettent des exercices de simulation s'avèrent opérationnels pour enclencher le processus d'apprentissage, alors que pour des étudiants avancés ce sont les connaissances antérieures qui peuvent être mises à profit pour le travail sur la prononciation et la diction.

affectif du locuteur⁶⁹. La prise en compte de *l'intention pragmatique* se doit donc d'être prioritaire dans l'intervention pédagogique.

La structuration rythmique et l'intonation constituent, du point de vue pédagogique, le deuxième volet de la didactique de l'oral. Conformément au statut de l'oralité, la primauté doit en effet être accordée aux composantes suprasegmentales, de par leur fonction structurante, et parce que tout l'être communicant est alors sollicité et appelé à agir ou réagir⁷⁰.

Dès que *l'intention pragmatique* manifestée par l'intonation se précise au plan lexical (« choix des mots ») et grammatical (structuration syntagmatique), le rythme et l'intonation en traduisent la structuration temporelle⁷¹. L'iconicité entre la structure rythmique et intonative du flux

⁶⁹ Elle distingue, par exemple : acte de parole 1 ou acte de parole 2, rituels communicatifs plus ou moins modélisés (saluer, aborder, solliciter, s'excuser, etc. ; cf. supra § 3.3), demande et interrogation, assertion, surprise, adhésion/rejet, etc. L'intonation en est la réalisation la plus immédiate et significative.

Certes, il n'est pas aisé de simuler en classe des situations où la charge affective soit clairement manifeste, mais les onomatopées, les exclamations « Oh ! », « Ah ! »..., certains rituels, les salutations, la prononciation expressive de certains noms de personnalités connues du monde des lettres, du spectacle, des arts, du sport, de la vie mondaine ou de la politique, etc. sont autant de manifestations orales de l'affectivité (qu'elles expriment l'adhésion ou le rejet) que la didactique des langues peut/doit inclure au plus tôt dans les matériels utilisés comme enclencheurs de la démarche d'enseignement-apprentissage.

⁷⁰ Par des mouvements plus ou moins spontanés ou volontaires (hochement de la tête, légère tape de la main, mouvement des bras, de la main, d'un doigt, du buste, etc.). Ces procédures requièrent bien évidemment l'utilisation d'enregistrements, et la perception de ces composantes suprasegmentales de la parole peut être stimulée et vérifiée par des questions apparemment anodines ou même hors de propos : « est-ce un homme, une femme ou un enfant qui parle ? », « Est-ce qu'on pose une question ou que l'on donne une information ? », « le personnage est-il autoritaire ? donne-t-il un ordre ? est-il content ? », etc.

Il va sans dire que ces enregistrements sont des échantillons de langue parlée, la spécificité de la démarche didactique à cette étape prime la perception et « tolère » les déviations (les « erreurs » phonématiques).

⁷¹ Le rythme manifeste une distinction capitale entre les éléments grammaticaux, en particulier les clitiques, qui ne sont jamais accentués et les éléments lexicaux porteurs de charge sémantique, toujours accentués. Cette distinction a son pendant au plan psycholinguistique : comportement automatisé (de type psychosensoriel et psychomoteur) pour la composante grammaticale, et valeur référentielle (donc plutôt de type psycho-cognitif et intellectualisé) pour la composante lexico-sémantique. L'accent à valeur démarcative et/ou culminative, absent des éléments grammaticaux, est déterminant dans les unités lexico-sémantiques, et associé à l'intonation produit des discontinuités essentielles pour la compréhension et pour la locution (cf. § 3.3 ; 4.2 et 4.3, notes 34 et 53).

Alors que les consignes articulatoires n'ont pas prise sur la prosodie de la parole, la Méthode verbo-tonale s'avère efficace car elle propose des procédures opérationnelles qui peuvent agir par le truchement de la gestualité que l'on peut activer d'une manière effective et délibérée.

phonique d'une part et la gestualité d'autre part constitue une donnée capitale pour définir les procédures didactiques.

L'intervention pédagogique sur la perception et production des unités phonématique a sa place en troisième lieu. La littérature consacrée à la phonétique verbo-tonale fournit une importante panoplie de démarches et de procédures concernant la perception et structuration de la matière phonique au plan segmental (cf. Renard, 1979, 2010 ; Billières, 2005 ; De Mann et De Vriendt, 2015 ; etc.) qui se sont avérées particulièrement efficaces, en proposant d'agir sur les principales *valeurs de la langue parlée* – tout spécialement sur la *tension* et la *tonalité* des réalisations phonématiques – au moyen de stimulations somato-sensorielles, d'activation de la macro-motricité et de stimulations par les mouvements⁷².

La perception et production de quelques logatomes, de brèves répliques, de formules ritualisées (aborder quelqu'un, téléphoner, demander l'heure, présenter quelqu'un, etc.) permet de constituer une sorte de corpus phonique de référence qui, dûment engrammé, constitue chez l'apprenant une sorte de diapason phonétique mental de la langue objet d'apprentissage. La poésie, les comptines dans le cas d'enfants (Roberge, 2002), ou les rituels de la communication quotidienne (salutations, demandes d'information, excuses, etc.), à la condition d'avoir été réellement mémorisés, s'avèrent également des références rythmiques, intonatives et phonématiques efficaces où convergent les facteurs psycho-linguistiques, psycho-moteurs (gestualité associée), psycho-sociaux et affectifs.

Mais pour que ces « références phoniques » soient réellement efficaces, leur engrammation est une condition *sine qua non*.

5.2 *La mise en mémoire – l'engrammation – des manifestations phoniques de la parole*

Todorov soulignait que la caractérisation de l'énonciation – qui est d'ailleurs essentiellement orale – s'avère particulièrement problématique du fait qu'il s'agit d'un phénomène fugace. Les manifestations phoniques de la parole, quelle que soit la modalité énonciative (interrogation,

⁷² Il y a lieu toutefois de souligner que certaines pratiques de type analytique qui se réclament de la verbo-tonale (par exemple, la correction de la voyelle isolée [y] prononcée *[u] par la technique dite de « prononciation nuancée ») rappellent par trop les moyens conductivistes dont les limites et les écueils sont notoires.

assertion, modalité jussive), quel que soit le genre de discours (descriptif, narratif, argumentatif, etc.) ou le type d'activité audio-phonatoire (parole spontanée, répétition, réplique, lecture, etc.) constituent de même des réalités par nature fugaces, éphémères et évanescences. Comment dès lors les appréhender et les engrammer ?

D'aucuns préconisent à cet effet la « simple répétition » de modèles : le principe du *crible phonologique* et la « surdité phonologique » (Polivanov, Troubetzkoy, Guberina) en ont montré les écueils et les limites (sans compter que le « réflexe conditionné » que l'on entend ainsi créer est loin d'en garantir la correction ; la critique ironique de Ionesco dans *La leçon* est à cet égard rédhibitoire...).

La « prononciation nuancée », la polysensorialité et l'activation macro-motrice – les mouvements, les rythmes phonétiques et les gestes, y compris lorsqu'il s'agit de signes (cf. *supra* § 2.1 et 4.3 : déictiques, imitations de formes, indications d'orientation, etc.) – s'avèrent des moyens efficaces de déstructuration, saisie et rétention du flux phonique. La valeur mnémonique de la composante affective de la parole est également un facteur essentiel de la mise en mémoire de l'oral.

Mais l'écriture – la *sription*, « l'outil dont on ne peut se passer » – s'avère également un moyen efficace et dans bien des cas incontournable pour fixer l'acquis phonique (cf. *supra* § 3.5).

Le fait que l'on en vienne à *donner plus d'importance à l'écriture qu'au signe vocal* (Saussure *dixit*) et le fait qu'il soit impossible « de faire abstraction de ce procédé » manifestent en effet des comportements qui peuvent être mis à profit, et qui au demeurant ont leur raison d'être : les caractéristiques de l'écrit et de l'oral, outre les différences formelles concernant les marques morphologiques ou textuelles, convoquent des capacités psychosensorielles, psychomotrices, somato-sensorielles et même psychocognitives distinctes et qui ont trait au premier chef à la perception et à la production de ces deux manifestations de la parole ; et il va sans dire, à leur engrammation, i.e., à la mise en mémoire puis à leur re-connaissance.

Il est symptomatique à cet égard que les méthodologues qui ont clairement proposé une approche spécifique de l'écrit soient précisément les « sgavistes » et « verbo-tonalistes » dont la méthodologie se base sur la primauté de l'oral :

Après une soixantaine d'heures de cours, (...) on peut passer à la langue écrite. À ce niveau, l'étudiant sait entendre, distinguer et répéter tous les sons du français ; l'orthographe ne risquera plus d'avoir trop d'influence sur la prononciation; le mot visuel pourra alors aider utilement la mémoire (P. Rivenc, op. cit., 95).

L'argumentaire est clairement explicité :

- pour *passer à la langue écrite*, attendre que l'étudiant « sache » *entendre, distinguer et répéter tous les sons du français* ;
- l'orthographe risque d'influencer négativement la prononciation ;
- le mot visuel [*sic*] aide utilement la mémoire ;
- la scription (l'écriture) précède le décodage (la lecture).

Il y aurait donc un risque à introduire l'écriture avant la maîtrise de « tous les sons » d'une langue. Mais, outre que l'orientation didactique ainsi libellée est à la fois vague et peu réaliste, le problème est posé de telle sorte qu'il traduit *lato sensu* une conception de la parole selon laquelle le flux phonique serait une chaîne de segments sonores discrets – les phonèmes – comme la chaîne des graphèmes. Il s'ensuivrait la possibilité d'établir une *équivalence* entre les manifestations phoniques et leur représentation écrite à l'aide du code de l'orthographe d'usage, ce qui s'avère infondé (cf. *supra* § 2.3)⁷³.

Quoi qu'il en soit, la formulation de la consigne est cependant remarquable en ce que la graphie est mentionnée comme une *aide à la mémoire*. En effet, contrairement aux manifestations phoniques, fugaces et éphémères, les « représentations du signe vocal » (i.e., les images écrites qui s'y réfèrent) sont des objets « permanents et solides » (Saussure, [1916] 1976, [95] 46).

[...] chez la plupart des individus les impressions visuelles sont plus nettes et plus durables que les impressions acoustiques, aussi s'attachent-ils de préférence aux premières (Saussure [1916] 1976, [95] 46-47).

La polysensorialité (vue + ouïe + gestualité) de tout acte langagier est un facteur mnémotechnique important, mais le fait qu'à l'écrit, les

⁷³ Le flux phonique « [n'est pas assimilable à] des perles enfilées sur l'axe du temps » (Boë, 1997), ce qui était dans les années 1950-1960, la position des partisans du fonctionnalisme). La démarche proposée par exemple dans *Voix et Images de France* s'inscrivait tout à fait dans cette perspective : les premières *dictées* consistaient à donner « l'équivalent en orthographe » non pas des énoncés prononcés mais de vocables dont la graphie est « proche » de la structure phonologique du segment correspondant de la chaîne parlée (« midi », « la table », « le bol », etc.).

discontinuités du flux phonique soient catégorisées et visualisées, et que la dimension temporelle de la langue soit transposée en dimensions spatiales (les relations de successivité des sons – par nature éphémères – sont substituées par des relations topologiques permanentes) permet d'asseoir la mémorisation du son. La simplicité des lois topologiques qui régissent l'alphabet fait que la perception visuelle, « chez la plupart des individus », est plus « nette et plus durable que les impressions acoustiques, « l'image graphique » étant perçue comme « un objet permanent et solide »⁷⁴.

Outre ces aspects psycho-sensoriels, il y a lieu de relever des facteurs socioculturels qui « expliquent » également la préséance souvent donnée à l'écrit par rapport à l'oral : l'importance donnée à l'orthographe, le fait que la maîtrise de l'écriture requière d'avoir été scolarisé, ou que la voix de l'autorité (administration, inventaires, grosses de jugement, de notaires, etc.) se manifeste en langue écrite, sont des éléments qui contribuent au « prestige de l'écriture » et jouent un rôle non négligeable dans le comportement « spontané » des apprenants.

Mais si « l'image graphique finit par s'imposer au son », c'est aussi parce que l'écriture est par définition analytique, et que « l'image graphique des mots » – leur représentation spatiale – est perçue comme « un objet permanent et solide » qui ne dépend plus de la mémoire immédiate, ce qui facilite l'explicitation du contenu affectif ou logique des messages et la hiérarchisation des arguments qui structurent un discours. L'écriture possède de ce fait « une valeur cognitive » (A. Prochiantz).

Le recours à l'écriture peut finalement contribuer à ce que l'apprenant et l'enseignant donnent à l'oralité son véritable statut au plan psychologique, dans l'élaboration de l'image mentale de la parole. Ainsi l'importance relative des « erreurs » en langue écrite sera déterminée en

⁷⁴ Les graphies répondent à des lois topologiques relativement simples ("o" = découpage de l'espace en un « intérieur et un extérieur ; "b" = "o" « latéralisé gauche en haut », etc.). Leur saisie se réalise à l'emporte-pièce (cf. la *Gestalt*), et ne requiert ni mémoire immédiate ni processus de structuration comme les manifestations phoniques.

Il y a lieu à cet égard de lever un malentendu : les « verbo-tonalistes » se sont souvent référés à la Psychologie de la Forme (*Gestalt*) pour souligner le caractère global de la perception, mais il va sans dire que celle-ci diffère notablement selon qu'elle concerne des objets représentés dans l'espace comme dans le graphisme (permanence de l'objet) ou qui se manifestent sur l'axe temporel, d'une manière fugace, comme dans la langue parlée.

fonction de l'image phonique (donc mentale) qu'elles manifestent : l'absence d'un pronom clitique sujet requiert une intervention didactique à l'oral, qui aide à créer l'automatisme de cet emploi⁷⁵ ; l'omission d'une marque d'accord (de personne, de genre, de nombre) qui n'a pas de manifestation orale requiert une intervention qui renvoie au plan psychocognitif (d'une manière explicite ou en commutant le syntagme ou l'accord qui ne se manifeste pas à l'oral avec une construction où il a une manifestation phonique, cf. *supra* note 27).

Au plan didactique, une approche qui postule l'analogie oral-écrit fausse certes la démarche d'apprentissage, mais c'est le principe d'analogie, la relation entre ces deux manifestations des langues ainsi posée, qui en est la cause, non la *scripture* :

L'emploi qu'on fait des mots « prononcer » et « prononciation » est une consécration de cet abus [la tendance à prendre l'écriture pour base] et renverse le rapport légitime et réel [sic] existant entre l'écriture et le langage (Saussure F. de, [1916] 1976, 52).

Il y a lieu, il va sans dire, de contrecarrer l'influence de la graphie sur la prononciation en particulier lors de la lecture, et des procédures didactiques restent à établir à cette fin⁷⁶. Mais si le linguiste – et donc l'enseignant et l'apprenant – doit « s'affranchir [sic] du mot écrit » (Saussure F. de, [1916] 1976, 55) dans ses recherches, cela n'implique pas pour autant qu'au plan didactique l'on doive/puisse renoncer à l'écriture, et à mettre à profit « le rapport légitime et réel existant entre l'écriture et le langage ».

⁷⁵ Par exemple, en recourant aux constructions clivées « Moi aussi, je... », « Elle aussi, elle... », qui permettent de bien différencier la valeur morphologique (la marque de la personne dans le syntagme verbal), qui n'est jamais accentuée, de l'emploi du pronom en tant que déictique ou substitut anaphorique.

⁷⁶ L'on peut poser, par exemple, que toutes les langues possédant un code écrit alphabétique fournissent d'innombrables exemples d'orthographe qui traduisent et illustrent le caractère éminemment symbolique de l'écriture et peuvent être mis à profit en didactique pour bien montrer qu'il s'agit d'un code conventionnel utilisé pour laisser une trace de l'image sonore mentale. Pour ce qui est du français : lettres « parasites » ([o] ⇨ "Oh !" ; [a:] ⇨ "Ah !" ; [*pari] ⇨ "Paris" ; [jɛ:R] ⇨ "hier"...) ; même symbole graphique, i.e., même lettre, pour des sons différents ("-- ille" ⇨ mille, famille → [mil] ≠ [famij] ; différentes graphies pour un même son ([sa] ⇨ "sa" ≠ "ça" ; [mɛ:R] ⇨ "mère" ≠ "mer" ≠ "maire" ; [farmasi]≠[fasil] ⇨ "pharmacie" ≠ "facile" ...), digraphes ("eu", "ou", "au", "oi", "ai", "gn"...) ; etc.

À la condition qu'il soit bien établi que la graphie ne « représente » pas la matière phonique, qu'elle n'est pas « la même chose » que le *dit* mais le *symbole* qui représente le *résultat* de la structuration du flux phonique, la *scripture* peut/doit compléter le processus de structuration phonique, elle peut/doit être un exercice pédagogique efficace. En tant que réalité polysensorielle et psycho-cognitive (analytique) qui assure et fixe les manifestations phoniques préalablement structurées et catégorisées en phonèmes, l'engrammation écrite – i.e., la *scripture* – fait partie du statut de l'oralité au plan didactique.

Il n'est pas évident, dès lors, qu'il faille surseoir à recourir à l'écrit même dans l'enseignement-apprentissage à des débutants.

La question est d'ordre méthodologique : puisque le risque vient de la tendance à « renverser le rapport légitime et réel existant entre l'écriture et le langage », il s'ensuit que les procédures d'intervention doivent être établies et ordonnées en fonction du rôle et de la place respective de l'écrit et de l'oral en didactique. Le protocole pédagogique à suivre se doit d'en préciser la séquenciation.

5.3 La séquenciation des procédures d'intervention didactique sur les manifestations phoniques

L'analyse du statut de l'oralité dans l'enseignement-apprentissage des langues a permis de poser que tout acte langagier se réalise d'abord en langue orale (qu'elle soit « pensée » ou « parlée »), mais qu'il n'est pas opérationnel du point de vue didactique d'en circonscrire la description au *dit* car c'est sur le *dire*, et seulement sur le *dire*, que l'intervention pédagogique peut agir. Les descriptions physicalistes – articulatoires ou acoustiques – de la matière phonique, en se limitant à l'analyse du *dit* et aux caractéristiques du flux phonique produit, ne prennent pas en compte les facteurs qui déterminent lesdites caractéristiques. Ces descriptions font ainsi l'impasse sur la dynamique du *dire*, et en particulier sur la temporalité de tout acte langagier ; ce faisant elles ne permettent pas d'établir une progression des procédures à appliquer.

Le statut de l'oralité au plan didactique permet par contre d'établir une telle progression⁷⁷, et d'établir le cadre d'un protocole qui prévoie

⁷⁷ Cette progression est induite de la séquenciation de l'activité de parole mais ne concerne bien évidemment que l'objectif prioritaire de l'intervention pédagogique et donc le niveau de correction visé : il va sans dire, par exemple, que la réalisation intonative, même si elle peut

trois types de procédures fondées sur la séquenciation du processus audio-phonatoire :

A) Niveau suprasegmental :

- *Sensibilisation au F₀ idiosyncrasique* (affects, voix d'homme, de femme...) ;
- *Modalités énonciatives expressives et grammaticales* (surprise, interrogation...) ;
- *L'iconicité rythmique phonique ↔ verbale ↔ corporelle* (accents, pauses...) ;
- *L'iconicité intonation ↔ gestualité⁷⁸, inflexions mélodiques ↔ macromotricités* (gestes) ;
- *Oppositions intonatives* (surprise ≠ interrogation ; discours communiquant ≠ interprétant, énoncés « parenthétiques », etc.).

B) Niveau phonématique :

- *Structuration fondée sur la tonalité : palette sonore, seuils, pôles de structuration phonologique* (aigus/graves/de meilleure audibilité), gammes optimales ;
- *Structuration fondée sur la tension : forte tension* (pôles de « consonanticité » maximale : [p-], [t-], [k-]) ≠ moindre tension ([-p], [-t], [-k] implatifs) ; allophones continus tendus ([f], [θ], [s] ...) ≠ non tendus ([v], [ð], [z]...) ; les approximantes ([-β-], [-ð-], [-ɣ-])⁷⁹ ; etc.

C) Engrammation et fixation des manifestations phoniques :

- *élaboration de « diapasons phoniques mentaux »* ;
- *gestualité associée, « transposition didactique », dramatisation* ;
- *relations d'opposition (corrélats phonologiques) / contraste (métaphonie)* ;
- *« scription », lecture, diction.*

recourir à des stimuli non lexicaux (allongement d'un phonème [aaaaa...], répétition d'une même syllabe [la-la-la-...], etc.) correspond normalement à un énoncé donc à la production et/ou perception d'une suite de phonèmes, et ne peut donc ignorer le niveau phonématique.

⁷⁸ {Intonème *progrédient/conclusif*} ⇔ {mouvements *centrifuges/centripètes*} ; {décrochage intonème *progrédient/conclusif mineur (incise)*} ⇔ {mouvement *direct/indirect*} ; {intonème *continuatif (hésitation)*} ⇔ {mouvement *continu légèrement ascendant*} ; etc.

⁷⁹ Cette proposition de protocole mentionne les *approximantes* (variantes non fonctionnelles). En espagnol et catalan, par exemple, il s'agit des réalisations obligées des consonnes [b], [d], [g] intervocaliques. Elles sont à prendre en compte non seulement lorsqu'elles apparaissent dans la langue cible, mais aussi lorsqu'elles font partie de la langue de base de l'apprenant et non de celle de l'enseignant : il n'est pas exceptionnel que celui-ci ne perçoive pas, dans ce cas, que l'apprenant ne respecte pas la discontinuité (l'occlusion) et réalise une approximante.

Cette suggestion de protocole didactique⁸⁰ n'est pas, bien évidemment, un répertoire de procédures, mais un canevas dans lequel insérer les différentes approches qu'implique le statut de l'oralité au plan didactique, en particulier de par la temporalité de l'activité langagière dont il manifeste la séquenciation.

*

Au terme de la présente approche du statut de l'oralité au plan didactique qui, il va sans dire, n'épuise pas le sujet, il y a lieu de souligner, comme il ressort de l'argumentaire suivi tout au long de cette contribution, que la plupart des procédures auxquelles renvoie le protocole ci-dessus répondent aux orientations de la Méthode Verbo-Tonale, et que leur mise en pratique est à préciser dans la recherche-action que réalise *nolens volens* tout enseignant. La démarche se fonde bien évidemment sur des critères d'efficacité pratique en didactique, mais son efficacité dépend d'abord de ce que la langue orale n'est pas circonscrite au flux sonore de la parole ou à l'activité des organes phonatoire qui le produisent, et de ce que l'enseignement-apprentissage des langues est situé dans la conception anthropologique du langage et de la communication.

Dans cette perspective, la prise en compte du *statut de l'oralité* dans l'enseignement-apprentissage des langues s'impose comme une exigence méthodologique dans la recherche et les applications didactiques afin de ne pas faire l'impasse sur le *dire*, de ne pas circonscire la parole à l'activité des organes phonatoire ou au *flux phonique* produit, et ne réduise pas l'apprenant à une « machine génératrice de sons ».

Au plan didactique, la prise en compte du *statut de l'oralité* s'avère incontournable dans une conception humaniste et éthique de l'être communicant et *a fortiori* de l'enseignement-apprentissage des langues.

⁸⁰ Certaines de ces suggestions sont spécifiques aux procédures pour débutants et applicables dès l'initiation de l'enseignement-apprentissage d'une langue étrangère (sensibilisation au F₀), d'autres concernent les procédures pour étudiants de niveau avancé (sujet communicant/interprétant, parenthétiques, etc.).

Références bibliographiques

- ALARCOS LLORACH, E., Les représentations graphiques du langage, in MARTINET, A. (dir.) *Le Langage*, Paris, Gallimard, Encyclopédie de la Pléiade, 1968, 513-568.
- ALCOBA RUEDA, S., LE BESNERAIS, M., MURILLO, J., Unité tonale et structure prosodique de l'espagnol », *Revue de Phonétique Appliquée, (RPA)* 1992, **105**, 261-286.
- ALCOBA, S., MURILLO, J., Intonation in Spanish, in HIRST, D. et DI CRISTO, A. (éds), *Intonation Systems: A Survey of Twenty Languages*, Cambridge University Press, 1998.
- APOTHELOZ, D., COMBETTES B., NEVEU F., *Les Linguistiques du détachement*, Berne, Peter Lang, 2008.
- ASP, C.W, GUBERINA P., PANSINI M., *Verbotonal Method for Rehabilitating People with Communication Problems*. New York, NY: World Rehabilitation Fund, 1981 ; 2011.
- ASP, C.W, KLINE M., KAZUNARI J. KOIKE, *Verbotonal Body Movements, Translational Speech-Language Pathology and Audiology*. En ligne, PDF, 2006 ; 2011.
- AVANZI, M., *Aspects prosodiques de la dislocation à gauche en français parlé. Étude prosodique, Proceedings of the prosody-discourse interface*, Paris- Jussieu, 2009, 77-91.
- BAQUÉ, L., *Les manifestations phoniques parenthétiques dans le discours spontané : une contribution à l'étude de la polyphonie*, Villeneuve d'Ascq, Presses Universitaires du Septentrion, 1998.
- BENSALAH, A., L'ancrage corporel et praxique du langage. Communauté corporelle et altérité dans l'enseignement d'une « langue » orale étrangère, *Langages*, 2013, **192**, 73-86.
- BERRÉ, M. (éd.), *Linguistique de la parole et apprentissage des langues : questions autour de la méthode verbo-tonale de P. Guberina*, Mons, CIPA, 2005.
- BESSE, H., *Méthodes et pratiques des manuels de langue*, Paris, Didier, 1985 ; 2004.
- BILLIÈRES, M., Les pratiques du verbo-tonal. Retour aux sources, in BERRÉ, M. (éd.), *Linguistique de la parole et apprentissage des langues : questions autour de la méthode verbo-tonale de P. Guberina*, Mons, CIPA, 2005, 67-87.
- BLANCHET, PH., CHARDENET, P. (éd.), *Guide pour la recherche en didactique des langues et des cultures : approches contextualisées*, A.U.F., Paris, Éditions des Archives Contemporaines, 2014.
- BOË, J.-J., Sciences phonétiques et relations forme/substance : I. Un siècle de ruptures, négociations et réorganisations. *Construction des Théories du son, Histoire, Épistémologie, Langage*, 1997, **XIX, Fasc. 1**, pp. 43-72.
- BOLINGER, D., *Intonation and Its Uses: Melody in Grammar and Discourse*, Stanford University Press - Language Arts & Disciplines, 1989.
- BOUCHEZ, J., De la théorie verbo-Tonale à la théorie des rythmes phonétiques, in CALVO, M-V. et MURILLO, J. (éds.), *Perception phonique et parole*. Mons, CIPA, 2009, 107-117.
- CALBRIS, G., PORCHER, L., *Geste et communication*, Paris, CRÉDIF-Hatier, 1989.

- CHARAUDEAU, P. (dir.), *Identités sociales et discursives du sujet parlant*, Paris, L'Harmattan, 2009.
- CORTES, J., Débats et combats, in BORG, S., POCHAT, L. (éd.) *Jacques Cortès, Linguiste, Didacticien et Humaniste*. Paris, Gerflint, 2011, 89-200.
- CREISSELS, D. *Syntaxe générale. Une introduction typologique 2*, Paris, Hermès.
- DAFF, M., MAIGA, A., Démarche convergente et intégrative de conception de bigrammaires, *Approches didactiques du bi-plurilinguisme en Afrique*, Paris, Archives Contemporaines/OIF, 2015.
- DAMASIO, A., L'esprit est modelé par le corps, *La Recherche*, 2003, **368**.
- DE MAN, M.-J., DE VRIENDT, S., RENARD, R., RIVENC, P., BARBÉ, G., COURTILLON, J. *Apprentissage d'une langue étrangère/seconde. 1. Parcours et procédures de construction du sens. 2. La Phonétique verbo-tonale. 3. La Méthodologie. 4. L'Évaluation*, Bruxelles, De Boeck université, 2001, 2002, 2003, 2004.
- DE MAN, M.-J., DE VRIENDT, S. (dir.), *Du son au sens*, Mons, CIPA, 2014.
- DE VRIENDT, S., *Trente ans de didactique des langues*, Bruxelles, Didier Erudition, 1997.
- DI CRISTO, A. Fr., *Les musiques du français parlé. Études de linguistique française*. Berlin/Boston, Gruyter GmbH (section consultable en ligne), 2016.
- DUCROT, O., *Le dire et le dit*, Paris, Éd de Minuit, 1985.
- FANT, G., *Speech Aoustics and Phonetics*. En ligne, 1962 ; 2004.
- FRANKOL D., PAVELIN LESIC, B., *Mouvement, geste, parole : les valeurs de la langue parlée dans la correction phonétique*, Zagreb, Polyclinique SUVAG, 2016. En ligne [PDF]
- GACHET F., AVANZI M., La prosodie des parenthèses en français spontané, *Verbum*, 2008 **XXX**, **1**, 53-84.
- GAUCHOLA R., MURILLO, J., *Iconicité entre macro-motricité et proprioception dans la structuration de la matière phonique en phonèmes*, Mons, CIPA, 2016 (sous presse).
- GERMAIN, C., *La notion de situation en linguistique*, Ottawa, éditions de l'université d'Ottawa.
- GUBERINA, P., La méthode audio-visuelle structuro-globale, *R.P.A.*, 1965, **1**, 35-64.
- GUBERINA, P., La parole dans la Méthode structuro-globale audio-visuelle, *Le Français dans le Monde*, 1974, **103**, 49-54.
- GUBERINA, P., *Rétrospection*. Zagreb, ArTrésor Naklada, 2003. Trad. Espagnole, *Retrospección*, CIPA, 2008.
- GUBERINA, P., RIVENC, P. (dir.), Avertissement, *Voix et Images de France. Livre de l'élève*, Paris-Bruxelles-Montréal, Didier, s.d.
- HAIMAN, J. (ed.), *Iconicity in Syntax*, Amsterdam/Philadelphia, John Benjamins Publ., 1985.
- HECAEN, H., ANGELERGUES, R., *Pathologie du langage*, Paris, Larousse, 1965.
- HJELMSLEV, L., *Prolégomènes à une théorie du langage*. Paris, Éd. Minuit, [1943] ; 1971.

- INTRAVAIA, P. *Formation des professeurs de langue en phonétique corrective. Le système verbo-tonal*. Mons, CIPA, 2003 ; 2007.
- JACOB, Fr., *La logique du vivant, une histoire de l'hérédité*. Paris, Gallimard, 1970 ; 2002.
- JAKOBSON, R., *Six leçons sur le son et le sens*. Paris, Éd. Minuit, 1976.
- JAKOBSON, R., *La charpente phonique du langage*, Paris, Éditions Minuit, 1980.
- KRISTEVA, J., *Polylogue*. Paris, Seuil, 1976.
- LANDERCY, A., Les paramètres acoustiques de l'audition et la perception des sons de la parole, *Revue de Phonétique Appliquée*, 1973, **25**, 3-23.
- NOBILE, L., Introduction. Formes de l'iconicité, *Le français moderne*, 2014, **2014-1** : 1-46
- MARTINET, J., Pour lire et écrire : l'Alfonic, *Communication et langages*, 1975, **25**.
- MELIGNE, D., *La sémantique de l'Action*, Thèse de doctorat Université Toulouse III, 2011. En ligne.
- MESTREIT, Cl., Segmentation du flux phonique en modalité interrogative et assertive, in CALVO, M-V., MURILLO, J. (éds), *Perception phonique et parole*. Mons, CIPA, 2009, 251-267.
- MURILLO, J., ¿A qué gramática recurrir en la enseñanza de la lengua en edad preescolar ?, *Filologija*, Zagreb, 1988, **15**, 153-168.
- MURILLO, J., Des réalisations phoniques dans l'enseignement/apprentissage des langues, in RENARD, R. (éd.), *Apprentissage d'une langue étrangère/seconde ; 2. La phonétique verbo-tonale*, Bruxelles, De Boeck Université, 2002, 279-315.
- MURILLO, J., Logique du vivant et logique de la physique en didactique des langues : l'approche verbo-tonale, *Cahiers du Centre Interdisciplinaire des Sciences du Langage*, 2003, **17**.
- MURILLO, J., Linguistique typologique fonctionnelle et Approche SGAV, *Synergies France*, **2**, 2005.
- MURILLO, J., Définir et évaluer les systèmes d'erreurs de prononciation à partir des pôles de structuration de la matière phonique, *Approche de l'expression orale en didactique des langues*, Université de Tunis, Agence de la Francophonie, 2007, 43-68.
- MURILLO, J., Du fait de style au fait de langue : une approche verbo-tonale de l'interrogation, in PAVELIN LESIC, B., *Francontraste : L'affectivité et la subjectivité dans le langage*. Mons, O.I.F., CIPA, 2013.
- MURILLO, J., Saisir et caractériser les manifestations phoniques de la parole, in MORRIS, D., éd., *La technologie aux limites de l'humain en didactique des langues*, Mons, CIPA, 2015.
- NEBRIJA, A., *Gramática sobre la lengua castellana*. Real Academia Española, PDF, et Barcelona, Galaxia Gutenberg, [1492]; 2011.
- NESPOULOUS, J-L., *Jacques Lordat (1773-1870), la naissance de la "neuropsychologie cognitive du langage"*, Académie des Sciences et Lettres de Montpellier, 2013.

- OSORIO, M., *Enseignement du français langue étrangère et macromotricité : une approche verbo-tonale de la prosodie*, Universitat Autònoma de Barcelona, recherche non publiée, 2016.
- PAVELIN, B., *Le geste et la parole*, Toulouse, Presses Universitaires du Mirail, 2002.
- PICCARDO, E., *Du communicatif à l'actionnel : un cheminement de recherche*. En ligne, PDF, 2014.
- PORQUIER, R., VIVES, R., Sur quatre méthodes audio-visuelles - Essai d'analyse critique, *Langue française* 1974, **24**, 105-122.
- PUREN, C., *Histoire des méthodologies de l'enseignement des langues*. [1988] 2012 éd. PDF.
- PUREN, C., De l'approche par les tâches à la perspective co-actionnelle *Les Cahiers de l'APLIUT* 2004 – **XXIII-1**, 10-26. PDF.
- PUREN, C., Histoire de la didactique des langues-cultures et histoire des idées, *Cuadernos de Filología Francesa Cáceres* 2007, **18**, 127-143.
- RENARD, R., *Introduction à la méthode verbo-tonale de correction phonétique*, Paris, Bruxelles, Mons, Didier, CIPA, 1974³.
- RENARD, R., Sur le recours à la tension en phonétique corrective, *RPA*, 1976, **17**, 71-88.
- RENARD, R., *Structuro-global et verbo-tonal – Variations 1962-2010*, Mons, CIPA, 2010.
- RENARD, R., Changement de paradigme culturel : de l'Un au Multiple, in BAQUE, L., ESTRADA, M. (éd.), *La Langue et l'être communicant*, Mons, CIPA, 2010, 263-277.
- RENARD, R., *Une éthique pour la Francophonie. Questions de géopolitique linguistique*, Mons, Centre International de Phonétique Appliquée – CIPA, 2015³.
- RIVENC, P., GUBERINA, P. (dir.), *Voix et Images de France*, CRÉDIF, Paris, Didier, 1960 ; 1971.
- RIVENC, P., *Pour aider à apprendre à communiquer dans une langue étrangère*, Paris, Didier Érudition ; Mons, CIPA, 2000.
- RIVENC, P., Julio Murillo, la Linguistique et la Didactique des Langues-Cultures, in BAQUE, L., ESTRADA, M. (éds.), *La Langue et l'être communicant. Hommage à Julio Murillo*, Mons, CIPA, 277-287.
- ROBERGE, C. Les intonations logique et affective, les grandes oubliées de l'éducation à la parole, in RENARD, R., (éd.), *Apprentissage d'une langue étrangère/seconde 2. La Phonétique verbo-tonale*, Bruxelles, De Boeck université, 131-154.
- ROSSI, M., *L'Intonation, le système du français. Description et modélisation*, Paris/Gap, Ophrys, 1999.
- ROUSSEAU, J.-J., *Essai sur l'origine des langues, Œuvres complètes II*, Paris, Gallimard, Col. La Pléiade, [1781]1981.
- SAUSSURE, F., *Cours de linguistique générale*. Paris. Payot, 1916, 1976.
- SCHNEIDER, S. *Reduced Parenthetical Clauses as Mitigators. A Corpus Study of Spoken French, Italian and Spanish*, Amsterdam/Philadelphia: John Benjamins, 2007.

SONESSON, G., Prolegomena to a general theory of iconicity, in KLAAS WILLEMS, K., DE CUYPERE, L., *Naturalness and Iconicity in Language*, Amsterdam/Philadelphia, John Benjamins Publishing Company, 2008, 47–72.

TROUBETZKOY, N.S., *Principes de phonologie*. Paris, Klincksieck, [1939] ; 1964.

WIOLAND, F., *Prononcer les mots du français*, Paris, Hachette, Autoformation, 1991.