

La langue, la voix, la parole

Ouvrage publié avec le concours de l'Université Paris-Sorbonne
Cellule de Recherche en Linguistique (Pages 10-16)
Textes édités par Prof. Amr Helmy Ibrahim
Paris 2014

Quand la prosodie précède le sens : De la contagion musicale à la didactique des langues

Branka Zei Pollermann

Directrice du [Vox Institute](#) – Genève

Anciennement chercheuse à l'Université de Genève (Département de psycholinguistique) et aux Hôpitaux universitaires de Genève (Département de médecine psychosomatique)

La parole - comme la musique - se déroule dans le temps. Sa structure prosodique la rend intelligible au niveau morpho-phonologique. Dans le flux verbal, le récepteur se sert des marques prosodiques et du contexte pour accéder au sens de l'énoncé entendu. On peut aisément émettre l'hypothèse que la perception des traits contrastifs d'un énoncé (accentuation, contours mélodiques, durées, césures) précède le repérage du contenu lexical. Or dans la didactique des langues, la prosodie reste encore négligée et l'enseignement de la prononciation est souvent centré sur des phonèmes isolés - indépendamment de leur contexte prosodique.

Dans cet article, je décrirai (1) quelques mécanismes cognitifs et neurologiques impliqués dans la perception de la parole et (2) mon adaptation d'une méthode de français langue étrangère.

Le rythme est un phénomène contagieux

Dans la musique aussi bien que dans la parole, le rythme représente une organisation des événements sonores dans le temps. Pour Platon « l'ordre dans le mouvement porte le nom de rythme » (Platon, 665a). Aussi, la composante motrice du rythme est implicite dans sa définition. Dès lors, le rythme sonore apparaît comme un phénomène acoustico-moteur et sa perception auditive implique une activation du système moteur lui correspondant.

Selon Schlaug (2009), les stimulations musicales peuvent activer le système locomoteur moyennant les réactions des neurones-miroir qui en font partie. Aussi, la découverte de neurones miroir (Gallese & Lakoff, 2005) nous donne une explication neurologique de la nature contagieuse du rythme musical. Trainor et Schmidt (2003) ont démontré que le rythme, l'intensité et la hauteur musicale sont d'abord traités séparément et ensuite recomposés pour fournir une « compréhension » globale de la phrase musicale (Peretz & Zatorre, 2005).

Pour Marin Mersenne (1636) l'organisation rythmique de la parole et celle de la musique sont intimement liées. Dans son ouvrage *l'Harmonie Universelle*, il insiste sur la nécessité de faire coïncider les mouvements prosodiques et les rythmes musicaux. Il y consacre deux parties du « Livre Sixième de l'art de bien chanter à la musique accentuelle ».

Dans la chaîne parlée - comme en musique - le rythme est la dimension temporelle sur laquelle se manifestent les traits contrastifs : des « temps forts » et des « temps faibles » ; les syllabes proéminentes vs. les syllabes non-proéminentes. La perception de la proéminence (accents lexicaux et accents secondaires) est liée aux contrastes acoustiques réalisés aux niveaux de la durée, de la hauteur mélodique et de l'intensité des syllabes (Martin 2009).

Comme le rythme musical, le rythme prosodique est contagieux. Nombreuses sont les études qui ont démontré que l'enfant (dès 6 semaines) synchronise/coordonne ses mouvements corporels avec le rythme de la parole de l'adulte (Kendon, 1972 ; Jaffe et coll. 2001). On constate ainsi que la perception des aspects rythmiques et musicaux de la parole précède l'acquisition du langage. L'enfant repère et s'approprie les éléments suprasegmentaux bien avant de parler. Condon (1984) a mis en évidence une isomorphie temporelle entre la parole du locuteur et le comportement kinésique de l'auditeur. Les études récentes sur l'émergence du langage (protolangage) ont montré que le langage a évolué parallèlement au développement des séquences complexes des mouvements corporels (Meyer, et al 2011).

Le rythme apparaît alors comme un *régulateur perceptif*, qui structure la production et la réception du message verbal (Bordas 2003). Henri Meschonnic (1982, pp. 69-70) conçoit le rythme comme « une structure », qui est l'organisation même du sens dans le discours.

La « contagion » rythmique est actuellement étudiée sous deux angles : (1) la physique des systèmes dynamiques non-linéaires et (2) les sciences neurologiques.

(1) La physique explique le phénomène de synchronisation des oscillateurs (métronomes par exemple) qui s'influencent mutuellement et finissent par se mettre progressivement en phase (Schöner G. 2002). Aussi, la synchronie interactionnelle est un phénomène semblable. Dans une interaction verbale, une production vocale relie un individu à un autre et le rythme de l'interaction structure leur temps commun.

(2) Les neurosciences éclairent le phénomène d'influence interactionnelle par la découverte des neurones miroirs - une catégorie de neurones qui s'activent non seulement lorsqu'on exécute une action mais aussi lorsqu'on observe ou entend un autre individu faire une action semblable. Aussi, les neurones-miroirs font partie des mécanismes perceptifs de l'homme (Gallese & Lakoff, 2005 ; Grafton, et al. 1997 ; Murata et al. 1997). Ce phénomène a déjà été signalé par Titchener (1909) qui a observé que le décodage des photographies de visages provoquait chez l'observateur-décodeur des mimiques analogues préalablement à la réponse verbale.

A l'époque actuelle, la découverte des neurones-miroirs apporte une « preuve matérielle » à l'épistémologie génétique de Jean Piaget (1961). Ce dernier conçoit la relation entre le sujet percevant et les caractères de l'objet perçu sous forme d'assimilation de l'objet

aux schèmes du sujet (schèmes sensorimoteurs pour commencer). Aussi selon Piaget, dès la naissance, un objet visuel et/ou sonore n'est perçu qu'en correspondance avec les schèmes sensori-moteurs du sujet. Les sons de la voix d'autrui sont perçus au moyen d'activation des schèmes audito-vocaux correspondants (par ex. la contagion vocale chez des bébés). De plus, les schèmes de la phonation et de l'ouïe s'assimilent réciproquement et « ...la voix des autres est perçue activement, c'est-à-dire assimilée aux schèmes de la phonation » du sujet percevant (Piaget, 1975, p. 83). C'est donc grâce à la découverte des neurones-miroirs qu'on comprend les bases neurologiques du fait que la perception des mouvements d'autrui et/ou de leurs concomitants auditifs implique une activation des schèmes sensorimoteurs du sujet.

Cette découverte apporte également un soutien décisif à la théorie motrice de la perception du langage : le récepteur perçoit les phonèmes d'une langue et les identifie par rapport aux mouvements articulatoires impliqués dans leur production (Libermann et al 1967 ; Libermann et Mattingly 1985 ; Liberman & Wahlen, 2000). Aussi le récepteur et l'émetteur partagent les représentations / images motrices des sons du langage.

Fadiga et Craighero (2003) ont mis en évidence le fait que pendant l'écoute des mots dont la prononciation implique les mouvements de la langue, on observe une activation accrue des potentiels évoqués (moteurs) provenant des muscles de la langue de l'auditeur. Ils concluent que la perception de la parole implique l'activation des centres moteurs spécifiques à la production du langage.

L'idée en soi n'est pas nouvelle, car le linguiste polonais Jan Niecisław Ignacy Baudouin de Courtenay (1845 –1929) a élaboré une *théorie du phonème et de ses variations* – théorie qui conçoit que la perception des phonèmes implique une reconstruction mentale des intentions articulatoires de l'émetteur – un programme mental qui coordonne les gestes articulatoires. Il semble alors que les accents étrangers résultent - au moins en partie - d'une *assimilation déformante* des caractéristiques phono-articulatoires des sons de la langue étrangère aux schèmes phono-articulatoires du sujet.

Comme expliqué plus haut, la perception de l'intonation et du rythme (proéminences, césures, durées) est immédiate et relativement facile. Cette facilité de la perception et de la reproduction des traits contrastifs de la parole trouve son application évidente dans la didactique des langues.

La didactique des langues

Pour apprendre à parler une langue étrangère il convient donc de commencer par l'apprentissage des traits contrastifs de la parole. Comme c'est le cadre suprasegmental qui détermine nombre de réalisations phoniques, l'acquisition de la prosodie devrait précéder l'apprentissage de la phonologie. L'accent lexical, l'accent secondaire, le groupe accentuel (qui indique les relations de dépendance morphosyntaxique), la vitesse d'élocution, les modalités interrogatives ou assertives des énoncés, tous dépendent de la prosodie. Prieto (1975) propose une conception originale des liens entre les traits contrastifs et les traits oppositionnels. Il conçoit les «traits contrastifs» comme des dimensions sur lesquelles se

manifestent les oppositions phonologiques : syllabe pré-tonique, syllabe centrale-tonique, syllabe post-tonique. Dans son exemple de l'opposition phonologique entre « *míro* » et « *miró* » (en espagnol) il montre que les deux mots ne s'opposent pas par rapport à la place de l'accent tonique (1^{ère} syllabe vs. 2^{ème} syllabe) mais par rapport à la dimension sur laquelle les traits phonologiques se manifestent. Pour Prieto une opposition ne peut s'établir qu'entre deux choses qui apparaissent sur la même dimension. Aussi « *míro* » et « *miró* » s'opposent du fait que la syllabe centrale comporte des phonèmes différents.

Syllabe pré-tonique	Syllabe centrale tonique	Syllabe post tonique
0	mí	ro
mi	ró	0

Le mot « *míro* » commence avec la syllabe centrale alors que le mot *miró* commence avec une syllabe pré-tonique : 0 | mí | ro | vs. mi | ró | 0 | (Prieto 1982). Aussi on en conclut que ce sont les traits contrastifs qui sous-tendent la manifestation des traits oppositionnels.

Dans la didactique des langues, la mise en valeur de la prosodie a été particulièrement développée et féconde déjà dans les années 60. Elle a donné naissance à la méthode dite Structuro-Globale-Audio-Visuelle (SGAV). Paul Rivenc (Rivenc et al. 1962, 1964) et Petar Guberina (1984) en étaient les premiers concepteurs.

Voici un exemple de cette approche réalisée dans mon adaptation de la méthode audiovisuelle « Pourquoi pas ! » (Sagot, 1990). Le point de départ est le dialogue, entendu par le haut-parleur et accompagné d'images (en PowerPoint) illustrant le sens des phrases. Chaque image comporte le tracé de la variation mélodique (courbe intonative) de chaque énoncé dont le rythme et l'intonation sont extraits et rendus audibles sous forme de chantonnement ou d'un instrument de musique.



Dans un premier temps l'élève imite le modèle rythmique et intonatif en fredonnant. Il accompagne sa production vocale du geste qui retrace la variation mélodique. Après avoir pratiqué la structure suprasegmentale en abstraction quasi-totale du sens, l'apprenant entend simultanément (en stéréo) l'énoncé parlé et la courbe mélodique chantonnée. On lui présente en suite la forme naturelle de l'énoncé parlé. C'est à ce moment qu'on lui demande de répéter l'énoncé. Ainsi avant de connaître le sens, l'apprenant maîtrise les marques prosodiques des liens de dépendance morphosyntaxiques, la réalisation des syllabes proéminentes (durée, variation mélodique, intensité) et des contours mélodiques (continuation, assertion, interrogation etc.). Aucune segmentation en mots n'est présentée à ce stade. Ce n'est

qu'après avoir reproduit correctement les énoncés du dialogue (accompagnés par des gestes et des mimiques faciales) que la forme écrite des mêmes énoncés est présentée.

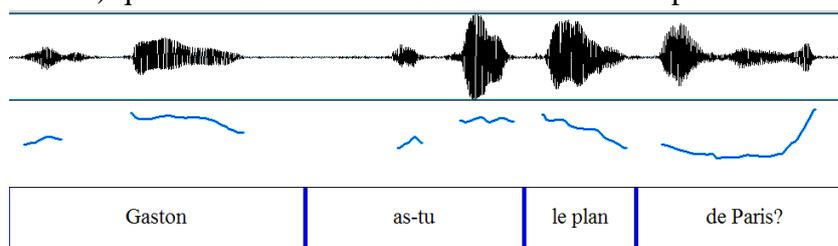


Fig.1 L'énoncé « Gaston as-tu le plan de Paris ? »

Cette didactique - basée sur le rythme, les variations mélodiques et accentuelles - porte l'empreinte du corps en mouvement (l'état de ses organes respiratoire, phonatoire et articulatoire) et de son état émotionnel. Aussi, l'énoncé acquiert une signification émotionnelle qui à son tour favorise la mémorisation (Bower, 1981, 1983, 1992). L'expression des émotions est particulièrement valorisée dans cette méthode, car souvent c'est l'émotion qui fournit le cadre interprétatif à l'énoncé et affine son sens.

Dans les méthodes du type SGAV, la perception et l'expression émotionnelles précèdent l'accès au sens lexical du message. Pour expliquer le mécanisme d'induction émotionnelle par voie vocale, une piste -souvent évoquée- est celle de *contagion émotionnelle*. Cette dernière est modélisé de façon suivante : l'émotion est un état de synchronisation des actions de 5 sous-systèmes de l'organisme : cognition, physiologie, expression vocale et non verbale, comportement, sentiment subjectif (Scherer, 2001, 2003). Il suffit alors de stimuler une de ces composantes pour déclencher l'ensemble des réactions et induire le sentiment subjectif d'une émotion. Ainsi en stimulant la composante vocale d'une émotion, on peut induire (à un degré plus faible) l'état émotionnel lui correspondant. Il est intéressant de noter de fortes ressemblances entre les moyens prosodiques inducteurs d'émotions et ceux utilisés dans la musique : des crescendos (en hauteur et/ou en volume) des tenues prolongées (de hauteur et/ou de volume) et autres.

Quand la prosodie précède le sens, l'apprenant perçoit les frontières des groupes syntaxiques du texte, la mise en emphase de certains mots et leurs valeurs émotionnelles – facteurs essentiels d'apprentissage d'une langue.

Références

- Bordas, E. (2003) « Le rythme de la prose Revue de sémio-linguistique des textes et des discours, 16». <http://semen.revues.org/2660>
- Bower, G. 1981. Mood and Memory. American psychologist.
- Bower, G. 1983. Affect and Cognition. Philosophical Transactions of the Royal Society of London B(302):387–402.
- Bower, G. 1992. How might emotions affect learning

- Condon W.S. (1984). Une analyse de l'organisation comportementale. In : La communication non-verbale, Eds : J. Cosnier et A. Brossard, Delachaux et Niestlé, Paris et Neuchâtel.
- Dessons, G. & Meschonnic, H. [1998] : *Traité du rythme. Des vers et des proses*, Paris, Dunod.
- Gallese & Lakoff, (2005). The brain's concepts: The role of the sensory-motor system in conceptual knowledge, *Cognitive neuropsychology*, 22 (3/4), 455–479.
- Grafton, S. T., Fadiga, L., Arbib, M. A., & Rizzolatti, G. (1997). Premotor cortex activation during observation and naming of familiar tools. *Neuroimage*, 6, 231–236.
- Guberina Petar (1984) Bases théoriques de la méthode audio-visuelle structuro- globale (méthode Saint Cloud-Zagreb), Une linguistique de la parole, in Aspects d'une politique de diffusion du français langue étrangère depuis 1945; Matériaux pour une histoire, Hatier, Paris.
- Jaffe J, Beebe B, Feldstein S, Crown CL, Jasnow MD. (2001) Rhythms of dialogue in infancy: coordinated timing in development. *Monogr Soc Res Child Dev*. 2001;66(2):i-viii, 1-132.
- Kendon, A. (1972) Some relationships between body motion and speech: An analysis of an example. In A. Siegman & B. Pope (Eds.), *Studies in dyadic communication*. New York: Pergamon Press.
- Kendon, A. (1980). "Gesticulation and speech: Two aspects of the process of utterance", In M. R. Key [Ed.], *The Relationship of Verbal and Nonverbal Communication*, pp. 207-227, Mouton, The Hague, the Netherlands.
- Levitin, D. & Tirovolas, A., (2009) Current advances in the cognitive neuroscience of music. *The Year in Cognitive Neuroscience 2009, N.Y. Academy of Sciences*, 1156, 211.–231.
- Lieberman, A.M., Cooper, F.S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967) Perception of the speech code, *Psychological Review*, 74, 431-461.
- Lieberman A.M. et Mattingly 1985, *Cognition*, 21, pp 1-36
- Lieberman, A. M., & Wahlen, D. H. (2000). On the relation of speech to language. *Trends in Cognitive Neuroscience*, 4, 187–196.
- Mersenne, M. (1636) *L'harmonie universelle*, Éditions du CNRS, Paris, 1986
- Martin, Ph. (2009) "Intonation du français", Armand Colin, Paris.
- Meschonnic, H.(1982) : *Critique du rythme. Anthropologie historique du langage*, Lagrasse, Verdier.
- Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997). Object representation in the ventral premotor cortex (area F5) of the monkey. *Journal of Neurophysiology*, 78, 2226–2230.
- Platon*, Introduction aux Lois - Livre II Nomoi , Traduction de Chambry. Oeuvre numérisée par J. P. Murcia : <http://remacle.org/bloodwolf/philosophes/platon/lois2fr.htm>
- Scherer, K. R. (2001). Appraisal considered as a process of multi-level sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.). *Appraisal processes in emotion: Theory, Methods, Research* (pp. 92-120). New York and Oxford: Oxford University Press.
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40, 227-256.

- Schöner G. (2002) Timing, clocks and dynamical systems, *Brain and Cognition*, 48, 31-51.
- Titchener, Edward B. 1909. *Lectures on the Experimental Psychology of the Thought Processes*. New York: The Macmillan Co.
- Trainor, L.J. & Schmidt, L.A. (2003) Processing emotions induced by music. In Peretz and Zatorre (Eds) *The Cognitive Neuroscience of Music*. 310-24. Oxford University Press.
- Peretz, I. & Zatorre, R. (2005) Brain organization for music processing. *Annual Review of Psychology*, 56: 89.-114.
- Piaget, J. 1961. *Les mécanismes perceptifs*, PUF.
- Piaget, J. (1975), *La naissance de l'intelligence chez l'enfant*, p 83.
- Prieto, L. (1975) *Études de linguistique et de sémiologie générales*. Droz, Genève.
- Rivenc P., Gauvenet, H., Guberina, P., Neveu P. & Vettier, R. (1962): *Le français par la méthode audiovisuelle (méthode d'enseignement du français langue étrangère)*, Paris, Didier.
- Rivenc, P. & Guberina, P. (1962 et 1964). *Voix et images de France (1r et 2e degrés)*. Cours et méthode audiovisuelle d'enseignement du français aux étrangers, réalisée par le CREDIF, Didier.
- Sagot, H. (1990). *Pourquoi Pas ! Pédagogia*, Saint-Herblain.
- Schlaug, G. (2009). Listening to and making music facilitates brain recovery processes. *Annals of the New York Academy of Sciences* 116