



HAL
open science

Sur les aspects multicritères et multidimensionnels de la saillance

Frédéric Landragin

► **To cite this version:**

Frédéric Landragin. Sur les aspects multicritères et multidimensionnels de la saillance. Maryvonne Boisseau; Albert Hamm. Saillance. La saillance en langue et en discours, Volume 2, 940, Presses universitaires de Franche-Comté, pp.15-29, 2015, Annales Littéraires de l'Université de Franche-Comté, 978-2-84867-513-8. <halshs-01154002>

HAL Id: halshs-01154002

<https://shs.hal.science/halshs-01154002v1>

Submitted on 20 May 2015

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.



HAL Authorization

Sur les aspects multicritères et multidimensionnels de la saillance

Frédéric Landragin
CNRS / ENS – Laboratoire Lattice
1, rue Maurice Arnoux
92120 Montrouge, France
frederic.landragin@ens.fr

DRAFT

Résumé :

De plus en plus considérée comme une notion multicritère, la saillance se caractérise par une mise en relief via des procédés prosodiques, syntaxiques, sémantiques... La diversité des facteurs invoqués et surtout des analyses d'un énoncé en termes de saillance conduit à envisager plusieurs échelles de saillance qui œuvrent en parallèle. Contrairement à la Théorie du Centrage et aux approches qui considèrent qu'à chaque instant il n'existe qu'un seul centre – qu'une seule entité saillante –, nous explorons les dimensions d'analyse permettant de modéliser plusieurs échelles de saillance simultanées. Cette modélisation fait de la saillance une notion multicritère et multidimensionnelle, ce qui entraîne des difficultés d'ordre méthodologique pour lesquelles nous amorçons une discussion. Avec un objectif à long terme de concrétisation de notre proposition, nous présentons quelques pistes de modélisation.

Abstract:

The notion of salience is more and more considered as a multi-criteria concept. It highlights a discourse entity with the help of prosodic, syntactic, and semantic features (among others). The diversity of salience features as well as the diversity of salience-oriented linguistic analyses leads us to think about several salience hierarchies that work together. Unlike Centering Theory and similar approaches that take only one center – one salient entity – into account at any time, we explore some dimensions of analysis in order to apprehend several simultaneous saliencies. This proposition makes salience a multi-criteria and multi-dimensional concept, which raises methodological difficulties. We discuss them and we present some modeling concerns, with the long-term aim of materializing our propositions.

1. Introduction

Comme le précédent colloque sur la saillance l'a mis en avant, la saillance est de plus en plus considérée comme un concept multicritère (Schnecker, 2009), qui repose à la fois sur des aspects prosodiques, lexicaux, morphologiques, syntaxiques, sémantiques, pragmatiques, ou encore stylistiques. Pour chacun de ces niveaux d'analyse du langage, plusieurs listes de critères ont été proposées, certains critères recouvrant parfois plusieurs niveaux (Landragin, 2004), le but étant de caractériser la notion de saillance en s'aidant du plus grand nombre possible de facteurs qui contribuent directement ou indirectement à mettre en avant une entité

linguistique par rapport aux autres. Une telle caractérisation permet ainsi d'identifier à tout moment l'entité du discours la plus saillante, c'est-à-dire celle qui est mise en relief le plus efficacement. Pour le moment, nous ne faisons qu'étendre la notion de centre de la Théorie du Centrage (Grosz *et al.*, 1995) et le rôle de la saillance tel qu'il est décrit dès la thèse de Candace L. Sidner (Sidner, 1979) : il s'agit d'identifier à toutes les étapes de la prise de connaissance d'un texte quel est le centre – l'entité la plus saillante –, et éventuellement quel est le classement – selon un critère de saillance – des entités mises en présence dans le texte.

Mais ne peut-il y avoir, selon le type de communication, selon le niveau d'analyse du langage, selon d'autres paramètres à déterminer, plusieurs saillances qui opèrent en parallèle ? Plutôt qu'une échelle de saillance classant les différentes entités du discours, nous aurions alors plusieurs échelles de saillance, chacune d'entre elles apportant un point de vue spécifique sur l'importance relative des entités. Ce sont ces aspects multidimensionnels que nous voulons étudier et discuter ici. Comment gérer plusieurs échelles de saillance ? Comment détermine-t-on les différentes échelles ? Est-ce selon les types d'unité : une échelle pour les référents, une autre pour les événements, etc. ? Est-ce selon les plans d'analyse : une échelle pour la syntaxe, une autre pour la sémantique, une encore pour la stylistique ? Est-ce selon les effets interprétatifs : une échelle pour les saillances globales qui perdurent tout au long du texte, une autre pour les mises en saillance locales ? Est-ce selon le genre textuel ?

Dans les sections 2 à 6, nous proposerons cinq échelles de saillance qui constitue notre modèle multidimensionnel de la saillance. Cette proposition tente de distinguer plusieurs facettes fonctionnelles de la notion qui nous intéresse, et met à plat ou regroupe des facteurs qui sont parfois de nature hétérogène. Les échelles qui en résultent ne sont pas classées par ordre d'importance, et il se peut que l'une d'entre elles joue un rôle beaucoup plus déterminant que les autres pour certains types de textes voire pour certains domaines de la linguistique. Ces aspects seront discutés dans la section 7, avec une illustration sur un exemple construit de l'application des cinq dimensions de saillance. Nous reviendrons sur quelques problèmes méthodologiques posés par la définition de la saillance en tant que concept multicritère et multidimensionnel : il est en effet tentant d'intégrer le maximum de facteurs sous le terme de « saillance », et de lui faire éventuellement perdre ainsi toute limite, donc toute possibilité de modélisation (cognitive, formelle ou computationnelle). Suite à ce constat, nous proposerons dans la section 8 quelques pistes pour la modélisation voire la formalisation de nos cinq dimensions de saillance, puis nous présenterons nos conclusions et nos perspectives de recherche autour de cette notion.

2. Saillance visuelle et saillance linguistique

Une première dimension repose sur la modalité de communication, avec une distinction entre la saillance d'une entité visuellement perçue, saillance qui influe sur la production linguistique des participants à la communication, et la saillance reposant sur des choix purement linguistiques. Les unités concernées sont les objets visibles pour la saillance visuelle et les entités du discours pour la saillance linguistique, avec des facteurs couvrant plusieurs niveaux d'analyse du langage (qu'il soit visuel ou verbal). Au niveau de la perception visuelle, les facteurs de saillance peuvent être intrinsèques aux unités (par exemple : objet visuel simple et lisible, proche de la notion de bonne forme de la Théorie de la Gestalt), syntaxiques (disposition des objets dans la scène visuelle ; placement à un endroit stratégique ; présence de symétries ou de répétitions) ou encore sémantiques : chaque objet apporte ses propres traits sémantiques qui contribuent à la sémantique du message visuel dans

son ensemble. Pour continuer l'analogie avec les niveaux d'analyse du langage, nous noterons également que certains facteurs comme un éclairage particulièrement dirigé sur un objet œuvrent à la manière de la prosodie. Au niveau du langage, on trouve des facteurs de saillance aussi bien dans les aspects prosodiques (proéminence, rupture dans le rythme d'élocution, intonation particulière), lexicaux (saillance intrinsèque des noms propres, manque d'autonomie référentielle des déictiques), syntaxiques (construction détachée, clivée, présentatif, position initiale, fonction grammaticale sujet), que dans les nombreux aspects sémantico-pragmatiques : sémantique lexicale, rôles thématiques, thèmes et rhèmes, topiques et commentaires, posés et présupposés, etc.

Dans notre modèle de saillance, cette première dimension est relativement facile à appréhender car elle est liée au type de signal, au type de perception. Un critère est visuel quand il repose sur le signal visuel ; un critère est linguistique quand il repose sur les traces linguistiques : le signal audio, le texte lui-même, éventuellement des contenus implicites, du moins si l'on se limite à des contenus directement inférables des traces linguistiques, comme c'est le cas avec les présuppositions les plus immédiates. Tout devient plus complexe lorsque l'on considère que le contexte de communication est multiforme : dans des exemples désormais classiques d'exophores tels que « *elle ne va pas rentrer* » en parlant d'une table que les interlocuteurs essaient de faire rentrer dans un coffre de voiture, ou « *attention, il risque de te mordre* », énoncé adressé à un enfant en train de poser ses mains sur les barreaux d'une cage où est enfermé un lion, la saillance opère à la fois au niveau visuel et au niveau linguistique. Au niveau visuel parce que le pronom « *elle* » ou « *il* » trouve un antécédent non pas linguistique mais visuel. Au niveau linguistique parce que l'énoncé, avec sa structure syntaxique et l'ensemble des choix qu'il matérialise, opère éventuellement une mise en relief linguistique, celle liée au terme « *attention* » par exemple. Tout devient encore plus complexe lorsqu'on considère que nos représentations cognitives sont multimodales et que la lecture d'un texte peut entraîner la construction mentale d'une scène visuelle. Dans un extrait de roman qui décrit un lieu, peut-on parler de saillance visuelle ? Le roman peut jouer avec cet aspect, en faire par exemple un élément de l'intrigue. Sans aller jusque-là, nous noterons que toute phrase lue peut évoquer en nous des représentations visuelles parfois précises, qui contribuent à la saillance des différents référents ou événements en présence (et qui donc s'additionnent à la saillance purement linguistique de ces entités du discours).

La distinction entre visuel et linguistique peut s'étendre à d'autres modalités de perception (tactile, olfactive, gustative, gestuelle) selon le type de communication considérée. Si la communication fait intervenir la manipulation d'objets en plus de leur perception visuelle, il est possible que les participants exploitent dans leurs messages cette dimension supplémentaire de saillance. C'est ce que nous avons exploré au niveau de la référence aux objets dans un système de dialogue homme-machine faisant intervenir un dispositif haptique à retour d'effort (Landragin *et al.*, 2002). Dans ce contexte, l'utilisateur du système pouvait exploiter dans les énoncés oraux qu'il produisait un élément (élément de texture, présence d'un objet) perçu via le retour d'effort. Autrement dit, l'exophore avec antécédent haptique était tout à fait possible, et par conséquent la saillance haptique avait un rôle essentiel, en plus des saillances visuelle et linguistique. Par ailleurs, une facette supplémentaire de la saillance est explorée dans la Langue des Signes Française, avec des facteurs propres à la nature gestuelle de cette langue et des langues signées en général (Blondel, 2003). Pour la suite de cet article, nous mettrons surtout en avant la facette linguistique de la saillance.

3. Saillance physique et saillance cognitive

P-saillance (saillance physique) et C-saillance (saillance cognitive) : une deuxième dimension de notre modèle distingue la forme du contenu, avec d'un côté des aspects de saillance détectables via une trace physique contenue dans le signal en tant que matériau physique, et de l'autre côté – plus immatériel – des aspects de saillance relevant d'inférences. Ces inférences se fondent sur l'interprétation du message par le sujet, avec des facteurs qui lui sont propres tels que l'attention, l'intention, la mémoire, les affects, et d'une manière générale tout ce qui est implicite mais fait partie du message communiqué et permet de l'expliquer en contexte. On est ici au-delà des présuppositions les plus immédiates, et la saillance cognitive fait plutôt intervenir les présuppositions que certains sujets font et d'autres pas, ou encore les sous-entendus et tous ces messages cachés que seuls certains sujets reçoivent. Il s'agit donc d'une saillance subjective, qui est par conséquent très difficile à formaliser : pour « calculer » la saillance cognitive, il est nécessaire de faire des hypothèses sur les états mentaux du sujet en plus d'inférences pragmatiques.

Cette distinction entre deux types de saillance a fait l'objet d'une précédente publication (Landragin, 2004) et a été à l'origine d'une classification systématique des facteurs de saillance selon leur degré d'abstraction. En complément, nous noterons que cette dimension peut s'étendre à ce que nous appellerons la saillance située (S-saillance) dans les situations de communication impliquant plusieurs interlocuteurs et l'ensemble des relations qu'ils entretiennent les uns avec les autres. Le terme « situé » étant sujet à des acceptions différentes selon les communautés qui l'utilisent, nous retiendrons qu'il s'agit ici d'une saillance contextuelle au sens large, qui prend en compte non seulement la pertinence de l'énoncé linguistique compte tenu du contexte immédiat et des connaissances partagées par les interlocuteurs – autrement dit des aspects inclus traditionnellement dans le champ de la pragmatique –, mais aussi d'une mise en perspective de cet énoncé compte tenu des relations sociales, socioprofessionnelles, psychologiques et relationnelles des interlocuteurs. Ainsi, une marque de politesse (ou son absence) peut être particulièrement saillante dans une situation de communication donnée entre des interlocuteurs entretenant des relations particulières. Plus généralement, tout acte de communication peut avoir une S-saillance de par ses conséquences sur les relations qu'entretiennent les interlocuteurs. Un énoncé « neutre » du point de vue linguistique peut fonctionner comme étape décisive dans les aspects relationnels, et être ainsi doté d'une forte S-saillance.

Dans le cas d'une modélisation computationnelle, seule la P-saillance peut être prise en compte de manière fiable : les C-saillance et S-saillance ne peuvent être modélisées que partiellement, par le biais d'hypothèses et d'inférences réalisées à partir des traces matérielles et d'éventuelles connaissances préalables sur les sujets.

4. Saillance à effet immédiat et saillance à effet continu

Une troisième dimension concerne la durée de l'effet de saillance, en opposant notamment la saillance rapide et immédiate (celle dénotée par ce qu'on appelle l'effet *pop-up*) à la saillance qui se construit petit à petit, de manière incrémentale, telle que celle qui s'applique au personnage principal d'un roman. Certains auteurs comme René Thom (Thom, 1993) utilisent le terme de saillance pour l'effet immédiat, et le terme de prégnance (impliquant une *imprégnation*) pour l'effet continu, bien que chez René Thom il soit difficile de présenter ainsi saillance et prégnance, dans la mesure où ces concepts sont décrits de manière beaucoup

plus complexe. On peut aussi parler, comme (Deleuze, 1969), de différence et de répétition, la saillance émergeant souvent par différence et la prégnance par répétition. Comme l'écrit par ailleurs Pierre Larrivée dans sa présentation de l'approche dynamique du langage de Bernard Pottier, « est latent tout ce qui dans les objets du monde est susceptible d'être perçu, est saillant dans les objets de la pensée tout élément qui se distingue nettement sur un fond pour un observateur et prégnant l'élément qui *pré*occupe cet observateur » (Larrivée, 2008).

La saillance à effet immédiat s'applique à un seul acte de communication, c'est-à-dire à une unité énonciative ou pragmatique unique. En considérant une prédication, elle relève alors par exemple de l'opposition entre thème et rhème. La saillance à effet continu est plus complexe et peut faire intervenir des macrostructures ou des superstructures. Elle se construit sur plusieurs actes de communication, et intervient donc dans le discours et la conversation (van Dijk & Kintsch, 1983).

Cette troisième dimension fait ainsi intervenir la taille des unités auxquelles elle s'applique, de l'unité communicative la plus petite à des portions de textes beaucoup plus conséquentes. Pour continuer dans cette voie, on peut étendre l'effet continu et lui adjoindre les cas où la saillance s'applique à un texte entier, par exemple un roman. C'est un peu l'approche de Karine Abiven avec le cas de l'anecdote (Abiven, 2010), et que nous pouvons appliquer par exemple à la totalité de la production romanesque d'un auteur. A titre d'illustration probablement discutable, la production du romancier américain Stephen King inclut un grand nombre de romans qui se terminent pour la plupart plutôt mal, mais sans faire mourir le personnage principal auquel le lecteur s'est attaché, ou alors c'est envisagé voire annoncé dès le début de l'histoire. Dans ce corpus, le roman *Cujo* dénote complètement et surprend ainsi les amateurs de Stephen King. Ce roman en devient saillant.

5. Saillance préalable et saillance nouvelle

Une quatrième dimension oppose l'exploitation d'une saillance existante à la mise en saillance d'une nouvelle entité. L'exploitation d'une saillance existante intervient notamment lors d'une production linguistique fondée sur une perception visuelle, par exemple lors d'une référence à un objet visible, avec les cas particuliers de pronoms exophoriques que nous avons déjà mentionnés. Elle permet également d'expliquer l'utilisation des pronoms anaphoriques. Dans les deux cas, la saillance de l'entité concernée est maintenue et il n'y a pas de nouvelle saillance. A l'inverse, la mise en saillance d'une nouvelle entité relève d'un mécanisme de composition du discours qui permet à une entité d'être mise en valeur, afin d'être reprise aisément dans la suite, en préparation d'une action future. Un procédé classique de mise en valeur est l'emploi d'une construction syntaxique dédiée, avec comme exemples les présentatifs ou encore certaines topicalisations. L'entité concernée devient ainsi saillante, quelles que soient les saillances préalables. Une mise en saillance, lorsqu'elle est très marquée, peut néanmoins annuler ou au moins atténuer les saillances préalables.

Cette distinction entre préalable et nouveau est essentielle dans la mesure où elle correspond à la différence fondamentale entre perception visuelle et langage : compte tenu de la linéarité du message linguistique, de son apparition progressive, et de la nécessité de faire constamment des choix lors de sa production, tout ce que l'on dit est construit et n'est pas neutre. Les saillances préalables et les mises en saillance sont permanentes. Elles peuvent même être involontaires et à l'origine de mécompréhensions, voire d'insatisfactions du locuteur par rapport à ce qu'il vient de dire et dans lequel il repère a posteriori des effets de saillance

indésirables ou mal contrôlés. Au contraire, tout ce que l'on voit n'est pas construit, est perçu immédiatement, n'est pas tributaire de choix, et ne fait donc pas forcément intervenir de saillance. L'analogie entre perception visuelle et langage fonctionne pour certains facteurs de saillance mais pas pour les modalités d'intervention de la saillance dans la production d'un message. Par contre, une autre analogie peut s'avérer plus parlante : celle entre le langage et une machine de Turing. Comme dans une machine de Turing ou plus simplement dans tout langage informatique de type procédural, on peut considérer chaque marqueur linguistique, qu'il s'agisse d'un mot, d'un syntagme, d'un morphème ou d'une phrase, comme un opérateur (ou une instruction) qui prend des données en entrée et qui produit en sortie d'autres données. Les saillances préalables nécessaires pour que le marqueur soit compris constituent les données en entrée, et les nouvelles saillances constituent les données en sortie. On peut alors classer les marqueurs en plusieurs catégories selon qu'ils contribuent surtout à exploiter des saillances préalables ou à produire de nouvelles saillances. D'une manière générale, les marqueurs les plus courts, comme les pronoms ou les partitifs, relèvent de l'exploitation d'une saillance préalable, alors que les marqueurs les plus longs indiquent plutôt un changement de saillance. Une revue détaillée des marqueurs s'avérerait nécessaire, par exemple dans le cas des expressions référentielles pour modéliser les évolutions de saillance des référents d'un texte, mais il semble qu'il y ait là une démarche linguistique réalisable de manière efficace.

6. Saillance informative et saillance rhétorique

Une cinquième dimension oppose les effets interprétatifs liés à l'apport d'information aux effets interprétatifs liés à la rhétorique. L'apport d'information intervient avant tout lors de l'apparition d'une nouvelle entité discursive : un élément nouveau est saillant de par l'information qu'apporte cette nouveauté, et l'on retrouve ici la distinction classique entre nouveau et déjà connu liée notamment à la distinction entre rhème et thème. De son côté, la saillance rhétorique relève de mécanismes de mise en scène et de composition du discours, et implique plusieurs types de saillance :

- la saillance liée à la mise en avant d'un argument fort, et plus généralement aux structures argumentatives (invention, disposition, élocution) ;
- la saillance liée à la poétique et aux effets des figures de style ;
- la saillance liée à la rhétorique normative ou prescriptive.

Cette dimension peut s'étendre à la stylistique de manière générale. Dans tous les cas, les aspects couverts sont nombreux, posant peut-être plus qu'ailleurs la question des limites de la saillance.

7. Discussion

Une fois posées ces dimensions, un premier constat (qui confirme de précédents constats) est que la saillance intervient dans tous les niveaux d'analyse du langage, dans toutes ses dimensions, et que les facteurs impliqués sont multiformes et parfois complexes. L'utilisation du langage implique constamment des mises en relief, et le terme de saillance convient bien pour regrouper ces aspects difficilement modélisables. Pour simplifier le problème, on peut s'intéresser à une dimension particulière et oublier les autres, au risque d'amputer la notion. Dans certains domaines de la linguistique, il est vraisemblable qu'oublier quelques dimensions permettra de travailler avec un modèle plus facile à appréhender, sans pour autant dénaturer complètement la notion. On peut aussi faire des raccourcis pratiques, comme celui consistant à élaguer les multiples cas de saillance portés par cette modélisation, pour ne garder

que quelques repères. (Sur la base de 5 dimensions au minimum binaires, il y a théoriquement au moins 32 cas de saillance qui s'avèrent de toute façon impossibles à conceptualiser). Par repères, nous entendons des régularités qui permettent de réaliser quelques rapprochements et ainsi de simplifier le modèle, avec par exemple :

- des liens entre différentes dimensions : par exemple, « saillance informative » (dimension 5) entraîne forcément « mise en saillance » (dimension 4), l'inverse n'étant pas vrai ;
- la détermination des recouvrements fréquents de dimensions pour un domaine donné : par exemple, il se peut que la dimension rhétorique revienne souvent à une mise en saillance, et se traduise par des facteurs physiques plutôt que cognitifs ;
- la redéfinition de notions sujettes à débat en les positionnant par rapport aux dimensions de notre modèle : par exemple, la redéfinition de la notion de « propos » telle qu'utilisée par Anne Grobet (Grobet, 2002), ou celle des notions de thème et de rhème avec des résultats variables selon l'école choisie, l'illustration ci-dessous mettant en avant la distinction entre connu et nouveau :
 - propos linguistique = C-saillance linguistique à effet continu ;
 - propos au sens large (visuel ou linguistique) = C-saillance à effet continu ;
 - thème = C-saillance linguistique préalable à effet immédiat ;
 - rhème = C-saillance linguistique nouvelle à effet immédiat ;
- la détermination des limites d'intervention des dimensions, ou du moins des dimensions privilégiées dans un domaine particulier :
 - pour le texte poétique = P-saillance avant toute autre forme de saillance ;
 - pour le texte de presse = saillance informative à effet immédiat avant toute autre forme ;
 - pour le texte de loi = inhibition de toute forme de saillance (dans la mesure du possible).

Néanmoins, restent cinq dimensions de saillance qui constituent autant de plans d'analyse face à un corpus donné. Ce principe implique cinq échelles de saillance, et, à long terme, pourrait se matérialiser par un système informatique capable de gérer simultanément cinq échelles de scores de saillance. Mais arriverait-on à exploiter de manière pertinente autant d'information ? A titre d'illustration, une phrase telle que « *attention, il risque de te mordre* », déjà mentionnée ci-dessus, relève, pour l'interprétation du pronom exophorique « *il* », de :

1. la saillance visuelle ;
2. la saillance cognitive ;
3. la saillance à effet immédiat ;
4. la saillance préalable ;
5. la saillance informative.

Il s'agit bien de plusieurs effets qui opèrent en parallèle, et il devient difficile de parler d'une saillance unique s'appliquant au référent « lion ». Maintenant, les cinq dimensions proposées sont-elles nécessaires ? Il nous semble qu'à partir du moment où l'on appréhende la saillance comme un concept multifactoriel, supprimer une de ces dimensions revient à restreindre le langage dans son pouvoir d'expression et dans ses effets interprétatifs. Cette remarque a bien entendu plus d'intérêt pour des exemples faisant intervenir plusieurs référents. Prenons le cas d'un roman qui fait intervenir un certain Pierre, et la phrase « *Pierre, son vélo est cassé* » qui apparaît après la présentation de Pierre et avant toute évocation de son vélo :

1. Première dimension : l'exemple relève de la saillance linguistique mais pas de la saillance visuelle car ni Pierre ni son vélo ne sont physiquement présents (sauf à

considérer une scène visuelle construite en temps réel par le texte, comme on l'a évoqué dans la section 2).

2. Deuxième dimension : la P-saillance se traduit par la mise en avant de Pierre (placement stratégique en début d'énoncé, dislocation à gauche) plutôt que de son vélo (fonction grammaticale sujet, mais aucun autre facteur vraiment déterminant). La C-saillance intervient car Pierre est activé dans la mémoire à court terme.
3. Troisième dimension : sans autre précision, il s'agit d'un effet immédiat. Si la narration détaille la vie de Pierre, toute mention supplémentaire contribue à le rendre saillant et on est alors dans l'effet continu.
4. Quatrième dimension : l'énoncé relève surtout de la saillance nouvelle avec la mise en thème de Pierre, qui pourra ensuite être repris par le pronom « *il* » plus facilement que le vélo. « *Pierre* » ne faisant pas appel à un antécédent linguistique (même si celui-ci existe), l'exemple ne fait pas intervenir de saillance préalable.
5. Cinquième dimension : l'exemple relève de la saillance informative avec l'introduction du vélo et du fait qu'il est cassé (rhème en tant qu'information nouvelle). Il relève par ailleurs de la saillance rhétorique avec la dislocation à gauche.

Ainsi, parmi les deux référents de cette phrase, le personnage Pierre est concerné par au moins trois dimensions et le vélo par au moins deux dimensions de saillance. Les deux référents sont donc affectés chacun d'une certaine saillance. Comme les dimensions opèrent en parallèle et rendent compte de différents effets des mécanismes de construction de discours, ces deux saillances peuvent, selon les dimensions sur lesquelles elles s'appliquent, se confronter ou influencer simultanément sur des plans différents, ce qui est tout l'intérêt d'une modélisation multidimensionnelle.

8. Premiers éléments de modélisation

Compte tenu du nombre de paramètres, il est illusoire d'aboutir à court terme à une modélisation formelle de la notion de saillance telle qu'elle est présentée ici. Notre but est pour l'instant de mettre en face de chaque dimension de saillance quelques pistes envisageables pour sa modélisation, et notamment sa quantification, le but à long terme étant d'intégrer aux modèles computationnels existants un calcul supplémentaire pour rendre compte des mises en relief opérées par la saillance. Si l'on considère par exemple la SDRT (Asher & Lascarides, 2003), théorie dans laquelle les contenus des phrases et les relations de discours sont représentés à l'aide d'outils logiques (variables, prédicats, etc.), une possibilité de prise en compte de notre modèle serait d'ajouter aux représentations de certaines phrases un voire cinq prédicats supplémentaires dédiés à la saillance.

Pour calculer le prédicat correspondant à notre première dimension de saillance, il est nécessaire de mettre en place une classification hiérarchique des facteurs de saillance linguistique et, d'une manière générale, des facteurs liés à chaque type de modalité perceptive intervenant dans la communication. Nous avons vu que ces facteurs sont de natures diverses. Qu'ils soient scalaires ou binaires (opposition entre deux valeurs possibles), leur résultat aboutit à un score normé entre 0 (aucune saillance) et 1 (saillance la plus forte), des valeurs intermédiaires étant possibles pour les facteurs scalaires. Comme la saillance fait intervenir plusieurs facteurs, nous faisons face à un ensemble de plusieurs scores entre 0 et 1. Or tous les facteurs n'ont pas la même contribution à la saillance. Un système de pondération, ou au moins un système gérant une priorité (à l'instar de la Théorie de l'Optimisation) s'avère donc utile : à chaque facteur est affecté un poids, et le calcul de saillance se fait en calculant la contribution de chacun des facteurs à la saillance de l'unité linguistique concernée. Un enjeu

réside dans la détermination des poids : souvent posés de manière intuitive dans les approches qui suivent ce principe, il faudrait au contraire les déduire d'expérimentations et d'études de corpus. Pour calculer le prédicat modélisant notre deuxième dimension, nous avons d'une part la P-saillance qui se rapproche de la méthodologie précédente avec le calcul, sur la base des traces physiques, de scores de saillance à partir d'une liste pondérée de facteurs, et d'autre part la C-saillance et la S-saillance pour lesquelles nous pouvons envisager un système de règles s'appliquant aux mêmes traces physiques. Ce système de règles vise à inférer des hypothèses sur les informations implicites et sur les états mentaux des interlocuteurs. La modélisation qui en résulte permet de mieux rendre compte des effets interprétatifs des messages échangés, en intégrant une formalisation de ces effets aux représentations logiques utilisées de manière plus classique pour les messages eux-mêmes. Le prédicat pour la troisième dimension exploite à nouveau la méthode d'une liste pondérée de facteurs. Ces facteurs de saillance peuvent être les mêmes que dans les dimensions précédentes, seules les pondérations variant de manière à obtenir un résultat qui reflète toute l'importance donnée aux procédés opérant dans le sens d'un effet immédiat ou d'un effet continu. Ce principe s'applique également au calcul du prédicat correspondant à la quatrième dimension, avec un ensemble de pondérations spécifiques. Pour calculer le prédicat correspondant à la cinquième dimension, il est en revanche nécessaire de mettre en œuvre des processus particuliers : d'une part la détection des premières apparitions de chacune des entités pour ce qui concerne la saillance informative, et d'autre part l'exploitation d'une bibliothèque de structures de phrases types pour ce qui concerne la saillance rhétorique. Cette bibliothèque regrouperait et stockerait l'ensemble des structures argumentatives, l'ensemble des figures de style, etc., ce qui permettrait au module d'analyse de mettre en rapport un énoncé avec une structure ou une figure connue, cette mise en rapport conduisant à la détermination d'un prédicat rendant compte de l'effet interprétatif lié à la structure ou la figure en question.

9. Conclusion et perspectives

La majeure partie du travail présenté ici relève d'une conceptualisation de la notion de saillance. Celle-ci vise à intégrer l'ensemble des facteurs intervenant dans la mise en relief d'une entité linguistique, qu'il s'agisse d'un référent, d'un événement, voire d'un texte entier. Les facteurs évoqués sont ainsi nombreux, de même que les points de vue du linguiste sur chacun de ces facteurs et sur leurs conséquences au niveau de la réception du message, de son interprétation, ou encore de sa mémorisation. Face à cette multiplicité des interventions de la saillance, nous en avons fait une notion multicritère et multidimensionnelle. Notre proposition de modélisation selon cinq dimensions principales ne s'attache pas directement aux types des unités auxquelles s'applique la saillance (référents, événements, etc.). Elle ne s'attache pas directement aux plans d'analyse évoqués dans l'introduction (syntaxe, sémantique, etc.). Par contre, elle se rapproche des effets interprétatifs (saillances globales versus saillances locales) et, indirectement, des spécificités du genre textuel, plus précisément de la nature et du but de la communication entre les interlocuteurs.

Avec notre modèle multiforme, nous contribuons à complexifier la notion de saillance, et, par conséquent, à retarder ses possibilités de modélisations formelles et computationnelles. Les quelques pistes que nous évoquons ici restent pour le moment très générales, et de nombreuses expériences sont à prévoir avant d'aboutir à un modèle formel qui intègre l'ensemble des aspects cités. Les perspectives de ce travail relèvent ainsi avant tout d'expérimentations : pour vérifier l'influence de tel facteur, il faudrait mettre en œuvre une expérimentation psycholinguistique qui oppose matérialisation et inhibition du facteur en

question ; pour vérifier l'importance de telle dimension, il faudrait procéder à des analyses et mettre en rapport saillance, structure informationnelle, chaînes de coréférences et transitions référentielles ; d'une manière générale, il faudrait exploiter les méthodes de la linguistique de corpus et annoter dans des textes l'ensemble des facteurs susceptibles d'intervenir pour procéder ensuite à des calculs de fréquences et à des recherches de corrélations, de manière à mettre en évidence l'intervention de la saillance. Ce sont des voies que nous explorons, avec notamment la clarification des procédures d'annotation envisageables dès que l'on s'intéresse à une notion multicritère. Enfin, la perspective computationnelle reste un objectif à long terme, dans la mesure où un système de calcul automatique de scores de saillance permettrait d'améliorer significativement les systèmes de compréhension automatique et de traitement automatique de la langue.

Références bibliographiques

- Abiven, K. (2010), Saillance et genre de discours : le cas de l'anecdote, *La saillance en langue et en discours, 5^e colloque du Réseau des linguistes du Grand Est*, Université de Strasbourg, pp. 2-3.
- Asher, N. & Lascarides, A. (2003), *Logics of Conversation*, Cambridge, Cambridge University Press.
- Blondel, M. (2003), Saillance linguistique dans une langue à modalité visuo-gestuelle, *Fonctions et moyens de la focalisation à travers les langues. Mémoires de la Société Linguistique de Paris, Tome XIII*, Leuven, Peeters, pp. 187-203.
- Caron, J. (1989), *Précis de psycholinguistique*, Paris, PUF.
- Deleuze, G. (1969), *Différence et Répétition*, réédition en 1997 : Paris, PUF.
- Grobet, A. (2002), *L'identification des topiques dans les dialogues*, Bruxelles, De-Boeck-Duculot.
- Grosz, B. J., Joshi, A. K. & Weinstein, S. (1995), Centering: A Framework for Modelling the Local Coherence of Discourse, *Computational Linguistics* 21-2, pp. 203-225.
- Landragin, F., Bellalem, N. & Romary, L. (2002), Referring to Objects with Spoken and Haptic Modalities. *Fourth IEEE International Conference on Multimodal Interfaces*, IEEE CS Press, Los Alamitos, CA, pp. 99-104.
- Landragin, F. (2004), Saillance physique et saillance cognitive, *Cognition, Représentation, Langage (CORELA)* 2(2), <http://edel.univ-poitiers.fr/corela>.
- Larrivée, P. (2008), *Une histoire du sens. Panorama de la sémantique linguistique depuis Bréal*, Bruxelles, Peter Lang.
- Neveu, F. (2009), *Lexique des notions linguistiques*, Paris, Armand Colin.
- Schnedecker, C. (2009), La notion de saillance : problème définitoires et avatars, *Aspects linguistiques et communicatifs de la mise en évidence dans un texte. Colloque Saillance*, Université de Genève, pp. 3-6.
- Sidner, C. L. (1979), *Towards a Computational Theory of Definite Anaphora in English Discourse*, Ph.D. Thesis, Massachusetts Institute of Technology.
- Thom, R. (1993), Saillance et prégnance, dans : Dorey, R. (éd.), *L'inconscient et la science*, Paris, Dunod, pp. 64-82.
- Van Dijk, T. & Kintsch, W. (1983), *Strategies of discourse comprehension*, New York, Academic Press.