

Ce formulaire (libellé Acronyme projet-nomcandidat) doit obligatoirement être envoyé  
sous **format pdf**  
à l'adresse électronique : [msh-app2020-contact@univ-lorraine.fr](mailto:msh-app2020-contact@univ-lorraine.fr).

**Date limite de candidature : le 17 janvier 2020 12 heures (Heure Paris)**

**A. TYPE DE DISPOSITIF DEMANDE**

MSHL « Projets 2020 »   
MSHL « Initiatives »

**B. RENSEIGNEMENTS D'ORDRE GENERAL**

Titre du projet (maximum : 150 caractères) :  
**Adjectifs, Verbes et Noms en -ance : héritage transcatégoriel des propriétés sémantiques**

Acronyme ou titre court (maximum : 12 caractères) : **ADVENANCE**

Nom et prénom du coordonnateur.e. du projet : Marie Laurence Knittel  
Statut et établissement : MCF Université de Lorraine  
Laboratoire d'appartenance (code unité et intitulé) : ATILF  
Téléphone : 06 73 72 36 59 Courriel : [marie-laurence.knittel@univ-lorraine.fr](mailto:marie-laurence.knittel@univ-lorraine.fr)

Indexation du projet (trois à cinq mots-clés) : Sciences du langage, sémantique, corpus,  
traitement de données, ressource lexicale trilingue

Proposition qui témoigne d'une intention d'incubation en vue d'un projet de plus grande  
envergure (émergence) impliquant un.e ou plusieurs jeunes chercheur.e.s

Doctorant  Master  Stage obligatoire  Stage non obligatoire

Proposition qui mobilise un partenaire extérieur public   
Proposition qui mobilise un partenaire extérieur privé   
Proposition qui aboutira à la mise en œuvre d'une thèse   
Proposition qui intègre un co-financement extérieur  (en attente)  
Si oui de quelle nature :  masse salariale,  fonctionnement  
Proposition à dimension internationale   
Proposition à forte dimension d'actualité politique et sociale   
Proposition liée aux défis sociétaux LUE

**Tableau 1 : Partenaires du projet** (les partenaires sont des laboratoires de recherche)

	Laboratoire ou autre groupe Nom et n° d'unité	Nom et prénom du correspondant du projet pour le laboratoire	Etablissement ou institution de rattachement	Nombre de chercheurs engagés par partenaire
Laboratoire Coordonnateur	ATILF (UMR 7118)	Marie Laurence Knittel	Université de Lorraine	4 - 3 Sciences du Langage - 1 Sciences du Langage et TAL
Partenaire 2	STL (UMR 8163)	Rafael Marín	Université de Lille	2 - 1 Sciences du Langage - 1 Sciences du Langage et TAL
Partenaire 3	MoDyCo (UMR 7114)	Florence Villoing	Université Paris-Nanterre	2 - 1 Sciences du Langage - 1 Sciences du Langage et TAL

**Tableau 2 : Budget du projet****Remarques :**

Un projet partenaire sur une thématique proche de celle d'AdVeNance a été déposé en novembre auprès de la MESH de Lille dans le cadre de l'appel à projet inter-MSH. Il est actuellement en phase de validation. S'il est accepté, des co-financements sont envisagés. En cas d'avis défavorable, il sera soumis à la MESH de Lille lors du prochain appel à projet.

- Le projet a été présenté à des membres du projet Impact OLKI (WP2 Langage, T2.1. Production de ressources linguistiques) pour déterminer si un éventuel soutien (logistique et/ou financier) est possible.

- Au fur et à mesure des appels, des cofinancements auprès de partenaires tels que le Pôle CLCS, les laboratoires, etc. seront demandés.

**Année 1 : Analyse des données du français**

Financements : Demandés (D) / Acquis (A)	Masse salariale (vacations)	Fonctionnement (déplacements, journée des travail stage)	Investissement	Total
ATILF : appel à projets internes 2020 (D) (déplacements)	1500	500		
UL : appel à projets du pôle CLCS 2020 (D) (journée de travail)		1000		
LUE-OLKI (D) (stage)		500		
MESH (Lille) (projet inter-MSH déposé à la MESH) (D) (déplacements)	500	500		
<b>Financement demandé à la MSH (déplacements)</b>	1000	500		
<b>Total</b>	<b>2500</b>	<b>3000</b>		<b>5500</b>

## Année 2 : Transposition à l'italien et à l'espagnol

Financements : Demandés (D) / Acquis (A)	Masse salariale (vacations )	Fonctionnement (déplacements, journée des travail, stage)	Investissement	Total
ATILF : appel à projets internes 2021 (journée de travail)	1500	1000		
UL : appel à projets du pôle CLCS 2021 (déplacements)	1500	1000		
<b>Financement demandé à la MSH (stage + équipement : Macbook Pro 13 pouces ssd 256 Go + mémoire 16 Go)</b>		500	1989	
<b>Total</b>	<b>3000</b>	<b>2500</b>	<b>1989</b>	<b>7489</b>

### C. Résumé de la proposition (de 1 à 5 pages maximum)

#### 1. RESUME DU PROJET

Le projet AdVeNance se fixe pour objectif d'étudier l'héritage des propriétés sémantiques à travers les catégories lexico-grammaticales, par l'observation de l'appariement des formes terminées par -ant et -ance (adjectifs, verbes et noms). Il sera fondé dans un premier temps sur l'analyse des propriétés morphosyntaxiques et sémantiques des unités lexicales du français, puis complété par l'étude des données de l'espagnol, et de l'italien. Ces données seront repérées dans des ressources telles que corpus (notamment FrCoW) et dictionnaires. Les objectifs finaux du projet sont doubles : progression des connaissances fondamentales et création d'un lexique trilingue annoté, qui pourra être diffusé en libre accès sur une plateforme (par exemple : ORTOLANG, CENHTOR).

#### 2. EXPOSE SCIENTIFIQUE DU PROJET

Notre projet porte sur la transmission des propriétés sémantiques entre trois catégories lexico-grammaticales majeures : verbes, noms et adjectifs.

##### 2.1. Propriétés typiques des catégories Nom, Verbe, Adjectif

De nombreux auteurs considèrent que chaque catégorie majeure dénote un certain type de concept (les verbes décriraient prototypiquement des événements, les noms des entités, et les adjectifs des propriétés).

En outre, chaque catégorie est caractérisée par une propriété sémantique typique :

- pour les verbes : l'aspect (caractère évolutif ou non de l'événance, l'existence d'une fin naturelle (télicité), extension temporelle) ;
- pour les noms : l'opposition massif / comptable ;
- pour les adjectifs : la scalarité, capacité à être situé et ordonné sur une échelle, et le caractère ouvert ou fermé de l'échelle.

##### 2.2. L'héritage transcatégoriel : état de la question

Certains auteurs ont suggéré des corrélations entre les propriétés d'items de catégories différentes mais morphologiquement apparentés. Ainsi, le caractère systématiquement massif des noms désadjectivaux, cf. *de la combativité* vs *\*une combativité*, a été mis en relation avec leur dénotation de propriété héritée des adjectifs (ici *combatif*) correspondants.

L'héritage des propriétés sémantiques des verbes par les noms correspondants qui a été étudié extensive et pour des langues diverses (Grimshaw 1991, Borer 2005<sup>1</sup>). Les observations convergent sur la conclusion que la télicité des verbes est corrélée à la comptabilité des noms.

### 2.3. Le projet AdVeNance : spécificités

Les résultats des travaux ci-dessus nous amènent à poursuivre nos recherches dans plusieurs directions.

**(i) La prise en compte de trois catégories dans les relations d'héritage** Comme mentionné ci-dessus, le caractère massif de certains noms déverbaux et des noms désadjectivaux met en évidence des relations d'héritage de V à N et de A à N. Puisque l'héritage entre deux catégories est bien établi, il nous semble intéressant d'examiner les relations d'héritage entre trois catégories. Il sera particulièrement utile de se pencher sur les relations des verbes et des noms avec les adjectifs, moins étudiés jusqu'à présent. Pour cela, nous examinerons les formes en *-ant / ance* (Dal & Namer 2010<sup>2</sup>, Knittel 2016<sup>3</sup>) qui, outre leur proximité formelle, peuvent se distribuer sur les trois catégories N, V, A (cf. *ignor(ant)<sub>V</sub>*, *ignorant<sub>Adj</sub>*, *ignorance<sub>N</sub>*).

L'examen de tels triplets sera centré sur trois problématiques.

(i) L'un des résultats du projet ANR Nomage (Balvet et al. 2012<sup>4</sup>) est l'observation que la plupart des noms en *-ance* sont construits sur des verbes statifs. Cette hypothèse demande à être vérifiée.

(ii) L'hypothèse que les adjectifs en *-ant* sont construits sur des verbes atéliques, qui établit un lien potentiel entre une propriété verbale et les adjectifs, sera mise à l'épreuve des données.

(iii) La relation entre scalarité et télicité. Notre hypothèse est que les adjectifs scalaires qui s'alignent sur une échelle ouverte entretiennent des relations plus étroites avec les verbes atéliques et les noms massifs, et vice-versa.

Si les hypothèses ci-dessus sont valides, nous attendons non seulement des appariements entre propriétés, mais aussi des régularités dans les appariements entre catégories, selon le modèle ci-dessous :

Verbes	Noms	Adjectifs
téliques	comptables	échelle fermée
atéliques	massifs	échelle ouverte

L'examen de l'appariement entre verbes, noms et adjectifs nous permettra également de mettre en évidence l'inexistence de certaines formes, telles que *\*bouillance* (vs. *bouillant<sub>V</sub>* ; *bouillant<sub>Adj</sub>*) ou *\*assurant* (vs. *assurant<sub>V</sub>* ; *assurance<sub>N</sub>*). On sera alors amené à interroger la cause de ces lacunes.

### (ii) L'adoption d'une approche comparative entre trois langues romanes

Le projet AdVeNance se développera dans un second temps vers l'étude de l'italien et l'espagnol, en se fondant sur les outils d'analyse développés et les connaissances acquises sur le français (cf. (iii) ci-dessus), et en mettant à profit les compétences linguistiques de plusieurs

<sup>1</sup> Borer, H. 2005. *Structuring sense*. Oxford: Oxford University Press. Grimshaw, J. 1990. *Argument structure*. Cambridge MA: MIT Press.

<sup>2</sup> Dal, G., Namer, F. 2010. Les noms en *ance/ ence* du français : quel(s) patron(s) constructionnel(s) ? In Durand, J., Habert B. & Laks, B. (eds) *Actes du 2eme Congrès Mondial de Linguistique Française*, 893-907.

<sup>3</sup> Knittel, M.L. 2016b. Les noms en *-ance* : un panorama. Actes du 5<sup>o</sup> CMLF (Tours 4-6 juillet 2016). F. Neveu, G. Bergounioux, M-H Côté, J.-M. Fournier, L. Hriba et S. Prevost (eds). ILF.

<sup>4</sup> Balvet A., Barque L., Condette M.-H., Haas P., Huyghe R., Marin R., Merlo A. (2012). La ressource Nomage, Confronter les attentes théoriques aux observations du comportement linguistique des nominalisations en corpus. *TAL*, 52-3/2012. 1-24

membres du projet (voir 3 ci-dessous). En effet, les trois catégories formellement proches décrites ci-dessus pour le français (cf. (5)) présentent des correspondances en espagnol (Fábregas & Marín 2017<sup>5</sup>) et en italien, langues romanes proches du français, cf. Esp.: *abundar<sub>V</sub>* / *abundancia<sub>N</sub>* / *abundante<sub>Adj</sub>*; It.: *abbondare<sub>V</sub>* / *abbondanza<sub>N</sub>* / *abbondante<sub>Adj</sub>*. Après validation, les hypothèses avancées pour le français (cf. 2.3.(ii)) seront vérifiées sur ces deux langues. Au vu des connaissances actuelles des phénomènes sémantiques et de leur relation avec les données morphosyntaxiques à travers les langues, on s'attend à trouver des parallélismes réguliers entre les trois langues étudiées. Cependant, l'existence de lacunes spécifiques à une ou deux des trois langues observée (Fr. \**délinquer<sub>V</sub>* / *délinquance<sub>N</sub>* / *délinquant<sub>Adj</sub>*; Esp. : *delinquir<sub>V</sub>* / *delincuencia<sub>N</sub>* / *délinquente<sub>Adj</sub>*; It. *delinquere<sub>V</sub>* / *delinquenza<sub>N</sub>* / *delinquente<sub>Adj</sub>*) nous amènera à nous interroger sur les différences de fonctionnement sémantiques et/ou syntaxiques qu'elles révèlent.

**(iii) Création d'un lexique annoté** Le projet AdVeNance comportera un volet orienté vers le traitement automatique des langues (TAL), puisqu'il débouchera sur la création d'un lexique annoté avec un alignement des 3 catégories étudiées. Actuellement, une phase de tests préparatoire portant sur 100 triplets extraits de Lexique.org (<http://www.lexique.org/>) est en cours. Celle-ci nous permettra d'élaborer un protocole de tests en affinant la grille d'analyse élaborée au cours du projet LexVAN (<https://www.meshs.fr/page/lexvan>), qui sera à son tour transposée à l'italien et à l'espagnol, si nécessaire après adaptation.

Dans un second temps, une ressource comportant 200 triplets français sera élaborée. Celle-ci permettra le repérage automatique des formes en *ance* construites sur des Verbes ou sur des Adjectifs, et fournira une aide à la catégorisation automatique des formes ambiguës (*énervant<sub>AdjMascSg</sub>* VS *énervant<sub>VPartPres</sub>*). La recherche des formes sur un grand corpus, comme le FrCoW (<https://corporafromtheweb.org/frcow16/#more-490>), nous permettra d'évaluer la productivité de la dérivation nominale en *ance* et adjectivale en *ant*. Des recherches similaires seront ensuite conduites sur les autres langues étudiées dans le projet, à partir de itWaC et ESCoW.

## 2.4. Apports scientifiques du projet

Du point de vue fondamental, la spécificité de ce projet est de se situer à l'interface entre trois disciplines des sciences du langage : la sémantique, la morphologie (puisque les catégories étudiées sont apparentées par la forme) et la syntaxe, puisque c'est le comportement syntaxique qui révèle les propriétés sémantiques et catégorielles des items étudiés. Parallèlement, il mettra en relation les manifestations morphosyntaxiques des propriétés sémantiques de trois langues proches. Il se distingue par l'étude parallèle de 3 catégories, au lieu de 2 comme habituellement, et par la part accordée à l'étude des adjectifs. En outre, il se situe dans une optique comparative entre trois langues romanes, puisque les résultats obtenus sur le français serviront de base à l'analyse de l'italien et de l'espagnol.

D'un point de vue méthodologique, les ressources produites ainsi que les batteries de tests utilisées en tant que grilles d'analyse seront mise à la disposition de la communauté, notamment pour l'élaboration ou l'enrichissement de ressources lexicales, comme par exemple la ressource jeuxdemots. (<http://www.jeuxdemots.org/jdm-accueil.php>). Les résultats obtenus seront complémentaires à ceux de l'ANR Demonext, avec lesquels ils pourront être mis en relation et confrontés.

## 3. METHODOLOGIE DU PROJET

---

<sup>5</sup> Fábregas, A., Marín, R. 2017. Lexical categories and aspectual primitives: the case of Spanish '-ncia'. In Bloch-Trojnar, M. & Malicka-Kleparska, A. (eds), *Aspect and Valency in Nominals*. Berlin: Mouton de Gruyter.

Le projet, d'une durée de 2 ans, sera organisé en deux phases successives d'une année chacune (phases 1 et 2). La phase 1 visera l'analyse des données du français (Tâches 1 à 6<sup>6</sup>), la phase 2 à la transposition de la méthodologie aux données de l'italien et de l'espagnol, à l'analyse puis à l'étude des résultats (Tâches 7 à 12).

### **Phase 1 : Etude du français**

#### **- T1 : Elaboration d'une liste L1 de 200 triplets V / Adjant / Nance**

- Repérage dans le corpus frCoW des adjectifs en –ant et des noms en –ance - Vérification de l'existence d'un verbe morphologiquement
- Nettoyage de L1 (suppression des triplets comportant des formes non apparentées sémantiquement)

*Ces tâches sont actuellement en cours de réalisation.*

#### **- T2: Elaboration d'un guide d'annotation (grilles d'analyses)** recensant les tests proposés dans la littérature et le projet LexVAN pour l'évaluation de :

- la scalarité des adjectifs
- la télélicité des verbes
- la massivité des noms

- rédaction d'un guide d'annotation (*tâche en cours de réalisation*).

**- T3: Annotation des propriétés sémantiques des items de L1** au moyen des grilles d'analyse. Cette tâche sera partiellement effectuée sous forme de vacations, sous la supervision des membres du projet.

**- T4: Mise en relation des propriétés des trois catégories impliquées dans les triplets**

**- T5: Vérification des hypothèses formulées** (cf. 2.3.(i) ci-dessus)

Si nécessaire, ajustement des hypothèses ou élaboration de nouvelles propositions

**- T6 : Elaboration d'un lexique annoté**

Cette tâche sera réalisée en vacations, supervisées par les membres du projet.

### **Phase 2 : Etude de l'italien et de l'espagnol**

#### **T7: Création de listes de 100 triplets du français à l'italien (L2) et à l'espagnol (L3)**

Pour chaque langue :

- évaluation de l'existence de triplets identiques
- repérage des lacunes
- complémentation de L2 et L3 jusqu'à 100 par des triplets non-attestés en français

**T8:** Transposition du guide d'annotation élaboré pour le français à l'italien et à l'espagnol :

- vérification de la validité des tests dans les trois langues
- si besoin, complémentation de la grille par des tests supplémentaires propres à chaque langue

**T9:** Annotation des propriétés sémantiques des items des listes au moyen des grilles d'analyse propres à chaque langue (vacations supervisées par les membres du projets locuteurs natifs des langues concernées)

**T10:** Mise en relation des propriétés des trois catégories impliquées dans les triplets de chaque langue ; élaboration du lexique trilingue aligné (vacations)

**T11:** Vérification sur l'italien et l'espagnol des hypothèses formulées sur le français. Si nécessaire, ajustement ou élaboration de nouvelles hypothèses pour chaque langue.

**T12:** Etude des lacunes :

- étude des lacunes propres à une langue et des lacunes communes à plusieurs langues ;
- mise en relation et analyse de leurs sources et de leurs implications intra- et inter-langues

---

<sup>6</sup> Les Tâches sont désignées par T dans ce qui suit, les listes par L.

<b>Année 1</b>	<b>Trimestre 1</b>	<b>Trimestre 2</b>	<b>Trimestre 3</b>	<b>Trimestre 4</b>
<b>Tâches</b>	T1-T2	T3-T4	T4-T5	T6
<b>Membres impliqués</b>	Knittel Marín Missud Villoing	Knittel Marín vacataire	Knittel Villoing Namer Lignon	Namer Balvet vacataire
<b>Année 2</b>	<b>Trimestre 1</b>	<b>Trimestre 2</b>	<b>Trimestre 3</b>	<b>Trimestre 4</b>
<b>Tâches</b>	T7-T8	T9-T10	T11	T12
<b>Membres impliqués</b>	Knittel Marín Namer Ruvoletto	Marín Namer Ruvoletto Balvet vacataire	Knittel Marín Villoing Namer Lignon Ruvoletto	Knittel Marín Villoing Namer

#### **4-5. RESULTATS ATTENDUS ET RETOMBEES ECONOMIQUES ET SOCIETALES POSSIBLES**

##### **4.1. Outils**

- Création de trois listes de (1 liste L1 de 200 triplets V/N/Adj alignés pour le français et 2 listes de 100 triplets V/N/Adj alignés pour l'italien (L2) et l'espagnol (L3))
- Elaboration de grilles d'annotations des propriétés verbales, adjectivales et nominales étudiées pour les trois langues concernées

##### **4.2. Connaissances fondamentales**

- Enrichissement des connaissances des phénomènes d'interface entre syntaxe, sémantique et morphologie
- Progression des connaissances en linguistique comparée des langues romanes.

##### **4.3. Diffusion des outils et des résultats**

Les listes et les grilles d'annotations seront mises à la disposition de la communauté des spécialistes (chercheurs, enseignants, traducteurs, linguistes, TAListes, etc), et, au-delà, du grand public (plateformes ORTOLANG et/ou CENHTOR).

Les observations faites au cours du projet seront diffusées lors de colloques et feront l'objet de publications scientifiques.

#### **6. DIMENSION INTERNATIONALE**

Au cours du projet, des collaborations avec des chercheurs et des universités italiennes et espagnoles seront envisagées. On pense notamment à A. Fábregas (UIT Tromsø), L. McNally (Pompeu Fabra), C. Melloni (Vérone).

A l'issue du projet, nous envisageons l'extension de la recherche à l'étude d'autres langues romanes, notamment le catalan, le portugais et le roumain. Des contacts seront pris avec de futurs partenaires.

#### **7. AUTRES ELEMENTS QUE LE PORTEUR SOUHAITE PORTER A LA CONNAISSANCE DES EVALUATEURS**


#### D. LES CHERCHEUR.E.S ENGAGÉ.E.S

Nom et prénom	Statut*	Discipline ou section CNU ou section CNRS	Laboratoire ou groupe d'appartenance	Etablissement	Courriel
Marie Laurence KNITTEL (sémantique, syntaxe)	MCF HDR	7°	ATILF	Université de Lorraine	marie-laurence.knittel@univ-lorraine.fr
Fiammetta NAMER (morphologie, sémantique, TAL, italien)	PR	7°	ATILF	Université de Lorraine	fiammetta.namer@univ-lorraine.fr
Stéphanie LIGNON (morphologie, phonologie)	MCF	7°	ATILF	Université de Lorraine	stephanie.lignon@univ-lorraine.fr
Samantha RUVOLETTO (italien, phonologie)	MCF (jeune chercheur)	7°	ATILF	Université de Lorraine	samantha.ruvoletto@univ-lorraine.fr
Rafael MARIN (sémantique, TAL, espagnol)	CR	7°	STL	Université de Lille	rafael.marin@univ-lille.fr
Antonio BALVET (linguistique, TAL)	MCF	7°	STL	Université de Lille	antonio.balvet@univ-lille.fr
Florence VILLOING (morphologie, sémantique)	Pr	7°	MoDyCo	Université de Nanterre	florence.villoing@parisnanterre.fr
Alice Missud (morphologie, TAL)	Doctorante (contrat doctoral)	7°	MoDyCo	Université de Nanterre	missud.a@parisnanterre.fr
Nombre total de chercheur.e.s	8				

\* Pr, MCF, CR, DR, IR, IE, Tech, doctorant.e, post-doc (dans ce dernier cas, préciser si salarié, boursier ou non au sein d'un laboratoire ou contractuel sur l'opération).

\*\*\*\*\*



<p>Nom, prénom du porteur.e de projet Marie Laurence KNITTEL Date : 16/01/2020</p>	<p>Nom, prénom du responsable du laboratoire Alex BOULTON Date : 16/01/20</p>
<p>Signature : M.L. Knittel</p> 	<p>Signature :</p> 