

## Génération intégrée localisée pour la production de documents

Pierre Hankach<sup>1, 2</sup>

(1) Orange Labs, 2 avenue Pierre Marzin, 22300 Lannion

(2) Université Paris 7, 5 rue Thomas Mann, 75013 PARIS

pierre.hankach@orange-ftgroup.com

**Résumé.** Dans cet article, nous proposons une approche intégrée localisée pour la génération. Dans cette approche, le traitement intégré des décisions linguistiques est limité à la production des propositions dont les décisions qui concernent leurs générations sont dépendantes. La génération se fait par groupes de propositions de tailles limitées avec traitement intégré des décisions linguistiques qui concernent la production des propositions qui appartiennent au même groupe. Notre approche apporte une solution pour le problème de complexité computationnelle de la génération intégrée classique. Elle fournit ainsi une alternative à la génération séparée (séquentielle ou interactive) qui présente plusieurs défauts mais qui est implémentée de manière répandue dans les systèmes de générations existants.

**Abstract.** In this paper, we propose a localized integrated approach for generation. In this approach, the integrated handling of linguistic decisions is limited to the production of propositions whose decisions concerning their generation are dependant. The generation is performed by groups of propositions of limited size with an integrated handling of linguistic decisions that concern the production of the propositions that belong to the same group. Our approach provides a solution for the computational complexity problem of classical integrated generation. Therefore, it provides an alternative to separated generation (sequential and interactive) that has many drawbacks but is widely implemented in today's generation systems.

**Mots-clés :** génération intégrée localisée - architectures de génération - SDRT - segmentation du discours.

**Keywords:** localized integrated generation - generation architectures - SDRT - discourse segmentation.

# 1 Introduction

Dans ce papier, nous proposons une nouvelle approche pour la gestion des décisions linguistiques de la génération basée sur l'identification des décisions linguistiques dépendantes. En effet, la façon dont les décisions linguistiques sont organisées et coordonnées a une grande influence sur les performances de la génération. La qualité de la sortie de la génération et la complexité de calcul en dépendent toutes les deux. Dans la littérature, deux types d'architectures pour gérer les décisions de la génération ont été proposés : L'architecture intégrée et l'architecture séparée. La première souffre du problème de la complexité alors que la deuxième présente des défauts qui influent négativement sur la qualité de la génération. Dans la suite de l'introduction, nous présentons les deux types d'architectures proposées dans la littérature et les implications de chacune sur les performances de la génération, ensuite nous introduisons brièvement notre approche.

Les systèmes à architecture intégrée (Appelt, 1985; Kantrowitz & Bates, 1992) implémentent une dépendance entre toutes les décisions linguistiques individuelles de la génération. Cette architecture est théoriquement séduisante mais représente un inconvénient majeur qui est sa grande complexité. Pour cette raison, les systèmes à architecture intégrée n'ont fonctionné que sur des applications de génération simples avec des sorties à taille réduite.

Les systèmes à architecture séparée sont divisés en deux catégories : les systèmes à architecture séquentielle (Reiter & Dale, 2000) et les systèmes à architecture interactive (Hovy, 1993). Dans les systèmes à architecture séparée, les décisions linguistiques individuelles du même type sont groupées dans des tâches comme la lexicalisation, l'agrégation ou l'ordonnancement. Dans l'architecture séquentielle de (Reiter & Dale, 2000), chaque tâche est allouée à un des trois niveaux : La macro-planification regroupe les tâches de sélection et d'organisation de l'information. La micro-planification regroupe les tâches pour choisir la meilleure perspective textuelle à l'expression construite par le macro-planificateur. La formulation regroupe les tâches pour réaliser un texte bien formé. L'exécution de la génération se fait séquentiellement de la macro-planification à la réalisation. Cette architecture possède l'avantage de réduire la complexité de la génération et de modulariser le traitement. En contre partie, elle ne prend pas en compte que les décisions d'une tâche de génération dépendent des décisions des autres tâches (voir (Danlos, 1984)). Notamment, le fait que les tâches du début de la séquence prennent des engagements précoces s'avère problématique. En effet, les décisions prises par les tâches du début de la séquence en tenant compte des informations disponibles à leur niveau, ne demeurent pas forcément les meilleures, après la réalisation des tâches qui suivent dans la séquence, quand à la qualité de la forme linguistique produite par la génération. Dans le pire des cas, les décisions prises par les tâches du début de la séquence vont engendrer des tâches irréalisables dans la suite, et donc un échec de la génération. Le phénomène est appelé l'écart de la génération. L'architecture interactive de (Hovy, 1993) tente de remédier à certaines limitations de la génération à architecture séquentielle. Cette architecture permet une alternance entre les tâches de la macro-planification et de la micro-planification. Elle permet au macro-planificateur de prendre en compte certaines décisions du micro-planificateur mais ne résout pas totalement les problèmes de l'architecture séquentielle tout en introduisant une complexité supplémentaire par rapport à cette dernière.

Dans cet article, nous proposons une nouvelle approche de construction d'un système de génération basée sur une génération intégrée localisée, où le traitement intégré des décisions linguistiques est limité à la production des propositions dont les décisions linguistiques qui concernent leur réalisation sont dépendantes. Notre approche est fondée sur l'intuition que la réalisation d'une proposition dans un texte ne dépend pas de la réalisation de toutes les autres propositions

du texte. Ceci entraîne une indépendance entre décisions linguistiques basée sur la localisation des propositions sur lesquelles elles portent (plutôt que sur leur type comme supposé dans les architectures séparées). Dans notre approche, la génération se fait par groupes de propositions de tailles limitées avec traitement intégré des décisions linguistiques pour la production des propositions d'un groupe. La génération complète d'un document se fait par la production successive de tous les groupes de propositions qui le composent. Dans la suite, nous appelons étape de génération la génération d'un groupe de propositions du document final. Notre approche de génération intégrée localisée permet d'implémenter une génération intégrée avec une complexité combinatoire acceptable. Ceci est une conséquence du fait qu'à chaque étape la dépendance est considérée entre un nombre limité de décisions linguistiques. Dans la section 2 nous présentons les bases théoriques d'une génération intégrée localisée. Dans la section 3 nous présentons les étapes de la génération intégrée localisée. Dans la section 4 nous présentons une brève description d'une étape de génération. Nous concluons dans la section 5.

## 2 Bases théoriques d'une génération intégrée localisée

Afin d'implémenter une génération intégrée localisée, nous avons besoin de définir des procédés pour :

- Identifier des groupes de propositions d'un discours dont les décisions linguistiques qui concernent leurs générations sont dépendantes. Pour cela, nous nous basons sur la structure des segments de discours de (Grosz & Sidner, 1986).
- Organiser les propositions d'un groupe entre elles et par rapport aux propositions d'autres groupes. Pour cela, nous nous basons sur la structure relationnelle de SDRT (Segmented Discourse Representation Theory, (Asher & Lascarides, 2003)) qui constitue une sorte de matrice où s'imbriquent les propositions du discours.

Dans la suite, nous présentons la segmentation du discours présentée par (Grosz & Sidner, 1986), puis la structure relationnelle de SDRT. Enfin, pour les besoins de notre approche de génération, nous superposons la segmentation établie par (Grosz & Sidner, 1986) sur la structure relationnelle de SDRT.

### 2.1 La segmentation du discours

#### 2.1.1 Grosz & Sidner

Une observation courante dans la littérature à propos du discours est sa nature segmentée. Les propositions du discours se regroupent en segments, où chaque segment réalise un but communicatif particulier. Les segments peuvent à leur tour se grouper avec d'autres segments et propositions dans des segments plus larges. (Grosz & Sidner, 1986) (dorénavant G&S) associe la segmentation du discours à sa structure intentionnelle. La structure intentionnelle est formée par les intentions de communication (dans la suite, nous interchangeons l'utilisation de "intention de communication" et "but communicatif") reliées entre elles par deux types de relations : dominance et satisfaction-précédence. Une intention  $I_m$  domine une intention  $I_n$  quand la satisfaction de  $I_n$  fait partie de la satisfaction de  $I_m$ . Une intention  $I_n$  satisfaction-précède une intention  $I_m$  quand  $I_n$  doit être satisfaite avant  $I_m$ .

La structure des segments (appelée structure linguistique dans G&S) est isomorphe à la structure intentionnelle. Chaque segment du discours réalise une intention de communication appelée

DSP (*Discourse Segment Purpose*) du segment. La relation entre les segments est une réflexion des relations qui existent entre leurs DSPs respectifs. Un segment de discours  $DS_1$  contient un autre segment de discours  $DS_2$  quand l'intention de communication  $DSP_1$  de  $DS_1$  domine l'intention de communication  $DSP_2$  de  $DS_2$ . Par exemple, considérons le discours de la figure 1 adapté de (Asher & Lascarides, 2003). La figure 2 représente les intentions communicatives du discours de la figure 1 et les relations de dominance qui les relient. Le discours global est le

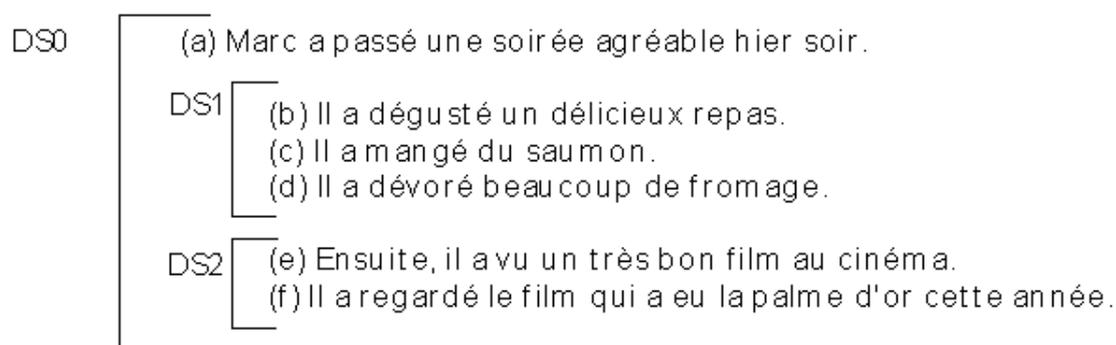


FIG. 1 – Soirée de Marc.

segment  $DS_0$ , qui réalise l'intention communicative  $I_0$ . La réalisation de  $I_1$  et de  $I_2$  fait partie du plan de la réalisation de  $I_0$ . Les intentions communicatives  $I_1$  et  $I_2$  sont par conséquent dominées par  $I_0$ . De fait de cette dominance, le segment  $DS_0$  contient les deux segments  $DS_1$  et  $DS_2$ .

La structure des segments n'est pas décompositionnelle. Un segment peut inclure d'autres segments ainsi que d'autres propositions qui appartiennent exclusivement à ce segment. Dans la suite de ce papier, nous utilisons le terme *noyau* pour désigner les propositions d'un segment qui appartiennent exclusivement à lui et à aucun des segments qu'il contient. Dans l'exemple de la figure 1, (a) constitue le noyau du segment  $DS_0$ . Le noyau d'un segment transmet le but principal de communication du segment, alors que les segments intermédiaires que contient le segment aident à la satisfaction de son but de communication.

### 2.1.2 Dépendances dans la réalisation du discours au vu de sa segmentation

La contribution d'un constituant (une proposition ou un segment) au discours se fait à travers sa contribution à la réalisation du but communicatif du segment auquel il appartient. Par conséquent, l'introduction d'une proposition, ainsi que la forme linguistique qu'elle prend dans le discours, va dépendre principalement de la manière avec laquelle elle se combine aux autres constituants du segment auquel elle appartient pour satisfaire son but communicatif. Du point de vue de la génération, la dépendance des décisions linguistiques est alors limitée aux décisions qui permettent de réaliser le but communicatif d'un segment. Ces décisions concernent la production du contenu propre (noyau) du segment et sa disposition avec les segments inclus qui réalisent des buts communicatifs partiels. Dans la section 3, nous présentons une génération par étapes des propositions du discours, où à chaque étape sont produites les propositions dont les décisions linguistiques concernant leur réalisation sont dépendantes.

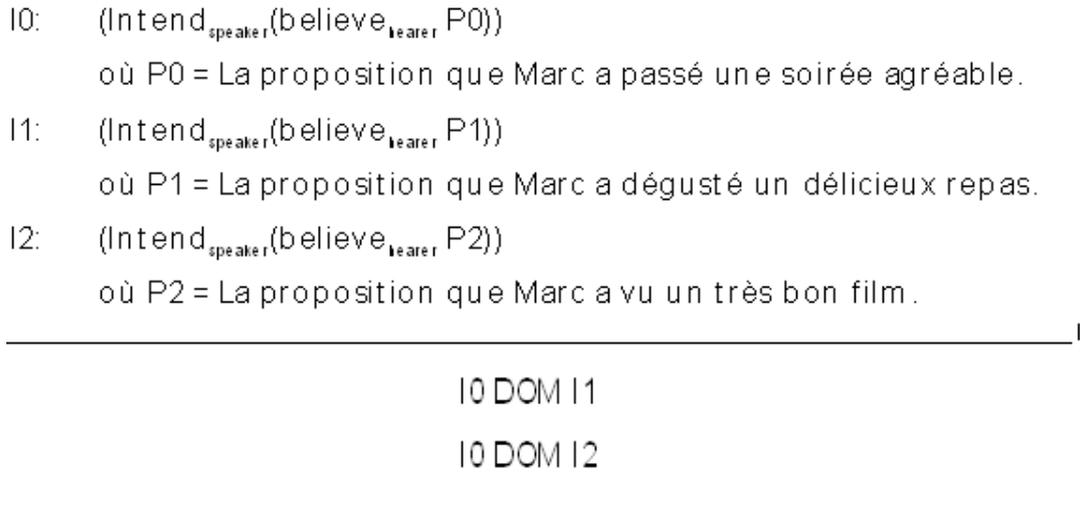


FIG. 2 – Les relations de dominance pour les DSPs du texte “la soirée de Marc”.

## 2.2 Structure relationnelle de SDRT

La structure relationnelle de SDRT décrit le rattachement des propositions d’un discours cohérent. Elle est construite à partir des propositions et des relations de discours (ou relations rhétoriques) qui les relient. Deux types de relations existent entre les propositions : les relations subordinantes et les relations coordinantes. Géométriquement, une relation subordinante est représentée par un trait vertical et une relation coordinante est représentée par un trait horizontal. Les propositions d’un discours cohérent se relient toutes entre elles pour former une structure hiérarchique unique. La forme de la structure hiérarchique découle des propriétés suivantes des relations subordinantes et coordinantes :

- Une relation subordinante laisse son premier argument accessible pour un nouvel attachement.
- Une relation coordinante bloque l’accessibilité à son premier argument pour de nouveaux attachements.
- Si deux propositions sont attachées par une relation coordinante, si l’une d’elle est rattachée par une relation subordinante à une troisième alors l’autre l’est aussi.

La figure 3 représente la structure relationnelle de SDRT du texte “soirée de Marc”

## 2.3 Alignement de la structure des segments sur la structure relationnelle

Dans ce qui suit, nous identifions les segments de discours et leurs rattachements dans la structure relationnelle de SDRT.

### Délimitation d’un segment

Soit une structure relationnelle  $S0$  et  $T0$  le noeud sommet de la structure  $S0$ . Selon les propriétés de la structure relationnelle,  $T0$  constitue le topique du contenu de la structure  $S0$ . La partie inférieure de la structure qui est subordonnée au topique, contribue à la satisfaction du but communicatif que le topique cherche à réaliser. Par conséquent, nous pouvons conclure que le but communicatif  $DSP0$  du segment  $DS0$  associé à  $S0$  est de réaliser le topique. Ceci est exprimé par la formule suivante :

$$DSP0 = I_{réaliser}(T0)$$

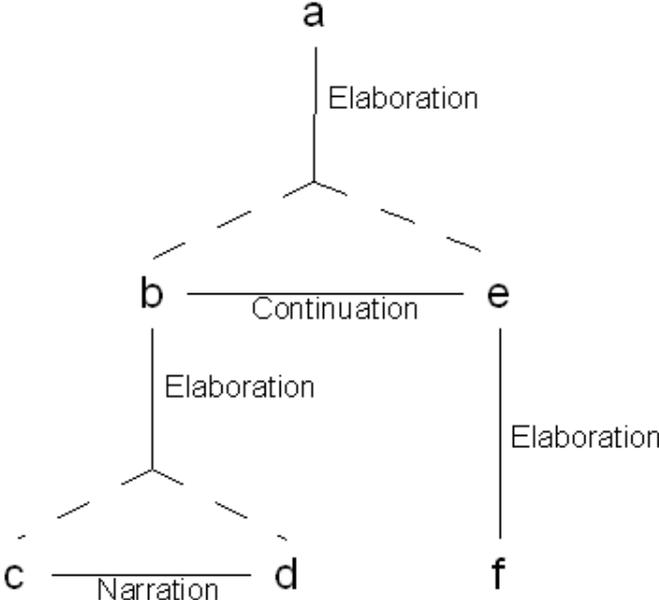


FIG. 3 – La structure relationnelle de SDRT du texte “soirée de Marc”.

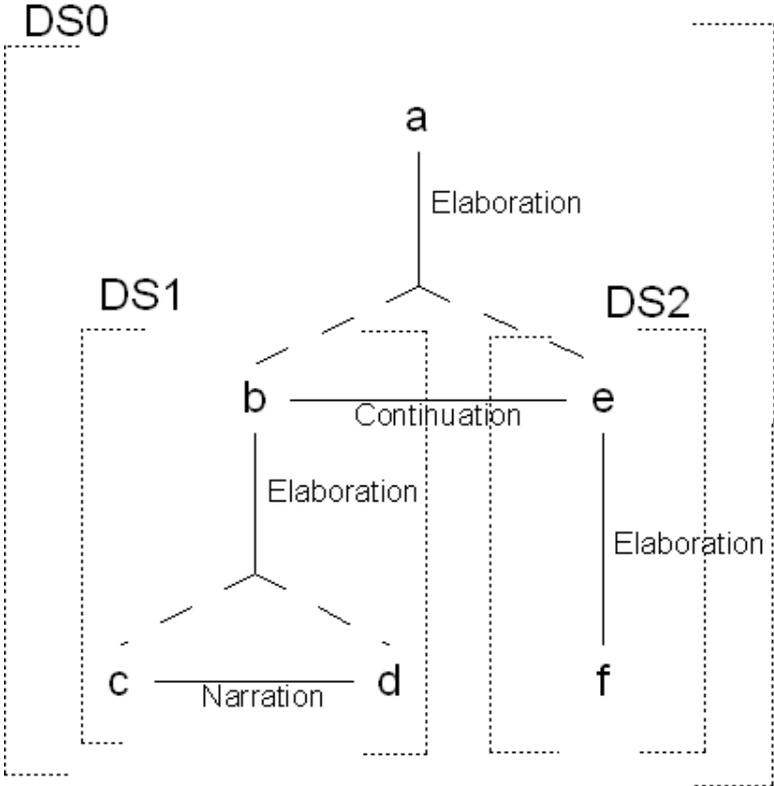


FIG. 4 – Les segments dans la structure relationnelle du texte “soirée de Marc”.

De plus, tous les éléments subordonnés au topique appartiennent au segment dont le but communicatif est  $DS_0$ . En conclusion, un segment de discours est représenté dans la structure relationnelle par le noeud-topique qui réalise son but communicatif et tous les éléments de la structure qui sont subordonnés à lui. La relation d'un segment avec les autres constituants du discours se fait à travers son topique. La figure 4 illustre les segments dans la structure relationnelle du texte de la figure 1.

### Enboitement des segments

Soit un segment  $DS_0$  qui contient un autre segment  $DS_1$ . La structure  $S_0$  correspondant à  $DS_0$  contient la structure  $S_1$  correspondant à  $DS_1$ , et le topique  $T_0$  de  $DS_0$  domine le topique  $T_1$  de  $DS_1$ . La dominance de  $T_1$  par  $T_0$  est concrétisée par une relation subordinante ou une chaîne de relations subordinantes reliant les deux topiques. La relation de dominance entre les topiques est le reflet de la relation de dominance entre les buts communicatifs des deux segments. Dans la figure 4 la structure de  $DS_1$  fait partie de la structure de  $DS_0$  ( $DS_0$  domine  $DS_1$ ) et (b) qui est le topique de  $DS_1$  et subordonné à (a) le topique de  $DS_0$ .

### Noyau d'un segment

Le noyau d'un segment  $DS_0$  est formé de la partie supérieure de la structure relationnelle  $S_0$  correspondante dont les éléments n'appartiennent à aucun des segments que contient  $DS_0$ . Cela est dû au fait que les segments inclus dans  $DS_0$  débutent en-dessous du sommet de la structure qui contient  $T_0$  et couvre tous les éléments qui sont subordonnés au noeud du début.

## 3 Etapes de la génération intégrée localisée

Dans cette section, nous présentons le déroulement de la génération d'un document par étapes. A chaque étape, les propositions dont les décisions linguistiques qui concernent leur réalisation sont dépendantes vont être produites. Le découpage en étapes est fondé sur la segmentation du discours présentée ci-dessus. L'attachement des propositions produites durant une étape entre elles et avec les autres propositions des différentes étapes de génération se base sur la structure relationnelle du discours cohérent de SDRT. La génération est implémentée en exploitant l'alignement des structures relationnelles de SDRT et celle des segments de G&S : la génération se réalise en construisant de manière descendante et par étapes la structure SDRT du document, où à chaque étape nous construisons une partie de la structure. Dans la suite, nous détaillons le processus.

### Entrée de la génération

Etant dans le contexte de génération de document, le point de départ de la génération est la réalisation d'un but communicatif. Nous considérons que l'entrée de la génération est un plan de buts communicatifs comparable à la structure de G&S : Le but communicatif principal domine des but communicatifs intermédiaires (second niveau de buts communicatifs) qui participent à sa satisfaction. A leur tour, chacun des buts communicatifs du second niveau peut dominer des buts communicatifs de troisième niveau et ainsi de suite.

### Description du processus

Dans la suite, nous appelons segment cible le segment (dans le sens de G&S) qu'on cherche à générer. Le segment cible peut contenir d'autres segments qu'on appelle segment(s) intermédiaire(s) (la structure des segments est connue car elle est réflexion du plan des buts communicatifs). Le déroulement de la génération d'un document est le suivant : La génération débute au niveau du but communicatif principal et se réalise de manière descendante. A ce niveau, le

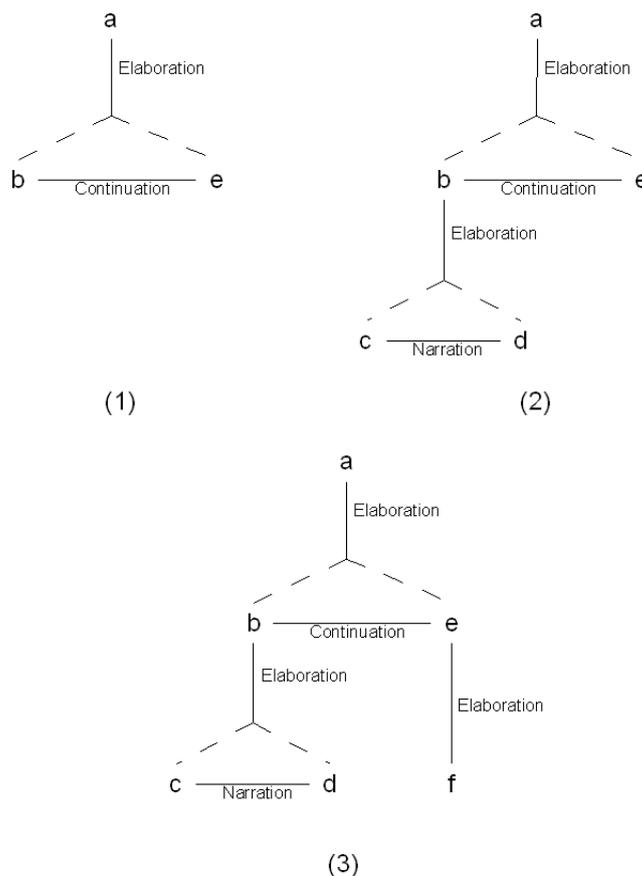


FIG. 5 – Génération progressive

segment cible de la génération représente le discours global. Le but du générateur est alors de produire les éléments linguistiques qui sont dépendants les uns des autres, et de laisser pour une étape ultérieure la production des éléments qui ne les affectent pas. Notamment, les éléments internes d'un segment intermédiaire n'ont pas d'incidence sur la génération des éléments du noyau du segment cible ou sur la génération des éléments d'un autre segment intermédiaire. Dans ce sens, le générateur va produire les propositions du noyau cible, et les propositions qui permettent d'ancrer les segments intermédiaires. Ceci revient à générer la partie supérieure de la structure relationnelle du segment qui contient les éléments du noyau du segment cible et les topiques des segments intermédiaires. Dans la suite, nous appelons ce groupe de propositions le noyau étendu du segment cible. Le processus de génération se poursuit en passant à un des segments intermédiaires (qui est déjà ancré à ce stade) qui devient le segment cible de la génération. La génération se complète en  $n$  tours, où  $n$  représente le nombre de segments dans le texte qui est le nombre des buts communicatifs dans le plan de communication.

La figure 5 représente la génération progressive du texte de l'exemple de la figure 1 en trois étapes (le texte ayant trois segments). Le premier groupe d'éléments générés contient le noyau étendu de  $DS_0$  (le noyau du segment  $DS_0$  et les topiques de  $DS_1$  et  $DS_2$ ). Les deuxième et troisième groupes constituent respectivement les suites des segments  $DS_1$  et  $DS_2$ .

## 4 Description d'une étape de génération

La génération du groupe de propositions correspondant au noyau étendu d'un segment cible se déroule en deux phases. Premièrement, le contenu sémantique pertinent est identifié. Il est représenté par un ensemble d'unités sémantiques candidates à l'inclusion dans le noyau étendu. Deuxièmement, la génération de la forme linguistique du noyau étendu est réalisée avec une architecture intégrée des tâches. Pour réaliser une génération à architecture intégrée, nous proposons d'implémenter l'ensemble des types de décisions linguistiques de la génération (sélection du contenu, la structuration, la lexicalisation, la pronominalisation) dans un CSP (problème de satisfaction de contraintes, (Tsang, 1993)). Les tâches de la génération sont alors modélisées par un ensemble de variables, les domaines sur lesquelles portent les variables et un ensemble des contraintes sur les variables. (HANKACH P., 2008) et (Moriceau & Saint-Dizier, 2005) implémentent une modélisation CSP des tâches de la génération. Suite à cette modélisation, un solveur CSP calcule la forme linguistique du noyau étendu (à partir des unités sémantiques candidates trouvés durant la première phase de sélection de contenu) en considérant l'ensemble des contraintes à la fois. Dans ce qui suit, nous allons présenter la première phase de sélection du contenu.

### **Sélection de contenu sémantique approprié pour une étape de génération**

La première étape de la génération d'un noyau étendu du segment cible est d'identifier le contenu sémantique pertinent à inclure. Nous proposons de réaliser une expansion relationnelle pour isoler l'ensemble d'unités sémantiques candidates à l'inclusion dans le noyau étendu d'un segment. La décision finale de l'inclusion d'une unité sémantique ainsi que son positionnement et sa verbalisation vont être prises durant la phase suivante de la génération du noyau étendu. L'expansion relationnelle est réalisée de la manière suivante : Le topique du segment cible constitue le centre d'expansion. En partant du centre d'expansion, nous sélectionnons toutes les unités sémantiques qui sont reliées à lui par une relation sémantique quelconque. Le même processus est répété avec les unités sémantiques nouvellement sélectionnées. Les unités retenues comme candidates à l'inclusion dans la forme finale du noyau étendu sont les suivantes :

- Les unités d'une chaîne d'expansion qui précède le topique d'un segment intermédiaire (le topique est inclus) si la distance entre le centre d'expansion et le topique est inférieure à une borne supérieure de  $N$  éléments.
- Les  $k$  premiers éléments des autres chaînes

L'expansion relationnelle permet d'identifier les unités de contenu sémantiquement reliées au topique du segment cible. Elle permet aussi d'établir la liaison avec les topiques des segments intermédiaires ce qui permet leur ancrage. La figure 6 illustre l'expansion relationnelle.

## 5 Conclusion

Nous avons présenté une nouvelle approche de construction d'un système de génération basée sur une génération intégrée localisée, où le traitement intégré des décisions linguistiques est limité à la production des propositions dont les décisions linguistiques qui concernent leur réalisation sont dépendantes. Notre approche permet de résoudre le problème de complexité de la génération intégrée classique. De ce fait, elle fournit ainsi une alternative viable à la génération séparée (séquentielle ou interactive) qui présente plusieurs défauts mais qui est implémentée de manière répandue dans les systèmes de générations existants. De plus notre approche de génération présente les caractéristiques suivantes : Elle permet de combiner les générations

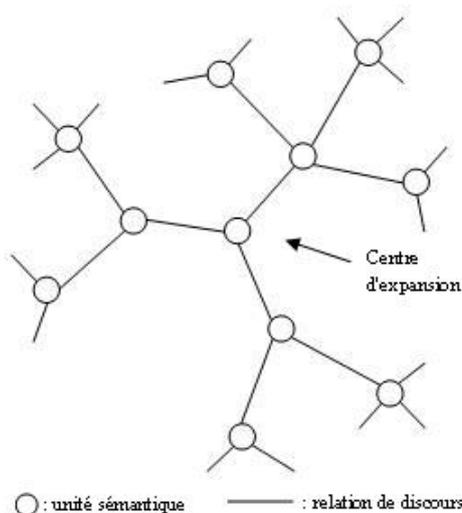


FIG. 6 – Expansion relationnelle.

descendante et ascendante. L'initiation des étapes de la génération se fait de manière descendante en partant des buts communicatifs, alors que la réalisation de la génération dans une étape est ascendante. Enfin, si besoin, la génération peut s'arrêter à n'importe quelle étape, ce qui peut être utile dans les cas d'applications en temps réel par exemple.

## Références

- APPELT D. (1985). *Planning English Sentences*. Cambridge, England : Cambridge University Press.
- ASHER N. & LASCARIDES A. (2003). *Logics of Conversation*. Cambridge : Cambridge University Press.
- DANLOS L. (1984). Conceptual and linguistic decisions in generation. *Proceedings of COLING-84*, p. 501–504.
- GROSZ B. J. & SIDNER C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, **12**(3), 175–204.
- HANKACH P. D. L. . P. F. (2008). Sdrt based csp problem planning approach for merging predefined messages in human-machine interaction systems. *Proceedings of INFOS 2008*.
- HOVY E. (1993). Automated discourse generation using discourse structure relations. *Artificial Intelligence*, **63**, 341–385.
- KANTROWITZ M. & BATES J. (1992). Integrated natural language generation systems. In R. DALE, E. HOVY, D. RÖSNER & O. STOCK, Eds., *Aspects of Automated Natural Language Generation*, p. 13–28. Berlin : Springer.
- MORICEAU V. & SAINT-DIZIER P. (2005). *A Constraint-Based Model for Lexical and Syntactic Choice in Natural Language Generation*, volume 3438 of LNCS - LNAI. <http://www.springerlink.com> : Springer. Revised selected papers of CSLP 2004.
- REITER E. & DALE R. (2000). *Building natural language generation systems*. Cambridge University Press.
- TSANG E. (1993). *Foundations of Constraint Satisfaction*. London : Academic Press.