

Vers une phonologie incarnée et distribuée, motrice, mémétique et énaïve

Didier Bottineau¹

Résumé

Il se développe depuis quelques décennies des théories linguistiques qui repensent la relation signifiant / signifié au prisme de la cognition incarnée : le signifiant lexical et grammatical est gestuel, et le signifié l'est également, en relation avec la « perçaction » de Berthoz (implication du rôle moteur du corps dans la production d'impressions du monde). Cette redéfinition du rapport signifiant / signifié, caractéristique de diverses approches (submorphémiques, linguistique énaïve), inaugure le champ d'une phonosémantique doublement incarnée, et énaïve, capable de parler de « l'inscription corporelle de l'esprit ». Toutefois, ces approches ne semblent pas disposer, pour l'heure, de modèles de représentations phonologiques métalinguistiques et cognitifs adéquats pour servir leur propos : la phonologie vise à produire des représentations abstraites des sons des langues et hésite à intégrer la dimension motrice ; la phonétique se concentre sur les sons de la parole dans ses rapports à la production et à la perception sans viser à caractériser la corporalité structurale de langue à l'origine des dynamiques phonosémantiques. Dans la présente étude, on porte un regard sur le débat autour de l'incarnation en phonologie et en phonétique, et on esquisse une première série de propositions en vue du développement d'une phonologie incarnée, utile à la modélisation technique des questions de phonosémantique dans un cadre énaïve.

Mots-clés : phonologie incarnée ; phonosémantique ; énaïve

Abstract

Some linguistic theories have been developing for several decades that revisit the signifier/signified relationship through the prism of embodied cognition: the lexical and grammatical signifier is gestural, and so is the signified meaning, in relation to Berthoz's "perçaction" (perception in action : the role of embodied motricity in the production of perceived objects as networks of affordances). This redefinition of the signifier/signified relationship, characteristic of various approaches (submorphemics, enactive linguistics), inaugurates the field of embodied and enactive phonosemantics. However, these approaches do not seem to have, for the time being, adequate metalinguistic and cognitive phonological representation models to serve their purpose: phonology aims to produce abstract representations of the sounds of languages and hesitates to integrate the motor dimension; phonetics focuses on speech sounds in their relationship to production and perception without aiming to characterize the structural corporeality of language supporting phonosemantic dynamics. The present study reviews some elements of the debate around embodiment in phonology and phonetics, and outlines a first series of proposals for the development of an embodied phonology, useful for the technical modeling of questions of phonosemantics in an active setting.

Keywords : embodied phonology ; phonosemantics ; enaction

¹ CNRS, UMR 5191 ICAR, ENS de Lyon. E-mail : didier.bottineau@ens-lyon.fr

Introduction

Il y a plus de vingt ans, aux Journées d'Etudes Linguistiques consacrées au thème des Universaux Sonores (JEL-3, Nantes, 21-23 mars 2002), Georges Bohas présentait la théorie des matrices et étymons pour les langues sémitiques, reprenant des données et analyses publiées par lui dans Bohas (2000 : 90-117). Dans l'un de ses exemples, il établissait un lien entre la courbure de la langue et la courbure de gestes ou d'objets évoqués par les étymons tirés de la matrice concernée ; et au moment des discussions, l'auteur de ces lignes lui posait la question suivante (restituée de mémoire) : l'hypothèse d'un lien entre la forme de l'articulation de la langue et la forme de l'objet est intéressante, mais elle relève d'une analogie opérée de l'extérieur, par le regard du linguiste ; comment les locuteurs pourraient-ils établir un tel lien ? Comment une similarité constatée du dehors pourrait-elle permettre d'inférer une solidarité attribuée à la cognition du sujet « en dedans » si elle demeure inobservable pour son performateur ? Aujourd'hui, on commence à connaître la réponse : l'analogie constatée par l'observation symptomatique repose sur des principes qui ne supposent pas un locuteur observateur et interprète conscient de sa propre activité. L'articulation phonatoire est gestuelle et motrice ; la perception et conceptualisation des objets l'est également selon les principes de la cognition incarnée ; les deux sont indissociablement liés. La forme et le sens ne sont pas des domaines abstraits qui travailleraient séparément et ne se relieraient que par des relations aléatoires dénuées de contenu ; au contraire, l'un et l'autre forment des domaines concrets animés et solidarisés par des principes analogues de motricité incarnée. Les locuteurs n'ont pas à inventer ni observer des relations et processus qui organisent d'emblée et conjointement leur activité biomécanique et conceptuelle : ils n'ont qu'à faire sens au son du corps et, par l'acte même de parler, faire jouer « l'inscription corporelle de l'esprit » (Varela et al., 1993). L'hypothèse d'une solidarité relative de la forme et du sens est fondée en droit dès lors qu'on la contextualise dans un paradigme cognitif incarné et émergentiste comme celui de l'énaction. En conséquence, pour interroger la manière dont le son linguistique en acte fait sens conceptuel en acte, il faut se munir d'une phonologie adaptée : une phonologie motrice (fondée sur le geste), incarnée (ancrée dans la biomécanique corporelle vivante générale), mémétique (intéressée par la distribution des motricités dans l'intersubjectivité), et énative (visant à étudier l'émergence du sens de l'expérience des formes en acte).

La présente étude se veut donc un jalon en vue de la résolution d'un problème qui se pose depuis longtemps dans l'histoire de la pensée linguistique. La question est notamment posée par Jakobson :

Bien entendu, on nous enseigne depuis longtemps que, comme tout signe verbal, le mot est une unité à deux faces. [...] Tout mot, et tout signe verbal en général, présente l'union du son et du sens, ou en d'autres termes l'union du signifiant et du signifié [...] Or, si le fait de cette union est absolument clair, sa structure reste fort peu connue. Une suite des sons se trouve être le véhicule du sens, mais comment les sons remplissent-ils cette fonction de véhicule ? Quels sont exactement les rapports entre les sons et le sens à l'intérieur du mot et de la langue en général ? (Jakobson, 1976 : 22-23).

Selon Le Tallec, la réponse est apportée par le cognème (Bottineau, 2002), phonème dont le schème phonatoire contribue à déterminer le profil abstrait de l'invariant d'opérateurs grammaticaux en lesquels il s'inscrit :

Le cognème est en quelque sorte le chaînon manquant entre le phonatoire et le sémantique, cette union du son et du sens que recherchait Jakobson. (Le Tallec-Lloret, 2012, reproduit dans 2017 : 164)

Dans le domaine lexical, on peut en dire autant de la TME de Bohas, et de diverses approches explicitement intéressées par l'incarnation des rapports signifiant / signifié. Ainsi, il existe des

théories de la forme du lexique et de la grammaire qui remettent en question « l'arbitraire du signe » en repérant des faits de cohérence forme / sens. Certaines d'entre elles adoptent, implicitement ou explicitement, la posture de l'incarnation radicale en pensant le signe linguistique comme activité sensorimotrice phonatoire et le sens linguistique comme ensemble de procédures elles-mêmes incarnées.

Ce positionnement implique une double rupture phénoménologique : il introduit la question du corps des deux côtés de la signifiance, inaugurant la possibilité de reposer la question des liens intégratifs en termes monistes de solidarités de motricités plutôt qu'en termes dualistes de similarités signifiant / signifié dans leur définition classique. La question, c'est de savoir comment les formes signifiantes sont capables, par leur motricité même, d'activer des schèmes moteurs autres que phonatoires et jouant un rôle constituant dans les représentations sémantiques, qu'il s'agisse des signifiés lexicaux munis de modèles gestuels implicites ou des signifiés grammaticaux en tant que procédures abstraites d'orchestration des signifiés lexicaux. Enfin, le problème : pour modéliser correctement ces rapports de solidarité entre forme et sens incarnés, il importe de disposer de représentations métalinguistiques des unités phoniques effectivement conçues en fonction des principes qu'il s'agit d'explorer, à savoir l'énergie langagière, que Bohas nomme la motivation intrinsèque (Bohas & Sager, 2012 : 38 ; Leeman, 2021). Or, jusqu'à présent, les travaux se sont essentiellement appuyés sur une phonologie structurale abstraite et par traits, insuffisamment pourvue en modélisation de motricité articulatoire et de multimodalité sensorielle.

Une précision terminologique s'impose. Actuellement, on appelle phonologie l'étude des sons des langues comme unités et représentations abstraites, les phonèmes, par opposition aux signaux acoustiques avérés et étudiés comme phénomènes empiriques, les phones. En toute rigueur, il faut réserver la phonologie à l'ensemble des phénomènes sonores de la langue et la diviser en domaines selon le degré de segmentabilité des représentations unités : la prosodie pour les unités non segmentables en seconde articulation et la phonématique pour les segments de seconde articulation, les phonèmes. On s'intéresse donc ici à la partie de la phonologie qu'il conviendrait de toujours désigner comme *phonématique* (étude des phonèmes d'une langue) en réservant l'adjectif *phonémique* au sens « relatif au phonème ». Dans cette étude, on n'emploie jamais *phonématique* au sens de « relatif au phonème » comme le fait l'usage courant (rapporté dans tous les dictionnaires consultés).

On se fixe donc ici pour objectif d'effectuer un petit pas en avant dans le développement d'un projet de phonologie éactive, fondée sur une lecture émergentiste des corrélations forme / sens conçues comme incarnées. Pour ce faire, on présente plus précisément la conception du langage dont il est question (section 1), on resitue la question de l'incarnation en phonétique et en phonologie (section 2), on repositionne la phonématique de la langue dans un cadre incarné élargi (section 3), puis on se concentre sur le système français avec les voyelles (section 4) et les consonnes (section 5). Le projet peut sembler ambitieux et suscitera sans doute plus de questions que d'éclaircissements conclusifs ; il ne vise pas à plus que cela et ne cherche qu'à poser qu'un jalon utile que l'on souhaite fertile.

1. Le langage humain entre *perçaction* et *sémaphonie*

Ces dernières décennies ont vu le développement de théories linguistiques dans le domaine de la submorphémie : l'analyse de formes lexicales et grammaticales qui reconnaissent des segments sémantiquement pertinents dans l'évocation du signifié d'ensemble. Pour le lexique, ceci concerne notamment des segments consonantiques nommés phonesthèmes (Firth, 1930), étymons (Guiraud, 1967), éléments idéophoniques (Tournier, 1985), marqueurs sublexicaux

(Philps, 2002), réputés exprimer une impression sensorielle ou motrice pertinente pour l'évocation de la notion lexicale, et aussi classer des ensembles de mots non liés ou non par source étymologique commune en diachronie. Un exemple bien étudié est *gl-*, lié à la luminosité en anglais et en allemand (*glint, glitter, glisten ; glühen, Glück*) (Sadowski, 2001), y compris pour *gl-* obtenu par réassemblage (*glauben* « croire », de *ge-lauben* « se faire aimer qqch à soi-même », « apprendre à apprécier qqch). Pour la grammaire, l'existence d'éléments formateurs est également identifiée depuis longtemps, comme le contraste *wh/th* en anglais chez Bloomfield (1933), auxquels la cognématique (Bottineau, 2000) attribue un invariant procédural abstrait correspondant à l'opposition amémoriel / mémoriel issu de la psychomécanique du langage de Gustave Guillaume (on retrouve une analyse comparable chez Lapaire et Rotgé (1993) avec l'opposition rhématique / thématique telle que formulée dans le cadre de la grammaire méta-opérationnelle d'Adamczewski & Delmas, 1982).

Parmi les approches submorphémiques de marqueurs lexicaux et grammaticaux, certaines établissent des correspondances explicites entre la motricité de l'articulation et la motricité de la conceptualisation.

Dans le domaine grammatical, Toussaint (1981) identifie dans les langues romanes un couple K/T (espagnol *cuanto / tanto*) et relie le couple amémoriel / mémoriel aux propriétés articulatoires des phonèmes : /k/, plosive vélaire, « antérieure » ou « précoce » si on considère son point d'articulation du point de vue de l'expérience du locuteur qui la réalise, serait apte à exprimer une saisie temporalisée elle-même antérieure, amémorielle ; /t/, plosive coronale, à point d'articulation postérieur ou tardif de ce même point de vue, serait apte à opérer une saisie temporelle ultérieure, mémorielle. Toussaint a d'abord considéré que l'articulation incarnée mimait l'articulation conceptuelle (1981) avant de se demander si ce n'était pas l'inverse (2002) : le contraste amémoriel / mémoriel émergerait de l'expérience de l'articulation de /k/ et t ; l'esprit humain « penserait » discursivement dans les termes de procédures d'élaboration du sens empruntés à la recapture sémiotisatrice des effets perceptuels de l'articulation vocale organisée, érigée en machine à se faire penser. Une telle approche revient à développer une *sémaphonie* incarnée : la parole dans une langue en tant qu'art vocal dont les formes articulées vectorisent et profilent la synthèse du sens ; une théorie des rapports forme / sens fondée sur la mise en correspondance, ou l'étude des solidarités, liant deux niveaux de motricité, celui de la phonation articulée et celui de la conceptualisation dans sa dimension procédurale et schémique. Parallèlement, et sans rapport direct a départ entre les deux, Bottineau (2002 et suivants) a défini la cognématique comme l'étude de la morphologie grammaticale fondée sur les rapports entre articulations phonologiques et invariants procéduraux dérivés, par exemple I/A et le couplage des degrés d'aperture vocalique avec leurs invariants respectifs (I : conjonction, A : disjonction, comme *di* et *da* « de » en italien ou *y* et *à* en français) (Bottineau, 2010).

Dans le domaine lexical, on trouve chez ces auteurs et de nombreux autres des analyses du « sens » des étymons lexicaux comme *st-* directement fondées sur la motricité phonatoire : par exemple, *st-* devrait son invariant (arrêt, immobilisation, fixation : *still, stay, stop* etc.) à la forme même de son articulation (apico-coronale d'abord fricative, continue et sifflante, puis occlusive et plosive, avec un « stop »). La théorie sémio-génétique de Philps (2002) étudie la manière dont ce couplage de motricités phonatoires et conceptuelles varie en diachronie depuis l'indo-européen et modélise des parcours morphosémantiques qui scénarisent la co-variation de la forme des étymons et de leur relation aux notions sémantiques des unités lexicales en lesquels ils s'inscrivent.

La forme des submorphèmes est une variable typologique liés à la phonologie de la syllabe, des phonèmes et des traits (Bottineau, 2014). Dans le domaine des langues sémitiques, la théorie

des matrices et étymons (TME) de Bohas (2016) étudie sous un angle panchronique la structuration des verbes arabes autour de « matrices » (couples réversibles de traits phonémiques) générateurs d'étymons au sens de Guiraud (couples de phonèmes, eux-mêmes réversibles) intervenant dans la production des radicaux verbaux. Bottineau (2021) interprète cette théorie comme une sorte de morphologie générative par laquelle le locuteur générerait des unités lexicales à partir d'une synthèse submorphémique en trois étapes : la matrice comme appariement de traits, l'étymon comme réalisation des couples de traits par des paires de phonèmes, le radical comme insertion des étymons dans des formes verbales lexicales constituées et énonçables. Cette génération de formes en trois étapes rappelle l'esprit de la chronogenèse de Guillaume (1929), qui voit dans les modes verbaux trois niveaux de formation de la représentation de l'image-temps : chez Bohas, les niveaux sont appliqués à la formation du signifiant lui-même ; la théorie de Bohas rappelle également la phonologie autosegmentale de Goldsmith (1990), qui propose que les généralisations phonologiques (phonèmes, tonèmes) résultent de la redistribution de traits autosegmentaux issus de « tiers » séparés, les traits étant eux-mêmes des représentations abstraites adossées à des organes (mais pas définies comme gestes). Dans ce cadre, Bohas introduit la notion de *motivation intrinsèque* : le sens lexical des radicaux verbaux comprend des éléments de sens à caractère impressif et moteur directement inspirés ou motivés au sens étymologique (mis en mouvement, animés) par les propriétés motrices des gestes articulatoires correspondant aux traits des matrices et leur orchestration phonémique. Dans le domaine de la linguistique italienne et des langues romanes, Rocchetti (1980) nomme « mouvements premiers » des simulacres de mouvements imprimés dans la structure articulatoire des signifiants en tant qu'actions motrices phonatoires évocatrices d'actions non phonatoires (Rocchetti, 2019 ; Saffi & Culoma, 2021). On ne parle donc pas de phonomimétisme (imitation phonatoire et articulatoire d'impressions sensorimotrices laissées par des objets et événements externes et mondains, préexistant indépendamment de l'activité langagière, par motivation intrinsèque) mais de *sémaphonie* (évocation directe de notions en jouant par le signifiant des propriétés de constitution procédurale, par motivation intrinsèque) : le corps parlant agit et s'agit comme un *sémaphone* (générateur de sens conçu par l'activité articulatoire) et procède par *sémaphonie*, terme volontairement proche de la *sémiophonie* de Derveaux (2003) dans le domaine de la poésie (Bottineau & Poirier, 2022) ; le *corps sémaphone* de la « parole orale-aurale » (Stewart 1995) des langues dites vocales contraste également avec le *corps sémaphore* de la parole gestuelle-visuelle des langues signées, qui font sens en mobilisant le corps complet dans l'évocation pantomimique des événements comme interactions entre personnes, objets et environnements. Le langage humain est une activité de *sémaphonie*, « une gestuelle articulatoire perfectionnée » (Rocchetti, 1991) ; les langues naturelles constituent des disciplines *sémaphoniques*, des domaines d'activité auto-régulée et structurée par un système évolutif² de motivations intrinsèques qui sous-tend l'organisation des solidarités liant phonation et expression en tout moment synchronique en lequel on capture comme par un instantané photographique un moment de la dynamique du parcours de l'ensemble.

Cette approche matérialise la notion aristotélicienne d'*energeia* (activité) avec sa capacité à fédérer, profiler et identifier les communautés humaines à travers leur activité de communication en tant que co-production du sens, avec son devenir et ses étapes historiques (la *sématogenèse* dans la *Science Nouvelle* de Vico (1744), cf. Trabant, 2004 : 34-36) et sa

² Dans ce type d'emploi, le mot *système* désigne la dynamique qui engendre les *structures* : à proprement parler, une langue est une structure de signes dont le devenir (le changement linguistique) est animée par un système, dont il existe diverses conceptions (la systémique guillaumienne, la théorie varélienne de l'autopoïèse, les systèmes sociaux de Luhmann, la répliation etc.). En italien, il existe un verbe *sistemare* (arranger, disposer, accommoder) qui signifie littéralement « animer les structures existantes en leur appliquant mes propres lois » : *Io sistemo* « c'est moi qui gère » (propos exaspéré d'une mère de famille constatant le désordre laissé dans la maison par ses fils).

vocation à structurer la forme interne des langues et les parcours discursifs (les *Sprachtypus* et *Redetypus* de Humboldt 1836). Elle se retrouve dans la phénoménologie du langage de Merleau-Ponty, qui conjoint l'incarnation de la parole et l'interactivité normative des rapports sociaux. Et surtout, elle réalise le programme de « l'inscription corporelle de l'esprit » (radical embodiment) propre au paradigme éenactif (Varela et al., 1991) : selon ce paradigme, le monde ambiant, assumé comme physique, objectif, matériel et prédonné (existant en l'état indépendamment du point de vue d'un observateur), est en réalité une construction relative au corps de l'observateur vivant (animal ou humain) qui en produit une observation en tant que réalisation au sens anglais (à la fois matérialisation et prise de conscience). Les objets visuels sont des évènements psychologiques et des actes de conscience élaborés en s'appuyant sur les réseaux de motricités corporelles habituellement mobilisés à l'occasion des productions antérieures et enregistrées d'objets analogues : la visualisation d'un ballon relie l'oculomotricité qui en capture la forme aux autres motricités (manuelles et bipédiques) qui ont participé aux visualisations antérieures de ballons ; le ballon visuel, en tant que construction synthétique, intègre le paradigme des modèles de schèmes moteurs humains utiles à la compréhension de l'objet et à l'engrammation de ses manipulations possibles : l'humain connaît le ballon visuel humain muni de propriétés anthropiques ; tout autre animal « voit » autre chose (le ballon félin, le ballon canin, munis de propriétés propres aux *enactions* de ces espèces), à quoi nous humains n'avons aucun accès (si ce n'est par le biais d'un exercice de pensée et de simulation imaginative). La perception relève d'une imagerie de synthèse biologique, elle procède par *apparitionnage* (production inconsciente d'apparition perceptuelle, de phénomène vécu comme objet)³ : elle élabore une impression figurale et multimodale du monde en s'appuyant sur une coordination de motricités sensorielles et autres, produisant à la conscience des objets construits en fonction du corps de l'individu selon son espèce d'appartenance, à la fois biologique et culturelle. Berthoz dans Berthoz & Andrieu (2011 : 10-12) renomme *perçaction* la perception en acte (*perception in action* en anglais), en tant que coordination motrice capable d'inscrire des représentations motrices cachées dans les phénomènes apparaissant à la conscience ; la perçaction reformule ainsi les « affordances » de Gibson (1976) dans le cadre de la cognition incarnée (Noe, 2004 parle de *action in perception*). Tous ces termes peuvent sembler abscons, mais il est difficile de les éviter : l'ensemble de ces travaux parle de ce que Frith (2009) nomme *the making of the mind* « la fabrique de l'esprit » (ou de la conscience), à savoir, l'ensemble des processus conscients qui concourent à produire des effets de conscience, d'expérience et de réalité dont les contenus sont radicalement différents des processus qui les engendrent, ce qui rend ces processus occultes, inaccessibles à l'introspection et à l'intuition immédiate, et rend nécessaire leur simulation par une terminologie adaptée, néologique et à caractère expérimental, donc elle-même difficile à manier.

Ainsi définie, la perçaction concerne directement le linguiste. D'un côté, le monde perçu est un apparitionnage de phénomènes élaborés par un corps vivant et agissant, qui charge les « objets »

³ Dans le roman *Voiles* (Cixous et Derrida, Galilée, 1998), Hélène Cixous introduit les néologismes *apparitionner* et *apparitionnement*, qui rendent compte d'expériences perceptuelles, psychologiques et poétiques non ordinaires. Le verbe désigne le processus conscient et vécu de formation du phénomène apparent, et le nom thématise l'expérience de ce processus dans la durée, qui survient notamment dans la perception de formes problématiques, ambiguës, ou relevant du mirage. On trouve également le mot *apparaissance* pour l'expérience dans la durée des variations de l'apparitionnement (terminologie photographique : l'apparaissance d'un visage sous divers angles ; le terme est utilisé en commentaire littéraire) : « Le lupeux est un démon dont la nature n'a jamais été bien définie et dont l'apparaissance varie suivant les localités. » (G. Sand, à propos du loup-garou vendômois). Pour notre part, nous réservons *apparitionnage* au processus ordinaire de production du phénomène par le corps vivant non réflexif, qui ne se vit pas comme peintre et figurateur du monde apparent. On pourrait préférer le terme plus simple *visualisation*, mais celui-ci désigne l'acte *intentionnel* de rendre visible ce qui ne l'est pas : mon corps visualise un monde, alors que moi je me vis comme voyant le monde du fait que mon corps réalise de l'apparitionnage qui ne se laisse pas vivre comme visualisation consciente d'elle-même.

de représentations motrices incarnées et de schémas d'action ancrés dans les pratiques culturelles. De l'autre, la parole humaine est elle-même une discipline motrice (phonoarticulatoire) capable d'élaborer des idées (le sens linguistique) associant des concepts (les signifiés lexicaux) eux-mêmes capables de référer à des référents (objets du monde construits comme actes de conscience par la perçaction avec sa motricité). De ce fait, forme et sens linguistiques sont tous deux régis par les mêmes principes de cognition incarnés que ceux qui sous-tendent la perçaction et l'enaction, la production de réalité apparemment matérielle et objective par le corps humain. Pour cette raison, il est nécessaire d'explorer les rapports liant les motricités de la phonation et de la perçaction soient mises en rapport. Allott (1973), s'appuyant sur la kinésique de Birdwhistell (1952), fait l'hypothèse d'une analogie ou d'une solidarité structurale liant les gestes articulatoires de la parole (phonèmes, étymon pluriconsonantique) aux gestes de manipulation des objets matériels. Selon la *Gestural Theory of the Origin of Speech* (Corballis, 2003), le langage humain résulterait d'un transfert de la motricité manuelle intentionnelle (manipulation d'objets et d'outils) à l'articulation phonatoire et non d'une articulation directe du cri pulsionnel, avec divers arguments : 1) sur les propriétés sémiotiques de la communication : la parole vocale s'accompagne habituellement d'une gestualité coverbale illustrant le rapport aux objets, actions et repérages ; les langues signées font sens avec la seule gestualité non phonatoire ; 2) en neurophysiologie fonctionnelle : le système des neurones miroirs chez les primates est structuré de manière analogue à celui du contrôle du langage humain (Rizzolatti & Arbib, 1998) ; et il existe une corrélation liant l'asymétrie manuelle (droitier / gaucher) et l'asymétrie cérébrale dans le traitement du langage ; 3) en apprentissage : enseigner une communication gestuelle à des primates expérimentalement est plus accessible qu'une communication articulée ; et pour l'enfant humain, le pointage manuel joue un rôle déterminant dans la mise en parole. Les théories vocale et gesturale des origines du langage sont actuellement mises en synergie par les théories multimodales (Gillespie-Lynch et al., 2014) et leur contextualisation par les théories multicausales (Prieur et al. 2020). La théorie sémio-génétique de Philips (2006) modélise l'historique du couplage des gestualités articulatoires et autres à travers l'analyse diachroniques des rapports forme / sens dans les transformations du lexique depuis l'indo-européen. On peut donc étudier le problème selon deux angles d'attaque : au plan évolutionnaire, on se demande comment la gestualité articulatoire en vient à se distinguer de la manuelle tout en lui restant relativement corrélée, et délimiter le périmètre de son domaine spécialisé de production sémantique, la sémaphonie ; et en sciences du langage, en reprenant le problème par le bout linguistique en diachronie et en synchronie, on interroge les corrélations formes / sens en tant que solidarités motrices liant les domaines hétérogènes de l'action verbale et l'action en général. Dans ce paradigme, on s'interroge sur le sens des représentations phonologiques articulatoires en rapport avec une conception du réel et du sens eux-mêmes kinésiques, inaugurant le domaine de la linguistique énaïve (Bottineau, 2010, 2011, 2012, 2013, 2018 ; Bottineau et al., 2017 ; Grégoire et al., 2017).

Un tel questionnement suppose une phonologie incarnée, dans l'esprit des propositions de Lafont (2004, 2007), et que l'on gestualise autant que nécessaire les traits invoqués dans les phonosémantiques telles que la TME, la cognématique etc. Baudouin de Courtenay (1915) définissait le phonème non pas comme une image acoustique, mais comme un *kinakème* (Jakobson & Waugh, 1979/2002 : 22 ; Adamski, 1990 : 71), mot-valise signifiant l'intégration en une unité indissociable d'un modèle d'articulation phonatoire, le *kinème*, et d'un modèle d'impression auditive associée, l'*acousmème*⁴ : une boucle sensorimotrice multimodale, qui

⁴ Saussure chasse le kinakème physiologique par la porte acoustique et le réintroduit par la fenêtre auditive : l'accent mis sur « l'impression acoustique » (Parret 2014, 112) n'est pas autre chose que l'expérience ressentie de la motricité sous-jacente à l'audition, que l'auteur refoule, confinant « l'impensé du corps ».

préfigure la cognition incarnée actuelle, en particulier les neurones miroirs de Gallese et Rizzolatti (distributeurs de schémas d'action partagés dans l'intersubjectivité). Dans une perspective interactionniste, il faudrait y ajouter la représentation de la gestualité faciale visible pour l'allocataire observant le locuteur, l'*optème*⁵, comme la protrusion labiale en français en tant que trait visuo-facial, l'ensemble formant le *kinakoptème*, une intégration motri-audio-visuelle⁶. Il faudrait également intégrer au modèle la modalité tactile avec la catégorie de l'*haptème* (Komel, 2019), unité de toucher, d'expérience intrareceptive caractérisant la production des phonèmes ; ce questionnement en est à ses débuts, et sera pour l'heure laissée de côté⁷. Le kinakème est présentée comme une entité relevant de la « psychophonétique », ce par quoi de Courtenay veut dire une phonologie psychologiquement orientée :

The terms “sounding,” “sound,” “resonance,” etc., designate the transient representations of linguistic thought and, therefore, belong to (natural) science; but insofar as they are representations of linguistic thought based on individual and individual-collective mental processes, they should be replaced by (proper linguistic) terms, such as:

the *phoneme*, the psychological equivalent of physical “sound,” the actual and reproducible phonetic unit of linguistic thought. The phoneme consists, in turn, of constituent elements of which we are not aware during linguistic intercourse but which can be obtained by analysis; they are:

the *kineme*, the articulatory, phonational element of linguistic thought;

the *acousmeme*, the simplest psychological element of audition or acoustic perception; and

the *kinakeme*, the complex representative of both the articulatory (phonational) and auditory elements.

This is the way things stand in psychologically oriented linguistic acoustics. (Baudouin de Courtenay, 1972: 279-280)

En termes actualisés, on dirait que de Courtenay envisage le phonème comme un *processus* mental « individuel-collectif » moteur et multimodal, ce qui en fait un grand précurseur de la cognition incarnée et distribuée, à rapprocher de Jousse et de Merleau-Ponty. Il faut souligner

⁵ Fisher (1968) nomme *visème* des catégories de gestes articulatoires qui se distinguent visuellement (en lecture faciale par un observateur ou auditeur). Il est rare qu'un visème caractérise un seul phonème (c'est le cas de la projection apico-post-dentale pour la ceta espagnole, qui n'existe que non voisée, à la différence de <th> anglais) : en général un visème caractérise un ensemble de phonèmes partageant le même geste visible, comme la constriction bilabiale pour /p/, /b/ et /m/ (ce qui empêche de distinguer ces phonèmes par la seule lecture faciale). Nous reformulons ici le visème (assemblage d'une racine latine et d'un suffixe grec) par le terme *optème* (racine et suffixe grec) de manière à produire une composition analogue aux termes créés par De Courtenay (*kinème*, *acousmème*, *kinakème*) et aboutir à une surcomposition cohérente sur le même modèle (*kinakoptème*).

⁶ Le visème est ordinairement mentionné dans les études consacrées à l'étude de la reconnaissance des sons langagiers dans une perspective de perception multimodale : acoustique-auditive et optique-visuelle ; on parle alors d'intégration audio-visuelle (Chen & Rao 1998, Chen 2001). Ces travaux ne référant pas à la théorie motrice de la perception des sons langagiers, ils n'intègrent pas directement le schème moteur au modèle perceptuel.

⁷ L'haptolinguistique (*haptolinguistics*) du philosophe slovène Mirt Komel est un développement récent au sein de l'école de psychanalyse de Ljubljana, dont les travaux portent notamment sur les études haptiques dans l'histoire de la philosophie et au prisme de la psychanalyse (Lacan, Anzieu, Merleau-Ponty, Derrida, Nancy, Levinas). L'haptème a été pensé d'abord comme unité de toucher hors du domaine linguistique, puis dernièrement comme composante de l'expérience du signe en acte, ce qui rejoint notre problématique, en particulier dans sa dimension phonosémantique : par exemple, quand Jünger (1934) qualifie l'invariant de /i/ par « l'étreinte », il s'appuie explicitement sur les composantes kinémiques (geste lingual) et haptique (proprioception tactile du rapport langue / palais). L'haptématique en est à ses débuts et fera l'objet d'une intégration méthodique en une étape ultérieure du développement de la phonématique incarnée.

que ce texte manipule la notion de « représentation » avec prudence : on parle de « représentations transitoires » et de processus mentaux « individuels-collectifs ». Baudouin de Courtenay semble se méfier de ce que l'on nomme aujourd'hui l'internalisme représentationnaliste radical (les représentations ne sont pas des modèles mentaux abstraits et stabilisés comme objets) et recherche une conception en acte de la représentation qui rende justice à ses deux facettes, incarnée (motrice et sensorielle multimodale) et distribuée (ancrée dans l'interaction normative). Son propos est compatible avec l'énaction (le corps parlant peut développer un savoir-faire phonatoire moteur autrement qu'en le représentant) et le principe de non-localisabilité (Cowley, 2011 : ce que le psychologue nomme représentations mentales est un dispositif processuel distribué sur un réseau d'interactivité impliquant des participants hétérogènes comme le corps vivant en acte et le milieu qu'il perturbe, se perturbant lui-même).

L'avenir confirmera sa proposition par des apports empiriques validés. Dans le domaine de l'écoute langagière, selon la *Motor Theory of Speech Perception* de Liberman et Mattingly (1985), l'audition d'un son langagier passe par l'activation du contrôle moteur de sa production phonatoire, non transmise à la zone phonatoire par inhibition : l'audition est une *intériorisation de la phonation d'autrui* par amorce d'écholalie refoulée, qui peut être perturbée expérimentalement en donnant à voir une parole faciale non conforme au signal acoustique (l'effet McGurk) ; même le discours intérieur réflexif (endophasie) opère ainsi, d'où la tendance à ne pas faire l'effort de confiner la somatisation quand on réfléchit sans témoin : on parle seul. Schwartz et al. (2012) complètent cette approche par un modèle du rôle de la perception auditive dans le contrôle de l'action, le PACT (*Perception-for-Action-Control Theory*) qui rééquilibre le rapport motricité phonatoire / perception auditive. Selon Tomatis (1991), l'extraction des formants acoustiques d'une voyelle atmosphérique (onde acoustique) est réalisée par un geste auriculomoteur de la membrane du tympan, qui ajuste son profil parabolique pour une collecte optimale des fréquences recherchées (fondamentale et harmoniques), à la manière d'une main qui se prépare à la capture d'une balle dont elle connaît les propriétés : la voyelle n'est pas captée, mais capturée, d'où la possibilité de solidariser les motricités auditive et phonatoire. Les activités de production du discours et de perception auditive s'intègrent mutuellement (Menard, 2010) : le son est dans le corps, et inversement.

Dans cette perspective, un phonème se réalise par un événement communicationnel par contact oral-aural (Stewart, 1995) réparti sur trois milieux : le locuteur qui le produit (kinème) et le perçoit (acousmème), le milieu qui le transmet (ondes acoustique et signal optique), et l'auditeur qui le perçoit (acousmème, optème) en reconstituant le geste producteur (kinème). Un phonème est donc un modèle de coordination phonatoire multimodal (gestuel et perceptuel) réparti sur l'espace hétérogène de l'interaction avec ses contenus hétérogènes (corps vivants des participants et milieu d'intervention), ce qui suppose des transductions dans les passages d'un domaine à l'autre (propagation des ondes acoustiques vs incarnation). En substance, un phonème est un *mème pratique* (Jouxte, 2005 ; 218 et 230-231)⁸, un événement mémétique interactionnel et distribué (le « mimisme » de Jousse 1936⁹, puis Dawkins 1976 revu par Steffensen & Cowley 2010 et Cowley 2012), un distributeur et diffuseur d'analogie : une unité de répliation éthologique interactionnelle qui coordonne les sujets par un modèle de

⁸ Sur l'ensemble des définitions du mème et des controverses autour de sa réception, cf. également Jouxte 2005, ch.7 (150-174).

⁹ Jousse (1936/1974, 511-512 et 720-750) : l'humain est un « mimeur » qui intègre la connaissance dans l'interaction en « incorporant » des « mimèmes » « modelants ». La connaissance est construite collectivement et circule par les mimèmes ; elle évolue au gré des contributions. Dans l'anthropos vivant, il n'y a pas d'un côté le corps et de l'autre côté l'âme. Il n'y a qu'un « Composé humain ». C'est ce composé humain que nous appelons l'anthropos. » Pour Jousse, la « communication » se fait par activation alternée de mimèmes incorporés bilatéralement par les participants. Un mimème est une unité de répliation comportementale signifiante ancrée dans l'interaction.

comportement générateur d'une impression sémiotique fédératrice¹⁰. La notion de répliqueur mémétique de Dawkins a été critiquée pour son application aux sciences sociales et culturelles : elle est démunie de l'ancrage empirique dont bénéficie le génome (Benitez-Bibriesca, 2001), ses implications politiques (Gray, 2008) et ses incohérences sémiotiques (Deacon (1999) et Kull (2000)). Mais en fondant le mimisme sur la cognition incarnée dans le paradigme des 4 E (Menary, 2010 : *embodied, embedded, enactive, extended* ; Gick (2015) en phonétique *embodied* et *embedded*), on l'arrime à un vecteur empirique robuste. La répliqueur mémétique prédit également l'ajustement des systèmes phonologiques de langues en contact (de communautés en interaction) dès lors qu'on les définit comme incarnés, ce qui renouvelle la question des « affinités phonologiques » (Jakobson, 1938).

La phonétique, par définition, s'intéresse à la production, alors que la phonologie tend à s'en détourner ; et on peut s'appuyer sur l'articulation pour classer les langues du monde (Bickford & Floyd, 1981). On va donc trouver des approches incarnées à divers niveaux en phonétique, mais rarement en phonologie. La phonétique articulatoire de Gick et al. (2013) étudie les conditions de production du discours par une approche multimodale de la motricité. La phonologie fonctionnelle de Boersma (1998), non mentionnée par le traité de théorie phonologique (Goldsmith et al., 2014), rapporte les phonèmes à un *modèle d'articulateur artificiel* où les articulateurs biologiques du corps phonatoire vivant sont modélisés par un tube virtuel où circule un flux régulé par des parois pouvant intervenir par positionnement et par à-coups : sans se poser comme mimesis du corps humain, le modèle peut être paramétré de manière à simuler des événements acoustiques de synthèse aux propriétés établies par calcul munies de valeurs conformes à celles constatées par le traitement du signal acoustique réel de la phonation humaine. Ce modèle intègre un début de simulation de la biomécanique phonatoire, mais il ne peut pas gérer la complexité correspondant à la reconnaissance auditive intersubjective des gestes phonatoires, d'où son inaptitude à simuler la voix avec sa polyphonie. Dans la phonologie articulatoire de Browman et Goldstein (1986, 1995ab, 2000 ; Goldstein & Fowler, 2003), on n'a plus affaire à un simulateur, mais à un modèle de la production vivante dont la formalisation remplace les traits par des gestes¹¹. Cette approche renégocie substantiellement la relation phonologie / phonétique en intégrant au modèle phonologique la biomécanique phonétique ; pour ce faire, elle s'appuie sur le modèle de production discursive de Saltzman (1986, 1995 ; Saltzman et Kelso, 1987), la dynamique des tâches (*task dynamic*), ce qui lui permet d'interpréter la phonologie autosegmentale de Goldsmith (1990) en termes incarnés malgré son abstraction initiale (Browman et Goldstein, 1990)¹².

Dans un esprit comparable, on trouve la *Phonology as Human Behavior* de Diver (1979), développée ultérieurement par les travaux linguistiques et cliniques de Tobin (1997), et absente du *Handbook*. Il est vrai que ces entreprises demeurent inabouties : malgré l'appellation

¹⁰ On trouve une métaphore remarquable de ce mimisme émergentiste dans l'acte IV du *Prometheus Unbound* de Shelley (1820) : la libération de Prométhée et son retour chez les humains ravis suscite la formation d'un ballet de célébration auquel s'adjoint un nombre croissant de participants qui, dansant ensemble, se réenchâinent harmoniquement et créativement, précipitant une « nation » au sens humboldtien, une communauté participative de production d'effet, de sens et d'identité par l'action conjointe.

¹¹ Cette théorie reste largement ignorée : aucun chapitre ne lui est consacré dans *The Handbook of Phonological Theory* (Goldsmith et al. 2014), et elle n'y est mentionnée qu'une fois avec un commentaire (Ladd 2014, 369), qui reconnaît l'importance de la rupture épistémologique trait / geste, mais l'interprète dans le cadre des représentations abstraites de la phonologie conventionnelle, et critique son désintérêt pour la question des *symmetries in phonological patterning* centrales dans la tradition issue de Troubetzkoi.

¹² Cette étude figure dans le premier volume des *Papers in Laboratory Phonology* de l'Université de Cambridge, série qui en principe étudie l'inscription de la phonologie de l'acte de parole ; mais sur les six volumes parus (de 1990 à 2004), seuls les deux premiers (1990 et 1992) intègrent la question du corps, avec le petit groupe d'auteurs habituels : Ohala, Browman et Goldstein, Docherty et Ladd.

phonology, elles ne fournissent pas une description formalisée des phonèmes de nature à compléter ou remplacer la phonologie abstraite en accomplissant le programme de la corporalisation. Mais la faiblesse de leur représentation, voire l'absence de toute mention, dans les traités de référence en phonologie, en dit long sur la réticence de la communauté à gesturaliser les traits et fonder une approche d'où émergerait conceptuellement la voix d'une langue. Ce problème touche très généralement l'*embodiment* en sciences du langage, comme le souligne la linguistique éactive (Bottineau & Grégoire, 2017). Mais dans le dialogue entre phonétique et phonologie, la divergence initiale des objectifs et méthodes est telle que la question de « l'interface » est un débat sans fin. Revenons sur quelques points essentiels pour notre problématique.

2. Le problème de l'incarnation en phonétique et en phonologie

La phonétique est l'étude de la parole humaine dans une langue donnée, avec trois niveaux d'analyse : la production du signal acoustique par le corps parlant, dont la phonation au sens étroit (production de la fréquence fondamentale par les plis vocaux) et l'articulation (tous les autres événements d'occlusion et constriction, y compris par les plis vocaux et le larynx) ; l'analyse des propriétés physiques du signal acoustique en tant qu'onde de variation de la pression atmosphérique (indépendamment de la perception auditive), notamment les formants mesurés en Hertz ; et l'analyse des processus auditifs, à savoir la production d'une impression sensorielle consciente (comme la voyelle « ah ») distincte de l'analyse acoustique (on entend une voyelle, pas une matrice de formants). L'objet de la phonétique inclut donc le corps agissant par motricité biomécanique et le corps percevant par captation sensorielle et traitement perceptuel, mais la discipline ne relève pas de la cognition incarnée, du moins pas pour les approches qui n'explorent pas l'écart entre activités biologiques et émergences sémiotiques. La phonétique intègre rarement cette approche radicalement incarnée de l'audition motrice, ce qui la conduit d'ailleurs à des impasses, clairement formulée par Tabain (2019) :

We have seen that there are many place of articulation contrasts that are encoded within the consonant portion itself – be it the stop burst, the nasal murmur, the lateral “ring,” or the fricative noise. We have gained some insight into the various acoustic models that are used to understand these sounds, and yet the mystery remains that despite the very different acoustic results produced by the articulatory gestures, the cues to consonant place of articulation remain surprisingly constant across manner. (Tabain, 2019: 268)

Le « mystère » est de comprendre comment l'information du point d'articulation est récupérable de manière constante à travers les langues malgré la disparité des signaux acoustiques générés par les gestes articulatoires. Ce problème naît de la problématique même : dans ce paradigme, le signal est censé encoder par son profil acoustique le geste qui l'a produit (Ohala 1983), et rendre le geste récupérable, ce qui concentrerait dans le signal un modèle informationnel abstrait de la phonation et de sa récupération. Dans la conception incarnée et mémétique du phonème, ce problème ne se pose pas : les interlocuteurs possèdent un savoir-faire moteur de production phonémique acquis dans leur historique par des exemplaires recyclables, aussi l'échange de signaux se contente-t-il de fournir des symptômes du modèle moteur à activer pour produire l'impression auditive par la perçaction adaptée. Ceci est vrai de la perception en général : quand on entend aboyer un chien que l'on ne voit pas, l'information du chien n'est pas dans le signal, et sa reconnaissance ne repose pas sur le déchiffrement des propriétés acoustiques, mais sur la mise en correspondance analogique de l'impression auditive avec un paradigme d'impressions analogues associées à l'image et idée du chien dans des situations où l'animal s'est manifesté de manière multimodale. De même, le phonème est moteur et multimodal, mais sa réalisation ne se distribue sur l'interactivité de l'ensemble du

réseau formé par le locuteur, le signal et l'auditeur, ce qui le rend non localisable dans sa seule partie physique, le signal : on y retrouvera des indices, mais pas un modèle complet de l'articulation. De ce fait, les indices peuvent sembler capricieux en typologie des langues en ne perdant rien de leur efficacité, et dans la communication, les aléas des conditions de production, transmission et audition compromettent beaucoup moins la compréhension que ne le laisse prévoir l'approche calculatoire : l'auditeur se contente d'indices partiels et variables pour rétablir une motricité globale qui lui permet de restituer une parole qu'il n'a pas captée en totalité. Cette souplesse est simplexe (Berthoz, 2009) en ce qu'elle résout par avance le problème de l'adaptation aux caprices de la transmission : l'incarnation distribuée permet au corps de faire ce que le signal à lui seul ne porte pas intégralement.

De son côté, la phonologie est une science spéculative : elle vise à décrire les « sons » d'une langue ramenés à un faisceau minimal de propriétés abstraites (Dell et al., 1984 ; Laks & Plénat, 1993 ; Laks, 1997) et organisé en structure de différenciations entre unités oppositives et distinctives ; elle vise également à formaliser et annoter ses modèles de manière à rendre les langues et dialectes comparables et positionnables dans le cadre d'une typologie des langues naturelles. Les représentations phonologiques sont en premier lieu des représentations métalinguistiques réflexives, de niveau méta-discursif, élaborées par des linguistes avec des méthodes précises et finalisées par des objectifs de modélisation à finalité descriptive et non explicative ni processuelle, ce que Putois (2006) nomme des « mensonges-pour-chercheurs ». En second lieu, elles sont régulièrement amalgamées avec des représentations mentales processuelles, intériorisées par les sujets parlants, et donc, constitutives de leur compétence en tant que savoir-faire procédural subconscient et non réflexif, épilinguistique ; ce qui confère au modèle un caractère explicatif et non plus descriptif. Sur la différence entre représentations phonologiques et phonétiques, Kingston (2022), synthétisant Pierrehumbert (1990), fournit un résumé exemplaire de la conception cognitiviste des « représentations » appliquées en phonologie, à savoir, des règles de manipulation de symboles abstraits, localisées « dans l'esprit » (paradigme internaliste) et relevant de la connaissance (*knowing, knowledge*), alors que les représentations phonétiques relèvent de l'action (*doing*) en tant qu'évènement « externe » (*events in the world*) :

This description of possible differences between phonological and phonetic representations is amplified in Pierrehumbert's (1990) catalogue. Phonological representations are in the mind, where they constitute a part of speakers' and listeners' implicit linguistic knowledge, differences between them are qualitative, and the rules and constraints that specify phonological well-formedness are syntactic, in specifying combinations and structural arrangements of discrete, symbolic categories. Phonetic representations are events in the world, differences between them are quantitative, and the gradient relationships between their constituents are expressed in calculus. The difference is also often characterized as one of knowing versus doing, where the phonology of a language represents what its speakers and listeners know about its sounds and the patterns they enter into, and its phonetics represents the behaviors of those speakers and listeners when actually producing or perceiving speech. Both this distinction between knowledge and behavior and this equation of phonology with knowledge and phonetics with behavior are, however, challenged by a number of scholars whose proposals are discussed in the following. (Kingston: 2022, 361)

L'action du corps est donc représentée en phonétique et exclue de la phonologie, où sont représentés et classés les « sons » par propriétés abstraites désincarnées, abiotiques et amotrices. Evidemment, tout pose problème dans ce passage : réduire la représentation phonologique mentale à la connaissance de règles de manipulations de symboles abstraits ne crée pas les conditions du passage à l'acte phonatoire par le corps, à moins que ces symboles ne représentent, justement, le corps en acte, et n'apportent un savoir-faire (*knowhow*) plutôt

qu'une connaissance (*knowledge*). Corollairement, on ne comprend pas comment les représentations phonétiques pourraient être des « événements du monde » (sauf à évacuer la production et la perception en ne se centrant que sur le signal acoustique) ni comment elles pourraient représenter l'action (*doing*) sans être incarnées au double sens de (i) contenir le corps dans la représentation figurale et (ii) êtres elles-mêmes localisées « dans le corps ». Le paradigme éactif évite le problème en pensant la représentation comme un modèle d'activité réalisée par le corps vivant et en interaction avec son extériorité (dont autrui pour le langage), en précisant que le corps se construit (comme structure biomécanique) autour du modèle d'activité fourni par les interactions à la communauté (comme système social dynamique) : le contrôle moteur de l'action de parole (intracortical) s'organise dans, par et avec la pratique phono-articulatoire (extracorticale) selon des liens d'intégration solidaire qui le dispensent de former des représentations mentales abstraites, autonomes et détachées jouant un rôle de poste de pilotage. Pour notre part, nous rechercherons donc des représentations phonologiques qui satisfassent à deux critères : (i) en tant que modèle, elles doivent être incarnées et figurer un savoir-faire de coordination biomécanique qui rende compte des activités de production et de perception (perçaction) ; (ii) en tant qu'explication, le modèle ne se prend pas lui-même pour la figuration d'entités mentales confinées « dans le cerveau » (paradigme internaliste) mais pour celles de processus distribués sur le corps vivant en interactivité avec d'autres corps dans le milieu écologique (paradigme de la cognition distribuée) (Gallina, 2011, Steiner, 2011). Par exemple, le « son » de la parole existe à la fois comme événement acoustique (étranger aux contenus de la perception) et comme événement auditif (l'événement sonore qui survient à la conscience, lui-même étranger à la description de l'événement acoustique), sachant que le « son auditif » est obtenu en réalisant inconsciemment le geste de sa production phonatoire : à l'époque actuelle, il n'est plus possible de modéliser des représentations abstraites des sons d'une langue dans un cerveau humain en ignorant le réseau phénoménologique impliqué dans la constitution du phénomène « son de parole » considéré sous ses divers angles interactif (corps parlant, milieu, corps écoutant – en sachant que le son acoustique du « milieu » n'est pas lui-même objectif ni autonome, mais bien une construction réalisée par un acousticien en acte et médiée par une instrumentation technique et conceptuelle).

Selon Ladd (2014 : 367-369), les idéalizations (représentations métadiscursives et mentales) en théorisation phonologique basées sur le phone sont de plus en plus remises en question par les avancées empiriques de la phonétique. Dans cette situation, Ladd évoque plusieurs solutions : le rejet de la phonologie formelle (Port et Leary, 2005, le langage comme système de symboles formels discrets) ; Ladd rappelle que cette critique est valable pour la phonologie du 20^e siècle en général et pas uniquement pour la phonologie générative. L'autre solution est le rejet de la phonologie en bloc, comme le fait la théorie exemplariste (toute production phonétique repose sur un train d'exemplaires antérieurs enregistrés et non sur une représentation mentale abstraite sous-jacente ; les exemplaires sont liés entre eux par réseaux analogiques : Goldinger, 1996 ; Coleman, 2002). La 3^e solution est donc la voie de l'*embodiment* : la phonologie de laboratoire (*Laboratory Phonology*) représentée par la Phonologie Articulatoire (*Articulatory Phonology*) de Browman & Goldstein (1986, 1989), Goldstein & Browman (1992), Gafos & Goldstein (2012) et beaucoup d'autres). Dans ce cadre, les représentations phonologiques ne sont pas des symboles formels abstraits, mais des gestes (*gestures*) définis comme des abstractions quantitatives.

La corporalisation des représentations phonologiques comprend deux étapes. D'une part, motoriser les traits, qui deviennent des gestes ancrés dans une biomécanique et perception multimodale, comme le fait le kinakème de de Courtenay (1915). D'autre part, contextualiser le phonème dans la syllabe, elle-même posée comme représentation motrice élaborée dans un cadre incarné. La phonétique motrice (*motor phonetics*) de Stetson (1951) pose radicalement le

problème en refusant aux consonnes et voyelles le statut de représentations unitaires abstraites, discrètes et autonomes : posant la syllabe comme unité minimale de séquence gestuelle, il y inscrit les phonèmes comme structures de gestes articulatoires solidaires et orchestrés, insérés en un moment chronologique du parcours syllabique (phonotaxique), et interdépendants au sein de cet enchaînement (la coarticulation). Sa phonétique motrice ne peut pas ne pas impacter la nature des représentations phonologiques, aussi hésite-t-il entre feindre l'existence d'un relatif consensus sur la question et laisser libre cours à l'expression du constat désabusé de son inexistence :

There has been some general recognition of the articulation as the fundamental event in phonology, rather than the acoustic pattern. (Stetson 1951, 3)

In phonemics identification and classification [of phonemes by features] are usually primary ; and the movement complex producing the significant trait is hardly considered or ostentatiously ignored, although it is obvious that uttering the phonemes is quite as important as perceiving the phonemes in this the most important form of communication. (Stetson 1951, 6)

Le premier extrait exprime en fait le vœu pieux de Stetson au début de sa présentation des principes de la phonétique motrice ; la seconde fait état de son constat désabusé, et de fait on observe toujours ce « ostentatiously ignored » plus de soixante ans plus tard, notamment dans la conclusion de Goldsmith dans le chapitre sur la syllabe du *Handbook of phonological theory* malgré son accord sur le statut de la syllabe comme unité de base :

My conjecture is that the syllable is ultimately best regarded as the lowest level (or one of the lowest levels) of rhythmic reoccurrence of possibilities in the language. Some might want to see this as the reflection of gestures made by the articulatory apparatus, a view that we have not surveyed in this chapter. For myself, I think that such a view analyzes language at the wrong level of abstraction or granularity : the correct level of abstractedness for the description of language is higher than that of jaw gestures. Sonority, and the wave-like occurrence of peaks of sonority, seems to me to be the fundamental pattern of syllabification in language. (Goldsmith et al.: 2014, 194)

Goldsmith assume d'avoir censuré les travaux sur l'inscription motrice des représentations phonologiques de la syllabe, semble prendre de très haut l'idée d'un ancrage gestuel de ces représentations, et ramène ces gestes à des « mouvements de la mâchoire », ce qui surprend quelque peu quand on connaît la complexité des travaux sur la coarticulation, la temporalité des matrices de gestes et plus généralement la qualité vocale, ou encore la théorie du Larynx Articulateur en phonologie (Esling, 2005 ; Esling et al., 2019), bien loin de la mâchoire (mais en relation avec elle). On ne voit guère pour quelle raison le corps serait le « mauvais niveau d'abstraction », et hormis l'expression autoritaire de l'acte de foi, aucune justification explicite ne vient étayer cette « conjecture ». Il faut garder à l'esprit que ce dualisme corps / esprit remonte au moins aux dialogues de Socrate et que la pensée dite occidentale a cultivé un rejet du corps sensible au profit d'un esprit immatériel rationnel ou religieux, relayé et amplifié par la pensée chrétienne, et finalement sublimé par le paradigme computationnel porté par le cognitivisme, qui voudrait voir dans la cognition humaine une activité comparable aux calculs numériques dans l'environnement abiotique de l'ordinateur métallique et électronique. De notre point de vue, il faut au contraire éviter le déni phénoménologique, reconnaître au corps son rôle actif dans la constitution des phénomènes, et accueillir des représentations motrices incarnées en phonologie à potentiel phonosémantique, en repartant des analyses de Stetson :

Every utterance is a movement consisting of the phrase which is the larger, inclusive unit of which the breath group, foot and syllable are organic parts. For this movement the posture is the adjustment of the abdomen-diaphragm and external chest muscles

for regulated expiration ; the movements of the breath group, and of the feet within the breath group are due to the abdomen-diaphragm movements of expiration. The series of small expiratory movements, the breath groups, constitute the phrase. The phrase is followed in the breathing cycle by a rapid inhalation. Although the utterance may consist of only one or two syllables, all the fundamental units are present. (Stetson : 1951, 4)

La phonétique motrice segmente l'énoncé (*utterance*) en niveaux d'analyse (groupe, pied et syllabe) correspondant à un travail du souffle expiratoire réalisé par les muscles abdominaux (diaphragme) et intercostaux : les segments sont définis en termes d'opérations musculaires et d'effets pneumatiques en amont du couple phonation / articulation et dans une perspective radicalement incarnée, dite motrice. L'implication est que l'objet empirique de la phonétique est la coordination de l'expiration et de la phono-articulation, dont l'étude devra s'appuyer sur un domaine anatomique élargi et couvrant toute la machine parlante (cf. Seikel et al., 2019).

Le plus petit segment, la syllabe, résulte d'un souffle impulsé par le diaphragme et découpé par des saccades des muscles intercostaux : il en résulte un schème syllabique constitué d'un transit borné par un amorçage et une clôture ; à ce niveau, ce schème syllabique, sorte de protosyllabe antérieure à l'instanciation par des phonèmes voisés et articulés, fournit le gabarit pneumatique dans le lit duquel intervient la production et l'introduction de phonèmes. Lorsque ce souffle présegmenté s'engage dans la zone phonatoire (du larynx aux lèvres en passant par les plis vocaux), vont être introduits des phonèmes correspondant à des gestes de conformation (voyelles) et de perturbation (consonnes) : la conformation orale (avec ou sans nasalisation) au cours du transit induit la production d'harmoniques F1 et F2 relatifs à la fréquence fondamentale F0 induite par la fréquence oscillatoire des plis vocaux, déterminant le phénomène acoustique de la voyelle en tant que traitement du transit de la protosyllabe. Des interventions perturbatrices de couples d'articulateurs actif/passif (apex et couronne), bi-actifs (lèvres) et plus (larynx) vont introduire une ou plusieurs actions motrices à effet acoustique au cours de l'amorce et/ou de la clôture qui, une fois investis, deviennent attaque et coda de la syllabe instanciée : l'attaque reçoit zéro, un ou plusieurs gestes générateurs de consonnes ; la coda, idem. Consonnes et voyelles ne sont donc pas des segments séparables, mais des spécifications des bornes de la syllabe. La représentation phonétique des phonèmes devrait les concevoir comme moments de la syllabe (ce que ne fait pas le *Cambridge Handbook of Phonetics* de Knight & Setter, 2022)¹³, et la phonologie de la syllabe (Angoujard, 1997 ; Scheer, 2015) aurait à intégrer les abstractions articulatoires (Kennedy 2022, 694) en corrélant la motricité syllabique intégrante et la motricité phonémique intégrée. Krakow (1989) étudie les consonnes comme gestes moteurs structurant la cinématique de la syllabe (mais sans insister sur le dialogue avec l'expiration).

Le plus souvent, le transit de la protosyllabe pneumatique est voisé et réverbéré par une configuration orale génératrice de voyelle, mais certaines langues comme le tchègue réalisent en cette phase une perturbation consonantique comme le r apical vibrant, livrant comme noyau un r dit syllabique. Si la clôture de la protosyllabe pneumatique ne subit pas l'intervention d'une perturbation articulée génératrice d'une (ou plusieurs) consonnes, la voyelle peut s'y poursuivre, comme le font les voyelles fermées du français (peut) ou les voyelles dites libres de

¹³ Dans le précis de Knight & Setter (2022), la section I (*Segmental Production*) comprend le chapitre 1 (*Physiological Foundations*), le chapitre 2 sur les voyelles et le chapitre 3 sur les consonnes (et d'autres chapitres sur la coarticulation et le discours enchaîné) : la double surprise est que (i) la syllabe n'est pas reconnue comme segment et ne fait pas l'objet d'un chapitre malgré son rôle intermédiaire entre phonèmes et discours enchaîné (vs Stetson, 1951), et (ii) le chapitre 1 sur la physiologie présente l'anatomie des systèmes respiratoire et phono-articulatoire, mais ne dérive pas de leurs activités une théorie de la syllabe (vs Stetson, 1951 et MacNeilage, 1998). Ce traité, le dernier en date, et bien que dédié à la phonétique, fait l'impasse totale sur une représentation incarnée de la syllabe : le message de Stetson ne passe décidément pas.

l'anglais (*free vowels*, par opposition aux *checked vowels*, abrégées et profilées en fonction du geste ultérieur en coda). Selon la Frame/Content Theory de la syllabe (MacNeilage, 1998), la structure syllabique s'inscrit dans le cycle des battements de la mâchoire inférieure : la voyelle de transit s'installe dans la phase descendante ouvrant l'espace où peuvent se gérer les degrés d'aperture langue / palais, alors que les consonnes d'amorçage et de clôture se réalisent par constrictions ou occlusions pendant les phases ascendantes intermédiaires (et on obtient des syllabes CV ou CVC selon que la clôture soit déléguée à la syllabe suivante ou intégrée anticipativement par la coda de sortie ; cette différence gestuelle stratégique est à l'origine de types syllabiques dans les langues du monde, par exemple le japonais comme langue à syllabe CV nasalisable, et doit donc être intégré aux représentations phonologiques, à la fois métalinguistiques et épilinguistiques).

Consonnes et voyelles sont coarticulés : /k/ français se positionne différemment selon qu'il intervient sur la préparation d'un transit conformé en /i/ ou /u/ (*qui* vs *cou*) ; l'anticipation coarticulatoire est variable selon les langues (ou entre leurs dialectes) : en anglais américain l'anticipation est forte (*new* prononcé *noo*) alors qu'en anglais britannique elle peut être bloquée par une transition (*new* prononcé yodisé, le yod retardant, mitigeant ou bloquant l'arrondissement des lèvres au moment de prononcer /n/). Les gestes articulatoires sont donc solidaires (coarticulation). Il existe des matrices de solidarisation générale de l'ensemble de l'appareil phonatoire et articulatoire qui ont reçu le nom de modes en phonétique corrective (Delattre 1953) et de bases articulatoires (*articulatory bases* ou *settings*) dans une longue tradition de travaux issus de la germanistique (Viëtor, 1887 ; Vockerad, 1925 ; Van Ginneken, 1907 et 1933 ; Gili Gaya, 1956 : 158-160 pour l'espagnol ; Honikman, 1964 ; Laver, 1978 ; Wadsworth, 1979 ; Schourup, 1981 ; Jenner, 2001 ; Gick et al., 2004 ; Roach, 2016) ; on parle parfois de *types phonétiques* sans référence aux *settings* (Gordon & Ladefoged, 2011), et la distinction *basis / setting* est elle-même discutée (Borissoff, 2012) et a des implications phonologiques (Gil Fernández, 2006), au risque de faire entrer le « diable dans les détails » de l'inventaire des sons des langues (Ohala, 2009). Dans le paradigme énonciatif, on définit ces bases articulatoires comme des conformations corporelles incarnées générales, propres à chaque langue, et génératrices d'une qualité vocale commune aux locuteurs qui l'utilisent, que l'on peut nommer par commodité « la voix des langues » (Bottineau, à paraître, 2023) : les locuteurs se co-ajustent par interaction mémétique, contraignant leurs corps parlants à s'adapter de manière convergente aux orientations musculaires et biomécanique d'une discipline sportive qui les fédère (raison pour laquelle la voix de la langue se manifeste de manière croissante avec le vieillissement du corps, et de manière exacerbée dans certains arts vocaux comme le chant flamenco pour l'espagnol).

Or, si l'on suit Stetson, consonnes et voyelles sont des moments de la syllabe, elle-même préparée par la segmentation de l'expiration. La démarche de Stetson implique une critique implicite de la double articulation de Martinet : entre le monème (ou morphème au sens général) et le phonème, se trouve la syllabe ; entre la morphématique et la phonématique, la syllabatique¹⁴. Cela implique que la représentation métalinguistique des sons de la langue ne peut être formalisée adéquatement que si l'on définit les phonèmes comme des matrices pondérées de gestes articulatoires orchestrés et survenant en cours de performance d'une protosyllabe pneumatique. Dans le paradigme structuraliste, la notion d'archiphonème correspondait précisément à la nécessité d'extraire des classes de voyelles et consonnes comme matrices de gestes articulatoires non profilées et à profiler ultérieurement au gré des inscriptions

¹⁴ En français, ce terme rare, absent des dictionnaires terminologiques, semble avoir été introduit par Schenkel (1981). En anglais, *syllabatics* est parfois utilisé pour décrire l'expérience de déclamation des haikus en poésie ou l'élocution des rappers : « Eminem's acrobatic syllabatics » ; mais il est également synonyme de *syllabaire* (écriture dont les symboles encodent des syllabes, comme les katakana et hiragana du japonais).

syllabiques, par exemple un archiphonème A réalisé différemment dans les mots *that* et *all* en anglais (avec, curieusement, le graphème comme marqueur et symbole de l'archiphonème dans la plupart des cas)¹⁵. En outre, l'orchestration de l'articulation des voyelles et consonnes dépend du régime énergétique de la protosyllabe pneumatique injectée : en français, la force et l'énergie des pulsations prosyllabiques varie peu, de même que la force d'articulation des phonèmes intervenant pour former la syllabe instanciée, ce qui produit l'effet général d'un nivellement accentuel et intonatif caractérisant pour une part le mode tendu de Delattre (1953, 1965, 1969 ; cf. Lambert-Drache, 1997) ; en anglais, la force et l'énergie des pulsations syllabiques varient avec les schèmes métriques inscrits dans le lexique, ce qui va faire covarier la force d'articulation vocalique et consonantique avec la force d'expiration protosyllabique et générer un dialogue entre les deux niveaux par lequel le déploiement des voyelles s'harmonise avec le tonus protosyllabique (voyelles accentuées vs inaccentuées ou amuïes ; consonnes occlusives non voisées d'attaque séparées de la voyelle par une expiration : *cat*, *top*, *pit*).

Ce dialogue entre la syllabation pulmonique et l'articulation phonatoire et articulatoire peut être rapproché du fonctionnement de la cornemuse d'Irlande, la *uilleann pipe*¹⁶ « tuyaux de coude » (du gaélique irlandais *píobáí uilleann*, *píobáí* « tuyaux » + *uilleann* « de coude », génitif de *uillinn* « coude ») : le sac (dans le rôle des poumons) est gonflé par le joueur (dans le rôle des muscles abdominaux), mais le souffle injecté dans les tuyaux pour la production des notes (la protosyllabe injectée dans la zone phonatoire-articulatoire) est pressurisé et contrôlé par une série de soufflets (*bellows*) localisés dans la région du sac que le joueur en position assise coince entre son ventre et son coude, régulant ce souffle par l'action du coude (*elbow*, dans le rôle des muscles intercostaux) indépendamment du gonflage buccal (et il peut donc parler tout en jouant, tant que le sac reste suffisamment pressurisé). Les notes résultent de la vibration produite dans la traversée du chalumeau (*chanter*, dans le rôle du phonateur) dont les huit trous sont contrôlés par les doigts (articulateurs). De même que la note émise par le chalumeau ne peut être décrite sans référer à l'injection silencieuse opérée par les soufflets contrôlés par les coudes, la voyelle phonatoire (audible) ne peut être décrite sans référer à l'activité silencieuse des abdominaux et intercostaux. Cette analogie permet de préciser un écart notable entre phonologie conventionnelle et phonologie motrice : la phonologie classique porte son attention sur la région du corps où se produisent les perturbations effectivement productrices de signal acoustique.

Une phonologie motrice élargit le champ en incluant des biomécanismes silencieux parce que localisés à l'extérieur de la zone où se produisent les interactions corps / air génératrice de bruit. Une telle inclusion permet à la fois de penser les phonèmes en tant que moments syllabiques et les traits gesturaux en tant que kinèmes articulatoires modélisables en amont de leur engagement syllabique. En la matière, les théories traitent différemment la question de l'ordre des phonèmes au sein du morphème mono- ou polysyllabique. Certaines approches reconnaissent une pertinence à l'ordre de progression linéaire : dans le cadre de la

¹⁵ En anglais, avant le grand changement phonétique (*great vowel shift*), on observait en gros une correspondance graphème / phonème vocalique, encore largement observé aujourd'hui en allemand en dehors du cas des digraphes *ei* et *eu*. L'archiphonème est en ce cas une représentation métalinguistique qui revient à rapatrier en synchronie le phonème commun correspondant à un état diachronique ancien, antérieur à la diversification des voyelles. La représentation phonologique épilinguistique et incarnée consisterait à supposer que la diversité des *o* et *a* phonémiques en anglais selon l'insertion syllabique correspond à un schème gestuel commun dont le profil varie avec les insertions syllabiques. L'hypothèse est particulièrement lourde, et peut-être superflue : les locuteurs continuent à classer les voyelles de *what* et *that* sous A parce que le graphème conservé A leur prescrit de le faire et non pas nécessairement parce qu'il y aurait un schème articulatoire commun sous-jacent. Pour l'anglais, la cognématique s'appuierait ainsi sur le couple phonographématique plutôt que sur le phonème seul : l'enjeu théorique et empirique est crucial, et affecte la manière de penser la cognématique en typologie des langues (les langues romanes ne posent guère ce type de problème).

¹⁶ <https://www.youtube.com/watch?v=dXJ3k5YXISs>

cognématique, Bottineau (2004) propose qu'un cognème joue un rôle de majeure cognitive en position initiale d'attaque syllabique et un rôle de mineure cognitive en coda (avec l'exemple de N de négation, initial dans *no* et final dans l'article indéfini *an*) ; dans le cadre de la chronophonétique (Macchi, 2018), l'ordre des phonèmes n'est pas une simple séquence linéaire, il atteste d'une « syntaxe temporelle » vectrice d'un processus d'engendrement des morphèmes par adjonction de submorphèmes. D'autres approches insistent sur la réversibilité des parcours submorphémiques : dans la TME de Bohas, la matrice est formée de l'appariement non ordonné de deux traits, et leur réalisation par des phonèmes au sein des étymons forme des couples réversibles. Bottineau (2014, 2021) suggère que ces divergences apparentes relèvent de l'altérité typologique des gabarits syllabiques et métriques au sein desquels se forment les notions lexicales selon les langues considérées : le lexique sémitique tend à re-générer le mot par le schème matrice-étymons-radicaux en chacun de ses réemplois par les locuteurs (d'où l'énorme productivité morphologique du dispositif, qui génère un nombre surprenant de mots plus ou moins synonymes à partir d'un nombre limité de matrices et d'étymons), tandis que les lexiques germanique et roman tendent à figer des étymons préconstruits à partir de matrices sous-jacentes, créant des phénomènes de figements ordonnés liés aux contraintes locales dans la syllabation, telles que l'ordre des consonnes au sein de l'attaque ou de la coda ; par exemple, l'étymon <l-p/b>, utilisé entre autres pour les objets sphériques ou perçus comme gonflables, est réversible (*balle, boule, bulle, bol, lobe, lippe, loupe*) sauf lorsqu'il se concentre en attaque ou en coda (*blague à tabac*). On aurait donc des matrices non ordonnées qui se projettent avec ou sans mise en ordre dans les étymons selon les formats d'insertion syllabique et les degrés de figement qu'ils impliquent. On n'explorera pas davantage cette question dans cette étude, mais il est clair qu'une phonologie incarnée contextualisée par la syllabe aura des conséquences pour l'étude des faits d'ordination et de réversibilité submorphémique au sein des unités lexicales et grammaticales dans un type de langue donnée.

Penser la voyelle dans le cadre d'un rapport de force entre expiration et phonation / articulation inclut par avance la problématique de la covariation du bras de fer qui préfigure, par exemple, la plasticité de la voyelle anglaise selon à la fois la structure syllabique (consonnes en attaque et en coda) et la relation entre le tonus expiratoire et les tonus phonatoire et articuloire (rapport de forces communément nommé « accent d'intensité »). C'est la compréhension de ces variations, de source biomécanique, qui permette *in fine* de maîtriser les difficultés d'appariement entre systèmes phonologiques et invariants phonosémantiques stylisés, comme ceux de la cognématique.

3. Vers un modèle incarné de la phonématique d'une langue

Pour modéliser la phonématique d'une langue, on esquisse ici un modèle phonatoire global qui considère le corps comme machine organique vivante à composantes interdépendantes. L'objectif est de produire une vue panoramique des rapports biomécaniques à l'origine des événements phonatoires responsables des phénomènes auditifs en partie ressentis comme « voix » en tant que propriété émergeant des segments de parole (tirades, syllabes, phonèmes). Un exemple de cette démarche est fourni par la dernière section du chapitre sur la phonétique articuloire de l'introduction à la linguistique descriptive de Gleason (1955 : 15.23 p. 255-256) : l'auteur distingue deux niveaux de segmentation, le groupe de souffle et la syllabe, et rapporte leur production à deux biomécanismes fonctionnels distincts. Les groupes de souffle sont produits par une traction continue des muscles abdominaux sur les poumons ; et pendant cette opération, les syllabes sont préparées par des impulsions discontinues produites par les muscles intercostaux sur les mêmes poumons, scandant le flux continu en amont de la vibration des plis vocaux et de l'articulation qui vont introduire les événements acoustiques dans le flux

prémodélé. Or l'observation des langues montre que le travail des poumons est lui-même une variable entre langues, et que la manière de se coordonner avec l'articulation l'est également, avec des antagonismes et rapports de forces contrastés. On a pour habitude d'exclure la « soufflerie » de l'analyse de la parole parce qu'elle-même ne génère pas de son, et on ne s'intéresse généralement qu'à la « bruyerie », à savoir le vibrateur laryngal (les plis vocaux) et les articulateurs buccaux et autres, qui réalisent la conversion de l'énergie cinétique du souffle phonatoire en énergie de propagation des ondes acoustiques. Notre démarche réinclut la soufflerie de manière à préciser son rôle propre et sa relation aux autres étages.

La machine à parole comprend deux blocs : la *soufflerie pulmonaire* et la *bruyerie* localisée dans le cou et la tête.

La *soufflerie* articule deux fonctions couplées :

- la *fluxation*, la production d'un flux continu et étendu par la pression des muscles intercostaux, qui prépare le lit du groupe de souffle ;
- la *pulsation*, la segmentation de ce flux continu et flux discontinu par l'action répétée des muscles intercostaux, qui découpent le flux en une chaîne d'impulsions qui fournira les maillons expiratoires à travailler pour produire des syllabes chargés de phonèmes en leurs différentes phases.

C'est ce couplage qui fait de l'ensemble une unité d'action contrôlée par le locuteur, exposée à la perception, réflexive, et dénombrable comme « unité d'autonomie vocale » (De Cornulier, 1982)¹⁷. La relation fluxation / pulsation est antagonique : la seconde intervient sur la première exactement comme la vibration et l'articulation interviennent comme antagonistes sur le souffle phonatoire au niveau supérieur. Huche et Ablali (1991 : 50) parlent « [d']expiration active régulée par activité inspiratrice persistante ». Les deux étages sont liés par une analogie fonctionnelle qui permettra de les solidariser : dans la soufflerie, un flux continu est brisé par un battage ; dans la bruyerie, ce flux discontinu est à son tour profilé par un battage vibratoire et articuloire. En français, la bruyerie est rythmée par la soufflerie : l'articulation installe la phonation dans le rythme stabilisé par la pulsation, ce qui en typologie phonologique fait du français une *syllabe-timed language* (langue cadencée par la syllabe), par opposition à d'autres cadencages (anglais : l'accent ; japonais : la more) (Pike, 1945 : 34-35 ; Roach, 1982), eux-même produits par d'autres hiérarchisations motrices. En outre, le français stabilise les voyelles en ajustant le tonus articuloire et phonatoire au tonus expiratoire, à la différence de l'anglais qui les fait covarier, solidarisant qualité vocalique et pression expiratoire.

La bruyerie coordonne diverses activités incarnées elles-mêmes solidaires :

- la *cavitation orale* : le maxillaire inférieur alterne des mouvements d'abaissement et de relèvement qui contrôle l'ouverture et la fermeture d'une *cavité*, une chambre phonatoire buccale à géométrie variable au sein de laquelle pourra jouer l'articulation, en particulier les mouvements linguaux et labiaux.

¹⁷ Angoujard (1997, 18) cite ce chapitre de De Cornulier (1982, 58-67) « Pourquoi dénombrer les syllabes » et en tire un objet phonologique abstrait. Or le dénombrement des pieds du vers dans l'exercice de métrique ne consiste pas à écouter des événements acoustiques mais à compter les pulsations phonatoires que l'on se sent produire soi-même en les déclamant, ce qui nous situe directement dans la cognition incarnée réflexivisée (action, proprioception focalisée et semiosis). De même, Angoujard remplace le terme de « pulsation syllabique » (Lotz 1972), qui concerne l'activité en première personne, par celui de « sommet syllabique », qui produit une représentation abstraite à partir de l'analyse du signal en troisième personne, réputée objectiver l'audition. Ce passage est emblématique du glissement phénoménologique responsable de la divergence entre cognition incarnée en acte et cognition représentationnelle calculatoire (computationnelle).

- *L'oscillation glottale* : les plis vocaux sont mis en vibration passive au passage du flux d'air expiré et accéléré par le rétrécissement (effet de Bernoulli) selon des configurations variables entre les langues (permettant par exemple le *vocal fry* optionnel ou systémique) et solidaires des autres activités. Comme dans les autres composantes, le régime énergétique de production d'effet est antagonique (Story, 2010 : 49-51) et co-articulé. En anglais, pour l'exclamation « Wow !!! », la puissance vibratoire des plis vocaux covarie avec la puissance musculaire de l'abaissement du maxillaire responsable de la *cavitation*, avec l'empan de l'abaissement laryngal (Esling, 2019), et aussi avec la puissance de la *fluxation*, le souffle phonatoire initial injecté par les poumons sous la pression des muscles abdominaux. Cette coordination ne s'observe pas dans « Waou !!! » en français et détermine un contraste important entre les effets obtenus, y compris la manière dont les pratiques et contraintes respectives profilent les corps parlant dans la durée selon la langue considérée.
- *La configuration linguale vocalique* : pendant l'expulsion d'une unité de pulsation oscillée par les plis vocaux produisant la fréquence fondamentale, et pendant une phase d'abaissement maxillaire (*cavitation*), la langue se positionne à diverses distances du plafond (degrés d'aperture) et tout en se repliant de manière à générer divers détroits (points de rétrécissement ou de pression), générant des configurations de réverbération harmonique de la fréquence fondamentale constituant les matrices formantiques F0/F1/F2 caractérisant chaque voyelle, et ce à tous les niveaux tonaux possibles (grave, médium, aigu). Le noyau de la syllabe est produit quand l'unité de fluxation est bruitée par le rapport de forces liant l'oscillation glottale, la cavitation buccale, la configuration linguale, le coulissage laryngal, la protrusion ou l'étirement labial, et d'autres, comme le pivotage de la luvette pour la nasalisation (lui-même à géométrie variable et solidaire) ou les divers types d'incurvation latérale du dos de la langue, qui concernent également les voyelles en général (Stone & al., 1988 ; Stone & Lundberg, 1996). L'énergie investie dans la configuration est liée aux autres : dans l'anglais *see*, la force de la pression dorso-vélaire est sans rapport avec celle de *si* français, mais elle augmente à mesure que décroît la pression de l'air de la pulsation : *see* sonne plus comme un *i* français en fin de voyelle longue qu'au début ; de même, dans *you*, l'arrondissement labial et l'abaissement laryngal progresse pendant la dépressurisation de la pulsation : le « ou » long passe progressivement d'un phonon initial ressemblant à /y/ français (sans la protrusion labiale) à un phonon final plus proche de « ou » ; la voyelle « ou » française ne pratique pas cela, même si l'allonge. C'est encore une affaire de coordination psycho-motrice liant des variations musculaires et énergétiques interactives.
- *La préhension linguale consonantique* : en début et fin de noyau vocalique, la langue peut « empoigner » ou « saisir » la pulsation par resserrement gradué (constriction) ou la bloquer en palpant brièvement le palais (occlusion : stop et plosion), générant les consonnes en attaque et coda en amont et en aval du noyau. La caractéristique audio-vocale des consonnes dans une langue donnée dépend de la coordination avec les autres activités : en français, pour les plosives /b/, /d/ et /g/, l'oscillation glottale (le voisement) précède l'occlusion (bilabiale, apico-coronale ou dorso-vélaire), créant l'impression d'un fort antagonisme entre pulsation et articulation, à l'origine de la saillance phonologique du contraste entre traits non voisés et voisés. En allemand, pour les consonnes correspondantes, le voisement est libéré par la plosion, avec une absence totale d'antagonisme pulsation / articulation et une saillance contrastive moindre pour le trait de voisement, d'où la faible importance de sa percevabilité, voire son évanescence dans certains dialectes. En anglais, pour les plosives non voisées /p/, /t/ et /k/, l'occlusion, très énergique, entre en antagonisme avec la pulsation, d'où un échappement d'air à l'instant de déblocage entre consonne et voyelle (*peer* [p^htə], *tea*

[^htɹə], key [k^hi:]], que l'on ne retrouve pas si la consonne est précédée de s (*speed, steal, scheme*) : la sifflante initiale produit un effet de fuite ou de « crevaison » qui réduit suffisamment la pression pour neutraliser l'antagonisme ; le voisement produit le même effet en convertissant une partie de l'énergie cinétique en énergie acoustique (*beat, deep, geese*). En français, l'antagonisme pulsation / occlusion s'opère à un niveau énergétique plus bas et stabilisé, dans un parcours syllabique non décroissant, d'où les effets décrits par le « mode tendu » de Delattre.

- Le *coulissage et les constrictions laryngales* : cet organe complexe intervient dans l'articulation de consonnes distinctes (langues sémitiques) et dans la co-articulation de voyelles comme /u/, et ce en relation avec les autres activités, donc dans le cadre d'orchestrations propres à chaque langue ; les laryngalisations de /u/ ne sont pas les mêmes en français, anglais, allemand et espagnol et relèvent de réseaux différents.
- Le *pivotage de la luette*, à des niveaux et forces variant selon les langues et la solidarité avec les autres activités (cf. la résonance remarquable de la nasalisation en chti, due à un abaissement renforcé de la luette en coarticulation avec une pulsation appuyée).

Une phonologie incarnée consiste donc à proposer des représentations incarnées et motrices des phonèmes en retenant les propriétés pertinentes pour leur distinctivité, y compris celles issues du rapport à l'expiration (cas des diphtongues de l'anglais), mais en excluant les phénomènes phonétiques non distinctifs et inexploités en phonosémantique (par exemple l'expiration après les plosives non voisées de l'anglais s'avère non pertinente). Une phonologie incarnée et motrice modélise et transcrit des réseaux de rapports de forces au sein d'orchestrations de gestes articulatoires (Ball, 1993 ; Esling, 2010) en prise avec l'expiration et susceptible de faire l'objet d'une exploitation phonosémantique. Dans ce cadre, un phonème n'est plus une matrice de traits, mais un *macro-geste phonatoire*, un orchestre de *micro-gestes articulatoires* hétérogènes et solidaires : un *kinome* formant un réseau de *kinimes*. Le terme *kinome* est formé sur la racine *kin-* « geste » + le suffixe *-ome*, aphérèse de *-some* « soma », « corps », « ensemble solidaire », « structure », comme dans *génome* et *connectome* « plan complet des connexions neuronales d'un cerveau » : il caractérise le phonème comme orchestre de gestes articulatoires. Le terme *kinime* (*kin-* + suffixe néologique *-ime* signifiant une micro-composante d'un réseau désigné par *-ome*, le <i> étant analogique de celui de *micro*) désigne les micro-gestes participant à l'orchestre, comme la protrusion labiale, l'abaissement laryngal ou la pression dorso-vélaire, ancrés dans des schèmes musculaires spécifiques (les muscles labiaux obiculaires pour la protrusion en français, distincte de l'arrondissement en anglais). Enfin, par un heureux hasard, la racine *kin-* elle-même contient l'allophonesthème <kn-> qui réalise l'archiphonesthème <CN-> évoquant la région bucco-nasale (Philps, 2001 : 213) : si le morphème lexical *kin-* évoque l'idée de geste en général, le phonesthème <kn-> la rattache à la région qui nous intéresse, les organes de la parole. Le kinome est un modèle analogique incarné qui inclut les redondances, par opposition au phonème du structuralisme.

4. Les voyelles du français

La qualité auditive des voyelles du français est liée à un équilibrage des rapports de force entre expiration, phonation et articulation, que décrit précisément la phonétique corrective de Delattre ; ce qui suit présuppose ce contexte et importe en phonologie des représentations motrices qui doivent beaucoup à la phonétique. Le système phonématique du français s'installe dans le cadre de la base articulatoire de cette langue, qui comprend peu ou pas de fluctuation de tonus syllabique (accent d'intensité), pas de covariation articulation / phonation / expiration

(vs l'anglais), et pas non plus de diduction (avancement de la mâchoire inférieure par pivotement comme en espagnol ibérique, cf. Bottineau, 2023).

Considérons la séquence de voyelles antérieures par degrés croissants d'aperture /i/-/e/-/ε/-/a/. Bouche fermée, tous les kinimes sont au degré zéro ; aucun geste, pas de parole. Sur cette base, pour produire une voyelle, il faut coordonner au moins

- La *cavitation* : abaisser la mâchoire de un ou plusieurs degrés ;
- La *configuration* : relever la langue d'un degré et aboutir au degré d'aperture correspondant à la configuration harmonique F1/F2
- La *pulsation*, l'injection du segment de souffle (fluxation) préparé par la soufflerie et segmenté en unité syllabique bornée ;
- L'*oscillation* : mise en vibration passive des plis vocaux dans la configuration (cavitation + palpation) selon une période contrôlée par la configuration génératrice de la matrice F0/F1/F2.

Le travail biomécanique de production de la voyelle est une action de la bruiteirie sur l'unité d'expiration produite par la soufflerie, la protosyllabe. Il consiste à injecter une unité de souffle syllabique segmenté (fluxation et pulsation) dans une bouche animée par la cavitation (maxillaire) et configuration (linguale) en laquelle l'oscillation (plus vocaux, F0) sera automatiquement réverbérée par les harmoniques (F1 et F2) ; le produit physique de ce travail est la voyelle acoustique en tant qu'évènement mêlant trois fréquences.

La voyelle /i/ forme un kinome formé de la matrice de kinimes suivantes : abaissement gnathique de degré 1, relèvement glossique dorso-vélaire de degré 1 avec compensation de l'abaissement gnathique, donc contact et pression (aperture 0, voyelle fermée) ; vibration glottique (voisement). Dans la syllabe française, la pulsation est généralement débloquée par un coup de glotte [ʔi]. L'aperture zéro (voyelle fermée) est obtenue par compensation antagonique de l'abaissement gnathique (+1) et du relèvement glossique (-1), deux kinimes d'orientations inverses.

Pour passer aux voyelles /e/, /ε/ et /a/, il suffit d'augmenter d'un cran pour chaque l'abaissement gnathique en maintenant le relèvement glossique d'un degré : pour /e/, abaissement gnathique de degré 2, relèvement glossique de degré 1 ; l'aperture de degré 1 (voyelle mi-fermée) est obtenue par deux kinimes d'envergures différenciées (+2 -1 = 1). Pour /ε/, abaissement gnathique de degré 3, relèvement glossique de degré 1, aperture +3 -1 = 2 (voyelle mi-ouverte). Pour /a/, abaissement gnathique de degré 3, relèvement glossique de degré 0, aperture +3 -0 = 3 (voyelle ouverte). Au plan gestuel, /a/ est la voyelle la plus simple au sens où elle combine un abaissement gnathique de degré 3 non corrigé par un relèvement glossique : le degré d'aperture langue / palais correspond à l'abaissement de la mâchoire. Il suffit d'ouvrir la bouche pour produire cette voyelle sans intervenir avec la langue, ce qui donne à ce degré d'aperture un kinime de moins qu'aux autres. L'implication pour le potentiel phonosémantique est que /a/ est peu, voir pas, préhensible : la langue ne fait aucun geste de nature à modifier la qualité vocale. Par contraste, les autres voyelles sont plus complexes en ce qu'elles compensent les degrés d'abaissement gnathique par un degré d'élévation glossique, ce qui les rend préhensibles à proportion de la réduction du degré d'aperture, jusqu'au contact avec pression. C'est l'analyse de ce ressenti qui fait définir /i/ comme la voyelle de « l'étreinte » par Jünger (1934), dans une perspective motrice et haptique (Komel, 2019).

Pour passer de /i/ à /y/ et /u/, c'est, pour le français, un couple de kinimes corrélatifs qu'il faut faire jouer : la rétraction linguale, qui fait reculer le dos de la langue et la région du palais impactée ; et la protrusion labiale, obtenue par contraction des muscles obiculaires, responsable de la sensibilité francophone au trait labial dans le modèle de Vaissière (2007 : 70). Pour /i/, le

couple protrusion / rétraction est au degré 0 : ni l'un, ni l'autre. Pour /y/, il passe au degré 1 : un cran de protrusion labiale est couplé à un cran de rétraction dorso-vélaire ; pour l'observateur extérieur, les lèvres se projettent partiellement en avant et se resserrent incomplètement. Pour /u/, ce couple passe au degré 2 : deux crans de protrusion labiale (par contraction obiculaire renforcée) et deux cran de rétraction dorso-vélaire. Le timbre de ces voyelles en français est la conséquence directe de ce couplage de kinimes solidaires. Il est également lié à la stabilité de la configuration, « plate » en français (mais sinusoïdale en anglais).

On distingue donc deux axes de variation kinimique coordonnés pour les voyelles du français :

- L'axe de pendulation gnatho-glossique, celui qui fait jouer ensemble la cavitation (abaissement de la mâchoire) et la configuration (relèvement lingual) en faisant varier les degrés d'aperture ; en le faisant jouer seul, on obtient l'axe vocalique /i/-/e/-/ε/-/a/.
- L'axe de coulissement dorso-labial (spécifiquement français pour sa composante obiculaire), celui qui fait jouer ensemble les kinimes corrélatifs de la rétraction dorso-vélaire et de la protrusion bilabiale ; en le faisant jouer seul, on obtient l'axe vocalique /i/-/y/-/u/.

Il reste donc la possibilité de faire jouer ensemble la pendulation et la coulisse, c'est-à-dire d'activer le couple protrusion-rétractions aux degrés 1, 2 et 3 de l'aperture (obtenus par les configurations +2-1, +3-1 et +4-1 du compas gnatho-glossique) : on obtient les axes intermédiaires /e/-/ø/-/o/ et /ε/-/œ/-/ɔ/. Dans l'antagonisme qui les relie, c'est la coulisse qui dépend du compas : moins le compas est ouvert (aperture minimale), plus l'empan des degrés du couple dorso-labial est étendu ; plus le compas est ouvert (aperture maximale), plus l'empan des degrés du couple dorso-labial est réduit : /a/ vs /ɑ/, voire neutralisé. L'extension de la coulisse labio-vélaire est inversement proportionnelle à l'ouverture du compas gnatho-glossique qui gère la cavitation. Pour les voyelles nasales, il faut ajouter comme kinimes solidaires (i) la réduction de l'aperture par augmentation d'un demi degré du relèvement lingual, (ii) l'abaissement corrélatif de la luette, et (iii) l'abaissement corrélatif du larynx. Le français diffère d'autres langues sur le profil précis de ces coordinations, qui restent à doser quantitativement de manière intégrée ; par exemple les nasalisation fines de *don't* /dɔ̃nt/ en anglais britannique (prononcé « donte » avec l'accent français) et *home* /hoom/ en anglais américain (vs britannique /həʊm/) se profilent dans les environnements respectifs de la décroissance corrélative des énergies de pulsation (syllabe), d'oscillation (voisement) et d'articulation (le mouvement de la luette covariant avec les couples bilabiale et apico-vélaire), sans rapport avec le type français.

Cette première définition du kinome des voyelles est étroite et incomplète : elle se base sur l'idée que pour rendre compte des voyelles du français, il suffit de se contenter du couple linguo-labial. Or l'hypothèse du Larynx articulateur d'Esling dit autre chose : les mouvements du larynx, solidaires de ceux du couple linguo-labial, jouent un rôle constituant et déterminant dans la qualité vocale de la voyelle, notamment /u/. Dans notre perspective phonosémantique, cela implique une relecture de l'expérience motrice, auditive et tactile du phonème : /u/ se caractérise par au moins cinq gestes (maxillaire, bilabial, lingual et laryngal, pour la phonation et l'articulation). La cognématique (Bottineau, 2002) a proposé que l'invariant du cognème U serait une « projection » (par contraste avec la conjonction en I et la disjonction en A). On peut ici préciser que cet invariant de projection a pour assise biomécanique (inscription corporelle au moins le couple labio-lingual, et éventuellement le trio labio-linguo-laryngal muni de kinimes corrélatifs de degré 2 (protrusion/rétraction/abaissement) : ensemble, les trois gestes étirent la zone de production vocalique en ses extrémités antérieure (larynx), médiane (langue) et postérieure (lèvres), sensibilisant l'ensemble de la trajectoire et ouvrant l'espace d'expérience

phonatoire vocalique dont se tire l'idée d'un mouvement de projection et de visée correspondant au cognème U.

Les voyelles /a/ et /u/ posent un autre problème : du fait du degré d'aperture (configuration) et d'abaissement laryngal, elles génèrent des formants de source orale mais susceptibles de faire résonner l'espace extérieur s'il lui-même est une cavité fermée de structure analogue à celle de la bouche humaine : on peut utiliser « hou ! » et « ah ! » pour tester l'écho d'une caverne ou les harmoniques d'une abbaye (cas de l'Abbaye de Timadeuc en Bretagne : 7 secondes de latence), mais pas la voyelle « hi ! », qui ne renverra aucun effet pertinent (ou très peu). Certaines voyelles sont *écologiques* au sens où leur matrice formantique générée par la configuration orale interne (biologique) peut être réverbérée, prolongée et reprofilée par des matrices harmoniques environnementales livrant des produits acoustiques secondaires, que l'auditeur pourra retraire comme expérience auditive complémentaire par perçaction et, éventuellement, interpréter, sémiotiser en un signifiant procédural, « cognématiser ». Par exemple /a/ induit des effets d'éloignement bien connus (disjonction) du fait de l'aperture, mais aussi des effets d'association intersubjective (*je suis là* au sens de « je suis ici », dans l'espace que nous partageons) imputables à sa résonance écologique (/a/ tend à mettre en valeur la résonance de l'espace partagé s'il est confiné). On retrouve ici le kinakème de De Courtenay : /a/ est une structure gestuelle (kinème, pour nous kinome) qui se solde par une impression acoustique (acousmème), éventuellement munie de propriétés environnementales distinctives (la résonance harmonisée d'une cavité confinée génératrice d'écho). La définition énaïve du phonème, incarnée (gestuelle et perceptuelle), suppose de tenir compte du rapport à l'environnement (la résonance éventuelle), et donc requiert une approche à la fois multimodale (motricités et sensorialités) et distribuée (intersubjective située dans un environnement susceptible de participer à l'interaction), ce qui menait également au kinakoptème (le phonème dans sa dimension visuelle pour la perception de la gestuelle visuo-faciale par l'auditeur-observateur participatif et attentif, l'allocutaire, cf. Baudouin de Courtenay *supra*). Ainsi, la représentation métalinguistique du phonème dans cette perspective incarnée, interactionnelle et située doit chercher à couvrir l'ensemble de la multimodalité distribuée sur la motricité et les sensorialités des interacteurs dans un environnement lui-même variable. La phonosémantique de /a/ et sa capacité à faire émerger le cognème A illustre bien le problème.

En didactique du FLE, certains groupes d'enseignants¹⁸ illustrent le rapport coulisse/compas intrabuccal en le jouant hors de la bouche, dans l'environnement écologique visuel de l'interaction verbale, au moyen des poings et des bras : le poing droit joue la protrusion labiale, le gauche la rétraction dorsale ; la voyelle /i/ est mimée par le contact des deux poings, /y/ par l'avancée d'un cran du poing droit et le recul d'un cran du poing gauche, et /u/ par l'avancée d'un second cran du poing droit associée à un second cran de recul du poing gauche, le tout sur un axe horizontal orienté de la bouche (poing gauche) vers l'horizon (poing droit), cet axe horizontal haut reproduisant à hauteur de tête la coulisse dorso-labiale sous aperture minimale. Pour les autres degrés d'aperture, il suffit de rejouer la coulisse horizontale en parallèle à trois étages inférieurs, en réduisant proportionnellement son extension à chaque niveau, jusqu'à son annulation avec /a/. Par ce jeu un peu complexe, mais rapidement appris en groupe, on apprend à figurer et *énacter* ensemble un modèle en 3D de bouche virtuelle extérieure, figurant la « chimère » commune que l'on cherche à s'approprier individuellement et incorporer corrélativement par *inverbations* interactives, solidaires et fédératrices (Ladmiral & Lipiansky, 1983). Selon Munhall et Jones (1995), le contrôle moteur de la phonation suppose la réalisation

¹⁸ Atelier « Techniques théâtrales et phonétique dans l'apprentissage du FLE » organisé le 24 mai 2013 par Françoise Herbet-Pain et Julie Montagu au 41^e Congrès de l'UPLEGESS (Union des Professeurs de Langues Etrangères des Grandes Ecoles et de l'Enseignement Supérieur), IÉSEG School of Management, campus Lille, 23-25 mai 2013.

de figures spatiales impliquant les organes comme trajecteurs parcourant des chemins (*paths*), travail qui s'inscrit dans le paradigme cognitiviste représentationnel et spectatoriel ; par le jeu, l'activité incarnée permet de réaliser les parcours des trajectoires en première personne actorielle multimotrice pour les intégrer au corps en acte dans sa cohérence psychomotrice. L'intérêt de ce mime ludique est considérable : d'une part, si on coordonne des mouvements hétérogènes, la plasticité neuronale et la simplicité corticale (Berthoz, 2009) mettent en place une commande conjointe plutôt que deux séparées (par un neurone pyramidal) ; si on asservit la coulisse dorso-labiale intrabuccal (incontrôlable pour l'apprenant) à la coulisse bimanuelle des deux poings (parfaitement contrôlée), la première va « obéir » à la seconde et se mettre en place « en pilotage manuel », quitte à s'autonomiser par la suite par le biais d'autres activités intermédiaires. D'autre part, si ce jeu manuel est également coordonné entre apprenants qui le pratiquent en s'observant mutuellement et se synchronisant (comme le ballet prométhéen de Shelley), le réseau de « kinakoptèmes » /i/-/y/-/u/ buccaux, rendu manifeste dans les champs visuels en interaction, se retrouve étendu et augmenté de kinèmes manuels corrélatifs inscrits à la fois dans les corps moteurs et percevants des participants et dans l'environnement où circulent les signaux optiques et acoustiques, conformément aux principes de la mémétique interactionnelle, incarnée et distribuée. Ces pratiques ludiques peuvent paraître fantaisistes, mais elles mobilisent, sans l'expliciter, l'ensemble des principes de l'architecture théorique que l'on expose dans cette étude : elles guident l'appropriation de la phonématique de la langue par un *détour simplexe* ; aussi s'avèrent-elles des « techniques efficaces » (Hagège, 1996 : 92-94).

5. Le système consonantique du français

Venons-en aux consonnes du français. Dans une perspective incarnée, la représentation métalinguistique des consonnes est plus complexe à construire que celles des voyelles : le travail de production des voyelles se limite à obtenir une configuration propre à dériver les formants harmoniques de la fréquence fondamentale ; la production de consonnes consiste à ajouter des matrices gestuelles, les kinomes comme orchestrations de kinimes, lors des phases d'amorçage et de clôture de la syllabe configurée par le « voyellage ». Concrètement, les consonnes réalisent des gestes de fermeture (constrictions) à l'aurore et au crépuscule du geste d'ouverture vocalique (maxillaire, cavitation) (théorie de l'inscription de la syllabe dans l'alternance maxillaire selon MacNeilage, 1998). Cette dynamique de fermeture implique un dialogue et un conflit entre deux partenaires biomécaniques et énergétiques : la fluxation (injection de l'unité protosyllabique comme segment de souffle) et l'articulation (intervention des couples d'articulateurs qui contrarient le passage de l'air). Une consonne pulmonique est un bras de fer opposant le geste pneumatique / pulmonique au geste articuloire, et l'issue du bras de fer, habituellement exprimé par les « modes d'articulation » (occlusif, fricatif, spirantisé), résulte du rapport des forces engagées dans l'interaction : l'occlusion/plosion est la signature d'une supériorité articuloire ; la spirantisation, celle d'une supériorité pulsatoire ; la friction, celle d'un antagonisme intermédiaire équilibré – une constriction suffisante pour induire une perturbation, mais pas de couple stop/plosion. Comme on ne parle ici que du français, on exclut les consonnes non pulmoniques (injectives, éjectives, clicks), qui produisent localement un événement d'air entrant ou sortant en associant l'ouverture/fermeture des plis vocaux, la remontée du larynx, et le plus souvent le voisement (plis vocaux encore) ; mais une phonologie incarnée adaptée à la typologie de langues devra les intégrer. La théorie incarnée de la consonne est donc l'histoire d'un rapport de forces entre fluxation et articulations aux seuils initial et terminal de la configuration vocalique (noyau syllabique), et l'extraction de cognèmes ancrés dans le potentiel phonosémantique des consonnes doit s'appuyer sur les modalités de cet antagonisme.

L'objectif est donc à présent de ramener tout phonème consonantique à un kinome comme ensemble structuré de kinimes : un orchestre de micro-gestes articulatoires structurés par des rapports de forces. Dans une perspective motrice, les points d'articulation sont à penser en termes de couples articulatoires : dorso-vélaire, apico-coronal et bilabial/labiodental (selon le mode d'articulation). Si le couple est formé d'un articulateur actif et d'un passif (apex et couronne), on distingue un organe préhensile palpeur (langue, apex) et un point fixe ciblé (couronne) dont l'atteinte dépend du rapport de force entre le geste et le souffle : si le geste l'emporte sur le souffle, on obtient une fermeture complète (stop) avec pression musculaire de l'articulateur actif sur le passif, suivie d'une libération (relâchement et plosion), donc une occlusive. Si le souffle l'emporte sur le geste, on obtient un rétrécissement limité, non générateur d'une signature acoustique saillante : une spirantisée. Entre les deux, si le rapport de force s'équilibre, on obtient un resserrement étroit générateur d'une signature acoustique saillante : une fricative.

Appelons kinime le schème gestuel de palpation de l'articulateur actif et dotons-le de trois forces (faible, moyenne et forte) exprimées en degrés (1, 2 et 3) correspondant à trois actions formulées en termes de préhension manuelle de la fluxation d'air (presser, serrer et lâcher) : en français, le kinime dorso-vélaire ne se présente qu'au degré 3 (occlusive pressée /k/) alors qu'en allemand il se présente également au degré 2 (fricative serrée /ç/ de *ich*), voire au degré 1 (spirantisée lâchée) dans certains dialectes. Pour le couple apico-coronal en français, le kinime de palpation se présente aux degrés 2 (fricative /s/) et 3 (occlusive /t/) : en termes de striction, /s/ est *inclosif* (non occlusif) alors que /t/ est *closif* (occlusif). En termes de motricité, /s/ et /t/ sont deux consonnes de même couple (point d'articulation) mais de degrés différents (mode d'articulation), ce qui aura une incidence sur leur potentiel phonosémantique, avec notamment le contraste S/T en cognématique (actualisation / interception) et leur mise en composition figée avec l'étymon lexical *st* et la notion gestuelle d'interception ou de fixation (*stare, stand*). Pour le couple labial, la situation est un peu plus complexe : le degré 3 occlusif est bilabial, avec deux articulateurs actifs pinceurs (/p/, closion des deux lèvres avec stop et plosion) alors que le degré 2 fricatif est labio-dental, avec un seul palpeur (/f/, inclusion labio-dentale). En l'absence de bilabiale fricative et de labiodentale occlusive, il faut considérer que la différence bilabial / labiodental est une variable d'ajustement et de coarticulation liée au changement de degré au niveau d'organes flexibles, les lèvres : en éliminant cette variable, on peut poser que /f/ est à /p/ ce que /s/ est à /t/, respectivement les degrés 2 et 3 des kinimes labial (réalisé labiodental au degré 2 et bilabial au degré 3) et apico-dental. Pour les trois couplages (points d'articulation), le premier réalise le degré 3 (/k/) et les deux autres les degrés 2 et 3 (/s/ et /t/, /f/ et /p/), ce qui livre 5 consonnes non voisées.

Sur cette base, les autres consonnes s'obtiennent en ajoutant des kinimes complémentaires. Chaque consonne est munie d'une valence correspondant au nombre et type de kinimes impliqué par le kinome, et certaines consonnes ont une valence plus lourde que d'autres du fait de l'ajout de kinimes optionnels comme le voisement ou la nasalisation. Certains kinimes sont binaires, comme le voisement, qui est activé ou pas. D'autres kinimes varient en force sur une échelle quantitative, le degré (inclusion occlusive vs closion pour une constriction en un seuil donné). La valence est donc hétérogène et complexe : elle regroupe des kinimes dont certains varient quantitativement et d'autres pas. Considérons donc l'ajout de kinimes simples. Le plus important est le voisement produit par les plis vocaux : au plan moteur, pour le couple dorso-vélaire, /g/ est la consonne /k/ (kinime : palpation dorso-vélaire au degré 3) augmentée du kinime de voisement. Pour le couple apico-coronal, /z/ et /d/ sont les consonnes /s/ et /t/ (degrés 2 et 3) augmentées du même kinime de voisement ; pour le couple labial, on trouve /v/ et /b/ comme labiales de degré 2 et 3 voisées. Ainsi, chaque consonne est à pondérer en fonction du nombre de kinimes coordonnés par son kinome et de leurs degrés pour les kinimes concernés :

chaque consonne est munie d'une valence motrice complexe (non exprimable par une valeur simple sur une échelle unique, du fait d'impliquer des motricités hétérogènes). La différence entre phonèmes permet donc de poser des relations d'oppositions porteuses de contrastes auditifs qui seront investis dans les différenciations sémantiques au sein du lexique (*seau* vs *zoo* vs *tôt* vs *dos*), mais ces oppositions diffèrent selon qu'elles jouent entre les degrés de phonèmes de même valence kinomique (*seau* et *tôt* : pour /s/, un seul kinime, palpation apico-dentale de degrés 2 et 3) ou entre phonèmes de valences kinimiques contrastées (*seau* et *dos* : /s/, un seul kinime, degré 2 ; pour /d/, deux kinimes, dont un de degré 3). Si on modélise une structure phonémique en pondérant les relations de distinction entre phonèmes (valeur) en les traitant comme kinomes gesturaux (kinimiques) en fonction des valences kinomiques et degrés kinimiques, on s'achemine vers un modèle plus raffiné qu'un simple jeu de distinctions, en principe apte à rendre compte des émergences phonosémantiques.

D'autres kinimes viennent compléter l'orchestre gestural et générer d'autres consonnes plus lourdes. L'un d'entre eux est l'incurvation linguale, qui scinde le flux aérien en deux canaux : consonnes dites latérales par opposition aux centrales. Le phonème /l/ est muni d'un kinome élaboré : palpation apico-coronale de degré 3 (comme /t/ et /d/), incurvation linguale et voisement. L'incurvation permet à l'air de s'écouler par les côtés du point de pression apico-coronal, ce qui rend la consonne « liquide » et non occlusive. /l/ contient donc le geste de formation de la plosive comme /t/ ou /d/, mais l'incurvation vient neutraliser l'effet par l'ouverture des chenaux latéraux. Ceci montre qu'il n'y a pas correspondance directe entre un geste et un trait : certains traits comme /occlusif/ catégorisent des effets simples résultant de gestes simples comme la palpation de degré 3, mais d'autres traits comme /liquide/ catégorisent des effets complexes résultant d'une coordination plurielle de gestes simples, comme la palpation de degré 3 (closion) et l'incurvation linguale pour la liquide. On rencontre le même problème avec la nasalisation : l'effet nasal est obtenu par un kinime, le geste d'abaissement de la luvette, qui scinde le flux d'air expiré en deux chenaux divergents, oral et nasal ; le chenal oral rencontre la palpation apico-dorsale de degré 3, ce qui devrait donner une occlusive (stop et plosion), mais on obtient une semi-occlusive (stop partiel et plosion amortie du côté oral, du fait de la perte de pression occasionnée par la fuite nasale et sa mise en résonance par le voisement). Ceci montre que /l/ et /n/ partagent des traits moteurs essentiels : /l/ réalise une bifurcation orale étroite du flux expiré par incurvation linguale ; /n/ réalise par abaissement de la luvette une bifurcation oralo-nasale élargie et divergente, ce qui en fait une « latérale verticale » sur l'ensemble du parcours cloisonné par le palais. En outre, ces consonnes bifurquées sont voisées, exploitant diverses résonances qui valorisent leur signature acoustique les mettent en exergue auditive du fait de la percussation. Autrement dit, /l/ et /n/ sont un /d/ augmenté du kinime de bifurcation orale (latérale /l/) ou de celui de bifurcation oralo-nasale divergente (nasale /n/), les deux restant des semi-occlusives (palpation de degré 3 avec closion, mais atténuation du stop et de la plosion par les bifurcations respectives) : au plan moteur, /l/ et /n/ sont un /d/ altéré ou compromis par deux avatars de la bifurcation. En outre, la nasalisation est un geste paradoxal : si la parole s'obtient ordinairement par déviation naso-buccale de l'air expiré et travaillé par phonation et articulation orale, la nasalisation opère momentanément une réviation partielle de l'expiration vers le conduit nasal, scindant la production acoustique en deux canaux et problématisant une phonation bifide. Dans le domaine de la phonosémantique, ceci jouera un rôle particulièrement important dans la modélisation des potentiels cognémiques de ces phonèmes et de leur investissement cognématique dans les systèmes grammaticaux des langues naturelles : par exemple, dans les langues romanes et en arabe classique, l'opposition défini / indéfini se forme sur le contraste L/N (*le/un*), et sa neutralisation par le partitif fait appel à D (*du*). En basque, les deux types de subordination de propositions à verbe conjugué passe par L (-la conjonctif) et N (-n relatif et interrogatif). On ne justifiera pas ici ces rapprochements trop rapides, on se contentera de suggérer que la spécification des liens entre invariants signifiés

et formes signifiantes passe par l'étude détaillée des relations sémaphoniques liant le signifiant en acte (comme phonème pour ce qui est du phonème) et le signifié en acte.

Enfin, on distingue un quatrième degré de palpation : les consonnes dites vibrantes, obtenues par passage en force du flux expiré et mise en vibration de l'organe palpeur (apex ou luvette pour R apical et uvulaire) ; on peut analyser la vibration comme le forçage de la closure et le renversement du couple stop/plosion par la vibration issue du forçage. Dans les langues romanes, ce geste est à la base du potentiel phonosémantique de R, lié à l'inchoation et au déclenchement l'action par passage en force (infinitif et divers suffixes dérivationnels). Il faut souligner que l'élément pertinent est le degré 4 (mode d'articulation) et non le point d'articulation : peu importe que le R soit réalisé uvulaire ou apical dans les dialectes du français (ou de l'italien) du moment que l'effet de forçage est amorcé par le degré 4 de l'antagonisme pression / expiration et la vibration qui en résulte.

Le tableau suivant résume les kinomes des consonnes apicales du français (dont /r/ dans sa version apicale). Chaque consonne est composée (i) d'un kinime obligatoire, la palpation, qui est aussi scalaire (variant par degrés 1, 2, 3 et 4) ; et (ii) de plusieurs kinimes supplémentaires, binaires (non scalaires, selon l'opposition 0/1 : inactifs ou actifs). En termes de composition motrice, chaque consonne du français est munie d'un kinime nucléaire, la palpation scalaire, et de kinimes périphériques binaires optionnels. Pour la palpation au seuil apico-coronal, on obtient la distribution suivante :

	Valence 1 Kinime 1 scalaire	Valence 2 + Kinime 2 binaire	Valence 3 + Kinime 3 binaire
seuil apico-coronal	palpation		
		+ voisement	
			+ incurvation + nasalisation
Degré 4 vibrant		r (apical)	
Degré 3 occlusif	t	d	l n
Degré 2 fricatif	s	z	
(Degré 1 spirantisé)			

Tableau 1. Les kinomes des consonnes apicales du français

Les phonèmes de valence 1 sont ceux qui ne mobilisent que le kinime scalaire de palpation avec : /s/ (au degré 2, fricatif) et /t/ (au degré 3, occlusif) ; le degré 1 (spirantisé) n'est pas réalisé en français, et le degré 4 (vibrant) ne l'est qu'avec voisement (valence 2). Les phonèmes de valence 2 sont ceux qui ajoutent le kinime de voisement (binaire) au kinime de palpation (à l'un de ses degrés), ce qui donne /z/ (degré 2 voisé), /d/ (degré 3 voisé) et /r/ (degré 4 voisé). Les phonèmes de valence 3 sont ceux qui surajoutent un kinime d'incurvation linguale (latéralisation) ou un kinime de nasalisation (l'un et l'autre s'excluant mutuellement en français), ce qui donne /l/ et /n/. Pour le français, on peut simplifier le tableau en éliminant le degré 1, inexploité. Le tableau fait apparaître un système assez limité : on ne peut pas alourdir les consonnes au-delà de 3 kinimes (dans certaines langues, l'incurvation et la nasalisation ne s'excluent pas, ce qui porte la valence à 4), et la diversité des kinimes est limitée qualitativement

(d'autres langues possèdent des kinimes variés, comme l'emphatisation par pharyngalisation, phonétique en anglais avec l'ombre de *call*, mais phonologique en arabe (de 4 à 11 consonnes emphatiques selon les auteurs¹⁹). Certains kinimes, comme l'incurvation et la nasalisation, sont binaires en français, mais scalaires dans d'autres langues. L'analyse du phonème comme kinome en tant que réseau de kinimes, binaires ou scalaires, centraux ou périphériques (obligatoires ou optionnels / complémentaires), et libres ou solidaires (comme la nasalisation et le voisement), et temporalisés (le moment du voisement diffère pour les occlusives en français et en allemand), permet de se faire une idée de comment une phonologie incarnée permet de modéliser le phonème comme coordination de motricités hétérogènes et interdépendantes. Dans le même temps, il ne faut pas imaginer pouvoir comparer les langues en ramenant le phonème à un kinome comme réseau de kinimes : comme on l'a dit, en toute langue, le kinome se réalise dans le contexte de la syllabation, aussi les kinimes sont-ils aussi interdépendants (coarticulation) et en prise avec la syllabation et l'expiration. De ce fait, des kinimes homologues entre deux langues peuvent recevoir la même dénomination, mais ils ne sont pas analogues pour autant.

Conclusion

Sur le principe, on a proposé un modèle phonologique incarné qui repose sur quelques idées directrices : l'unité d'intervention est la syllabe ; la forme d'intervention est le phonème défini comme kinome, ou ensemble de kinimes, défini comme unités d'interventions associables en phonèmes (kinomes) mais également dissociables et recomposables autrement (autosegmentalité incarnée). Cette conception est à la fois motrice et sensorielle et multimodale, et peut supposer d'ajouter au phonème des propriétés visuelles, voire environnementale, comme la résonance. Cela suppose une approche incarnée et distribuée : le phonème comme processus épilinguistique se distribue sur la communication située comme coordination orale-aurale écologique ; et un regard énonciatif sur tous les termes de la relation : même « l'environnement » et le « signal acoustique » sont des énonciats réalisés de divers points de vue et non des entités autonomes relevant d'un monde prédonné. L'appareil conceptuel est complexe, mais il est nécessaire pour parler du logos comme vecteur énonciatif dans la condition humaine. On a donc trouvé dans la littérature existante des ressources considérables pour contribuer à échafauder une phonématique énonciative, mais sur le fond et dans ce cadre conceptuel, en vue de s'outiller pour la linguistique énonciative, le plus gros reste à faire. Il s'agit en outre de réfléchir à la formalisation et l'annotation des principes évoqués, qui présente une réelle difficulté : ramener les traits à des oppositions binaires permet de réaliser des tableaux complexes, mais gérables, pour inventorier les phonèmes au sein des systèmes phonologiques des langues ; étudier les phonèmes comme kinomes en rapport biomécanique et énergétique avec la respiration et syllabation est une autre affaire, qui demandera du temps.

Pour terminer, évoquons un exemple anecdotique assez curieux. Pour évoquer la première personne « moi », on dit *oum* en dioula (Burkina Faso), sans attribution de fonction, et *ego* ou *me* (*mihi* etc.) en latin, avec une consonne vélaire /g/ pour la forme au cas régissant l'accord verbal (nominatif) et une consonne bilabiale nasale /m/ pour les formes aux cas régis par la valence verbale (accusatif, datif) : une forme non casualisée en dioula, deux groupes de formes casualisées en latin. Amusons-nous à observer les rapports forme / sens en faisant abstraction de toute considération diachronique et aréale. En comparant *oum* à *ego* et *me*, on s'aperçoit que

¹⁹ La diversité des inventaires s'explique par le fait que certains auteurs classent les consonnes en fonction de l'impression auditive qualifiée d'emphatique alors que d'autres décomposent plus précisément la coordination motrice correspondant à l'analyse en kinimes gestuels à l'origine de cette impression.

oum regroupe dans une syllabe les traits distribués par *ego* et *me* : *oum* intègre à une syllabe une conformation vocalique fermante /u/ qui se prolonge en coda par une occlusion bilabiale compensée par nasalisation /m/ ; si on « casse » *oum* en deux composantes, on obtient d'un côté *go* (/u/ moins fermé en /o/ du fait de la perte de la coda, cf. *amicus* / *amigo*) (*ego* étant un composé), et de l'autre, en repartant du /m/ coda en situation d'attaque d'une autre syllabe demandant appelant une voyelle, on obtient *me*, *mihi* etc. Ceci n'explique rien ; il s'agit juste d'un modèle métalinguistique heuristique qui suggère la possibilité d'une correspondance entre le dégroupement des kinomes de *oum* et leur regroupements partiels, séparés et contrastifs *ego* et *me*, avec la question de savoir si ce jeu de boulier de Newton a quelque chose à voir avec l'émergence sémantique associée, à savoir la distinction morphosyntaxique entre cas régissant et cas régis. On ne prétend pas que le couple latin *ego* / *oum* serait issu de la dislocation d'une forme antérieure de type *oum* : à notre connaissance, les connaissances actuelles en linguistique diachronique, typologique et aréale ne sont aucunement favorables à une hypothèse de ce genre. Et pourtant, l'analyse phonosémantique fait apparaître entre *oum* et *ego* / *me* un air de famille dans leur capacité à évoquer la première personne sans et avec distinction de cas. Une phonologie incarnée et énaïve aide au moins à prendre conscience que ce genre de question se pose et peut se formuler précisément.

Bibliographie

- ADAMCZEWSKI, Henri & DELMAS, Claude. (1982). *Grammaire Linguistique de l'Anglais*. Paris : Colin.
- ADAMSKI, Dariusz. (1990). Baudouin de Courtenay et la linguistique générale. *Linx* 23, *Traductions de textes peu ou mal connus* : 67-80.
- ALLOTT, Robin. (1973/2012). *The Physical Foundation of Language : Exploration of a Hypothesis*. Hertfordshire, Able publishing / Xlibris Corporation, UK.
- ANGOUJARD, Jean-Pierre. (1997). *Théorie de la syllabe : rythme et qualité*. Paris : CNRS Editions.
- BALL, Martin J. (1993). Further to Articulatory Force and the IPA Revisions. *Journal of the International Phonetic Association* 23(01), June 1993: 39-41.
- BENITEZ BRIBIESCA, Luis. (2001). Memetics: A dangerous idea. *Interciencia: Revista de Ciencia y Tecnología de América*, 26, 1: 29-31.
- BERTHOZ, Alain. (2009). *La simplicité*. Paris : Odile Jacob.
- BERTHOZ, Alain & ANDRIEU, Bernard. (2011). *Le Corps en Acte. Centenaire de Merleau-Ponty*. Nancy : Presses Universitaires de Nancy.
- BICKFORD, Anita C. & FLOYD, Rick. (1981/2006). *Articulatory Phonetics – Tools for Analyzing the World's Languages*. Dallas, Texas : SIL International.
- BIRDWHISTELL, Ray L. (1952). *Introduction to Kinesics : An Annotation System for Analysis of Body Motion and Gesture*. Washington, DC : Department of State, Foreign Service Institute.
- BLOOMFIELD, Leonard. (1933). *Language*. New York : Henry Holt.
- BOERSMA, Paul. (1998). *Functional Phonology – Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- BOHAS, Georges. (2000). *Matrices et étymons, développements de la théorie*. Lausanne: Éditions du Zèbre.

- BOHAS, Georges. (2016). *L'illusion de l'arbitraire du signe*. Rennes : Presses Universitaires de Rennes.
- BOHAS, Georges & DAT Mihai. (2007). *Une théorie de l'organisation du lexique de langues sémitiques : matrices et étymons*. Lyon : ENS Editions.
- BOHAS, Georges & SAGUER, Abderrahim. (2012). *Le son et le sens, Fragments d'un dictionnaire étymologique de l'arabe classique*. Damas : Presses de l'IFLPO.
- BORISSOFF, Constantine. (2012). Basis of articulation or articulatory setting ? IATEFL Pronunciation SIG newsletter. <https://borissoff.files.wordpress.com/2012/01/basis-of-articulation-or-articulatory-setting.pdf>
- BOTTINEAU, Didier. (2002). Les cognèmes de l'anglais : principes théoriques. Dans R. Lowe (dir.), en collaboration avec J. Pattee et R. Tremblay, *Le système des parties du discours, Sémantique et syntaxe, Actes du IXe colloque de l'Association internationale de psychomécanique du langage*. Laval : Les Presses de l'Université Laval, Québec, Canada : 423-437.
- BOTTINEAU, Didier. (2004a). Le problème de la négation et sa solution dans la langue anglaise : le cognème N. Dans C. Delmas, L. Roux (dir.), *La contradiction en anglais, C.I.E.R.E.C. Travaux 116*, Saint-Etienne : Publications de l'Université de Saint-Etienne : 27-53.
- BOTTINEAU, Didier. (2010). Language and enaction. Dans J. Stewart, O. Gapenne & E. Di Paolo (eds), *Enaction: toward a new paradigm for cognitive science*. Cambridge: MIT, 267-306.
- BOTTINEAU, Didier. (2010). La théorie des cognèmes et les langues romanes : l'alternance i/a. La submorphologie grammaticale en espagnol et italien. Dans G. Luquet & W. Nowikow (eds.), *La Recherche En Langues Romanes: théories et applications*. Łódź : Łódź Academy of International Studies & Université Sorbonne Nouvelle – Paris 3: 11-45.
- BOTTINEAU, Didier. (2011). Parole, corporéité, individu et société : l'embodiment entre le représentationnalisme et la cognition incarnée, distribuée, biosémiotique et enactive dans les linguistiques cognitives. Dans J.-B. Guignard (dir.), *Linguistique cognitive : une exploration critique, Intellectica 56, 2011/2* : 187-220.
- BOTTINEAU, Didier. (2012). La parole comme technique cognitive incarnée et sociale . *Linguistique et phénoménologie du langage, La Tribune Internationale des Langues Vivantes*, 52-53, Perros-Guirec : Anagrammes, 44-55.
- BOTTINEAU, Didier. (2013). Pour une approche enactive de la parole dans les langues. Dans G. Louÿs & D. Leeman (dir.), *Langages*, 192, 4/2013 (décembre), *Le vécu corporel dans la pratique d'une langue* : 11-27.
- BOTTINEAU, Didier. (2014). Explorer l'iconicité des signifiants lexicaux et grammaticaux en langue française dans une perspective contrastive (anglais, arabe). Dans L. Nobile (éd.), *Le français moderne*, 82/2, *Formes de l'iconicité en langue française. Vers une linguistique analogique (Tome 2)* : 243-270.
- BOTTINEAU, Didier & GREGOIRE, Michaël (éds.) (2017). *Langage et éaction : corporéité, environnements, expériences, apprentissages, Intellectica 68*. <http://intellectica.org/fr/numeros/langage-et-enaction-corporeite-environnements-experiences-apprentissages>
- BOTTINEAU, Didier. (2018). Incarnation langagière et grammaire des langues naturelles: vers la fin d'un clivage. Dans J. Dokic & D. Perrin (éd.), *La cognition incarnée, Recherches sur la philosophie et le langage*, 33. Paris : Librairie philosophique, Vrin : 251-294.

- BOTTINEAU, Didier. (2021). La théorie des matrices et étymons (TME) de Georges Bohas comme morphophonosémantique lexicale générative, opérative et incarnée. Dans D. Leeman (éd.), *La submorphologie motivée de Georges Bohas : vers un nouveau paradigme en sciences du langage – Hommages à Georges Bohas*. Paris : Champion : 99-132.
- BOTTINEAU, Didier. (2023 à paraître). De la phonation interactive incarnée aux voix des langues naturelles. Dans M. Embarki & C. Pillot-Loiseau, *Voix, langues, remédiations*, revue *Langages*.
- BOTTINEAU, Didier & GREGOIRE, Michael (éds.). (2017). *Langage et énonciation : corporéité, environnements, expériences, apprentissages*, *Intellectica* 68.
- BOTTINEAU, Didier & POIRIER, Marine. (2022). Les submorphémies : constructions et discussions - Éléments de constitution poétique des submorphémies linguistiques. Dans F. Bravo (éd.), *Poétique du signifiant. Approche submorphémique de l'espagnol*. Rennes : Presses Universitaires de Rennes : 39-57.
- BROWMAN, Catherine P. & GOLDSTEIN, Louis. (1986). Towards an articulatory phonology. Dans C. Ewen & J. Anderson (eds.), *Phonology Yearbook 3*. Cambridge : Cambridge University Press: 219-252.
- BROWMAN, Catherine P. & GOLDSTEIN, Louis. (1990). Tiers in articulatory phonology, with some implications for casual speech. *Papers in Laboratory Phonology I, Between the Grammar and the Physics of Speech*. Cambridge : Cambridge University Press: 341-376.
- BROWMAN, Catherine P. & GOLDSTEIN, Louis. (1995). Dynamics and articulatory phonology. Dans R. F. Port & T. Van Gelder (eds), *Mind as Motion*. Cambridge : MIT Press: 174-193.
- BROWMAN, Catherine P. & GOLDSTEIN, Louis. (1995). Gestural syllable position Effects in American English. Dans F. Bell-Berti & L. J. Raphael (eds.), *Producing Speech : Contemporary Issues – for Katherine Safford Harris, AIP Series in Modern Acoustics and Signal Processing*. New York : American Institute of Physics: 19-33.
- BROWMAN, Catherine P. & GOLDSTEIN, Louis. (2000). Competing constraints on intergestural coordination and self-organization of phonological structures. *Les Cahiers de l'ICP, Bulletin de la Communication Parlée* 5: 25-34.
- CHEN, Tsuhan. (2001). Audiovisual speech processing. *IEEE Signal Processing Magazine* 18: 9–21.
- CHEN, Tsuhan & RAO, Ram, R. (1998). Audio-visual integration in multi-modal communication. *Proceedings of the IEEE* 86: 837–852.
- COLEMAN, John. (2002). Phonetic representations in the mental lexicon. Dans J. Durand & B. Laks (eds), *Phonetics, Phonology, and Cognition*. Oxford : Oxford University Press: 96-130.
- CORBALLIS, Michael. (2003). From hand to mouth: the gestural origins of language. Dans M. H. Christiansen & S. Kirby, *Language evolution*. Oxford : Oxford University Press: 201-218.
- CORNULIER, Benoît de. (1982). *Théorie du vers*. Paris : Le Seuil.
- COURTENAY, Baudouin de. (1972). The Difference between Phonetics and Psychophonetics. Dans E. Stankiewicz (ed. & transl.), *A Baudouin de Courtenay Anthology*. Bloomington : Indiana University Press: 278-283.
- COURTENAY, Baudouin de. (1972). An Outline of the History of the Polish Language. Dans E. Stankiewicz (ed. & transl.), *A Baudouin de Courtenay Anthology*. Bloomington : Indiana University Press: 323-355.
- COWLEY, Stephen J. (2011). *Distributed Language*. Amsterdam : Benjamins.

- COWLEY, Stephen J. (2012). Mimesis and language: a distributed view. *Interaction Studies* 13: 17-40.
- DAWKINS, Richard. (1976). *The Selfish Gene*. Oxford : Oxford University Press.
- DEACON, Terrence. (1999). The Trouble with Memes (and what to do about it). *The Semiotic Review of Books* 10(3): <http://projects.chass.utoronto.ca/semiotics/srb/10-3edit.html>.
- DELATTRE, Pierre. (1953), Les modes phonétiques du français. *The French Review*, 27/1 : 59-63.
- DELATTRE, Pierre. (1965). *Comparing the Prosodic Features of English, German, Spanish and French*. Heidelberg : Julius Groos Verlag.
- DELATTRE, Pierre. (1969). *The general phonetic characteristics of languages*, final report, Technical report, US Department of Health, Education and Welfare. ERIC: ED051725.
- DELL, François, HIRST, Daniel & VERGNAUD, Jean-Roger. (1984). *Forme sonore du langage – Structure des représentations en phonologie*. Paris : Hermann.
- DERVEAUX, Michel. (2003). *La poétique sémiophone – Etude sur la sonorité du langage dans la modernité littéraire et musicale*, Limoges : L’Harmattan.
- William. (1979). Phonology as human behavior. Dans D. Aaronson et R. Rieber (eds), *Psycholinguistic research: implications and applications*. New York : Hillsdale, Lawrence Erlbaum Assoc.: 161-182.
- ESLING, John H. (2005), There are no back vowels: The laryngeal articulator model. *Canadian Journal of Linguistics*, 50: 13-44.
- ESLING, John H. (2010), Phonetic notation. Dans W. J. Hardcastle, J. Laver & F. E. Gibbon (eds), *The Handbook of Phonetic Sciences* (2nd ed.). Chichester (UK) : Wiley-Blackwell: 678-702.
- ESLING, John H., MOISIK, Scott R., BENNER, Allison & CREVIER-BUCHMAN, Lise. (2019). *Voice Quality, The Laryngeal Articulatory Model*. Cambridge : Cambridge University Press.
- FIRTH, John, Rupert. (1930). *The tongues of men, and speech*. London : Oxford University Press.
- FISHER, Cletus. G. (1968). Confusions among visually perceived consonants. *Journal of Speech and Hearing Research*, 11(4): 796–804.
- GAFOS, Adamantios I. & GOLDSTEIN, Louis M. (2012). Articulatory representation and organization. Dans A. C. Cohn, C. Fougerson & M. K. Huffman (Eds.), *Oxford handbook of laboratory phonology*. Oxford : Oxford University Press: 220–231.
- GALLINA, Jean-Marie. (2011). Les représentations : un enjeu pour les sciences cognitives. Dans N. Bault, V. Chambon, N. Maïonchi-Pino, F.-X. Pénicaud, B. Putois & J.-M. Roy, *Peut-on se passer de représentations en sciences cognitives ?* Préface d’O. Koenig. Bruxelles : De Boeck: 13-31.
- GIBSON, James J. (1979). *The Ecological Approach to Visual Perception*. London : Routledge.
- GICK, Bryan. (2015), « Towards an Embodied and Embedded Phonetics », *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*, Paper Number 1046, August 2015, The Scottish Consortium for ICPhS 2015 (ed.), Glasgow, UK: the University of Glasgow.
- GICK, Bryan, WILSON, Ian, KOCH, Karsten & COOK, Clare. (2004). Language-specific articulatory settings : evidence from inter-utterance rest position. *Phonetica*, 61: 220-233.

- GICK, Bryan, WILSON, Ian & DERRICK, Donald. (2013). *Articulatory Phonetics*. Chichester (UK) : Wiley-Blackwell.
- GIL FERNÁNDEZ, Juana. (2006). Implicaciones fonológicas de la base de articulación. *Filología y lingüística : estudios ofrecidos a Antonio Quilis, Volumen 1*. Valladolid : CSIC-UNED: 219-252.
- GILI GAYA, Samuel. (1956), « Base de articulación », *Elementos de fonética general* (5e ed.), Madrid : Editorial Gredos.
- GILLESPIE-LYNCH, Kristen, GREENFIELD, Patricia M., LYN, Heidi & SAVAGE-RUMBAUGH, Sue. (2014). Gestural and symbolic development among apes and humans : support for a multimodal theory of language evolution. *Frontiers in Psychology* 5: 1228, <https://www.frontiersin.org/articles/10.3389/fpsyg.2014.01228/full>
- GLEASON, Henry A. (1955). *An Introduction to Descriptive Linguistics*. London : Holt, Rinehart & Winston.
- GOLDINGER, Stephen. (1996). Words and voices : Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology : Learning, Memory, and Cognition* 22: 1166-1182.
- GOLDSMITH, John. (1990), *Autosegmental and Metrical Phonology*. Oxford : Basil Blackwell.
- GOLDSMITH, John, RIGGLE, Jason & YU, Alan C. L. (eds) (2014). *The Handbook of Phonological Theory, 2nd edition*. Chichester (UK) : Wiley Blackwell.
- GOLDSTEIN, Louis P. & BROWMAN, Catherine. (1992). Articulatory Phonology : An Overview. *Phonetica* 49(3-4): 155-180.
- GOLDSTEIN, Louis. P. & FOWLER, Carol A. (2003). Articulatory phonology: a phonology for public language use. Dans A. S. Meyer & N. O. Schiller (eds.), *Phonetics and Phonology in Language Comprehension and Production : Differences and Similarities*. Berlin: Mouton de Gruyter.
- GORDON, Matthew & LADEFOGED, Peter. (2001). Phonation types: A cross-linguistic overview. *Journal of Phonetics*, 29: 383-406.
- GRAY, John (2008). John Gray on secular fundamentalists. *The Guardian* (London), 15 mars 2008.
- GREGOIRE, Michaël, BARNABÉ Aurélie, BOTTINEAU, Didier & MAÏONCHI-PINO, Norbert (éds.). (2017). *Langage et énonciation : problématiques, approches linguistiques et interdisciplinaires // Énonciation, émergence du langage, production du sens, Signifianges (Signifying)*, [S.l.], v. 1, n. 1, i-iv, nov. 2017.
<http://revues.clermont-universite.fr/index.php/Signifianges/issue/view/18>
- GUIRAUD, Pierre. (1967). *Structures étymologiques du lexique français*. Paris : Payot (1967), Larousse (1986).
- HONIKMAN, Beatrice. (1964). Articulatory setting. Dans D. Abercrombie, D. B. Fry, P. A. D. MacCarthy, N. C. Scott & J. L. M. Trim (eds), *In Honour of Daniel Jones*. London : Longman: 73-84.
- HUMBOLDT, WILHELM VON. (1836/1974). La différence de construction du langage dans l'humanité et l'influence qu'elle exerce sur le développement spirituel de l'espèce humaine ou Introduction à l'œuvre sur le kavi. *Introduction à l'œuvre sur le kavi et autres essais* (Traduction de l'allemand par P. Caussat). Paris : Seuil : 143-419.

- JAKOBSON, Roman. (1938). Sur la théorie des affinités phonologiques entre les langues. *Actes du Quatrième Congrès International des Linguistes tenu à Copenhague du 27 août au 1^{er} septembre 1936*, Copenhague : Einar Munksgaard, 38-48. Reproduit dans Troubetzkoy N. S. (1949 / 1957), *Principes de phonologie*, traduits par J. Cantineau. Paris : Klincksieck : 351-365.
- JAKOBSON, Roman. (1976). *Six leçons sur le son et le sens*. Paris : Editions de Minuit.
- JAKOBSON, Roman & WAUGH, Linda. (1979/2002). *The Sound Shape of Language*. Berlin : Mouton De Gruyter.
- JENNER Bryan. (2001). Articulatory setting: genealogies of an idea. *Historiographia Linguistica* 28: 121-141.
- JOUSSE, Marcel. (1936/1974). *L'anthropologie du geste*. Paris : Gallimard, TEL.
- JOUXTEL, Pascal. (2005). *Comment les systèmes pondent: Une introduction à la mémétique*. Paris : Editions le Pommier.
- JÜNGER, Ernst. ([1934]1979). *Éloge des voyelles* (traduction : J.-L. Evard). Monaco : Éditions du Rocher.
- KINGSTON, John. (2019). The interface between phonetics and phonology. Dans W. F. Katz & P. F. Assmann (eds.), *The Routledge Handbook of Phonetics*. London & New York : Routledge: 359-400.
- KNIGHT, Rachael-Anne & SETTER, Jane (eds). (2021). *The Cambridge Handbook of Phonetics*. Cambridge: Cambridge University Press.
- KOMEL, Mirt. (ed.) (2019). *The Language of Touch: Philosophical Examinations in Linguistics and Haptic Studies*. New York: Bloomsbury Publishing.
- KRAKOW, Rena A. (1989). *The articulatory organization of syllables : a kinematic analysis of labial and velar gestures*. PhD dissertation, Yale University.
- KULL, Kalevi. (2000). Copy versus translate, meme versus sign: development of biological textuality. *European Journal for Semiotic Studies*, 12(1): 101-120.
- LADD, Dwight, Robert. (2014). Phonetics in Phonology. Dans J. Goldsmith, J. Riggle & A. C. L. Yu (eds), *The Handbook of Phonological Theory, 2nd edition*. Chichester (UK) : Wiley Blackwell: 348-373.
- LADEFOGED, Peter. (1975). *A course in Phonetics*. New York: Harcourt Brace Jovanovich.
- LADMIRAL, Jean-René & LIPIANSKY, Edmond-Marc. (1989). *La Communication interculturelle*. Paris : Armand Colin.
- LAFONT, Robert. (2004). *L'Être de Langage – Pour une anthropologie linguistique*. Limoges : Lambert-Lucas.
- LAFONT, Robert. (2007). *Il Y a Quelqu'un – la Parole et le Corps*. Paris : L'Harmattan.
- LAKS, Bernard. (1997). *Phonologie accentuelle – Métrique, autosegmentalité et constituance*. Paris : CNRS Editions.
- LAKS, Bernard & PLENAT, Marc. (1993). L'objet de la phonologie : les objets phonologiques. Dans B. Laks & M. Plénat (éds.), *De Natura Sonorum – Essais de phonologie*. Vincennes : Presses Universitaires de Vincennes : 5-23.
- LAMBERT-DRACHE, Marilyn. (1997). *Sur le bout de la langue – introduction au phonétisme du français*. Toronto : Canadian Scholars' Press Inc.

- LAPAIRE, Jean-Rémi & ROTGÉ, WILFRID. (1993). *Linguistique et Grammaire de l'Anglais*. Toulouse : Presses universitaires du Mirail.
- LAVER, John. (1978). The concept of articulatory settings: an historical survey. *Historiographia Linguistica*, 5: 1-14.
- LAVER, John. (1994). *Principles of phonetics*. Cambridge : Cambridge University Press.
- LE HUCHE, François & ALLALI, André (1991). *La voix (tome 1, 3^e éd.)*, Anatomie et physiologie des organes de la voix et de la parole. Paris : Masson.
- LEEMAN, Danielle. (éd.). 2021. *La submorphologie motivée de Georges Bohas : vers un nouveau paradigme en sciences du langage – Hommages à Georges Bohas*. Paris : Champion.
- LE TALLEC-LLORET, Gabrielle. (2017). *Linguistique du signifiant – Diachronie et synchronie de l'espagnol. Préface de Didier Bottineau*. Limoges : Lambert-Lucas.
- LIBERMAN, Alvin M. & MATTINGLY, Ignatius. G. (1985). The motor theory of speech perception revised. *Cognition*, 21: 1-36.
- MACCHI, Yves. (2018). Chronophonétique (I). Esquisse d'embryologie du mot. Dans C. Fortineau-Brémond & E. Blestel (éds.), *Le signifiant sens dessus-dessous. Submorphémie et chrono-analyse en linguistique hispanique*. Limoges : Lambert-Lucas: 169-200.
- MACNEILAGE, Peter F. (1998). The Frame/Content Theory of Evolution of Speech Production. *Behavioral and Brain Sciences* 21: 499-511.
- MACNEILAGE, Peter F. (2015). Central Tenets of the Frame/Content Theory of Evolution and Acquisition of Speech Production. Dans M. A. Redford (ed), *The Handbook of Speech Production*. Chichester (UK) : Wiley Blackwell: 353-378.
- MÉNARD, Lucie. (2010). Multimodal Speech Production. Dans M. A. Redford (ed.), *The Handbook of Speech Production*. Chichester (UK) : Wiley Blackwell: 200-221.
- MENARY, Richard. (2010). Introduction to the special issue on 4E cognition. *Phenomenology and the Cognitive Sciences*, 2010, 9: 459–463.
- MUNHALL, Kevin G. & JONES, Jeffery. A. (1995). The Spatial Control of Speech Movements. Dans F. Bell-Berti & L. J. Raphael (eds.), *Producing Speech : Contemporary Issues – for Katherine Safford Harris, AIP Series in Modern Acoustics and Signal Processing*. New York : American Institute of Physics: 521-537.
- OHALA, John. J. (1983). The Origin of Sound Patterns in Vocal Tract Constraints. Dans P. F. MacNeilage (ed), *The Production of Speech*. Berlin: Springer Verlag: 189-216.
- OHALA, John. J. (2009). Languages' sound inventories: the devil in the details. Dans F. Pellegrino, E. Marsico, I. Chitoran & C. Coupé (eds), *Approaches to Phonological Complexity*. Berlin / New York : Mouton de Gruyter: 47-58.
- NOË, Alva. (2004). *Action in Perception*. Cambridge, MA : MIT Press.
- PARRET, Herman. (2014). *Le Son et l'Oreille, Six essais sur les manuscrits saussuriens de Harvard*. Limoges : Lambert-Lucas.
- PHILPS, Dennis (2000). Le sens retrouvé? De la nomination de certaines des parties du corps: le témoignage des marqueurs sub-lexicaux de l'anglais e <CN->. *Anglophonia* 8: 207-232.
- PHILPS, Dennis (2006). From mouth to hand. Dans A. Cangelosi, A. Smith & K. Smith (eds.), *The Evolution of Language*. London/Singapore : World Scientific Publishing: 247-254.

- PIERREHUMBERT, Janet. (1990). Phonological and phonetic representation. *Journal of Phonetics* Volume 18, Issue 3, July 1990: 375-394.
- PORT, Robert & LEARY, Adam. (2005). Against Formal Phonology. *Language* Vol. 81, No. 4 (Dec., 2005): 927-964.
- PRIEUR, Jacques, BARBU, Stéphanie, BLOIS-HEULIN, Catherine & LEMASSON, Alban. (2020). The origins of gestures and language: history, current advances and proposed theories. *Biological Reviews* 95: 531-554.
- RIZZOLATTI, Giacomo & ARBIB, Michael. (1998). Language within our grasp. *Trends in neurosciences*, Volume 21, Issue 5: 188-194.
- ROACH, Peter. (2016). Articulatory settings. Peter Roach's Home Page, <https://www.peterroach.net/blog/march-09th-2016>
- ROCCHETTI, Álvaro. (1980). *Sens et forme en linguistique italienne : études psychosystématiques dans la perspective romane*. Thèse de Doctorat d'État, Paris : Paris 3 Sorbonne Nouvelle.
- ROCCHETTI, Álvaro. (1991). La langue, une gestuelle articulatoire perfectionnée ? *Geste et image* 8-9, Paris : Éditions du CNRS : 63-78.
- ROCCHETTI, Álvaro. (2019). Arbitraire ou motivation du signe ? Le geste articulatoire est-il le fondement de l'expression linguistique ? *Studii de Știință și Cultură XV/2019*, Arad : Editura Universității Vasile Goldiș : 9-18.
- SADOWSKI Piotr. (2001). The sound as an echo to the sense: the iconicity of English gl- words. Dans Fischer O. & Nänny M. (eds), *The Motivated Sign, Iconicity in Language and Literature* 2. Amsterdam: Benjamins: 69-88.
- SAFFI Sophie & CULOMA Virginie. (2021). Réflexions sur le lien entre sens et forme au niveau submorphémique. Dans C. Fortineau-Brémond & S. Pagès (eds.), *Le morphème en question : exemples multilingues d'analyse submorphologique (anglais, arabe, français, guarani, italien)*. Aix en Provence : Presses Universitaires de Provence : 21-37.
- SALTZMAN, Elliot & KELSO, J. A. Scott. (1987). Skilled actions: A task dynamic approach. *Psychological Review* 94 (1): 84-106.
- SALTZMAN, Elliot. (1995). Dynamics and Coordinate Systems in Skilled Sensorimotor Activity. Dans R. F. Port & T. Van Gelder (eds), *Mind as Motion*. Cambridge : Cambridge, MIT Press: 148-173.
- SCHENKEL, Roger. (1981). *Introduction à La Syllabatique ou gymnastique phonétique basée sur les structures syllabiques du langage*. Éditeur : Roger Schenkel.
- SCHEER, Tobias. (2015). *Précis De Structure Syllabique*. Lyon : ENS Editions.
- SCHOURUP, Lawrence. (1981). The Basis of Articulation. *Ohio State University Working Papers in Linguistics* 25, Columbus, Ohio: 1-13.
- SCHWARTZ, Jean-Luc, BASIRAT, Anahita, MENARD, Lucie & SATO, Marc. (2012). The Perception-for-Action-Control Theory (PACT) : A perceptuo-motor theory of speech perception. *Journal of Neurolinguistics*, Volume 25, Issue 5, September 2012: 336-354.
- SEIKEL, Anthony, DRUMRIGHT, David & HUDOCK, Daniel. (2019). *Anatomy & Physiology for Speech, Language, and Hearing (6th ed.)*. San Diego : Plural Publishing Inc.

- STEFFENSEN, Sune V. & COWLEY, Stephen. J. (2010). Signifying bodies and health: a non-local aftermath. Dans S. J. Cowley, J. C. Major, S. V. Steffensen & A. Dinis (eds.), *Signifying Bodies : Biosemiosis, Interaction and Health*. Braga : Portuguese Catholic University: 331-355.
- STEINER, Pierre. (2011). Enaction, pragmatisme et minimalisme représentationnel. Dans N. Bault, V. Chambon, N. Maïonchi-Pino, F.-X. Pénicaud, B. Putois & J.-M. Roy, *Peut-on se passer de représentations en sciences cognitives ?* Préface d'O. Koenig. Bruxelles : De Boeck: 191-210.
- STETSON, Raymond, Herbert. (1951). *Motor Phonetics – a study of speech movements in action*. Amsterdam : North Holland Publishing Company.
- STEWART, John. (1995). *Language as Articulate Contact – Towards a post-semiotic Philosophy of Communication*. New York, State University of New York Press.
- STONE, Maureen & LUNDBERG Andrew. (1996). Three-dimensional tongue surface shapes of English consonants and vowels. *Journal of The Acoustical Society of America* 99(6): 3728-3737.
- STONE, Maureen, SHAWKER, Thomas. H., TALBOT, Thomas. L. & RICH, Alan. H. (1988). Cross-sectional tongue shape during the production of vowels. *The Journal of the Acoustical Society of America* 83(4): 1586-1596.
- STORY, Brad. H. (2010). Mechanism of voice production. Dans M. A. Redford (ed.), *The Handbook of Speech Production*. Chichester (UK) : Wiley Blackwell: 34-58.
- TABAIN, Marija. (2019). The phonetic properties of consonants. Dans W. F. Katz & P. F. Assmann (eds), *The Routledge Handbook of Phonetics*. London & New York : Routledge: 264-288.
- TOBIN, Yishai. (1997). *Phonology as Human Behavior – Theoretical Implications and Clinical Applications*. Durham and London : Duke University Press.
- TOMATIS, Alfred. (1991). *Nous sommes tous nés polyglottes*. Paris : Fixot.
- TOURNIER, Jean. (1985). *Introduction descriptive à la lexicogénétique de l'anglais contemporain*. Paris-Genève : Champion-Slatkine.
- TOUSSAINT, Maurice. (1981). *Contre l'arbitraire du signe* [préface de M. Arrivé]. Paris : Didier Érudition.
- TOUSSAINT, Maurice. (2003). Analogiques. *Cahiers de linguistique analogique*, 1 : 331-350.
- TRABANT, Jürgen. (2004). *Vico's New Science of Ancient Signs – A study of sematology*, Translated from the German by Sean Ward. London & New York : Routledge.
- UEXKÜLL, Jakob. von (1934/2010). *A foray into the world of animals and humans ; with a theory of meaning*. Minneapolis : University of Minnesota Press.
- VAISSIÈRE, Jacqueline. (2007). Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic, and perceptual properties of sounds across languages. Dans M. J. Solé, P. S. Beddor & M. Ohala (eds.), *Experimental approaches to phonology*. Oxford : Oxford University Press: 54-71.
- VAN GINNEKEN, Jacobus. (1907). *Principes de linguistique psychologique : Essai de synthèse*. Paris : Marcel Rivière.
- VAN GINNEKEN Jacobus. (1933),. La biologie de la base d'articulation. Dans H. Delacroix (éd.), *Psychologie du Langage, Journal de psychologie normale et pathologique* 30. Paris : Félix Alcan : 266–320.

VARELA, Francisco, THOMPSON, Evan & ROSCH, Eleanor. (1991). *The embodied mind. Cognitive Science and Human Experience*. Cambridge : The MIT Press. (Expanded edition in 2015).

VICO, Giambattista. (1744/2001). *La science nouvelle*. Traduit et présenté par Alain Pons. Paris : Fayard.

VIËTOR, Wilhelm. (1887). *Elemente der Phonetik und Orthoepie des Deutschen, Englischen und Französischen*. Heilbronn: Gebr. Henninger.

VOCKERADT, Werner. (1925). *Die Deutsche und die Englische Artikulationsbasis*. Doctoral Dissertation, Greifswald.

WADSWORTH, Stafford. (1979). Articulatory Settings: The mechanics of variety. *Revue de Phonétique Appliquée* 52: 255–278.