Chute de schwa interne

Ce document étudie la chute des schwas internes, en hiatus et hors hiatus. Des lemmes ont été sélectionnés à partir de ceux réunis par Philippe Ségéral dans la bibliographie, et ont été interrogés dans le NCA3. La lemmatisation étant parfois en échec sur cette variation, l’ensemble des variantes graphique de chaque lemme a été réunis, puis une interrogation sur formes graphiques de la variation a eu lieu, en vérifiant que les cas relevés ne tiennent pas à l’homographie avec d’autres lemmes. Par hasard, le nombre de lemmes est le même pour les schwas en hiatus et les schwas hors hiatus : 8 lemmes pour chaque catégorie.

# Préambule — Dees, R et la cartographie

Les données de ce document sont ont été produite à partir du NCA 3, *via* une version simplifiée et customisée de Comparalem, programme R d’analyse de la variation graphique. Dans les métadonnées du NCA, 3 dates et 3 localisation sont disponibles :

1. Dates et loc. de Dees
   * Loc. par interpolation sur l’espace dialectal virtuel défini par les chartes
   * Date par ??
2. Date et loc. de composition
   * Définies par Glessgen et son équipe à partir de travaux philologiques traditionnels
3. Date et loc. du ms.
   * Idem

Le principal problème pratique des métadonnées de Dees est qu’elles sont défectives : sur 295 textes, seuls 98 sont localisés. Les régions Dees et le nombre de textes de chaque région sont représentés en Figure 1. Quatre régions ne disposent d’aucun texte : la Bretagne, la Mayenne et la Sarthe, le Maine-et-Loire, et la Vienne. Ce défaut reproduit ici celui de l’Atlas de Dees, qui rend difficile l’analyse dialectologique dans le Sud-Ouest du domaine. Par ailleurs, vis-à-vis du corpus du NCA, plusieurs régions sont sous-représentées, les textes n’étant pas localisés : Dees n’attribue ainsi que 8 textes à l’Angleterre (41 selon loc. compo.), 15 à la Normandie (37 selon loc. compo), 5 à la Vendée et aux Deux-Sèvres (10 selon loc. compo), etc. Ainsi, dans les localisations de Dees, ne dépassent 10 textes que les régions Somme et Pas-de-Calais, Haute-Marne, Nièvre et Allier, Normandie, Région Parisienne, Aisne et Wallonie.

En complétant les localisations de Dees à partir des données philologiques (2) et (3), on parvient à suppléer une localisation pour 88 textes ; seuls dix restent non localisés (cf. *infra*). Le défaut du Sud-Ouest demeure : seules les Vendée et Deux-Sèvres et l’Indre-et-Loire voient leur nombre de texte augmenter (resp. de 5 à 10 et de 1 à 2). L’Angleterre, elle, gagne 33 textes, passant de 8 à 41 textes, ce qui est un gain massif. Les autres gains sont plus modérés, avec la Normandie (de 15 à 28 textes), l’Aube (de 9 à 23 textes), la Somme et le Pas-de-Calais (de 37 à 47 textes), la région parisienne (de 14 à 18 textes), le Hainaut (de 3 à 6 textes), la Moselle et Meurthe-et-Moselle et le Nord (de 2 à 3 textes).

**METTRE À JOUR LES CARTES ET LES COMMENTAIRES !!!!**

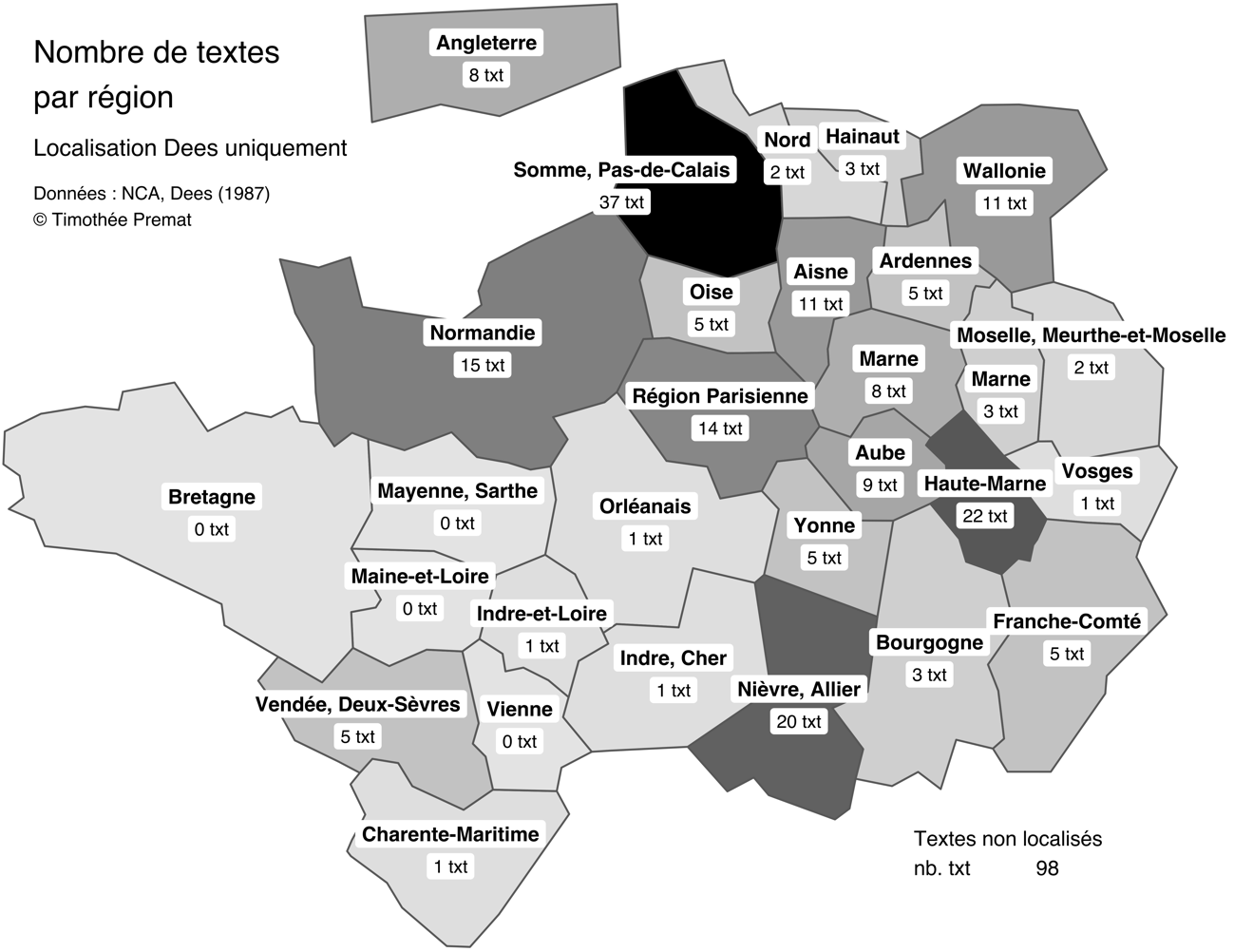


Figure . Nombre de textes par région selon les localisations de Dees

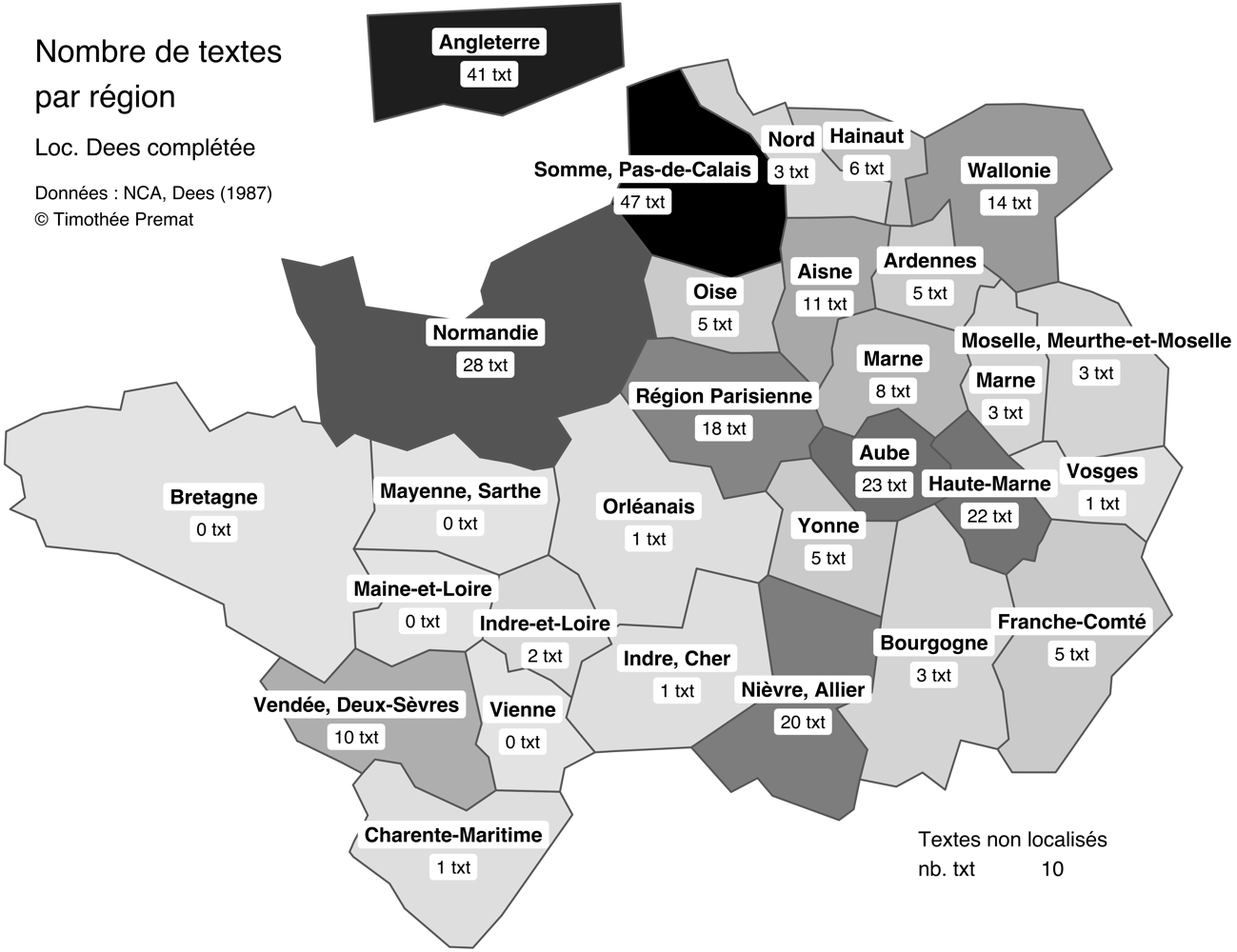


Figure . Nombre de textes par région selon les localisations complétées

Par ailleurs, les datemoyennedees (1) ne correspondent pas, comme leur nom l’indique, à une réalité historique : sauf en cas de composition et de copie simultanée, elles correspondent à une abstraction.

Cela dit, les localisations de compo. (2) et ms. (3) ne sont pas propres à être projetées sur une carte directement : elles cumulent plusieurs niveaux de granularité (p.ex. Canterbury et Angleterre sur le même plan), et confondent géographie et dialectologie (p.ex. Canterbury et agn. : l’un est un lieu, l’autre un dialecte). Enfin, certaines sont trop vagues pour être cartographiées (p.ex. Ouest, Est, etc.), ou, lorsqu’elles correspondent à des aires dialectales, ne sont pas clairement identifiable à des objets géographiques situés en termes de coordonnées géospatiales, tandis que nous disposons de fichiers géospatiaux qui correspondent aux régions définies par Dees en géographie moderne (fichiers transmis par Guylaine Brun-Trigaud).

Concernant les localisations, nous avons adopté l’approche suivante : nous conservons la localisation de Dees (1) lorsque celle-ci est disponible (199 / 296 txt). À défaut, nous utilisons la localisation de composition (2), sous une forme normée ( ??). Lorsque la localisation de composition est également manquante, nous utilisons la localisation du ms (3). Ce faisant, seuls 10 textes demeurent sans localisation. Cette méthode, qui combine des localisations de trois natures différentes, ne semble pas poser problème, dans la mesure où, comme illustré *infra*, elle ne change fondamentalement les résultats que lorsqu’une région dispose de très peu de textes ; lorsqu’une région dispose d’un nombre de textes supérieurs à 5 ou 10, les résultats ne varient que marginalement. Il y a donc une cohérence linguistique assez bonne entre les trois jeux de données, ce qui autorise à appuyer une analyse sur leur mélange.

Concernant la date, nous avons retenu la date de composition (2). Tel que développé *infra*, l’analyse des premières apparitions de la forme à hiatus interne réduit montre que c’est bien la date de composition qui offre la meilleure homogénéité, tandis que la date des ms. peut varier amplement.

# Schwas en hiatus internes

Les 8 lemmes sélectionnés sont :

* Abeesse (24 occ.) ~ abesse (14 occ.)
* Armeure (82 occ.) ~ armure (9 occ.)
* Cheveleure (11 occ.) ~ chevelure (5 occ.)
* Marcheant (179 occ.) ~ marchant (21 occ.)
* Mienuit (54 occ.) ~ minuit (1 occ.)
* Mireoir (15 occ.) ~ miroir (8 occ.)
* Pareïs (76 occ.) ~ parvis (2 occ.)
* Pecheeur (185 occ.) ~ pecheur (84 occ.)

Soit 770 occurrences avant traitement par le script R. Celui-ci retire 5 occurrences par groupe situées dans les textes les plus tardifs (=760). Soit, avant traitement, 626 occ. hiatiques et 144 occ. non hiatiques.

## Distribution diachronique

Les premières occurrences de réduction du hiatus apparaissent dans *psautier*, un texte anglo-normand du début du XIIe s. (compo. 1125, ms 1150), avec 9 formes *pecheur(s)* pour *pecheeur(s)*. La réduction apparaît ensuite :

* pour le lemme *marcheant* avec quatre formes *marchant* ou *marchans* dans *nimc* et *nimd*, deux ms. du *Charroi de Nîmes* ;
  + *nimc* est artois et date de 1150 compo. (1295 ms) ;
  + *nimd* est meusien et date de 1150 compo. (1275 ms.).
* de nouveau pour le lemme *pecheeur* avec
  + 11 formes *pecheur(s)* dans teo2, texte anglo-normand, compo. 1150, ms. 1210.
  + 2 formes *pecheur(s)* dans yonecH[[1]](#footnote-1), compo. 1165, ms. 1275.
* Pour les lemmes *armeure* et *cheveleure* avec reis[[2]](#footnote-2), texte anglo-normand, compo. 1175, ms. 1190. Ce ms. contient aussi 2 formes *marchant* et une forme *pecheurs*.
  + Il faut donc retenir la date de compo. et non la date du ms., puisque ce texte a un faible écart compo-ms. et est cohérent avec la date de compo des autres textes mais non avec leur date de ms. Même pattern pour *psautier*, texte des premières occurrences.
* Le lemme *mienuit* n’apparaît syncopé qu’à partir de gerv[[3]](#footnote-3), texte bourguignon, compo. 1210, ms. 1275. Ce texte ne contient qu’une seule forme syncopée, *minuit*.
* Le lemme *pareis* n’apparaît syncopé qu’à partir de *barril*[[4]](#footnote-4), texte orléanais, compo. 1218, ms. 1317. Ce texte ne contient qu’une seule forme syncopée, *parvis*.
* Le lemme *abeesse* n’apparaît syncopé qu’à partir de modw[[5]](#footnote-5), texte anglo-normand, compo. 1230, ms. 1250.
  + En plus de 13 formes *abesse*, le texte contient aussi 10 formes *pecheurs* et 1 formes *chevelure*
* Le lemme *mireoir* n’apparaît syncopé qu’à partir de fabb[[6]](#footnote-6), texte picard, compo 1250, ms. 1310. Ce texte ne contient qu’une seule forme syncopée, *miroir*.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Lemme | Première occurrence | | | |
| Date compo. | Date ms. | Dialecte | ID |
| *Pecheeur* | 1125 | 1150 | a.n. | psautier |
| *Marcheant* | 1150 | 1175-90 | meus./art. | nimd/nimc |
| *Armeure* | 1175 | 1190 | a.n. | reis |
| *Cheveleure* | 1175 | 1190 | a.n. | reis |
| *Mienuit* | 1210 | 1275 | bourg. | gerv |
| *Pareis* | 1218 | 1317 | orl. | Barril |
| *Abeesse* | 1230 | 1250 | a.n. | Modw |
| *Mireoir* | 1250 | 1310 | pic. | fabb |



L’apparition des formes à hiatus réduit s’étire sur 125 ans. 181 textes contiennent les lemmes en question, sous leur forme hiatique et/ou simplifiée. Hormis dans 10 textes (de 1 à 9 occ. par texte), les formes à hiatus réduit ne se substituent jamais parfaitement aux formes aux formes hiatiques, et, de fait, 152 textes (84%) ne contiennent que des formes hiatiques.[[7]](#footnote-7)

## Distribution diatopique

### Localisation Dees : R\_Code



Figure . Réduction des hiatus internes (localisation Dees)

En se fondant uniquement sur les localisations proposées par Dees (R\_Code), on obtient la carte en Figure 1.

La sur-représentation de l’a.n. parmi les premières attestation (cf. tableau) peut être retenue : l’a.n. n’est pas majoritaire en nombre de textes sur cette période (3 textes a.n. sur 15 textes entre 1125 et 1150 (bornes comprises) ; 37 textes a.n. sur 246 textes entre 1125 et 1250 (bornes comprises)). Néanmoins, la réduction du hiatus interne apparaît sur le continent dès 1150 pour *marcheant* (Meuse et Artois). Par ailleurs, sur les 10 textes ne contenant que des formes réduites, 6 sont anglo-normands.[[8]](#footnote-8)

De nombreuses régions échappent totalement à l’apparition de formes à hiatus réduit, alors qu’elles attestent les formes non réduites. Dans certains cas, ces régions disposent de peu de textes attestant ces lemmes, donc l’analyse est incertaine (p.ex. Charente Maritime, Indre et Cher, Moselle et Meurthe-et-Moselle, 1 texte par région attestant les lemmes), mais à l’inverse certaines régions disposent d’un nombre satisfaisant de textes et n’attestent pas les formes réduites (Normandie : 10 textes, Wallonie : 8 textes) ou ne les attestent que très marginalement (Nièvre et Allier : 1%, 17 textes ; Haute-Marne : 3%, 16 textes).

Si l’on veut dégager un patron géographique, il faut commencer par noter que la situation est assez hétérogène. Si l’on peut noter que les formes à hiatus simplifié semblent absente d’une petite moitié Sud-Ouest de la région d’oïl, il faut néanmoins considérer l’absence totale d’occurrences dans la moitié des régions de cette région, ce qui ne rend pas l’analyse assurée. Ainsi, les formes réduites des hiatus semblent se concentrer aux extrémités d’un axe Nord-Ouest–Sud-Est, et, dans une moindre mesure, au centre de cet axe.

### Localisation suppléante : R\_Code\_custom

Figure . Réduction des hiatus internes (localisation Dees + Glessgen)

Les données de localisation fournies par Dees ne concernent que 198 textes sur 296. Nous avons été en mesure d’établir une localisation pour 88 textes, à partir des informations philologiques collectées par Glessgen et son équipe. Pour cela, lorsque Dees n’a pas indiqué de région (et donc de R\_Code), nous avons retenu le R\_Code correspondant à la région de composition du texte ou, à défaut, à la région de copie du texte (une minorité de cas). Seuls dix textes n’ont pas pu être dotés d’un R\_Code\_custom, dans la mesure où les informations philologiques sont absentes ou trop vagues (p.ex. « Ouest » ne peut correspondre à une seule région de Dees). Enfin, des ajustements ont été nécessaires, dans la mesure où les indications géographiques de Dees ne correspondent pas toujours aux étiquettes habituelles de la philologie ; ainsi, nous avons toujours fait correspondre l’indication « pic. » (dialecte picard) aux départements de la Somme et du Pas-de-Calais utilisés par Dees, ou les indications « champ. » (champenois) et leur sous-catégorie « champ. mérid. » (champenois méridional) au département de l’Aube. L’utilisation de ces données supplétives ne change guère la carte. On peut noter les points suivants :

* Le nombre de textes normands passe de 10 à 16, mais le taux de réduction des hiatus interne demeure 0% : cela renforce l’hypothèse selon laquelle certaines régions ne sont pas atteintes par cette réduction.
  + De même pour l’Aube, qui passe de 5 à 17 textes, mais demeure à 0% de hiatus réduits
  + De même pour la Wallonie, qui passe de 8 à 9 textes mais demeure à 0% de hiatus réduits
  + De même, dans une moindre mesure, pour la Somme et le Pas-de-Calais, qui passent de 22 à 25 textes, et de 6% à 5% de hiatus internes réduits.
* Le nombre de textes anglo-normands passe de 5 à 22, et le taux de hiatus réduits passe de 59% à 48%. Cette réduction du taux reste raisonnable et cohérente ; l’Angleterre continue à être une des deux régions où les hiatus sont le plus réduits.
* Enfin, ces localisations supplétives permettent de combler deux trous dans les données, qui s’avèrent cohérents avec l’analyse proposée *supra* : l’Indre-et-Loire et le Nord disposent chacun d’un texte attestant uniquement des variantes non réduites des hiatus interne.

Ainsi, l’ajout de localisation philologiques sur la carte ne change guère la lecture que l’on peut en faire : la répartition continue à ignorer une petite moitié Sud-Ouest de la carte (mais marquée également par une absence totale de ces lemmes sur une partie du territoire), la Normandie, la frange Nord-Est de la région d’oïl ainsi qu’une bande de territoires comprenant les régions de Dees Indre et Cher, Nièvre et Allier, Yonne et Aube. En positif, les hiatus réduits se concentrent donc aux deux extrémités d’un large axe allant du Nord-Ouest (Angleterre) au Sud-Est (Bourgogne, Franche-Comté), et, dans une moindre mesure, au centre de cet axe (Orléanais, 17% ; région parisienne, 18% ; Aisne, 12% ; Meuse, 38%).

Deux taux sont à relativiser par le faible nombre de textes sur lesquels ils portent : le taux de 50% de la Bourgogne (2 textes), et le taux de 38% de la Meuse. Si la réduction des hiatus internes est bien attestée dans ces deux régions, la valeur exacte de leur taux peut être surévaluée.

1. Marie de France, *Lai de Yonec*, ms. H. [↑](#footnote-ref-1)
2. *Les quatre livres des rois*, ms. M. [↑](#footnote-ref-2)
3. *Le bestiaire de Gervaise* [↑](#footnote-ref-3)
4. *Le conte dou barril, par Jouham de la Chapele de Blois*, ms. P. [↑](#footnote-ref-4)
5. *Saint Modwenna* (hagiographie : *Vie de Saint Modwenna* ?), ms. O [↑](#footnote-ref-5)
6. Coll. de fabliaux [↑](#footnote-ref-6)
7. length(which(POS\_EnonE\_count$n.nonE == 0)) [↑](#footnote-ref-7)
8. aileg, edmond, fablesB, gerv, gibv, joinv, miro, plainte, psautier, yonecH. [↑](#footnote-ref-8)