

Les Dictionnaires du Brabançon (WBD) et du Limbourgeois (WLD) sur l'Internet

Joep Kruijsen et Jos Swanenberg

Département de Linguistique

Université Radboud, Postbus 9103, NL 6500 HD Nijmegen

j.kruijsen@let.kun.nl - j.swanenberg@let.kun.nl

Résumé

Le Dictionnaire des Dialectes Brabançons (*Woordenboek van de Brabantse Dialecten, WBD*) et le Dictionnaire des Dialectes Limbourgeois (*Woordenboek van de Limburgse Dialecten, WLD*) couvrent le sud-est du territoire où l'on parle le néerlandais, les provinces de Brabant du Nord et de Limbourg dans les Pays-Bas et les provinces d'Anvers, de Brabant flamand et l'autre province de Limbourg dans la Belgique. C'est le territoire où l'on trouve les variétés du bas-franconien méridional, entre le flamand des Flandres dans l'ouest, le cléverois dans le nord, le ripuarien dans l'est et les dialectes wallons dans le sud.

Depuis la fin de 2002 les deux dictionnaires se trouvent sur internet. La banque des données complète (enquêtes propres, autres enquêtes depuis la fin du XIX^e siècle, dépouillement de monographies, etc.) qui est à la base des fascicules parus récemment, peut être consultée sur le web, aussi bien article par article du dictionnaire, simulé par le programme conçu à ce but, que dans la forme d'un fichier qu'on peut parcourir selon des paramètres spécifiques.

Pendant la conférence on pourra prendre connaissance des possibilités nouvelles pour la lexicographie dialectale, telle qu'on les a mis en oeuvre dans les *WBD* et *WLD*.

1. Introduction

Un trésor de renseignements dialectaux s'est ouvert naguère sur l'internet dans le domaine des dictionnaires dialectaux du néerlandais septentrional. Ce sont

- (a) les index cumulatifs de la terminologie rurale (volume I) et de la terminologie artisanale (volume II) des dictionnaires des dialectes limbourgeois et des dialectes brabançon;
- (b) la banque des données complètes des huit fascicules parus du vocabulaire général (volume III) de ces deux dictionnaires et
- (c) les réponses complètes de la grande enquête dialectale dans le sud des Pays-Bas entreprise en 1914.

Les trois banques de données peuvent être consultées sur <http://www.kun.nl/dialect/>.

2. Les dictionnaires régionaux

Le Dictionnaire des Dialectes Brabançons (WBD) et *Le Dictionnaire des Dialectes Limbourgeois (WLD)*, les deux fondés par Toon Weijnen, fondateur, lui aussi, de *l'Atlas Linguarum Europae*, recouvrent la partie germanophone des l'ancien duché de Brabant (ce sont les provinces de Brabant du Nord dans les Pays-Bas en les provinces d'Anvers et de Brabant flamand dans la Belgique) de même que les deux provinces de Limbourg, l'un dans la Belgique et l'autre dans les Pays-Bas. À ces deux dictionnaires, rédigés dans les universités de Nimègue et de Louvain, se joint à l'ouest *Le Dictionnaire des Dialectes Flamands (WVD)*, conçu de la même façon et recouvrant les provinces de la Flandre

orientale et la Flandre occidentale en Belgique, les Flandres françaises en France en La Flandre zélandaise dans les Pays-Bas, voir la carte dans Figure 1.

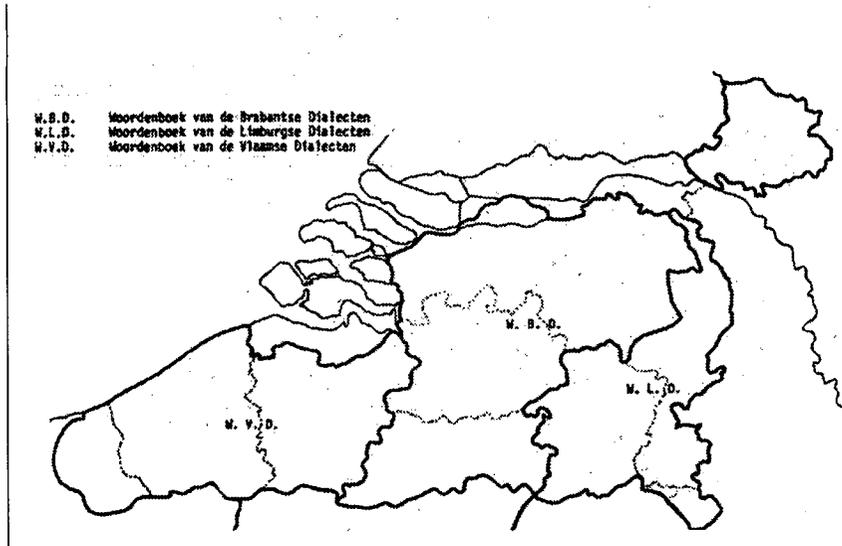


Figure 1: Les trois dictionnaires septentrionaux du néerlandais

Les trois dictionnaires prennent une position spéciale dans la lexicographie dialectale. L'ordre des articles ne suit pas l'alphabet, mais l'« emploi », l'ordre logique du monde virtuel des dialectisants. Ainsi, dans la terminologie rurale, les fascicules suivent les activités de l'agriculteur qui prépare ses terres, soigne les prés, sème et entretient les champs, nourrit son bétail en entretient la basse-cour, récolte et range ses produits. On trouve, dans la partie des terminologies artisanales, les métiers du boulanger et du boucher, du mineur et du cordonnier, ou encore, dans le vocabulaire général, le noms des oiseaux et de la flore.

Un article du dictionnaire donne toutes les dénominations avec les lieux d'occurrence, qui ont été rapportées pour la notion en cause, pourvues d'une carte, d'un commentaire et d'une illustration si cela est utile. L'ensemble d'un fascicule est rendu accessible par des index alphabétiques.

Les sources sont d'abord les enquêtes propres du projet, plus de 140 questionnaires distribués depuis 1960 dans quelque 200 villages ; puis les enquêtes d'autres instituts de dialectologie (d'Amsterdam, de Louvain, etc.) et les enquêtes historiques, les dictionnaires locaux, les monographies, etc. Somme toute, les projets visent à la description exhaustive des vocabulaires dialectaux.

Les Volumes I et II des deux dictionnaires contiennent les terminologies rurales et artisanales et la publication des quelque 40 fascicules est quasiment terminée. La publication des fascicules du Volume III, qui contiendra le vocabulaire général, est en cours.

Le site dont il est question contient donc les index des Volumes I et II en les données complètes des fascicules publiés des volumes III.

2.1 Index des Volumes I et II.

Les quelque 40 fascicules des deux premiers volumes des deux dictionnaires (agriculture et métiers et artisanats) contiennent tous un index alphabétique sur les titres des articles et les mots de référence (interprétations des termes dialectaux relevés).

Le site contient les index cumulatifs pour chacun des volumes.

Si l'on cherche sous le mot de référence **spie** ('clavette') dans l'index sur volume I (terminologie rurale) du dictionnaire limbourgeois, on trouve :

Les index des Volumes I et II			
Spie 'clavette'			
forme de référence	fascicule	chapitre	
page			
Spie		WLD.I.01	labourage 71
Spie		WLD.I.03	fenaison 30
Spie	WLD.I.03	fenaison	45
Spie		WLD.I.06	ferme 127
Spie	WLD.I.10	harnais des chevaux	26

Figure 2: **Spie** 'clavette' dans le volume I du dictionnaire limbourgeois (WLD)

et cela veut dire que **spie** figure dans les fascicules sur le traitement de la terre, dans celui sur la fenaison, celui sur la ferme et dans celui sur le harnais des chevaux.

2.2 La banque des données des volumes III

La publication des fascicules des deux dictionnaires a commencé dans les années soixante. Toutes les variantes relevées pendant les enquêtes se retrouvent imprimées dans ces premières publications, comme c'était dans ce temps la seule moyen de les publier, de les « rendre publiques ». Aujourd'hui d'autres moyens de publication sont à notre disposition. Le début de la publication des fascicules sur le vocabulaire général a été un point de repère approprié pour changer la macrostructure des dictionnaires.

Dans le volume III, le texte du dictionnaire publié « sur papier » ne contient que les formes de référence avec l'indication de la fréquence et des régions ou des villages où l'on les a trouvées. Au lieu de faire paraître toutes les variantes phonétiques et tous les codes des

localités dans le texte des fascicules mêmes, nous avons installé un site dans l'internet qui contient la banque organisée des données et que l'on peut consulter à côté du livre.

Chaque fascicule connaît de la sorte sa propre « base de matériaux » qui contient toutes les attestations des enquêtes et des sources anciennes, avec leurs interprétations qui ont trouvé abri dans les fascicules. De la sorte ces bases de matériaux peuvent être considérées comme des versions très détaillées des fascicules publiés, où les formes de référence sont pourvues de la forme relevée même, telles que les informateurs les ont notées, en notation phonétique ou autre, et de la localisation précise de chaque attestation, avec source, date de l'enquête, commentaire de celui qui a rempli le questionnaire, etc.

Cette opération rend d'un côté le texte publié beaucoup plus accessible pour un grand public intéressé et pourvoit de l'autre les dialectologues d'infiniment plus de détails sur les enquêtes que dans les anciens fascicules contenant les vocabulaires techniques.

Les bases de matériaux des fascicules du Volume III sont donc à consulter sur l'internet ; à ce moment il s'agit des huit fascicules sur la maison, la flore et la faune des dialectes brabançons et limbourgeois et ces bases contiennent plus de 500.000 données dialectales. Dans une année le nombre de fascicules disponibles sera triplé.

Deux façons de consultation du site s'imposent: (1) faire simuler un article du dictionnaire et enrichir de la sorte la lecture du fascicule ; et (2) chercher directement sur des formes, des significations, des lieux d'attestation, des sources etc., dans la base complète.

Je présente ici quelques exemples des deux sentiers.

2.2.1 Simulation du dictionnaire. On peut faire simuler un article du dictionnaire directement dans la base des matériaux : rassembler tous matériaux pour un concept et les faire organiser en formes de référence, parallèlement au texte publié.

Prenons comme exemple le geai (**Vlaamse gaai**, WLD Volume III, section 4,1 p. 146-8) et comparons les deux versions; d'abord la version imprimée (Figure 3).

Dans le dictionnaire, l'article du geai comprend deux pages et demie, dont voici la première. On y discerne le titre, la carte avec la légende et l'indication des fréquences en chiffres, l'énumération des sources; l'explication et la précision du concept présenté en ensuite la série des formes de référence, (éventuellement les variantes lexicales en italiques), dans un ordre libre (non-alphabétique) avec les noms des régions (avec l'indication de la fréquence relative) ou les noms des localités. Dans la figure 3 on voit les quatre premières formes de référence (sur 60 au total), avec celle de **markolf** à la fin. Ensuite on trouve dans le dictionnaire l'indication des régions (Ripuarien) avec la fréquence relative (alg. 'général'), ou les nom des localités (comme Hasselt).

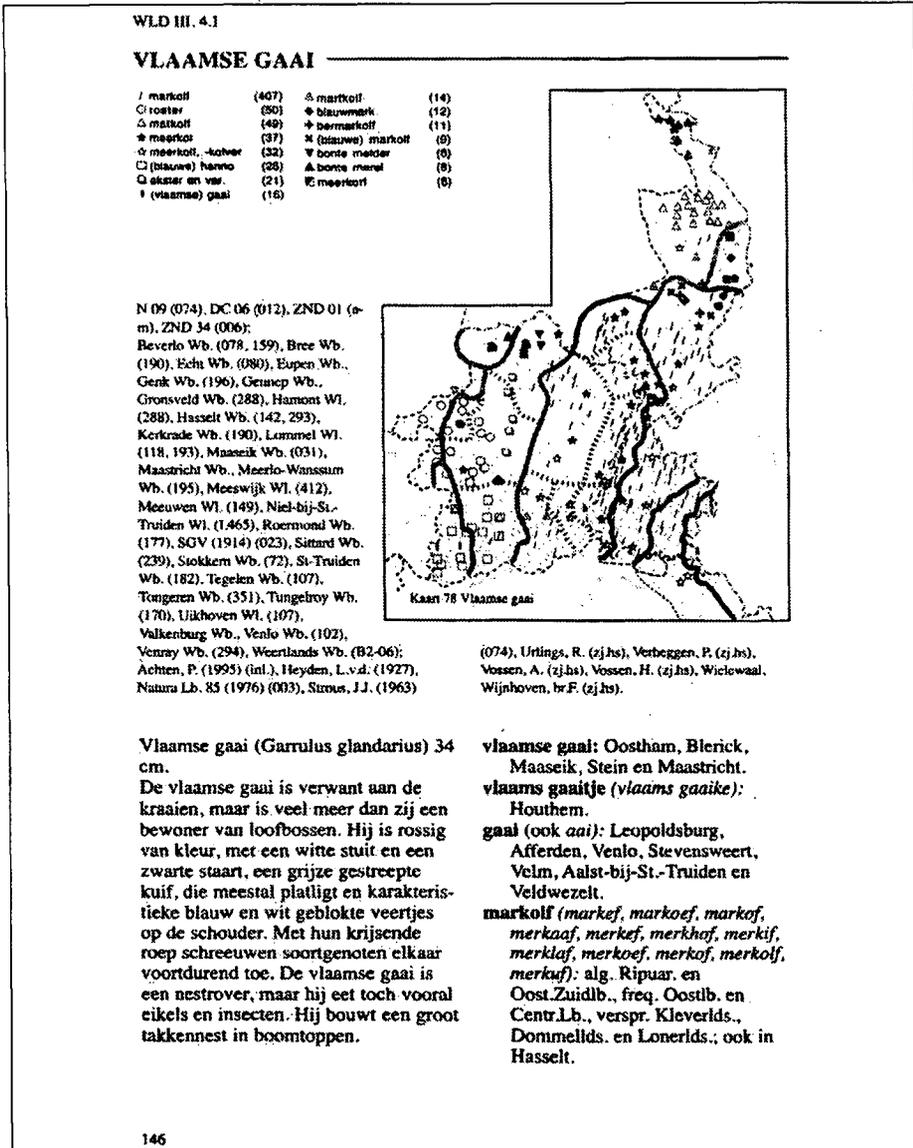


Figure 3: Page 146 du fasc. des oiseaux : début du traitement du gaai dans le WLD

Retenons comme exemple de traitement le cas de **markolf**, la dénomination la plus fréquente dans les deux provinces de Limbourg du gaai: plus de 400 attestations dans le limbourgeois de l'est et du centre). L'étymologie du terme l'identifie avec le nom d'homme Marculph, ou Marculphus, personnage connu dans les vaudevilles médiévaux, où il figure comme le grand

bavard. La série de variantes lexicales (douze dans ce cas) réfère à la fantaisie et à l'étymologie populaire.

Dans la base de matériaux des oiseaux, on retrouve la table des matières de la publication, et donc aussi l'article du geai. Une simple touche suffit pour faire apparaître toutes les formes de référence du même geai, avec toutes les autres informations dont il était question.

L'ensemble de l'article du geai comprend 6 pages; les huit lignes sur markolf du dictionnaire de la Figure 3 sont devenues deux pages dans la simulation de la base des matériaux. Nous avons repris ici les dix premières lignes de la forme de référence **markolf**:

Markolf Markef mark´f ZND 01 (1922) Q 089; *markef* ZND 01 (1922) Q 168; ZND 34 (1940) Q 168a; Markoef *ma.rkoef* N 09 (1961) L 331; *ma_rkoef* SGV (1914) L 323; *markoe.f* N 09 (1961) L 330; *markoef* DC 06 (1938) L 322; N 09 (1961) L 322; Strous, J. (1963) L 322; N 09 (1961) L 322a; Strous, J. (1963) L 322a; N 09 (1961) L 325; Strous, J. (1963) L 325; Roermond Wb. L 329; Strous, J. (1963) L 329, L 330, L 331, L 332; Verheggen, P. (zj.hs) L 332; SGV (1914) L 333; Markof *ma.rkoef* N 09 (1961) L 331b; *ma_rkof* ZND 01 (1922) L 419; SGV (1914) Q 113; Merkaaf *merkoaf* N 09 (1961) Q 113; Merkef [*me7rk´f* Hamont Wl. L 286; [*mErk´f* Meeuwen Wl. L 364; [*mErk´f* N 09 (1961) Q 072; *mIerkef* N 09 (1961) Q 083; *me\$:rkef* N 09 (1961) Q 086; *me@rekef* N 09 (1961) Q 198b; Strous, J. (1963) Q 198b; *me@reko_f* Strous, J. (1963) Q 196; *me(rk´f* DC 06 (1938) Q 034; *me(rke(f* SGV (1914) Q 103;

Figure 4: Détail de la base des matériaux du geai du dictionnaire limbourgeois (WLD)

On y distingue la forme de référence **Markolf** (en gras), les variantes lexicales (en romain) Markef, Markoef, Markof, etc.; les variantes (en italiques) telles qu'on les a relevées dans les enquêtes : *markef*, *ma.rkoef*, etc.; puis les sources avec la date de la relevée: l'enquête telle que «ZND» (Louvain) ou «N» (Nimègue) ou l'indication d'une monographie telle que «Strous, J. (1963)» en enfin les codes des localités, selon le registre généralement accepté dans les Pays-Bas et la Belgique néerlandophone, celui de Kloeke (voir *Systematisch* etc. sous les Références).

Il est évident que les informations fournies par les base des matériaux de l'internet sont de beaucoup plus détaillées en ce qui concerne la documentation phonétique et la finesse des indications géographiques que ne le donne le dictionnaire publié, qui, lui de son côté, est plus accessible au lecteur.

2.2.2 Recherche dans la banque des données. On peut s'approcher aussi directement à la banque des données, en se passant des fascicules parus, et parcourir, par exemple, la base des matériaux des noms dialectaux des oiseaux, à la recherche d'une certaine forme de référence, qui pourrait se présenter sous plusieurs concepts. On peut aussi vérifier quels noms d'oiseaux se présentent dans une certaine source (questionnaire ou monographie), ou parmi toutes les attestations provenant d'un certain village, etc.

De cette façon on peut trouver la réponse aux questions telles que:

(a) quelles sont les significations du mot **markolf** en limbourgeois? Et on trouvera, hors des 407 attestations pour le geai dont nous avons parlé, encore deux autres oiseaux qui se sont désignés par le mot **markolf**: la pie grièche et la merle.

Cette polysémie est encore plus fréquente dans la flore. Si l'on cherche, dans la base de données de la flore en brabançon, les occurrences de la dénomination **hondskruid** ('fleur de chien') on obtient 4 attestations, réparties sur 4 plantes, respectivement la tanaïs, l'aunée, la violette des chiens et le chiendent, voir Figure 5.

LEMMATITEL (titre)	TREFWOORD PLAATSCODE (forme de réf.) (variante)	OPGAVE (source)	BRON (code de localité)
Boerenwormkruid	Hondskruid <i>hondskruid</i>	DC 60A (1985)	K 147
Donderkruid	Hondskruid <i>onskruid</i>	DC 46 (1971)	K 188a
Hondsviooltje	Hondskruid <i>hondskro't</i>	N 092 (1982)	P 002
Kweek	Hondskruid <i>honskruid</i>	Antwerps Wb. 3	K 244

Figure 5: Forme de référence **hondskruid** dans la base brabançonne des matériaux de la Flore (WBD), détail

(b) quels sont les noms d'oiseau dans le dictionnaire de Maastricht (Endepols 1955)? Le résultat de recherche comprend 36 attestations; ce ne sont que --à côté de 16 dénominations pour les pigeons (preuve que les colombophiles sont nombreux à Maastricht!) et quelques parties du corps comme le bec-- sept oiseaux : pinson, hirondelle, corneille, geai, bécasse, canard et cygne.

(c) ou la même question sur les noms d'oiseau dans la ville de Maastricht dans toutes les sources de la banque des données; voir Figure 6 (page suivante).

3. La banque des données

Jusqu'ici il a été question de données interprétées par les rédacteurs des dictionnaires, qui ont déterminé les formes de référence qui apparaissent dans les fascicules et qui sont transférées dans les «base de matériaux» sur l'internet. Ces bases de matériaux se limitent, provisoirement, aux fascicules parus.

LEMMATITEL (titre) localité)	TREFWOORD PLAATSCODE (forme de réf.)	OPGAVE (variante)	BRON (source)	BRON (code de
GELE KWIKSTAART Q 095	Akkermannetje	<i>akkermenke</i>		DC 26 (1954)
VINK	Amourette	<i>amourette</i>	DC 06 (1938)	Q 095

APPELVINK	Appelvink	<i>appelvink</i>	N 09 (1961)	Q 095
ROOFVOGEL, ALG.	Arend	<i>arend</i>	Veldeke 11 (1936)	Q 095
JONG EN KAAL				
VOGELTJE, ADJ.	Bang	<i>bang</i>	N 83 (1981)	Q 095
HEGGENMUS	Bastaardnachtegaal	<i>bastaard-</i> <i>nachtegaal</i>		DC 06 (1938)
Q 095				
BARMSIJS	Berpje	<i>be[^]rbke</i>	N 09 (1961)	Q 095

Figure 6: Toutes les attestations dans la ville de Maastricht dans la base des matériaux des oiseaux en limbourgeois (WLD), détail

BEGRIP (concept)	BLZ. (page)	PLAATS (localité)	OPGAVE (variante)
Aardappel	001	Q 116	(<i>erpel</i>)
Aardappel	001	L 387	[<i>E#rpel</i>
Aardappel	001	L 266	[<i>patat</i>
Aardappel	001	L 266	[<i>petat</i>
Aardappel	001	Q 118	<i>a_a_rpel</i>
Aardappel	001	Q 030	<i>a_a_rpel</i>
Aardappel	001	Q 113	<i>a_dappel</i>
Aardappel	001	Q 203	
		<i>a_e(dappel</i>	
Aardappel	001	L 429	<i>a_rpel</i>
Aardappel	001	Q 105	
		<i>e#e#rappel</i>	
Aardappel	001	Q 117b	<i>e\$dappel</i>
Aardappel	001	L 164	<i>e\$rpel</i>
Aardappel	001	Q 033	<i>e\$rpel</i>
Aardappel	001	L 426	<i>e\$rpel</i>
Aardappel	001	Q 110	<i>e\$rpel</i>
Aardappel	001	L 382	<i>e\$rpel</i>
Aardappel	001	L 434a	<i>e\$rpel</i>
Aardappel	001	L 215	<i>e\$rpel</i>

Figure 7: Exemple de pomme de terre dans l'enquête SGV (détail)

En attendant le souhait s'est imposé pour construire d'abord une banque des données commune aux trois dictionnaires septentrionaux, limbourgeois, brabançon et flamand, contenant les *matériaux bruts*, de première main, dans la notation des informateurs.

Les premiers pas dans cette direction ont été mis et ils sont présentés sur le site indiqué. En 1914 une grande enquête a été initiée dans la meilleure tradition gilliéronienne, dans le sud-est des Pays-Bas (2280 items dans 179 localités), par trois fondateurs de la

linguistique à Nimègue: Schrijnen, Van Ginneken en Verbeeten (l'enquête se désigne d'après les initiales: l'enquête SGV). Les quelque 400.000 données brutes de cette enquête, utilisée toujours dans cahiers originaux des initiateurs pour la rédaction de la partie historique des dictionnaires du brabançon et du limbourgeois, ont été digitalisées et se trouvent sur le site et sont à consulter, voir Figure 7, détail des données de **pomme de terre** (page précédente).

Depuis 1960, le centre de dialectologie de Nimègue a distribué 60 questionnaires sur le vocabulaire général, dans quelque 200 points d'enquête. Les données complètes de ces questionnaires ont été introduites dans l'ordinateur en on prépare actuellement ces données pour les placer sur le site à côté des données de 1914. Après, les matériaux bruts des questionnaires nimégeois des volumes I en II (terminologie rurale et artisanale) et utilisés pour les premiers fascicules parus, seront préparés pour le site.

Ainsi, on prépare à l'université de Louvain l'introduction dans la banque des données des réponses réunies pendant une grande enquête dialectale («ZND») menée depuis 1922 à l'institut de dialectologie de l'université de Louvain.

4. Perspectives

Les projets de recherche vont encore plus loin. Au cours de 2004 le dictionnaire du brabançon sera terminé, celui du limbourgeois suivra quelques années plus tard et le dictionnaire du Flamand encore quelques années plus tard. Dans le cadre des trois dictionnaires, beaucoup de programmes de tri et de cartographie ont été projetés et les rédactions cherchent à divulguer ces outils sur le site, au profit des lecteurs et des utilisateurs des dictionnaires. A ce moment nous faisons l'inventaire des possibilités techniques pour introduire automatiquement les articles parus dans les publications anciennes «sur papier» et d'autres recueils de matériaux dans la banque des données et d'y accoupler les programmes de tri et de cartographie.

Finalement, le dialectologue pourra remanier les résultats de ses enquêtes dans cette banque de données dans des articles nouveaux, selon ses préférences et à son gré.

5. Références

- Endepols, H.J.** (1955, 1985 4^e éd.). *Woordenboek of Diksjaener van 't Mestreechs*. Maastricht: Boosten & Stols.
- Systematisch en alfabetisch register van plaatsnamen voor Nederland, de Nederlands-sprekende delen van België en Noord-Frankrijk en het noordwesten der Duitse Bondsrepubliek*. Amsterdam/Antwerpen 1962.
- WBD: Weijnen, A., J. van Bakel e.a.** (1967 --). *Woordenboek van de Brabantse Dialecten*. Assen: Van Gorcum. (paraît en volumes et fascicules).
- WLD: Weijnen, A., J. Goossens e.a.** (1993 --). *Woordenboek van de Limburgse Dialecten*. Assen: Van Gorcum. (paraît en volumes et fascicules).
- WVD: Devos, M., H. Ryckeboer e.a.** (1979 --). *Woordenboek van de Vlaamse Dialecten*. Tongeren: Michiels. (paraît en volumes et fascicules).
- www.kun.nl/dialect