

# *Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle*

---

Dans ce dernier chapitre, nous proposons un regard croisé sur un certain nombre des facteurs examinés dans cette thèse. Nous avons analysé l'influence de différents facteurs sur le choix d'une expression référentielle. Ces facteurs se situent à différents niveaux : (lexico-)syntaxique, référentiel, discursif, informationnel et interactionnel. L'objectif de ce chapitre est d'évaluer le poids respectif de certains de ces facteurs. Pour ce faire, nous avons eu recours à des modèles statistiques probabilistes, qui mesurent l'influence de différents facteurs ainsi que l'interaction entre ces facteurs. Nous présenterons d'abord ces modèles et les facteurs retenus pour l'analyse en section 1, avant de présenter et discuter les résultats dans les sections 2, pour le français, et 3, pour l'allemand. La section 4 propose une synthèse des résultats obtenus.

## **1. Évaluation statistique des facteurs dans un modèle probabiliste**

### **1.1 Circonscription de l'approche choisie**

Comme noté ci-dessus, nous sommes intéressée dans une mesure du poids respectif des différents facteurs relevés dans cette thèse, dont nous avons vu qu'ils jouent un rôle dans le choix des expressions référentielles. Nous avons constaté et montré à l'aide de statistiques que, par exemple, la

dislocation en français est parmi les expressions préférées dans l'expression du topic en français, et qu'elle joue également souvent un rôle dans des mouvements interactionnels qui effectuent un retour ou un décrochage par rapport au discours précédent. Mais quelle est l'interaction entre ces facteurs ? Est-ce que l'un prime sur l'autre ? Est-ce que les deux se combinent ? C'est à ce type de question que nous tenterons de répondre ici. Puisqu'il ne nous sera pas possible de traiter l'ensemble des expressions linguistiques, nous avons opéré des choix. Nous nous sommes focalisés dans cette thèse davantage sur la référence aux entités, et nous allons adopter également ce focus dans le présent chapitre. En ce qui concerne les noms, il aurait bien sûr été intéressant de comparer les emplois référentiels et non référentiels, mais cela n'aurait pas été pertinent pour les autres expressions considérées ici. Nous devons, pour des raisons de place, laisser de côté ici également la référence à la personne. Parmi les expressions en référence aux entités, nous avons choisi de porter notre intérêt sur les dislocations en français (en allemand, les occurrences sont trop peu fréquentes bien sûr pour une analyse statistique). Nous allons nous intéresser également aux noms, dont nous avons vu dans le chapitre précédent qu'ils peuvent aussi remplir les mêmes fonctions de mouvement interactionnel de RETOUR ou DECROCHAGE et dont nous supposons qu'ils pourraient jouer un rôle plus important en allemand, à la place des dislocations. Nous avons choisi de regarder également les D-Pro de l'allemand, catégorie qui apparaît pour le moment comme intermédiaire ou passe-partout : employée aussi bien pour la continuité que la discontinuité dans la chaîne topicale, nous allons examiner ici son emploi comme fonction de différents facteurs à la fois. Pour le français, nous allons examiner les pronoms démonstratifs et les pronoms personnels clitiques, catégories en construction/acquisition pour les jeunes enfants, et que nous souhaitons opposer aux D-Pro en allemand, qui lui ne fait que très peu usage de pronoms personnels dans la référence aux entités. Enfin, nous allons nous intéresser de plus près aux référents non-verbalisés, dont nous avons vu dans le chapitre précédent que leur emploi dans des mouvements interactionnels qui effectuent un retour sur ce qui précède mérite quelques affinements. Ce sont donc les expressions les plus importantes pour l'encodage du topic auxquels nous allons nous intéresser ici.

Pour évaluer l'influence des différents facteurs étudiés sur le choix des expressions référentielles, nous ne pouvons pas nous tourner vers des modèles de régression paramétriques, puisque nos données ne suivent pas une distribution normale. Nous nous sommes alors tournée vers des modèles mixtes (*Generalized Linear Mixed-Effects Models*), que nous présenterons dans la section suivante.

## 1.2 Introduction du modèle et des facteurs

L'un des intérêts principaux des modèles mixtes est, comme le suggère leur nom, de pouvoir combiner des effets dits « fixes » et des effets aléatoires. Nous pouvons ainsi tester les effets pour lesquels nous pouvons établir une hypothèse forte, i.e. dont nous supposons qu'ils ont une influence

## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

---

sur le choix de l'expression référentielle,<sup>253</sup> tout en prenant en compte la variabilité individuelle. En outre, ce test statistique ne requiert pas un nombre identique d'occurrences entre les différentes observations. Nous nous sommes servi de la fonction *glmer* du programme de statistiques *R* (R Core Team, 2018), implémentée dans le package *lme4* (D. Bates, Maechler, Bolker, & Walker, 2015). Nous sommes partie, pour chaque expression référentielle étudiée dans ce chapitre, du modèle le plus complet théoriquement plausible, incluant des facteurs aléatoires et fixes, et avons procédé à une sélection du modèle le plus adapté par élimination successive des facteurs non significatifs (Zuur, Ieno, Walker, Saveliev, & Smith, 2009). Cette procédure sera précisée plus loin, après l'introduction des facteurs considérés.

### 1.2.1 Facteurs et variables retenus pour le calcul des effets aléatoires

Nous avons testé comme effet aléatoire la session d'enregistrement, l'enfant cible ainsi que le type de jeu (puzzle, dînette, maison de poupées..). L'influence du type de jeu a d'abord été testée comme effet fixe, mais n'apportait pas de pouvoir explicatif significatif aux modèles, les situations étant finalement trop hétérogènes. En effet, nous avons vu dans les chapitres précédents que le choix des expressions référentielles semble entre autres dépendre de certains genres et mouvements discursifs (nommer et localiser notamment pour les dislocations), et nos exemples ont montré que nous pouvions trouver des dénominations et demandes de dénominations dans les puzzles, mais aussi par exemple dans le jeu avec la maison de poupées. La façon dont nos participants ont investi les différentes situations de jeu semble alors trop hétérogène pour servir de prédicteur pour l'emploi d'une expression plutôt que d'une autre. Nous avons alors tenté de poser le jeu comme variable aléatoire, mais la variance observée pour les jeux dans ces modèles était Zero, lorsque nous prenions en compte aussi les autres effets aléatoires que sont l'enfant étudié et la session d'enregistrement. Pour une partie des enfants, nous disposons d'une seule session, et inclure à la fois la session et les enfants comme facteur aléatoire a également eu comme résultat de renvoyer une variance proche de Zero pour un des facteurs (Enfants), et d'empêcher le modèle à estimer correctement les effets. En effet, ces trois facteurs aléatoires se redoublent ou mesurent peu ou prou la même chose. Il fallait alors n'en retenir qu'un seul. La prise en compte de la session d'enregistrement donnait globalement les modèles les plus forts : c'est alors ce facteur que nous avons retenu comme variable aléatoire dans tous les modèles. Cela nous permet de tenir compte et de contrôler la variation individuelle, et de montrer quels effets fixes expliquent au mieux l'emploi d'une expression, au-delà des différences aléatoires. Les effets fixes considérés sont décrits dans la section suivante.

---

<sup>253</sup> Ces hypothèses reposent sur des considérations théoriques, mais sont aussi déjà étayées par les tests statistiques de base conduites dans les chapitres précédents. Pour chacun des facteurs qui apparaissent ainsi comme pertinents, nous étudierons alors ici l'impact relatif, en interaction avec les autres facteurs.

### 1.2.2 Facteurs et variables retenus pour le calcul des effets fixes

Nous avons considéré pour chaque modèle différents facteurs formels, informationnels et interactionnels. Certains des facteurs analysés dans les chapitres précédents ont été exclus d'embrée : Nous n'avons ainsi pas retenu la position de l'expression par rapport au verbe. D'une part, ce facteur est lié à d'autres contraintes formelles, et, comme nous l'avons expliqué dans le CHAPITRE VII-4.1, un véritable choix entre position pré- et postverbale est seulement donné dans des énoncés déclaratifs avec verbe en position V2 en allemand. D'autre part, il nous semble que la position dans l'énoncé relève davantage d'une variable dépendante que d'un facteur : il s'agit d'une coréalisation, conjointe au choix de la catégorie grammaticale, en fonction d'autres facteurs que nous étudions ici. Ces raisons valent aussi pour la réalisation prosodique, que nous n'avons pas incluse non plus dans les facteurs à considérer, mais plutôt étudié en tant que variable dépendante. Au vu des résultats des chapitres précédents, il nous a semblé plus intéressant de nous pencher sur la réalisation prosodique des expressions que sur la position par rapport au verbe, pour laquelle nous avons vu qu'elle dépend de différents facteurs à la fois, et qui donne des modèles peu parlants.

Parmi les facteurs formels, nous avons considéré la fonction syntaxique. Nous avons considéré également comme facteur formel, pour le français, le cadre ou schème lexico-syntaxique étudié pour les dislocations (voir les CHAPITRE VIII-1.5 et CHAPITRE X). Nous nous sommes basée pour cela sur les schèmes identifiés comme fréquemment associés aux dislocations dans notre corpus, et nous les avons relevés aussi pour les autres expressions référentielles.

Pour les facteurs fonctionnels, nous avons retenu plusieurs facteurs liés à la dimension pragmato-discursive. Nous avons considéré le statut de topic, la place dans la chaîne topicale, et les différents statuts attentionnels ou discursifs d'un référent. Pour le niveau interactionnel, nous avons regroupé différents phénomènes observés dans le chapitre précédent en une seule catégorie, prenant en compte tous les phénomènes de décrochage ou de RETOUR sur un TdP ou énoncé précédent. Nous avons également considéré à chaque fois le groupe MLU des enfants.

Le principe de comparaison dans les modèles employés est un système binaire : toutes les catégories ont été converties en système binaire, i.e. une expression correspond au topic ou non, à la discontinuité topicale ou non, etc., exprimé par 0 ou 1. Le Tableau XI-1 ci-dessous détaille l'ensemble des facteurs retenus :

Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

Tableau XI-1 – Facteurs et variables retenus pour l'analyse statistique

Facteurs	Variables	Description
<b>Syntaxique</b>		
Fonction syntaxique	SUJET_THEME	Expressions en fonction syntaxique sujet ou catégorisé comme Thème dans le sens logico-syntaxique de Zemb (SUJET_THEME), par opposition aux autres fonctions syntaxiques (AUTRE)
Schèmes lexico-syntactiques	SCHEMELFREQ	Expression référentielle produite dans un des schèmes identifiés comme fréquemment associé aux dislocations (FREQSCHEM), par opposition à celles qui ne sont pas produites dans un de ces cadres (AUTRE)
<b>Pragmatico-discursif</b>		
Statut topical	TOPIC	Expression correspondant au topic de l'énoncé (TOPIC), opposé aux expressions qui n'ont pas ce statut (AUTRE)
	TOPDISCONT	Cas de topic disruptif, ensemble des topics Nouveau et Réactivé (TOPDISCONT), par opposition à toutes les autres expressions, topic ou non (AUTRE)
	TOPCONTINU	Cas de maintien d'un topic (TOPCONTINU), opposé au reste des expressions référentielles, qu'ils soient topic ou non (AUTRE)
Statut attentionnel	DONNE	Le référent est donné dans le discours (DONNE), par opposition aux autres statuts attentionnels (AUTRE)
	ACTIV	Le référent est activé (découle d'un autre référent mentionné ou sous le focus de l'attention conjointe ; ACTIV), par opposition aux autres statuts attentionnels (AUTRE)
	NOUVEAU	Le référent est nouvellement introduit (NOUVEAU), par opposition aux autres statuts attentionnels (AUTRE)
	REINT	Le référent est réintroduit dans le discours après abandon momentané (REINT), par opposition aux autres statuts attentionnels (AUTRE)
<b>Interactionnel</b>		
	MVMT_RETUR	Le référent est concerné par un mouvement (inter)actionnel de retour ou décrochage (MVMT_RETUR) qui correspond à un ou plusieurs des caractéristiques étudiées dans le chapitre 11 : <ul style="list-style-type: none"> <li>- Qui crée un lien avec un énoncé ou un TdP précédent en reprenant un élément lexical comme nouveau point de départ</li> <li>- Qui constitue une confirmation, demande de confirmation ou de clarification d'un énoncé/tour précédent</li> <li>- Qui constitue une réitération d'une intention communicative</li> <li>- Qui crée une activité d'énumération (liste)</li> <li>- Qui crée un contraste ou un changement de point de vue</li> </ul> Cela est opposé aux cas où un référent constitue un ajout par rapport à ce qui précède ou fait l'objet d'une continuité simple (AUTRE)
<b>Développemental</b>	GroupeMLU	Groupe MLU 2 ou Groupe MLU 3

Notons tout de suite que la colinéarité peut poser un problème pour les modèles mixtes, et qu'il n'était pas toujours possible de considérer tous les facteurs à la fois. Il nous fallait à chaque fois trouver un compromis pour les facteurs à considérer. Il faut comprendre la colinéarité comme le fait de mesurer plusieurs fois la même chose. Imaginons une catégorie qui corresponde toujours au topic de l'énoncé. Inclure dans ce cas les facteurs du statut de topic, de la continuité et de la discontinuité topicale revient alors de mesurer plusieurs fois la même chose, et empêche le modèle d'estimer correctement les variances dans les données. Nous avons donc essayé à chaque fois de trouver le meilleur compromis pour un modèle théoriquement valable et statistiquement plausible. Concrètement, cela veut dire que nous avons considéré le facteur TOPDISCONT plutôt que TOPCONTINU dans nos modèles, puisque l'inclusion des deux à la fois posait des problèmes pour une estimation correcte. Dans un souci de cohérence et de comparabilité des différents modèles, nous avons alors choisi de garder les mesures de discontinuité (topicale (TOPDISCONT) et interactionnelle (MVMT\_RETUR)), dont nous avons vu qu'ils sont pertinents pour l'emploi des dislocations.

Les modèles que nous présenterons dans les sections suivantes sont le résultat d'une comparaison entre différents modèles possibles pour une catégorie d'expression donnée, pour trouver le modèle le mieux ajusté aux données. La procédure employée était de commencer avec le modèle le plus complexe théoriquement plausible (i.e., considérer tous les facteurs pour lesquels nous supposons une influence sur la variable dépendante, le choix de l'expression référentielle étudiée (plutôt que d'une autre expression). Les modèles sont évalués avec un index de concordance (valeur de C dans la première ligne de chaque table). Cette valeur indique la qualité d'ajustement (*goodness of fit*) du modèle en question.<sup>254</sup> Nous avons procédé à des méthodes d'élimination (et parfois d'ajout) de facteurs par étapes, en évaluant la significativité de chaque facteur et en testant la qualité d'ajustement après suppression des variables non significatives. Différents modèles ont été comparés entre eux à l'aide d'un test du rapport de vraisemblance (*likelihood-ratio test*, fonction `anova()` dans R<sup>255</sup>). Ce test nous permet d'évaluer si un modèle plus simple est statistiquement comparable au modèle plus complexe, et de choisir le modèle le mieux ajusté, i.e. qui a la meilleure probabilité de générer des distributions de données telles que nous les avons observées.

Il faut tout de même ajouter que nous nous intéressons ici avant tout à expliquer les distributions observées, et nous nous servons des statistiques en ce sens, plutôt que de prétendre pouvoir établir des modèles avec un véritable pouvoir prédictif. Pour cela, des données plus homogènes, telles qu'obtenues dans un protocole expérimental seraient probablement nécessaires. Cela n'invalide pas les conclusions que nous allons tirer des modèles retenus, mais justifie que nous avons parfois privilégié l'ajustement le plus plausible en termes de théorie à l'ajustement purement mathématique.

## 2. Résultats pour le corpus français

### 2.1 Les dislocations

#### 2.1.1 Les enfants

Pour les enfants, le Tableau XI-2 ci-dessous donne les résultats de l'analyse statistique pour les dislocations.

---

<sup>254</sup> Les valeurs de l'index de concordance peuvent aller de 0 à 0.5 :

C<0.5	mauvais modèle	C>0.8	modèle fort
C=0.5	modèle égal au hasard	C=1	modèle parfait, prédit toutes les occurrences
C>0.7	bon modèle		

<sup>255</sup> Fonction `anova()` dans R avec *a* minuscule, différent de l'Anova classique : ce n'est pas un test de régression paramétrique, mais bien un test de rapport de vraisemblance entre deux modèles `glmer`.

Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

Tableau XI-2 – Résultats du modèle linéaire mixte pour les dislocations (enfants, français)

Dislocations, C=0.9550577				
Effets aléatoires :	Variance	SD		
Session	0.3872	0.6222		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-7.8385	0.6508	-12.044	<.00001***
SUJET_THEME	1.2064	0.3702	3.259	.00112 **
SCHEMEEFREQ	4.3705	0.3531	12.377	<.00001***
TOPDISCONT	1.8162	0.2660	6.828	<.00001***
MVMT_RETUR	2.2948	0.3512	6.534	<.00001***
Groupe MLU (MLU3)	1.9195	0.4724	4.063	<.00001***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1304

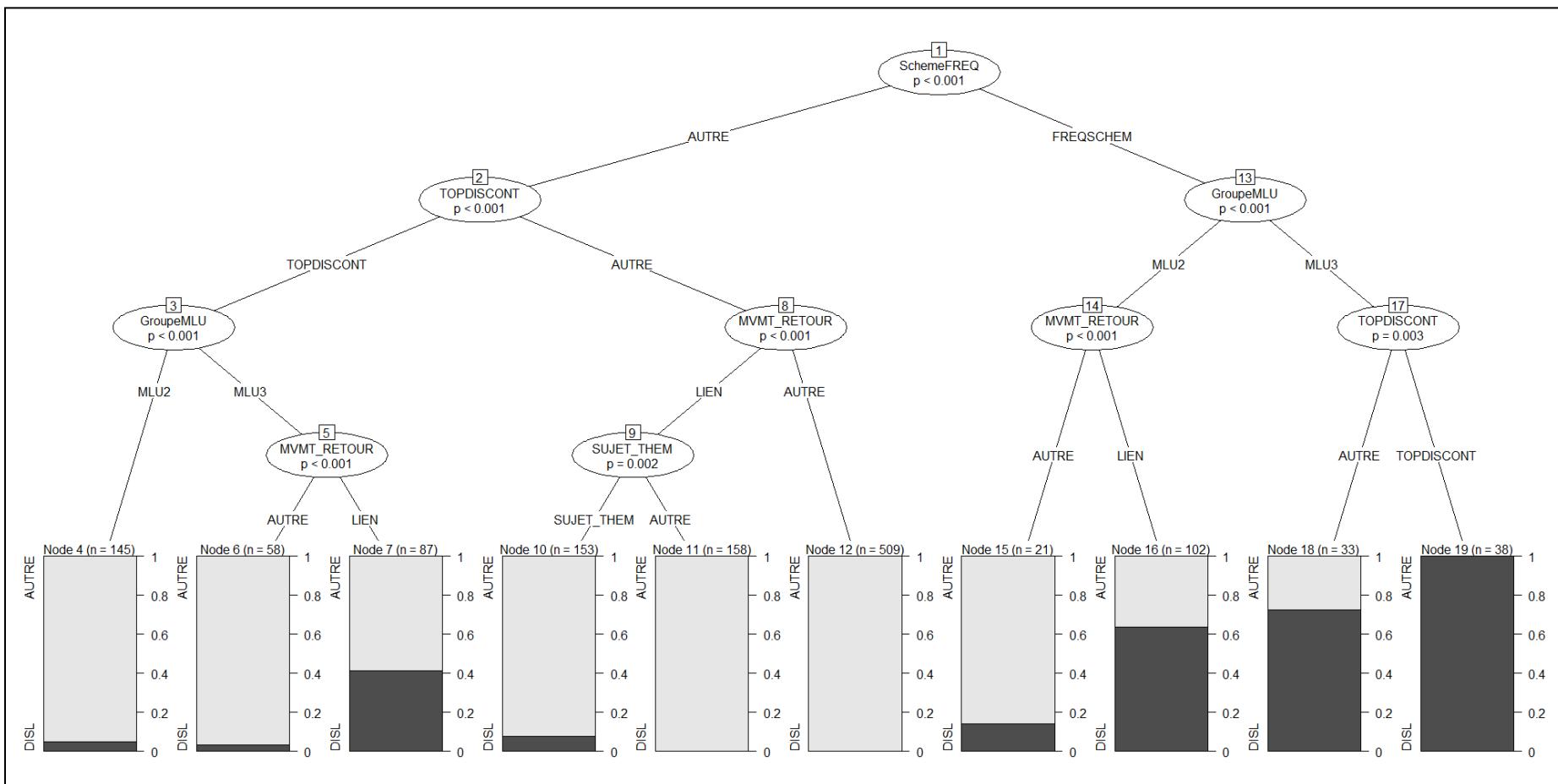
Il peut paraître surprenant de ne pas trouver le statut de topic parmi les effets fixes. Comme toutes les dislocations dans la référence aux entités correspondent au topic chez nos enfants (pour rappel, les cas indécidables ne sont pas inclus dans le modèle), inclure ce facteur donne une séparation parfaite (toutes les dislocations ont le statut de topic), n'a que très peu de pouvoir explicatif supplémentaire concernant la distribution des dislocations, et minimise le poids des autres facteurs (ou plutôt, cela empêche en fait le modèle de les estimer correctement). Compte tenu du fait que nous gardons tout de même une mesure liée au statut de topic avec TOPDISCONT, nous avons alors choisi un modèle sans le facteur TOPIC, ce qui permet d'évaluer le poids respectif des autres facteurs de manière plus stable. Ce n'est pas le modèle idéal, mais nous semble le meilleur compromis théorique et technique possible.

Comme nous nous y attendions, l'emploi des dislocations dépend de plusieurs facteurs à la fois. Ces facteurs sont tous significatifs, et dans une relation positive (signe positif des valeurs de z) avec la variable dépendante, la dislocation, et permettent d'expliquer 95% des occurrences des dislocations (valeur de C du modèle). L'emploi d'une dislocation est plus probable lorsque l'expression est en fonction sujet. En effet, nous avons vu dans les CHAPITRE VII-4.2 et CHAPITRE VIII-1.3 que la majeure partie des dislocations concerne le sujet. Nous avons constaté également qu'une grande partie des dislocations apparaît avec certains schèmes concrets, notamment des dénominations et localisations, et notre modèle confirme que la présence d'un tel schème favorise l'emploi d'une dislocation. Le contexte de topic discontinu (i.e., nouveau ou réactivé) favorise également l'emploi d'une dislocation ; comme le contexte d'un mouvement interactionnel qui fait un retour sur un TdP ou énoncé précédent. Enfin, nous observons un effet du MLU que nous n'avions pas observé dans les chapitres précédents (pas de corrélation ni de différence significative entre les groupes MLU pour la seule fréquence des dislocations). De fait, les pourcentages moyens entre les groupes MLU2 et MLU3 sont presque identiques (13% en MLU2, 14,46% en MLU3), mais le modèle suggère qu'un MLU plus élevé favorise l'emploi des dislocations (rappelons toutefois qu'un énoncé composé de seulement 2 termes peut en principe contenir une (proto-)dislocation, puisque les proto-dislocations, sans résomptif, sont inclus dans cette catégorie).

Aucun des différents statuts attentionnels n'était significativement lié à l'emploi d'une dislocation, ce qui n'est pas surprenant au vu des résultats du CHAPITRE VII, qui ont montré une distribution relativement homogène. Enlever ces facteurs du modèle n'a pas altéré significativement les variances observées, et nous avons donc retenu le modèle ci-dessus comme le modèle le plus simple expliquant au mieux les données.

Nous avons donc observé que ces différents facteurs favorisent conjointement l'emploi d'une dislocation. Quel est alors l'interaction de ces différents facteurs, ou quel est leur poids respectif ? Pour répondre à cette question, nous avons réalisé des arbres de partition (*conditional inference trees*), sur la base des facteurs retenus dans le modèle ci-dessus. Ces arbres ont été réalisés avec la fonction `cmtree()` dans R (Hothorn, Hornik, Strobl, & Zeileis, 2019; Hothorn, Hornik, & Zeileis, 2006). Ces arbres de partition rendent compte de la distribution, et, crucialement, de l'interaction des différents facteurs. L'arbre est construit de manière récursive, en partant du facteur le plus fortement associé à la variable de réponse. Un des avantages des arbres de partition est qu'ils permettent facilement d'interpréter et de communiquer des résultats statistiques sous forme visuelle. Chaque nœud de l'arbre représente un facteur qui influence significativement la distribution de la variable dépendante. Le nœud le plus haut correspond ainsi au facteur le plus fortement associé à la variable associée. Les différents nœuds sont numérotés de gauche à droite, en partant du haut. Dans chacun des branchements obtenus par la première coupe binaire, d'autres facteurs peuvent apparaître et créer des branches supplémentaires. Les histogrammes au bas de l'arbre représentent alors la distribution relative de la variable de réponse pour chaque combinaison de facteurs (dans le présent cas, le taux de dislocation, en gris foncé dans les histogrammes).

Arbre de partition 1 – Distribution et interaction des effets fixes pour les dislocations chez les enfants francophones



Nous pouvons constater que les différents facteurs entretiennent des relations complexes. Le facteur le plus important (en dehors du statut de topic, qui, rappelons-le, n'entre pas dans le modèle, puisque les dislocations correspondent toujours au topic chez nos enfants) semble être l'association avec un des schèmes fréquemment observées avec les dislocations (ça c'est X, c'est quoi ça, il/elle est où X, ...), représenté par le noeud 1 tout en haut. Pour ces schèmes fréquents (branche droite), le deuxième facteur est ensuite le groupe MLU (noeud 13), les enfants dans le groupe MLU3 produisant plus de dislocations.

Nous constatons alors que la discontinuité topicale joue un rôle à plusieurs niveaux, pour les enfants du groupe MLU 3 (noeud 17), et plus généralement lorsque l'expression n'est pas associée à un des schèmes fréquents (noeud 2). La présence d'un mouvement interactionnel, où la dislocation crée un lien de retour sur un TdP ou énoncé précédent, et prise en compte elle aussi à plusieurs niveaux, pour les enfants du MLU 2 dans les schèmes fréquents (noeud 14), et pour les enfants du MLU 3 dans les topics discontinus qui n'appartiennent pas un schème (noeud 5). La discontinuité topicale est aussi le facteur qui compte pour la majorité des dislocations n'appartenant à un des schèmes privilégiés, associées là aussi à un MLU plus élevé et la présence d'un mouvement interactionnel. La relation entre dislocation et fonction sujet (ou THEM<sub>Z</sub>) apparaît alors comme moins importante, en interaction avec un mouvement interactionnel de RETOUR, dans des cas non concernés par ni par un schème lexico-syntaxique ni par une discontinuité topicale. La hiérarchie de l'arbre nous permet donc, malgré sa complexité relative, de conclure que l'association à un des schèmes lexico-syntaxiques fréquents (noeud 1), suivi de la présence d'un mouvement interactionnel de RETOUR et/ou d'un contexte de discontinuité topicale sont les facteurs les plus importants, dont l'interaction complexe explique, conjointement à un MLU plus élevé, l'emploi d'une dislocation.

### 2.1.2 Les adultes

Pour les adultes, nous avons observé une seule occurrence de dislocation qui ne correspond pas au topic de l'énoncé. Comme pour les enfants ci-dessus, cela nous semble peu pertinent, et pose également le problème de la séparation (quasi-)complète. En outre, cela risque de compliquer le modèle pour le bénéfice d'expliquer la différence d'une seule occurrence. Nous avons alors choisi, comme pour les enfants, de ne pas inclure le statut de topic dans le modèle pour les dislocations chez l'adulte. Le Tableau XI-3 ci-dessous présente les effets retenus dans le modèle qui correspond le mieux à nos données.

Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

Tableau XI-3 - Résultats du modèle linéaire mixte pour les dislocations (adultes, français)

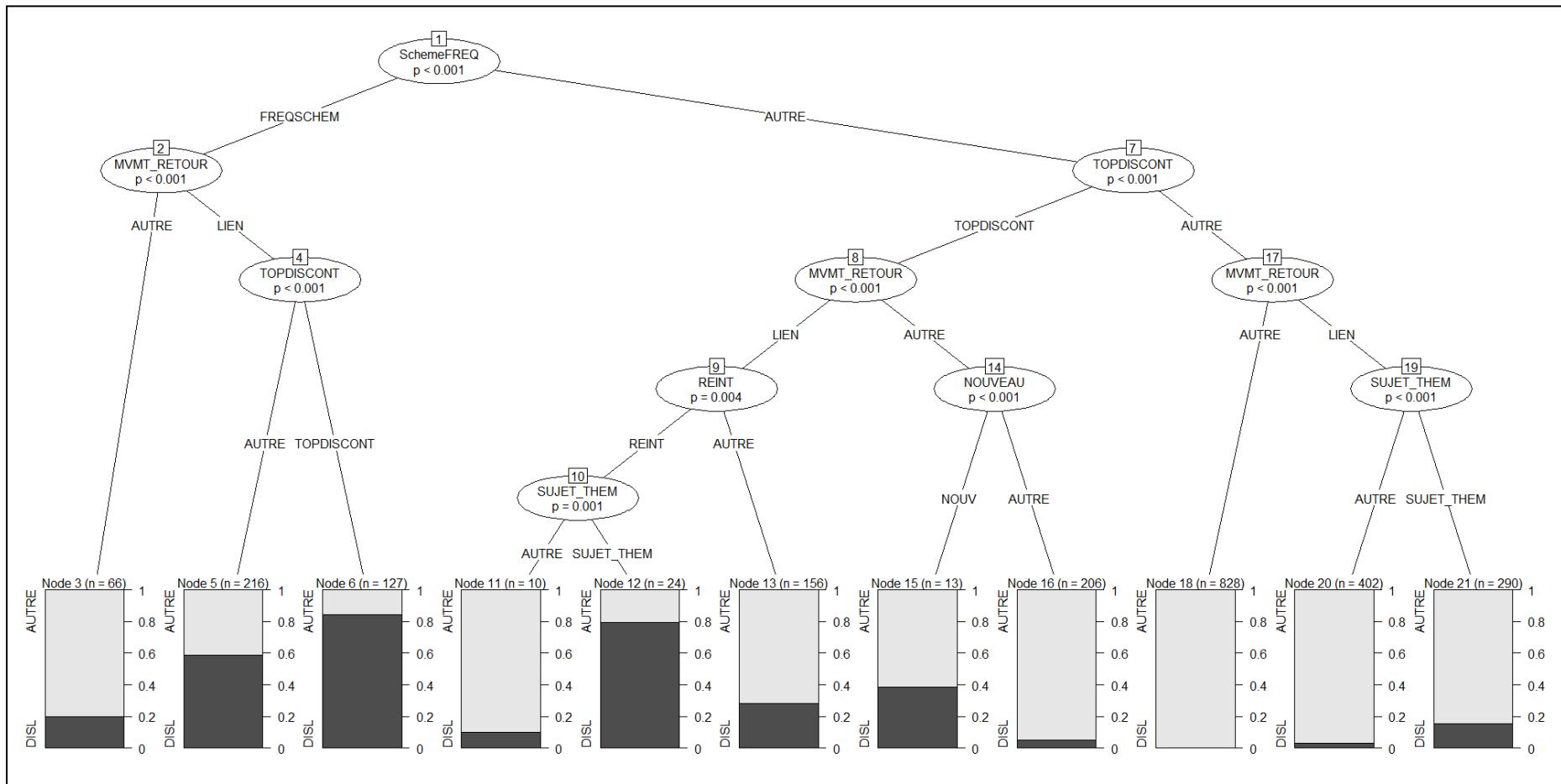
Dislocations, C=0.9236687				
Effets aléatoires :	Variance	SD		
Session	0.2018	0.4492		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-5.6330	0.3045	-18.502	<.00001 ***
SUJET_THEME	1.2501	0.1932	6.471	<.00001 ***
SCHEMEEFREQ	2.4547	0.1746	14.059	<.00001 ***
TOPDISCONT	1.5504	0.1668	9.293	<.00001 ***
MVMT_RETUR	2.3435	0.2238	10.471	<.00001 ***
Statut attentionnel (Réactivé)	0.6583	0.2870	2.294	0.00218 **
Statut attentionnel (Nouveau)	1.0702	0.4412	2.426	0.0153 *

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 2338

Comme pour les enfants, nous observons un effet de la fonction sujet/THEM<sub>Z</sub> et de l'association avec les schèmes lexico-syntactiques fréquents. L'emploi d'une dislocation est également positivement lié aux contextes de topic discontinu ainsi qu'à la présence d'un mouvement de RETOUR sur un TdP ou énoncé précédent. Contrairement au modèle pour les enfants, nous constatons aussi un effet du statut attentionnel, l'emploi d'une dislocation étant plus probable pour introduire ou réactiver un référent (par opposition aux référents activés ou donnés dans le discours). Le groupe MLU des enfants avec lesquels interagissent les adultes n'apparaît pas comme pertinent dans ce modèle.

Comme le montre l'arbre de partition ci-dessous, l'interaction entre ces différents facteurs est relativement complexe et suggère que souvent, un cumul de plusieurs des facteurs considérés motive l'emploi d'une dislocation. Le facteur principal, comme pour les enfants, est l'association avec un des schèmes identifiés dans nos données comme fréquemment associés avec les dislocations. Les dislocations associées à un de ces schèmes sont aussi fréquemment impliquées dans un mouvement interactionnel qui crée un RETOUR sur un énoncé précédent, que ce soit une activité d'énumération, la réitération d'une demande, une (demande de) confirmation ou une dénomination contrastante avec celle proposée par l'enfant, pour rappeler quelques exemples. Dans ce cas de figure, la dislocation effectue aussi fréquemment un changement de topic (cas des listes typiquement, mais aussi des réitérations/relances après que l'enfant a parlé d'autre chose, pour le rappeler à la tâche interactionnelle en cours). Dans d'autres cas, le mouvement interactionnel se fait avec maintien du topic, et ce sont là essentiellement des cas où la mère réitère plusieurs fois une question, soit pour reformuler, soit parce que l'enfant ne réagit pas ou semble avoir généralement perdu l'intérêt pour l'activité, comme nous l'avons montré dans le chapitre précédent.

## Arbre de partition 2 – Distribution et interaction des effets fixes pour les dislocations chez les adultes francophones



## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

---

Dans les cas où la dislocation n'est pas associée à l'un de ces schèmes fréquents, elle sert à changer de topic dans la majorité des cas, où alors à effectuer un mouvement interactionnel de RETOUR. Nous observons dans ces catégories des sous-groupes, avec des interactions entre topic discontinu, mouvement interactionnel et introduction ou réactivation d'un référent. L'arbre de partition montre aussi que les facteurs du statut attentionnel et de la fonction syntaxique sont secondaires, et comme pour les enfants, la présence d'un schème lexico-syntaxique ainsi que le changement de topic et le mouvement de RETOUR, ainsi que leur combinaison, expliquent une grande partie des données.

### 2.2 Les noms

#### 2.2.1 Les enfants

Passons maintenant aux facteurs influençant la production un nom pour référer à une entité. Le modèle final nous donne une très forte qualité d'ajustement à la distribution observé, avec une valeur de C proche de 0.90 :

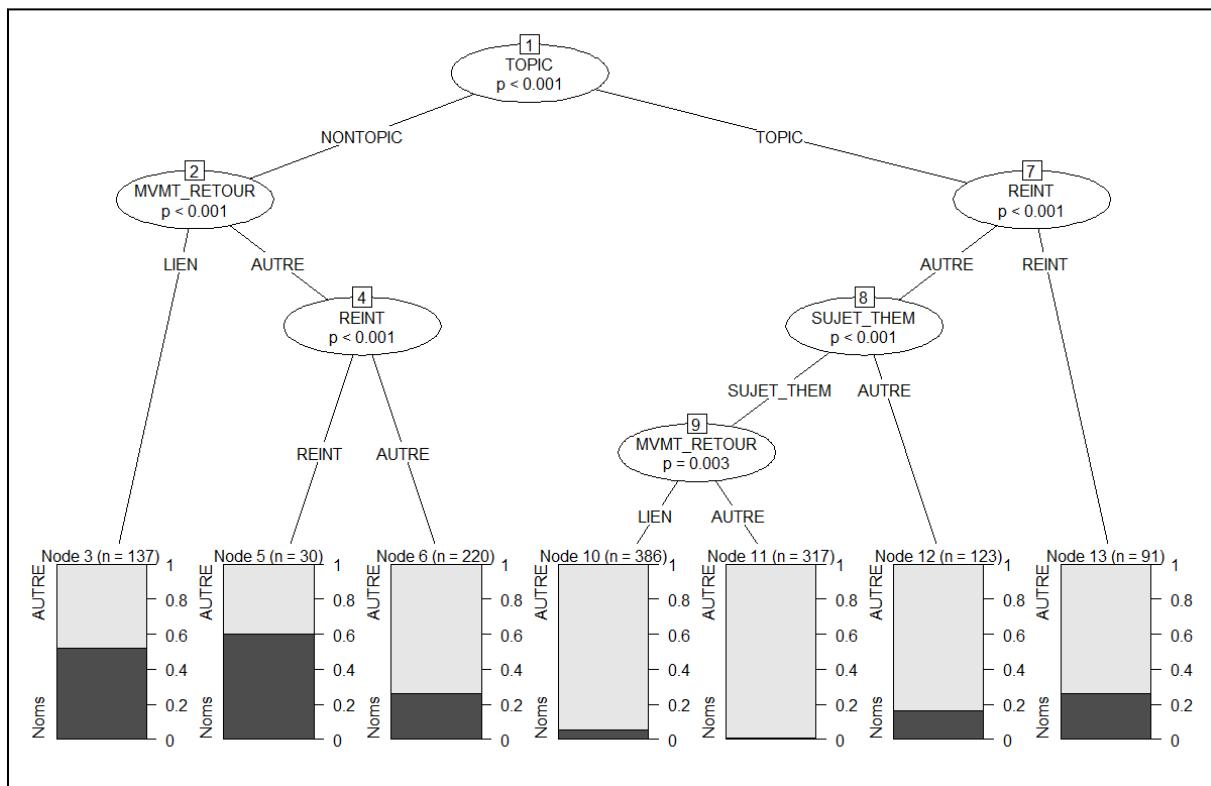
Tableau XI-4 - Résultats du modèle linéaire mixte pour les noms (enfants, français)

Noms, C=0.8914292				
Effets aléatoires :	Variance	SD		
Session	1.407	1.186		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-0.6865	0.3775	-1.818	0.06902 .
TOPIC	-1.8309	0.3254	-5.626	<.00001 ***
SUJET_THEME	-1.9978	0.2917	-6.848	<.00001 ***
TOPDISCONT	0.6402	0.3132	2.044	0.04093 *
MVMT_RETUR	1.0108	0.2012	5.023	<.00001 ***
Statut attentionnel (Réintroduit)	0.9172	0.2666	3.441	0.00058 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1304

Nous constatons, en concordance avec nos résultats dans les CHAPITRE VII et CHAPITRE IX, que les noms ne sont pas favorisés dans l'expression du topic (valeur de z négative), ni en fonction sujet. En revanche, lorsqu'ils correspondent au topic, ils sont favorisés dans des contextes de discontinuité topicale. De plus, nous observons une relation positive avec des mouvements de RETOUR sur un énoncé ou TdP précédent, et avec le statut attentionnel : les noms sont favorisés pour la réintroduction d'un référent. L'arbre de partition ci-dessous permet de visualiser l'interaction des différents facteurs :

## Arbre de partition 3 – Distribution et interaction des effets fixes pour les noms chez les enfants francophones



Cet arbre nous montre que le facteur principal pour avoir un nom dans la référence aux entités, est le statut de non-topic (autrement dit, le topic est exprimé préférentiellement par autre chose qu'un nom, ce que nous avons vu dans le CHAPITRE IX-2.1.1.1). En même temps, nous observons que ces noms qui ne correspondent pas au topic font fréquemment l'objet d'un mouvement de RETOUR sur un énoncé précédent. C'est typiquement le cas lorsqu'un énoncé est réitéré en cas d'absence de réponse, ou reformulé pour préciser :

## Exemple XI-1 - [FRA] Lola/3;00.01/MLU3+/Maison Poupées

- |       |  |
|-------|--|
| MER72 | ah@i il y a un téléphone ?                                 |
| MER72 | fais voir où il est  |
| MER72 | je le vois pas moi .                                       |
| ENF71 | il est là . {a toujours son bras dans la maison, pointe ?} |
| MER73 | fais voir ?  |
| ENF72 | là dans <b>la cuisine</b> là .                             |

Nous le trouvons aussi dans des séquences de reformulations métalinguistiques qui ne font pas partie d'une séquence de dénomination (nous y reviendrons dans la section concernant les référents non-verbalisés ci-dessous) :

## Exemple XI-2 - [FRA] Garance/2;04.19/MLU2/Maison

- |       |   |
|-------|---|
| MER51 | elle marche pas la sonnette.                |
| ENF54 | [nolet] ‘sonnette’ ?                        |
| MER52 | <b>la sonnette.</b>                         |
| ENF55 | [sonet] ‘sonnette’ +... {cherche du regard} |

Dans les autres cas, le nom correspond plus souvent à une réintroduction d'un référent qu'à un autre statut attentionnel. Lorsque le nom a le statut de topic, là aussi il sert le plus souvent à réintroduire

## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

un référent momentanément abandonné (ce qui, au niveau de la chaîne topicale, peut correspondre à un topic nouveau ou réactivé). Du reste, pour les topics qui ne correspondent pas au statut attentionnel Réintroduit, le nom encode soit un référent en fonction Non-Sujet, soit il est concerné par un mouvement interactionnel de RETOUR.

### 2.2.2 Les adultes

Chez les adultes, nous observons également une relation négative entre l'emploi d'un nom et le statut de topic et la fonction sujet :

Tableau XI-5 - Résultats du modèle linéaire mixte pour les noms (adultes, français)

Noms, C=0.8862056				
Effets aléatoires :	Variance	SD		
Session	0.1207	0.3474		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-0.5983	0.1446	-4.137	<.00001 ***
TOPIC	-1.7848	0.1802	-9.903	<.00001 ***
SUJET_THEME	-2.2364	0.1867	-11.980	<.00001 ***
SCHEMEEFREQ	-1.8760	0.3926	-4.778	<.00001 ***
TOPDISCONT	0.6704	0.2030	3.303	0.000958 ***
MVMT_RETUR	1.7060	0.1392	12.259	<.00001 ***
Statut attentionnel (Nouveau)	0.9239	0.2161	4.276	<.00001 ***
Statut attentionnel (Réintroduit)	1.1323	0.1926	5.880	<.00001 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 2338

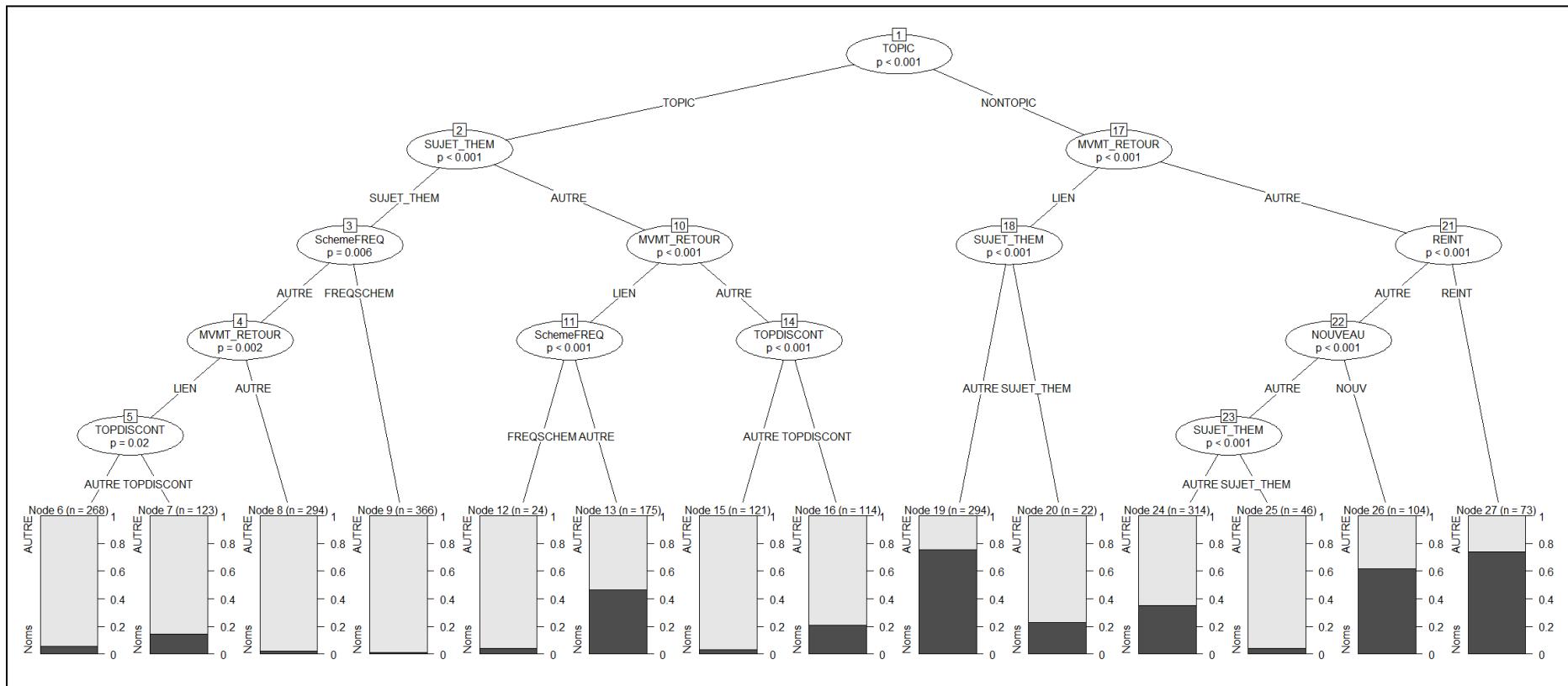
De plus, les adultes produisent également des noms dans les schèmes identifiés comme fréquents avec les dislocations (voir Exemple XI-3 ci-dessous), mais les noms n'y sont pas favorisés, contrairement à ce que nous avons vu pour les dislocations ci-dessus.

#### Exemple XI-3

- |                  |        |                                      |
|------------------|--------|--------------------------------------|
| Garance (Puzzle) | MER123 | I'escargot est là. {pointe escargot} |
|                  | MER144 | où est la vache ?                    |
| Côme (Puzzle)    | MER110 | où sont les poussins ?               |

Dans les contextes favorisant l'emploi d'un nom, nous avons, comme pour les enfants, la discontinuité topicale, la présence de mouvement de RETOUR, et le statut attentionnel. Les noms sont favorisés pour réintroduire, mais aussi pour introduire un référent nouveau. Considérons maintenant l'interaction des facteurs : le statut de topic apparaît comme le facteur le plus important, comme pour les enfants. Les noms qui ne correspondent pas au topic sont fréquemment associés à un mouvement de RETOUR sur un énoncé précédent (reformulation ou réitération), et fréquemment en fonction syntaxique autre que sujet (ce dernier facteur vaut aussi pour les noms qui encodent le topic, nœud 2). Lorsqu'un nom encode le topic, il a alors tendance à être concerné par un mouvement de RETOUR et/ou de changer de topic.

## Arbre de partition 4 – Distribution et interaction des effets fixes pour les noms chez les adultes francophones



## 2.3 Les pronoms démonstratifs

Pour le calcul du modèle des pronoms démonstratifs du français, nous avons choisi de regrouper les formes clitiques (*c'est*) et les formes toniques ou fortes (*ça*), pour deux raisons : la fréquence plutôt faible des démonstratifs toniques (en dehors de ceux employés dans une dislocation) d'une part, et le souci de comparabilité avec l'allemand d'autre part, où le D-Pro semble fonctionner à la fois comme une forme faible et une forme forte. Comme dans le CHAPITRE VII-2.3, les deux types de pronom démonstratif français sont donc regroupés dans une seule catégorie GLOBDEMO.

### 2.3.1 Les enfants

Les facteurs pertinents pour l'emploi des démonstratifs chez les enfants sont donnés dans le modèle ci-dessous :

Tableau XI-6 - Résultats du modèle linéaire mixte pour les pronoms démonstratifs (enfants, français)

GLOBDEMO, C=0.8082302				
Effets aléatoires :	Variance	SD		
Session	0.9826	0.9912		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-3.7212	0.3749	-9.925	<.00001 ***
SUJET_THEME	1.1591	0.2144	5.407	<.00001 ***
MVMT_RETUR	0.8578	0.1860	4.612	<.00001 ***
Statut attentionnel (Nouveau)	1.0456	0.3445	3.035	0.00241 **
Statut attentionnel (Activé)	1.0118	0.2040	4.959	<.00001 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1304

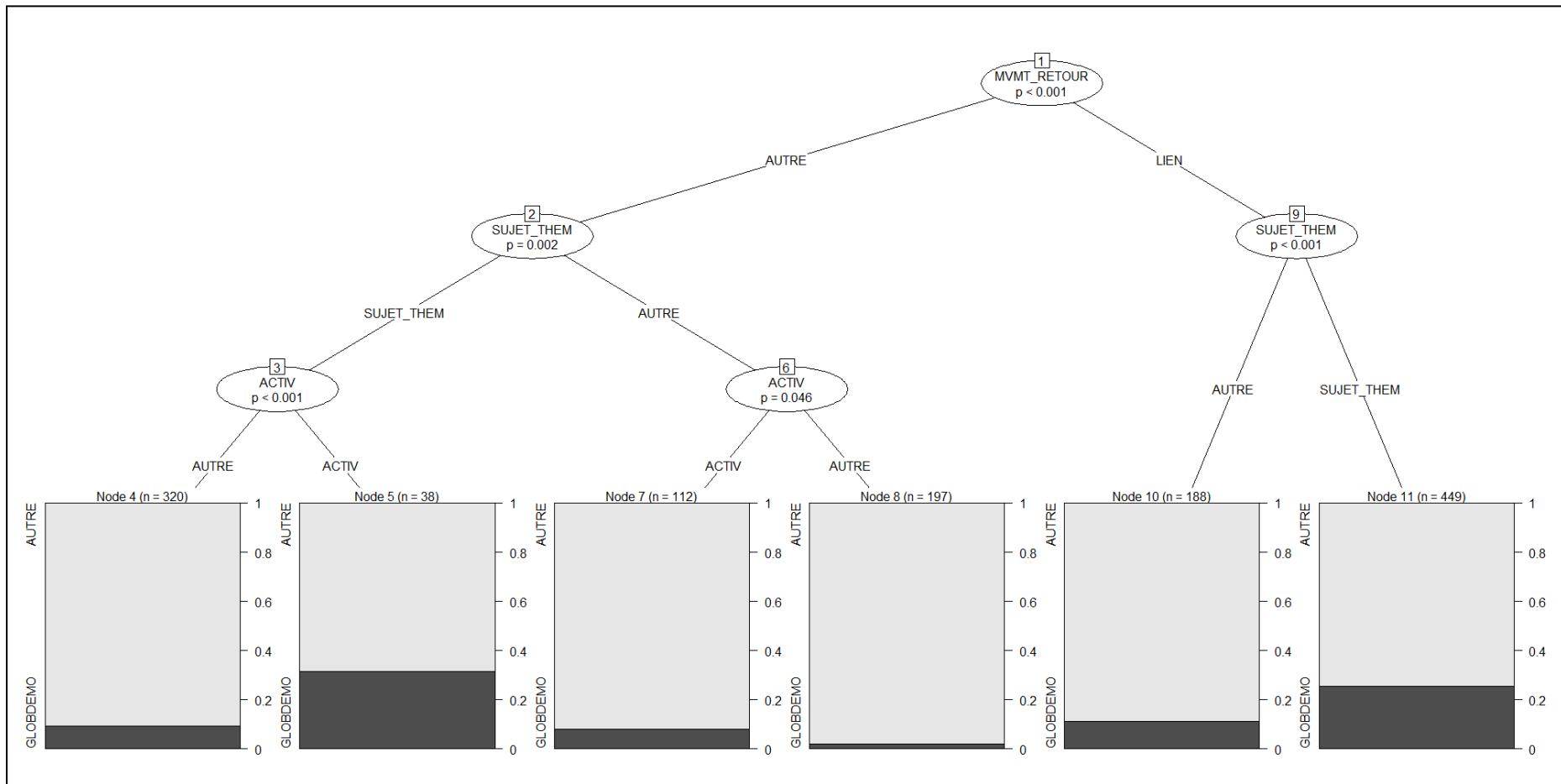
Dans ce modèle, aucune des mesures du statut de topicalité n'était pertinente.<sup>256</sup> Pour les enfants français, le démonstratif n'est donc pas spécialement lié aux statuts de topic. Nous verrons plus loin que ce facteur influence la production chez les adultes, mais d'un degré bien moindre que ce n'est le cas qu'en allemand. La production d'un démonstratif était en revanche positivement influencée par la fonction sujet, la présence de mouvement interactionnel de RETOUR, et pour des référents aux statuts attentionnels Nouveau et Activé.

L'arbre de partition ci-dessous révèle que le facteur principal chez nos enfants francophones est la présence d'un mouvement RETOUR ; cela correspond largement aux conduites de (demande de) confirmation ou d'opposition dans la dénomination d'un objet, décrits dans le chapitre précédent. La préférence pour la fonction sujet semble être un facteur principal, indépendamment de la présence ou non d'un mouvement RETOUR, les démonstratifs sont plus fréquemment en fonction sujet. Cela est essentiellement dû à la grande fréquence du démonstratif clitique dans le présentatif *c'est* en français.

---

<sup>256</sup> Ce résultat ne changeait pas lorsque nous avons vérifié séparément les démonstratifs clitiques et toniques.

## Arbre de partition 5 – Distribution et interaction des effets fixes pour les pronoms démonstratifs chez les enfants francophones



## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

Enfin, contrairement au modèle *glmer* présenté dans le Tableau XI-6, qui montre une influence des statuts attentionnels Nouveau et Activé, seulement le statut Activé apparaît dans l'arbre de partition, là encore, favorisant l'emploi d'un démonstratif.

### 2.3.2 Les adultes

L'emploi d'un démonstratif chez nos adultes semble répondre à un ensemble de facteurs plus complexes que ce que nous venons d'observer chez les enfants, et nous observons notamment une influence du statut de topic.

Tableau XI-7 - Résultats du modèle linéaire mixte pour les pronoms démonstratifs (adultes, français)

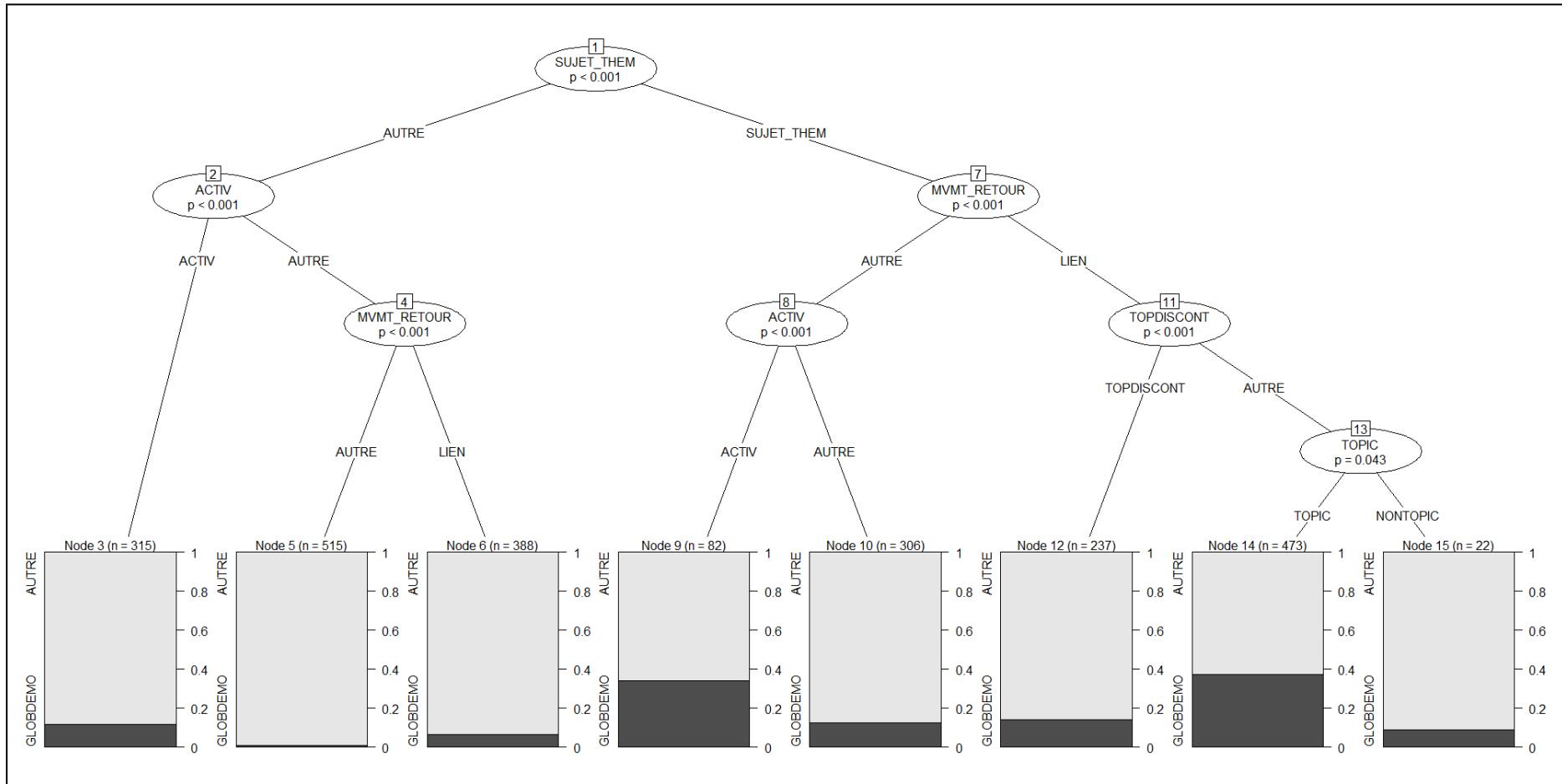
GLOBDEMO, C=0.7807048				
Effets aléatoires :	Variance	SD		
Session	0.2083	0.4564		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-4.4329	0.3772	-11.751	<.00001 ***
TOPIC	0.4654	0.2302	2.021	0.043257 *
SUJET_THEME	1.5487	0.1822	8.498	<.00001 ***
TOPDISCONT	-0.6601	0.1936	-3.410	0.000650 ***
MVMT_RETUR	0.5295	0.1386	3.820	0.000133 ***
Statut attentionnel (Donné)	1.0851	0.3580	3.031	0.002437 **
Statut attentionnel (Activé)	1.7672	0.3492	5.061	<.00001 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 2338

Le statut de topic, la fonction sujet, la présence d'un mouvement de RETOUR ainsi qu'un statut attentionnel de Donné ou d'Activé apparaissent comme les facteurs favorisant l'emploi d'un démonstratif. En revanche, nous observons une relation négative avec la discontinuité topicale. Observons maintenant l'interaction et la distribution de ces facteurs dans l'arbre de partition :

Ici, le facteur principal semble être la fonction de sujet (ou THEM), contrainte structurelle du français. Lorsque le démonstratif n'est pas en fonction sujet, il encode soit un référent déjà activé, soit il correspond à un mouvement de RETOUR (nœuds 3 et 6). Dans la majorité des cas, en revanche, le démonstratif est en fonction sujet, impliqué dans un mouvement de RETOUR (confirmation/opposition des dénominations notamment, comme évoqué pour les enfants) et en continuité topicale (nœud 14). Nous verrons alors en section 3 que les facteurs favorisant la production d'un D-Pro en allemand sont assez différents.

## Arbre de partition 6 – Distribution et interaction des effets fixes pour les pronoms démonstratifs chez les adultes francophones



## 2.4 Les pronoms personnels clitiques

Passons maintenant aux modèles pour les pronoms personnels clitiques, expression par excellence de la continuité dans le discours.

### 2.4.1 Les enfants

Nous avons vu dans le CHAPITRE VII que nos enfants francophones ne produisent pas encore beaucoup de pronoms personnels clitiques (80 occurrences, 6% des expressions en référence aux entités). Nous avons voulu quand même estimer les facteurs influençant leur emploi et les comparer à ceux pertinents pour les adultes. Nous n'avons pas regroupé les pronoms personnels clitiques avec les fillers à ce stade, puisque les fillers peuvent aussi préfigurer le *c'est*, que nous avons groupé avec les formes faibles, mais dont nous avons vu que la distribution en fonction des différents facteurs considérés dans les chapitres précédents n'est pas identique à celle des pronoms personnels clitiques. Le démonstratif clitic se situe quelque part à mi-chemin entre les démonstratifs forts et les pronoms personnels clitiques de troisième personne. Nous observons donc ici seulement la distribution des pronoms personnels clitiques, en gardant à l'esprit la taille relativement faible de l'échantillon.

Le meilleur modèle pour expliquer la distribution de ces pronoms personnels clitiques est donné dans le Tableau XI-8 :

Tableau XI-8 - Résultats du modèle linéaire mixte pour les pronoms personnels clitiques (enfants, français)

Pronoms personnels clitiques, C=0.8635236				
Effets aléatoires :	Variance	SD		
Session	1.265	1.125		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-4.5970	0.5572	-8.250	<.00001 ***
TOPIC	2.7852	0.4529	6.150	<.00001 ***
TOPDISCONT	-1.0866	0.3214	-3.380	0.000724 ***
MVMT_RETUR	-1.7909	0.3061	-5.851	<.00001 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1304

Cela nous montre que la production d'un pronom personnel clitic semble être favorisée lorsque l'expression encode le topic de l'énoncé, et au contraire défavorisée dans un contexte de discontinuité topicale et d'un mouvement interactionnel de RETOUR, ce qui correspond à nos observations des chapitres précédents. Le fait que le facteur de la fonction syntaxique n'apparaît pas dans ce modèle ne veut pas dire qu'il n'est pas pertinent dans le choix d'un pronom clitic, de même pour le statut attentionnel. Lorsque ce sont les seuls facteurs que nous entrons dans le modèle et le comparons à un modèle sans aucun facteur, la fonction syntaxique et le statut attentionnel de Donné apparaissent comme significatifs :

Tableau XI-9 - Modèle linéaire mixte alternatif pour les pronoms personnels clitiques (enfants, français)

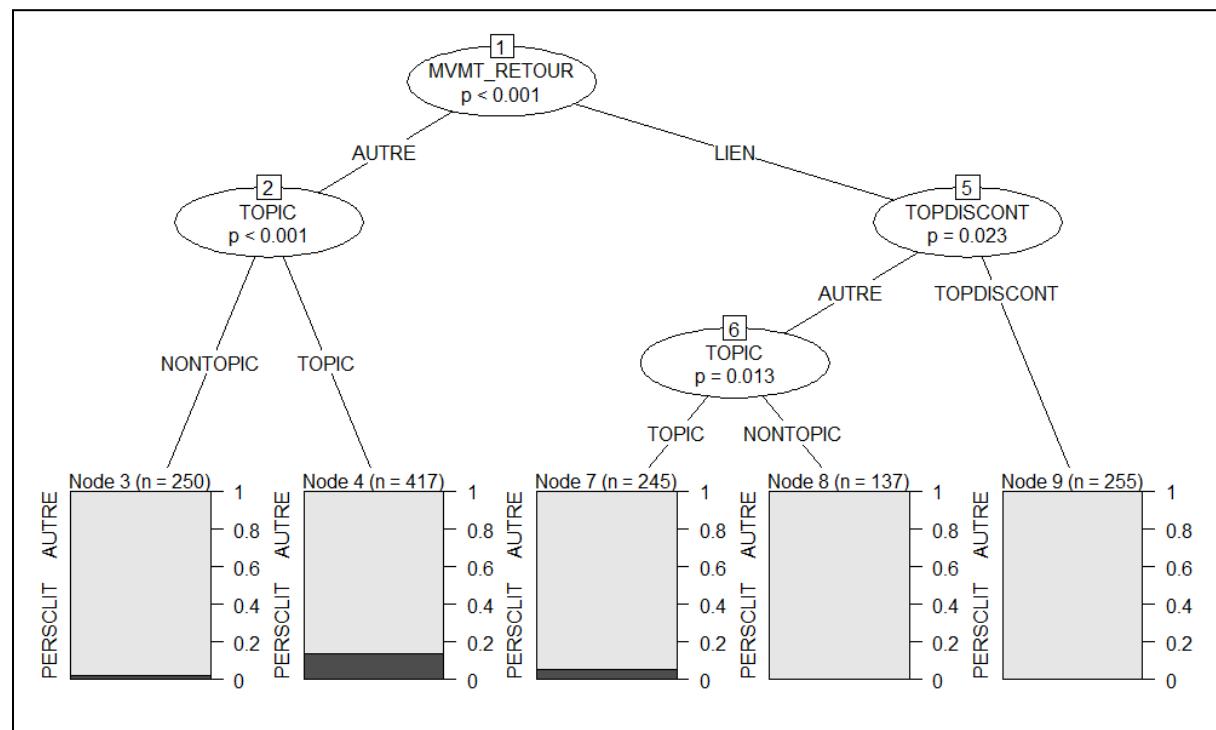
Model 1 : Pronoms personnels clitiques, C=0.8040506				
Effets aléatoires :	Variance	SD		
Session	1.265	1.125		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-4.9565	0.4957	-9.998	<.00001 ***
SUJET_THEME	1.2743	0.3269	3.898	<.00001 ***
Statut attentionnel (Donné)	1.2212	0.3099	3.941	<.00001 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1304

Dès lors que le facteur du statut de topic et celui de la discontinuité entrent dans le modèle cependant, SUJET\_THEME et le statut Donné n'apportent pas de pouvoir explicatif supplémentaire, autrement dit, il ne reste plus beaucoup de variance à expliquer par ces facteurs une fois que le topic est pris en compte. Donc, les facteurs retenus dans le modèle final, en Tableau XI-8 ci-dessus, semblent être ceux qui expliquent au mieux la distribution observée. L'absence d'un effet du MLU peut paraître surprenant aussi, étant donné que les pronoms personnels clitiques sont une catégorie sous acquisition. Toutefois, nous avons montré dans le CHAPITRE VII que nous constatons une corrélation positive entre MLU et les pronoms personnels clitiques dans la référence à la personne, mais non pas dans la référence aux entités. Il n'est donc en ce sens pas étonnant que le MLU ne figure pas dans ce modèle, calculé pour les expressions en référence aux entités.

L'interaction entre les effets fixes retenus est donnée dans l'arbre de partition ci-dessous :

Arbre de partition 7 – Distribution et interaction des effets fixes pour les pronoms personnels clitiques chez les enfants francophones



La majorité des pronoms personnels clitiques ne sont pas concernés par un mouvement de RETOUR et ont le statut de topic. Ceux qui sont concernés par un mouvement sont des topics qui ne

## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

correspondent pas à un changement ; i.e., ce sont des cas de maintien de topic (nœud 7). Le type de mouvement en question est alors une réitération à l'identique d'un énoncé qui ne reçoit pas de réaction, comme dans cet exemple de Clément, déjà discuté dans le CHAPITRE X (Exemple X-54), et qui peut consister en une « stratégie » de répétition à l'identique d'un même énoncé pour une même intention communicative :

### Exemple XI-4 - [FRA] Clément/2;03/MLU2/Puzzle

MER63	oh qu'il est content papa ! {CHI et MOT se tournent pour regarder PER entrer dans la pièce}
MER63	qu'est-ce que c'est qu'il a <dans la main> [>] ?
PER1	<0 [=! rit]> [<].
MER64	oh dis don(c) dis don(c).
MER64	oh il est drôlement <content> [=! rit].
ENF65	[le la] ‘(i)l est là’ ? {CHI se met debout sur sa chaise pour mieux voir ce que fait FAT}
MER65	<il trouve des trésors papa> [>].
ENF66	<[le la] ‘(i)l est là.’> [<]
	{CHI montre PER du doigt}
ENF66	<[le la] ‘(i)l est là .’> [<]
	{CHI montre PER du doigt}
MER66	il cherche des trésors
MER66	et <il les trouve> [>].
ENF67	[il e la] <il est là> [<] !
MER67	il est là papa.

### 2.4.2 Les adultes

Le Tableau XI-10 présente le modèle retenu pour les pronoms personnels clitiques chez nos adultes francophones :

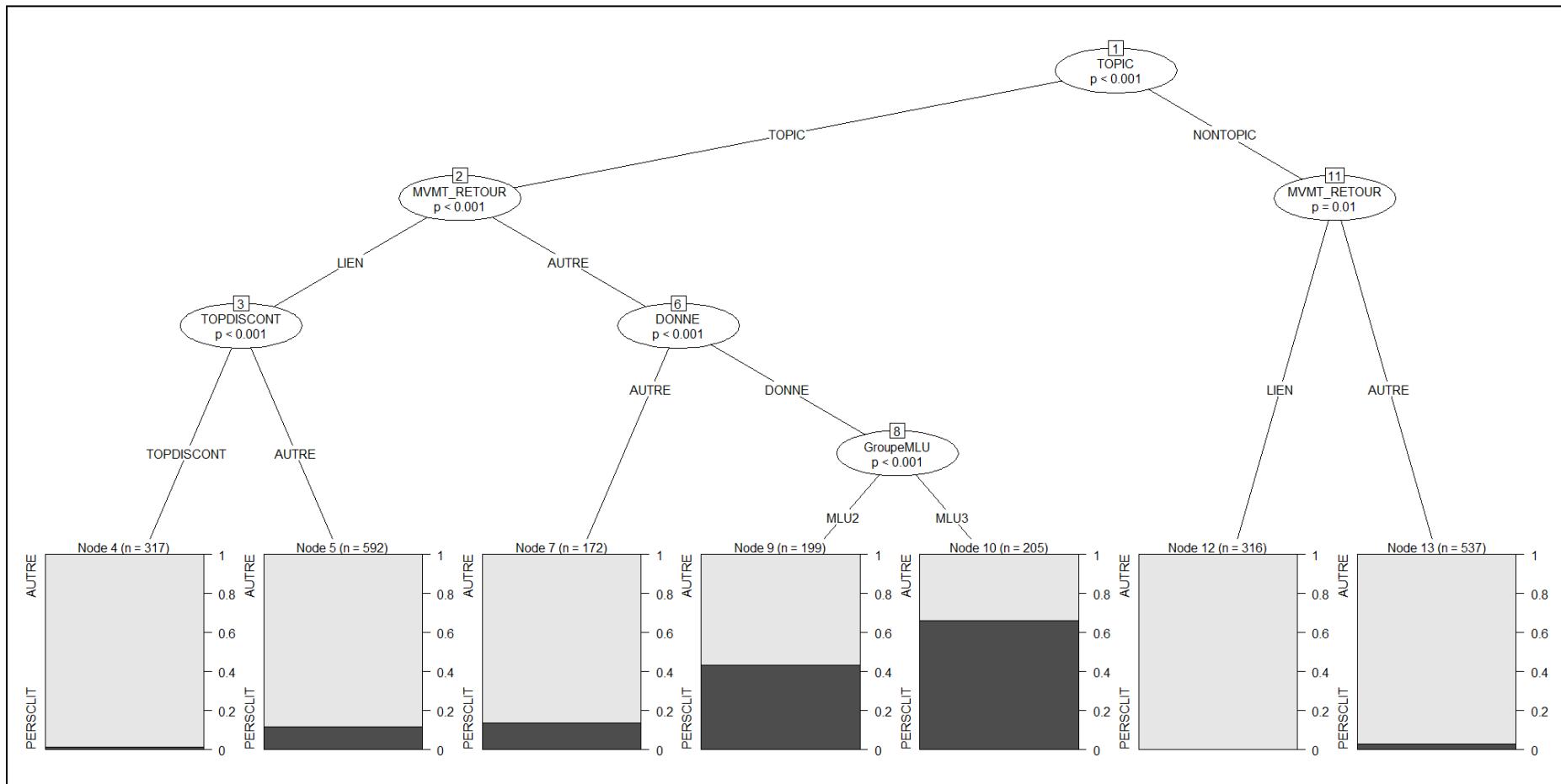
Tableau XI-10 - Résultats du modèle linéaire mixte pour les pronoms personnels clitiques (adultes, français)

Pronoms personnels clitiques, C=0.889128				
Effets aléatoires :	Variance	SD		
Session	0.2884	0.537		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-4.8547	0.3892	-12.474	<.00001 ***
TOPIC	3.3234	0.2914	11.403	<.00001 ***
TOPDISCONT	-0.4787	0.2023	-2.367	0.0180 *
MVMT_RETUR	-2.5592	0.1658	-15.438	<.00001 ***
Statut attentionnel (Donné)	1.4861	0.2380	6.244	<.00001 ***
Groupe MLU (MLU3)	0.6928	0.3446	2.010	0.0444 *
SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 2338				

Pour les adultes, nous obtenons un modèle avec les mêmes effets que pour les enfants : une relation positive avec le statut de topic, et une relation négative avec les topics discontinus et les mouvements de RETOUR. De plus, nous observons pour les adultes que le statut attentionnel « donné dans le discours » est cette fois-ci parmi les facteurs favorisants significatifs, résultat attendu pour les pronoms personnels clitiques. Les mères des enfants du groupe MLU 3 semblent également employer plus de pronoms personnels clitiques que celle dont les enfants ont un MLU plus bas.

Comment s'organise alors l'interaction de ces facteurs ? L'arbre de partition montre que le statut de topic est de loin le facteur principal, avec la grande majorité des pronoms correspondant au topic, suivi de l'absence d'un mouvement de RETOUR. La production d'un pronom personnel clitique est alors plus favorisée encore si le référent est donné, et davantage produit si l'enfant appartient au groupe MLU 3. Lorsque le pronom personnel clitique est concerné par un mouvement de RETOUR, c'est alors, comme pour les enfants, majoritairement dans un contexte de continuité topicale (nœud 5), et correspond, comme pour les enfants, à des répétitions (confirmation, demande de confirmation) ou réitérations d'un énoncé du même locuteur à l'identique.

**Arbre de partition 8 – Distribution et interaction des effets fixes pour les pronoms personnels clitiques chez les adultes francophones**



## 2.5 Les référents non-verbalisés

Enfin, nous avons estimé les effets pertinents et leur interaction pour l'ensemble des référents non-verbalisés (Zero et Implicite pris ensemble), puisqu'un nombre important de topics ne sont de fait pas verbalisés. Fonctionnellement, les Zero et les Implicite semblent très proches, comme l'ont montré les analyses des chapitres précédents. Pour cette raison, et vu le nombre relativement plus faible des Zero, il nous a semblé préférable de regrouper ces catégories.

### 2.5.1 Les enfants

Le tableau ci-dessous présente le modèle final pour les référents non-verbalisés chez les enfants francophones.

**Tableau XI-11 - Résultats du modèle linéaire mixte pour les référents non-verbalisés (enfants, français)**

Non-verbalisés, C=0.8628874				
Effets aléatoires :	Variance	SD		
Session	0.5514	0.7426		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-3.8255	0.4034	-9.483	<.00001 ***
TOPIC	4.2440	0.4041	10.502	<.00001 ***
SUJET_THEME	-0.4632	0.2163	-2.142	0.032228 *
TOPDISCONT	-0.7801	0.2156	-3.618	0.000297 ***
MVMT_RETUR	-1.3305	0.1566	-8.497	<.00001 ***
Statut attentionnel (Given)	0.7135	0.2346	3.042	0.002352 **

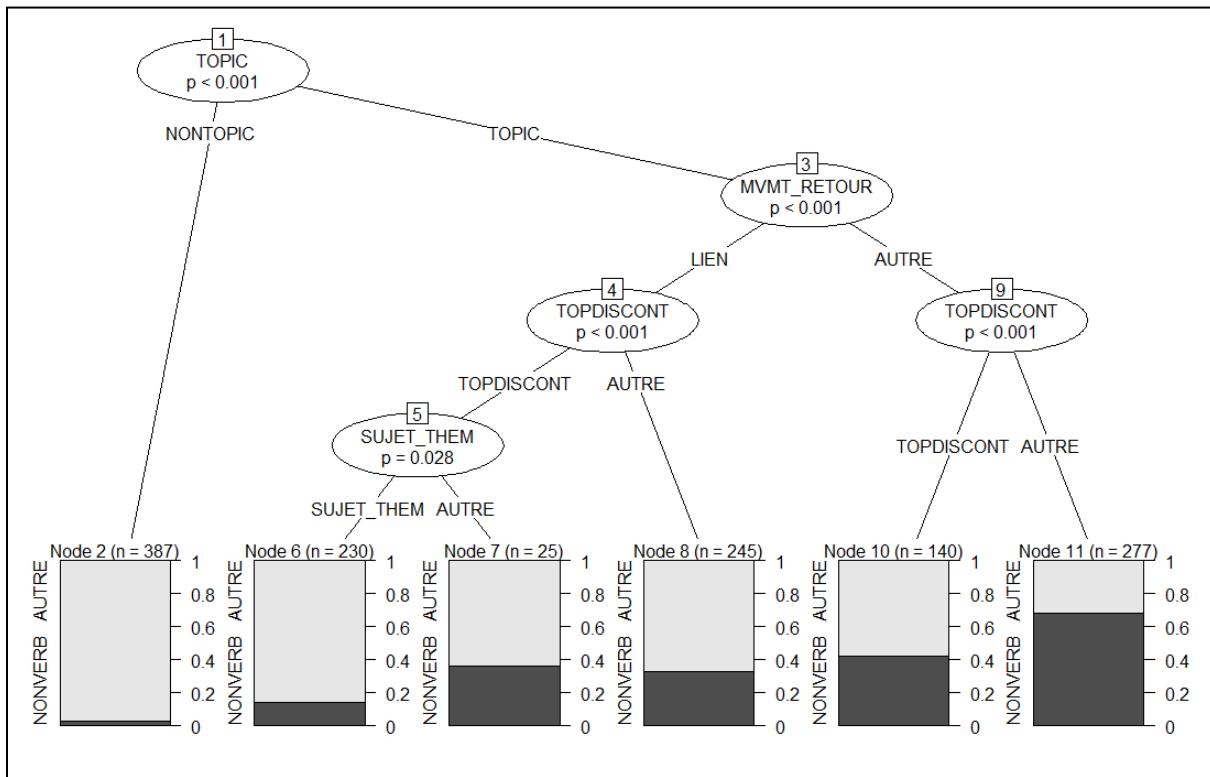
SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1304

Comme attendu, le statut de topic et le fait que le référent soit déjà donné dans le discours favorisent la non-verbalisation d'un référent, alors qu'au contraire, la discontinuité topicale et la présence d'un mouvement de RETOUR sont des contextes dans lesquels la non-verbalisation est moins probable. Nous constatons que le MLU de l'enfant ne joue pas de rôle dans cette configuration. L'on pourrait s'attendre à ce que les non-verbalisations diminuent, d'autant plus que nous avons constaté dans le CHAPITRE VII une augmentation progressive du nombre de pronoms personnels clitiques. Or, comme nous l'avons rappelé plus haut, la différence entre les groupes MLU constatée pour la réalisation des clitiques concerne la référence à la personne, et non pas la référence aux entités, où les pronoms personnels clitiques sont généralement peu fréquents et dont le taux semble, dans notre corpus, varier plutôt avec l'enfant et la session qu'avec le niveau de MLU. Pour la fonction sujet, nous constatons une relation négative. Nous savons que la relation est positive pour les seuls Implicite, et négative pour les vraies formes Zero (compléments du verbe *mettre* dans un grand nombre de cas). Étant donné que ces différences n'affectent pas les autres facteurs, nous acceptons cette conséquence du regroupement des Implicite et des Zero.

L'interaction entre les facteurs retenus se présente comme ceci :

Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

**Arbre de partition 9 – Distribution et interaction des effets fixes pour les référents non-verbalisés chez les enfants francophones**



Le facteur principal est donc le statut de topic, en conformité avec nos observations dans le CHAPITRE IX. Les topics non-verbalisés ne sont en majorité pas impliqués dans un mouvement de RETOUR, ni dans un changement de topic.<sup>257</sup> Cela dit, il reste tout de même une quantité non-négligeable de cas où un topic n'est pas verbalisé dans des contextes de changement de topic (nœud 8), d'un mouvement interactionnel de RETOUR (nœud 4), voire les deux (nœud 5). Nous avons montré dans le chapitre précédent que les enfants dans nos données semblent s'appuyer davantage que les adultes sur le contexte physique, et peuvent ainsi changer de topic sans le verbaliser si le référent en question est sous leur attention (et souvent, sous attention conjointe). De même, nous avons discuté le fait qu'un grand nombre de référents Implicites sont le support d'une dénomination, et qu'il faudrait considérer que pour un topic non-verbalisé, la mention du nom du référent permet d'effectuer un retour sur ce qui précède. Répétons ici pour le confort du lecteur l'exemple cité dans le chapitre précédent (Exemple X-62) :

<sup>257</sup> Comme nous l'avons déjà constaté pour d'autres modèles, les facteurs identifiés comme significatifs dans le modèle n'apparaissent pas nécessairement tous dans l'analyse de l'arbre de partition, c'est ici le cas du statut attentionnel (Donné). Ceci semble être du au fait d'un calcul légèrement différent dans l'arbre de partition, qui ne prend pas en compte les effets aléatoires et qui effectue un calcul récursif. Toutefois, les effets qui n'apparaissent pas dans l'arbre sont ceux qui ont la relation la moins forte dans le modèle (valeur de z dans les tableaux), et la perte d'information entre le modèle et l'arbre de partition nous semble négligeable.

**Exemple XI-5 - [FRA] Lola/3;00.01/MLU3+/Maison Poupées**

ENF58            et ça c' est [/] ça c' est son cartable  
 ENF58            et ça une petite DS<sup>258</sup> je crois. { pointe accessoires de personnage}  
 MER60            **une petite quoi ?**  
 ENF59            **une petite DS.**

**2.5.2 Les adultes**

Pour les formes non-verbalisés chez les adultes, nous obtenons un modèle avec les mêmes effets principaux que pour les enfants, et des valeurs assez proches pour les effets fixes. Le facteur de la fonction syntaxique n'est pas significatif dans ce modèle, et n'y figure donc pas. Comme nous l'avons discuté ci-dessus, cela est lié au fait qu'en français, les Implicite correspondent à la fonction sujet ou THEM en majorité, mais non pas les Zero, qui sont essentiellement des compléments de certains verbes. Nous verrons dans la section sur l'allemand que là, les formes non-verbalisés sont plus clairement favorisés par la fonction sujet dans l'ensemble.

**Tableau XI-12 - Résultats du modèle linéaire mixte pour les référents non-verbalisés (adultes, français)**

Pronoms personnels clitiques, C=0.8263079				
Effets aléatoires :	Variance	SD		
Session	0.2334	0.4832		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-4.0097	0.3540	-11.326	<.00001 ***
TOPIC	3.7621	0.3798	9.905	<.00001 ***
TOPDISCONT	-1.1044	0.2365	-4.670	<.00001 ***
MVMT_RETUR	-1.4917	0.1560	-9.560	<.00001 ***
Statut attentionnel (Donné)	-0.5314	0.2457	-2.163	0.0306 *

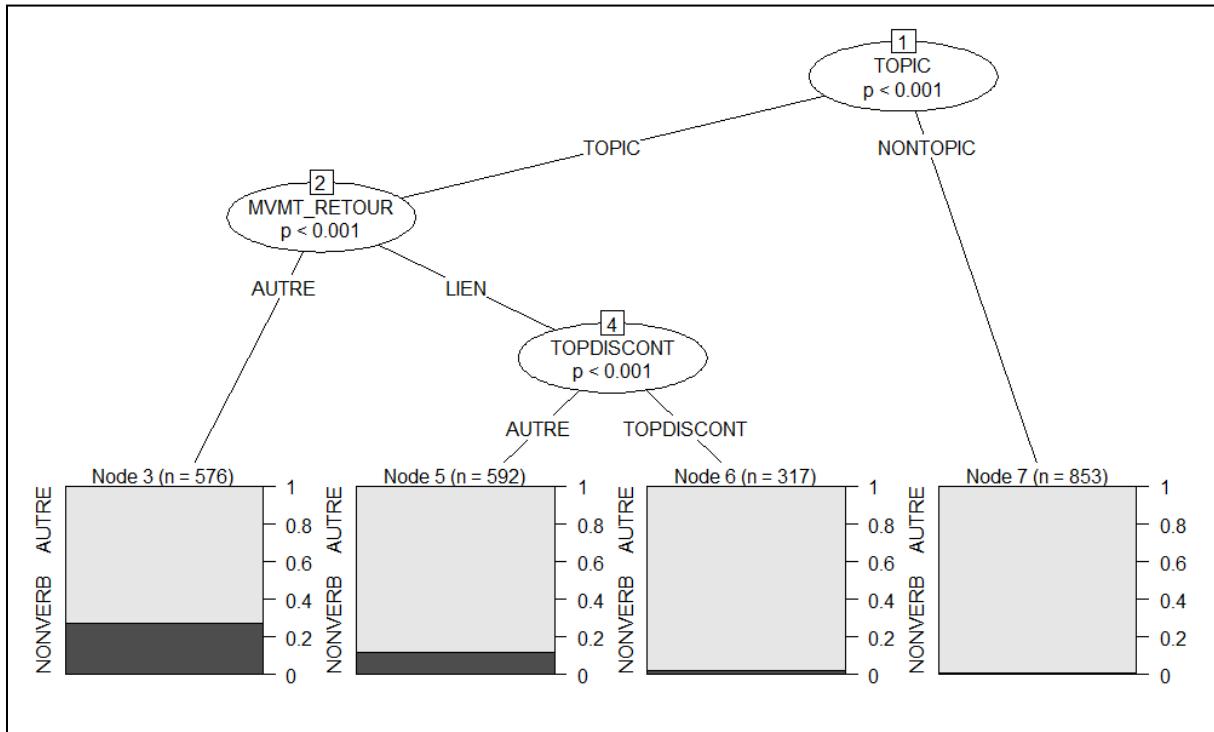
SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 2338

L'interaction entre les facteurs, en revanche, montre que la distribution des référents non-verbalisés est bien moins hétérogène chez les adultes :

<sup>258</sup> DS : il s'agit de la console de jeu électronique portable.

Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

**Arbre de partition 10 – Distribution et interaction des effets fixes pour les référents non-verbalisés chez les adultes francophones**



Nous venons de voir que chez les enfants, il y a un taux non négligeable de référents non-verbalisés dans des contextes de rupture, topicale et/ou interactionnelle. Ces cas existent aussi chez les adultes, mais sont plus rares globalement. Notamment les changements de topic sont plus rarement effectués sans verbalisation que ce n'était le cas chez les enfants, et la plupart des non-verbalisations concernent le topic, dans des contextes autres qu'un mouvement de RETOUR.

## 2.6 Résumé et discussion

Nous avons donc pu constater jusqu'ici qu'en français, les différents facteurs que nous avons étudiés au cours des chapitres précédents interagissent de manière complexe et permettent de modéliser et expliquer les contextes dans lesquels une expression référentielle donnée sera favorisée ou non. Pour les dislocations, outre le fait qu'elles sont (quasiment) toujours au statut de topic, l'association à un des schèmes lexico-syntaxiques (autour des verbes *être* et *mettre*, pour nommer et localiser ; notamment *être X*, *être où* et *mettre où/là*) semble être le principal facteur qui favorise leur emploi. Toutes les activités/jeux sur lesquels est basé notre corpus permettent ce type d'activités langagières, et certaines, comme les puzzles ou les lectures d'album, les favorisent sans doute particulièrement. Dans d'autres types d'interactions routinisé, il nous semble possible que ce lien soit moins fort. Nous pensons par exemple à des routines ou activités qui seraient basé davantage sur des actions successives : s'habiller, faire un gâteau, des séquences davantage narratives dans le jeu symbolique. Dans nos données, dans le jeu avec la maison de poupées, qui se prête au jeu symbolique narratif, il y a toujours une première partie où les participants installent les accessoires dans la maison et les

nomment. Cela est possiblement le cas parce que les participants découvrent le matériel de jeu pour la première fois. Toutefois, les schèmes lexico-syntaxiques trouvés correspondent en grande partie à ceux décrits par Jourdain & Canut (2018) dans le suivi longitudinal de deux enfants entre 1 et 2 ans (voir notre compte rendu dans le CHAPITRE VIII-1.5). Parisse (2008) n'a pas trouvé de relation privilégiée entre dislocation et des verbes spécifiques, mais dans son étude, seulement les expressions nominales ont été considérées, et comparées pour les seuls verbes, non pas pour des schèmes plus spécifiques autour de ces verbes. Les sujets lexicaux disloqués et non disloqués y apparaissent très fréquemment aussi avec les verbes *aller* et *être* (2008 : 24), mais nous ne savons donc pas si pour des schèmes plus concrets les dislocations étaient plus fréquentes que les sujets non disloqués. Nous savons en revanche que les actes de nommer et localiser font partie d'un grand nombre de jeux et autres routines dans lesquels sont régulièrement engagés les jeunes enfants, et que les demandes de localisation et de catégorisation d'un objet font partie des premières questions adressées aux jeunes enfants (Kraft, 2007; Ninio & Snow, 1996), et il nous semble intéressant de vérifier cette association sur des corpus plus grands et plus spécifiquement contrôlés pour les activités et genres discursifs déployés.

Nous avons vu en outre, pour les dislocations et les noms en usage référentiel, que la discontinuité topicale ainsi que la présence d'un mouvement interactionnel de RETOUR étaient tous les deux parmi les contextes favorisants. La relation contraire est observée pour les pronoms personnels clitiques et référents non-verbalisés, ce qui confirme nos observations sur les formes fortes et formes faibles dans les chapitres précédents. Cela nous a permis de montrer que les facteurs de la dimension informationnelle, i.e. de la gestion des topics, et le facteur lié à la gestion de l'interaction jouent tous les deux un rôle, qui semble être complémentaire et convergeant en ce qui concerne les formes fortes et faibles : l'emploi d'une expression référentielle peut servir à la fois la gestion du topic et de l'interaction. Nos résultats pointent également vers une certaine convergence de la gestion du topic et du statut attentionnel et discursif : les dislocations et noms sont favorisés en (ré)introduction d'un référent, alors que les pronoms personnels et référents non-verbalisés sont favorisés pour des référents déjà donnés dans le discours. Cela n'est pas vrai cependant des pronoms clitiques chez les enfants, où le statut attentionnel n'apparaît pas comme facteur pertinent. D'une part, les pronoms personnels sont peut-être trop peu fréquents à ces âges pour en tirer des conclusions très robustes, d'autre part, nous avons constaté, comme pour d'autres expressions, que le facteur du statut de topic semble souvent supplanter d'autres facteurs : lorsque nous avons calculé un modèle sans le statut de topic, le statut Donné figurait aussi comme facteur favorisant dans le modèle des enfants. Nos résultats confirment les considérations théoriques du CHAPITRE II, que topicalité, fonction sujet et information donnée ont tendance d'aller de pair, et il est alors difficile de savoir auquel des facteurs les locuteurs sont le plus sensibles, s'ils sont de fait perçus comme des facteurs isolés ou plutôt justement comme un faisceau de facteurs ayant tendance à converger. Les modèles construits semblent toutefois indiquer qu'en fonction de l'expression considérée, les différents groupes de

## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

---

facteurs n'ont pas le même poids : le statut du topic apparaît comme crucial pour les dislocations (au point que nous avons dû l'exclure pour éviter l'auto-corrélation du modèle), les référents non-verbalisés et les pronoms clitiques pour les adultes, et le statut de non-topic pour les noms. Le statut attentionnel apparaît, dans tous nos modèles, comme un facteur secondaire ou ne figure pas du tout, alors que dans l'étude de da Silva et al. (soumis), qui ne prennent pas en compte la dimension du topic, le rôle du statut attentionnel et la fonction syntaxique apparaissent comme plus important comparé à nos modèles.

### 3. Résultats pour le corpus allemand

Passons maintenant à la présentation des modèles statistiques établis pour notre corpus allemand. Nous nous sommes intéressée ici aux D-Pro, aux noms et aux référents non-verbalisés, ainsi qu'aux expressions réalisées avec une proéminence prosodique ou accentuation.

#### 3.1 Les D-Pro

Nous avons constaté dans le CHAPITRE IX-2.1 que seulement une petite part des D-Pro n'encode pas le topic (21 occurrences, 11% chez les enfants, 12 occurrences, 2,5% chez les adultes). Nous savons donc que le statut du topic est un contexte qui favorise fortement les D-Pro, et chez les adultes, ils constituent la catégorie la plus fréquente dans le topic. Pour les arbres de partition, cela implique pour les adultes que le poids des autres facteurs est écrasé, et ceux-ci ne figurent pas dans l'arbre (même si, dans le modèle, leur effet est significatif, voir nos remarques plus haut sur cette question). Comme pour les dislocations en français, nous avons alors calculé les modèles avec et sans topic. Pour les enfants, le lien entre topic et D-Pro n'est pas aussi fort que chez les adultes, aussi, le modèle correspondait moins bien aux données en enlevant ce facteur ( $C=0.74$ ), nous l'avons donc gardé.

### 3.1.1 Les enfants

Nous observons que pour les enfants, le statut de topic, le changement d'un topic à l'autre ainsi que la présence d'un mouvement de RETOUR sont les contextes qui favorisent l'emploi d'un D-Pro (alors que nous verrons plus bas que ce n'est pas le cas pour les adultes).

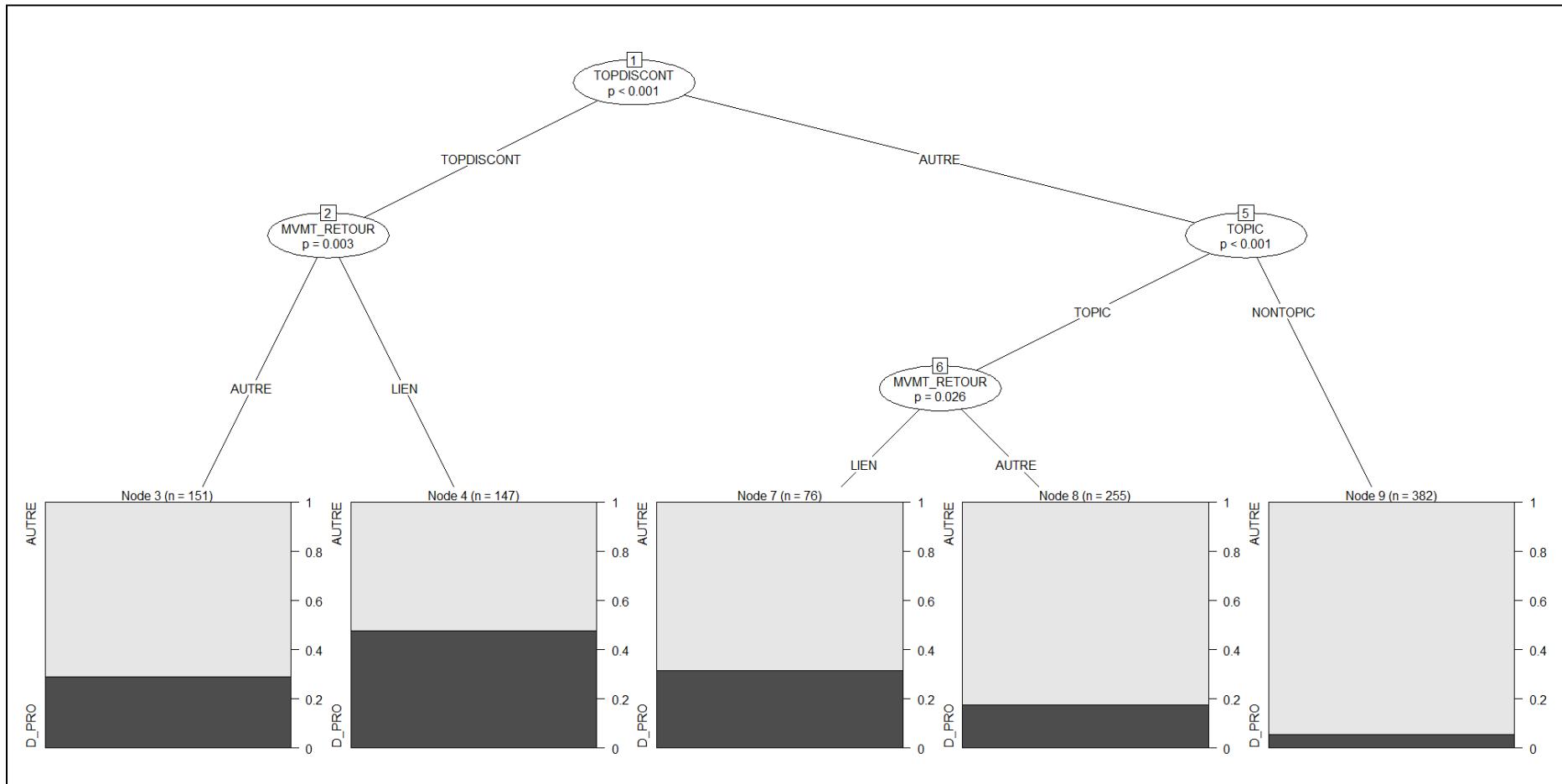
**Tableau XI-13 - Résultats du modèle linéaire mixte pour les D-Pro (enfants, allemand)**

D-Pro, C=0.7727999				
Effets aléatoires :	Variance	SD		
Session	0.07046	0.2655		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-3.1979	0.2562	-12.483	<.00001 ***
TOPIC	1.6504	0.2679	6.161	<.00001 ***
TOPDISCONT	0.6519	0.1890	3.448	0.000564 ***
MVMT_RETUR	0.7425	0.1763	4.211	0.000660 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1011

Dans ce modèle, ni la fonction syntaxique, le statut attentionnel ni le groupe MLU n'apportaient de pouvoir explicatif supplémentaire. Par rapport à nos observations des chapitres précédents, le fait que la fonction syntaxique ne montre pas d'effet peut s'expliquer par le fait que d'une part, les D-Pro n'apparaissent pas seulement en fonction sujet. D'autre part, les référents non-verbalisés en appui syntaxique, correspondant au sujet ou au THEM<sub>Z</sub>, sont particulièrement fréquents chez les enfants, et en comparaison, le taux des D-Pro en fonction sujet est moins important que cela n'est le cas chez les adultes, comme nous le verrons ci-dessous. Il reste bien sur possible qu'avec plus de données, nous observons un effet de la fonction syntaxique (l'absence d'une évidence n'est pas l'évidence d'une absence), mais en tout cas ce résultat ne nous paraît pas étonnant au vu de nos données. La même observation vaut pour le statut attentionnel, qui ne figure pas dans le modèle final (aucun des statuts attentionnels n'avait un effet significatif, et leur suppression du modèle ne changeait pas le « goodness of fit »). En effet, nous avons vu dans le CHAPITRE VII-5 que les D-Pro sont employés pour des référents donnés dans le discours, activés, ou réintroduits. Ils sont moins fréquents en comparaison dans les référents nouveaux, mais ceux-ci sont généralement peu fréquents, et exprimés essentiellement par des noms (ou des pronoms interrogatifs pour les adultes). Dans le contexte des référents nouveaux, les D-Pro se fondent donc dans la masse des autres expressions généralement peu fréquentes pour introduire un référent.

**Arbre de partition 11 – Distribution et interaction des effets fixes pour les D-Pro chez les enfants germanophones**



Considérons à présent l'arbre de partition pour observer l'interaction entre les effets retenus dans le modèle. Les D-Pro sont majoritairement employés pour exprimer un changement de topic, et aussi fréquemment pour signaler en plus un mouvement de RETOUR sur un tour ou énoncé précédent. De fait, dans nos données, cela correspond essentiellement à des cas de topics contrastées dans des activités d'énumération, pour les différentes parties du Monsieur Patate à assembler ou les différentes figurines dans la maison de poupées par exemple. Le reste des occurrences concerne des topics dans des contextes non-disruptifs. Cela correspond essentiellement au maintien d'un topic (nœuds 6, 7 et 8), avec ou sans lien RETOUR. Enfin, une petite partie des D-Pro ne correspondent pas au topic de l'énoncé : ce sont surtout des cas de focus étroit, élicités par une question :

#### Exemple XI-6 - [GER] Nadja/2;05.27/MLU2/Puzzle

MER18	was nimmst du jetzt für ein Tier ?	qu'est ce que tu prends maintenant comme animal ?
ENF18	und +...	et...
MER19	hm@i ?	hm ?
ENF19	<b>das !</b> {saisit une pièce}	<b>ça !</b>

Même si l'interaction entre les différents facteurs s'organise de manière différente de ce que nous avons observé pour les dislocations chez les enfants francophones, nous pouvons tout de même conclure que produire un énoncé qui présente un lien de RETOUR vers une action précédente et/ou changer de topic (par (ré-)introduction) font partie des facteurs qui favorisent l'emploi d'une dislocation en français, et d'un D-Pro en allemand, en ce qui concerne les enfants. Nous verrons alors que pour les adultes, les résultats sont moins contrastés. Nous pensons que cela correspond bien à la caractéristique observée des D-Pro de servir aussi bien pour les continuités que pour les ruptures.

### 3.1.2 Les adultes

Les facteurs qui sont apparus comme les plus pertinents pour expliquer l'emploi d'un D-Pro chez les adultes germanophones sont donnés dans le Tableau XI-14 ci-dessous. Nous l'avons déjà dit plus haut, nous avons toujours le problème que l'emploi d'un D-Pro est très fortement corrélé avec le statut de topic. Cette relation n'est pas aussi systématique que pour les dislocations dans la référence aux entités cependant. Choisir le meilleur modèle pour les D-Pro au niveau statistique, et qui soit pertinent au niveau théorique, est particulièrement compliqué de ce fait. Si nous incluons le statut de topic dans ce modèle, la qualité d'ajustement (*goodness of fit*) du modèle aux données observées est meilleure, mais nous perdons en pouvoir explicatif pour les autres variables.

## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

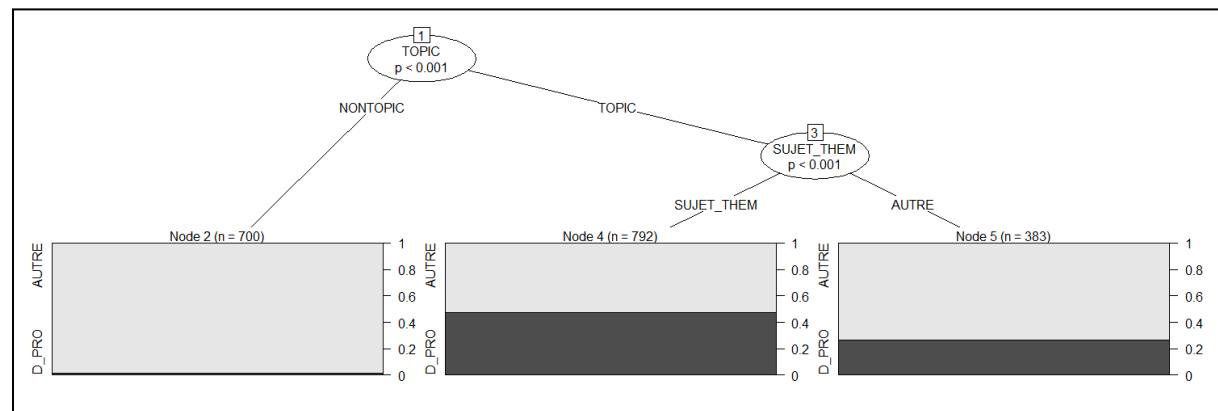
Tableau XI-14 - Résultats du modèle linéaire mixte pour les D-Pro (adultes, allemand)

D-Pro, C=0.8062661		Variance	SD	
Effets aléatoires :				
Session	0.05259	0.2293		
Effets fixes :	Estimate	SE	z	
(Intercept)	-4.5133	0.3129	-14.422	<.00001 ***
TOPIC	3.4111	0.3022	11.288	<.00001 ***
SUJET_THEME	0.9122	0.1352	6.746	<.00001 ***
Statut attentionnel (Activé)	0.3676	0.1518	2.422	0.000566 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1875

Ce modèle nous dit alors seulement que les D-Pro chez les adultes sont favorisés pour l'expression du topic, en fonction sujet et pour des référents activés. Les autres facteurs considérés ne sont pas significatifs lorsque le statut de topic est dans le modèle, et l'interaction avec le statut attentionnel semble marginale, puisqu'elle n'apparaît plus dans l'Arbre de partition 12 ci-dessous :

Arbre de partition 12 – Distribution et interaction des effets fixes pour les D-Pro chez les adultes germanophones



Nous avons alors testé de calculer aussi un modèle moins bien ajusté, mais qui permet de mieux comprendre l'interaction des autres facteurs (ou plutôt, visualiser l'absence d'un impact clair) lorsque nous ignorons le statut de topic, que nous savons en outre très fort. Le tableau ci-dessous présente les résultats pour le meilleur modèle obtenu sans l'effet du statut de topic :

Tableau XI-15 - Modèle linéaire mixte alternatif pour les D-Pro (adultes, allemand)

D-Pro, C=0.7675923		Variance	SD	
Effets aléatoires :				
Session	0.09501	0.3082		
Effets fixes :	Estimate	SE	z	
(Intercept)	-3.3572	0.2379	-14.111	<.00001 ***
SUJET_THEME	1.3972	0.1279	10.928	<.00001 ***
TOPDISCONT	1.1866	0.1318	9.002	<.00001 ***
Statut attentionnel (Donné)	1.3610	0.1992	6.833	<.00001 ***
Statut attentionnel (Activé)	0.7102	0.2164	3.281	0.00103 **

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1875

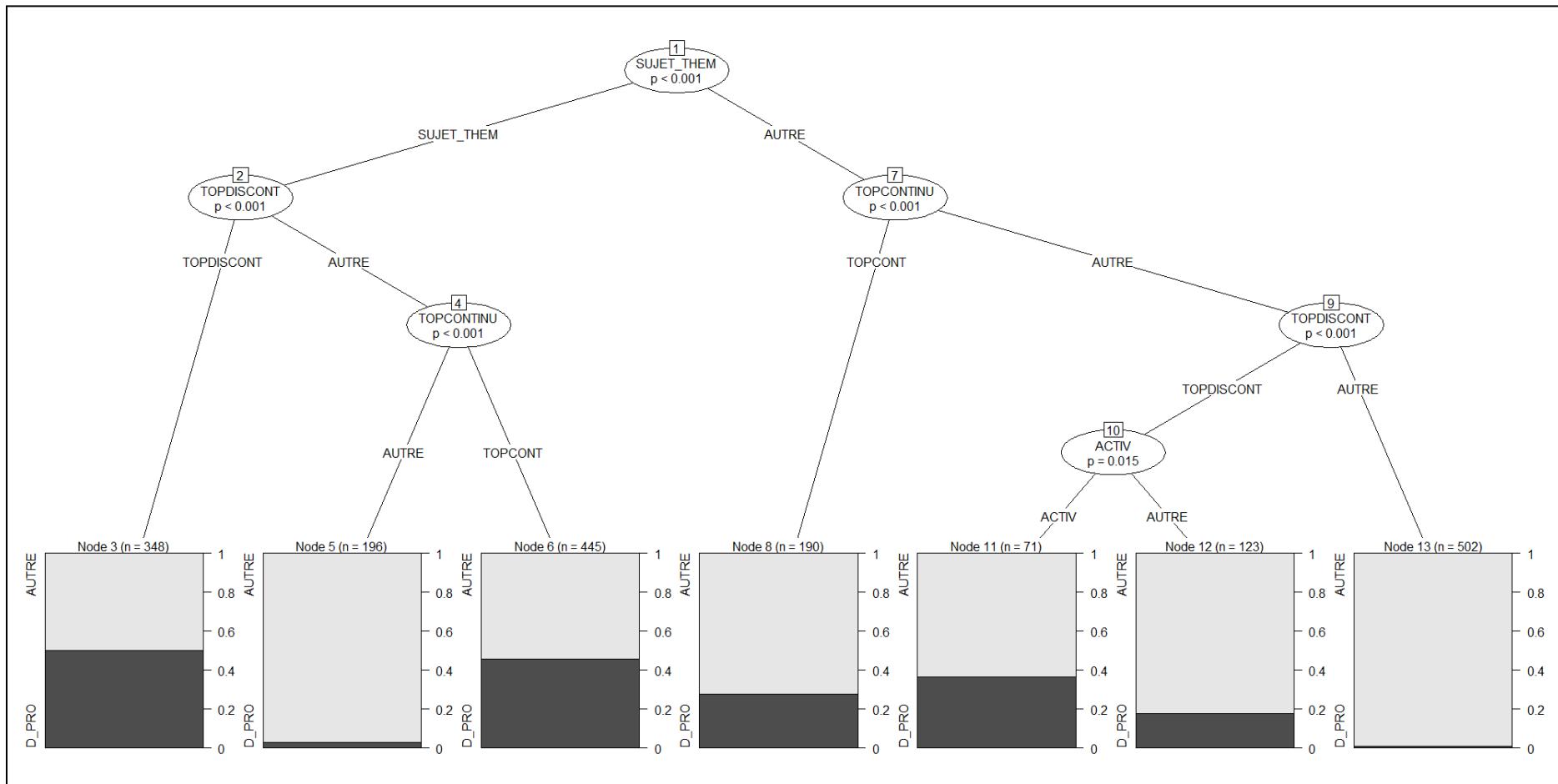
Nous avons observé alors un effet positif de la fonction syntaxique sujet, de la discontinuité topicale ainsi que pour le statut attentionnel : l'emploi d'un D-Pro est favorisé en fonction sujet (ou THEMZ) et pour changer de topic, et pour des référents donnés dans le discours ou activés. Le facteur interactionnel n'apparaît pas comme significatif, et il ne l'est pas non plus lorsque nous l'avons

considéré comme seul et unique facteur. Nous ne pouvons donc pas conclure que pour les adultes, il joue un rôle dans le choix d'un D-Pro. Cette différence avec le modèle obtenu pour les enfants nous semble être liée en grande partie au fait que dans notre corpus allemand, ce sont davantage les enfants qui s'engagent dans des activités d'énumération, alors que les adultes ne le font pas (ou beaucoup moins), et cela pourrait relever essentiellement des contingences des dyades et situations observées. Cela expliquerait aussi les différences dans les modèles statistiques, alors que les valeurs moyennes d'un D-Pro pour un mouvement RETOUR ne sont pas très différentes pour les enfants et les adultes (autour de 50%, de fait). Alors que le lien avec la fonction sujet et les statuts « Donné » et « Activé » confirme les résultats du CHAPITRE VII, le lien avec les topics discontinus paraît surprenant, étant donné que nous n'avons pas détecté de différence pour la position des D-Pro dans la chaîne topicale (CHAPITRE IX-2.2). De fait, lorsque, comme pour les D-Pro et les dislocations, toutes les occurrences (ou presque) ont le statut de topic, cela veut dire qu'ils sont soit des topics continus, soit discontinus. Considérer dans le modèle et les topics discontinus et les topics continus reviendrait alors à mesurer deux fois la même chose et pose les problèmes de colinéarité déjà mentionnés. Nous avons donc testé un premier modèle avec le facteur « topic continu », et un deuxième avec le facteur « topic discontinu ». L'un ou l'autre des facteurs étaient significatifs dans le modèle, mais le modèle calculé avec TOPCONTINU était moins bien ajusté ( $C=0.74$ ).

Tous ces faits pointent vers l'observation déjà faite dans les chapitres précédents : les D-Pro sont une catégorie quelque peu passe-partout, liés surtout à l'expression du topic, mais au-delà aussi bien employés pour la continuité que pour la discontinuité. En ce sens, le premier modèle simplifié ci-dessus (Tableau XI-14) est peut-être le plus réaliste. L'arbre de partition correspondant aux facteurs retenus dans le second modèle nous permettra de mieux expliciter ce que cela signifie. L'arbre montre que la fonction syntaxique est maintenant le facteur le plus important, la fonction sujet ou thème favorisant l'emploi d'un D-Pro. Discontinuité et continuité topicale sont tous les deux des facteurs significatifs ensuite, reflétant le fait que les D-Pro sont majoritairement au statut de topic, et, lorsqu'ils le sont, correspondent soit à la discontinuité, soit à la continuité dans la chaîne topicale. Alors que pour les enfants, comme nous l'avons vu ci-dessus, les mouvements de RETOUR sur un énoncé précédent jouaient un rôle (positif) significatif, ce facteur n'a pas d'influence significative (ni positive, ni négative) pour les adultes.

Les D-Pro apparaissent donc clairement comme une catégorie hétérogène, employée aussi bien pour la continuité topicale que pour la discontinuité, et pas spécialement liée aux décrochages ou mouvements de RETOUR pour les adultes. Pour les enfants, l'influence favorisante d'un changement de topic et d'un décrochage interactionnel nous semblent tous les deux liés, nous l'avons déjà dit, aux activités d'énumération dans lesquelles s'engagent les enfants.

**Arbre de partition 13 – Calcul alternatif : Distribution et interaction bis des effets fixes pour les D-Pro chez les adultes germanophones**



## 3.2 Les noms

### 3.2.1 Les enfants

Considérons maintenant les contextes dans lesquels des noms sont produits par les enfants germanophones. Le Tableau XI-16 ci-dessous présente le meilleur modèle obtenu :

**Tableau XI-16 - Résultats du modèle linéaire mixte pour les noms (enfants, allemand)**

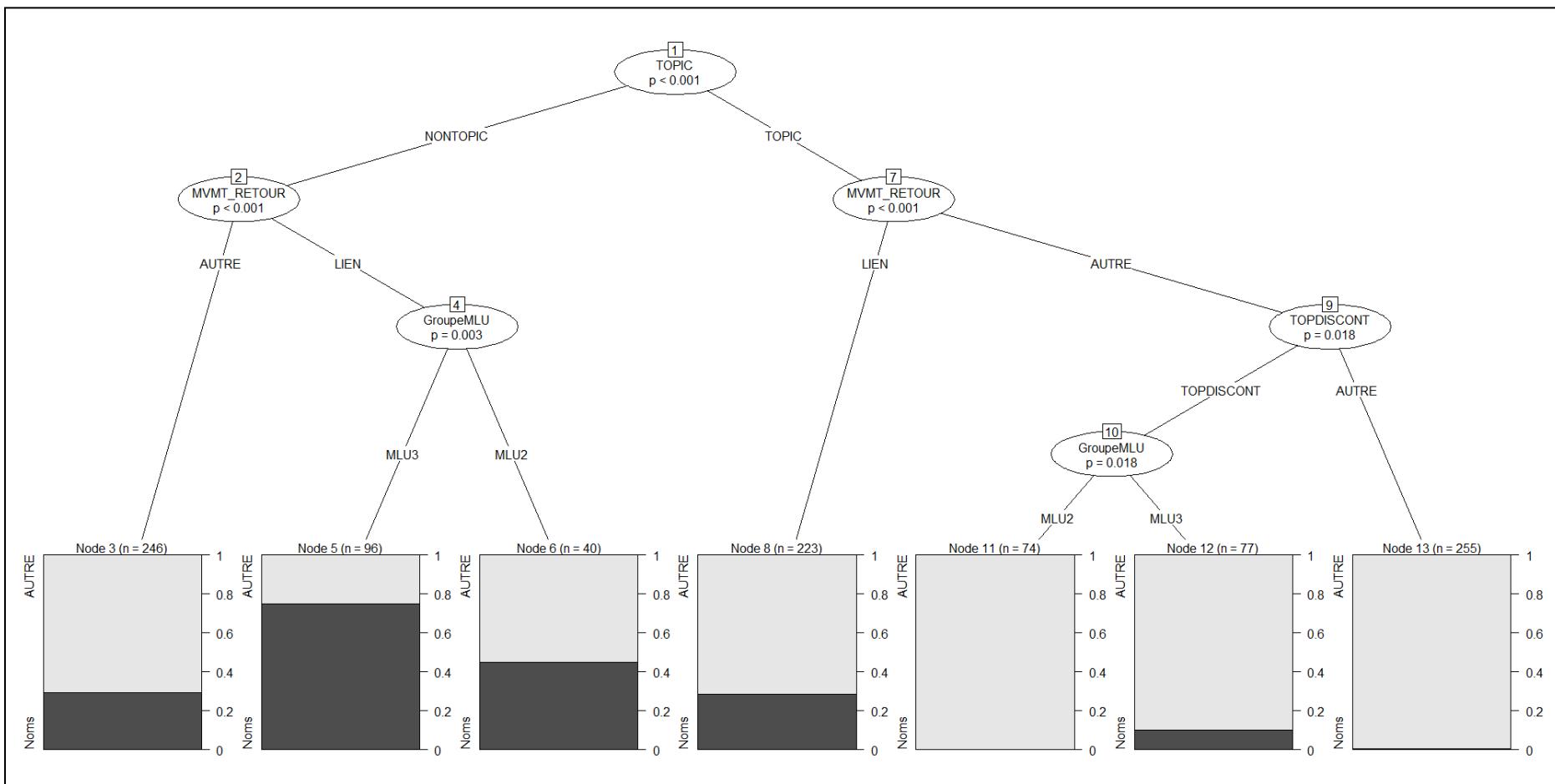
Noms, C=0.8428376				
Effets aléatoires :	Variance	SD		
Session	0.1783	0.4222		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-1.6950	0.2769	-6.121	<.00001 ***
TOPIC	-2.3251	0.2673	-8.699	<.00001 ***
TOPDISCONT	0.6117	0.2886	2.119	0.03406 *
MVMT_RETUR	1.9147	0.1934	9.900	<.00001 ***
Groupe MLU (MLU3)	0.8710	0.3237	2.691	0.00713 **

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1011

Conformément à nos observations dans les chapitres précédents, les noms ne sont pas favorisés comme expression du topic en général. Nous observons en revanche un effet positif des contextes de changement de topic, confirmant là aussi les constats précédents : lorsqu'un nom encode le topic, cela correspond plus fréquemment à une introduction ou réactivation d'un topic qu'au maintien du topic précédent. Nous constatons également un effet du mouvement interactionnel de RETOUR, résultat attendu aussi vu les observations faites dans le CHAPITRE X. L'effet positif du groupe MLU plus avancé correspond à la tendance générale de nos enfants de produire moins de référents implicite et plus de noms en groupe MLU 3, mais, contrairement aux enfants francophones, cette différence est significative ici. S'il s'agit d'un effet d'échantillon ou si cette différence entre les enfants francophones et germanophones correspond au fait qu'en français, il y a en plus la possibilité d'employer une expression nominale disloquée est une question à laquelle nous ne pourrons répondre ici, mais cela nous semble une piste intéressante pour des futures recherches avec un corpus plus important.

L'interaction entre les facteurs retenus telle que représentée dans l'arbre de partition ci-dessous indique que le facteur principal pour le choix d'un nom est, comme dans le corpus français, le statut de Non-Topic. Les noms référentiels sont ensuite favorisés par un mouvement de RETOUR (reformulation, confirmation, demande de confirmation, énumération), que l'expression corresponde au topic ou non. La discontinuité topicale et le groupe MLU apparaissent comme facteurs secondaires.

**Arbre de partition 14 – Distribution et interaction des effets fixes pour les noms chez les enfants germanophones**



### 3.2.2 Les adultes

Le tableau ci-dessous donne le modèle le mieux ajusté pour les noms chez les adultes germanophones :

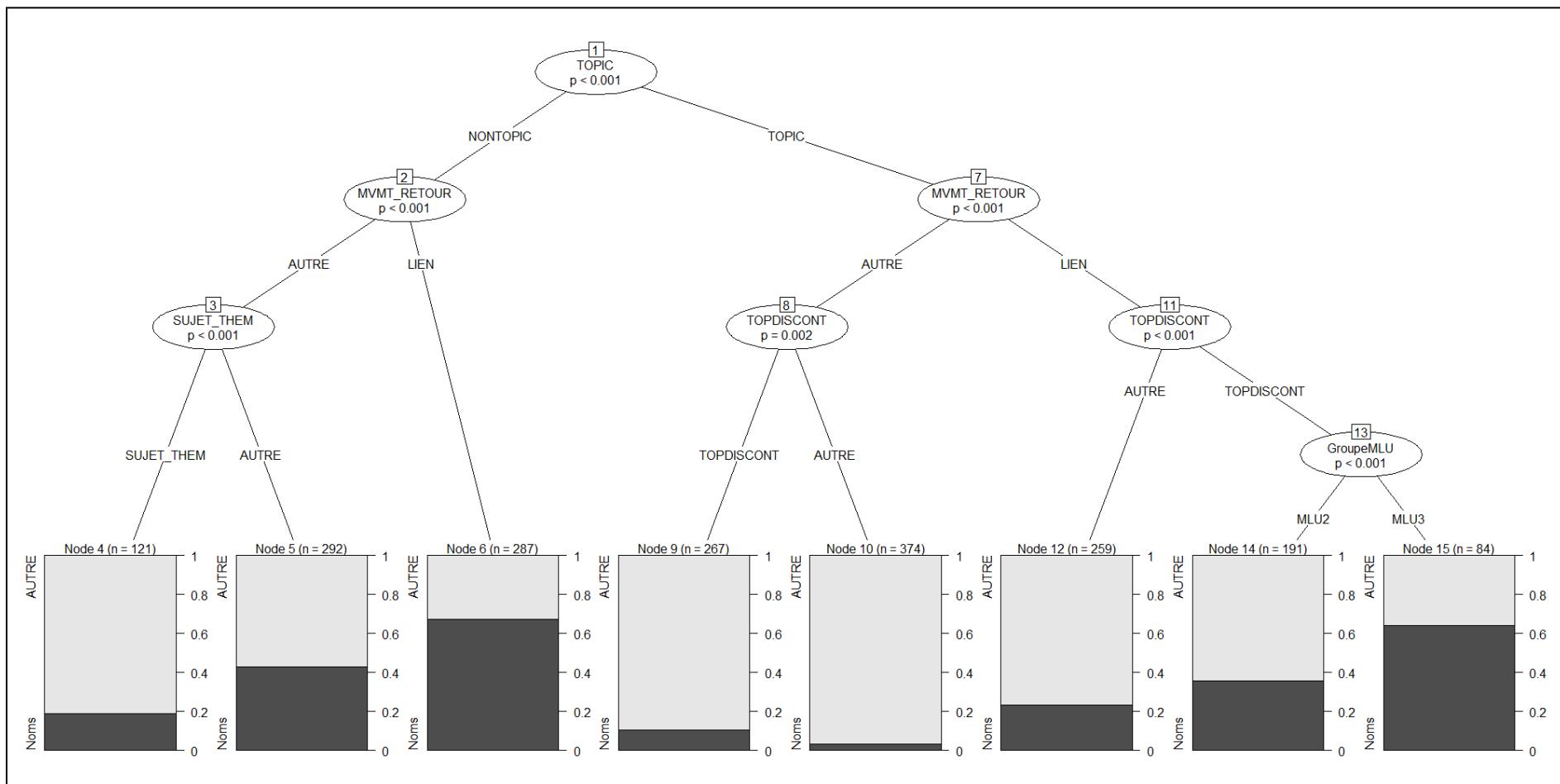
**Tableau XI-17 - Résultats du modèle linéaire mixte pour les noms (adultes, allemand)**

Noms, C=0.8119678				
Effets aléatoires :	Variance	SD		
Session	0.0814	0.2853		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-1.0189	0.1728	-5.896	<.00001 ***
TOPIC	-1.9176	0.1684	-11.387	<.00001 ***
SUJET_THEME	-0.5277	0.1298	-4.064	<.00001 ***
TOPDISCONT	0.8401	0.1727	4.865	<.00001 ***
MVMT_RETUR	1.7639	0.1267	13.927	<.00001 ***
Statut attentionnel (Réintroduit)	0.6705	0.1812	3.701	0.000215 ***
Groupe MLU (MLU3)	0.6196	0.2072	2.990	0.002785 **

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1875

Comparé au modèle retenu pour les enfants, plus de facteurs semblent jouer un rôle chez les adultes pour le choix d'un nom. Comme pour les enfants, le statut de topic ne favorise pas l'emploi d'un nom, mais les contextes de discontinuité topicale et d'un mouvement de RETOUR le font. En plus, chez les adultes les noms sont significativement moins probables en fonction sujet, et favorisés pour la réintroduction d'un référent. Enfin, nous constatons que les mères des enfants du groupe MLU 3 produisent aussi plus de noms que celles dont les enfants ont un MLU plus bas. Cette observation pourrait alors indiquer soit une influence du développement linguistique, soit être un effet de l'échantillon, comme indiqué plus haut. De toute façon, comme le montre l'arbre de partition ci-dessous, la différence entre les groupes MLU n'est pas parmi les facteurs les plus importants. Comme pour les enfants, le statut de topic apparaît comme principal facteur, suivi du décrochage interactionnel et du changement de topic, dont la présence simultanée concerne une grande partie des occurrences. Lorsqu'un nom ne correspond pas à un topic est n'est pas non plus concerné par un mouvement de RETOUR, c'est la fonction syntaxique qui prédit la probabilité d'avoir un nom plutôt qu'une autre expression.

Arbre de partition 15 – Distribution et interaction des effets fixes pour les noms chez les adultes germanophones



### 3.3 Les référents non-verbalisés

Enfin, nous allons examiner les modèles pertinents pour les référents non-verbalisés dans le corpus allemand.

#### 3.3.1 Les enfants

Pour les enfants, nous obtenons un modèle qui est quelque peu plus complexe que les modèles pour les référents non-verbalisés dans le corpus français.

**Tableau XI-18 - Résultats du modèle linéaire mixte pour les référents non-verbalisés (enfants, allemand)**

Non-verbalisés, C=0.9115095		Variance	SD	z	p
Effets aléatoires :	Estimate				
Session	0.07044	0.2654			
Effets fixes :					
(Intercept)	-3.0324	0.3974	-7.631	<.00001 ***	
TOPIC	4.0977	0.4104	9.986	<.00001 ***	
SUJET_THEME	0.9202	0.2447	3.760	0.00017 ***	
TOPDISCONT	-1.0624	0.1996	-5.324	<.00001 ***	
MVMT_RETUR	-2.1585	0.2203	-9.797	<.00001 ***	
Groupe MLU (MLU3)	-1.0175	0.2474	-4.112	<.00001 ***	

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1011

Nous constatons premièrement qu'en allemand, contrairement aux modèles pour le français, la fonction syntaxique sujet ou THEMZ joue un rôle significatif positif. Cela est un résultat attendu vu la différence entre les langues discutée dans les chapitres précédents, l'allemand permettant plus facilement la non-verbalisation du sujet d'une part, et du fait que les vraies formes Zero correspondent soit au sujet, soit à l'objet en allemand, alors qu'en français, elles sont essentiellement des compléments du verbe, non sujet. Nous avons discuté aussi le fait que dans notre corpus allemand, nous observons plus d'énoncés averbaux à sujet (ou THEMZ) non verbalisé qu'en français, où des pronoms clitiques apparaissent plus facilement (i.e., par exemple, à un *c'est beau* en français correspond plus (ou du moins aussi) facilement un *schön* en allemand qu'un *das ist schön*).

Le statut de topic est également un contexte favorisant, alors que la discontinuité topicale et des mouvements de RETOUR ne favorisent au contraire pas la non-verbalisation, comme attendu vu les résultats et observations des chapitres précédents. Enfin, comme pour les noms, nous observons un effet du MLU, cette fois ci dans l'autre sens : alors que l'emploi d'un nom est favorisé en groupe MLU3, la non-verbalisation y est moins probable qu'en groupe MLU2. Cela veut dire que dans les deux langues, nos enfants verbalisent plus souvent les référents en groupe MLU3, mais d'une part la distribution des formes produites en MLU3 est plus variée chez nos enfants francophones que ce n'est le cas chez nos enfants germanophones. Ces derniers ne produisent presque pas de fillers, les pronoms personnels ne jouent pas un grand rôle dans la référence aux entités, ni les dislocations, et ce sont donc plutôt les D-Pro et les noms qui sont plus fréquents en MLU3 globalement. D'autre part, cela veut dire que l'écart entre les groupes MLU semble être plus marqué dans notre corpus

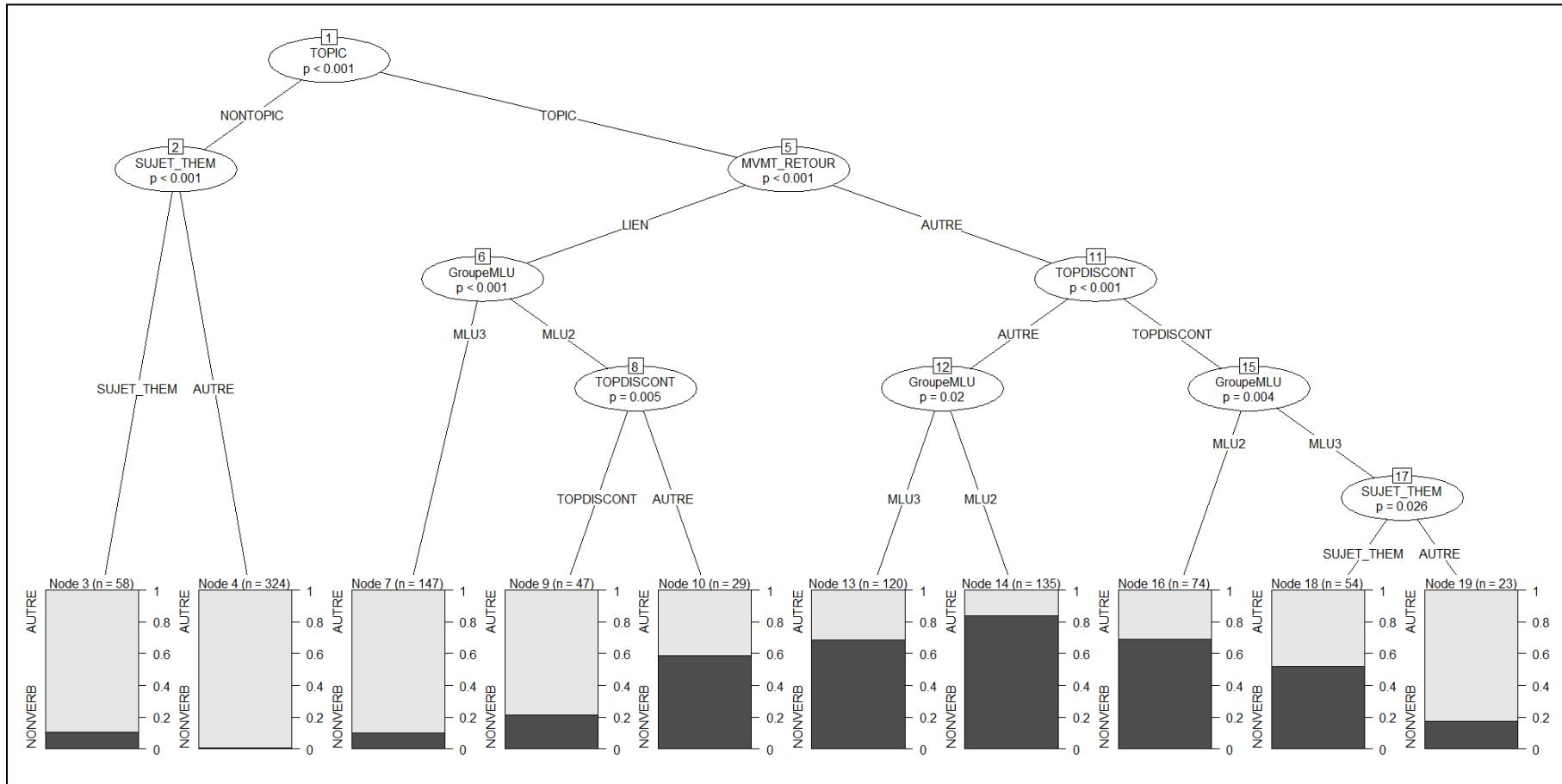
## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

---

allemand que dans notre corpus français, et il pourrait bien s'agir là d'un effet d'échantillon. Si nous prenons donc les différences entre les groupes MLU avec une bonne dose de précaution, il nous semble qu'au contraire la préférence pour les non-verbalisations sujet est un effet pertinent et différent entre les deux langues.

Considérons à présent l'arbre de partition ci-dessous. Cet arbre montre que les référents non-verbalisés sont effectivement de préférence produits dans des contextes où le référent à le statut de topic, et n'est pas concerné par un mouvement interactionnel de RETOUR ni par un changement de topic. A plusieurs niveaux, le facteur MLU 2 apparaît comme favorisant la non-verbalisation, de même que la fonction syntaxique sujet ou THEMz. Il y a toutefois, comme pour les enfants francophones, un certain nombre de cas où la non-verbalisation d'un référent correspond à un contexte de RETOUR, d'un changement de topic, voire les deux simultanément. Nous verrons ci-dessous que ces cas sont moins fréquents chez les adultes, comme dans le corpus français. Toutefois, l'écart entre enfants et adultes et moins flagrant dans le corpus allemand, nous y reviendrons lorsque nous discuterons le modèle pour les adultes.

## Arbre de partition 16 – Distribution et interaction des effets fixes pour les référents non-verbalisés chez les enfants germanophones



## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

---

Pour revenir aux enfants, comme en français, un bon nombre de ces cas inattendus correspondent à des énoncés averbaux où un syntagme nominal est répété pour confirmation ou demande de confirmation. Au niveau de la structure topic-commentaire, nous avons donc un topic non-verbalisé qui est impliqué dans un retour sur un référent, alors que dans d'autres cas, un D-Pro ou un nom est employé. Ce qui explique alors ces décrochages avec non-verbalisation est l'emploi justement de l'expression nominale, qui reprend l'énoncé précédent en mention :

### Exemple XI-7 - [GER] Lili/2;05.12/MLU2/Maison Poupees

MER19	was ist das ? {regarde CHI puis toilettes}	qu'est ce que c'est (ça) ?
ENF20	weiss i(ch) nich(t) .	je sais pas.
MER20	eine toilette . {pointe}	des toilettes.
ENF21	<b>eine toilette .</b>	<b>des toilettes.</b>

Mais les enfants vont aussi fréquemment introduire un nouveau référent comme topic sans le verbaliser lorsqu'il est sous le focus de leur attention, qu'ils sont en train de manipuler et/ou lorsqu'il fait partie de la routine en cours. Dans l'activité du puzzle, Nadja saisit une nouvelle pièce de puzzle et la place sans verbaliser le référent (ENF118). Comme dans l'exemple de Clément discuté en section 2.5 ci-dessus, cette conduite est suivie par une demande de dénomination justement de la part de la mère :

### Exemple XI-8 - [GER] Nadja/2;05.27/MLU2/Puzzle

ENF118	<b>glaub ich da.</b> {prend dernière pièce et la place}	<b>je crois là.</b>
MER119	un(d) was is(t) das ?	et qu'est-ce que c'est ?
ENF119	0. {regarde caméra}	
MER120	was is(t) das für ein Tier ? {se penche pour regarder CHI et pointe la pièce}	qu'est-ce que c'est comme animal ?
ENF120	ein Pferd. {se penche sur la pièce et la touche}	un cheval.

Comme les enfants francophones, les enfants germanophones s'appuient donc aussi beaucoup sur le contexte physique et ne verbalisent pas nécessairement tous les changements au niveau référentiel ou topical. Au niveau interactionnel, un bon nombre de décrochages non-verbalisés s'appuie donc aussi sur le contexte physique (énumération), et la plupart sont de fait le topic non-verbalisé d'un énoncé averbal à prédicat nominal, comme dans l'Exemple XI-7 ci-dessus. Enfin, il faut également prendre en compte le fait discuté dans le CHAPITRE IX que les formes Zero correspondant au topic-drop ne dépendent pas tant d'un contexte de continuité topique ni même référentielle ; la haute saillance du référent (par exemple l'attention conjointe sur un objet ou une action) semble être un critère suffisant.

### 3.3.2 Les adultes

Pour les référents non-verbalisés chez les adultes, nous obtenons un modèle avec une relation positive pour le statut de topic et la fonction syntaxique de sujet ou THEMz, et une relation négative avec les topics discontinus et les mouvements de RETOUR, comme pour les enfants. Pour les adultes, en revanche, le groupe MLU de l'enfant n'était pas un facteur pertinent :

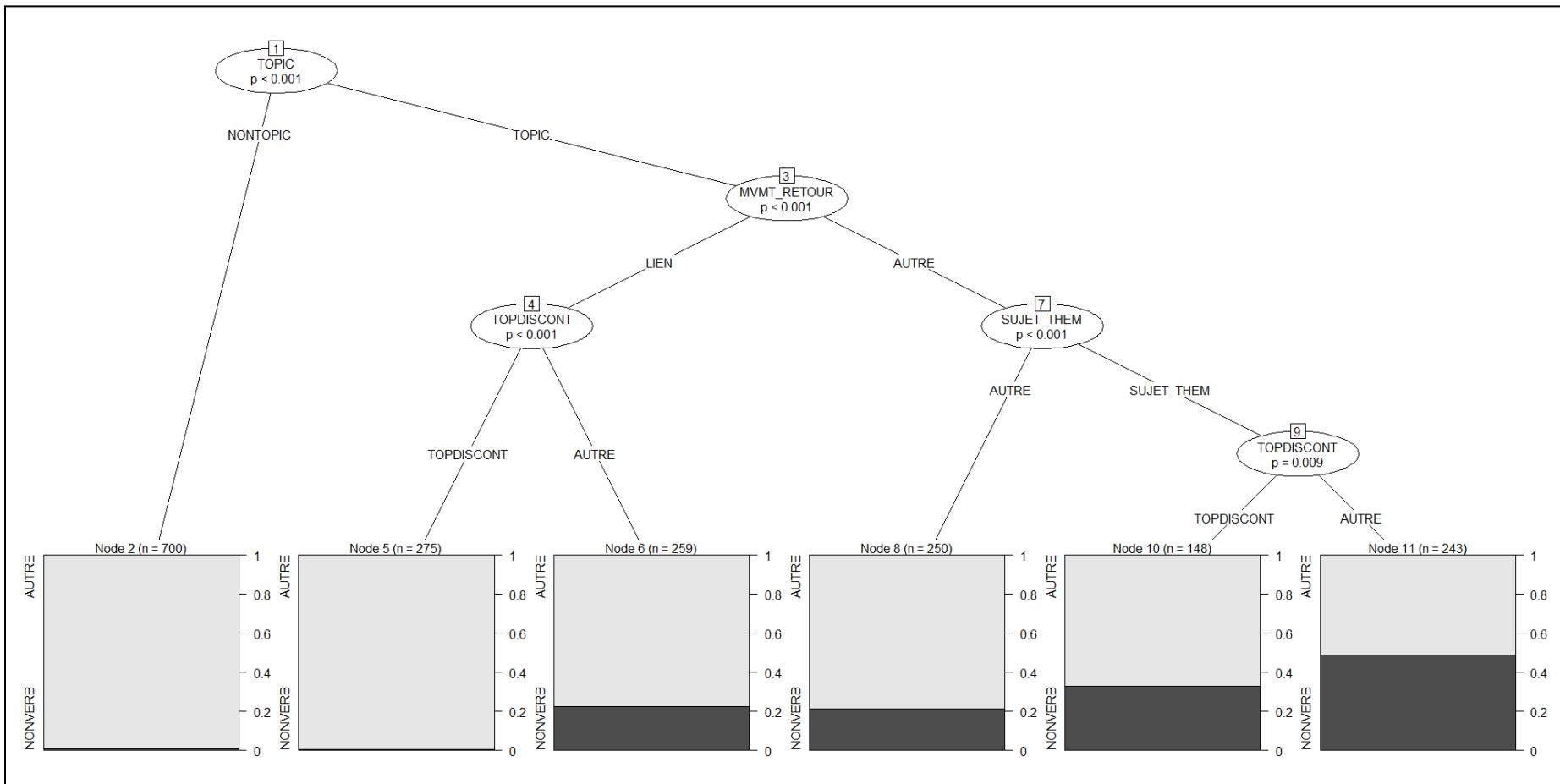
**Tableau XI-19 - Résultats du modèle linéaire mixte pour les référents non-verbalisés (adultes, allemand)**

Non-verbalisés, C=0.8504286				
Effets aléatoires :	Variance	SD		
Session	0.07133	0.2671		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-4.5203	0.3795	-11.912	<.00001 ***
TOPIC	3.6440	0.3716	9.806	<.00001 ***
SUJET_THEME	0.9114	0.1687	5.402	<.00001 ***
TOPDISCONT	-1.0768	0.1597	-6.744	<.00001 ***
MVMT_RETUR	-1.4583	0.1655	-8.811	<.00001 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1875

L'arbre de partition ci-dessous montre que, comparé aux enfants, le taux des référents non-verbalisés dans des contextes de rupture (changement de topic ou mouvement de décrochage) est plus faible, même si cette différence n'est pas aussi nette qu'entre enfants et adultes francophones (section 2.5 ci-dessus). Les adultes marquent donc les changements ou décrochages, au niveau informationnel et au niveau interactionnel, plus souvent de manière explicite que ne le font les enfants.

**Arbre de partition 17 – Distribution et interaction des effets fixes pour les référents non-verbalisés chez les adultes germanophones**



### 3.4 Les proéminences prosodiques

Enfin, nous avons voulu tester l'impact des différents facteurs étudiés sur la réalisation prosodique d'une expression référentielle. Rappelons que l'annotation prosodique du corpus a seulement été réalisée pour les expressions au statut de topic, et uniquement pour le corpus allemand. Nous ne nous intéressons ici donc non pas à la proéminence prosodique en tant que telle, qui peut aussi marquer le focus d'un énoncé, mais bien en tant que marque possible d'un changement du statut topical, conformément à nos hypothèses de départ. Rappelons également que la proéminence prosodique a été annotée seulement pour les énoncés comprenant plusieurs éléments linguistiques. L'échantillon étudié dans cette dernière section est donc plus petit que celui des sections précédentes.

#### 3.4.1 Les enfants

Le modèle obtenu pour les enfants, présenté ci-dessous, confirme les observations du CHAPITRE IX-3.4 sur le lien entre prosodie et topicalité :

**Tableau XI-20 - Résultats du modèle linéaire mixte pour les accentuations ou proéminences prosodiques (enfants, allemand)**

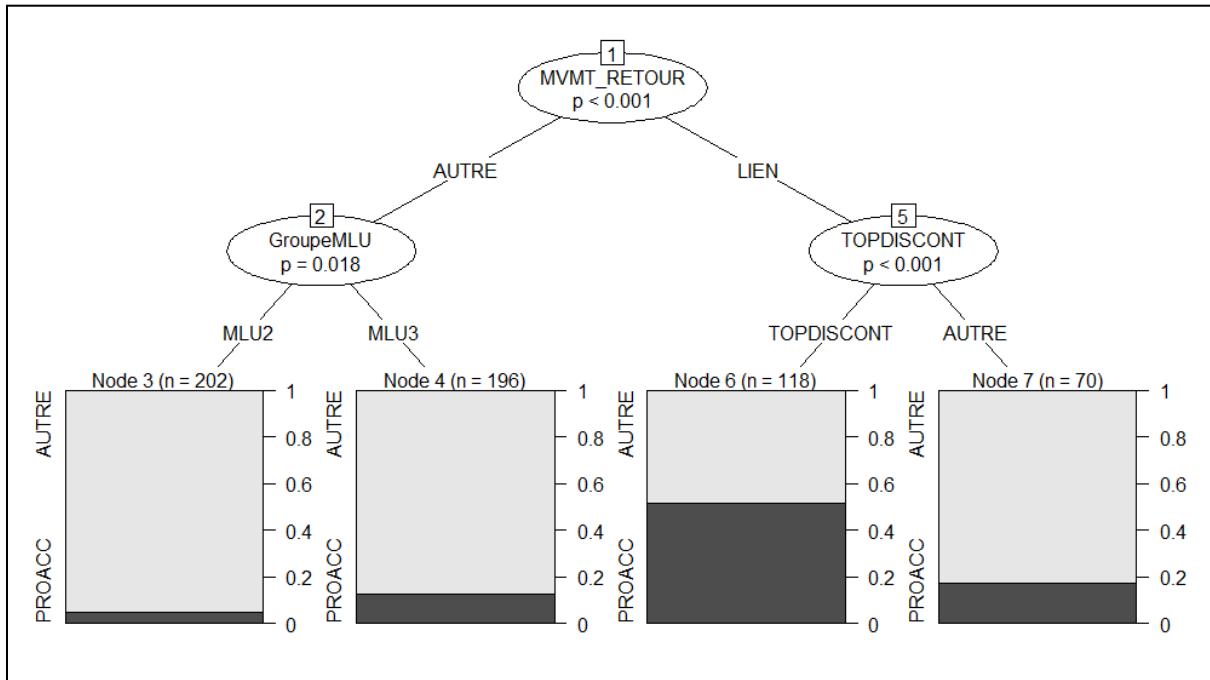
Proéminences prosodiques, C=0.8058171				
Effets aléatoires :	Variance	SD		
Session	0.04375	0.2092		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-3.5179	0.3192	-11.020	<.00001 ***
TOPDISCONT	1.1959	0.2566	4.660	<.00001 ***
MVMT_RETUR	1.5733	0.2443	6.441	<.00001 ***
Groupe MLU (MLU3)	0.9525	0.2931	3.249	0.00116 **

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 586

La proéminence prosodique ne marque pas le topic en tant que tel, qui peut aussi être non-accentué, mais semble effectivement liée aux changements de topic. Nous constatons de plus que la présence d'un mouvement RETOUR favorise également l'accentuation, tout comme l'appartenance au groupe MLU 3 plutôt que 2, ce qui est lié au fait que les enfants moins avancés produisent davantage d'énoncés à un seul terme, dont la proéminence relative par rapport à d'autres termes ne peut être mesurée. L'arbre de partition indique que c'est l'interaction entre un mouvement de RETOUR (et notamment les cas de contraste discutés dans les chapitres précédents) et une discontinuité topicale qui favorise le plus l'accentuation prosodique (nommée ici PROACC ; nœud 6). Malgré la taille plus petite de l'échantillon, ce résultat nous semble assez net et correspond aux observations faites dans les chapitres précédents.

## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

### Arbre de partition 18 – Distribution et interaction des effets fixes pour les proéminences prosodiques (PROACC) chez les enfants germanophones



### 3.4.2 Les adultes

Dans les données des adultes, ce sont la discontinuité topicale et le mouvement interactionnel de RETOUR qui apparaissent comme seuls facteurs pertinents ; le Groupe MLU des enfants n'apparaît ici pas comme pertinent.

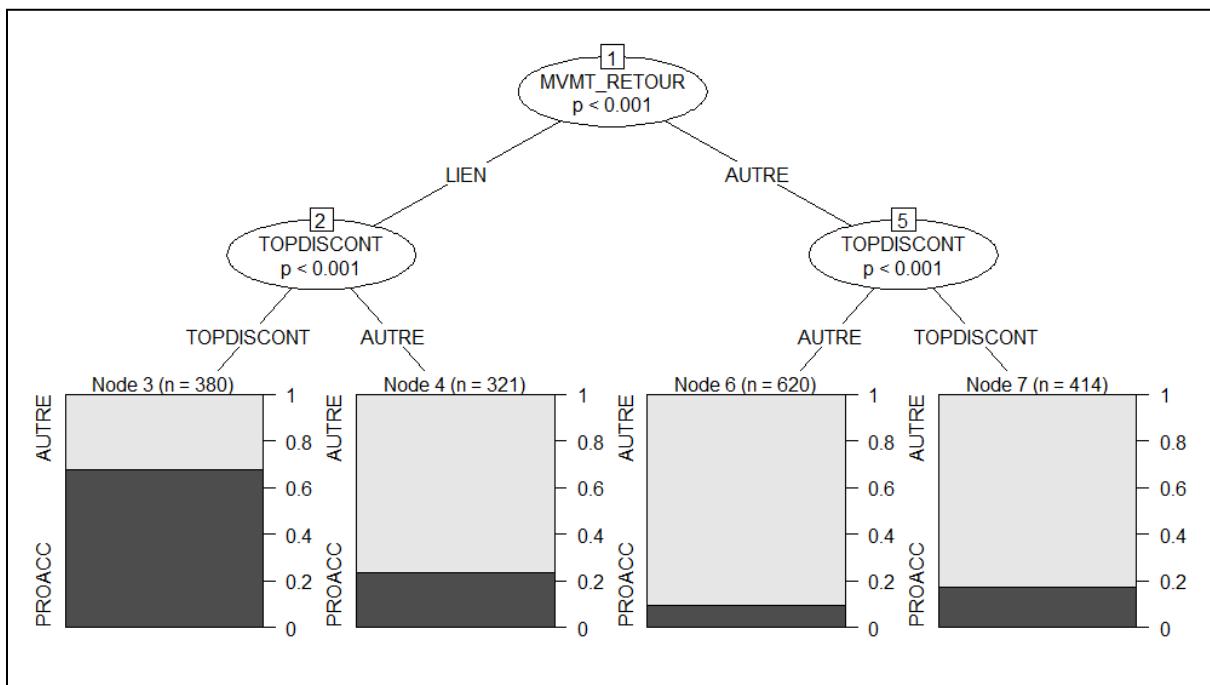
**Tableau XI-21 - Résultats du modèle linéaire mixte pour les accentuations ou proéminences prosodiques (adultes, allemand)**

Proéminences prosodiques, C=0.7912248				
Effets aléatoires :	Variance	SD		
Session	0.0508	0.2254		
Effets fixes :	Estimate	SE	z	p
(Intercept)	-2.6465	0.1420	-18.64	<.00001 ***
TOPDISCONT	1.3827	0.1266	10.92	<.00001 ***
MVMT RETOUR	1.7768	0.1259	14.11	<.00001 ***

SD= Standard Deviation ; SE= Standard Error ; Nombre de sessions : 12 ; Nombre d'observations : 1735

Comme pour les enfants, les contextes les plus favorables pour l'accentuation prosodique sont l'interaction entre une discontinuité topicale et un mouvement de RETOUR (nœud 3) ; ou la présence d'au moins un des deux facteurs (nœuds 4 et 7).

**Arbre de partition 19 – Distribution et interaction des effets fixes pour les proéminences prosodiques (PROACC) chez les adultes germanophones**



### 3.5 Résumé et discussion

En allemand, nous nous sommes intéressée aux D-Pro, aux noms, aux référents non-verbalisés ainsi qu'aux accentuations prosodiques. Nous avons constaté que les noms sont impliqués dans des mouvements de RETOUR, et également dans la discontinuité topicale lorsqu'ils ont le statut de topic. Pour les D-Pro, nous avons obtenu des résultats plus hétérogènes entre enfants et adultes. Les D-Pro semblent, pour les adultes, essentiellement servir à marquer le topic, et ne semblent pas favorisés pour signaler des changements de topics plutôt que leur maintien, ni jouer un rôle privilégié dans les mouvements de RETOUR. Pour les enfants, c'est davantage le cas, et nous l'avons mis en lien avec le fait que dans notre corpus allemand, ce sont davantage les enfants qui s'engagent dans des activités d'énumération de référents dans des listes, en employant des D-Pro. Pour les référents non-verbalisés, nous avons obtenu des modèles globalement comparables pour le corpus allemand et pour le corpus français, les non-verbalisations ne sont pas préférées pour effectuer des changements au niveau informationnel ou interactionnel. Dans les deux langues, les enfants les emploient quand même dans ce type de contexte avec une plus grande fréquence que les adultes, s'appuyant davantage sur le contexte physique et l'interprétation de l'adulte de leurs productions, dont la continuité/disruptivité n'est pas toujours clairement signalée. Toutefois, les adultes allemands, comparé aux adultes français, semblent se servir des non-verbalisations un peu plus souvent pour changer de topic et/ou effectuer un RETOUR, et cela est dû, à notre avis, à la caractéristique du topic-drop, qui permet de signaler une continuité, là où formellement intervient un changement de topic et/ou est effectué un lien vers un référent d'un énoncé ou tour précédent. L'analyse des

## Évaluation de l'importance relative des différents facteurs examinés sur le choix d'une expression référentielle

proéminences prosodiques enfin a permis de confirmer que cette réalisation d'une expression référentielle est effectivement favorisée dans des contextes qui marquent une rupture, que ce soit au niveau des topics ou au niveau de l'interaction. Notamment les topics contrastifs sont marqués, dans nos données, par une proéminence prosodique, là où en français nos locuteurs emploient une dislocation.

## **4. Synthèse générale**

Nous avons donc pu constater dans ce chapitre qu'effectivement, les facteurs liés à la gestion du topic et ceux liés à la gestion de l'interaction interagissent fréquemment et leur impact sur l'emploi d'une expression référentielle semble cumulative. En revanche, nous avons noté que d'autres facteurs apparaissent davantage comme complémentaires, ou plutôt coextensifs : ainsi par exemple, statut de topic, fonction syntaxique sujet et statut attentionnel Donné semblent effectivement souvent aller de pair, et n'apparaissent en conséquence pas tous comme pertinents dans les modèles. Le statut de topic semble l'emporter souvent sur d'autres facteurs (en tant que contexte défavorisant pour les noms, favorisant pour les autres expressions), mais ce n'est pas le cas pour les démonstratifs français ni pour les proéminences prosodiques de l'allemand. Les différents facteurs que nous avons étudiés au cours des chapitres précédents interagissent donc de manière complexe. Les dislocations du français correspondent soit à un changement de topic, soit à un mouvement interactionnel, soit les deux dans la majorité des cas. Comme nous avons seulement 39 occurrences en tout pour l'allemand, nous n'avons pas construit de modèle statistique, mais de fait les dislocations de l'allemand sont également produites dans ces contextes (avec deux seuls exceptions). Pour le français, nous avons en outre constaté une association privilégiée des dislocations avec un nombre réduit de schèmes lexico-syntaxiques. Ces schèmes sont aussi attestés dans nos données avec des noms ou pronom non disloqués, mais ces expressions apparaissent aussi en dehors de ces cadres. En ce sens, la présence d'un de ces cadres ou schèmes rend l'emploi d'une dislocation hautement prévisible, et correspond à une forme privilégiée pour exprimer une intention communicative comme nommer ou localiser.

Les modèles de ce chapitre ont également permis de voir que si les noms ne sont pas favorisés pour exprimer le topic, ils peuvent être employés pour signaler un changement de topic et un mouvement interactionnel de RETOUR. Nous avons choisi de ne pas comparer directement les deux langues dans nos modèles, pour éviter de brouiller ainsi d'autres facteurs, mais la comparaison des différents modèles de noms pour le français et pour l'allemand indique que l'allemand emploie davantage de noms dans ces deux contextes que le français (position relative des facteurs dans les modèles, visible dans les arbres notamment).

Pour ce qui est des pronoms démonstratifs, nous avons constaté que les démonstratifs du français sont favorisés en contexte de fonction sujet, présence d'un mouvement de RETOUR (dans les dénominations notamment) et pour le statut attentionnel d'Activé. Pour l'allemand, en revanche, ce semble être essentiellement le statut de topic qui favorise l'emploi d'un D-Pro. Le D-Pro allemand apparaît par ailleurs comme une forme passe-partout, qui sert aussi bien à exprimer la continuité que la discontinuité, dans les productions des adultes du moins : chez les enfants, cela est plus contrasté, et les D-Pro sont davantage liés à une rupture informationnelle et/ou interactionnelle. Ce résultat doit être mis en lien avec la grande fréquence de non-verbalisations chez l'enfant généralement, et avec le fait que les chaînes référentielles et topicales dans les interactions avec des jeunes enfants ont généralement tendance à être plutôt courtes, et que ce sont plutôt les mères qui contribuent à enchaîner sur un topic donné, alors que les enfants changent plus facilement d'un centre d'attention à un autre.

Les discussions des exemples du chapitre précédent ainsi que des modèles présentés ici nous a également permis de mettre en avant le fait que les productions des enfants semblent globalement correspondre aux contextes attendus pour une expression donnée. Cependant, et notamment dans l'emploi des formes non-verbalisées, nous avons pu constater aussi des écarts de ces emplois attendus. Si les enfants utilisent globalement des formes fortes pour signaler des ruptures (au niveau informationnel et interactionnel), toutes les ruptures ne sont en revanche pas signalées ainsi. Les enfants emploient plus souvent que les adultes des formes nulles pour changer de topic ou pour des énoncés qui constituent fonctionnellement un décrochage ou RETOUR interactionnel. Ces instances passent cependant souvent presque inaperçues ou sont interprétées et reformulés par l'adulte.

Il semble alors que les discontinuités informationnelles et interactionnelles sont en français essentiellement exprimés par des dislocations, suivi des démonstratifs et des noms, alors qu'en allemand, ces ruptures sont marquées par les proéminences prosodiques, les noms et les D-Pro.