



# Graphes patrons pour représenter des parcours d'analyse de corpus audiovisuels. Exemple : Le domaine DLC (= Diversité Linguistique et Culturelle)

Peter Stockinger

## ► To cite this version:

Peter Stockinger. Graphes patrons pour représenter des parcours d'analyse de corpus audiovisuels. Exemple : Le domaine DLC (= Diversité Linguistique et Culturelle). [Rapport de recherche] 12, ESCoM-AAR (Equipe Sémiotique Cognitive et Nouveaux Médias - Archives Audiovisuelles de la Recherche). 2006, pp.22. <hal-01330721>

**HAL Id: hal-01330721**

**<https://hal.archives-ouvertes.fr/hal-01330721>**

Submitted on 12 Jun 2016

**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.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0  
International License

## Programme RIAM et ANR

### Projet de R&D

**SAPHIR** - Système d'Assistance à la Publication Hypermédia  
par spécification Intentionnelle et modélisation Rhétorique -

# **Graphes patrons pour représenter des parcours d'analyse de corpus audiovisuels.**

**Exemple : Le domaine DLC (= Diversité Linguistique et  
Culturelle)**

(Rapport de travail)

**Peter Stockinger**

Institut National des Langues et Civilisations Orientales (INALCO)

Paris, ESCoM-AAR, Octobre 2006

## Table des matières

<b>INTRODUCTION .....</b>	<b>3</b>
<b>PARTIE I : DESCRIPTION DES SITUATIONS DE COMMUNICATION (SCIENTIFIQUE).....</b>	<b>5</b>
1/ LA SITUATION « ÉLÉMENTAIRE » .....	5
2/ INTÉGRATION D'UN DISCOURS DI DANS UN DISCOURS D PLUS GLOBAL FAISANT PARTIE D'UNE COMMUNICATION SCIENTIFIQUE .....	7
<b>PARTIE II : PARCOURS « PARLER D'UNE LANGUE <i>EN ELLE-MÊME</i> ».....</b>	<b>9</b>
1/ ETAPES DU PARCOURS DE DESCRIPTION .....	9
2/ GRAPHES PATRONS .....	10
<b>PARTIE III : PARCOURS « PARLER DE LA DIVERSITÉ LINGUISTIQUE DANS LE MONDE ACTUEL ET DANS L'HISTOIRE » .....</b>	<b>15</b>
1/ ETAPES DU PARCOURS DE DESCRIPTION .....	15
2/ GRAPHES PATRONS .....	16
<b>PARTIE IV : TROIS OPTIONS POUR LA DESCRIPTION – INDEXATION D'UN CORPUS AUDIOVISUEL À L'AIDE D'UN OU D'UN ENSEMBLE DE GRAPHES PATRONS.....</b>	<b>21</b>

## Introduction

A l'aide du **vocabulaire des objets d'analyse**<sup>1</sup> élaboré pour parler du monde « Diversité Linguistique et Culturelle » (DLC)<sup>2</sup>, nous sommes en mesure de spécifier des situations types ou encore des types de situations assez récurrentes qui caractérisent le domaine de connaissances DLC ainsi que les « attitudes », c'est-à-dire les manières selon lesquelles les chercheurs (témoins, ...) les voient, les « mettent en scène » et en parlent ....

Ces *situations* et *attitudes* nous servent à élaborer un « catalogue raisonné » de **graphes patrons** pouvant nous servir à une description-indexation plus facile et rapide du corpus audiovisuel PCI qui fait partie du [programme AAR \(Archives Audiovisuelles de la Recherche\)](#).

Nous distinguons donc entre 1) les graphes patron qui explicitent les « **façons** » **de parler** (i.e. les situations de discours) et 2) les graphes patron qui explicite les **situations dont on parle**.

Les graphes patron, eux, font partie de **parcours** typiques de **description-indexation**. Comme son nom l'indique, un parcours de description-indexation propose à un interprète une succession de questions qu'il/qu'on peut poser à un

---

<sup>1</sup> Il s'agit de la première version de l'**ontologie ASA** (Atelier de Sémiotique Audiovisuelle pour l'analyse de corpus de vidéos en SHS), réalisée dans le cadre des deux projets de R&D – [LOGOS](#) (projet européen financé dans le cadre du 6<sup>ème</sup> FP entre 2006 et 2009) et [SAPHIR](#) (projet français financé par l'ANR entre 2006 et 2009). Cette première version est décrite dans le rapport suivant : P. Stockinger, [The PCI Ontology. Presentation and discussion of an ontology for the description and indexing of audiovisual resources in the field of cultural heritage of minorities and indigenous people](#); Paris, ESCoM-AAR 2007 (DOI: 10.13140/2.1.1180.5129). L'ontologie ASA était l'objet de plusieurs projets de recherche et de publications et servait surtout à l'analyse (classification, description, interprétation, ...) de corpus audiovisuels dans le cadre du [Programme AAR \(Archives Audiovisuelles de la Recherche\)](#). Depuis 2015, elle fait partie de la [plateforme Campus AAR](#), déployée sur les serveurs du CCSD/HAL et mise à la disposition de la communauté des chercheurs et enseignants en SHS.

<sup>2</sup> Une réalisation expérimentale sous forme d'un portail d'analyse de corpus audiovisuels consacré au patrimoine culturel des minorités et des peuples indigènes est hébergée sur un des serveurs de la Fondation MSH à Paris à l'adresse suivante : <http://semioweb.msh-paris.fr/corpus/dlc/FR/>. Le maintien et la diffusion en ligne de cette réalisation dépend de la bonne volonté de l'institution qui l'héberge. Même s'il s'agit d'un établissement de valorisation et de diffusion des connaissances scientifique, l'institution en question ne garantit pas son archivage et encore moins sa diffusion pérenne.

document audiovisuel tandis que le graphe patron précise le contenu en lui-même de la question.

Un « **parcours de description** » prend donc en compte

- 1/ de quoi on parle (= situations types),
- 2/ comment on en parle (= façons d'en parler : exposé > *description* ou *exemple*)
- 3/ étapes de décrire 1 et 2

Avec un peu d'imagination peut-on affirmer que l'activité de création d'une part des graphes patron et d'autre part des parcours de description, s'apparente à celle de la spécification de l'**univers sémantique** d'un ensemble de (nouveaux) documents tandis que l'activité de l'indexation d'un corpus de documents audiovisuels à l'aide des graphes patron s'apparente à celle de l'**instanciation** de l'univers sémantique dans un (nouveau) document.

Il y a une très grande diversité de sujets à travers lesquels les chercheurs thématisent tel ou tel aspect de l'univers DLC. En effet, il y a au moins autant de types de sujets qu'il y a des thèmes composant l'ontologie du domaine ci-dessus.

Nous identifierons, pour commencer quelques types de sujets qui nous paraissent particulièrement intéressants sachant que ce « catalogue » de sujets peut être parfaitement compléter par des nouvelles définitions plus tard. Voici le premier catalogue de parcours :

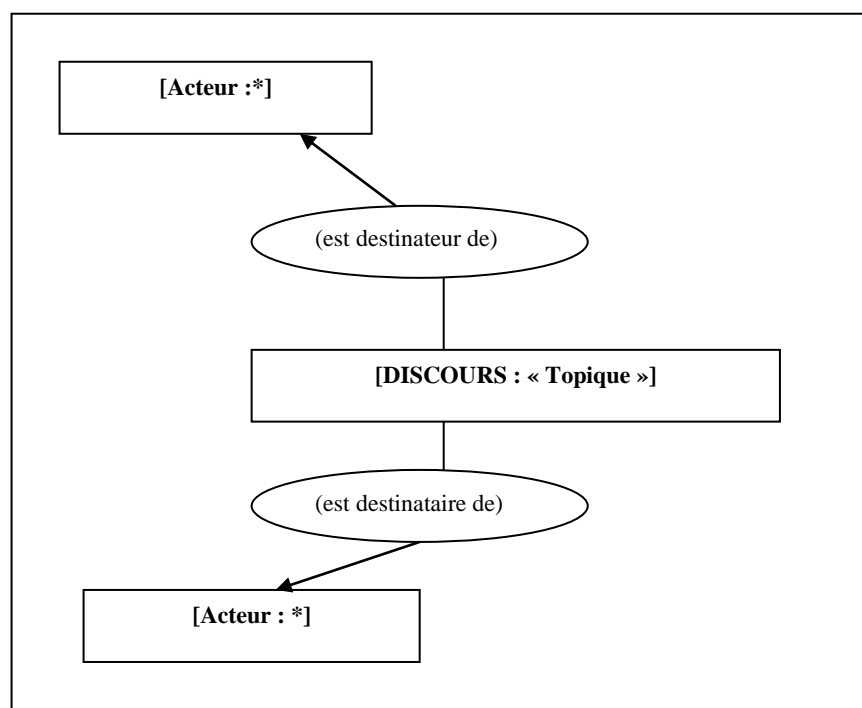
1. Parcours « Parler d'une langue *en elle-même* ».
2. Parcours « Parler de la diversité linguistique dans le monde actuel et/ou dans l'histoire ».
3. Parcours « Parler de l'évolution linguistique ».
4. Parcours « Parler du patrimoine linguistique ».
5. Parcours « Parler du rapport entre langue et culture ».

Pour le moment ne sont élaborés que les deux premiers parcours.

## Partie I : Description des situations de communication (scientifique)

Il s'agit de décrire les situations de discours dans lesquelles on thématise, on parle d'un domaine de connaissances.

### 1/ La situation « élémentaire »



(figure 1)

Figure 1 représente une situation évidente pour tout type de communication (scientifique) – elle n'est pas particulière au domaine DLC :

1. le chercheur X
2. produit un discours D
3. ayant comme objet la topique T
4. et dont le destinataire est Y.

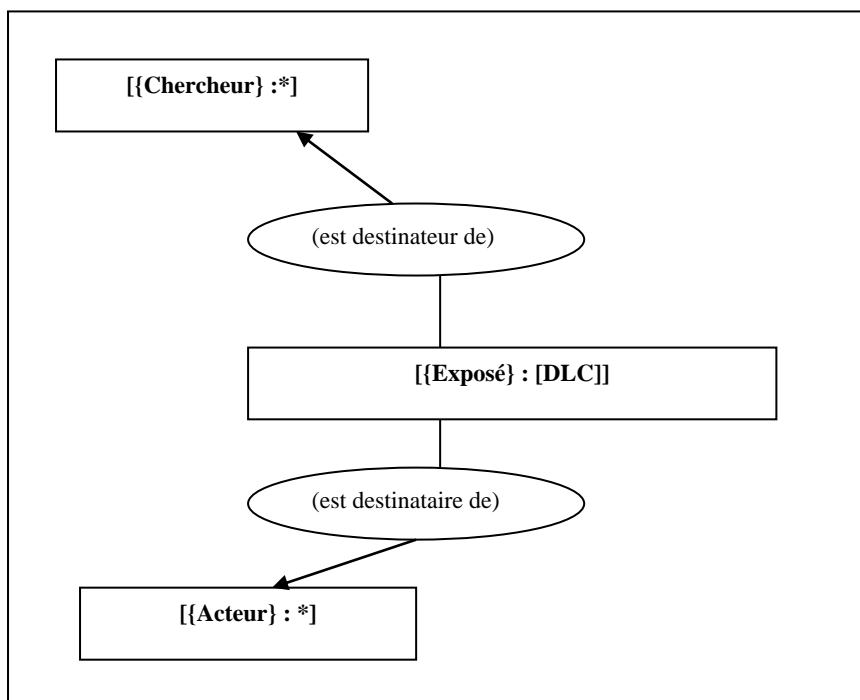
Par défaut, l'énonciateur est le **chercheur** produisant une conférence ou un entretien, sauf dans le cas des citations (l'énonciateur est la personne citée) et du discours indirect (libre).

Le destinataire est, par défaut le « on », i.e. une *communauté* pas plus définie mais, pragmatiquement parlant, composées d'acteurs dont on suppose qu'ils ont un intérêt à priori et les compétences nécessaires pour suivre le discours du chercheur.

Le discours D, lui, est par défaut **expositif**, c'est-à-dire un **exposé**. La figure (2) montre la situation « par défaut ».

Ce graphe peut être très utile pour différentes raisons :

- constitution d'une bibliothèque de citations particulièrement intéressantes de chercheurs interviewés ;
- extraction, dans un interview, de citations ou encore de références à d'autres discours ;
- etc.



(figure 2)

**Note :** Les accolades « { » et « } » symbolisent le fait que l'indexeur a le choix entre plusieurs thèmes/concepts. Par exemple, quand il s'agit d'un discours qui exemplifie, il doit avoir la possibilité de modifier le thème [Exposé] et de choisir le thème [Exemple]. De même le thème [Chercheur] doit être remplaçable par d'autres thèmes tels que [Témoin], [Enseignant], ...

**Note :** Comme déjà dit, la figure 1 ne représente que les critères les plus élémentaires pour décrire la situation de discours. Il faudra donc prévoir la possibilité de pouvoir utiliser d'autres critères pour, si nécessaire, produire une description plus fine de la situation du discours. Nous pensons notamment aux critères suivants :

1. **modalité épistémique** (par défaut : neutre) : certitude, doute, ...
2. **modalité appréciative** (par défaut : neutre) : réserve, adhésion, rejet, ...
3. **portée de la validité** (par défaut : neutre) :
4. **temps T** (par défaut : le moment de l'entretien, de la conférence ; i.e. de la production de l'événement scientifique)
5. **lieu L** (par défaut : le lieu, « l'adresse » de production de la conférence)
6. **référence C** (par défaut : connaissances et travaux antérieurs du chercheur lui-même) : connaissances scientifiques, personnalités.



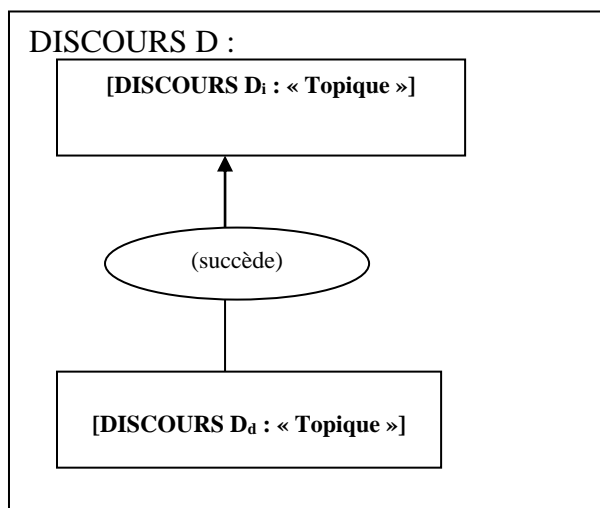
Note : même si l'indexeur ne s'en sert pas (régulièrement, systématiquement), ces sept critères doivent au moins être disponibles (par exemple : sous forme d'un choix « description avancée »).

## 2/ Intégration d'un discours $D_i$ dans un discours $D$ plus global faisant partie d'une communication scientifique

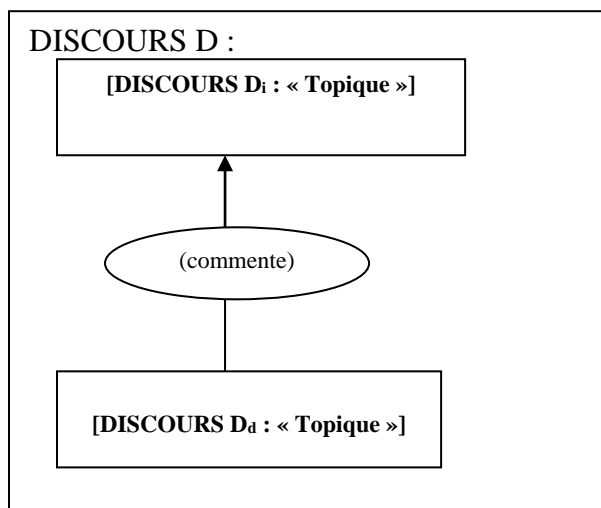
Comme déjà dit, le *discours expositif* (figure 1) peut être considéré en lui seul. Mais, il ne faut pas oublier qu'il « répond » à d'autres discours tout en les prolongeant. Nous distinguons entre deux types d'intégration de plusieurs situations de discours en un discours global faisant partie d'une communication scientifique du genre « entretien », « conférence », etc. (cf. *supra*, partie V) :

- 1/ l'intégration narrative lato sensu qui concerne avant tout le développement d'une topique et qui s'articule à travers les relations narratives utilisées pour « mettre en scène » un domaine ;
- 2/ l'intégration rhétorique sous forme d'une « prise de position » - d'une « commentaire » - par rapport à un autre discours déjà tenu (par le même auteur ou par un autre auteur)

La figure 3 montre un des cas les plus typiques pour (1), à savoir la succession de plusieurs discours formant ensemble un discours global consacrée à une topique donnée. La figure 4 visualise la configuration rhétorique entre deux discours formant un discours global commun – la relation « commenter » pouvant être remplacée, spécialisée par les relations narratives de l'argumentation (cf. l'ontologie des relations).



(figure 3)



(figure 4)

## Partie II : Parcours « Parler d'une langue *en elle-même* »

Il s'agit d'un parcours typique de description visant l'extraction systématique des informations relatives à une langue donnée.

### 1/ Etapes du parcours de description

#### Parcours en 6 étapes

**Etape 1 - Présentation synthétique de la langue :** identifier la langue, en faire une présentation globale et générale, produire quelques données ...

**Etape 2 – Aspects linguistiques de la langue :** chercher des informations relatives à la structure « interne » de la langue (de sa grammaire, de son lexique, ...) et de son « être linguistique » (appartenance typologique, géographie linguistique, ...)

**Etape 3 – Aspects sociolinguistiques de la langue :** chercher des informations relatives à la place de la langue dans la vie et la culture d'une communauté, d'un acteur, ...

**Etape 4 – Aspects historiques de la langue :** chercher des informations relatives à l'évolution de la langue, à son histoire, ...

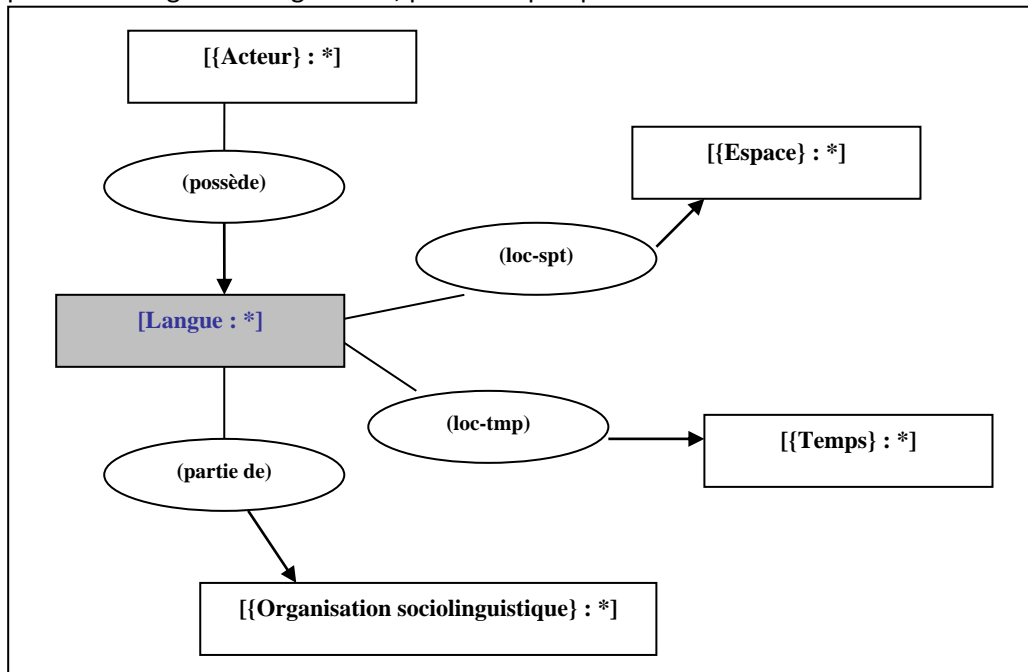
**Etape 5 – Aspects didactiques de la langue :** chercher des informations relatives à l'enseignement et à l'apprentissage de la langue ...

**Etape 6 – Connaissances scientifiques de la langue :** enfin, chercher des informations relatives à la recherche linguistique elle-même consacré à la langue

Ces étapes peuvent être complétées par d'autres ce qui signifie qu'un indexeur « spécialiste » doit avoir la possibilité de définir des nouveaux graphes patrons et les ajouter à la liste des graphes patrons déjà existants ...

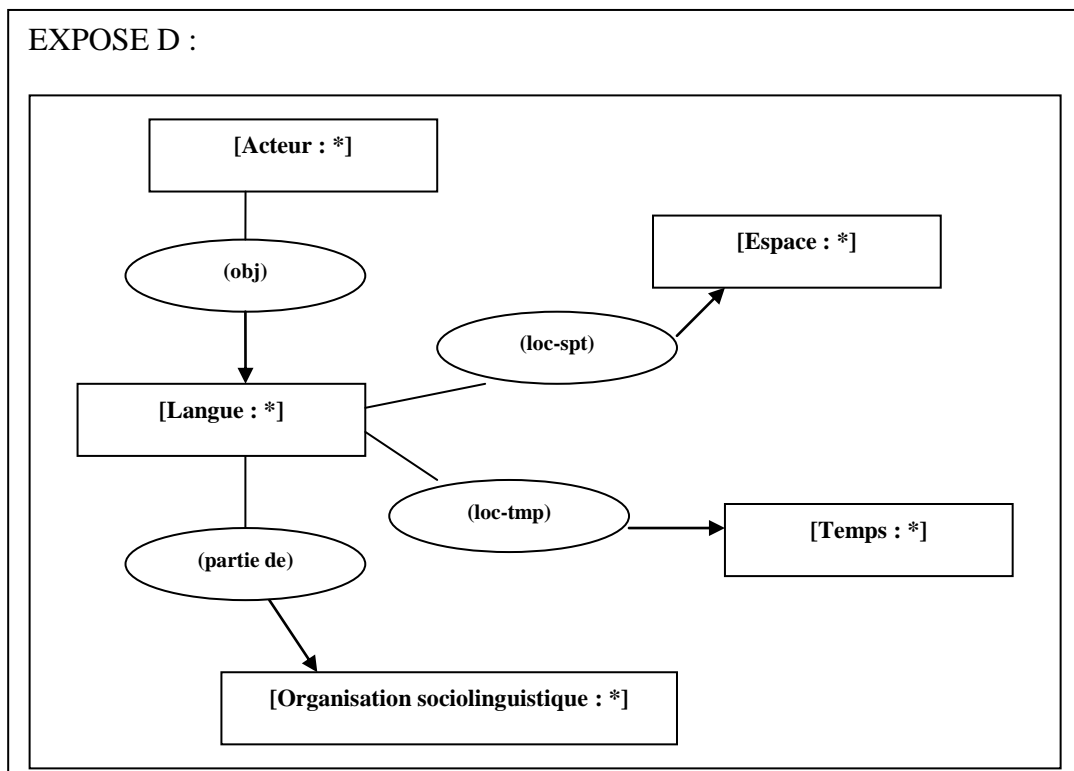
## 2/ Graphes patrons

**Etape 1 - Présentation synthétique de la langue :** identifier la langue, en faire une présentation globale et générale, produire quelques données ...



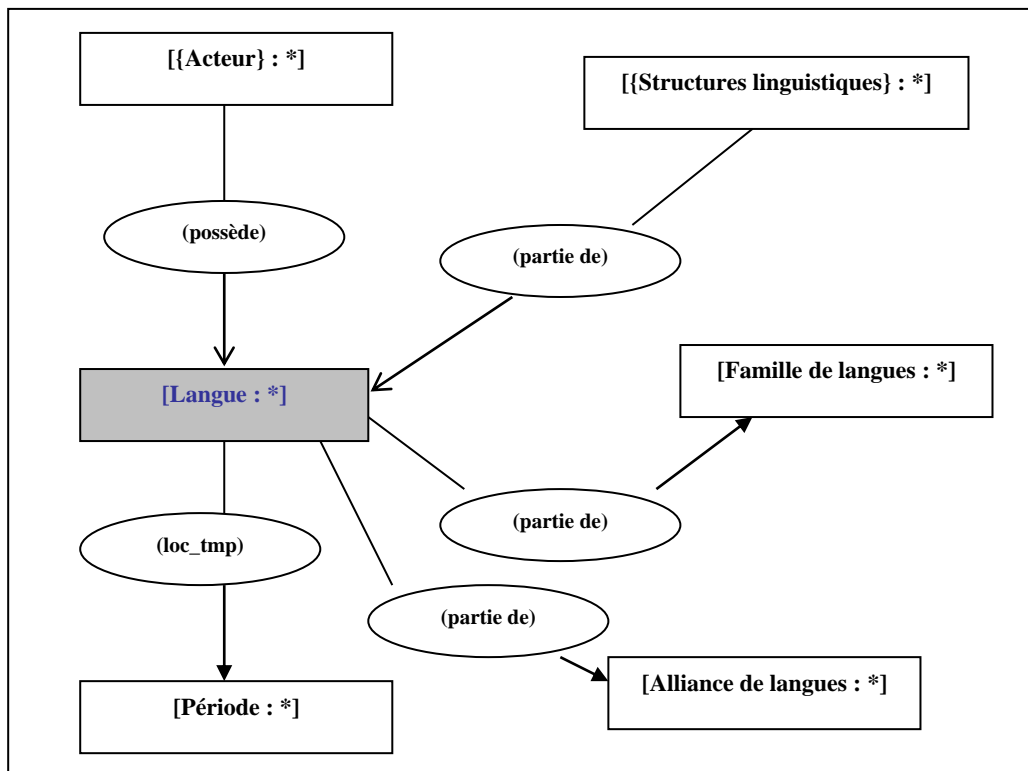
(figure 5)

**Note :** la boîte grisée symbolise le concept/thème central du graphe patron ...



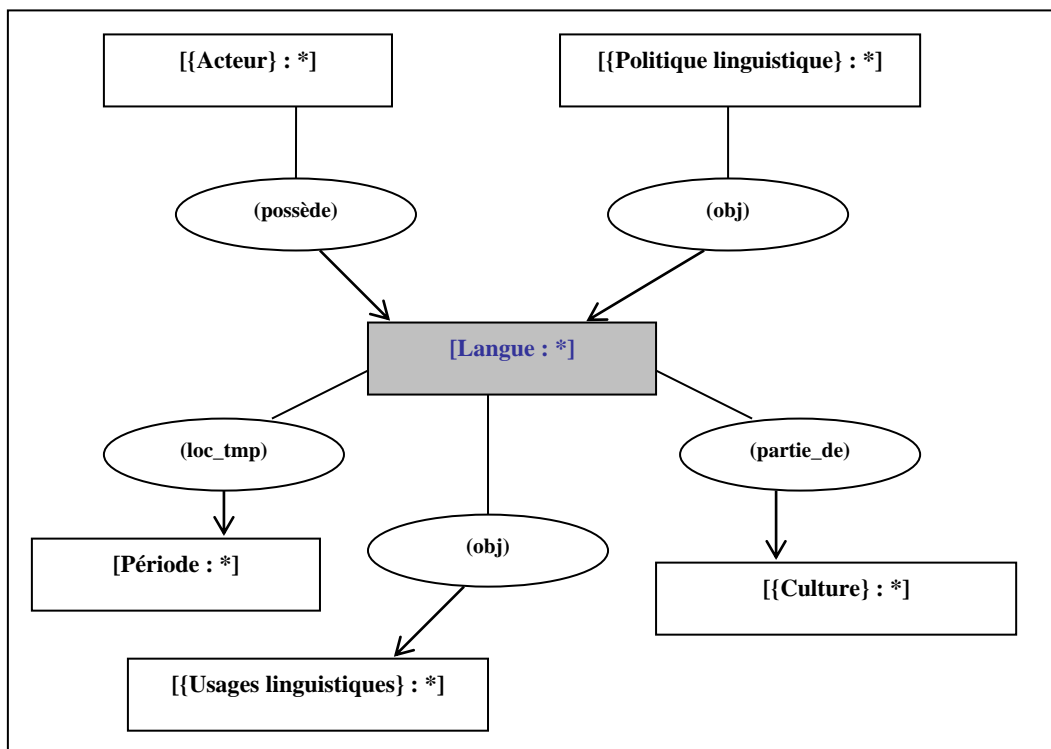
(figure 6)

**Etape 2 – Aspects linguistiques de la langue :** chercher des informations relatives à la structure « interne » de la langue (de sa grammaire, de son lexique, ...) et de son « être linguistique » (appartenance typologique, géographie linguistique, ...).



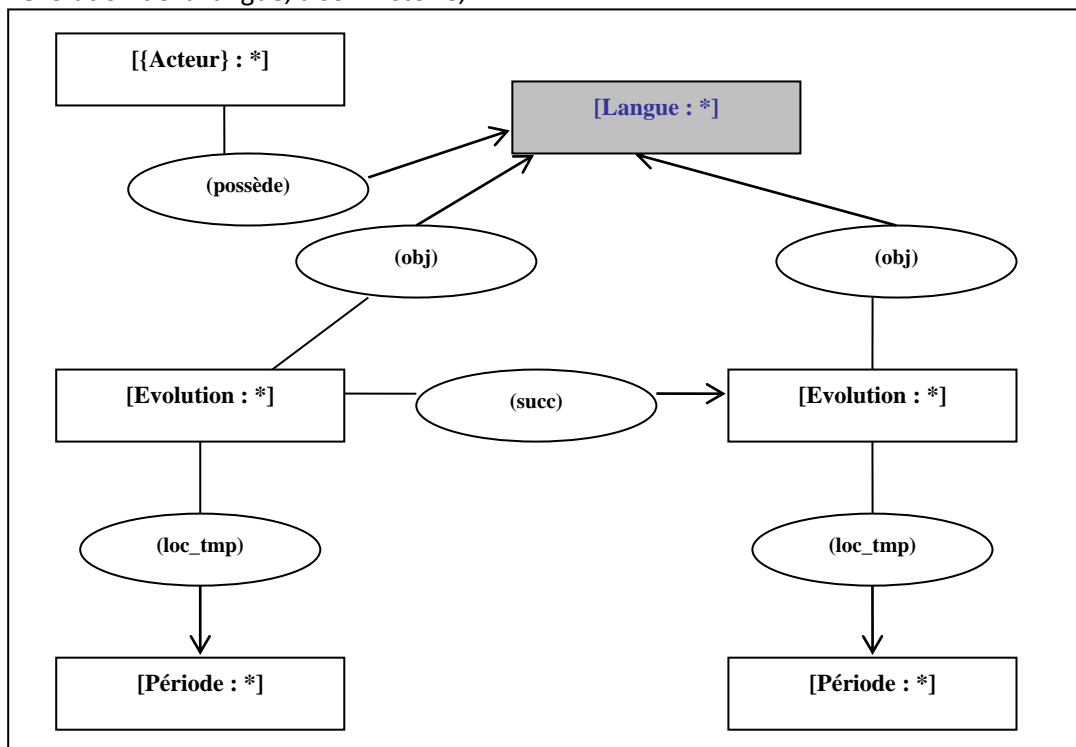
(figure 7)

**Etape 3 – Aspects sociolinguistiques de la langue :** chercher des informations relatives à la place de la langue dans la vie et la culture d'une communauté, d'une personne, ...



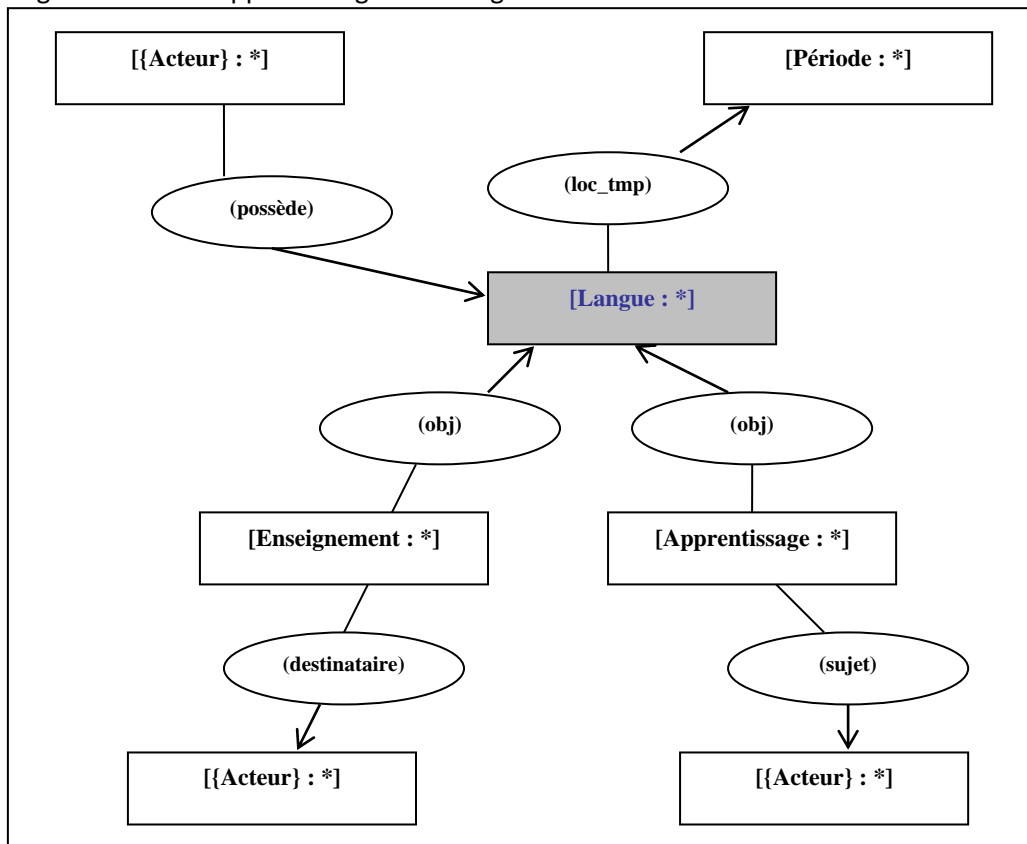
(figure 8)

**Etape 4 – Aspects historiques de la langue :** chercher des informations relatives à l'évolution de la langue, à son histoire, ...



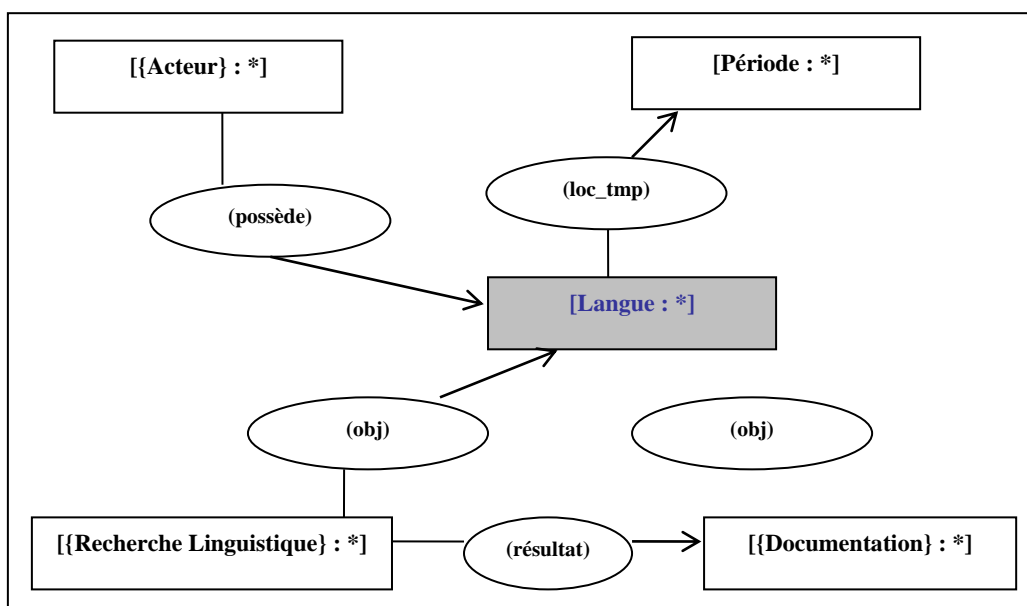
(figure 9)

**Etape 5 – Aspects didactiques de la langue :** chercher des informations relatives à l'enseignement et à l'apprentissage de la langue ...



(figure 10)

**Etape 6 – Recherche linguistique sur la langue :** enfin, chercher des informations relatives à la recherche linguistique elle-même consacré à la langue



(figure 11)

**Note :** Comme déjà dit, la fonction de ces graphes patrons est de faciliter et aussi « standardiser » la description d'un corpus audiovisuel selon un certain point de vue. Le principe est qu'ils peuvent être modifiés par l'auteur-interprète et qu'ils peuvent être enrichis par d'autres graphes patrons ....



## Partie III : Parcours « Parler de la diversité linguistique dans le monde actuel et dans l'histoire »

1/ Etapes du parcours de description

### Parcours : 6 Etapes (typiques)

**Etape 1 – De la diversité en général :** identifier la langue, en faire une présentation globale et générale, produire quelques données ...

**Etape 2 – Aspects linguistiques de la langue :** chercher des informations relatives à la structure « interne » de la langue (de sa grammaire, de son lexique, ...) et de son « être linguistique » (appartenance typologique, géographie linguistique, ...)

**Etape 3 – Aspects sociolinguistiques de la langue :** chercher des informations relatives à la place de la langue dans la vie et la culture d'une communauté, d'un acteur.

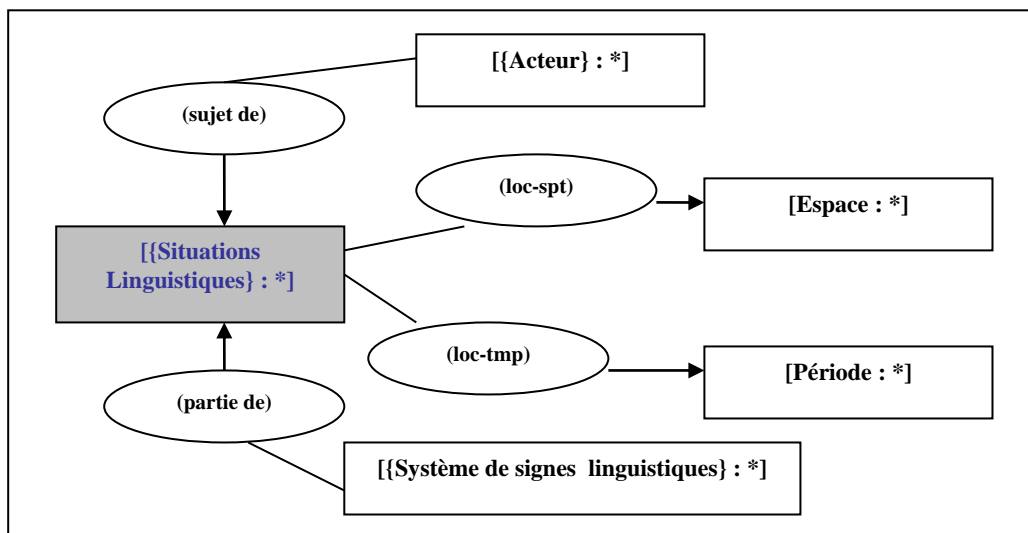
**Etape 4 – Aspects historiques de la langue :** chercher des informations relatives à l'évolution de la langue, à son histoire, ...

**Etape 5 – Aspects didactiques de la langue :** chercher des informations relatives à l'enseignement et à l'apprentissage de la langue ...

**Etape 6 – Connaissances scientifiques sur la langue :** enfin, chercher des informations relatives à la recherche linguistique elle-même consacré à la langue

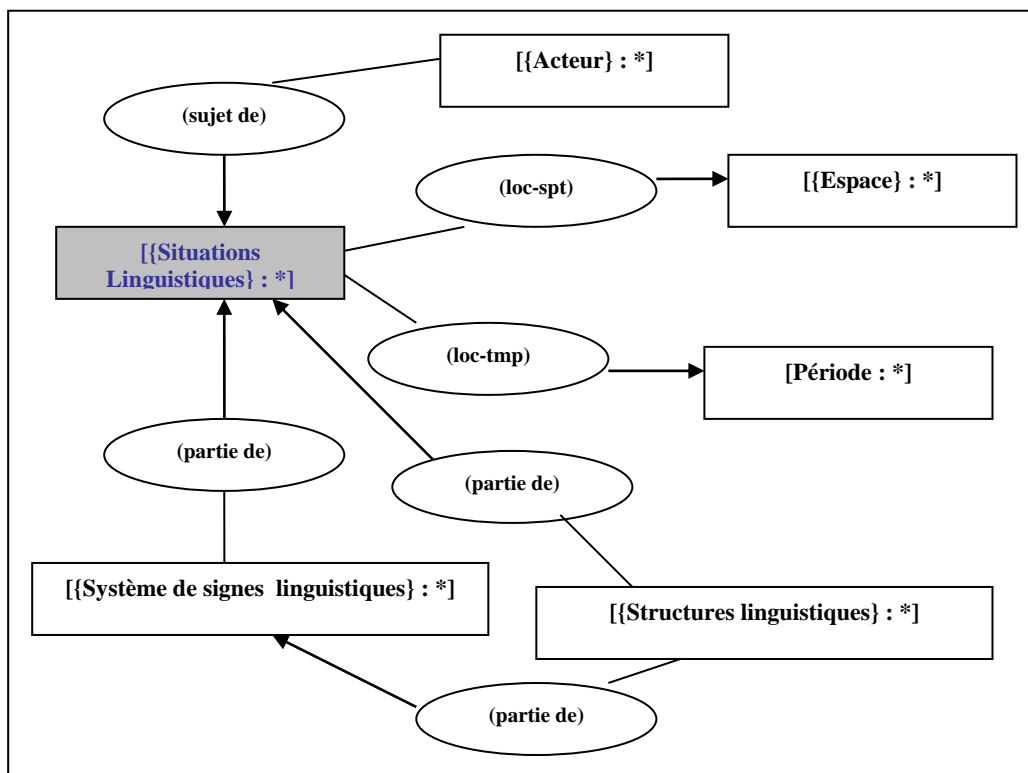
## 2/ Graphes patrons

**Etape 1 – De la diversité en général :** les situations linguistiques (familles de langues, alliances linguistiques, ...) dans l'espace, le temps et les groupes sociaux...



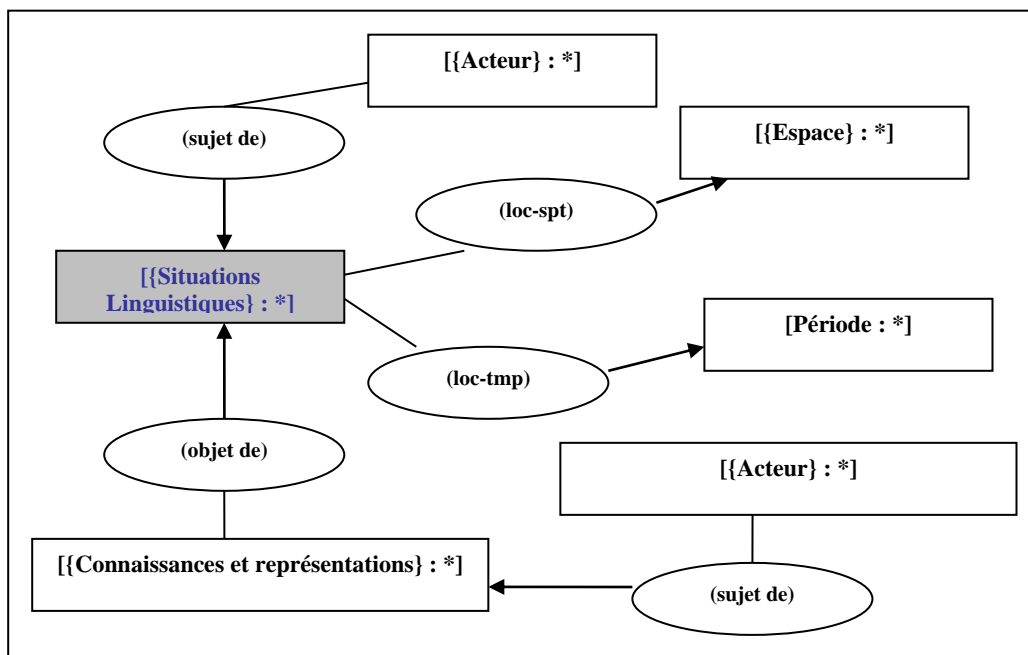
(figure 12)

**Etape 2 – Spécificités linguistiques de la diversité des langues :** les aspects grammaticaux, lexicaux, phonologiques, phonétiques, ... que les langues, à un moment donné, partagent ou possèdent en propre



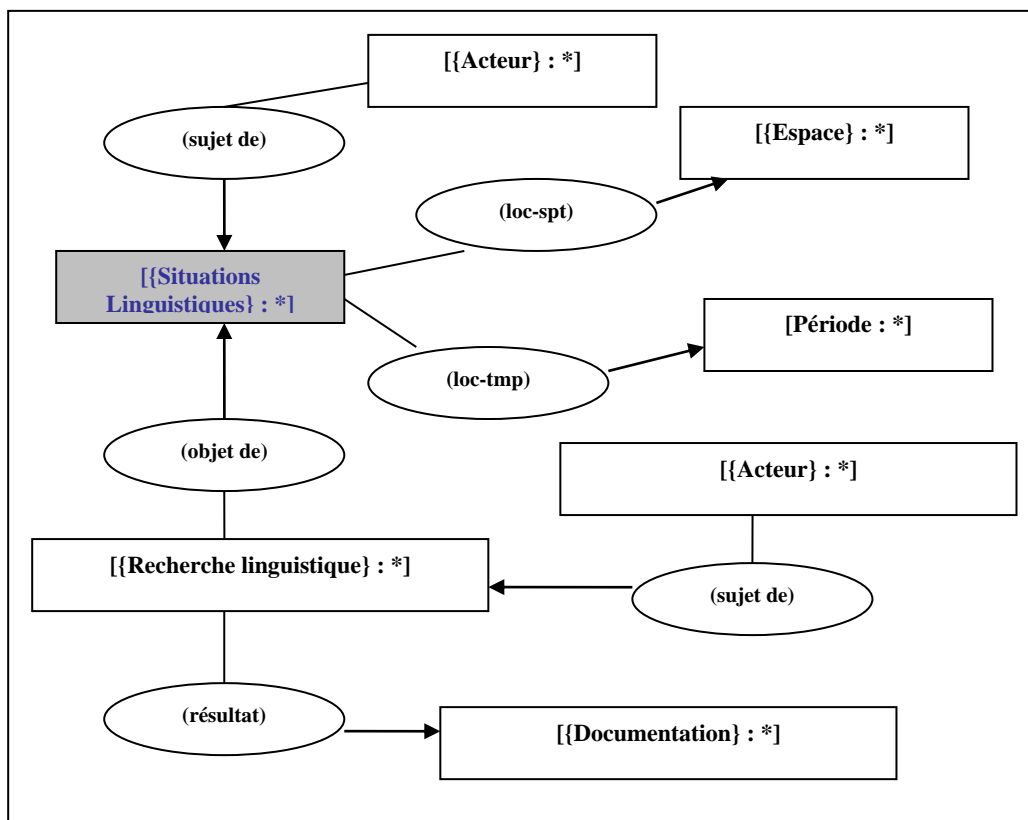
(figure 13)

**Etape 3 – Représentations (culturelles, mythologiques, populaires, ...) de la diversité des langues :** mythes, légendes, visions populaires, ... de la diversité linguistique ....



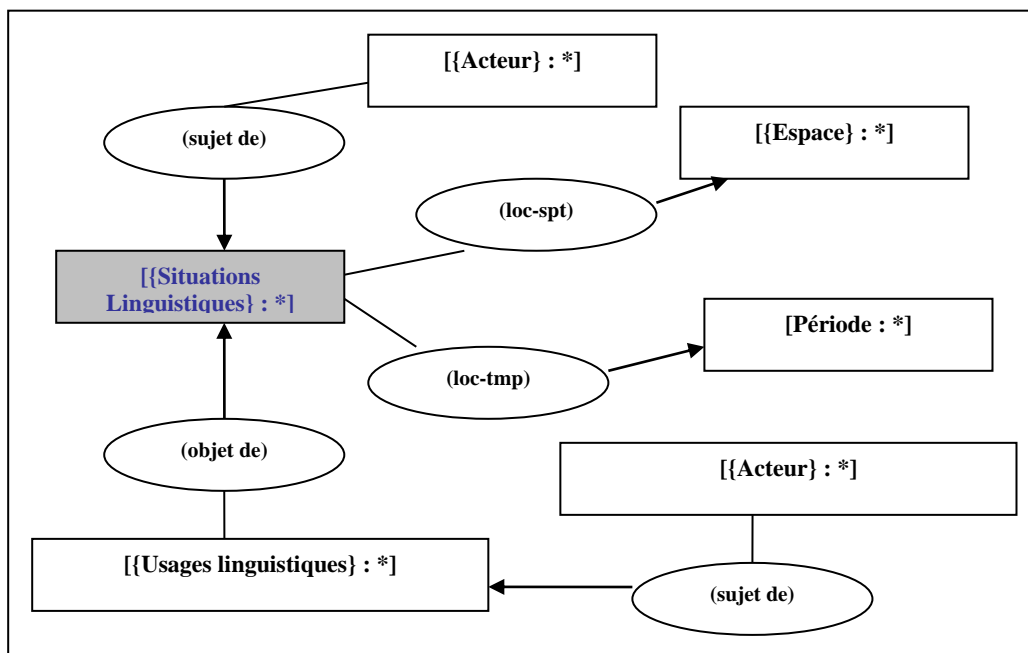
(figure 14)

**Etape 4 – Connaissances scientifiques et recherche sur la diversité linguistique:** mythes, légendes, visions populaires, ... de la diversité linguistique ....



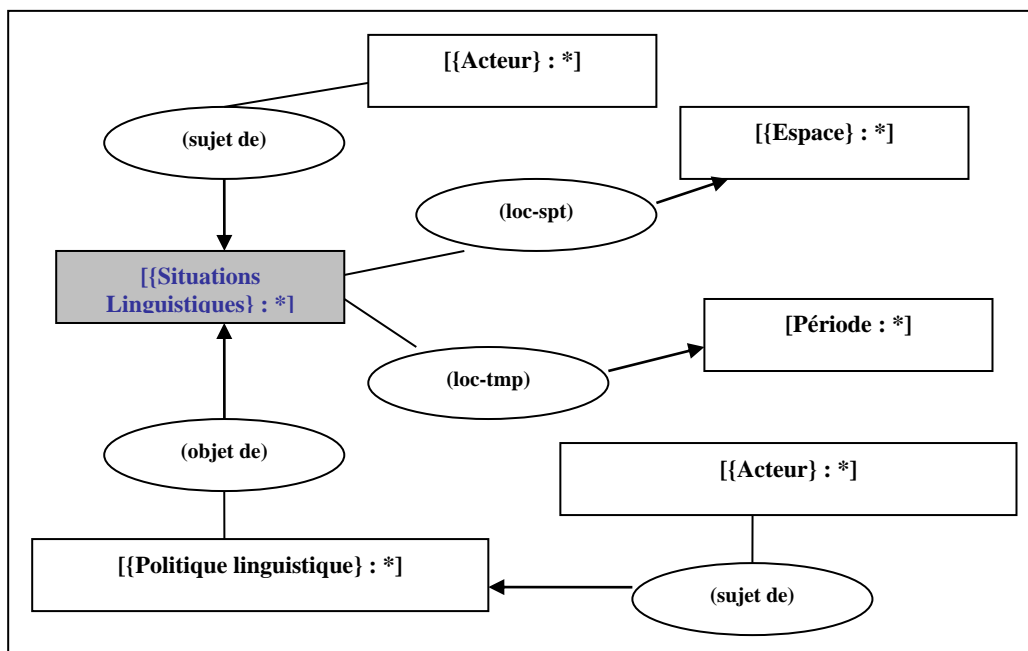
(figure 15)

**Etape 5 – Pratiques d’expression et de communication dans un contexte de diversité linguistique:**



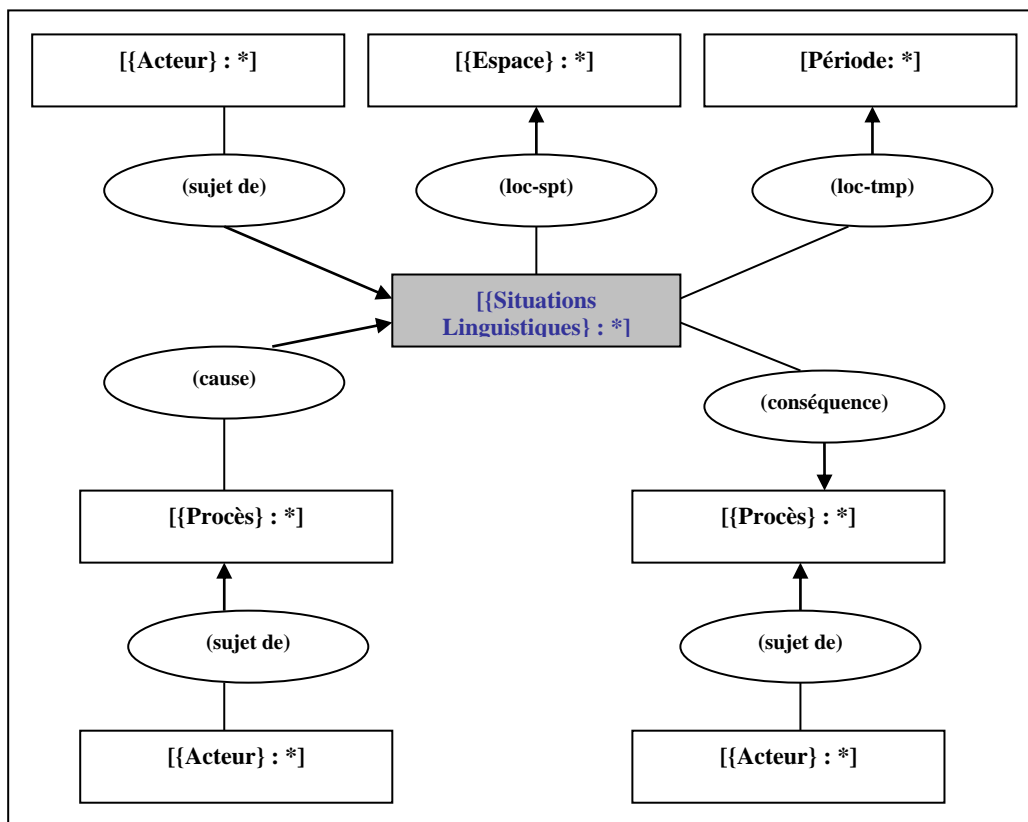
(figure 16)

**Etape 6 – Politique linguistique et diversité linguistique : interventions intentionnelles, pilotage, aménagement, valorisation et sauvegarde, ....**



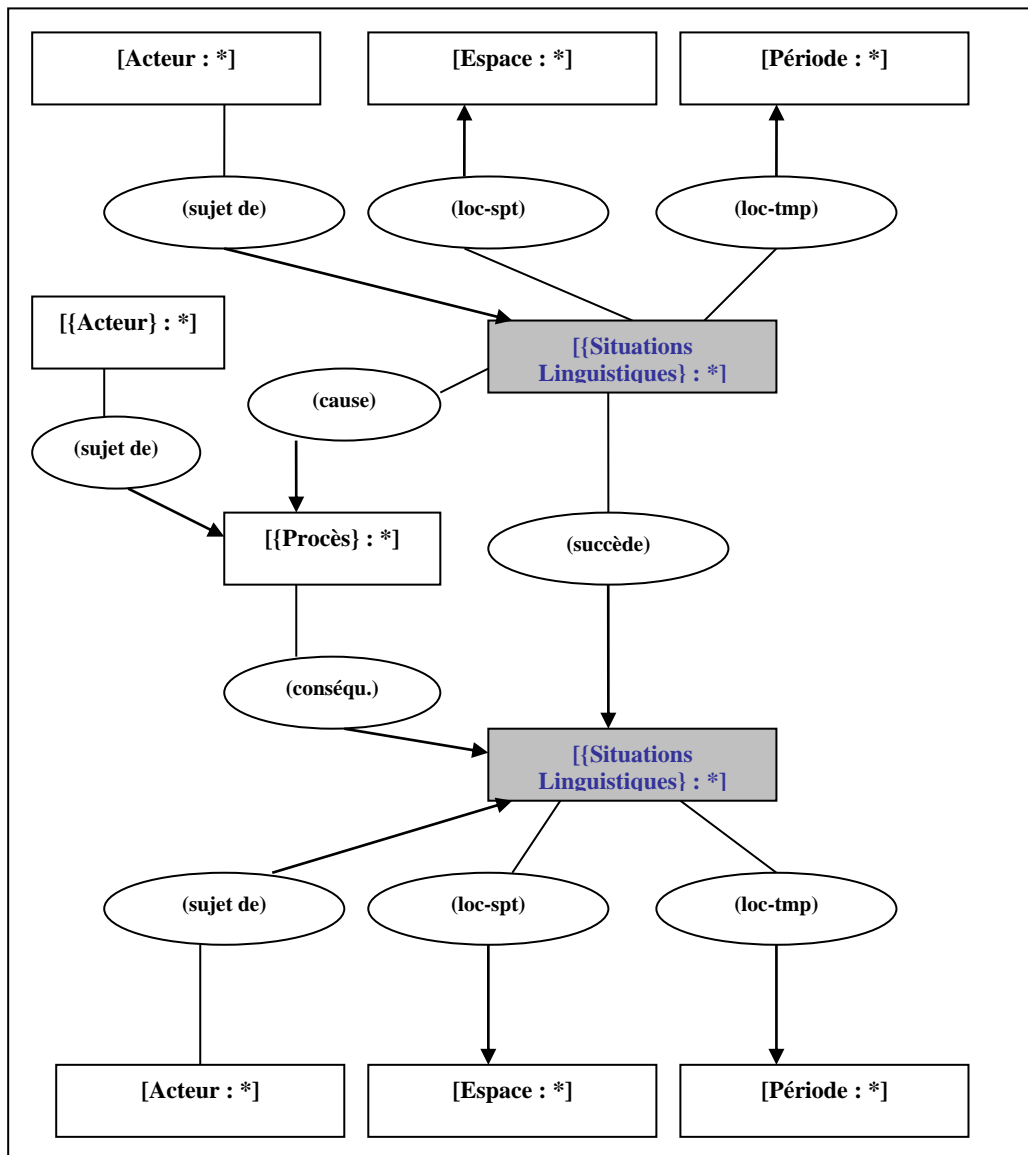
(figure 17)

**Etape 7 – Causes et conséquences de la diversité linguistique : mythes, légendes, visions populaires, ... de la diversité linguistique ....**



(figure 18)

**Etape 8 – Evolution et diversité linguistique : mythes, légendes, visions populaires, ... de la diversité linguistique ....**



(figure 19)

## Partie IV : Trois options pour la description – indexation d'un corpus audiovisuel à l'aide d'un ou d'un ensemble de graphes patrons

La description-indexation d'un corpus de documents audiovisuels peut se concevoir (à mon avis ...) selon les trois options suivantes :

<b>Première option : description – indexation « ponctuelle » (en dehors d'un parcours)</b>
--

1. Dans la liste des graphes patrons disponibles, l'indexeur sélectionne le graphe patron qui l'intéresse.
2. Il sélectionne et visionne le ou les fichiers vidéos composant son corpus.
3. Il spécifie la situation de discours qui contextualise, enchâsse la topique traitée, sélectionne les {thèmes/concepts} qui sont pertinents pour son propos et les désigne par des expressions linguistiques, des termes (entités nommées, noms communs, expressions verbales, ...).
4. Le cas échéant, il sélectionne dans la liste des paramètres plus spécialisés caractérisant la situation de discours (modalités épistémiques, appréciatives, ...) ceux qui l'intéressent et les désigne par des expressions linguistiques.
5. Puis, il sélectionne et « instancie » le thème/concept central du graphe patron sélectionné ({thèmes/concepts} en *couleur grisée*) en le désignant par une expression linguistique, un terme.
6. En visionnant une vidéo, il « marque » les moments temporels pertinents et les réfère aux {thèmes/concepts} correspondant dans le graphe patron.
7. Chaque moment pertinent correspondant à un thème/concept dans le graphe patron constitue donc un référent du graphe patron instancié.
8. L'indexeur désigne, le cas échéant, les {thèmes/concepts} du graphe autres que celui central, par des expressions linguistiques, par des termes.
9. A la fin de son indexation, l'indexeur voit affiché les résultats de son travail sur lesquels il pourra éventuellement intervenir (coquilles, ...).

**Note :** un {thème/concept} peut posséder comme référent plusieurs termes, plusieurs expressions linguistiques mais s'il s'agit du thème/concept central en *couleur grisée*, il faut instancier un deuxième graphe patron (par exemple, un chercheur peut parler de la langue française et puis de la langue allemande ...).

#### **Deuxième option : Un parcours de description**

1. Dans la liste des parcours des graphes patrons disponibles, l'indexeur sélectionne le parcours qui l'intéresse.
2. Parmi les graphes patrons composant le parcours sélectionné, il coche/décoche les graphes patrons qu'il souhaite utiliser (il pourra revenir sur son choix !) et, éventuellement, modifie l'ordre d'apparition des graphes sélectionnés.
3. Ensuite, le premier graphe patron sélectionné s'affiche et l'indexeur procédera comme dans la première option (cf. ci-dessus) et, cela, pour chacun des graphes sélectionnés. Il pourra arrêter son travail de description quand il veut.
4. A la fin de son travail, l'indexeur verra afficher les résultats de celui-ci pour pouvoir y intervenir, le cas échéant.

#### **Troisième option : création de graphes patrons « ex nihilo » ou en s'appuyant sur des graphes patrons déjà existants ....**

Il s'agit ici du « mode expert », i.e. d'une **nouvelle création** d'un ou de plusieurs graphes patrons organisés en un parcours par un spécialiste de spécification de configurations définissant le format d'une ou d'un ensemble de descriptions à proprement parler d'un corpus audiovisuels ...

1. L'auteur-interprète spécialiste choisit entre la création soit d'un nouveau graphe patron soit d'un parcours de description composé par un ensemble de graphes patrons. Il explicite en langage naturel de quoi il s'agit (exemple : graphe patron pour décrire des situations référant à l'usage de la langue dans les pratiques quotidiennes de la communication ; parcours de graphes patrons pour décrire des situations référant aux rapports entre langue et culture, ...)
2. A l'aide de l'ontologie (du domaine, rhétorique, ....) il spécifie le ou les graphes patrons et les enregistre.
3. Il continue ensuite comme dans l'option 2 ci-dessus.