

La saillance comme point de départ pour l'interprétation et la génération

Frédéric Landragin

LORIA

Campus scientifique — BP 239
54506 Vandœuvre-lès-Nancy CEDEX
Frederic.Landragin@loria.fr

1. Introduction

La saillance intervient fortement dans l'interprétation d'un énoncé en situation de dialogue ou dans la compréhension d'un texte : mettant en avant un élément du contexte, elle dirige l'attention sur cet élément et rend sa prise en compte prioritaire dans le processus de résolution des références et des coréférences. Elle permet ainsi d'attribuer un référent à un pronom ou à une expression ambiguë telle que « *le N* » lorsque le contexte contient plusieurs objets du type N (Clark *et al.* 1983). La saillance linguistique constitue par exemple une aide à la résolution des anaphores : le N qui vient d'être mentionné est saillant et peut être repris par un pronom (Lappin & Leass 1994). Dans le dialogue homme-machine sur écran, la saillance visuelle constitue un critère d'identification du N perçu de manière prioritaire et sur lequel est basé l'expression « *le N* » (cf. par exemple Kievit *et al.* 2001). La saillance due à la situation d'interaction permet la compréhension d'un pronom sans antécédent linguistique, comme dans l'exemple de (Isard 1975) où un enfant tente de caresser un lion à travers une cage et se voit prévenir du danger par « *attention, il risque de te mordre* ». La saillance due au contexte de tâche et à l'historique de l'interaction joue le même rôle de désambiguïsation : « *le N* » peut se comprendre non pas comme « *le seul N* » mais comme « *le N qui vient d'être manipulé* » ou « *le N suivant dans la succession des tâches* » (Wright 1990). Nous voyons en cela la saillance comme critère de détection des indices sur lesquels baser l'identification de l'implicite.

L'idée que nous défendons ici est que la saillance est un phénomène global, qui intègre les différentes facettes du contexte (perception visuelle, langage, tâche applicative, historique de l'interaction), et qu'une caractérisation de la saillance en vue de son calcul par un système de compréhension doit reposer sur des critères plus généraux que les distinctions classiques entre thème et rhème ou entre topique et commentaire. Nous tentons d'identifier ces critères et de proposer une caractérisation générale aussi complète que possible. Pour cela, nous explorons les principaux travaux sur la saillance visuelle et la saillance linguistique, et nous fusionnons les facteurs de saillance avancés, ce qui nous permet d'identifier des types de critères. Nous explorons alors systématiquement ces types de critères et nous analysons les instances manquantes. Nous en déduisons une caractérisation générale que nous illustrons ici essentiellement sur la saillance linguistique. Nous proposons enfin un calcul numérique des saillances visuelle et linguistique, le but étant d'apporter à un système de compréhension ou de génération automatique une base pour la modélisation de l'attention.

2. Comment appréhender la saillance ?

Une première définition de ce qui est saillant est ce qui arrive en premier à l'esprit, ce qui capte l'attention. Ce point de vue correspond souvent à considérer comme saillant ce qui est naturel, simple, clair. Selon Stevenson (2002), les premiers travaux dans les années 1970 se sont focalisés sur la saillance visuelle avec cette idée de simplicité naturelle. Osgood et Bock (1977) parlent par exemple de simplicité naturelle (*naturalness*) et de clarté (*vividness*). Ainsi, l'ordre naturel des constituants dans une phrase reflète souvent l'ordre naturel des événements (agent–action–patient) et par conséquent une certaine hiérarchie de saillance. Du côté de la perception visuelle, ce point de vue rejoint celui de la Gestalt à propos de bonne forme : quelque chose de simple, d'immédiat, de percutant (cf. Guillaume 1979).

Deux autres définitions sont souvent avancées. En suivant la conception de la Gestalt selon laquelle la saillance dénote la force de résistance aux perturbations, est saillant ce qui est stable. D'un autre côté, est saillant ce qui est original ou nouvellement introduit dans la situation. Ce dernier facteur rejoint le critère de non-familiarité de Loftus et Mackworth (1978) : des objets non familiers tendent à être fixés plus longtemps et en deviennent saillants. Ce facteur rejoint également celui d'inattendu : est saillant l'élément perturbateur, inattendu, curieux, intrigant, énigmatique. Il s'agit en effet de l'élément sur lequel on s'interroge, ou sur lequel le regard va s'attarder pour résoudre le problème qu'il pose. Par exemple, tout élément visuel ou langagier pour lequel l'activité perceptive de reconnaissance s'avère difficile en devient saillant. Les deux définitions se contredisent ainsi totalement.

Un autre problème apparaît lorsque la saillance est définie par rapport à des facteurs cognitifs tels que l'attention, la mémoire ou la familiarité. Ces facteurs sont propres à l'utilisateur et montrent en quoi la saillance est subjective. Citons la familiarité culturelle (par exemple la présence d'un être humain ou d'un lion dans notre champ de vision est saillante et fait qu'on va se tourner vers eux) et la familiarité individuelle (au cours de notre éducation et de notre vie passée, nous acquérons tous nos propres sensibilités, notre propre vision des couleurs, etc.). La saillance dépend de plus de l'attention de l'utilisateur au moment de l'interaction, de son intérêt, de son intention communicative : est saillant ce qui a de l'intérêt compte tenu de l'objectif de la communication. Osgood et Bock (1977) parlent ainsi de saillance liée à l'intérêt du locuteur (*motivation-of-speaker*). Nous pouvons déduire de ce facteur une dépendance de la saillance envers la tâche applicative : quand on invite des collègues à entrer dans un bureau, les chaises sont saillantes car ils vont vouloir s'asseoir. D'un autre côté, il n'y a pas équivalence entre saillance et intention communicative. La saillance peut au contraire prendre un rôle perturbateur : on peut se focaliser intentionnellement sur un objet précis tout en étant perturbé par un objet fortement saillant.

Tous ces problèmes se retrouvent lors du passage des définitions aux formalisations. Une illustration immédiate concerne la distinction entre « donné » (*given*) et « nouveau » (*new*) : le donné est saillant car il est stable et connu, mais le nouveau peut être tout autant voire plus saillant, étant justement nouveau et susceptible d'intriguer. Nous retrouvons la même impossibilité à déterminer l'élément saillant dans la distinction entre « thème » et « rhème » : selon Caron (1989), la séquence la plus naturelle de deux phrases est celle où le rhème de la première phrase est repris comme thème de la seconde. Le rhème est repris et est donc saillant. Quant au thème de la nouvelle phrase, on peut aussi considérer qu'il est saillant. Caron conclut qu'il est impossible de savoir si c'est le thème ou le rhème qui est l'élément le plus saillant dans un énoncé. Même chose à propos de la distinction entre « présupposé » et « posé » : une information présupposée, donc implicite, peut s'avérer aussi saillante qu'une information explicite. Même chose à propos de la saillance de l'« agent » et de la saillance du « patient » : classiquement, par exemple dans (Sidner 1979) ou dans la Théorie du Centrage (Grosz *et al.* 1995), l'agent est considéré comme plus saillant que le patient, lui-même considéré avant les autres rôles thématiques. Il n'y a cependant pas unanimité à propos de cette hiérarchie. Dans des travaux plus récents et basés sur des expérimentations, (Stevenson *et al.* 1994) montrent qu'entre agent et patient, la préférence est significativement pour le patient, du moins dans certaines phrases. Ainsi, pour celles qui décrivent un événement, les conséquences de l'événement sont plus présentes dans les représentations mentales que les conditions initiales. Ces conséquences s'appliquant au patient, celui-ci en devient plus saillant que l'agent. Pour les phrases qui ne décrivent pas d'événement, tout dépend des composants de la phrase et rien ne peut être conclu. Pour déterminer quelle est l'entité saillante entre agent et patient, il faudrait détailler chaque type d'action, donc chaque type de verbe, en étudiant sa sémantique. Le problème s'avère complexe, du fait de la multiplicité des critères qui entrent en jeu. (Stevenson *et al.* 1994) ainsi que (Pearson *et al.* 2001) détaillent par exemple le cas des verbes de transfert (donner quelque chose à quelqu'un) et montrent que le receveur est plus saillant que le donneur et que l'objet transféré. Bref, dans l'état actuel des recherches, il s'avère impossible de faire des liens définitifs entre la saillance et les distinctions classiques que nous avons citées. Les facteurs de saillance souvent avancés se recouvrent, se contredisent, et rien de précis ne peut être dégagé.

Une dernière considération donnera le fil directeur de notre approche : dans le cadre d'une tâche consistant en la description téléphonique d'un itinéraire routier, Edmonds (1993) montre que la saillance dépend du contexte incluant le référent. Il donne l'exemple d'un immeuble *a priori* saillant par sa taille importante, et qui perd toute saillance lorsqu'il est entouré d'immeubles encore plus grands. Il considère également que certaines caractéristiques sont saillantes pour certains objets et pas pour d'autres, de même que certaines caractéristiques sont saillantes dans un but précis du dialogue et non dans un autre but. Par exemple, la caractéristique « taille » est saillante lorsque le but du dialogue est la désignation d'un immeuble, mais ne l'est pas lorsqu'il s'agit de la désignation d'une intersection de rues. Ainsi, plutôt que de chercher une définition propre à un objet ou à une structure en particulier, nous nous tournons vers l'émergence de saillance dans un ensemble contextuel. Les propriétés de l'énoncé et les propriétés physiques de l'ensemble des objets de la scène permettent de distinguer certains éléments, et de les considérer comme saillants. Considérant qu'il est plus facile d'appréhender les facteurs de saillance visuelle que ceux de saillance linguistique, nous avons analysé les rôles des propriétés physiques des objets et des particularités visuelles d'une scène pour en dégager des critères applicables à la saillance en général. Nous retenons ainsi les critères suivants :

- critères intrinsèques aux unités (unités visuelles : objets ; unités linguistiques : mots) ;
- placement à un endroit stratégique (dans le cadre de l'image ; dans la structure de l'énoncé) ;
- isolement (singleton dans une partition en groupes perceptifs ; groupe de mots en apposition) ;

- rupture dans une continuité (par exemple dans le rythme de l'image ou de l'énoncé) ;
- répétition (dans la disposition des objets ; dans l'apparition des mots ou expressions) ;
- symétrie (dans la disposition des objets ; par une figure de style telle qu'un chiasme).

Ces critères mettent l'accent sur la structure de l'image ou de l'énoncé. C'est avec ces aspects structurels que nous commencerons notre caractérisation.

3. Une caractérisation de la saillance linguistique

Des critères liés à la forme de l'énoncé. A la suite de Stevenson (2002), nous séparons les aspects formels, c'est-à-dire liés aux caractéristiques prosodiques et grammaticales de l'énoncé oral, des aspects sémantiques liés au contenu du message. Dans la majorité des travaux en linguistique computationnelle, par exemple dans (Alshawi 1987), (Lappin & Leass 1994), (Grosz *et al.* 1995), ou encore dans les travaux de Hajičová cités dans (Krahmer & Theune 2002), ce sont essentiellement des critères formels qui définissent la saillance. Comme le soulignent Krahmer et Theune (2002), la récence est souvent mise en avant, les entités les plus saillantes étant définies comme étant les plus récemment mentionnées. A partir de ces travaux et de nos critères, nous proposons la classification suivante :

1. La saillance intrinsèque au mot : mots constitués de phonèmes sonores ; noms propres ; déictiques purs du fait de leur manque d'autonomie référentielle et de l'habitude qu'ils entraînent chez l'interlocuteur à faire particulièrement attention aux conditions de leur énonciation.
2. La saillance par une mise en avant explicite lors de l'énonciation : prosodie particulière ; présence d'une pause avant et après la prononciation d'un mot ; erreur de prononciation.
3. La saillance par une construction syntaxique dédiée : constructions clivées (« *c'est ... qui ...* ») ; détachements en tête de phrase (« *le triangle, le rouge, tu dois le mettre à côté du bleu* »).
4. La saillance syntaxique liée à l'ordre d'apparition des mots : le début et la fin sont prédisposés pour rendre saillant le mot ou le groupe de mot qui y prend place ; les répétitions et les symétries.
5. La saillance grammaticale, c'est-à-dire liée aux fonctions grammaticales des mots : le sujet est souvent considéré comme l'élément le plus saillant (ce qui justifie les constructions passives).
6. La saillance indirecte ou transfert grammatical de saillance : une entité du discours en lien grammatical direct avec l'entité focalisée en devient saillante.

Des critères liés au sens de l'énoncé. C'est quand nous abordons la sémantique que nous considérons les notions de thème, de focus ou encore de topic :

1. La saillance liée à la sémantique des mots : agent-patient.
2. La saillance liée à la sémantique de l'énoncé : thème-rhème (concerne le repérage de l'énoncé par le locuteur) ; présupposé-posé.
3. La saillance liée à la sémantique de la conversation (se construisant au cours du dialogue) : topique-commentaire (à un niveau stylistique, le topique est l'entité dont parle la conversation et le commentaire est ce qui en est dit) ; donné-nouveau (concerne le repérage de l'énoncé par l'interlocuteur).
4. La saillance indirecte ou transfert sémantique de saillance : un référent très lié au topique est plus saillant qu'un référent qui ne lui est pas apparenté.

4. Un calcul numérique de la saillance

A.	« <i>le triangle rouge</i> » dans « <i>le triangle rouge se met à côté du bleu</i> »	
B.	« <i>le triangle rouge</i> » dans « <i>c'est le triangle rouge que tu dois mettre à côté du bleu</i> »	
	cas A	cas B
accent tonique :	1	0
construction syntaxique dédiée :	0	1
place stratégique en début d'énoncé :	1	1
fonction grammaticale sujet :	1	0
saillance :	0.75	0.5

Le principe que nous suivons est le suivant : un objet ou une entité du discours qui partage une propriété avec d'autres obtient un score de 0 pour cette propriété ; un objet qui est le seul à posséder une propriété obtient un score de 1 pour cette propriété. Cette méthode basée sur la singularité nous semble opérationnelle, dans la mesure où repérer des singularités dans une liste de propriétés s'avère possible d'un point de vue computationnel, du moins à partir du moment où ces propriétés peuvent se représenter formellement. Les figures présentent des scores numériques pour un calcul opérationnel des saillances visuelle et linguistique.

A. Par la couleur et la taille					B. Par l'isolement				
	a	b	c	d		a	b	c	d
forme :	0	0	0	1	forme :	0	0	0	0
couleur :	0	1	0	0	couleur :	0	0	0	0
taille :	0	1	0	0	taille :	0	0	0	0
isolement :	0	0	0	0	isolement :	1	0	0	0
saillance :	0	0.5	0	0.25	saillance :	0.25	0	0	0

5. Conclusion et perspectives

La saillance ne fait pas partie du message communiqué, mais tout le message se base sur elle, s'explique par elle, se structure en fonction d'elle. Nous avons voulu montrer ici que cette notion pouvait être appréhendée d'une manière efficace en considérant les différentes modalités de la communication. A partir d'une abstraction faite sur les facteurs de saillance visuelle, nous avons pu préciser certaines notions propres à la saillance linguistique. Notre principale perspective est de spécifier des algorithmes d'exploitation des scores proposés en tenant compte des autres paramètres mis en jeu dans l'énoncé. Nous aurons en particulier à spécifier des stratégies pour la confrontation entre la saillance visuelle et la saillance linguistique.

Bibliographie

- ALSHAWI, H. (1987), *Memory and Context for Language Interpretation*, Cambridge University Press.
- CARON, J. (1989), *Précis de psycholinguistique*, PUF, Paris.
- CLARK, H.H., SCHREUDER, R. et BUTTRICK, S. (1983), Common Ground and the Understanding of Demonstrative Reference, *Journal of Verbal Learning and Verbal Behavior* 22.
- EDMONDS, P.G. (1993), *A Computational Model of Collaboration on Reference in Direction-Giving Dialogues*, Ms. Thesis, University of Toronto, Canada.
- GARROD, S. & SANFORD, A.J. (1988), Thematic Subjecthood and Cognitive Constraints on Discourse Structure, *Journal of Pragmatics* 12.
- GROSZ, B.J., JOSHI, A.K. et WEINSTEIN, S. (1995), Centering: A Framework for Modelling the Local Coherence of Discourse, *Computational Linguistics* 21(2).
- GUILLAUME, P. (1979), *La psychologie de la forme*, Flammarion (collection Champs), Paris.
- ISARD, S. (1975), Changing the Context, In: KEENAN, E.L. (Ed.) *Formal Semantics of Natural Language*, Cambridge University Press, London & New York.
- KIEVIT, L., PIWEK, P., BEUN, R.J. & BUNT, H. (2001), Multimodal Cooperative Resolution of Referential Expressions in the DENK System, In: BUNT, H. & BEUN, R.J. (Eds.), *Cooperative Multimodal Communication*, Springer, Berlin & Heidelberg.
- KRAHMER, E. & THEUNE, M. (2002), Efficient Context-Sensitive Generation of Referring Expressions, In: VAN DEEMTER, K. & KIBBLE, R. (Eds.), *Information Sharing: Reference and Presupposition in Language Generation and Interpretation*, CSLI Publications, Stanford.
- LAPPIN, S. & LEASS, H.J. (1994), A Syntactically Based Algorithm for Pronominal Anaphora Resolution, *Computational Linguistics* 20(4).
- LOFTUS, G.R. & MACKWORTH, N.H. (1978), Cognitive Determinants of Fixation Location during Picture Viewing, *Journal of Experimental Psychology: Human Perception and Performance* 4.
- OSGOOD, C.E. & BOCK, J.K. (1977), Saliency and Sentencing: Some Production Principles, In: ROSENBERG, S. (Ed.), *Sentence Production: Developments in Research and Theory*, Erlbaum, Hillsdale, NJ.
- PEARSON, J., POESIO, M. & STEVENSON, R. (2001), The Effects of Animacy, Thematic Role and Surface Position on the Focusing of Entities in Discourse, In: *Proceedings of the First Workshop on Cognitively Plausible Models of Semantic Processing*.
- SIDNER, C.L. (1979), *Towards a Computational Theory of Definite Anaphora in English Discourse*, Ph.D. Thesis, MIT.
- STEVENSON, R.J., CRAWLEY, R.A. & KLEINMAN, D. (1994), Thematic Roles, Focus and the Representation of Events, *Language and Cognitive Processes* 9(4).
- STEVENSON, R.J. (2002), The Role of Saliency in the Production of Referring Expressions, In: VAN DEEMTER, K. & KIBBLE, R. (Eds.), *Information Sharing: Reference and Presupposition in Language Generation and Interpretation*, CSLI Publications.
- WRIGHT, P. (1990), Using Constraints and Reference in Task-Oriented Dialogue, *Journal of Semantics* 7.