

Didier Bottineau

La linguistique enactive pour des avatars signeurs



Introduction

- Ma position dans cette journée:
 - linguiste des langues vocales
 - intéressé par la diversité des langues dans la typologie des langues naturelles
 - langues indo-européennes romanes, germaniques, celtiques, slaves
 - langues plus « exotiques »: basque, japonais, inuktitut, guarani
- Mon objectif: tenter d'apporter un regard sur la notion d'avatar signeur à partir du travail que j'ai réalisé sur les langues naturelles
- Plan:
 - 1) comment j'en viens à une linguistique énaïve à partir de la grammaire des langues naturelles (cela ne va pas de soi)
 - 2) quelques propriétés qui fondent une « linguistique énaïve »
 - 3) quelques propriétés phénoménologiques des langues vocales
 - 4) vers les propriétés phénoménologiques des langues signées ou « manuo-faciales »
 - 5) conséquences pour les avatars signeurs envisagés dans le contexte d'interfaces

Comment un grammairien des formes abstraites en vient-il à une linguistique énaactive?

- La « submorphémie grammaticale »
- Mon cadre théorique de référence comme étudiant: la psychomécanique du langage de Gustave Guillaume
 - idée centrale: les systèmes grammaticaux organisent des formes linguistiques qui correspondent à des étapes mentales dans des « actes de représentation »
 - élément de méthodologie: on s'appuie sur des analogies de niveau « submorphémique » (au-dessous de ce qui est considéré comme morphème)
 - Ex: march-A / marche-R-A, march-AIT / marche-R-AIT
- En appliquant cette méthode à la grammaire de l'anglais de manière généralisée, j'ai mis en évidence l'existence d'un répertoire d'éléments formateurs de niveau phonémique qui entrent dans la composition de marqueurs grammaticaux
 - * Ex: WH « rechercher une entité non disponible en mémoire de travail » vs TH « récupérer une entité disponible en mémoire de travail »
 - d'où des oppositions simples: WHEN = « moment amémoriel », THEN = « moment mémoriel », cf WHERE / THERE etc.
 - j'ai nommé « cognèmes » ces unités submorphémiques de base munies d'un invariant schématique ou procédural que l'on peut formuler en termes d'instructions.
 - il est possible d'expliquer entièrement la valeur opératoire des morphèmes grammaticaux de l'anglais en exploitant les cognèmes, mais le phénomène est très général dans les langues vocales du monde, à différents degrés de « finesse » ou de « résolution » (cas extrême: le guarani)

Comment un grammairien des formes abstraites en vient-il à une linguistique énactive? (2)

- La « submorphémie lexicale »
- Caractériser les « cognèmes » dans le domaine grammatical supposait de les distinguer des « submorphèmes lexicaux » (« idéophones », « phonesthèmes ») selon les auteurs
= matrices de consonnes qui évoquent un geste pertinent pour la manipulation de l'objet
* Ex1: SW anglais = notion de balancement commune à sweep « balayer », switch « commuter », swing « balancer », sway « vaciller », swoon « se pâmer » etc.
* Ex2: SP anglais = notion de centrifugation = rotation et/ou éjection par force centrifuge
Remarque à quel point SP est compliqué à expliquer en langue vocale mais facile à réaliser par un geste, comme: lancer une pierre à l'aide d'une fronde
- Dans les deux submorphémies (grammaticale et lexicale) il existe des relations fortes entre
 - l'expérience de l'articulation phonatoire et de l'audition du signifiant vocal
 - l'expérience sensorimotrice de l'objet ou de l'action qu'il s'agit d'évoquer au niveau du mot lexical

→ Ce travail m'a amené à proposer que les formes langagières vocales n'encodent pas des représentations séparées; au contraire, c'est l'expérience de la vocalisation (et de l'audition) qui suscite la prise de conscience de la notion recherchée

→ la submorphémie m'a amené accidentellement au paradigme de l'énaction: l'expérience incarnée du signifiant vocal donne accès à des notions abstraites via des ressentis corporels

Propriétés d'une linguistique éactive (L.E.)

- Une L.E. étudie le langage humain en tant qu'activité éthologique incarnée: *energeia* (Aristote, Humboldt, Coseriu), *linguaging* (Maturana)
 - Sur la question de la « production du sens linguistique »:
 - une L.E. étudie ce que cette éthologie peut faire « éacter » = la manière dont la pratique de la parole fait advenir « l'expérience » en tant que « micro-monde vécu »
 - *Ex. extrême: dans une langue des sorciers du pays Dogon, les modes verbaux indiquent quel esprit est à l'origine de votre prise de parole (tel esprit pour la connaissance, tel autre pour le désir, tel autre pour la peur ou la colère, etc.);
 - ces modes vous obligent à vous éacter à la manière d'un masque visible posé sur un esprit invisible: le corps parlant est le porte-parole d'un esprit autre que le sien, mais l'acte de parole le fait s'identifier à l'esprit-source supposé
 - phénomènes analogues en langues gestuelles, beaucoup plus systématiques: signer l'idée de « tristesse » en jouant la tristesse par une mimique du visage
 - ➔ dire le mot vocal « triste » ne vous fait pas vivre la tristesse
- ALORS QUE
- ➔ signer la « tristesse » par un signe complexe qui relie un geste manuel, une mimique faciale et un regard de partage avec autrui, va inévitablement vous donner un ressenti du vécu de la tristesse

Propriétés d'une linguistique éactive (L.E.) (2)

- Une L.E. ne considère pas le sens linguistique comme représentation abstraite encodée par des formes symboliques.
→ Au contraire, une L.E. se concentre sur la manière dont l'expérience de l'acte signifiant contribue à faire émerger le sens en tant qu'acte mental, psychologique, et expérience vécue.
- Une phénoménologie de l'expérience de l'acte signifiant est indispensable pour décrire ce que l'on vit exactement lorsque l'on dit ou signe quelque chose dans un système linguistique donné, vocal ou « signal ».
→ à quoi faut-il opposer une « langue vocale »? Une langue « manuo-faciale »?
→ on recherche ici des termes qui ne décrivent pas l'interprétation abstraite de l'évènement, le « signe », mais qui s'intéressent à la dynamique incarnée de l'acte expressif dans sa dimension motrice, sensorielle et interactive, donc visuelle.
- Une construction phénoménologique précise de l'expérience des signifiants vocaux et manuo-faciaux a des conséquences très importantes pour le développement d'avatars signeurs: elle vise à spécifier les paramètres d'expérience vécue que l'on aimerait faire entrer dans l'interface et à sélectionner ceux qui sont techniquement gérables à un moment donné de savoir-faire technologique.

Propriétés du sens linguistique vocal

- Le psychologue Pierre Janet a proposé que
 - la parole humaine survient en cas de blocage dans une situation pratique ou de problème à résoudre
 - formellement, la parole vocale réside dans le minimum gestuel possible: faire sens au moyen de la phonation seule (en ne comptant pas sur le reste du corps)

NB – ceci est une variable typologique et culturelle: le danois et le finnois sont connus pour être des langues vocales à très faible accompagnement gestuel, par opposition à l'italien
- Conséquence:
 - locomotion: on se déplace avec les jambes
 - manipulation: on intervient sur les objets avec les mains
 - interaction: on parle à autrui avec la bouche; chez l'adulte, la bouche n'est pas l'organe de navigation dans l'espace ni l'organe d'intervention sur les objets.

→ La parole vocale se réalise par un système centré sur la bouche, lequel est quasiment exclu de l'engagement dans le monde, de la « palpation » au sens de Merleau-Ponty.
- D'où l'impression de séparation du sens linguistique de source vocale:
 - la parole vocale se réalise par un comportement « écologiquement neutre », ou presque
 - elle produit un « sens linguistique » vécu comme une représentation séparée du monde, une idée qui réfère à une situation, la commente, ou éventuellement l'évoque à partir de rien

Propriétés du sens linguistique manuo-facial (langues signées)

- Le signeur produit le sens en coordonnant une série de comportements interdépendants, solidarisés par des relations:
 - le geste manuel, avec sa variation (vitesse, extension, force)
 - la position du geste manuel par rapport au corps du signeur et à celui du destinataire
 - l'orientation du regard du signeur, tourné vers le destinataire, ou vers son propre geste, ou vers un autre espace (sol, ciel, plafond, horizon...)
 - la capacité du signeur à capter le regard et l'attention du destinataire
 - * langues vocales: il n'est pas nécessaire d'écouter ni de regarder pour entendre
 - * langues manuo-faciales: il est indispensable de regarder pour comprendre et partager
- Conséquences: en langue des signes,
 - la dynamique d'expression est en prise avec les organes qui assurent la manipulation (mains), la locomotion (jambes, déplacements dans l'espace) et l'interaction (bouche: imitation de la parole, et affichage des émotions et affects), ce qui rend possible l'iconicité (les relations d'analogie entre les gestes de parole et les gestes de la vie ordinaire)
 - le système d'expression est multicanal et investit pleinement le corps dans toutes ses dimensions motrices et sensorielles, ce qui fait du « sens » une expérience plus sportive, physique et émotionnelle qu'en langue vocale
 - le système d'expression est hautement interactif avec le destinataire, ce qui fait du sens une coordination intersubjective vécue avec plus d'intensité que dans une langue vocale.

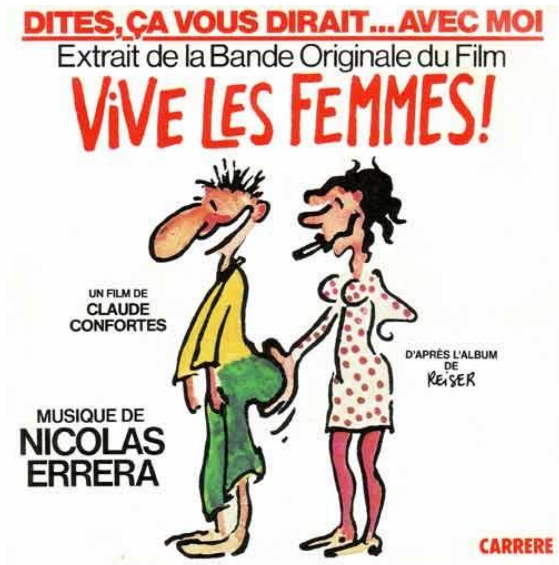
Conséquences pour la production d'avatars signeurs

- Le paradigme éactif nous amène à nous interroger non pas sur l'avatar lui-même en tant qu'objet, mais sur l'expérience vécue de l'interface homme-avatar considérée du point de vue de son participant humain.
- L'écueil à éviter (me semble-t-il) est de produire des avatars signeurs dont les signes sont conçus comme des unités formelles strictement segmentales ou linéaires, et aussi conçus comme des symboles qui renvoient à une unité de sens ou de représentation séparée.
- Dans l'idéal, l'avatar signeur gagne en efficacité
 - (i) en faisant varier la mimique
 - (ii) en faisant varier l'orientation du regard vers le destinataire, le geste ou une zone du cadre ou du champ
 - (iii) en donnant des indices de la manière dont l'acte de signer est éprouvé ou ressenti en termes de fatigue, de force, d'émotion jouée.→ tout le défi est de faire entrer les composantes de l'interactivité incarnée dans un monde virtuel, par définition immatériel, mais dans lequel on doit se reconnaître un minimum.

Une piste de réflexion: la bande dessinée

- Certains dessinateurs de BD ont développé un style graphique qui concentre un « maximum d'humanité » dans un « minimum graphique »:

→ Reiser



Une piste de réflexion: la bande dessinée

- Certains dessinateurs de BD ont développé un style graphique qui concentre un « maximum d'humanité » dans un « minimum graphique »:
→ Jul / Silex and the City



- Ce type de bande dessinée fixe dans un graphisme simple et limité une impression forte de l'intention, de l'émotion, du ressenti dans le vécu.
→ très probablement, le renforcement des « indices d'humanité vécue » dans la simulation par les avatars signeurs pourrait s'appuyer sur certaines techniques graphiques de ce type, en environnement informatique régulé. De manière révélatrice, la BD Silex and the City conserve toute sa puissance dans le dessin animé, sans aucune perte, parce que cette BD était un candidat idéal à la mise en animation.

Conclusion

- **L'énaction est une théorie de l'expérience vécue = de la production de la réalité ressentie (qui masque l'invisibilité du monde en soi)**
 - porter un regard éenactif sur les avatars signeurs revient à orienter leur conception en fonction de certains paramètres de la cognition incarnée, essentiels en langue des signes
 - toute langue des signes fait produire le sens dans l'interactivité et en s'appuyant sur l'ensemble du corps comme ressource, dans son rapport à l'autre, à l'espace et à lui-même
 - peut-on imaginer des avatars qui puisse simuler un « vécu de signeur » dans leur apparence et leur attitude?
 - variation des points de vue autour de l'interface: quel rôle attribue-t-on à l'observateur? Celui de témoin en 3^e personne, qui observe la performance et la « décode » sans se vivre comme participant, ou celui de destinataire en 2^e personne, qui se sent interpellé comme si l'avatar prenait une consistance relative?
 - finalement, l'interface avec l'avatar repose la question du test de Turing: jusqu'à quel point adhère-t-on à l'avatar, non pas en tant qu'agent psychologique autonome, mais en tant que modèle virtuel d'un signeur interactif?

Conclusion (remarques finales)

- L'énaction est également une théorie de la socialité et de l'interactivité
 - le corps vivant énonce des micro-mondes vécus en fonction du moment d'évolution ou d'histoire autopoïétique de l'espèce et de la communauté, qu'il réalise à un moment donné dans l'espace et dans le temps
 - les communautés humaines qui parlent une langue donnée développent des traits culturels, physiques et attitudeux qui les caractérisent de l'intérieur (sentiment d'appartenir à une tribu) et de l'extérieur (formation des clichés et stéréotypes des groupes humains les uns sur les autres: les Italiens sont ceci ou cela selon les Français, qui sont ceci ou cela selon les Allemands, etc.).
- or les communautés de signeurs ont connu ces dernières années une mise en observation à travers des fictions cinématographiques ou autres, comme La Famille Bélier, dont le réalisme est sujet à caution (cf. Bienvenue chez les Ch'ti, et j'en suis un moi-même) mais contribue à fixer une « identité » de la communauté face au public extérieur.
- D'où une question technique concernant l'avatar signeur: cherche-t-on une simulation relativement immatérielle d'un signeur anonyme, ou une simulation relativement consistante d'un signeur muni d'un vécu virtuel et d'un ancrage social virtuel? La question est un peu curieuse mais elle impacte en premier lieu l'idée de l'interface et au-delà les orientations de la conception dans sa dimension technique.