



Corela

Cognition, représentation, langage

HS-30 | 2020

**Phonétique, littérature et enseignement du FLE :
théories et recherches**

Enseignement de l'intonation et neurolinguistique

Philippe Martin



Édition électronique

URL : <http://journals.openedition.org/corela/10076>

DOI : [10.4000/corela.10076](https://doi.org/10.4000/corela.10076)

ISSN : 1638-573X

Éditeur

Cercle linguistique du Centre et de l'Ouest - CerLICO

Référence électronique

Philippe Martin, « Enseignement de l'intonation et neurolinguistique », *Corela* [En ligne], HS-30 | 2020, mis en ligne le 14 avril 2020, consulté le 23 mai 2020. URL : <http://journals.openedition.org/corela/10076> ; DOI : <https://doi.org/10.4000/corela.10076>

Ce document a été généré automatiquement le 23 mai 2020.



Corela – cognition, représentation, langage est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

Enseignement de l'intonation et neurolinguistique

Philippe Martin

Introduction

- 1 Comme le soulignent Élisabeth Guimbretière (1994) et Bernard Dufeu (2008), l'enseignement de l'oral, et en particulier de l'intonation, reste une partie négligée de l'apprentissage du français langue étrangère. Si les manuels spécialisés de FLE consacrent parfois quelques chapitres à la prononciation des syllabes et de mots particuliers, l'intonation ne fait l'objet, au mieux, que de quelques paragraphes, décrivant les fonctions traditionnelles d'indication de la modalité déclarative et interrogative de l'énoncé et de ses variantes. En réalité, la description linguistique, traditionnellement fondée sur la syntaxe de l'écrit, ne fait pas beaucoup mieux et ne parvient pas à déterminer clairement le rôle, pourtant essentiel, de l'intonation dans la compréhension de la parole. La dominance actuelle de l'approche autosegmentale-métrique (Frota et Prieto, 2015) mélangeant les aspects phonétiques et phonologiques de l'intonation ne contribue pas à clarifier la question. L'absence de consensus sur les modèles phonologiques de l'intonation semble également décourager les auteurs de manuels de FLE, qui se résignent à reproduire les mêmes observations et conseils déjà publiés il y a plus de 50 ans. On peut toutefois citer au moins une exception notable, le manuel « Improving French Pronunciation » publié en 2012.
- 2 Et pourtant, l'intonation de la phrase est essentielle à la fois pour le locuteur et pour l'auditeur. On peut montrer en effet, que loin d'être la cerise sur le gâteau syntaxique, l'intonation dans sa production et dans sa perception précède l'élaboration et le décodage de la structure syntaxique, du moins pour la fenêtre de travail habituelle du locuteur, de l'ordre d'une dizaine de mots (Keating et Shattuck-Hufnager, 2002 ; Blanche-Benveniste, 2003 ; Krovokapic, 2012).

- 3 La description phonologique de l'intonation, tout comme l'acquisition dans le cadre du FLE, se doit de procéder en deux étapes, dont chacune est liée à des propriétés cognitives spécifiques :
- 4 1. Le phrasé, c'est-à-dire la segmentation en groupes accentuels, regroupant une ou plusieurs syllabes dont la dernière est accentuée. Ces groupes peuvent contenir plusieurs mots et leurs caractéristiques temporelles, leur durée et leur débit, sont liées à l'identification des unités morphologiques, dont la temporalité paraît liée à une onde cérébrale spécifique (delta).
- 5 2. Le regroupement progressif incrémental des groupes accentuels pour former des groupes plus grands au cours du déroulement de la lecture ou de l'énonciation de l'énoncé, de manière à constituer la structure prosodique au final. Ce processus de hiérarchisation est rendu nécessaire par les capacités limitées de mémorisation de listes de groupes accentuels, donc les signifiés partiels successifs doivent être réévalués de manière plus compacte pour ne pas être perdus par l'auditeur dans le traitement de leur signification, l'image acoustique d'un segment de parole s'estompant en quelque 3 secondes s'il est suivi d'autres segments de parole (de 20 à 30 secondes s'il est suivi de silence).

Groupes accentuels

- 6 Normalement, on ne lit pas mot à mot, mais par groupes accentuels, c'est-à-dire par séquences de mots dont la dernière syllabe du dernier mot est accentuée en français. Ces groupes accentuels constituent des unités de production et de perception de la parole. Ils peuvent contenir une seule syllabe, un seul mot, un mot de contenu et les mots grammaticaux associés, ou encore des unités plus grandes contenant plus d'un mot de contenu, c'est-à-dire un nom, un verbe, un adverbe ou un adjectif. Dans le cas où ils ne contiennent qu'un seul mot de contenu, ils correspondent aux « accent phrases » de l'anglais ou de l'italien, langues à accent lexical, alors que le français en est dépourvu, ne disposant que d'un accent de groupe. On verra plus loin que la composition des groupes accentuels ne dépend que du débit de parole, pouvant varier dans un intervalle de durées minimale et maximale déterminé.
- 7 L'observation de la parole spontanée montre bien que les groupes accentuels constituent les unités de production et de perception utilisés par le locuteur ou l'auditeur. En effet, comme l'a montré Claire Blanche-Benveniste (2003), les reprises et reformulations d'un groupe se font toujours par un groupe accentuel complet, même s'il n'y a pas de modification des éléments initiaux. Ainsi, dans un exemple tel que « ... votre livre porte sur... sur le Vatican », le locuteur ne finalise pas la prononciation du groupe après le mot « sur » pour permettre la sélection de l'unité lexicale « Vatican », puis reprend l'entièreté du groupe accentuel « sur le Vatican » au lieu de reprendre en ajoutant simplement « le Vatican » (enregistrement Historiquement show, du 14 avril 2018).

Enseignement de la lecture et groupes de sens

- 8 Pourtant, après l'apprentissage de la lecture par syllabes puis par mots, on enseigne encore souvent aux jeunes lecteurs natifs la segmentation des phrases en « groupe de

sens », notion qui se définit comme « constituant un sens ». Or, cette définition pose problème, non seulement pour les linguistes, mais aussi pour les apprenants. La notion de « groupe de sens » dans le contexte de l'apprentissage de la lecture est déroutante car très approximative. Ainsi si *l'armoire* est un groupe de sens, *la petite armoire verte* l'est aussi, et de même *la petite armoire vert-bouteille*. Pourtant, si on considère la prononciation même silencieuse de ce dernier exemple et non pas sa seule représentation orthographique, on s'aperçoit qu'il ne peut être prononcé en une fois, et se segmente en deux groupes comme par exemple *la petite armoire* et *vert-bouteille* dont les dernières syllabes sont accentuées. On retrouve là l'obligation apparente de prononcer plus d'une syllabe accentuée pour des mots très longs, comme *paraskévidekatriaphobie* (« la peur du vendredi 13 »), ou *intergouvernementalisation* prononcés (par exemple) *paraskévidekatriaphobie* et *intergouvernementalisation*.

Durée des groupes accentuels

- 9 Il n'est pas difficile de montrer expérimentalement que les durées des groupes accentuels s'inscrivent dans une gamme allant de 250 ms à 1250 ms environ (Martin, 2014). La durée minimale de l'ordre de 250 ms est déterminée en remarquant que deux syllabes accentuées successives, comme par exemple dans *par le fait que* ou *le travail de nuit nuit*, nécessitent pour être perçues comme accentuées un espacement temporel d'au moins 250 ms. En réduisant suffisamment la seule partie silencieuse de cet espace par un éditeur de signal, la première syllabe cesse d'être perçue comme accentuée, bien qu'aucune de ses caractéristiques acoustiques n'ait été modifiée. La durée minimale d'un groupe accentuel, ne comprenant donc qu'une seule syllabe, doit intégrer cette distance par rapport à la syllabe précédente, nécessairement accentuée en français, alors que sa syllabe unique peut avoir une durée plus courte.
- 10 De même, la durée maximale des groupes accentuels s'observe par exemple dans « le parler jeune » à grand débit de parole (Lehka et Le Gac, 2004 ; Martin, 2014), et dépasse rarement quelque 1250 ms, au prix d'une forte compression de durée syllabique. C'est cette durée maximale qui nous contraint à accentuer plus d'une syllabe dans les mots longs, comme évoqué plus haut.

Groupe accentuel et débit de parole

- 11 Comme on l'aura compris, non seulement le groupe accentuel ne correspond pas nécessairement à un groupe de sens, mais, de plus, ces groupes dépendent du débit de parole. D'un débit très lent résulte des groupes accentuels d'une seule syllabe, un débit un peu moins lent au mot à mot, un peu plus rapide à des groupes accentuels ne comprenant qu'un seul mot de contenu et un débit moyen à plusieurs mots de contenu inclus dans un seul groupe accentuel, la limite étant la durée maximale de 1250 ms. Dans ce dernier cas, il se produit une compression des durées syllabiques pour des groupes accentuels avec plusieurs mots de contenu (Martin, 2014).

Eurythmie

- 12 On a remarqué depuis longtemps (Wioland, 1985), que la segmentation en groupes accentuels avait tendance à être eurythmique, c'est-à-dire à équilibrer la durée des groupes successifs, soit par expansion ou compression syllabique pour compenser les variations du nombre de syllabes d'un groupe au suivant (cas de la parole spontanée), soit en regroupant les mots pour obtenir des nombres de syllabes comparables parfois au prix de la congruence avec les unités syntaxiques (cas de la lecture). Une compression des durées syllabiques moyennes permet de réaliser des énoncés comprenant moins de groupes et plus de mots dans chaque groupe en français, un ralentissement du débit permet au contraire de réaliser des groupes plus longs contenant moins de mots.

Groupes de souffle

- 13 On peut aussi noter que les groupes accentuels et les groupes de souffle ne se correspondent pas nécessairement, bien que dans l'apprentissage de la lecture les deux puissent être liés. Le groupe de souffle est une séquence de mots (ou de syllabes si l'énonciation d'un mot n'est pas terminée) entre deux phases d'inspiration dans le cycle de respiration. Tout l'apprentissage de la lecture orale, outre les groupes accentuels, porte sur le contrôle du cycle de respiration, en allongeant le plus possible la phase d'expiration pendant laquelle on peut produire de la parole, et la phase d'inspiration, où la phonation n'est pas possible. Cet apprentissage vise à permettre la réalisation de plusieurs groupes accentuels à l'intérieur d'un même groupe de souffle, et doit donc viser le contrôle et l'ajustement du débit de parole dans l'intervalle de chaque groupe de souffle, dont la durée est déterminée par de nombreux facteurs physiologiques (fatigue, état émotionnel, etc.).

Un exemple

- 14 Pour illustrer ces observations, examinons un exemple destiné à l'apprentissage de la lecture et découpé en groupes de sens délimités par des barres obliques (disponible sur :

<http://sylvain.obholtz.free.fr/crbst_235_m.html>).

Au cirque, / un petit lion / vient d'arriver. / Tous les matins, / les enfants viennent / le voir dans sa cage. / Le directeur a dit / que le lion / devait présenter un numéro : / il peut marcher, / danser / et faire des équilibres. / Dans le cirque, / il s'entraîne / et les enfants l'applaudissent. /

- 15 Il ne semble pas évident d'expliquer que *les enfants viennent* constitue un seul groupe de sens, plutôt que de considérer *les enfants* et *viennent le voir*, qui semblent pourtant aussi constituer des groupes de sens. Dans la phrase suivante, *Le directeur a dit / que le lion / devait présenter un numéro*, *devait présenter un numéro* forme sans doute un groupe de sens, mais ne peut être prononcé, même silencieusement, sans le segmenter en deux parties : *devait présenter* et *un numéro*, parties qui constituent également des groupes de sens. La raison tient à la durée d'énonciation de ce groupe (la durée virtuelle en voix intérieure), durée qui comme on l'a vu ne peut dépasser 1250 ms environ.

- 16 On peut faire la même constatation dans *il s'entraîne / et les enfants l'applaudissent*. Le groupe de sens *et les enfants l'applaudissent* ne peut être prononcé qu'en deux groupes *et les enfants* et *l'applaudissent*, sauf à procéder à une brusque accélération de débit pour ce dernier « groupe de sens ».
- 17 Dans la phrase *Le directeur a dit que le lion devait présenter un numéro*, une prononciation très lente conduit à une segmentation syllabe par syllabe dont chacune est accentuée : Le / di / rec / teur / a / dit / que / le / lion / de / vaît / pré / sen / ter / un / nu / mé / ro, ou si le tempo est un petit peu plus rapide, une prononciation mot à mot / Le / directeur / a / dit / que / le / lion / devait / présenter / un / numéro /.
- 18 Un débit moyen de parole conduit à segmenter en groupes accentuels comprenant chacun un mot lexical (i.e. verbe, adverbe, nom ou adjectif) : / Le directeur / a dit / que le lion / devait / présenter / un numéro /.
- 19 Une prononciation encore plus rapide mène au regroupement en un seul groupe accentuel de plusieurs mots du contenu, la limite étant déterminée par la durée d'énonciation de chacun des groupes, soit 1250 ms environ, ce qui correspond à 10 syllabes par seconde / Le directeur a dit que le lion / devait présenter un numéro /.

La structure prosodique incrémentale de dépendance

- 20 On comprend l'importance du phrasé dans la réalisation de l'intonation de l'énoncé. En effet, l'apprentissage de l'intonation de la phrase dans le cadre du FLE devrait se faire en deux phases : a) le phrasé, c'est-à-dire la segmentation en groupes accentuels, en procédant par débit de parole progressif, du plus lent (peut-être même le mot à mot), en passant par le débit moyen (groupes accentuels canoniques, contenant un seul mot lexical), au plus rapide, impliquant plus d'un mot de contenu dans les groupes accentuels, la limite étant donnée par la durée maximale de 1250 ms ; et b) l'assemblage incrémental des groupes accentuels successifs en une structure prosodique élaborée au cours du déroulement de la lecture. Cette seconde partie de l'acquisition présuppose évidemment que l'apprenant sache segmenter en groupes accentuels, unités minimales de la structure prosodique.
- 21 La structure prosodique est une construction incrémentale élaborée par le locuteur au fur et à mesure de la production de l'énoncé, et reconstituée par l'auditeur au cours du déroulement de la parole. Cette reconstruction fait appel à un codage implémenté par des mouvements mélodiques à l'endroit des syllabes accentuées (en position finale des groupes accentuels en français). Il en va de même pour le lecteur, qui tend à reconstituer au fur et à mesure de sa lecture la structure prosodique voulue par le scripteur.
- 22 C'est ainsi que se construit l'accès progressif et incrémental au sens pour l'auditeur. La congruence avec la syntaxe, avec les unités syntaxiques ne peut être totale, ne fût-ce que par les différences de profondeur des structures syntaxiques et prosodiques, cette dernière étant nécessairement limitée à trois niveaux (correspondant aux IP « Intonation Phrase », ip « intermediate intonation phrase » et AP « accent phrase » dans la terminologie autosegmental-métrique).
- 23 L'assemblage des groupes accentuels successifs constituant la structure prosodique est encodé par le locuteur par des caractéristiques prosodiques spécifiques portant sur les syllabes accentuées nécessairement en position finale de ces groupes. Ces

caractéristiques sont instanciées par des contours mélodiques présentant en français des différences de hauteur et de pente mélodique (sauf bien sûr dans le cas de la parole chuchotée où d'autres caractéristiques interviennent).

- 24 Dans ce mécanisme, le contour terminal conclusif, en position finale dans l'énoncé, sauf en cas de complément rapporté (Bally, 1944) ou de postnoyau prosodique, constitue la balise de référence en ce qu'elle est perçue par tous les auditeurs comme signalant effectivement une fin d'énoncé, toujours reconnu comme terminal, et ce malgré que les réalisations phonétiques en termes de fréquence fondamentale puissent varier. Toutefois, on peut retenir comme paramètres habituels une chute d'intensité de 6 à 12 dB par rapport aux syllabes accentuées précédentes, le niveau de fréquence fondamentale le plus bas de l'énoncé, et une chute mélodique plus ou moins marquée.
- 25 Les contours terminaux présentent des variantes, habituellement désignées dans la catégorie déclarative par commandement (contour impératif) et implication (contour implicatif, présentant une « bosse » dans le mouvement descendant), et dans la catégorie interrogative par surprise (contour d'exclamation) et doute (contour de doute montant interrogatif suivi d'une courte descente).
- 26 Les autres contours mélodiques, appelés C1, C2 (C0 dénotant le contour terminal) sont caractérisés respectivement par une montée (C1) et une descente (C2) dépassant le seuil de *glissando* (Rossi, 1971), c'est-à-dire dont la variation mélodique est perçue comme telle par les auditeurs, par opposition aux mouvements mélodiques montant ou descendant du contour neutralisé Cn, inférieur au seuil de *glissando*, qui sont perçus comme des tons statiques. Dans ce modèle, il existe également un contour neutralisé C0n plat, terminant les « postnoyaux » prosodiques (Martin, 2015 ; 2018) rencontrés dans des constructions telles que *à la caisse C0 ils se pèsent C0n* ou le classique *c'est le château C0 que j'ai acheté C0n* (dans le sens de « c'est le château et ce n'est pas la maison d'à côté que j'ai acheté »).
- 27 Pour représenter ces différents contours phonologiques, on utilise des flèches d'interprétation plus intuitive : C0 ↓ terminal déclaratif, Ci ↑ terminal interrogatif, C1 ↗ continuité majeure à montée mélodique perçue, C2 ↘ continuité mineure à descente mélodique perçue, Cn → contour neutralisé perçu comme ton statique, C0n ← contour neutralisé à descente mélodique non perçue. La direction « à droite » de ces relations de dépendance est symbolisée par la direction des flèches représentant les différents contours, celle relative au contour C0n étant une dépendance « à gauche » envers un contour terminal C0 déclaratif qui le précède. La continuité mineure présente une descente mélodique perçue et non une montée, qui serait alors identifiée comme un contour C1, c'est-à-dire une continuité majeure. Ce cas se rencontre parfois dans certains styles de parole, en particulier dans le discours politique devant un grand nombre d'auditeurs. On l'a comparé alors aux réalisations des enseignants dans la dictée.

Retour sur l'exemple

- 28 Les regroupements successifs des groupes accentuels sont déterminés par les relations de dépendance « à droite », c'est-à-dire envers un groupe accentuel qui suit dans l'énoncé pour indiquer un regroupement des groupes accentuels concernés. Dans l'exemple suivant, extrait des mêmes exercices de lecture déjà cités :

Dans un château, un magicien vient d'entrer. Le roi l'amène tous les jours dans sa tour d'argent. La reine a dit que le magicien était habile : il peut transformer, réduire les géants, faire rire ou pleurer.

- 29 Les groupes accentuels successifs sont déterminés dans l'hypothèse d'une vitesse de lecture moyenne, avec donc un seul mot lexical par groupe (soulignés dans l'exemple).

Dans un château / un magicien / vient / d'entrer /
 Le roi / l'amène / tous / les jours / dans sa tour / d'argent /
 La reine / a dit / que le magicien / était habile /
 il peut / transformer / réduire / les géants / faire rire / ou pleurer /

- 30 Dans une deuxième étape, on peut faire intervenir l'eurythmie, c'est-à-dire la tendance à équilibrer la durée de lecture de groupes accentuels successifs, ainsi que l'évitement de deux syllabes accentuées successives (ici entre *l'amène* et *tous*), dont la réalisation entraînerait une pause d'au moins 250 ms entre les deux voyelles accentuées successives.

Dans un château 4 syl / un magicien 4 syl / vient d'entrer 3 syl / (vient / d'entrer
 Le roi 2 syl / l'amène 2 syl / tous les jours 3 syl / dans sa tour 3 syl / d'argent 2
 syl / (tous les jours
 La reine 2 syl / a dit 2 syl / que le magicien 5 syl / était habile 4 syl /
 il peut 2 syl / transformer 3 syl / réduire 2 syl / les géants 3 syl / faire rire 2
 syl / ou pleurer 3 syl /

- 31 La troisième et dernière étape établit des regroupements des groupes définis précédemment par la réalisation de contours mélodiques appropriés sur les syllabes accentuées qui terminent chaque groupe.
- 32 Les relations de dépendance indiquées par $C2 \searrow$ envers $C1 \nearrow$ peuvent aussi l'être par les contours neutralisés $C_n \rightarrow$ (les regroupements résultants sont indiqués par des crochets) :

[Dans un château $C1 \nearrow$] [un magicien $C_n \rightarrow$ vient d'entrer $C0 \downarrow$]
 [Le roi $C1 \nearrow$] [l'amène $C2 \searrow$ tous les jours $C1 \nearrow$] [dans sa tour $C_n \rightarrow$ d'argent $C0 \downarrow$]
 [La reine $C1 \nearrow$] [a dit $C_n \rightarrow$ que le magicien $C_n \rightarrow$ était habile $C0 \downarrow$]
 [Il peut $C_n \rightarrow$ transformer $C1 \nearrow$] [réduire $C2 \searrow$ les géants $C1 \nearrow$] [faire rire $C_n \rightarrow$
 ou pleurer $C0 \downarrow$]

- 33 Une autre solution fait appel à la neutralisation de certains contours $C2$ réalisés comme C_n , n'entraînant pas de différence dans la structure prosodique résultante. Ainsi $C_n \rightarrow$ indique une dépendance envers $C2 \searrow$, $C2 \searrow$ envers $C1 \nearrow$ et $C1 \nearrow$ envers $C0 \downarrow$.

[Dans un château $C1 \nearrow$] [un magicien $C_n \rightarrow$ vient d'entrer $C0 \downarrow$]
 [Le roi $C1 \nearrow$] [l'amène $C_n \rightarrow$ tous les jours $C1 \nearrow$] [dans sa tour $C_n \rightarrow$ d'argent
 $C0 \downarrow$]
 [La reine $C1 \nearrow$] [a dit $C_n \rightarrow$ que le magicien $C_n \rightarrow$ était habile $C0 \downarrow$]
 [Il peut $C_n \rightarrow$ transformer $C1 \nearrow$] [réduire $C_n \rightarrow$ les géants $C1 \nearrow$] [faire rire $C_n \rightarrow$
 ou pleurer $C0 \downarrow$]

- 34 Les regroupements font appel à la syntaxe et devraient correspondre aux différentes unités syntaxiques. La première solution est évidemment la plus modulée prosodiquement, car la plus redondante du point de vue de l'indication des relations de dépendance entre groupes accentuels.

Planification de la structure prosodique

- 35 La structure prosodique précède les autres structures, syntaxique, informationnelle, du moins dans un empan restreint de deux ou trois groupes accentuels. En parole

spontanée, même silencieuse, le locuteur élabore une structure prosodique partielle avant d'y insérer une tournure syntaxique et sélectionner ensuite les éléments lexicaux qui vont s'y intégrer (cf. Blanche-Benveniste, 2003). En effet, la réalisation d'un contour descendant par exemple $C2 \searrow$ implique nécessairement l'occurrence planifiée ultérieure d'un contour montant $C1 \nearrow$, celle d'un contour $C1 \nearrow$ présuppose un contour terminal conclusif $C0 \downarrow$ qui doit apparaître plus tard, le contour $Cn \rightarrow$ étant neutre par rapport à cette attente de contour dans l'énoncé, qui peut être aussi bien $C2 \searrow$, $C1 \nearrow$ que $C0 \downarrow$. Ces réalisations spécifiques ne peuvent se faire sans planification préalable de la part du locuteur.

- 36 En lecture, l'exemple Le roi l'amène tous les jours dans sa tour d'argent peut donner lieu à différentes solutions de regroupements incrémentaux, tels que

[Le roi l'amène] $C1 \nearrow$ [tous les jours dans sa tour] $C1 \nearrow$ [d'argent] $C0 \downarrow$

- 37 ou

[Le roi l'amène] [tous les jours] $C1 \nearrow$ [dans sa tour d'argent] $C0 \downarrow$

- 38 la congruence syntaxique étant sans doute la moins probable, car elle implique que le lecteur connaisse déjà le texte de la phrase, et résiste à la tentation de regrouper au moins les deux premiers groupes accentuels *Le roi* et *l'amène* :

[Le roi] [l'amène tous les jours] [dans sa tour d'argent] $C0 \downarrow$

La structure prosodique et la lecture

- 39 En lecture, la structure syntaxique et le lexique sont évidemment imposés par le texte, mais, sauf si le texte a déjà été analysé auparavant, le lecteur ne découvre la syntaxe que progressivement et seulement pour les deux ou trois groupes accentuels à venir lors de son énonciation. Il en résulte une possible non congruence due à la nature incrémentale du processus de regroupement des groupes accentuels le long de l'axe temporel.
- 40 Les stratégies prosodiques, différentes en lecture et en spontané, peuvent s'expliquer par les mouvements oculaires en lecture, qui ne dépassent pas quelque 20 à 30 caractères dans l'exploration du texte. N'ayant pas nécessairement connaissance de l'entièreté d'une phrase relativement longue au cours de sa lecture, le lecteur a tendance à rapprocher les unités les plus proches à droite au fur et à mesure de leur découverte visuelle, et regrouper dans un même groupe accentuel des mots qui ne feraient pas partie d'une même unité syntaxique. On a ainsi observé expérimentalement (Martin, 2018) que lors de la lecture de la phrase *deux alpinistes allemands ont trouvé le cadavre d'un homme dans un glacier*, les lecteurs préfèrent regrouper *ont trouvé le cadavre* et *d'un homme*, alors que syntaxiquement les groupes devraient être *ont trouvé* et *le cadavre d'un homme*.

Structure prosodique et neurolinguistique

- 41 Le cerveau est constitué d'un nombre considérable de neurones (de l'ordre de 100 milliards) interconnectés par groupes (les ensembles neuronaux) dans des régions spécifiques de la masse cérébrale. Ces interconnexions réalisées par des chaînes de synapses permettent un transfert de l'information stockée chimiquement dans chaque neurone. Ce transfert induit des variations de potentiel électrique faible (de l'ordre du

μV , le millionième de Volt) observables par des capteurs positionnés sur le crâne des sujets (électroencéphalographie, EEG). Ces transferts ont lieu en tout temps, même pendant le sommeil, mais lorsqu'ils résultent d'une stimulation tactile, auditive, visuelle, olfactive ou gustative, on parle de potentiels évoqués, c'est-à-dire provoqués par une stimulation extérieure.

- 42 Les oscillations neuronales résultent d'un mécanisme d'excitation-inhibition d'ensembles neuronaux qui les lie à d'autres ensembles neuronaux par, entre autres mécanismes, un système de couplage en boucle à verrouillage de phase. Ce couplage synchronise les différentes oscillations des zones neuronales de fréquences différentes, une des fréquences étant asservie à une autre de manière à ce que les périodes de l'une soient un multiple entier des périodes de l'autre.
- 43 L'activité électrique de transfert de groupes de neurones à d'autres groupes de neurones ne se fait pas de manière désordonnée avec des fréquences d'oscillation neuronales réparties au hasard. D'une part, ces transferts opèrent selon des gammes de fréquences spécifiques liées à des activités cognitives déterminées, et d'autre part ils sont synchronisés en phase dans chaque gamme de variation.

Fréquence et période

- 44 Chaque gamme d'onde cérébrale est désignée par une lettre grecque : ondes alpha, bêta, delta, gamma, etc., non seulement d'après leur gamme de fréquence, mais aussi selon les fonctions de transmission des informations observées entre zones neuronales. Les gammes d'ondes cérébrales dans le domaine des stimuli de parole sont les ondes delta, variant de 0,8 à 4 Hz, les ondes thêta, variant de 4 à 10 Hz et les ondes gamma de 30 à 70 Hz (ordre de grandeur).
- 45 Pour les ondes delta, oscillant entre 0,8 Hz et 4 Hz, la période et donc le délai de transmission entre masses cérébrales concernées est de l'ordre de 250 ms à 1250 ms, pour les ondes thêta, variant de 4 Hz à 10 Hz, de l'ordre de 100 ms à 250 ms, pour les ondes gamma, variant de 30 Hz à 70 Hz, de 15 ms à 33 ms, valeurs qui peuvent varier légèrement d'un individu à l'autre.
- 46 L'étendue des variations de période de thêta correspond donc relativement bien à celle des intervalles entre syllabes inaccentuées, de 100 ms à 250 ms, et celles de delta à celle des intervalles entre syllabes accentuées, de 250 ms à 1250 ms. Ces valeurs dépendent de délais de transmission entre zones neuronales, propres à chaque individu.
- 47 Les paragraphes qui suivent examinent (succinctement) les liens entre les ondes cérébrales thêta et delta et les unités phonologiques que sont les syllabes et les groupes accentuels en parole spontanée et en lecture (orale ou silencieuse).

Les syllabes

- 48 Les syllabes ont une durée minimale de 100 ms et maximale de 250 ms, durée au-delà de laquelle elles tendent à être perçues comme accentuées.
- 49 Ces valeurs correspondent à celles des durées minimales et maximales des oscillations thêta. La valeur minimale de 100 ms est la durée minimale de traitement de l'information syllabique synchronisée par thêta, comme l'ont montré Ghitza (2011)

ainsi que, entre autres, Gilbert et Boucher (2007), Giraud and Poeppel (2012). D'autre part, en excédant la valeur maximale de 250 ms, la syllabe ne peut plus être prise en compte par thêta et entre dans la fenêtre temporelle de delta pour être perçue comme syllabe accentuée.

- 50 Des syllabes successives doivent être séparées d'au moins 250 ms pour être perçues comme accentuées. Si elles sont séparées par moins de 250 ms, la première des deux est perçue comme inaccentuée (Martin, 2015 ; 2018). Les syllabes accentuées (donc acoustiquement différentes par un trait quelconque des autres syllabes) sont liées aux oscillations delta. Si une syllabe accentuée survient avant la fin d'un cycle de delta en cours, ce cycle est interrompu, et un nouveau cycle commence, synchronisé par la deuxième syllabe accentuée. Le traitement de l'accentuation de la première syllabe n'aboutira pas quand bien même les facteurs acoustiques seraient présents, et sera remplacé par celui relatif à la deuxième syllabe accentuée.
- 51 Les syllabes sont perçues comme accentuées soit directement à partir d'une combinaison de facteurs acoustiques (durée syllabique, mouvement mélodique, intensité), soit par prédiction par l'auditeur de leur position attendue dans le groupe accentuel, suivi d'une validation à partir des facteurs acoustiques. On peut remarquer que la perception des syllabes accentuées a lieu même en l'absence de traits acoustiques (en lecture silencieuse par exemple). À l'inverse, la présence des traits acoustiques ne suffit pas toujours dans la parole brouillée ou reproduite à l'envers en inversant le sens de l'enregistrement. La prédiction de l'occurrence d'un certain nombre de groupes accentuels probables en cours d'audition semble donc nécessaire, ce nombre dépendant des informations relevant du contexte et de la situation connues par l'auditeur, nombre qui peut donc être plus important en début de tour de parole. La connaissance des règles phonologiques permet à l'auditeur de réduire la combinatoire des possibles, la perception de la syllabe accentuée résultant en définitive d'une validation des traits acoustiques perçus, ou à défaut de ceux qui, absents, devraient être présents.
- 52 On peut citer à ce sujet l'expérience sur la perception des syllabes accentuées du berbère par des sujets qui ne connaissaient absolument pas cette langue (Mettouchi et al., 2007). Les traits acoustiques sont présents dans le signal de parole, mais les auditeurs n'identifient presque aucune syllabe accentuée.

Les groupes accentuels

- 53 Une durée minimale de 250 ms (si on intègre la pause précédant la syllabe accentuée) pour des groupes accentuels d'une seule syllabe (nécessairement accentuée), et une durée maximale de 1250 ms avec un nombre de syllabes ne pouvant excéder 9 ou 10. Les périodes des ondes delta varient de 250 ms à 1250 ms. Les ondes delta étant entraînées par les syllabes accentuées, celles-ci doivent également être espacées de plus de 250 ms et de moins de 1250 ms. Les groupes accentuels ne contenant qu'une seule syllabe accentuée (hors accent d'insistance), en position finale du groupe en français, doivent donc avoir une durée comprise entre 250 ms et 1250 ms, quel que soit le nombre de syllabes qu'elles contiennent (Martin, 2014 ; 2015).
- 54 Un débit syllabique minimal de 4 syllabes / s (au-delà duquel toutes les syllabes sont perçues comme accentuées) et maximal d'environ 10 syllabes / s. Ces valeurs minimales et maximales de débit syllabique (hors syllabes accentuées) correspondent à l'empan de

variation des périodes des oscillations thêta de 4 à 10 Hz, assurant le traitement des informations syllabiques. En deçà, les syllabes sont perçues comme accentuées, au-delà elles ne sont plus identifiées, le débit de parole étant trop rapide.

Lecture

- 55 En lecture silencieuse des groupes accentuels, les syllabes finales sont perçues comme accentuées par le lecteur bien que les corrélats acoustiques de l'accent soient absents. Lors de la lecture, la prédiction de la position des syllabes accentuées par le lecteur est toujours possible du fait de sa connaissance de la règle d'accentuation sur la dernière syllabe du groupe accentuel. La vitesse de lecture va déterminer le nombre de syllabes des groupes accentuels lus et dès lors la position de la syllabe accentuée. Il en va de même lors de l'audition de parole lue ou spontanée, si les conditions de prédiction sont remplies, c'est-à-dire s'il est possible d'estimer la composition d'un groupe accentuel en cours de réalisation. Cette prédiction est plus difficile au début d'un tour de parole, ou impossible pour une langue inconnue ou lorsqu'on brouille ou inverse le signal de parole (cf. Arnal et Giraud, 2017).
- 56 La lecture aussi bien orale que silencieuse procède non par mots mais par groupes accentuels dont la durée totale de prononciation, s'ils étaient oralisés, ne dépasse pas 1250 ms.
- 57 Ceci est une conséquence directe de la lecture par groupes accentuels. La lecture, tout comme le monologue intérieur ou la parole spontanée, ne peut se faire sans production régulière de syllabes accentuées, en voix intérieure ou non, et qui ne peut donc avoir lieu sans synchronisation périodique des oscillations delta, qui s'avère probablement être l'horloge maîtresse dans le système des ondes cérébrales impliquées dans la production et la perception de la parole (voir discussion sur ce point dans Martin, 2018).
- 58 En lecture silencieuse, la durée minimale de lecture d'un groupe accentuel est de 250 ms.
- 59 Bien que les mouvements oculaires par saccades permettent la détection de groupes accentuels plus rapidement (Quercia, 2010), la lecture de chaque groupe ne peut se faire en moins de 250 ms même en lecture silencieuse « rapide » (speed reading) en lisant tous les mots. Cette observation suggère que le traitement linguistique de chaque groupe accentuel par le lecteur implique les ondes delta, dont la période la plus courte est précisément de 250 ms (4 Hz).
- 60 En lecture silencieuse comme en lecture orale, il y a restitution des syllabes accentuées dans chaque groupe accentuel lu, on peut faire l'hypothèse que ce sont les ondes delta qui, en plus de la synchronisation des ondes thêta, assurent la synchronisation du traitement linguistique des groupes accentuels, peut-être sous la forme d'une validation dans une comparaison du groupe accentuel lu avec celui prédit par le lecteur.

Eurythmie

- 61 L'eurythmie se manifeste de manière à limiter les variations de durée des groupes accentuels successifs soit par expansion ou compression des durées syllabiques

moyennes de ces groupes (parole spontanée), soit en regroupant les mots du texte de manière à obtenir un nombre de syllabes variant peu d'un groupe au suivant (parole lue ; cf. Wioland, 1985). Il est possible, c'est une hypothèse, que cet effet soit dû à la tendance des ondes delta à peu varier la durée des périodes d'oscillation successives après détection d'un certain nombre (2 ou 3 ?) groupes accentuels en début d'énoncé.

Conclusion

- 62 La description phonologique de l'intonation de la phrase telle qu'étayée par les découvertes récentes relatives aux capacités cognitives des locuteurs et, en particulier, aux ondes cérébrales, conduit à considérer un enseignement dans le cadre du FLE ou celui, plus large, de l'apprentissage de la lecture, en deux phases :
- 63 1. Acquisition de la segmentation en groupes accentuels fondée, au départ, sur l'identification des mots lexicaux (verbes, adverbes, noms et adjectifs) et sur les effets du débit de parole (d'un débit lent segmentant la phrase en mots, à un débit rapide groupant plusieurs mots lexicaux dans un seul groupe accentuels).
- 64 2. Réalisation des contours mélodiques appropriés à l'endroit des syllabes accentuées finales des groupes accentuels pour indiquer la construction itérative d'une structure prosodique au cours de l'énonciation.
- 65 Bien plus qu'apparaître comme un module d'apprentissage superflu, l'enseignement de l'intonation de l'énoncé s'avère être pour les apprenants indispensable à la compréhension et à la production de l'oral bien plus sans doute que bien d'autres faits de morphologie et de syntaxe.

BIBLIOGRAPHIE

BALLY C., 1944, Linguistique générale et linguistique française, Francke, Berne.

BLANCHE-BENVENISTE C., 2003, « La naissance des syntagmes dans les hésitations et répétitions du parler », in J.-L. ARAOUI (dir.), *Le sens et la mesure. De la pragmatique à la métrique Hommages à Benoît de Cornulier*, Paris, Honoré Champion, coll. « Colloques, congrès et conférences sur les Sciences du langage », p. 40-55.

DUFEU B., 2008, « L'importance de la prononciation dans l'apprentissage d'une langue étrangère », Franc-parler.org. Disponible sur : <http://www.franccparler-oif.org/images/stories/dossiers/phonetique_dufeu3.htm>.

FROTA S. et PRIETO P. (dir.), 2015, *Intonation in Romance*, Oxford, Oxford University Press.

GHITZA O., 2011, « Linking speech perception and neurophysiology : speech decoding guided by cascaded oscillators locked to the input rhythm », *Frontiers in Psychology*, n° 2, p. 130.

GILBERT A. et BOUCHER V., 2007, « What do listeners attend to in hearing prosodic structures ? Investigating the human speech-parser using short-term recall », *Interspeech 2007*, p. 430-433.

Disponible sur : <https://www.isca-speech.org/archive/archive_papers/interspeech_2007/i07_0430.pdf>.

GIRAUD A-L. et POEPEL D., 2012, « Cortical oscillations and speech processing : emerging computational principles », *Nature neuroscience*, vol. 15, p. 511-517.

GUIMBRETIERE É., 1994, *Phonétique et enseignement de l'oral*, Paris : Didier.

KEATING P. et SHATTUCK-HUFNAGE S., 2002, « A Prosodic View of Word Form Encoding for Speech Production », *UCLA Working Papers in Phonetics*, n° 101, p. 112-156. Disponible sur : <<https://linguistics.ucla.edu/people/pakssh/pdf>>.

KRIVOKAPIC J., 2012, « Prosodic planning in speech production », in S. FUCHS, M. WEIHRICH, D. PAPE et P. PERRIER (dir.), *Speech planning and dynamics*, Berne, Peter Lang, coll. « Speech Production and Perception », p. 157-190.

LEHKA I. et LE GAC D., 2004, « Étude d'un marqueur prosodique de l'accent de banlieue », *Actes des XXIII^e Journées d'Études sur la Parole*, Fès, Maroc. Disponible sur : <http://www.afcp-parole.org/doc/Archives_JEP/2004_XXVe_JEP_Fes/actes/jep2004/Lehka-LeGac.pdf>.

LÉON M., GEZUNDHAJT, H., MARTIN, P. et MICHAËL E., 2012, *Improving French Pronunciation*, Toronto, New Millenium Edition.

LINNELL P., 2005, *The Written Language Bias in Linguistics*, London, Routledge.

MARTIN P., 1975, « Analyse phonologique de la phrase française », *Linguistics*, n° 146, p. 35-68.

—, 2014, « Spontaneous speech corpus data validates prosodic constraints », in N. Campbell, D. Gibbon et D. Hirst (dir.), *Proceedings of the 6th conference on speech prosody*, p. 525-529.

—, 2015, *The Structure of Spoken Language*, Cambridge, Cambridge University Press.

—, 2018, *Intonation, structure prosodique et ondes cérébrales*, London, ISTE, coll. « Sciences cognitives ».

METTOUCHI A., LACHERET-DUJOUR A., SILBER-VAROD V. et IZRE'EL S., 2007, « Only Prosody ? Perception of speech segmentation in Kabyle and Hebrew », *Nouveaux cahiers de linguistique française*, vol. 28, p. 207-218. Disponible sur : <http://clf.unige.ch/files/6614/4102/7522/17-Mettouchi_nclf28.pdf>.

QUERCIA P., 2010, « Ocular movements and reading : a review », *J. Fr. Ophtalmologie*, vol. 33, n° 6, p. 416-423.

ROSSI M., 1971, « Le seuil de glissando ou seuil de perception des variations tonales pour la parole », *Phonetica*, n° 23, p. 1-33.

WIOLAND F., 1985, *Les structures rythmiques du français*, Genève-Paris, Slatkine-Champion.

RÉSUMÉS

Comme le soulignent Élisabeth Guimbretière (1994) et Bernard Dufeu (2008), l'enseignement de l'oral, et en particulier de l'intonation, reste une partie négligée de l'apprentissage du français langue étrangère. Si les manuels spécialisés de FLE consacrent parfois quelques chapitres à la prononciation des syllabes et de quelques locutions particulières, l'intonation ne fait l'objet, au mieux, que de quelques paragraphes, décrivant les fonctions traditionnelles d'indication de la modalité déclarative et interrogative de l'énoncé et de ses variantes. En réalité, la description linguistique, toujours fondée sur la syntaxe de l'écrit, ne fait pas beaucoup mieux et ne parvient pas toujours à déterminer clairement le rôle, pourtant essentiel, de l'intonation dans la

compréhension de la parole. La dominance actuelle de l'approche autosegmentale-métrique mélangeant les aspects phonétiques et phonologiques de l'intonation ne contribue pas à clarifier la question.

Or, les découvertes récentes portant sur les contraintes déterminées par les ondes cérébrales, et, en particulier, les oscillations delta (variant de 0,8 à 4 Hz), thêta (de 4 à 10 Hz) et gamma (de 30 à 80 Hz) permettent non seulement de rendre compte des variations rythmiques des groupes accentuels en français, mais aussi de mieux comprendre la nécessité d'avoir une structure prosodique encodée selon le système spécifique de chaque langue. Le fait que cette structure prosodique apparaisse avant la syntaxe et la sélection lexicale autant en production qu'en perception, et ce, aussi bien en parole spontanée qu'en parole lue, justifie amplement la nécessité impérieuse d'enseigner le système prosodique de la langue en question avant même le lexique et la syntaxe.

As emphasized by Élisabeth Guimbretière (1994) and Bernard Dufeu (2008), the teaching of pronunciation, and in particular intonation, remains a neglected part of learning French as a foreign language. While FSL textbooks sometimes devote a few chapters to the pronunciation of syllables and words, intonation takes, at best, only a few paragraphs, describing the traditional functions of intonation indicating the declarative or interrogative modality of the sentence, and their variants. In fact, linguistic description, still based on written text syntax, does not do much better and fails to clearly determine the role, however essential, of intonation in understanding speech. The current dominance of the autosegmental-metrical approach, mixing the phonetic and phonological aspects of intonation, does not help to clarify the issue.

However, recent discoveries concerning the correlation between stress groups stressed syllables and brain waves, and in particular delta (varying from 0.8 to 4 Hz), theta (from 4 to 10 Hz) and gamma (from 30 to 80 Hz) accounts not only for the rhythmic variations of stress groups in French, but also allows to better understand the reasons to have a prosodic structure specific to every language. The fact that the prosodic structure pre-exists syntax and lexical selection both in production and perception, in spontaneous and read speech, justifies the imperative need to teach the prosodic system of the language in question before lexicon and syntax.

INDEX

Mots-clés : Intonation, structure prosodique, ondes cérébrales, groupes accentuels, contours mélodiques.

Keywords : Intonation, prosodic structure, brain waves, accent phrase, melodic contours.

AUTEUR

PHILIPPE MARTIN

LLF, UFRL, Université Paris Diderot