



HAL
open science

Prosodie et émergence du sens. Propositions pour une étude cognitive de l'intonation

Gilles Col

► **To cite this version:**

Gilles Col. Prosodie et émergence du sens. Propositions pour une étude cognitive de l'intonation. Canadian Journal of Linguistics / Revue canadienne de linguistique, 2007, 52 (3), pp.255-277. halshs-00602544

HAL Id: halshs-00602544

<https://shs.hal.science/halshs-00602544>

Submitted on 27 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PROSODIE ET EMERGENCE DU SENS
PROPOSITIONS POUR UNE ETUDE COGNITIVE DE L'INTONATION

Gilles Col

Université François-Rabelais, Tours (France)
Laboratoire FORELL, MSHS, Poitiers (France)

Canadian Journal of Linguistics/Revue Canadienne de Linguistique,
Numéro 52(3): 255-277 – 2007

Version auteur

Résumé :

Cet article vise à montrer que l'intonation (dans le cas de l'anglais) participe activement à la construction et à l'émergence du sens. Il présente quatre propositions qui vont de la perception de l'intonation jusqu'à son rôle dans la représentation du sens. On défend tout d'abord le concept de « forme intonative » qui repose sur le modèle gestaltiste de « bonne forme » (proposition 1) et dont la caractéristique fondamentale est d'être « dynamique » (proposition 2). Dans la proposition 3, l'intonation est replacée dans l'épissure sémantique avec les autres composantes linguistiques (syntaxe, lexique, grammaire). Enfin la proposition 4 donne à l'intonation un rôle fondamental dans l'évolution de la « scène verbale ».

Abstract :

This paper aims at giving to intonation (in the case of English) a driving role in the building and the emergence of meaning. It presents four propositions, going from the perception of intonation to its role in the representation of meaning. First, the concept of « intonational form », based on the gestaltist model of « good form » is stated (proposition 1) as well as this of « dynamic form » (proposition 2). In proposition 3, intonation is equaled to the other linguistic components (syntax, lexicon, grammar) in the semantic splicing. Finally, proposition 4 gives intonation a fundamental role in the evolution of the « verbal scene ».

1. Objectif et cadre théorique

Je commencerai cette étude par un préambule méthodologique. Les analyses qui suivent s'inscrivent dans un travail plus global de recherches en sémantique. L'objectif de cette étude est précisément d'essayer d'intégrer des données prosodiques dans le domaine spécifique qu'est la sémantique « cognitive »¹. Par sémantique « cognitive », j'entends une sémantique qui cherche à dépasser les distinctions entre phrase et discours, entre grammaire et lexique, et qui postule globalement une approche « constructiviste » et non-compositionnelle du sens.

Ainsi, les propositions qui vont suivre sont pour l'instant programmatiques (et très largement théoriques...), et définissent un axe de recherches². Leur but est d'ouvrir un débat, dans les études anglaises, sur une approche cognitive de l'intonation de l'anglais, en établissant des liens entre des domaines comme la grammaire, le lexique, la pragmatique, la logique, et dans une certaine mesure, la phénoménologie.

Dans cette optique, « cognitif » prend alors le sens suivant : prendre en compte « les interlocuteurs en tant que sujets pensants, engagés dans une activité cognitive de production et/ou de compréhension de l'énoncé » (Victorri et Fuchs 1996 : 199). Pour ces auteurs, comprendre un énoncé, c'est construire quelque chose, à partir de l'activité d'énonciation et de locution.

L'arrière-plan théorique de cette étude est double³. Il s'agit d'envisager la construction du sens en termes de spatialisation, c'est-à-dire de représentation par des espaces. Je m'appuie autant sur la théorie des « espaces mentaux » de Fauconnier (1984, 1994 1997), que sur celle de la « scène verbale » de Victorri. Le mot « espace », dans ces deux approches, n'est pas à prendre dans une perspective représentationnelle au sens de « représentation du monde extralinguistique ». Le postulat de Fauconnier est de « voir le déroulement du discours comme une série de configurations cognitives produites successivement les unes à partir des autres » (1991:231). Ces configurations ont comme caractéristique de « subdiviser l'information en la relativisant à différents domaines » (ibid:232). La finalité de la théorie des espaces mentaux est de rendre compte du travail de compréhension de l'énoncé, travail que Fauconnier exprime en termes d'espace. Ces espaces sont définis par Fauconnier de manière mathématique, comme des « ensembles structurés et modifiables —des ensembles avec des éléments a, b, c, ..., des relations satisfaites par ces éléments (R1ab, R2a, R3cbf, ...) et tels que l'on puisse ajouter de nouveaux éléments, ou établir de nouvelles relations entre leurs éléments » (1984:32). Dans une perspective sémantique et pragmatique, les espaces mentaux sont mis en place à partir d'expressions linguistiques qui jouent le rôle d'« introducteur d'espace mental » : « on appellera « introducteur » les expressions qui établissent un nouvel espace ou qui renvoient à un espace

Je tiens à remercier très chaleureusement Jean-Louis Duchet qui a consacré beaucoup de temps à relire cet article et à répondre à mes questions. Je remercie également Anne Lacheret-Dujour et Bernard Victorri pour leurs remarques sur ce travail.

Une première version de cet article a été publiée dans les actes du XI^e colloque de l'Association des Anglicistes pour les Etudes de Langue Orale dans l'Enseignement Supérieur, Secondaire et Élémentaire organisé à l'Université Paris 13 (Villetaneuse, France). Je remercie l'association pour autoriser la publication de cette version modifiée de l'article.

¹ Pour l'utilisation de ce terme, je renvoie au numéro 53 de la revue Communications, (1991).

² D'autres domaines de la langue anglaise ont déjà été abordés dans la même perspective, comme la grammaire (Col 2001, 2006), le style indirect libre (Col 2003, 2004), et la sémantique temporelle (Col, à paraître).

³ Voir les différentes présentations de ces deux théories dans Col 2001, 2003.

déjà introduit dans le discours » (Fauconnier, 1984:32⁴). Un introducteur sert plus particulièrement à mettre en relation deux termes de deux espaces mentaux par l'intermédiaire d'une fonction (appelée « fonction pragmatique »). Il permet ainsi l'accès à un autre espace, nouveau ou déjà construit. De manière simplifiée, on dira qu'on a accès à un espace, M, à partir d'un autre, B (pour Base). L'accès à proprement parler se fait par l'intermédiaire d'éléments de chaque espace mis en relation tel que : « If two elements *a* and *b* are linked by a connector *F* ($b = F(a)$), then element *b* can be identified by naming, describing, or pointing to its counterpart *a* » (Fauconnier 1997:41). Ainsi, on peut distinguer (i) un espace-point de vue (Viewpoint) à partir duquel les autres sont structurés, mis en place ou atteints, (ii) un espace-focus, c'est-à-dire l'espace en cours de « structuration interne » (celui sur lequel est portée l'attention), et (iii) un espace de base, le point de départ de la construction, vers lequel on peut toujours revenir. L'idée fondamentale étant qu'une phrase en tant que telle n'a pas un nombre défini de sens (« a sentence in itself has no fixed number of readings », 1997:54), mais plutôt un potentiel pour générer du sens : « It has a potential for generating connections in mental-space configurations » (1997:54). Les espaces mentaux constituent ainsi un réseau qui s'enrichit dans le déroulement du discours et permet le déploiement du sens : « Une expression de langue E n'a pas de sens en soi ; elle a plutôt un potentiel de sens et c'est dans un discours complet en contexte qu'il y aura production et actualisation de sens. [...] Le déroulement du discours [met en place des] domaines rattachés les uns aux autres par des connecteurs, et structurés de façon interne, aussi bien à partir d'indices linguistiques que d'éléments situationnels et contextuels. [...] Une expression de langue qui intervient dans le discours au stade n place alors un ensemble de contraintes sur la nouvelle configuration produite, cela en fonction de la configuration déjà engendrée au stade n-1. » (Fauconnier 1991:231). Pour compléter cette brève présentation, deux notions sont fondamentales dans la théorie des espaces mentaux : la correspondance (« mapping ») d'un espace à un autre espace, à travers l'identification d'éléments des différents espaces entre eux, et le « mixage » (« blending ») de structures partielles qui crée une structure nouvelle et enrichie. Le sens se développe ainsi à travers le réseau d'espaces mentaux et non pas dans un espace isolé : « In the many-space model of conceptual projection, meaning is not constructed in any single space, but resides in the entire array and its connections. » (1998:24)

Le postulat de Victorri se démarque de celui de Fauconnier. Son objectif est tout d'abord de rendre compte de l'activité d'énonciation elle-même, et pas seulement de la compréhension de l'énoncé. Pour lui, « toute énonciation construit un espace d'une nature particulière, que nous appelons le champ intersubjectif, et qui est doté d'un statut phénoménologique propre : il est le résultat de l'activité de parole, qui possède cette propriété de faire surgir devant les interlocuteurs des entités, des procès, etc. qui ne doivent leur mode d'existence qu'à leur évocation par la parole et le sentiment partagé par les interlocuteurs de pouvoir les « percevoir » en tant que produits de cette évocation » (Victorri 1999:89). Le point de vue de Victorri est proche des travaux effectués en phénoménologie, et associe clairement la sémantique et un certain type de perception. L'espace construit dans le discours est un espace partagé par les interlocuteurs, et se trouve être « irréductible à de simples images mentales » (Victorri et Fuchs 1996:201). Ce que Victorri appelle « scènes verbales » sont les structures de base du champ intersubjectif créé par la parole : « les entités et les procès évoqués par la parole s'organisent en scènes, dotées d'une dimension temporelle qui permet d'évoquer de manière très directe le déroulement des procès. On peut comparer ces scènes à des séquences cinématographiques, avec leur cadrage, leur angle de prise de vue, leur avant plan et leur arrière plan, la possibilité de construire une scène à l'intérieur d'une autre, etc. Mais ce n'est là qu'une analogie. Les scènes verbales ne sont pas des scènes cinématographiques : elles ont des caractéristiques propres qui les différencient radicalement de scènes visuelles. » (1999:89). Les caractéristiques des scènes

⁴ Voir aussi 1997:40 et chapitre 2 en général, où l'auteur parle également de Space Builder à propos des prépositions, adverbes, conjonctives, etc.

verbales ne seront pas développées ici⁵, mais on retiendra qu'elles sont de nature essentiellement topologique et associent une dynamique, des dimensions, un point de vue, et qu'elle permettent des relations entre elles. Dans cette perspective, le sens global des énoncés sera « identifié avec les scènes verbales qu'ils ont contribué à évoquer dans le champ intersubjectif » (1999:90). Contrairement à l'approche de Fauconnier, la prosodie est présente dans ce cadre théorique. Elle joue effectivement un rôle, avec les autres données linguistiques, dans la construction de la scène verbale (voir la proposition 4 dans la section 3 pour de plus amples développements sur ce point). Globalement, elle contribue « soit à évoquer des entités ou des événements sur la scène, soit à modifier le point de vue sur cette scène » (Lacheret-Dujour, Ploux, Victorri, 1998:99).

L'intérêt de ces deux approches cognitives est de concevoir le sens d'une part comme représentable (et là je m'inscris aussi dans la Théorie des Opérations Énonciatives de Culioli qui sert d'ailleurs de point de départ à la théorie des scènes verbales⁶) et spatialisable, et d'autre part comme une construction complexe et interactive (je reviens plus loin sur la question de la « construction »). Ces deux cadres théoriques me permettent de faire des propositions concernant la place et le rôle de l'intonation dans le développement du sens. La théorie de Fauconnier permet en effet de mieux caractériser l'intonation du point de vue perceptif, en relation avec les autres composantes linguistiques et en en proposant une visualisation complexe. La théorie de Victorri, de son côté, permet de mieux intégrer la prosodie dans l'activité d'énonciation et l'évolution de la scène intersubjective. L'objet de cette théorie est effectivement l'analyse de l'interaction langagière, et toutes les données linguistiques sont prises en compte, y compris l'intonation, dans la mesure où elles servent à évoquer ou à modifier des entités/événements sur la scène verbale.

1.2 L'avant-plan théorique (c'est-à-dire ce qui va me permettre plus directement de faire des propositions concernant l'intonation) est double également. Il prend appui sur la théorie de la Forme (Gestalttheorie) et sur une conception pragmatique des relations entre formes de l'énoncé et structures informatives (Lambrecht 1994). La théorie de la Forme, historiquement théorie psychologique et philosophique, rejette l'idée d'« atomisme mental » : « il n'y a pas de sensation, d'images, de sentiments qui puissent être isolés du tout » (Guillaume [1937], 2000:14). Cette approche va me permettre de développer une vision unifiée de la prosodie et de la sémantique, en m'appuyant sur des mécanismes de perception simples, comme la ségrégation fond / figure, et le principe de la « bonne forme » (voir plus bas). L'intérêt de ce que propose Lambrecht va dans le même sens. Il cherche à faire le lien entre les formes des énoncés et la prosodie, notamment en étudiant les structures syntaxiques focalisantes. Son propos est de revoir la dichotomie accentuation de l'information rhématique / désaccentuation de l'information thématique, et de suggérer un principe général (General Phrasal Accent Principle) où syntaxe et prosodie participent de manière égale à la définition d'un domaine sémantique. Lambrecht considère effectivement que la syntaxe marque l'ouverture du domaine, alors que la prosodie signale sa fermeture, plaçant du même coup les éléments (grammaticaux ou autres) situés après une tête ou un noyau intonatif en dehors de ce domaine : « a sentence accent marks the END of a semantic domain, whose BEGINNING is marked by non-prosodic means » (1994:247). Cette conception va en définitive à l'encontre d'une conception « instrumentale » de la prosodie, par laquelle le signal acoustique marquerait le début d'un événement : « [It is] in fact the opposite of the natural mechanism whereby a noise signals the beginning of an event, as e.g. the mechanism whereby the sound of a starting pistol marks the beginning of a dash » (1994:248).

⁵ L'article de Victorri (1999) est accessible dans les archives en ligne Sciences Humaines et Sociales du CNRS à : <http://halshs.ccsd.cnrs.fr>

⁶ « Nous rappellerons qu'énoncer, c'est construire un espace, orienter, déterminer, établir un réseau de valeurs référentielles, bref un système de repérage. » (Culioli 1999a:49).

L'objectif général de cette étude étant de faire des propositions spécifiques sur le rôle de l'intonation d'un point de vue cognitif (dans la perspective de ce que j'appelle « l'émergence du sens »), je me contenterai de rappeler ma propre perspective de travail, déjà développée ailleurs (Col 2001, 2003, 2004). Elle se résume à une métaphore (avec le risque de simplification que comporte l'usage des métaphores dans le domaine scientifique...), celle de l'épissure. En termes nautiques, une épissure est la jonction de deux cordages par les brins qui les composent. Il semble ainsi que les différents éléments comme la prosodie, la syntaxe, la grammaire, le lexique se rejoignent par entrelacement pour former un énoncé, et ne se superposent pas. Le sens ne serait alors pas un tout unifié : du sens serait produit et ce sens serait représentable par des opérations, et par ailleurs, du sens émergerait, par exemple de relations conflictuelles entre différents repères énonciatifs, ou différents niveaux d'énonciation comme dans le cas du style indirect libre (voir Col 2003, 2004). C'est expressément dans cette perspective que les propositions qui suivent s'inscrivent.

2. Problématique

A part quelques exemples isolés (cf. par exemple Huart 2002), les données prosodiques ne sont pas toujours présentes dans les analyses grammaticales et syntaxiques, et quand elles le sont, c'est généralement sous la forme d'argument supplémentaire pour aller dans le sens de la syntaxe. On situe assez facilement l'intonation au-dessus et parallèlement à la syntaxe et à la grammaire. Rossi définit ainsi l'intonation comme « [relevant] en premier lieu d'un niveau supérieur parallèle à la syntaxe et à la sémantique, dont elle reçoit l'information pour se constituer » (1999:197). On peut encore renvoyer à la définition déjà plus ancienne de l'intonation comme élément « suprasegmental » (Hockett 1942). La place de l'intonation répond ainsi à une vision des langues comme des architectures complexes, vision qu'elle contribue à conforter. Je souhaiterais plutôt faire de l'intonation un élément autre que situé vaguement à part du reste des systèmes linguistiques, par conséquent donner à l'intonation un autre statut qu'un simple ajout à l'édifice. Je rejoins ici entièrement le point de vue de Nicaise quand il déclare que « l'intonation ne véhicule pas un message qui se surajoute au message du texte [...] [Il y a] interaction entre mélodie et texte » (2000:130). On concevra donc l'intonation comme un moteur de l'interprétation et une cause de l'émergence du sens.

Face à une conception hiérarchisée de l'intonation, on note l'hypothèse de type lexicaliste adoptée par Nicaise (2000:121-122), dans la lignée des travaux de Liberman ou de Ladd (1996). Cette hypothèse permet de caractériser les unités intonatives sur le modèle du lexique et de sa morphologie. On distingue de cette façon racines et affixes. Les mélodies simples, ou nucléaires, se définissent comme les racines, et les mélodies complexes, avec introduction d'une tête dans l'unité intonative, comme une mélodie constituée d'une racine et d'un préfixe. On a ainsi une liste fermée de mélodies qui prennent différentes valeurs selon les contextes. Chaque partie, tête/préfixe et noyau/base, contribue à l'interprétation des mélodies. Je renvoie à l'ouvrage de Nicaise et Gray (1999) pour l'exposé complet de cette approche. Celle-ci met en valeur le rôle du contexte dans l'interprétation de l'intonation, et c'est un point central sur lequel je vais revenir. Elle permet de mettre en relation les différents éléments qui composent une unité intonative, en particulier la tête (ou « préfixe » dans cette approche) et le noyau (« racine »). Ainsi, une égale importance est donnée à chaque partie de l'unité intonative, qui contribue chacune à l'interprétation des mélodies. La priorité n'est pas donnée à un élément en particulier, comme c'est généralement le cas avec la courbe finale, y compris dans l'approche suggérée par Ladd pour qui la tête joue un rôle limité : « the head ... is merely a substring of the contour » (1996:211).

Une telle approche permet avant tout de prendre en compte dans l'interprétation des mélodies, la morphologie globale d'une unité intonative et « l'interaction entre mélodie et texte » (Nicaise 2000:130) mais pas de rendre compte de la constitution même de la mélodie. Pourtant,

parler de « préfixe », de « racine » d'une mélodie permet non seulement de décrire ce lexique particulier, mais aussi de poser la question de sa constitution. C'est justement sur ce point que je vais faire des propositions (cf. 3.2).

Dans une autre type d'approche, issue des neurosciences, l'intonation joue un rôle central dans l'élaboration du sens. Elle joue plus particulièrement un rôle dans l'accès au lexique. Les travaux en neuropsychologie (Lambert et Nespoulos 1997 ; Lechevallier et al. 1995, Segui, Dupoux, Mehler 1990, Segui 1992) font même apparaître que ce rôle est double. Segui (1992) montre que la prosodie permet la détection du phonème initial des mots plurisyllabiques et que cette détection se fait grâce au code accoustico-phonétique avant l'accès au lexique ; la détection du phonème initial d'un mot monosyllabique quant à elle « pourrait s'effectuer à partir du code lexical correspondant [...] En effet, la présentation d'un mot monosyllabique donne lieu à un accès « irrépissible » au lexique mental » (1992 : 144). La perception (acoustique ou visuelle) est, dans le cadre des neurosciences, un système de traitement et d'interprétation de l'information à partir de représentations déjà acquises (cf. Segui 1992, Bonnet 1995). Le traitement des sons linguistiques est néanmoins complexe et s'effectue à différents niveaux : phonétique, phonologique, lexical, sémantique et syntaxique (Eustache 1995 : 248). Cette approche, que l'on pourrait qualifier de « fonctionnelle » vu qu'elle cherche à doter l'intonation d'un certain type de fonction dans l'activité neuronale, permet de faire un lien entre lexique et intonation du point de vue de la perception du signal sonore. L'intonation permet ainsi l'accès au sens. Il semble maintenant essentiel, à partir des observations des neurolinguistes, de se demander comment se construisent les unités intonatives, et au-delà, comment s'élabore le sens, et non pas seulement comment on y accède⁷.

3. Propositions

Proposition 1 : pour des formes intonatives

L'intonation, dans une perspective centrée sur l'activité du sujet, est à redéfinir comme un cas de perception de formes acoustiques, d'objets sensibles, qui mettent en jeu les mécanismes perceptifs traditionnels. La théorie de la forme⁸ repose en effet sur l'idée que « il n'y a pas de coupure entre le corps et l'esprit. [...] L'esprit n'est pas une force organisatrice qui, de manière mystérieuse, par une activité spontanée et inconditionnelle, ferait surgir, d'un chaos de processus physiologiques, un ordre qui leur serait complètement étranger » (Guillaume 2000:26). Il y a donc des organisations (« Gestalt » en allemand) qui ne relèvent pas directement de l'activité de l'esprit. Cette théorie « part [ainsi] des formes ou structures considérées comme des données premières. Elle ne se donne pas une matière sans forme, une pure multiplicité chaotique pour chercher ensuite par le jeu de quelles forces extérieures à ces matériaux indifférents ceux-ci se grouperaient et s'organiseraient. Il n'y a pas de matière sans forme » (ibid:24). Chaque organisation / forme est donc perçue grâce à un mécanisme général (la ségrégation entre un fond et une figure) et grâce à plusieurs facteurs qui permettent d'aboutir à une « bonne forme ». La ségrégation entre un fond et une figure repose sur l'idée que « tout objet sensible n'existe qu'en relation avec un certain « fond » (ibid:65). Figure et fond se caractérisent ainsi : « [Ils] ont tous deux leur unité, mais il y a deux types d'unités [...] : celle de la figure, qui possède forme, contour, organisation, et celle du fond, qui est une continuité amorphe, indéfinie, inorganique » (ibid:67). Une figure est alors une bonne forme en tant que la meilleure forme possible, c'est-à-dire une forme « régulière, simple et symétrique » (ibid:63). Elle n'est pas la somme de plusieurs éléments, mais « une fonction de plusieurs variables » (2000:24). Guillaume résume la

⁷ Il semble également important de dépasser l'impression donnée par la neurolinguistique que le sens est plus ou moins donné, et que le langage permet d'y accéder.

⁸ La présentation qui suit s'inspire essentiellement de l'ouvrage de Guillaume 2000.

théorie par le principe suivant : « une forme est autre chose, ou quelque chose de plus que la somme de ces parties. Elle a des propriétés qui ne résultent pas de la simple addition des propriétés de ses éléments ». (ibid : 18).

Cette conception non-compositionnelle⁹ nous encourage à définir une unité intonative comme une « forme » intonative. La Gestalttheorie s'est essentiellement intéressée à la perception de l'espace, mais de nombreuses analogies ont été faites avec l'étude des mélodies : « L'étude de la mélodie nous [montre] comment des sons [...] indépendants les uns des autres engendrent dans la conscience de l'auditeur un « phénomène » qui présente les caractères des formes » (ibid:25). La ségrégation fond / figure repose effectivement sur une différence de « grain » et de granularité pour la perception visuelle, mais plus largement sur une différence d'« intensité » : « il n'y a donc de perception d'objet que si des différences d'intensité existent entre les excitations provenant de plusieurs parties du champs » (ibid:64). C'est ce principe de ségrégation par intensité, ou hauteur, qui nous permet de proposer une redéfinition des unités intonatives en « formes » intonatives.

Redéfinir une unité intonative en forme intonative nous amène par extension à reconsidérer les éléments qui composent la forme en question et qui sont, pour simplifier, les saillances intonatives. Ces saillances correspondent à des syllabes qui se distinguent des autres par leur hauteur, si on schématise l'intonation comme une série de hauteurs et de creux, c'est-à-dire comme une série de syllabes marquée par une hauteur de la fréquence fondamentale (Fo) et d'autres avec une fréquence plus basse. On pourrait ainsi considérer ces syllabes, suivant le principe gestaltiste, comme des « bonnes formes », c'est-à-dire, des formes « régulières, simples, symétriques » et unifiées.

On sait effectivement, grâce aux travaux en neuropsychologie (voir, pour un exposé général, Segui 1992), que le signal de parole est continu, alors que la perception de la parole est de nature discrète. Segui fait également l'hypothèse que le premier niveau de segmentation du signal sonore est la syllabe, et que celle-ci serait perçue de manière unitaire. La syllabe est analysée dans ce cadre méthodologique comme une structure, avec une hiérarchie propre. Cette conception n'est pas sans poser certains problèmes, en particulier concernant les syllabes de l'anglais dont la segmentation est loin d'être aussi nette que dans d'autres langues (voir Cruttenden 1994, Duchet 1993)¹⁰. Elle permet néanmoins de postuler une conception gestaltiste de l'intonation. En effet, on considère en perception visuelle que certains éléments se détachent des autres parce qu'il y a derrière eux d'autres éléments qui servent à les distinguer, ce qui crée un contraste qui aide à « l'émergence d'une figure structurée » (Bonnet 1995:42). De la même manière, je formulerais l'hypothèse d'une corrélation entre la hauteur de Fo et la perception unitaire de la syllabe : plus Fo sera élevée, plus la syllabe apparaîtra unifiée, de même qu'en perception visuelle « la « fermeture » des contours [d'une figure] est une opération nécessairement préalable à l'« émergence » d'une figure structurée » (Bonnet ibid). Ainsi, les variations de Fo créent un système de contrastes entre syllabes saillantes et présentant un degré d'unité important, et d'autres qui le sont moins¹¹. Par extension, la perception unifiée de la syllabe constitue un élément primordial de l'unification de ce que Rossi (1999) ou Hirst et Di Cristo (1986) appellent l'« unité intonative » : la saillance intonative est définissable, dans la

⁹ Celle-ci se caractérise par ailleurs par son originalité : Guillaume précise (2000:18) que les promoteurs de la Gestalttheorie se réclament de von Ehrenfels dont l'ouvrage fondateur est paru en 1890. On lira également avec intérêt l'ouvrage consacré à un autre animateur de la Gestalttheorie, Wolfgang Köhler (voir Références).

¹⁰ La segmentation syllabique pose certains problèmes dans le cas des consonnes intervocaliques, comme par exemple : extra, balance, butter... (Voir Cruttenden 1994:50-51 pour une présentation simple). La question est de savoir à quelle syllabe appartient le /s/, le /l/, ou le /t/ dans ces trois exemples.

¹¹ Montrer la corrélation saillance / perception unitaire de la syllabe demande une analyse phonétique fine qui est en cours de réalisation.

conception gestaltiste de forme défendue ici, comme un élément dynamique de la construction de l'unité intonative.

Proposition 2 : pour une conception relationnelle de l'intonation

On vient de voir qu'une approche gestaltiste de l'intonation nous amène à formuler l'idée de « forme intonative » dans laquelle les éléments saillants (certaines syllabes, on va y revenir) semblent perçus de manière unifiée. Cette approche implique du même coup une conception relationnelle des éléments saillants dans la forme intonative, que je développe dans ma seconde proposition.

Une figure structurée émerge d'un fond parce qu'elle est en relation avec ce fond, dans la continuité de ce fond « amorphe, indéfini, inorganique » (Guillaume 2000:67). Ceci nous invite à promouvoir une approche « contextualisée » de l'intonation et à étudier en priorité les relations entre le noyau et la tête de la forme intonative. Noyau et tête sont des parties plus saillantes d'un tout organisé. Mais à l'intérieur de ce tout s'établissent des relations entre les différentes parties : noyau, tête et autres syllabes. Le noyau et la tête sont par ailleurs en relation de contraste entre eux-mêmes, dans la mesure où ce sont des syllabes perçue de manière unitaire.

L'approche contextualisée de l'intonation que je défends me fait prendre en compte les diverses saillances qui apparaissent dans le discours. Ces saillances appartiennent soit à la forme intonative considérée (tête et noyau), soit, en l'absence de tête, à la forme intonative la plus proche (le critère de proximité jouant aussi un rôle essentiel dans la perception de la forme, selon Guillaume 2000:61). La tête est effectivement un élément facultatif de la forme intonative, et en son absence, un contraste s'établit entre le noyau d'une forme et le noyau de la forme qui précède.

Prenons un exemple pour illustrer cette conception relationnelle des saillances dans une forme intonative. C'est un exemple tiré d'un reportage sur une petite entreprise du Pays de Galles qui propose des plats prêts à consommer, qui se veulent de bonne qualité. C'est la journaliste qui parle, et qui commente la démarche de l'entreprise. L'exemple est le suivant, présenté d'une part avec une transcription prosodique et d'autre part sous forme de visualisation de la courbe intonative réalisée avec le logiciel PRATT :

(1) It sounds like a simple enough idea but setting up a scheme hasn't been easy.
 'L'idée semble assez simple, mais sa réalisation n'a pas été facile.'

<i>It</i>	<i>sounds</i>	<i>like a simple enough idea</i>	<i>but</i>	<i>setting up a scheme hasn't been</i>	<i>easy</i>
	H	*ML%	H		*ML%

La notation utilisée propose de découper l'énoncé en différents tons. On va distinguer les tons suivants : ton haut (H), c'est-à-dire le haut de la tessiture, ton moyen (M) pour la partie moyenne de la tessiture, et ton bas (L) pour la partie basse. La notation « % » marque ce que Liberman (1975) appelle le « boundary tone » ou « ton de limite » qui correspond à la dernière syllabe d'une proposition syntaxique. La dernière notation « * » correspond au « ton nucléaire » (« nuclear tone ») associé à la syllabe qui porte l'accent nucléaire, c'est-à-dire l'accent principal d'une unité lexicale de l'anglais. Ce système de transcription intonative tonale, utilisé dans la continuité des travaux de Nicaise (Nicaise 1987, 2000 ; Nicaise et Gray 1997), s'inspire des travaux de Pierrehumbert sur l'anglais (Pierrehumbert 1980 ; Pierrehumbert et Hirschberg 1990), ainsi que de Liberman (1975). Il se rapproche également du Tonal and Break Indices

(TOBI) de Silverman et al. (1992) qui proposent des critères de transcription prosodique de l'anglais comparable à l'Alphabet Phonétique International (IPA)¹².

Pour ce qui est de la courbe intonative correspondant à l'énoncé, le logiciel utilisé (PRAAT) en donne la visualisation suivante : voir Figure 1.

Le premier commentaire concerne l'aspect contrastif de l'intonation. Notre séquence associe ici trois formes intonatives. Le choix de traiter ces trois formes à la fois tient au fait qu'elles forment un ensemble dans le discours, ensemble syntaxique, thématique et prosodique. On relève ainsi que les deux grandes unités (it sounds like a simple enough idea et setting up a scheme hasn't been easy) contiennent chacune une tête et un accent nucléaire. On relève aussi que le contraste entre formes intonatives est double : tête et noyau d'une même forme ensemble (sound v. idea, setting v. easy) et les formes entre elles. Ainsi, le noyau idea contraste d'une part avec sounds à sa « gauche » et d'autre part avec setting à sa « droite ». Easy, de son côté, contraste essentiellement avec setting. Les relations entre têtes et noyau dans cet exemple ne sont pas vraiment linéaires : l'accent nucléaire est pris dans un réseau de hauteurs qui du coup le définissent, même si idea est directement suivi d'une pause. Idea est également suivi de but, ce qui complexifie encore les choses. But sert en fait de pivot syntaxique et prosodique à cette séquence. Sémantiquement, but est un « inverseur » exprimant une idée d'extériorisation (Garnier et Guimier 1997:181), et il peut être facilement isolé par la syntaxe et par la prosodie. Mais c'est aussi un connecteur syntaxique, dont la fonction est de contraster des éléments entre eux. Ici, il permet à plusieurs formes intonatives de contraster et, en même temps, il émerge du contraste créé entre la tête de la forme qui le suit (setting) et la tête de la forme précédente (sounds). La hauteur de but se situe entre sound et setting. Parler de contraste veut dire en définitive parler de contraste avec un contexte composé de différentes saillances de hauteurs différentes. C'est un contraste multidimensionnel, qui fait qu'un noyau est pris dans un réseau de saillances non seulement avec la tête de sa propre forme mais avec celle d'autres formes.

On voit ainsi comment le principe gestaltiste semble à l'œuvre dans la constitution de l'intonation. Dans cette perspective, une forme intonative est considérée comme une forme dynamique qui repose sur un ensemble de relations de type contrastif. L'aspect dynamique des formes intonatives est particulièrement visible dans la physionomie de la courbe finale, qui est un point crucial de l'étude de l'intonation tant du point de vue descriptif que sémantique (que l'on développe dans les propositions qui suivent).

Les questions que soulèvent les études de l'intonation portent généralement sur la courbe finale, et plus particulièrement sur le placement de l'accent nucléaire. C'est sur ce point que Huart conclut son exposé sur la mélodie : « en matière d'intonation, il importe de garder à l'esprit que ce qui compte le plus pour la compréhension est le placement des accents, en particulier de l'accent nucléaire » (2002:61). Ce point de vue est développé dans le cadre d'une grammaire de l'anglais oral, et de sa didactique. Cependant, il semble que l'étude de la courbe finale limite l'objet d'étude ; tout se passe comme si, dans la perspective d'un lexique des mélodies, on n'étudiait que le radical des mots sans les éléments d'affixation. A mes yeux, la place de l'accent nucléaire est un enjeu parmi d'autres de l'étude de l'intonation, avec la physionomie de la courbe, la place et la hauteur de la tête, la présence d'une avant-tête, etc, d'autant que le noyau n'est en fait pas un élément statique mais dynamique. Dans le cas des « question-tags » par exemple, la hauteur et donc le placement du noyau, ne posent pas vraiment de problème : la courbe part d'une position basse. C'est ce qui suit, soit une descente soit une montée (« Low Fall » ou « Low Rise ») qui entraîne un changement d'interprétation. Pour Nicaise et Gray (1999:91-92), le locuteur attend une confirmation de ce qu'il vient de dire dans le cas du Low Fall, alors que dans le cas du Low Rise, il invite l'interlocuteur à confirmer ou infirmer. C'est l'orientation de la courbe — qui, ici, n'est pas affectée par la hauteur du noyau —

¹² TOBI est cependant limité pour l'instant à la transcription de l'anglais américain et de ses variétés.

qui modifie l'interprétation de l'énoncé. On voit que l'interprétation du fait intonatif (sur lequel on reviendra plus bas) se fait dans ce cas à partir des éléments qui suivent l'accent nucléaire, le « contexte à droite » pourrait-on dire. Ceci m'amène donc à considérer que c'est le contexte en son entier, c'est-à-dire en plus du noyau, l'avant-tête et la tête quand elle apparaissent, ainsi que la queue de la courbe, qu'il faut prendre en compte, et pas uniquement l'accent nucléaire.

Prenons un autre type de configuration mélodique : le cas d'une même courbe finale, mais avec changement de hauteur de tête. D'après Nicaise et Gray (1999), les mélodies High Head + Low Rise et Low Head + Low Rise par exemple, n'ont pas la même interprétation. Dans le premier cas (1999:83), le locuteur donne l'impression d'une grande assurance :

(2) A : *Where are you going, Dad?*

B : *You'll see.*

H *LM%

Dans le second (1999:84), il peut exprimer une certaine réserve :

(3) A : *Can I have another apple?*

B : *I suppose so.*

L *LM%

On voit ainsi que la présence et la hauteur d'une tête dans l'unité intonative ont des conséquences au niveau de l'interprétation, ce qui est relativement prévisible. Il paraît donc essentiel de prendre en compte autant la courbe que l'environnement. Comme les travaux en sémantique grammaticale le montrent, une unité prend du sens à partir de son environnement et donne du sens à son environnement (voir le principe de convocation / évocation développé dans Victorri 1999). Il semble en être de même pour l'intonation. Une étude récente (Gaudy 2002) montre d'ailleurs comment dans le cas des question-tags descendants, ce qui précède la courbe annonce véritablement la suite. Gaudy défend effectivement l'idée que « le contour intonatif de la base [ce qui précède la courbe] comprend par anticipation le modèle de celui de la reprise [la courbe elle-même] » (2002:182). Cette conception rejoint celle de certains psycholinguistes qui insistent sur le rôle des contrastes intonatifs « comme support des processus attentionnels et d'anticipation » (Bacri 1994:52).

Ma seconde proposition suggère en définitive que c'est en termes de relations contrastives et de dynamique que se définit une forme intonative. Je vais, dans ma troisième proposition, voir ce qu'il en est alors des éléments moins saillants, dans la mesure où ils font partie de la physionomie générale de la courbe (constituée de manière continue d'une série de points) et qu'ils permettent aux saillances d'émerger.

Proposition 3 : place et rôle des éléments non-saillants

Partant du principe que l'intonation appartient à un ensemble linguistique plus large, mais qu'elle n'est ni au-dessus ni au-dessous des autres composantes¹³ que sont la syntaxe et le lexique (pour simplifier), je souhaiterais intégrer dans mes propositions les éléments qui ne sont pas « classiquement » pris en compte dans les études sur l'intonation : les éléments moins saillants que le noyau et la tête, essentiellement les autres unités lexicales. Ils constituent fondamentalement le « fond » qui permet aux éléments saillants d'être distingués, comme je l'ai suggéré dans ma première proposition. Je suggère maintenant qu'ils sont également partie prenante dans le processus général dit d'« intégration conceptuelle » (blending). Je souhaiterais proposer une vision unifiée des composantes linguistiques d'un énoncé, dans laquelle l'intonation est entrelacée avec la syntaxe, la grammaire, et le lexique.

Dans le cadre de la linguistique cognitive développée par Fauconnier et Turner, le blending se présente comme une opération cognitive qui permet de traiter des phénomènes linguistiques et sémantiques comme certaines constructions grammaticales, les métaphores, les constructions contrefactuelles (cf. Fauconnier 1997:182). Elle consiste à établir une projection partielle entre deux espaces initiaux (« inputs ») par une relation d'identité, causalité, intention, ... Une correspondance (« mapping ») est ainsi établie. Les éléments communs des espaces d'entrée sont réunis dans un « espace générique ». Les autres éléments des deux espaces sont projetés sur un « espace intégrant » (« blend »), qui regroupe donc une partie des éléments de chacun des espaces initiaux. C'est par conséquent une projection sélective. C'est dans cet espace intégrant qu'est produite une structure dite « structure émergente » qui n'est pas fournie par les espaces d'entrée, mais par la projection / « mapping » des espaces et par leurs relations¹⁴. La structure émergente se constitue à partir de certains éléments des espaces d'entrée. Elle n'existe donc pas à l'origine dans ces espaces, elle émerge de ces derniers. On obtient la schématisation suivante (d'après Fauconnier 1997a et Fauconnier et Turner 1998) : voir Figure 2.

Un des cas de traitement de blending les plus convaincants est sans doute celui de la métaphore. Si on prend l'exemple cité par Turner (2000), « Ce chirurgien est un boucher », et sans entrer dans les détails, on va pouvoir distinguer tout d'abord deux espaces d'entrée, l'un qui comprend le rôle de « chirurgien typique » ainsi que quelqu'un qui est chirurgien (Input 1), et l'autre qui comprend le rôle de « boucher typique » (Input 2). Dans l'espace générique, on trouve la correspondance établie par l'énoncé entre le chirurgien et le boucher. On trouve ici les éléments communs aux deux espaces d'entrée : le sang, les outils coupants, des corps, ... Le sens même d'incompétence véhiculé par l'énoncé ne se trouve donc pas dans les espaces d'entrée. Il se constitue dans l'« espace intégrant », c'est-à-dire l'espace issu d'une projection sélective d'éléments à partir des deux inputs. Les éléments « sélectionnés » sont, dans notre exemple, les qualités spécifiques du chirurgien (sauver des vies, pour aller vite) et celles d'un boucher (découper de la viande morte). Ces qualités sont mises en correspondance par leur projection dans l'espace intégrant. Et c'est cette projection qui fait émerger le sens d'incompétence qui n'existe pas au départ dans les espaces initiaux. Le sens d'incompétence s'élabore par conséquent sous la forme d'une « structure émergente », et plus précisément sous la forme d'une structure sémantique émergente.

Il ne s'agit pas ici d'appliquer tel quel le modèle Fauconnier / Turner à l'intonation de l'anglais. Il s'agit plutôt de retenir du modèle le principe de projection sélective d'éléments et de leur fusion produisant ainsi une structure sémantique dans laquelle l'intonation a une vraie fonction. Les éléments projetés sont issus des composantes de l'énoncé, composante

¹³ Le terme « composante » doit s'entendre au sens de « élément dynamique (force) entrant en composition » et qui produit un « ensemble complexe » (Dictionnaire Le Robert).

¹⁴ Sans entrer dans les détails, signalons que Fauconnier distingue trois types de relation : relation de composition, d'achèvement (« completion ») et d'élaboration (1997a:150-151). Toutefois, l'origine même de ces relations n'est guère explicitée dans la littérature. A noter enfin que cette présentation reprend en partie celle faite dans Col 2003 et 2004.

grammaticale, syntaxique, lexicale et intonative. Reprenons comme illustration l'exemple commenté plus haut : voir Figure 1.

Cet énoncé véhicule globalement le sens de difficulté. Plus exactement, il véhicule l'idée que la difficulté de l'entreprise n'est pas évidente à première vue, qu'il ne faut pas s'en tenir à la première impression si l'on veut apprécier à sa juste valeur tout le mérite de l'entreprise. En somme, il est difficile de se rendre compte que c'est difficile. Ma proposition vise à montrer que l'idée de difficulté non apparente se développe ici dans les différentes composantes de l'énoncé, et surtout dans leur fusion dans l'« espace intégrant » du modèle Fauconnier. Le domaine sémantique complexe de difficulté s'élabore en fait à partir d'une série de contrastes.

On relève un premier contraste au niveau des marqueurs de temps et d'aspect : on a d'un côté un temps simple (le présent de sounds), et de l'autre un marqueur aspectuel (present perfect de hasn't been). On a ainsi deux points de vue sur les procès, avec deux intervalles associés qui indiquent une forme de « conflit » (au sens de Gosselin 1996) : le présent simple ne limite pas la portée sémantique de l'assertion au moment d'énonciation comme le fait de son côté le present perfect. Les deux procès ne sont pas « perçus » de la même façon, pourrait-on dire, dans la mesure où ils ne renvoient pas au même type d'intervalle temporel. Leur corrélation (marquée par but) indique ainsi un contraste qui va permettre à l'idée de difficulté de se développer par rejet du premier procès et de son intervalle, à l'aide bien entendu d'autres marqueurs dans l'énoncé. En effet le verbe sound donne un point de vue davantage modal à l'assertion. Il véhicule un point de vue épistémique sur la validation de la relation sujet / prédicat que l'on notera <it - be a simple enough idea> (dans la cadre de la théorie de Culioli). Ce point de vue épistémique s'inscrit dans le cadre d'une comparaison (sounds like). Sound marque un choix lexical dans la mesure où ce verbe n'est pas grammaticalisé comme marqueur de modalité, comme peut l'être un auxiliaire modal par exemple. Mais sa présence dénote un point de vue qui contraste avec l'assertion de la seconde partie de l'énoncé (setting up a scheme hasn't been easy) dans laquelle ne figure aucun marqueur de modalité épistémique. Ce contraste entre sounds like d'un côté et l'absence de marqueur épistémique de l'autre renforce finalement l'assertion de « not be easy » en la rendant sémantiquement plus saillante (modalité v. absence de modalité), et prépare l'émergence de l'idée de difficulté. C'est ainsi qu'un premier groupe de marqueurs, ceux de temps et aspect, contribue au passage de la facilité apparente à la vraie difficulté.

On relève un second contraste qui va aussi permettre de faire émerger l'idée de difficulté, au niveau de la composante syntaxique. On a effectivement vu (dans notre seconde proposition) que l'énoncé avait comme pivot but à sens concessif. Cette conjonction contribue bien entendu à développer le contraste (voire l'opposition) entre les deux parties de l'énoncé.

J'en viens maintenant au troisième contraste et à la composante lexicale, et plus précisément à l'opposition sémantico-intonative simple enough idea et not easy. La visualisation de l'exemple sous PRAAT nous permet de constater que iDEa et EASy¹⁵ sont porteurs du noyau de chacune des formes intonatives, et que le premier (idea) a sensiblement la même hauteur que le second. Cette régularité au niveau du choix des noyaux permet d'associer deux unités lexicales dans une relation de conformité (« même forme que »). L'intonation, et plus exactement la saillance sur idea et easy, tend ainsi vers une harmonisation, mais les unités intonatives dans lesquels se trouvent les deux unités concernées s'opposent par leur sens, et plus largement par leur syntaxe et leurs marqueurs aspecto-temporels. On a ainsi affaire à une forme de « conflit » entre composantes, et pas seulement au sein d'une même composante comme dans le cas de la syntaxe et de la grammaire. Le rôle contrastif de l'intonation est finalement à voir aussi dans les relations plus complexes entre composantes et pas seulement au sein d'une forme intonative.

Les relations de contraste au sein d'une même forme intonative prennent aussi en compte les éléments non saillants. Un syntagme, quelle que soit sa nature (nominale ou verbale), se compose en effet d'éléments non saillants mais indispensables à sa constitution ; ces éléments-là

¹⁵ Les majuscules marquent les éléments saillants dans la forme prosodique (voir Figure 1).

participent aussi à l'élaboration du sens. On a ainsi commenté la présence de hasn't been (marquée par une certaine hauteur dans la courbe intonative) qui contraste avec le présent simple du verbe sounds qui est intonativement saillant. L'adjectif simple n'a pas non plus d'intonation saillante, mais il contraste néanmoins avec (not) easy. On voit ainsi que ces éléments font partie de l'épissure globale du sens : la plupart des éléments fusionnent dans l'idée de facilité apparente et de difficulté bien réelle, y compris les éléments non saillants.

On retrouve ainsi dans la composante intonation et ses relations avec les autres composantes ce que l'on a déjà observé dans les autres composantes, c'est-à-dire, le rôle contrastif et la participation à l'émergence de l'idée de « difficulté réelle ». Cette idée émerge vraiment au niveau de la courbe finale de la dernière unité intonative, mais elle est anticipée (ou « profilée ») par les autres composantes de l'énoncé, c'est-à-dire par les éléments moins saillants des deux formes intonatives. Je reprends finalement la conception de Lambrecht (voir première partie de cette étude) sur le rôle clôturant de l'intonation — « a sentence accent marks the **END** of a semantic domain, whose **BEGINNING** is marked by non-prosodic means » (1994:247). Mais j'y ajoute que le rôle de l'intonation ne se limite pas à la physionomie de la courbe finale ; il est à appréhender dans l'épissure de l'énoncé, et donc dans ses relations complexes aussi bien entre les composantes de l'énoncé qu'au niveau de la forme intonative.

Avec le modèle Fauconnier / Turner, on peut formaliser l'émergence de l'idée de difficulté de notre exemple dans la « structure émergente » du blending grâce à la projection de deux domaines sémantiques (domaine de la simplicité apparente et domaine de difficulté réelle) structurés par la grammaire, la syntaxe, le lexique et l'intonation. Ces deux domaines partagent des éléments linguistiques communs (tête et noyau, focus syntaxique) que l'on place dans l'espace générique (voir Figure 2). Ils ont aussi des éléments spécifiques (sites des têtes et noyaux respectifs, choix des marqueurs aspecto-temporels) qui permettent d'établir une projection partielle entre ces deux domaines. Mais c'est dans l'espace intégrant qu'émerge l'idée de difficulté « difficile à percevoir » : celle-ci est en quelque sorte en gestation dans l'ensemble de l'énoncé, dans ses différentes composantes, mais elle n'est visible qu'au niveau de l'énoncé comme fusion des différentes composantes sur la base d'une dynamique conflictuelle. Une telle conception permet ainsi d'intégrer l'intonation de manière active dans l'élaboration du sens, et non plus comme une confirmation de ce que dit la syntaxe¹⁶.

Proposition 4 : place et rôle de l'intonation dans la construction et l'évolution de la scène verbale

La précédente proposition porte sur les relations de la composante intonative par rapport aux autres composantes d'un énoncé, ainsi que sur la fusion des différentes composantes dans l'émergence du sens. Il s'agit maintenant de proposer une place et un rôle à l'intonation dans l'élaboration du champs intersubjectif, c'est-à-dire l'élaboration de ce que Victorri appelle la scène verbale. Dans la conception de Victorri, la scène verbale est construite par les éléments grammaticaux et syntaxiques et le lexique sert à faire évoluer la scène (voir Victorri 1999). Qu'en est-il alors du rôle de l'intonation, dans la mesure où on vient de voir que celle-ci concerne autant les éléments grammaticaux que lexicaux ? Je propose de situer le rôle de l'intonation, dans la perspective intersubjective de la scène verbale, au niveau de la perception du sens et non pas uniquement au niveau de son élaboration, même si ces deux activités sémantiques sont nécessairement liées.

¹⁶ Dans un cadre théorique différent comme l'approche modulaire, Nølke (1994) arrive à la même conclusion en refusant de donner à l'intonation une place particulière : « L'intonation intervient à plusieurs niveaux ; et c'est pour cela que j'ai préféré ne pas lui accorder un module particulier » (1994:72).

Pour avancer, donnons quelques précisions sur ce qu'est la « scène verbale », ainsi que la place de la prosodie dans ce cadre théorique. Ces remarques s'inspirent de deux articles : Victorri 1999, et Lacheret-Dujour, Ploux et Victorri 1998, ainsi que de la postface de Gosselin, 2005). La scène verbale désigne dans la terminologie de son auteur, le champ intersubjectif résultant de l'activité de parole. La parole sert à évoquer des scènes que le locuteur cherche à faire partager à ses interlocuteurs en faisant comme si elles se déroulaient devant eux, et qu'on pouvait les « percevoir ». On a vu dans la sous-section 1.1 que ces « scènes verbales » ont un statut phénoménologique bien particulier, dont le caractère intersubjectif et simulé les distingue à la fois des perceptions véritables et des pensées non verbalisées. Les propriétés de la scène pertinentes dans l'analyse de l'intonation sont essentiellement les dimensions et la dynamique; le point de vue sur la scène créé par l'intonation joue un rôle relativement prévisible, comme on va le voir. Je rappelle que dans l'optique de cette théorie, « les données linguistiques énoncées (prosodie comprise) sont des éléments de construction de cette scène verbale, qui servent soit à évoquer des entités ou des événements sur la scène, soit à modifier le point de vue sur cette scène » (Lacheret-Dujour, Ploux et Victorri 1998:99). Pour ces auteurs, « toute montée mélodique serait le signe d'une tension dans la relation intersubjective à propos de la construction de la scène, les descentes mélodiques étant au contraire le signe d'une résolution de cette tension » (ibid). La schématisation du français parlé donne dans cette perspective une courbe qui comporte une tête (montée mélodique de début de période), un corps (descente régulière entrecoupée de crêtes), et une queue en fin de période. Une période est suivie d'un continuatif final qui délimite deux périodes. En termes cognitifs, la tête correspond à « une demande de changement de point de vue sur la scène verbale : nouveau cadrage ou déplacement pour se centrer sur une entité de la scène. » (Lacheret-Dujour, Ploux et Victorri 1998:100). Le corps, c'est-à-dire la descente qui suit la tête, correspond « à une stabilisation du point de vue sur la scène verbale » (ibid). Les crêtes qui peuvent apparaître dans le corps de la mélodie « relèvent d'une évocation ponctuelle sur la scène verbale d'entités ou d'événements qui réclament un effort de la part de l'interlocuteur » (ibid) et les plats correspondent plutôt à « l'évocation d'éléments dont la présence est présumée par le locuteur ». Enfin la montée finale (continuatif) indique une « déstabilisation dans l'interlocution [...] sans apporter de modification à la scène verbale » (ibid).

Si nous reprenons notre exemple en (4) :

(4) It sounds like a simple enough idea but setting up a scheme hasn't been easy.

'L'idée semble assez simple, mais sa réalisation n'a pas été facile.'

<i>It</i>	<i>sounds</i>	<i>like a simple enough idea</i>	<i>but</i>	<i>setting up a scheme hasn't been</i>	<i>easy</i>
	H	*ML%	H		*ML%

on remarque ainsi que la tête sur sounds correspond bien à un cadrage de l'énonciateur qui veut mettre en évidence à la fois le caractère modal (voir plus haut) de son point de vue et la facilité apparente de l'entreprise. Le plat qui suit présente, par sa physionomie régulière, la stabilisation de ce point de vue et l'accès à l'idée de difficulté, et enfin la mise au point sur la courbe finale permet un cadrage sur idea. Ce cadrage final crée aussi, on va le voir, une dynamique. De manière plus précise, le plat contient une crête sur simple qui correspond bien à l'évocation ponctuelle d'une entité qui réclame un effort de la part de l'interlocuteur : c'est effectivement la notion de simplicité qui va être mise en cause dans la suite de l'énoncé. La montée sur but est également un cadrage sur le point de vue dans la mesure où ce connecteur apporte une contradiction. La forme intonative qui suit présente à quelque chose près les mêmes caractéristiques. La tête sur setting correspond à un déplacement sur une entité préconstruite, conjointement au rôle sémantique de -ing. Elle signale aussi une forme de restriction sémantique, entre le projet en apparence simple, et dont la difficulté n'apparaît pas directement, et sa mise en pratique qui s'avère justement difficile. La crête sur scheme marque que l'attention de l'interlocuteur est requise. Scheme renvoie à une entité qui fait écho à idea; leur hauteur (environ 220 hz) est par ailleurs sensiblement équivalente. Ces observations mettent

essentiellement en évidence le rôle des saillances dans l'évolution de la scène. Elles contribuent à montrer et à faire évoluer le point de vue sur elle, point de vue qu'elles concentrent ou au contraire élargissent pour l'interlocuteur.

L'intonation a également un rôle dans l'attribution de dimensions à la scène. Par « dimension », il faut entendre, avec Victorri, les dimensions qualitative et quantitative telles qu'elles sont définies dans la théorie des opérations énonciatives (voir Culioli 1999b). Dans notre exemple, on remarque que a simple enough idea a une valeur quantitative, liée à l'article indéfini a, qui marque l'opération énonciative d'« extraction » dans la perspective de Culioli. Ceci étant, l'énoncé est nuancé par le prédicat sounds like, qui donne le point de vue de l'énonciateur. Le prédicat fait que l'occurrence de la notion /idea/ est placée sur une trajectoire orientée vers l'extérieur du domaine notionnel. Il semble en fait que c'est surtout l'intonation, c'est-à-dire ici la tête de l'unité (sounds) qui donne une dimension qualitative à la relation prédicative. Sound n'est effectivement pas défini en priorité comme un verbe modal comme on l'a vu dans la troisième proposition, mais plutôt comme un verbe de perception. C'est l'attention créée par l'intonation qui contribue à exprimer le point de vue et la modalisation portée simple idea. L'intonation ne se contente pas de souligner le point de vue modal, elle contribue à la créer, dans la mesure où l'intonation n'est pas un simple ajout à l'édifice syntactico-sémantique. Au niveau lexico-grammatical également on trouve un marqueur de modalisation : like. Celui-ci donne à sound sa véritable caractéristique modale et le fait passer d'un verbe de perception à un verbe épistémique portant un point de vue sur la validation de la relation prédicative. Dans la suite de la séquence, a scheme présente les mêmes caractéristiques. L'article indéfini marqueur de l'opération d'extraction, a, donne une dimension quantitative à la scène, comme dans le cas de idea. Mais scheme correspond à une légère remontée de la courbe intonative. Cette légère saillance confère à cette entité une certaine importance, on l'a vu. L'intonation lui donne même une dimension qualitative : son ancrage aspectuo-temporel n'est pas mis à l'avant-plan, ce qui donnerait une dimension quantitative à la scène, c'est son identification qualitative qui semble être en jeu. Scheme est effectivement mis en parallèle avec idea, qui lui aussi est doté d'une saillance. On peut aller jusqu'à dire que ces deux termes renvoient au domaine notionnel de projet, et qu'ils sont mis en résonance par l'intonation. C'est but, lui aussi porteur d'un accent nucléaire, qui les met en relation. Il contribue à faire de simple enough idea une occurrence orientée vers l'extérieur du domaine notionnel¹⁷.

Je voudrais finir par la question de la courbe finale. J'ai suggéré plus haut qu'elle crée une dynamique. Dans la conception de Victorri, la dynamique est surtout liée à la nature des procès, et de leur catégorisation. Il semble qu'en fait, la dynamique de la scène verbale ne tienne pas seulement à la nature des procès. Pour beaucoup de linguistes, l'enjeu de l'intonation est la courbe finale, en particulier dans le cas de l'anglais. Elle indique par exemple la position de l'énonciateur et du co-énonciateur par rapport au domaine notionnel. Ainsi, « les mélodies montantes renvoient à une prise en compte de l'interlocuteur ou du moins sont compatibles avec cet appel [...] Les mélodies descendantes sont centrées sur le locuteur » (Nicaise et Gray 1999:96). On peut aussi parler en termes de stabilisation / déstabilisation dans l'interlocution, comme le font Lacheret-Dujour, Ploux et Victorri à propos du continuatif final. Quels que soient les termes choisis, il est clair que c'est la courbe finale qui crée une relation entre les interlocuteurs et qui donc produit une dynamique sur la scène verbale. Cet espace d'une nature particulière n'est effectivement pas stable, mais en constante stabilisation ; la prosodie, et en particulier l'intonation, contribue à le rendre mouvant, à montrer et cacher des éléments, à les définir ou à les indéterminer, à faire de la scène un véritable espace sémantique où règne la

¹⁷ L'exemple analysé ne permet pas de voir toutes les possibilités de dimensions de la scène. Je me suis concentré sur la dimension qualitative, mais, sans pouvoir le montrer ici, l'intonation devrait donner également des dimensions quantitatives à la scène verbale. La relation saillance / QIt semble toutefois plus évidente, et pourrait être généralisée.

polysémie et qui laisse émerger du sens. La notion d'émergence du sens implique en définitive une double activité d'élaboration et de perception du sens.

Annexes

It sounds like a simple enough idea, but setting up a scheme hasn't been easy

Figure 1 : Visualisation de la courbe intonative (réalisée avec le logiciel PRAAT)

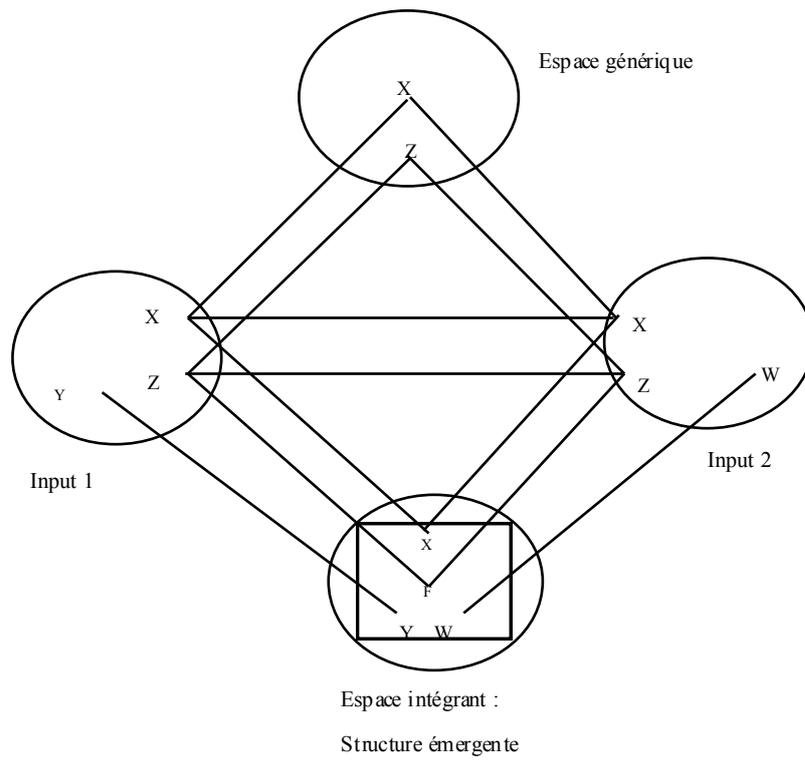


Figure 2 : schéma de l'Intégration Conceptuelle, d'après G. Fauconnier et M. Turner.

Références :

- Bacri, Nicole. 1994. Valeur fonctionnelle de l'intonation et acquisition du langage. *Cahiers d'Acquisition et de Pathologie du Langage* 11:51-61.
- Bonnet, Claude. 1995. Processus cognitifs de la perception. In *Perception et agnosies*, sous la dir. de Bernard Lechevallier, Francis Eustache et Fausto Viader, 31-54. Bruxelles : De Boeck.
- Col, Gilles. 2001. Infinitif en TO et accès à une scène verbale. *Anglophonia* 10:109-123.
- Col, Gilles. 2003. Style indirect libre et intégration conceptuelle. In *Bulletin de la Société de Stylistique Anglaise*, numéro spécial : *Stylistique et énonciation : le Cas du discours indirect libre*, sous la dir. de Gilles Mathis, Monique De Mattia-Viviès, et Claire Pégon, 185-203.
- Col, Gilles. 2004. Théories cognitives et l'hypothèse de l'émergence du sens. *Tropismes* 12:115-140.
- Col, Gilles. 2006. Sous presse. Relation ou intégration prédicative ? La prédication comme principe d'émergence du sens. In *La Prédication*, sous la dir. de Jean-Marie Merle, *Bibliothèque de Faits de Langues*, Paris : Ophrys.
- Col, Gilles. A paraître. Blending, Windowing and Access in Future Time Evocation. In *From Gram to Mind : Grammar as Cognition*, sous la dir. de Jean-Rémi Lapaire, Guillaume Desagulier et Jean-Baptiste Guignard. 2007. Bordeaux: Presses Universitaires de Bordeaux.
- Col, Gilles et Duchet, Jean-Louis. 2000. Eléments pour une définition des valeurs de la forme *gonna* en anglais, à partir du corpus Électronique C.O.L.T. *Cahiers FORELL* 14:167-186.
- Cruttenden, Alan. 1994. *Gimson's Pronunciation of English*. London : Arnold.
- Culioli, Antoine. 1999a. *Pour une linguistique de l'énonciation*, vol. 2. Paris : Ophrys.
- Culioli, Antoine. 1999b. Des façons de qualifier. In *Les Opérations de détermination*, sous la dir. d'Alain Deschamps et Jacqueline Guillemain-Flescher, Paris : Ophrys. 3-12.
- Duchet, Jean-Louis. 1993. La coupe syllabique en anglais. In *Wodan*, Bd. 8 : *Études de linguistique et de littérature en l'honneur d'André Crépin*, sous la dir. de Danielle Buschinger et W. Spiewok eds. Greifswald : Reinneke Verlag.
- Eustache, Francis. 1995. Identification et discrimination auditive : données neuropsychologiques. In *Perception et agnosies*, sous la dir. de Bernard Lechevallier, Francis Eustache et Fausto Viader, 243-271. Bruxelles : De Boeck.
- Fauconnier, Gilles. 1984. *Espaces mentaux. Aspects de la construction du sens dans les langues naturelles*. Paris : Editions de Minuit.
- Fauconnier, Gilles. 1994. *Mental Spaces*. Nouv. éd. rév. New York : Cambridge University Press. [1985]
- Fauconnier, Gilles. 1991. Subdivision cognitive. *Communications* 53:229-248.
- Fauconnier, Gilles. 1997. *Mappings in Thought and Language*. Cambridge : Cambridge University Press.
- Fauconnier, Gilles et Sweetser, Eve. 1996. *Spaces, Worlds and Grammar*. Chicago : The University of Chicago Press.
- Fauconnier, Gilles et Turner, Mark. 1998. Conceptual Integration Networks. *Cognitive Science* 22(2):133-187.
- Garnier, George et Guimier, Claude. 1997. *L'Épreuve de linguistique au Capes et à l'agrégation d'anglais*. Paris : Nathan.
- Gaudy, Isabelle. 2002. Le question-tag descendant : juxtaposition de deux unités ? *Travaux du C.I.E.R.E.C.* 107:169-183.
- Gosselin, Laurent. 1996. *Sémantique de la temporalité en français*. Louvain : Duculot.
- Gosselin, Laurent. 2005. *Temporalité et modalité*. Louvain : Duculot.
- Guillaume, Paul. 2000. *La Psychologie de la forme*. Paris : Flammarion. [1937].
- Hirst Daniel J. et Albert Di Cristo. 1986. Unités tonales et unités rythmiques dans la représentation de l'intonation, In *XVèmes Journées d'Étude sur la Parole (GALF)*, 93-95. Université d'Aix en Provence.

- Hockett. 1942. A System of Descriptive Phonology. *Language*, 18:3-21
- Huart, Ruth. 2002. *Grammaire orale de l'anglais*. Paris : Ophrys.
- Lacheret-Dujour, Anne, Sabine Ploux et Bernard Victorri. 1998. Prosodie et thématization en français parlé. *Cahiers de Praxématique* 30:89-111.
- Ladd, Robert. 1996. *Intonational Phonology*. Cambridge : Cambridge University Press.
- Lambrecht, Knut. 1994. *Information Structure and Sentence Form*. Cambridge : Cambridge University Press.
- Lechevallier, Bernard, Francis Eustache et Fausto Viader, dir. 1995. *Perception et agnosies*. Bruxelles : De Boeck.
- Lieberman, Mark. 1975. *The Intonational System of English*. Ph Dissertation, Massachusetts Institute of Technology.
- Mehler, Jacques, Emmanuel Dupoux et Juan Segui. 1990. Constraining Models of Lexical Access : the Onset of Word Recognition. In *Cognitive Models of Speech Processing. Psycholinguistic and Computational Perspectives*, sous la dir. de Gerry T. Altmann, 236-262. Cambridge, Mass. : The M.I.T. Press
- Morel, Marie-Annick et Laurent Danon-Boileau. 1998. *Grammaire de l'intonation. L'exemple du français*. Paris : Ophrys.
- Nicaise, Alain. 1987. *Phénomènes intonatifs en français. De la perception à l'interprétation*. Thèse de Doctorat d'Etat, Université Paris 7.
- Nicaise, Alain. 2000. Pourquoi l'intonation ? In *Points d'interrogation. Phonétique et phonologie de l'anglais*, sous la dir. de Pierre Busuttil, 111-133. Pau : Presses Universitaires.
- Nicaise, Alain et Mark Gray. 1999. *L'Intonation de l'anglais*. Paris : Nathan.
- Nølke, Henning. 1994. *Linguistique modulaire : de la Forme au sens*. Louvain, Paris : Peeters.
- Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation*. Ph Dissertation, Massachusetts Institute of Technology.
- Pierrehumbert, Janet et Julia Hirschberg. 1990. The Meaning of Intonational Contours in the Interpretation of Discourse. In *Intentions in Communication* sous la dir. de Jerry Morgan, Philip R. Cohen et Martha E. Pollack, 271-311. Cambridge, Mass. : The M.I.T. Press.
- Rosenthal, Victor et Yves-Marie Visetti. 2003. *Köhler*. Paris : Les Belles Lettres.
- Rossi, Mario. 1999. *L'Intonation, le système du français : description et modélisation*. Paris : Ophrys.
- Segui, Juan, Emmanuel Dupoux et Jacques Mehler. 1990. The Role of the Syllable in Speech Segmentation, Phoneme Identification and Lexical Access. In *Cognitive Models of Speech Processing. Psycholinguistic and Computational Perspectives*, sous la dir. de Gerry T. Altmann, 263-280. Cambridge, Mass. : The M.I.T. Press.
- Segui, Juan. 1992. Perception du langage et modularité. In *Introduction aux sciences cognitives*, sous la dir. de Daniel Andler, 131-152. Paris : Gallimard.
- Silverman, Kim et al. 1992. TOBI : A Standard for Labeling English Prosody. In *Proceedings of the 1992 International Conference on Spoken Language Processing*, Volume 2 : 867-870.
- Victorri, Bernard. 1999. Le sens grammatical. *Langages* 136:85-105.
- Victorri, Bernard et Catherine Fuchs. 1996. *La Polysémie*. Paris : Hermès.