

Shoebox

la boîte à chaussures du linguiste

POUR
WINDOWS ET MACINTOSH

Alan Buseman
Karen Buseman

Traduction en cours : **Christian Chanard**
chanard@cnrs-bellevue.fr

Version 4.0

Summer Institute of Linguistics
Waxhaw, North Carolina
April 1998

Cette documentation est vendue avec le logiciel Shoebox. Et ce logiciel est une propriété, protégée par copyright, de la Société Internationale de Linguistique (Summer Institute of Linguistics). Dans le souci de partager le fruit de notre recherche avec la communauté scientifique la plus large, le propriétaire de ce logiciel est autorisé à distribuer des copies de ce jeu de disquettes à des amis ou associés, dans la mesure où il n'en tire pas fait de profit commercial. Les bénéficiaires de ce logiciel, s'ils décident de l'utiliser, peuvent en devenir propriétaires en achetant ce livre avec la dernière version du logiciel.

Shoebox est en cours de développement. En diffusant ce logiciel, la SIL ne prend aucun engagement à en assurer le suivi. Il est, en revanche, permis à l'utilisateur d'adresser ses commentaires aux auteurs du logiciel, qui pourront choisir d'aller plus loin ou non dans son développement.

Microsoft Word, Microsoft Windows, Windows 95, Windows NT, et MS-DOS sont des marques déposées de Microsoft Corporation. IBM est une marque déposée de International Business Machines Corporation. Macintosh est une marque déposée de Apple Computer Corporation.

Ce manuel a été saisi sous Microsoft Word pour Windows 8.0 avec des polices TrueType Microsoft Windows. L'original a été imprimé sur une imprimante Hewlett-Packard LaserJet Series 6 sans réduction.

La couverture est de Linda Lackey.

©1989, 1990, 1991, 1993, 1996, 1998 by the Summer Institute of Linguistics, Inc.

All rights reserved

Printed in the United States of America

Shoebox for Windows and Macintosh

Version 3.0	November, 1996	200 copies
	December, 1996	200 copies
	May, 1997	200 copies
Version 4.0	April, 1998	200 copies ***

Produced and Distributed by:

JAARS, Inc.
International Computing and Telecommunications Services
PO Box 248, JAARS Road
Waxhaw, NC 28173-0248
USA

Telephone: (704) 843-6085

FAX: (704) 843-6500

Email: computer_sales_jaars@sil.org

Product support e-mail:

shoebox_support@sil.org

Table des matières

Remerciements	v
Introduction	7
Shoebox pour Windows et Macintosh	7
Survol de la documentation	7
APPRENTISSAGE	9
Fonctionnalités de base	10
Premiers pas	10
Projet Shoebox	11
Démarrer l'apprentissage	11
Sortir de Shoebox	12
Affichage	12
Le système d'Aide	13
Les Marqueurs	15
Tri, Index et Hiérarchie des marqueurs	21
Rechercher, Retrouver, Sauter à, Filtrer	26
Quelques autres commandes	35
Créer un nouveau projet et une base de données	37
Présentation	37
Partir de zéro: pas à pas	39
*Installer la police Axininca	39
*Keyman (pour Windows seulement)	40
Démarrer un nouveau projet	41
Créer une nouvelle base de données	41
Texte interaligné	49
Introduction	49
Principe de base de l'inter-alignement	49
Paramétrage de l'inter-alignement	51
Un peu plus sur l'Inter-alignement	55
Affixes	57
Morpho-phonologie	58
Restructuration du texte interaligné	63
Un peu plus sur la morpho-phonologie	63
Traduction dans des langues apparentées	64
RÉFÉRENCES ET	67
QUESTIONS FRÉQUENTES	67
Principes de base	69
Les Projets	69
Les Marqueurs de Champs	70
Classes	72

Impression	73
A propos de l'exportation de fichier	74
L'enregistreur de Windows pour créer des macros	77
Enregistrer votre Presse-papiers	80
Questions fréquentes	81
Démarrer avec Shoebox.....	81
Marqueurs	82
Polices.....	83
Gérer des bases de données	84
Questions diverses	85
Notes sur l'Interalignement.....	87
Définir un découpage dans le champ de structure sous-jacente.....	87
Les morceaux les plus longs ont priorité sur les plus courts	87
Réduire les ambiguïtés.....	88
Valeurs forcées.....	89
Réduplication.....	90
Infixes	92
Racines composées	92
Les tirets dans le lexique	92
Caractères séparateurs de morphèmes	93
Lignes de segmentation multiples	93
Mots multiples.....	94
Gloses multiples	94
Contrôle orthographique.....	95
Editer	95
Coller dans d'autres applications.....	95
APPENDICES	97
Appendice A: Importer de Shoebox 2.0	99
Présentation	99
Importer un Projet complet —les étapes	99
Importer vos propres données.....	110
Appendice B: version SH2.0, comparaisons.....	112
Les différences les plus importantes avec la version DOS de Shoebox.....	112
Filtres.....	113
Où est passé ?	114
INDEX.....	ERREUR ! SIGNET NON DÉFINI.

Remerciements

John Wimbish et ses collègues de la SIL en Indonésie sont à l'origine de la réflexion sur le logiciel Shoebox et en ont développé la version DOS en 1988. Depuis, leur travail a migré sur les plates-formes Windows et Mac, et s'est développé de façon significative.

Le logiciel a été écrit par Alan Buseman, Rod Early, Mark Pedrotti, Brian Yoder, et Bryan Wussow. Le fichier d'Aide a été écrit par Barbara Altork. Le directeur de Projet a été Karen Buseman. Doug Casse et Doug Rintoul ont apporté une contribution significative à la programmation (pour la partie Table de remplacements CCT) de même que Timm Erickson (pour le rendu contextuel utilisant les fichiers de définition de Script). Doug Eberman, Josh Kelley et Wolfgang Witteveen ont également apporté leur soutien, de façon plus ponctuelle, par leur connaissance en programmation.

Cette documentation a été écrite par Karen Buseman, avec la contribution de Neal Brinneman et Mark Pedrotti. Elle a été revue et significativement augmentée par David Coward. La relecture a été faite par Don Horneman, Jim Albright, et un groupe de testeurs. Le support du logiciel Shoebox est presque intégralement dû à Jon Coblenz. Phil Baer, Jr., Chet Matheson, John Dunn et Michael Buchanan ont développé le cours d'entraînement sur ordinateur (CBT). Neal Brinneman a produit les cours d'entraînement sur Shoebox.

Ces différentes personnes ont élaboré le produit sur lequel vous allez travailler avec votre ordinateur. Cependant ceci n'aurait pas été possible sans l'aide inestimable de notre équipe de testeurs. Nombre d'entre eux ont contribué au produit final en reportant des problèmes et en donnant des idées par courrier électronique, occasionnant de multiples échanges pendant un certain temps. Rien de tout cela n'aurait été possible sans l'équipe du courrier électronique de JAARS qui a su maintenir en état ce lien vital de par le monde. Merci à tous.

Introduction

Shoebox pour Windows et Macintosh

Shoebox est un logiciel de gestion de bases de données spécialement conçu pour gérer les données d'un linguiste de terrain. Il est particulièrement adapté à l'interalignement de texte et au développement de lexiques. Il est également utile pour les notes de terrain, les informations anthropologiques, et grammaticales.

Cette version de Shoebox tourne sur Windows 3.1, Windows 95, et Macintosh.¹ Elle tourne également sur Windows NT (bien que Shoebox ne gère pas Unicode). En utilisant le même code source, nous avons compilé trois versions différentes : SHW.EXE (pour Windows 3.1), SHW32.EXE (pour Windows 95), et une version pour Macintosh. Les ordinateurs sous Windows 95 peuvent faire tourner l'une ou l'autre version Windows, et la version Windows 95 tourne sur NT. Seule la version Mac tourne sur un Mac.

Cette documentation sert à la fois pour la version Mac et Windows du logiciel. Cependant il y a quelques légères différences entre les deux versions Mac et Windows. Les raccourcis claviers en sont une. Sur le Mac, c'est la touche *commande* qui est utilisée, alors que sous Windows c'est la touche *Alt*. Certaines touches de raccourcis ne peuvent avoir la même valeur. Tous les choix des menus qui autorisent un raccourci clavier l'affichent dans le menu, aussi, généralement dans cette documentation nous ne les mentionneront pas explicitement. Si vous souhaitez utiliser les raccourcis clavier, reportez-vous au menu ou au fichier d'Aide.

Une autre différence consiste en l'utilisation du *Clic-droit*. Une souris Windows comporte au moins deux boutons. La souris Mac n'en a qu'un. Nous avons utilisé le Clic-droit parcimonieusement. Dans ces cas, *Option+clic* est utilisé sur Mac.

Une différence terminologique que nous ne pouvions éviter concerne ce qui est désigné sous le nom de *répertoire* dans Windows 3.1 (et DOS) et qui correspond au *dossier* sur Mac et Windows 95. Ces termes sont équivalents et les deux seront utilisés dans cette document. Autre différence terminologique, le *groupe de programmes* sous Windows 3.1 qui est appelé *fenêtre de dossier* sous Mac et Windows 95.

Survol de la documentation

La première section de cette documentation se veut être un didacticiel ou "parcours" à effectuer sur l'ordinateur. Il constituera un bien meilleur outil de formation si vous le suivez sur l'ordinateur plutôt que si vous vous contentez de le lire. Ce document devrait vous familiariser avec le logiciel et les différentes façons de s'en servir. Vous pourrez trouver plus de détails dans les fichiers d'Aide qui sont accessibles en ligne. Les instructions sur les différentes façons d'accéder à l'aide de façon efficace sont présentées dans ce parcours.

NOTE: Nous vous recommandons *vivement* de réaliser le parcours. Certains points et concepts ne sont malheureusement pas si intuitifs. Nous avons essayé de rendre chaque section indépendante des autres de façon à pouvoir faire le parcours par morceaux suivant vos besoins ou votre intérêt personnel.

Voici les différentes sections abordées dans ce document :

Fonctionnalités de base fait une introduction à la fenêtre d'affichage de Shoebox, au fichier d'Aide, aux Marqueurs, au Tri, aux principales fonctions de déplacement dans la base de données, comme Rechercher, Retrouver, Sauter vers, et les Filtres.

¹ Pour les utilisateurs de Shoebox pour DOS (SH2), nous avons essayé de conserver la plus grande compatibilité possible. Cette version lit et écrit des fichiers de données en Format Standard, tout comme SH2 le faisait. Nous ne pouvions pas conserver les fichiers de paramétrage mais nous avons fourni des fichiers d'aide pour une conversion semi-automatique. Voir *Trois concepts*, p. 37, pour apprendre à la faire.

Démarrer un nouveau projet montre comment créer un nouveau Projet et des bases de données. Il décrit les notions de Projet, Encodage de langue, et Type de base de données.

Texte interaligné est une introduction aux fonctions d'interalignement de texte, y compris le découpage morphologique et l'élaboration de lexique à partir de texte.

Principes de base est plutôt un index de références. Il contient des informations sur l'impression et l'exportation de bases de données, l'utilisation des Classes et Filtres, les Propriétés des marqueurs, les Projets, etc. Il discute également de *Recorder*, l'enregistreur de macros de Windows, et du Presse-livres qui peuvent rendre la vie plus facile.

Questions fréquentes contient certaines questions qui nous ont été posées, et leurs réponses, ainsi que quelques problèmes courants rencontrés, avec leurs solutions.

L'Appendice comporte deux chapitres :

Trois concepts décrit comment importer des fichiers provenant de SH2 dans cette version-ci de Shoebox.

Appendice B: version SH2.0, comparaisons comporte des informations mettant en relief les différences entre la version SH2 et cette version-ci de Shoebox, comme les différences dans l'utilisation des Filtres et une section "Où est passée ?" qui compare les menus des deux versions.

Conventions de notation

Dans ce document et dans le fichier d'Aide, un + entre deux ou trois touches signifie qu'elles doivent être tapées ensemble. Cela signifie qu'il faut maintenir appuyé la première touche et taper la seconde (ex. Ctrl+N). Nous avons choisi cette notation pour représenter les combinaisons Option+clic sur Mac utilisées comme équivalent au Clic-droit sous Windows.

Lorsqu'un raccourci clavier est mentionné, ses versions Windows et Mac sont étiquetés par l'initiale des plates-formes ainsi **w**:Alt+S, **m**:commande+M signifie : sous Windows utiliser *Alt+S* et sous Mac utiliser *commande+M*.

Cette police est utilisée pour marquer les items qui doivent être tapés exactement comme affichés. Elle est également utilisée pour la dénomination des choses, spécialement lorsqu'elles sont nouvelles ou en cours d'explicitation.

Cette police est réservée aux Menus que vous aurez à choisir ou aux Boutons sur lequel vous devrez cliquer (pas pour les Menus ou Boutons auxquels il est simplement fait référence dans le texte).

Les marqueurs de champs sont identifiés par cette police.

Parfois l'*italique* ou le **gras** est utilisé pour une mise en relief (parfois les *deux*).

SECTION 1

APPRENTISSAGE

Fonctionnalités de base

Premiers pas

Présentation

Nous allons décrire et illustrer l'utilisation des fonctionnalités principales, en mettant en évidence le comportement à la fois pratique et spécifique du logiciel. Il ne s'agira pas ici de montrer comment utiliser Shoebox pour un projet particulier, mais plutôt de présenter une série d'illustrations rapides de fonctionnalités importantes. Dans une large mesure, les différents sujets seront indépendants les uns des autres, bien qu'à l'intérieur d'un même sujet, une sous-section aura tendance à s'appuyer sur les résultats des précédentes. Le système d'Aide sera quelque peu détaillé, en particulier les marqueurs, le tri, ainsi que les différentes commandes Rechercher/Retrouver; d'autres sujets seront traités plus superficiellement.

Si vous n'avez pas ce manuel

Si vous n'avez pas un exemplaire papier de ce document, vous pouvez exécuter le parcours d'apprentissage dans la partie haute de l'écran tout en affichant ce document dans la partie basse. Vous êtes sans doute en train de lire ce document sous Microsoft Word 6 ou Word Viewer 6 (ou encore Microsoft Word 5 ou 6 sous Mac). Pour réserver de la place pour Shoebox dans la partie haute de l'écran, nous allons réduire la taille de la fenêtre d'affichage de ce document à la partie basse de l'écran. Pour maximiser la quantité de texte affiché, nous allons enlever les ascenseurs, les barres d'outils et les barres d'état.

Dans Word (ou Word Viewer):

- Choisissez Affichage, Normal.
- Maximiser la sous-fenêtre de ce document.
- Choisissez Outils, Options (dans Word Viewer choisissez Affichage, Options).
- Choisissez l'onglet *Affichage* s'il n'apparaît pas en premier plan.
- Sous *Fenêtre*, décochez Barre d'état et Barre de défilement horizontal. (Mais pas la barre de défilement vertical. elle sera nécessaire.)
- Sous *fenêtre*, mettez Largeur de la zone de style à zéro.
- Sous *Affichage*, cochez Ajuster à la fenêtre.

Il peut être intéressant de désactiver l'option d'*Enregistrement automatique* vu que vous n'aurez pas à modifier ce document. Ceci se trouve sous l'onglet *Enregistrement*.

- Cliquez sur OK pour fermer la boîte *Options*.
- Choisissez Affichage, Barre d'outils.
- Décochez toutes les barres d'outils.
- Cliquez sur OK pour fermer la boîte *Barre d'outils* (sous Word'97 celle-ci se ferme automatiquement, aussi il vous faudra répéter ce processus pour chaque barre que vous voudrez enlever).
- Si la règle s'affiche, Cliquez sur Affichage, Règle pour la faire disparaître.

Ceci devrait vous ramener à une fenêtre contenant juste la barre de titre et le menu en haut, l'ascenseur à droite, et rien à gauche ni en bas. On a maximisé ainsi la quantité de texte visible dans un espace limité d'affichage.

- Redimensionnez la fenêtre générale de Word afin qu'elle ne couvre plus que le tiers inférieur de l'écran.

Vous devriez maintenant pouvoir voir environ 8 lignes (avec le zoom à 100%). Si vous supportez un caractère plus petit, vous pouvez voir quelques lignes de plus en choisissant une valeur de zoom (dans *Affichage*) plus petite.

Les notes de bas de page de ce document ne sont pas critiques pour la compréhension, mais si vous devez lire une note de bas de page, vous pouvez. Commencez par déplacer le curseur sur le numéro de la note. Puis choisissez *Affichage, Notes*. Puis tirer le haut de la fenêtre des notes jusqu'à ce que la note soit lisible. Par exemple, voici un échantillon de note.² Lorsque vous aurez terminé de lire la note, choisissez *Fermer* dans la fenêtre des notes pour la fermer.

Projet Shoebox

Dans Shoebox, vous pouvez définir différentes configurations de travail appelées Projets. Un projet s'appuie sur un fichier de configuration de projet qui mémorise quels étaient les fichiers chargés, combien de fenêtres vous aviez ouvertes, où elles étaient placées sur l'écran et quelles données elles affichaient (ainsi que divers autres paramètres).

Les Projets sont tout à fait intéressants pour organiser votre espace de travail Shoebox par domaines de recherche. Chaque domaine peut avoir son propre projet qui conservera la trace de ce que vous faisiez la fois précédente dans ce domaine. De cette façon, passer d'un domaine à l'autre est aisé et rapide.

Démarrer l'apprentissage

Dans cet apprentissage, nous lancerons différents projets pour ouvrir des fichiers d'échantillons de données spécifiquement confitures.

Nous utiliserons quelques échantillons de données provenant d'un lexique de la langue Selaru d'Indonésie.³ Ce lexique utilise le jeu standard de marqueurs du progiciel MDF (Multi-Dictionary Formatter).⁴

- Si Shoebox n'est pas encore lancé, démarrez-le en le sélectionnant avec *Démarrer, Programmes*, menu *Shoebox* (ou en double-cliquant sur son icône).
- Si vous n'avez jamais lancé Shoebox ou si la dernière fois que vous l'avez lancé, vous avez fermé votre projet avant de quitter, vous verrez la boîte de dialogue *Pas de projet ouvert*; choisissez *Ouvrir un projet existant* et cliquez sur OK.
- Si vous avez déjà lancé Shoebox auparavant et étiez sorti sans fermer le projet, Shoebox ouvrira ce que vous aviez ouvert la dernière fois. Choisissez *Projet, Ouvrir* dans la menu.
- Ceci vous amène à la boîte de dialogue *Choisissez le fichier de Projet à ouvrir*. Dans cette boîte, naviguez dans le dossier SAMP_SET. Il se trouve en dessous du dossier SAMPLES (qui est lui-même en dessous du dossier dans lequel Shoebox a été installé).
- Sélectionnez SHLEX.PRJ.
- Choisissez *Ouvrir*.⁵

Ceci ouvrira le fichier de projet SHLEX ainsi que tous les fichiers qui lui sont associés. (Un fichier de projet mémorise les noms et chemins de tous les fichiers que vous utilisiez la dernière fois que le projet a été ouvert.) Lorsque le projet s'ouvre correctement, une fenêtre affiche le premier enregistrement du fichier SHLEX.DB.

Pour ceux qui n'ont pas ce manuel

Si vous êtes en train de lire ce document sur écran, Shoebox a peut-être recouvert un bout de cette fenêtre. Pour un affichage optimal, positionnez le haut de la fenêtre de ce document de façon à ce qu'elle recouvre tout juste la barre d'état de Shoebox. Cette barre d'état

² Ceci est un échantillon de note de bas de page pour les utilisateurs de *viewer*.

³ Ce sont là des données non publiées, collectées par David et Naomi Coward.

⁴ Ce apprentissage n'aborde pas le programme MDF. Pour plus de détails sur l'utilisation de ce programme, reportez-vous à la documentation de MDF incluse avec Shoebox ou bien le manuel : *Making Dictionaries: A guide to lexicography and the Multi-Dictionary Formatter*. David F. Coward, Charles E. Grimes, et Mark R. Pedrotti. Waxhaw, NC: SIL, 1998 (2nd édition).

⁵ Ce bouton est intitulé OK sous Windows 3.1.

apparaîtra lorsque Shoebox sera actif, mais elle ne masquera que la barre de titre de la fenêtre du document, pas le texte. Sur Mac, la barre d'état de Shoebox apparaît en bas de la fenêtre, aussi il sera judicieux de positionner la fenêtre du document juste en dessous de la barre d'état.

Si vous lisez ce document sur l'écran, vous devrez régulièrement changer de fenêtre entre le document et Shoebox, en cliquant avec la souris; pour revenir dans Shoebox, cliquez toujours sur sa barre de titre (ou bien dans la zone de fond à l'extérieur de ses fenêtres si elle est visible.) Ceci évitera de perdre la position initiale du curseur dans Shoebox.

Il peut arriver qu'une boîte de dialogue vienne recouvrir une partie de la fenêtre de ce document. Si cela arrive, cliquez sur la barre de titre de cette boîte de dialogue et tirez-la vers le haut de l'écran.

Sortir de Shoebox

A n'importe quel moment de cet apprentissage, si vous devez sortir de Shoebox, faites Fichier, Fermer en laissant ouvertes les bases de données et fenêtres. Ceci conservera dans l'état toutes les données. Lorsque vous lancerez de nouveau le programme, celui-ci chargera automatiquement vos fichiers et les affichera exactement dans l'état où ils étaient précédemment vous permettant ainsi de continuer là où vous en étiez arrivé.

Affichage Format Standard

Shoebox lit et écrit des fichiers dans ce qu'on appelle le Format Standard de la SIL—*Standard Format* (SF). Ce format est utilisé par la SIL pour stocker des données sur les langues depuis les années 70. Il consiste en des balises génériques appelées marqueurs qui étiquettent chaque unité d'information dans un fichier. Ce format était utilisé avant que les traitements de texte n'existent, mais les marqueurs ont la même fonction que les styles dans un traitement de texte et ainsi ils peuvent être convertis en styles pour la publication ou l'impression à travers un traitement de texte. Cependant un marqueur fait bien plus qu'agir comme un style, car il renseigne sur le type de l'information qui le suit. Ce balisage est utilisé par de nombreux autres logiciels de la SIL pour structurer l'information dans les fichiers en format standard.

Un *marqueur* au format standard consiste en un caractère antislash (\) en début de ligne, suivi par un ou plusieurs caractères constituant une abréviation du type de contenu du marqueur. Le marqueur est séparé du texte qui le suit par un espace.

Enregistrements et Champs

Shoebox organise les données en enregistrements et en champs. Un *champ* est constitué de son marqueur et du texte associé. Le *marqueur d'enregistrement* est le marqueur qui débute chaque nouvel enregistrement.

Une *base de données* est une collection d'enregistrements. Par exemple, un lexique contient de nombreux enregistrements, généralement un enregistrement par mot ou morphème de la langue. A son tour, chaque enregistrement contient divers éléments d'information sur le mot ou le morphème en question; ces divers éléments d'information constituent les champs.

Ce qu'il faut remarquer

Observez qu'il y a une barre verticale entre les marqueurs (sur la gauche) et le contenu des champs; et ce contenu est aligné sur la droite de la ligne. Nous avons la zone des marqueurs et la zone des contenus.⁶

Dans la zone des données, vous pouvez déplacer le curseur de texte au moyen des touches flèches. Vous pouvez également positionner le curseur avec la souris. Vous pouvez éditer du texte en insérant ou effaçant des caractères comme d'habitude ; Remarquez que lorsque vous déplacez le curseur de la souris dans la zone des marqueurs, il se transforme en une flèche.

La barre d'outils

Au dessus de la zone des données et en dessous de la barre des menus se trouve la barre d'outils. Il y a de nombreux outils dans cette barre. Ce sont des raccourcis pour les

⁶ Si certaines données apparaissent en gras, en italique ou en couleur, c'est que ce parcours aura déjà été fait par une autre personne. Vous devrez réinstaller les exemples. Pour cela, fermez Shoebox et relancez son installation en sélectionnant uniquement les exemples. Ensuite relancez Shoebox, et si nécessaire ouvrez le projet SHLEX comme précédemment.

différentes commandes. De la gauche vers la droite⁷ dans la barre d'outils, on trouve Ouvrir un fichier, Sauvegarder, Couper, Copier, Coller, Précédent, Suivant, Interalignement, et Index (affiche une liste des enregistrements). Puis on trouve une case pour un mot à rechercher, suivie d'une petite flèche vers le bas qui servira à afficher la liste des derniers mots recherchés, et ensuite deux paires de jumelles signifiant respectivement Rechercher le Suivant et Rechercher le Précédent. Enfin on trouve la case pour le filtre actuellement appliqué à l'affichage et sa petite flèche qui sert à présenter la liste des filtres disponibles pour cette base de données. Certains de ces points sont décrits plus en détail ci-dessous. Les boutons Copier et Coller seront inactivables tant qu'aucun texte n'aura été préalablement sélectionné.

Si vous préférez ne pas voir ou utiliser la barre d'outils, vous pouvez la supprimer par le menu Affichage, Barre d'outils. (Elle pourra être restaurée de la même manière.)

La barre d'état

La barre d'état en bas de l'écran affiche différents types d'informations. Sur la gauche se trouve une *zone de message*. Elle sert à décrire l'action demandée lorsqu'on clique sur un bouton du menu. Ensuite vous pouvez voir le *marqueur d'enregistrement en cours* ainsi que le contenu du champ correspondant. Ensuite on trouve le *marqueur de tri primaire* et son contenu. Ceci permet de savoir exactement où l'on se trouve dans la séquence des enregistrements. (En ce point du parcours, cette zone de la barre d'état devrait être vide. Ceci parce que Shoebox n'affiche le marqueur de tri primaire (et le contenu de ce champ) que s'il est différent du marqueur d'enregistrement, en effet comme ici SHLEX est trié par rapport au marqueur d'enregistrement, cet affichage serait redondant.) Ainsi, ces deux champs dans la barre d'état vous aideront à déterminer où vous vous situez dans la séquence des enregistrements.

La barre d'état affiche également le numéro de l'enregistrement en cours et le nombre total d'enregistrements dans la base de données, ce qui devrait être ici 1/57 si vous êtes au début de la base. (En fait il ne s'agit du nombre total d'enregistrements que lorsque la base est triée par le marqueur d'enregistrement et que le Filtrage n'est pas actif. Ceci est détaillé plus bas sous le titre *Tri*, p.)

Dans la dernière zone de la barre d'état, est affichée le nom du fichier de projet, ici *SHLEX.PRJ*.

Si vous préférez ignorer la barre d'états, elle peut être masquée par Affichage, Barre d'Etat. (Elle peut être restaurée de la même façon.)

Le système d'Aide

Le système d'aide est un volet très important du programme car il contient bien plus d'informations qu'il n'y en a dans le manuel. Le système d'aide contient des aides pour débiter, des aides sur l'utilisation du programme ainsi que des aides techniques pour les utilisateurs avancés. Aussi ayez le réflexe de faire appel à l'Aide chaque fois que vous rencontrerez un problème. Il y a quatre manières d'accéder à l'Aide, à savoir l'aide contextuelle, le sommaire, les liens et la recherche.

L'aide contextuelle

L'aide contextuelle sert à expliciter les fonctions des outils et les choix des menus.

Choisissez Aide.⁸

Choisissez Aide contextuelle.

Vous voyez apparaître à côté du pointeur de la souris, un point d'interrogation qui vous signale qu'il est prêt à vous donner une aide contextuelle.

Déplacez le pointeur de la souris sur le bouton de l'Index (le bouton de la barre d'outils représentant une grille) puis cliquez dessus.

L'aide devrait s'activer et vous fournir des informations sur l'Index. (Si avec Mac vous obtenez un message disant *Fichier d'Aide introuvable*, reportez-vous à la note ci-dessous.⁹)

⁷ Si vous travaillez sous Win95, vous devrez peut-être maximiser (élargir) l'écran pour voir tous les outils.

⁸ Sur Mac, c'est le point d'interrogation dans le coin en haut à droite de l'écran.

Fermez la fenêtre d'aide.

Vous pouvez aussi utiliser l'aide contextuelle sur un choix de menu

Choisissez Aide, Aide contextuelle.

Choisissez Base de Données, Trier.

Vous verrez apparaître des informations sur le tri.

Fermez la fenêtre d'aide.

Vous pouvez aussi avoir une aide contextuelle sur n'importe quelle boîte de dialogue.

Choisissez Base de Données, Trier.

Vous verrez apparaître une boîte de dialogue intitulée *Champs de tri*. Un des boutons de cette boîte s'appelle *Aide*.

Cliquez sur le bouton Aide.

Vous verrez apparaître la même boîte de dialogue que précédemment.

Fermez la fenêtre d'aide.

Cliquez sur Annuler dans la boîte de dialogue *Champs de tri*.

Les utilisateurs sous Windows peuvent accéder à l'aide contextuelle en tapant la touche de fonction F1.

Le sommaire de l'aide

Le sommaire est une autre façon d'accéder à des informations d'aide.

Choisissez Aide, Index.

Vous verrez une table des matières comportant divers sujets. Ces derniers devraient vous permettre de trouver comment faire ce que vous cherchez à faire. Par exemple sous *Overviews of Features* (survol des fonctionnalités) vous verrez un sujet intitulé *Index*.

Cliquez sur Index.

Vous verrez de nouveaux sujets apparaître parmi lesquels *Moving around in the database*

Cliquez sur Moving around in the database.

Vous verrez un sommaire sur les différents moyens de se déplacer dans une base de données.

Ne refermez pas encore la fenêtre d'Aide.

Les liens dans l'aide

Les liens dans l'aide sont un moyen d'accéder à partir d'un sujet donné, à un sujet qui lui est relié. Tout mot ou phrase souligné dans une fenêtre d'aide est un lien vers des informations supplémentaires. Ainsi, la fenêtre d'aide actuellement ouverte contient de nombreux mots soulignés (comme *Enregistrement suivant*, *Rechercher*, et *Retrouver*) .

Cliquez sur le mot souligné *Browse* au bas du texte d'aide.

Vous verrez la même information sur *Browse* que précédemment. Vous verrez également de nombreux liens vers d'autres sujets.

Si un mot ou une phrase est souligné de façon continue, il conduira à un sujet différent. Le bouton *Back* permet de revenir en arrière sur le sujet précédent. Si un mot ou une phrase est souligné en pointillés, cliquer dessus fera apparaître une petite fenêtre descriptive tout en laissant affichée la fenêtre précédente ; un second clic de souris fermera cette petite fenêtre.

⁹ Sous Mac, l'aide tourne dans un programme séparé qui nécessite 500 Ko de mémoire. Si le système ne dispose pas de cette mémoire, l'aide ne se lancera pas. Il y a plusieurs façons de s'en sortir. Si vous avez lancé d'autres programmes, vous pourriez les fermer pour libérer de la mémoire. Un moyen sûr d'avoir plus de mémoire consiste à valider la mémoire virtuelle, ou à l'étendre si elle est déjà activée. Vous devriez également contrôler les chiffres de la mémoire affectée à Shoebox. Vous pouvez y accéder en sélectionnant l'icône de Shoebox et en choisissant Fichier, Lire les informations. Si la taille souhaitée est plus grande que la taille disponible lorsque Shoebox est lancé, c'est qu'il utilise toute la mémoire.

Dans la fenêtre d'aide sur *Browse*, cliquez sur *empty field*. Vous verrez apparaître une petite boîte contenant une définition.

Cliquez n'importe où dans la fenêtre de définition pour la refermer.

Comme autre bouton utile dans le haut de la fenêtre d'aide remarquez *Contents*, qui permet d'accéder à la table des matières du fichier d'aide.

Ne refermez pas encore la fenêtre d'aide.

La recherche d'aide

Un des moyens les plus efficaces pour rechercher de l'aide consiste à rechercher un mot comme dans un index de livre. L'un des boutons en haut de la fenêtre d'aide s'intitule *Rechercher*

Cliquez sur *Rechercher*.

Vous verrez une boîte de dialogue comportant une case à remplir avec le sujet à rechercher, et une liste alphabétique des sujets traités.

Tapez *sor* et remarquez qu'au fur et à mesure que vous tapez une lettre, la liste des items saute au premier d'entre eux qui commence par ce que vous avez tapé.

Sélectionnez *Sorting Command* (*Base de Données* menu).

Sous *Windows 95*, cliquez sur *Afficher*. Sous *Win 3.1* ou *Mac*, cliquez sur *Montrer l'item*, puis cliquez sur *Go To*.¹⁰

Vous verrez apparaître la même fenêtre d'aide sur le tri que précédemment.

Fermez la fenêtre d'Aide.

Choisissez *Aide, Rechercher* for *Aide on*.

Vous verrez la même fenêtre de recherche d'aide.

Annuler *Recherche* pour revenir à *Shoebox*.

Les Marqueurs

Les noms de marqueurs de champs

Bien que les marqueurs soient utiles et même indispensables, ils n'en sont pas moins inesthétiques et sont souvent trop brefs pour être aussi explicites qu'on le souhaiterait. *Shoebox* est assez souple dans l'affichage des marqueurs. Par exemple,

Choisissez *Affichage, Noms de champs*.

Vous verrez la ligne se déplacer vers la droite et les noms des champs se substituer à celui des marqueurs.

La ligne séparant les champs des contenus peut être déplacée. Pour déplacer cette ligne,

Placez le pointeur de la souris sur la ligne de séparation verticale.

La flèche devrait se transformer en une double flèche horizontale.

Maintenez le bouton de la souris appuyé en tirant la ligne vers la droite de deux ou trois centimètres.

Vous constaterez que certains Noms de champs sont très longs. Par exemple,

Example free trans. (E)

Pour voir ce champ, cliquez sur l'ascenseur vertical pour dérouler le texte vers le bas.

Pour modifier le nom de ce champ, faites **w**:Clic-droit ou **m**:Option+clic sur le nom du champ.

Vous verrez apparaître une boîte de dialogue contenant de nombreuses propriétés du marqueur **\xe**.

¹⁰ Un double-clic sur la *Commande de tri* (menu *Database*) revient au même que cliquer sur *Afficher* (ou *Montrer les items*). De la même façon, double-cliquer sur un item revient à cliquer sur *Aller à*.

- Remplacer le nom de champ par :
Trad. libre Exemple (E).
- Cliquez sur OK.
- Changer l'affichage pour voir en même temps les marqueurs et les noms des champs.
(Noms des marqueurs et des champs.)
- Changer l'affichage pour ne plus voir que les marqueurs. (Choose Noms des marqueurs.)
- Maintenant remettez le nom de champ d'origine pour le marqueur **\xe** :
Example free trans. (E)

Insérer et supprimer des marqueurs

Vous pouvez insérer et supprimer des marqueurs grâce aux mêmes touches que vous utilisez pour insérer et effacer dans un texte. Par exemple,

- Déplacez le curseur à la fin du mot *listen* dans le premier champ **\ge**.
- Tapez **w**:Entrée ou **m**:Retour
- Tapez un antislash (\).

Une boîte s'ouvre dans la zone des marqueurs.

- Tapez **ge** suivi d'un espace.

Notez que lorsque vous avez tapé l'espace, la boîte s'est fermée et le curseur est passé dans la zone des contenus.

- Tapez **w**:Suppr_arrière ou **m**:Suppr.

Remarquez que vous êtes de nouveau dans la boîte des marqueurs avec **ge** en surbrillance.

- Tapez **w**:Suppr_arrière ou **m**:Suppr pour supprimer le **ge**.
- Tapez **w**:Suppr_arrière ou **m**:Suppr de nouveau et la boîte des marqueurs disparaît et le curseur se retrouve à la fin du champ précédent..

La liste déroulante des marqueurs

Shoebox maintient une liste de tous les marqueurs. Il y a deux endroits où cette liste peut être consultée. La petite boîte d'édition affichée pour saisir un marqueur comporte une flèche qui fait apparaître la liste de tous les marqueurs. Vous pouvez choisir un marqueur dans cette liste.

La liste des marqueurs est également utilisée pour détecter un nouveau marqueur au moment de son insertion. Par exemple,

- Essayez d'insérer le marqueur **\xx**.

Vous verrez apparaître une boîte disant: *Ce marqueur n'est pas dans la liste des marqueurs. L'ajouter ??*

- Choisissez No.

Remarquez que le curseur reste dans la boîte d'édition du marqueur vous permettant de corriger **\xx** en autre chose.

- Changer le marqueur en **yy** et tapez à nouveau une espace.

Vous verrez apparaître à nouveau la boîte disant: *Ce marqueur n'est pas dans la liste des marqueurs. L'ajouter ??*

- Choisissez Oui pour ajouter le marqueur à la liste.

Vous verrez alors la boîte de dialogue pour les propriétés du marqueur **yy**.

- Donnez-lui le nom de champ *provisoire*.
- Cliquez sur OK.

Vous avez maintenant le champ **\yy** dans l'enregistrement.

La liste des marqueurs

Une table des marqueurs comportant plus d'informations que dans la boîte déroulante précédente, est accessible dans la boîte de dialogue des *Marqueurs*.

- Choisissez Base de données, Propriétés.

Vous verrez apparaître une boîte intitulée *Propriétés du Type de Base de Données* présentant une liste complète de tous les marqueurs utilisés dans le fichier.

Remarquez que le marqueur qui est en surbrillance est celui sur lequel se trouvait le curseur au moment où vous avez appelé la boîte de dialogue des Marqueurs.

Notez également que certaines des lignes de marqueur sont en gras et d'autres en normal. Ceux qui ne sont pas en gras ne sont utilisés dans aucun enregistrement de la base de données. La liste inclut tous les marqueurs du progiciel MDF (Multi-Dictionary Formatter), mais cette base de données ne les utilise pas tous. D'avoir la liste complète de tous les marqueurs vous permet de vous souvenir des marqueurs valides lorsque vous voulez vous restreindre à un jeu standard, comme celui de MDF ou tout autre jeu standard. (Le jeu de champs MDF n'est qu'une option. Vous n'êtes pas tenu d'utiliser cette liste pour votre lexique.)

Remarquez le jeu de boutons sur la droite de la liste des marqueurs. Le bouton Supprimer (Supprimer) est grisé. Vous ne pouvez pas effacer un marqueur qui est déjà utilisé dans un enregistrement.

- Cliquez sur OK pour revenir à la base de données.
- Supprimez le marqueur **\yy** dans l'enregistrement (utilisez la touche Supp_arrière, ou bien cliquez une fois sur le marqueur de champ et tapez la touche Suppr).
- Choisissez Base de données, Propriétés.

Comme le curseur est en ce moment sur un marqueur différent, c'est ce marqueur qui sera en surbrillance.

- Sélectionnez le marqueur **\yy**.

Constatez que la ligne de ce marqueur n'est pas en gras. Cela signifie qu'aucun enregistrement n'utilise actuellement ce marqueur. Lorsque vous cliquez sur la ligne **\yy**, le bouton Supprimer n'est plus grisé.

- Supprimez le marqueur **\yy** de la liste en cliquant sur le bouton Supprimer.
- Cliquez sur OK.

Remplacer un Marqueur

Pour remplacer un marqueur, on tape Supp_arrière, comme précédemment pour supprimer un marqueur.

- Placez le curseur derrière le mot *listen*.
- Tapez Supp_arrière. La boîte des marqueurs s'ouvre avec **ge** en surbrillance.
- Tapez **gn** suivi d'un espace. Le marqueur est remplacé.
- Remettez le marqueur d'origine **ge**.

Remplacer un marqueur partout

Un marqueur peut être remplacé partout. Ceci doit être manié avec précaution, car le changement n'affectera que les bases (du même type) actuellement chargées, et il se peut que vous ayez un autre fichier du même type non-ouvert, néanmoins cela peut se faire. Par exemple, nous pourrions vouloir remplacer le marqueur **lge** en **leng** (pour English). Pour cela,

- Choisissez Base de données, Propriétés.
- Sélectionnez le marqueur **ge** et cliquez sur Modifier (ou bien double-cliquez sur la ligne **ge**).

(Il s'agit de la même boîte de dialogue que celle obtenue précédemment par **w:Clic-droit** ou **m:Option+clic** sur la marqueur de la fenêtre de texte.)

- Cliquez dans la boîte d'édition *Marqueur* et remplacez le marqueur **ge** par **eng**.
- Cliquez sur OK.

Vous allez voir apparaître une boîte de dialogue demandant si vous voulez remplacer le marqueur **\ge** par **\eng** partout où il apparaît.

- Choisissez Oui.
- Cliquez sur OK dans la boîte *Propriétés du Type de Base de Données*.

Comme nous ne voulons pas remplacer définitivement ce marqueur, nous allons le restituer comme à l'origine.

- Remplacez le marqueur **\eng** par **\ge** partout.

Fin de l'étape

Cet apprentissage est organisé en une série d'étapes. Chaque étape démarre avec un nouvel échantillon de fichiers. Ainsi si vous avez un problème avec une section quelconque de cet apprentissage, vous pouvez démarrer à zéro la section suivante. Il est également possible de refaire la section précédente. Mais pour revenir en arrière, vous devrez réinstaller les fichiers d'exemples pour les restaurer. (Lorsque vous faites une leçon, les fichiers d'origine s'en trouvent modifiés ; ceci signifie que si vous essayez de refaire une section, les choses qu'on vous demande de faire auront déjà été faites.) Pour restaurer l'état initial des exemples, lancez l'installation de Shoebox et choisissez uniquement celle des exemples. La même technique sera utilisée pour restaurer les exemples lorsqu'une personne aura achevé l'apprentissage sur votre système.

C'était donc la première étape. A ce point, vous avez annulé tous les changements effectués, aussi si vous désirez recommencer cet apprentissage une prochaine fois, vous pouvez le faire simplement en relançant le projet SHLEX.PRJ.

NOTE: S'il se trouve des données dans ce projet en gras, en italique ou en couleur, c'est que la section suivante a déjà été faite. Dans ce cas, vous devez réinstaller les exemples.

Marqueur et Encodage de langue

Les données sont saisies dans une langue. Dans un dictionnaire (multilingue), les différents champs contiendront des données dans des langues différentes.¹¹ Dans Shoebox, une information relative à la langue peut être associée à chaque marqueur. Cette information recouvre l'ordre de tri alphabétique, la police par défaut, et le clavier associé. (Il s'agit des points les plus importants dans le cadre de cet exposé). Dans Shoebox, ce jeu d'informations sur une langue est ce qu'on appelle "l'encodage de la langue".¹²

Un encodage de langue est associé à un marqueur par l'intermédiaire de la boîte de dialogue *Propriétés des Marqueurs*.

- w**:Clic-droit or **m**:Option+clic sur le marqueur **lx** pour faire apparaître la boîte de dialogue *Propriétés des Marqueurs*.

Sur la gauche de la boîte de dialogue, vous verrez l'étiquette *Encodage de langue*. A côté d'elle se trouve une boîte déroulante contenant le nom de l'encodage de langue associé au marqueur **lx**. Elle devrait afficher *vernacular*. (Nous préférons utiliser des noms génériques dans la base SHLEX.DB.)

- Cliquez sur la petite flèche pour voir les autres encodages de langues disponibles.
- Fermez la boîte de dialogue *Propriétés des Marqueurs* en choisissant Annuler.
- w**:Clic-droit ou **m**:Option+clic sur le marqueur **ps** pour faire apparaître la boîte de dialogue *Propriétés des Marqueurs*.

Ce marqueur devrait être affecté de l'encodage de langue *English*.

- Fermez la boîte de dialogue *Propriétés des Marqueurs* en choisissant Annuler.

¹¹ Il est parfois nécessaire de pouvoir changer de langue à l'intérieur même de certains champs. C'est un problème autre qui est discuté en réponse à l'une des questions dans le chapitre *Reference*.

¹² L'Encodage des langues est abordé dans le chapitre *Démarrer un nouveau projet et base de données*, ainsi que dans le chapitre *Importer de Shoebox 2.0*.

Langue et police de caractères

Maintenant voyons de plus près les polices de caractères. Il y a deux types d'association marqueur-police dans Shoebox¹³. Tous les champs peuvent se voir affecter la même police. C'est la relation la plus courante et la plus recommandée. Cependant il est possible d'associer une police donnée à un champ particulier.

Traditionnellement, un dictionnaire multilingue présente la langue source (celle du lexème) en gras. Dans notre exemple il s'agit de la langue "vernaculaire". Pour changer la police de la langue vernaculaire en gras,

Choisissez Projet, Encodage de langues.

Vous verrez une liste des encodages de langues. Elles ne sont pas intégrées dans le logiciel, vous pourrez ainsi créer les vôtres lorsque vous démarrerez avec vos propres données.¹⁴

Sélectionnez *vernacular*.

Choisissez Modifier.

Ceci vous conduit à une boîte de dialogue à onglets intitulée *Propriétés de l'Encodage de langue - vernacular*. Les différents onglets permettent d'accéder à toutes les caractéristiques de l'encodage de la langue. La police standard pour les données dans cette langue se trouve définie dans l'onglet Options.

Choisissez l'onglet Options.

Cliquez sur Choisir la Police.

Comme style de police, choisissez Gras.

Faites OK deux fois, puis choisissez Fermer dans la boîte *Encodage de langues*.

Remarquez que maintenant tous les champs en vernaculaire sont en gras. Vous vous en apercevrez mieux en faisant dérouler l'enregistrement vers le bas de façon à voir les phrase illustratives étiquetées **lxv**.

Un moyen pratique d'accéder à l'encodage d'une langue consiste à demander les propriétés d'un marqueur dans cette langue. Nous allons utiliser cette approche pour mettre les champs de langue nationale en italique.

w:Clic-droit ou **m**:Option+clic sur le marqueur **gn** pour faire apparaître la boîte de dialogue *Propriétés des Marqueurs*. Elle a l'encodage de langue *national*.

Cliquez sur le bouton Propriétés de la langue pour faire apparaître les propriétés de l'encodage de langue *national*.

Choisissez l'onglet Options.

Cliquez sur Choisir la Police.

Comme style de police, choisissez Italique.

Faites OK jusqu'à ce que toutes les boîtes de dialogues soient fermées.

Si vous déroulez l'enregistrement, vous verrez que tous les champs en langue nationale sont maintenant en italique.

Police spécifique pour un champ

Il peut être intéressant d'associer une police spéciale (ou à un enrichissement de police) à un champ particulier, pour rendre ses données plus visibles. Par exemple, vous pourriez vouloir que les lexèmes apparaissent en caractères plus grands que les autres pour pouvoir les lire d'un rapide coup d'œil.

Faites un **w**:Clic-droit ou **m**:Option+clic sur le marqueur **lx** pour faire apparaître la boîte de dialogue *Propriétés des Marqueurs*.

¹³ En fait, il y en a trois. Le menu *Affichage, Polices des marqueurs* vous permet de modifier la police dans laquelle les marqueurs sont affichés. Cette fonctionnalité n'est pas abordée dans ce document.

¹⁴ Un avantage à conserver tous les fichiers de projets dans un même dossier, c'est qu'ils peuvent ainsi partager les divers encodages de langue que vous aurez définis—vous n'aurez pas à les recréer à chaque nouveaux projets.

Sur la droite de la boîte Propriétés des Marqueurs se trouve une case à cocher intitulée *Utiliser police d'encodage*. Elle sert à dire à Shoebox d'utiliser la police par défaut associée à la langue. Si vous voulez autre chose que cette police par défaut, vous devez invalider cette option et spécifier la police que vous voulez.

Cliquez sur la case à cocher Utiliser police d'encodage pour la désactiver.

Cliquez sur le bouton Choisir la Police.

Cela fait apparaître la boîte de dialogue Police.

Changez la taille en 16 points.

Faites OK, OK.

Vous verrez le lexème s'afficher dans un corps bien plus grand que les autres champs.

Faites revenir le lexème à la police par défaut de la langue.¹⁵

Il est également possible d'utiliser le souligné ou la couleur pour mettre en évidence le contenu d'un champ particulier.

Maintenant, nous allons parcourir la base de données. Tout d'abord, remarquez la série de boutons comportant des flèches dans la barre d'outils (nous en avons déjà parlé). L'une pointe à gauche, l'autre à droite et deux autres pointent sur des barres verticales. Elles servent à se déplacer respectivement, vers l'enregistrement *Précédent*, l'enregistrement *Suivant*, le *Premier* enregistrement et le *Dernier* enregistrement. Shoebox n'affiche normalement qu'un enregistrement à la fois dans une fenêtre, et il s'agit là de moyens pratiques pour changer l'enregistrement en cours. (Ces commandes sont également accessibles par le menu Base de Données.)

Retournez au premier enregistrement de la base de données, si vous vous étiez promener ailleurs (utilisez le bouton Premier enregistrement).

Faites Suivant deux fois pour aller au troisième enregistrement (le second enregistrement -a).

Localisez la catégorie grammaticale nationale (champ `\pn`).

Ce n'est pas facile à trouver dans cette jungle de marqueurs n'est-ce pas ? (c'est à la quatrième ligne, si vous ne l'avez pas encore trouvé.)

Faites un **w**:Clic-droit ou **m**:Option+clic sur le marqueur **pn**.

Cliquez dans la case à cocher Utiliser police d'encodage pour la désactiver.

Cliquez sur Choisir la Police.

Localisez la boîte intitulée *Couleur*. La couleur en cours est le noir.

Cliquez sur le triangle à côté de la couleur.

Choisissez Rouge.¹⁶

Cliquez sur OK, OK.

Remarquez comment le mot "vous saute aux yeux" maintenant. Déplacez vous en avant de plusieurs enregistrements (**w**:Alt+N ou **m**:commande+]) et voyez comme il est aisé de voir si cet enregistrement particulier comporte ou non une catégorie grammaticale. (Certains enregistrements comporte le marqueur mais pas de contenu.)

Ceci est très efficace, particulièrement comme moyen provisoire de localiser un champ sur lequel vous vous focalisez. Lorsque ce champ ne fera plus l'objet d'une attention particulière, vous pourrez le remettre dans la police standard de la langue. Vous pouvez aussi choisir de mettre en relief certains champs de façon permanente, par exemple de

¹⁵ Pour cela, faites **w**:clic-droit ou **m**:option+clic sur `\lx` puis cliquez sur *Utiliser police d'encodage*.

¹⁶ Si vous avez un moniteur monochrome, choisissez des couleurs qui donnent le plus de contraste, ou faites autrement pour apporter du relief, comme souligner, augmenter la taille des caractères ou choisir une police différente.

marquer les champs de sous-entrées pour qu'ils soient bien en évidence—comme un titre de chapitre dans un livre.

Tri, Index et Hiérarchie des marqueurs

Tri et vues multiples

Shoebox peut afficher plusieurs représentations de la même base de données dans différentes fenêtres, leur contenu étant trié suivant des champs différents. Par exemple,

- Revenez au début de la base de données (cliquez sur le bouton Premier enregistrement Record).
- Choisissez Fenêtre, Dupliquer.

Vous verrez apparaître une nouvelle fenêtre intitulée SHLEX.DB:2, présentant le même enregistrement.¹⁷

- Choisissez Fenêtre, mosaïque horizontale.

Ce choix de présentation est pratique pour voir plusieurs enregistrements en même temps. Remarquez que si vous éditez l'enregistrement en cours, il change dans les deux fenêtres. Remarquez également que vous pouvez passer à un autre enregistrement dans une fenêtre, sans que l'autre ne change.

- Si la fenêtre 1 est à droite de la fenêtre 2, cliquez n'importe où dans la fenêtre 1 puis refaites mosaïque horizontale, cela la fera passer à gauche. (Il est plus naturel d'avoir la fenêtre 1 à gauche, aussi pour la suite nous nous référerons aux fenêtres dans cet ordre-là.)

Maintenant, c'est là que c'est génial ! Nous allons trier la fenêtre 2 par rapport au champ *glose* alors que la fenêtre 1 restera triée par le *lexème*.

- Cliquez n'importe où dans la fenêtre 2.
- Choisissez Base de Données, Trier.

Remarquez que le tri est relatif au champ *\lx* (le marqueur d'enregistrement).

- Cliquez sur Vider.
- Sélectionnez *ge* dans la liste *Marqueurs disponibles*.
- Cliquez sur le bouton Premier pour le déplacer dans la liste *Champs de tri*.

Comme résultat, vous devriez avoir *ge* dans la liste *Champs de tri*. (Au fait, vous pouvez directement déplacer les items d'une liste vers l'autre en double-cliquant dessus.)

- Cliquez sur OK.
- Faites Suivant et Précédent plusieurs fois pour constater que la fenêtre 2 est maintenant triée suivant le contenu du champ *glose* anglaise.

Remarquez aussi que lorsque vous vous déplacez vers un nouvel enregistrement le curseur se positionne maintenant au début du champ *glose* (et non plus au niveau du marqueur d'enregistrement comme avant). Ceci est dû à ce que le curseur se place toujours au début du *champ de tri primaire*, lors d'un déplacement vers un nouvel enregistrement. Remarquez également que la barre d'état affiche maintenant le marqueur d'enregistrement et aussi le champ de tri.

Indexation multiple d'un enregistrement

Dans notre exemple, la base de données dans la fenêtre 2 est triée suivant le champ *glose*, mais il peut y avoir plusieurs champs *gloses* pour un même *lexème*. Que fait donc Shoebox lorsqu'il y a plusieurs occurrences du champ de tri primaire ? Shoebox affiche alors un enregistrement pour chacune des occurrences du champ de tri primaire. C'est un peu

¹⁷ Si la fenêtre SHLEX est grande ouverte, la seconde fenêtre apparaîtra directement par dessus. Il est difficile de s'apercevoir qu'elle est là, mais la barre en haut de la fenêtre devrait afficher SHLEX.DB:2.

difficile à décrire sans illustration. Commençons par positionner la base sur le bon enregistrement.

- Cliquez dans la fenêtre 2 (celle triée relativement au champ **ge**).
- Choisissez Base de Données, Premier enregistrement (Top) ou cliquez sur le bouton Premier enregistrement dans la barre des outils.

Le curseur va se positionner dans l'enregistrement dont le marqueur d'enregistrement contient *ebnomatruma*, mais il se place près du mot *ancestors* dans la champ **ge**. Ceci parce que la base de données est triée maintenant par rapport au champ **ge** et que le premier mot est *ancestors*. Remarquez que le curseur se place toujours au début du *champ de tri primaire* lors d'un déplacement vers un autre enregistrement. Remarquez également que la barre d'état affiche le marqueur d'enregistrement `\lx ebnomatruma` et le champ de tri primaire en cours `\ge ancestors`. La barre d'état signale aussi qu'il s'agit de l'enregistrement 1/100. Vous vous souvenez qu'il n'y avait que 57 enregistrements lorsque vous avez démarré—y en a-t-il eu de nouveaux créés ? Non—mais il y a 100 champs **ge**. Le second chiffre ne décompte pas ici le nombre d'enregistrements, mais plutôt le nombre d'occurrences du champ de tri.

Lorsque le champ de tri primaire apparaît plusieurs fois dans un enregistrement donné, cet enregistrement-là apparaîtra en différents endroits dans cette représentation de la base. Nous allons nous déplacer dans l'enregistrement *ahwahw*, qui contient quatre gloses anglaises.

- Toujours dans la fenêtre 2, déplacez-vous de quatre enregistrements (cliquez quatre fois sur le bouton Suivant).

Le curseur se positionne devant *bad eyes*, qui est une des gloses anglaises de l'enregistrement *ahwahw* (cette glose correspond à la sous-entrée *mata ahwahw*). Le curseur est peut-être trop loin dans le texte pour que vous puissiez voir *le marqueur d'enregistrement*, mais vous pouvez le vérifier en regardant dans la barre d'état (c.à.d. "**\lx ahwahw**").

- Déplacez-vous au début de l'enregistrement (utilisez l'ascenseur ou tapez Ctrl+Début).

Remarquez que cet enregistrement contient également les gloses *dusty*, *dirty*, et *cloudy*.

- Déplacez-vous neuf enregistrements plus loin dans la glose *cloudy* (qui figure dans la même mot Selaru *ahwahw*).

En vous déplaçant six enregistrements plus loin, toujours en avant, vous devriez arriver à *dirty* et encore sept enregistrements plus loin à *dusty*. Ils constituent tous le *même enregistrement*, vous le retrouvez identique ici et là parce que le champ **ge** apparaît plusieurs fois dans cet enregistrement. Cet affichage de la base de données triée par la glose est très utile pour un travail de références, car cela vous donne un accès aux données inverses rapide pour ces mots difficiles à retenir que vous "savez être là, quelque part."

Cet affichage produit l'équivalent de l'*index inverse* d'une base de données¹⁸ (basée sur le champ **ge**), mais il ne s'agit pas d'un véritable index. Dans un véritable index, les champs sont réordonnés, le champ glose prend la place de champ lexème, etc. Mais même comme ça, la capacité multi-tri de Shoebox présente un avantage sur d'autres processus d'inversion véritable, en ce que cet affichage, tout en étant trié différemment de la base de données d'origine, reste liée à la base d'origine (en fait, *c'est* la base de données d'origine), ainsi toute modification faite dans l'une se répercute dans l'autre.

Tri sur plusieurs champs

Il est également possible de trier suivant plusieurs champs. Par exemple, nous pouvons trier par catégorie grammaticale, puis par lexème. Pour cela,

- Cliquez n'importe où dans la fenêtre 2.

¹⁸ Dans un dictionnaire bilingue, disons Français-Anglais, les mots français sont donnés avec des définitions en anglais. Un dictionnaire inverse Anglais-Français donnerait des mots anglais avec des définitions en français.

- Choisissez Base de Données, Trier.
- Vider la case *Champs de tri*. Déplacez **ps** dans *Champs de tri* en le sélectionnant et en cliquant Premier.
- Déplacez **lx** comme deuxième champ de tri en le sélectionnant et en cliquant Dernier.

Le résultat est **ps** puis **lx** comme *Champs de tri*.

- Cliquez sur OK.
- Allez au premier enregistrement de la base de données (cliquez sur Premier enregistrement).

Remarquez que cela vous conduit dans l'enregistrement *babaf*. Il s'agit d'un adjectif et par conséquent il est classé en premier.

- Allez 7 enregistrements plus loin au début des *noms*.

Remarquez que le tri se fait d'abord par la catégorie grammaticale, puis par le lexème, à l'intérieur d'une même catégorie grammaticale. Par exemple, les adjectifs sont classés avec *babaf* en premier, puis *babah*. Les premiers noms que vous avez sont *abat*, *abu*, *adar*.

- Choisissez Base de Données, Trier.
- Enlevez les marqueurs **\ps** et **\lx** de la liste *Champs de tri* puis remplacez le marqueur **ge** pour que la base de données soit classée de nouveau suivant ce champ.
- Cliquez sur OK.

Tri par la droite

L'étude des suffixes et de la structure des mots peut être facilitée par le classement des mots de la droite vers la gauche pour en repérer les schèmes.

- Cliquez dans la fenêtre 1 (triée suivant le champ **lx**)
- Allez dans le premier enregistrement de la base.
- Déplacez-vous de 4 enregistrements, jusqu'à *abat*.

Remarquez que la barre d'état mentionne 5/57.

- Choisissez Base de Données, Trier.
- Validez la case de droite Trier premier champ à partir de la fin.
- Cliquez sur OK.

Remarquez que maintenant la barre d'état mentionne 39/57. La raison en est que *abat* est maintenant triée en commençant par la lettre **t** de la fin au lieu de la lettre **a** du début. (Notez que ce n'est pas la bonne méthode pour classer des mots en écriture droite-gauche. De telles écritures sont généralement stockées de façon interne dans l'ordre de prononciation, mais sont inversées à l'affichage, ce qui fait que le tri normal, gauche-droite, s'applique correctement.)

- Déplacez-vous quelques enregistrements plus loin pour vérifier que vous êtes dans les lexèmes se terminant par la lettre *t*.
- Revenez à *abat*.

Il est toutefois assez malaisé d'avoir une bonne idée des schèmes des mots en passant d'un enregistrement à l'autre. Le mode d'affichage *Index* décrit dans la section suivante est plus approprié.

Index

Pour voir les choses différemment,

- Choisissez Index dans le menu Affichage.

Cela vous place en mode Index, qui vous présente un enregistrement par ligne. Vous pouvez voir maintenant que les enregistrements sont ordonnés par les fins de mots. Remarquez également que les mots sont alignés sur la marge de droite. Ceci parce que les mots sont triés par la fin ; cet alignement permet de repérer facilement les schèmes recherchés.

- Choisissez Base de Données, Trier.

- Cliquez dans la case Trier premier champ à partir de la fin pour la désactiver.
- Cliquez sur OK.

L'affichage devrait avoir changé et *abat* devrait être revenu à sa place alphabétique normale. Remarquez que les mots sont maintenant alignés normalement (pour cette langue).

- Choisissez Affichage, Champs affichés dans l'Index.

Ceci fait apparaître la boîte de dialogue permettant de choisir les champs à afficher dans l'Index. Cela fonctionne comme la boîte de dialogue des *Champs de tri*. Au départ, seul le champ \lx, le marqueur d'enregistrement, est affiché dans la fenêtre de l'Index.

- Ajoutez les champs \ge puis \xv après le champ \lx.
- Cliquez sur OK.

Vous devriez avoir maintenant 3 champs: le lexème, la glose anglaise et l'exemple vernaculaire.

- Dans l'Index, tirez vers la gauche la ligne de séparation entre les deux premiers champs, jusqu'à ce qu'elle soit proche de la fin des entrées lexicales.

De la même façon vous pouvez tirer la ligne de séparation entre la glose anglaise et l'exemple vernaculaire vers la gauche ou la droite.

Il y a trois façons d'entrer et de sortir de la fenêtre de l'Index : par le menu Affichage, Index, par le raccourci clavier, et par le bouton Index dans la barre d'outils. (Il s'agit du bouton comportant une sorte de grille comme dessin, juste avant la boîte de Recherche.) De plus, depuis la fenêtre d'Index, taper sur **w**:Entrée ou **m**:Retour vous fera passer à l'affichage normal de l'enregistrement sélectionné dans l'Index.

- Sélectionnez la fenêtre 2.
- Dans le menu Affichage, choisissez Champs affichés dans l'Index.
- Insérez le champ **ge** à la fin de la liste des champs à afficher.
- Cliquez sur OK.

Maintenant, on devrait mieux voir que la fenêtre est triée par rapport au champ de glose anglaise. En déroulant la fenêtre vers le haut et vers le bas, on reconnaît les enregistrements comportant des champs multiples (car ils ont plusieurs gloses).

- Cliquez sur le bouton Premier enregistrement dans chacune des fenêtres, afin que les premiers enregistrements s'affichent dans chaque fenêtre.

Remarquez que la fenêtre triée par la glose affiche l'enregistrement *ebnomatruma* avec comme première glose *ancestors*, alors que la fenêtre triée par le lexème affiche le premier enregistrement *a* (avec comme glose *listen*).

Conserver ouverte une fenêtre Glose anglaise de votre lexique permet de disposer d'une sorte de lexique inverse en ligne.

- Revenez pour chacune des fenêtres à un affichage normal des enregistrements (en cliquant sur le bouton Affichage Index pour chacune des fenêtres).

Hiérarchie des marqueurs

La hiérarchie des marqueurs est un moyen d'indiquer la relation structurelle entre les différents champs d'un enregistrement. La hiérarchie est utile dans le mode d'affichage Index, et pour l'interalignement.

- Comme la base de données SHLEX a déjà une hiérarchie établie, et qu'elle comporte trop de marqueurs à reconfigurer pour voir ce qui se passe sans elle, nous allons illustrer la hiérarchie à travers une autre base de données.
- Choisissez Fichier, Ouvrir
- Naviguez dans le répertoire SAMP-SET si vous n'y êtes pas déjà
- Ouvrez la base de données *HIERARCH.DB*.

- Tirez le haut de la fenêtre HIERARCH.DB pour faire apparaître plus de données

Elle comporte l'unique enregistrement :

```

\lx escape
\ps n
\ge avoidance of reality
\ps adj
\ge providing an escape
\ps vt
\ge to elude
\ps vi
\ge to get away

```

(Notez que bien que ce fichier utilise des marqueurs ressemblant à ceux de SHLEX.DB, il est en fait lié à un autre type de base de données qui lui n'a pas de hiérarchie.)

Nous allons voir ces données en utilisant trois paramétrages différents.

- Choisissez Affichage, Champs affichés dans l'Index.
- Spécifiez les champs **ge** et **ps** pour l'Index. (Double-cliquez sur les champs **ge** et **ps** pour les déplacer dans la colonne de droite; double-cliquez sur le champ **lx** dans la colonne de droite pour l'enlever de la colonne de droite.)
- Cliquez sur OK

Une seule ligne s'affiche. Ceci parce que la base de données ne comporte qu'un enregistrement, et que l'affichage est trié par le champ **lx** (même si ce champ n'est pas affiché). Remarquez que l'Index n'affiche que la première occurrence des champs **ge** et **ps** de l'enregistrement.

- Choisissez Base de Données, Trier
- Spécifiez le champ **ge** comme seul champ de tri. (Double-cliquez sur **ge** pour le déplacer dans la colonne de droite; double-cliquez sur **lx** pour l'enlever de là.)
- Cliquez sur OK

Vous voyez maintenant s'afficher :

avoidance of reality	n
providing an escape	n
to elude	n
to get away	n

C'est mieux, sauf qu'ils affichent tous la même catégorie grammaticale! Ce résultat aberrant résulte du fait que la programme ne sait pas qu'il y a une relation entre les champs **ps** et **ge**. Shoebox se focalise sur le champ de glose, lequel constitue la clé de tri primaire, mais il se contente d'inclure le premier champ **ps** rencontré.

Le même résultat serait obtenu avec un tri par les deux champs **ps** et **ge**. (Si vous ne trie *que* par le champ **ps**, vous verrez toutes les catégories grammaticales, mais avec la même glose pour chacune d'elles!)

Maintenant nous allons configurer la hiérarchie.

- w**:Clic-droit or **m**:Option+clic sur le champ **ps** dans la zone marqueur.

Par défaut, Shoebox place les différents marqueurs d'un enregistrement hiérarchiquement en dessous du marqueur d'enregistrement. Aussi, dans cet exemple les champs **ps** et **ge** ont déjà été placés sous le champ **lx**. Pour le vérifier :

- w**:Clic-droit or **m**:Option+clic sur le champ **ps** dans la zone marqueur.

Ceci ouvre la boîte *Propriétés des Marqueurs* pour le marqueur **ps**.

- Constatez la présence de **lx** dans la case intitulée *Sous quoi dans la hiérarchie*.
- Cliquez sur OK

Maintenant changeons la hiérarchie par défaut en précisant, dans *Propriétés des Marqueurs*, que le marqueur **ge** soit dorénavant en dessous du champ **ps** :

- w**:Clic-droit ou **m**:Option+clic sur le champ **\ge** dans la zone marqueur pour faire apparaître les *Propriétés des Marqueurs* du marqueur **ge**
- Dans la case intitulée *Sous quoi dans la hiérarchie*, remplacez le champ **lx** par **ps**.
- Cliquez sur OK

Ceci définit une hiérarchie simple dans la base de données. Pour voir les marqueurs avec leur hiérarchie :

- Choisissez Affichage, Hiérarchie des marqueurs.

Ceci devrait amener à l'affichage de l'enregistrement suivant :

```

\lx      escape
. \ps    n
. . \ge  avoidance of reality
. \ps    adj
. . \ge  providing an escape
. \ps    vt
. . \ge  to elude
. \ps    vi
. . \ge  to get away

```

Remarquez que les marqueurs sont maintenant renfoncés de différentes profondeurs. Cela ressemble à une présentation de plan, et montre graphiquement quel champ est en dessous de l'autre. Cette représentation vous permet de vérifier si la hiérarchie est bien configurée. Il n'est pas nécessaire de l'afficher pour qu'elle soit opérationnelle.

- Choisissez Affichage, Index (ou cliquez sur le bouton Index).

L'Index devrait vous afficher :

avoidance of reality		n
providing an escape		adj
to elude		vt
to get away		vi

Grâce à la hiérarchisation, le programme peut maintenant savoir que les différentes gloses appartiennent à différentes catégories grammaticales, et cette information se reflète dans l'Index.

Comme les bases de données comportent généralement plus de trois champs, il vaut mieux, si vous établissez une hiérarchie, le faire pour tous les champs de la base de données une fois pour toute.

- Fermez la base de données. (Choisissez Fichier, Fermer.)

Rechercher, Retrouver, Sauter à, Filtrer

Rechercher

Rechercher est un moyen rapide de parvenir à un enregistrement particulier. *Rechercher* remplacera l'affichage de l'enregistrement en cours par celui recherché. Avec *Rechercher*, vous ne pouvez rechercher des données que dans le champ de tri primaire. La séquence que vous recherchez doit correspondre au début (ou la totalité) du contenu du champ de tri primaire.¹⁹

¹⁹ Si le champ de tri primaire trie par la fin, alors la séquence est également recherchée par la fin, par exemple 'ma' retrouverait *ebnomatruma* (si la recherche ne porte pas sur le contenu entier du champ).

- Cliquez quelque part dans la fenêtre 1 (sortez du mode Index si vous y êtes).²⁰
- Choisissez Base de Données, Rechercher ou le raccourci clavier (voir l'Aide pour trouver la liste des raccourcis).
- Tapez le mot *abat* puis cliquez sur OK.

Remarquez que l'enregistrement est retrouvé immédiatement. Un index sur le champ de tri est créé pour chaque fenêtre, ce qui explique la vitesse de recherche.

Nous pouvons aussi rechercher dans une base de données triée par un autre champ que le champ d'enregistrement.

- Cliquez dans la fenêtre 2, qui est triée par rapport au champ **\ge** (sortez du mode l'Index si vous y êtes).
- Choisissez Base de Données, Rechercher.

Remarquez que la boîte de dialogue *Rechercher* indique qu'elle recherche dans le champ **\ge**.

- Rechercher *dog*.

Vous devriez aboutir à l'enregistrement *asw*, avec le curseur devant *dog*.

- Rechercher *dip*.

Une boîte de dialogue précisant que *dip* and *dipper* sont des correspondances possibles. (Si nous avons choisi l'option *Champ entier* alors seul *dip* aurait été trouvé.)

La boîte Plusieurs Résultats comporte une option permettant d'afficher un autre champ en plus de celui recherché. Par défaut, le champ **\ge** est choisi ici.

- Cliquez dans la liste déroulante à droite de la boîte de dialogue Plusieurs Résultats, sous l'étiquette *Champ Additionnel*.
- Tapez *lx*, ou sélectionnez **lx** dans la liste déroulante.

Le programme affiche maintenant le contenu du champ **lx** en face des correspondances trouvées. Ceci peut vous aider à choisir entre *dip* or *dipper*. (Sélectionnez *dip* et cliquez sur OK.)

Retrouver Alors que *Rechercher* ne recherche des données que dans le champ de tri primaire, *Retrouver* peut retrouver des données *n'importe où* dans un enregistrement.

- Allez dans la fenêtre 1— celle triée par le champ **\lx**.
- Déplacez-vous dans le premier enregistrement de la base de données.
- Choisissez Edition, Retrouver.
- Tapez la seule lettre *a* dans la case *Séquence à retrouver* (mais ne lancez pas encore la commande Retrouver).

Bien sûr, vous pouvez imaginer le nombre de fois qu'une lettre seule peut être trouvée, particulièrement une lettre aussi commune que *a*. Il y a plusieurs façons de réduire le nombre de correspondances non souhaitées.

Remarquez qu'il y a de nombreuses options regroupées en deux colonnes. Juste en dessous de la case *Séquence à retrouver*, il y a un groupe d'options intitulé *Trouver dans*. Ici vous pouvez choisir de limiter la recherche par champ ou par langue.

- Choisissez Tout champ de langue, et sélectionnez *vernacular* dans la liste déroulante.

Remarquez que vous pouvez limiter la recherche en la restreignant à l'enregistrement en cours ou au contraire rechercher dans toute la base de données.

²⁰ On suppose ici la même configuration de fenêtres que dans la section du tri. Si vous avez sauté cette section, veuillez suivre les étapes au début de *Tri sur plusieurs champs*, p.22 pour configurer deux fenêtres du même fichier SHLEX.DB, la première triée par rapport au champ (**lx**) et la seconde par rapport au champ de glose anglaise (**ge**).

- Choisissez La base entière (filtrée).

Maintenant remarquez qu'à droite de la case *Trouver dans*, en haut de la seconde colonne d'options, il y a une case à cocher intitulée *Mot complet*.

- Choisissez Mot complet. Un X ou une coche devrait apparaître dans la case. Ceci évitera la recherche de *a* à l'intérieur des mots.

En dessous de cette case, il y a encore une série d'options, intitulé *Correspondance des caractères*.

- Choisissez l'option Même ceux qui sont ordin. ignorés. (Le tiret est un des caractères normalement ignorés. Choisir cette option empêchera la commande Retrouver de « voir » les tirets, ce qui évitera l'affichage des entrées *-a*.)
- Cliquez sur OK. Le *a* en haut de l'enregistrement sera mis en surbrillance pour montrer qu'il a été trouvé.
- Choisissez Edition, Retrouver le suivant (ou bien utilisez le raccourci clavier).

Remarquez qu'au lieu de s'arrêter sur tous les *a* de son chemin, le programme saute à un *a* dans une des phrases illustratives. Observez qu'il se trouve dans le champ `\xv` qui est un champ en langue vernaculaire, comme spécifié dans la requête.

A présent regardez la barre d'outils, juste à droite du bouton `Index`. Vous verrez une petite case de texte qui contient maintenant la lettre *a*. Cette case comporte une liste déroulante des dernières séquences pour lesquelles vous avez effectué une commande *Retrouver*— et qui est vide actuellement, à part le *a*. A sa droite se trouve une paire de jumelles avec une flèche pointant vers la droite. Il s'agit du bouton *Retrouver le suivant*.

- Cliquez sur le bouton *Retrouver le suivant*.

Il y aura une pause pendant que *Retrouver* examine les données. Il s'agit d'une recherche séquentielle, le programme parcourt rapidement chaque enregistrement de la base de données, et cela peut prendre un certain temps s'il n'y a qu'un petit nombre d'occurrences dans une grande base de données. Il devrait s'arrêter dans le champ `\xv` de l'enregistrement `\x baba`. Remarquez que la traduction anglaise dans le champ suivant contient également le mot (anglais) *a*. La prochaine fois que vous demanderez *Retrouver le suivant* (ci-dessous), le programme *ne s'arrêtera pas* sur ce *a* parce qu'il n'est pas dans la langue demandée:

- Cliquez sur le bouton *Retrouver le suivant*.

Une nouvelle pause se fait sentir pendant que *Retrouver* recherche plus avant dans les données, pour s'arrêter finalement à l'entrée *-he tel*. La fenêtre s'est déroulée horizontalement pour montrer l'enregistrement.

Juste après le bouton *Retrouver le suivant*, se trouve une autre paire de jumelle avec une flèche pointant vers la gauche. Il s'agit du bouton *Retrouver le précédent*.

- Cliquez sur le bouton *Retrouver le précédent*. Vous verrez de nouveau l'entrée *baba*.

Sauter vers

La commande *Sauter vers* est encore une autre façon de retrouver des données. *Sauter vers* affichera dans une nouvelle fenêtre l'enregistrement correspondant aux critères que vous avez spécifiés. *Sauter vers* se comporte comme *Rechercher* dans la mesure où il ne recherchera les données qu'au début d'un champ (ou à la fin s'il trie par la droite). Mais il diffère de *Rechercher* car il ne limitera pas sa recherche au champ de tri primaire.

- Dans la fenêtre 1, Recherchez l'enregistrement *bain*
- Placez le curseur dans la première phrase illustrative `\xv Bain e? Mlay`. juste devant le mot *e*.
- Choisissez Edition, Sauter vers.

Ceci fera apparaître une boîte de dialogue avec le mot *e* déjà inscrit dans la case d'édition *Enregistrement à rechercher*. En dessous de cette case, vous noterez que l'Encodage de langue est également déterminé par le champ contenant la donnée recherchée. (Si le mot

recherché devait être dans une autre langue, vous pourriez choisir celle-ci dans la liste déroulante.) En dessous se trouve la case à cocher *Champ entier*.

Cochez *Champ entier* (si ce n'est pas déjà fait).

Sur la droite, sous les boutons *OK* et *Annuler*, vous verrez un autre bouton intitulé *Dest.(ination de) saut*

Choisissez *Dest. Saut*.

Ceci nous amène à une boîte de dialogue qui fonctionne comme les boîtes *Champs de tri* ou *Champs affichés* dans l'*Index*. La colonne de gauche contient une liste des bases de données actuellement chargées (*bases de données disponibles*), et celle de droite contient une liste des *bases destinataires* (elle est probablement vide en ce moment).

Les destinations de saut sont les champs (des bases de données) qui seront à comparer avec la séquence demandée dans *Sauter vers*. Contrairement à *Rechercher*, qui ne regarde que dans le champ de tri primaire, *Sauter vers* examinera tous les champs spécifiés. Comme *Rechercher*, *Sauter vers* ne compare que le début des champs. Un index sera construit pour chaque couple base de données/champ spécifié dans *Dest. saut*.

Si la colonne de droite (*Bases destinataires*) comporte quelque chose, cliquez sur le bouton *Vider*, entre les deux cases, pour la vider.

Dans la colonne de gauche, sélectionnez *SHLEX.DB*.

En dessous de la colonne de gauche, se trouve une petite case déroulante intitulé *Champs à explorer*. Déroulez la liste en cliquant sur la flèche pointant vers le bas. Remarquez qu'elle présente tous les champs de la base de données. Elle sélectionne par défaut le marqueur d'enregistrement, *\X*, ce qui nous convient.

Cliquez sur *Premier*.

Cliquez sur *Champs à explorer*, et choisissez le champ *ge*.

Cliquez sur le bouton *Dernier*.

En regardant un peu plus loin dans la boîte de dialogue, on trouve deux autres cases à cocher, en dessous de la liste de gauche, qui sont loin d'être inutiles. La première annonce *Afficher le chemin complet*.

Cocher *Afficher le chemin complet*.

Le chemin des bases de données, sur le disque dur, s'affichera dans les boîtes de droite et de gauche. (NOTE: sur certains systèmes, le chemin peut s'avérer trop long pour que la fenêtre puisse l'afficher, et le nom des bases de données risque de disparaître de l'écran.) Montrer le chemin peut être utile pour ceux qui travaillent avec des configurations complexes (les consultants, par exemple).²¹

²¹ Par exemple lorsque vous expérimentez à partir de différents fichiers de test, il est possible d'appeler une copie d'une base de données se trouvant quelque part et de l'ajouter au *Jump path*. Puis, plus tard, vous pourriez télécharger cette base de données pour en charger une autre du même nom se trouvant dans un autre dossier. La base de données dans le *Jump Path* semble correct, mais ça ne l'est pas (car il utilise un autre chemin). *Shoobox* sait que la base de données d'origine n'est pas chargée, aussi il ne sautera pas vers cette nouvelle base, même si elle a le même nom. Cela peut être déconcertant. Faire afficher le chemin permet de clarifier la situation. C'est un peu pour cela qu'il vous a été demandé de commencer par vider le *Jump Path*.

La configuration des destinations de saut retient les spécifications des fichiers que vous aurez placés dans le chemin de recherche, même si certains sont fermés par la suite. Retenir les spécifications de la destination de saut pour des fichiers déchargés est cependant une bonne chose. En effet, vous n'aurez pas besoin de reconstruire une destination de saut à chaque modification mineure de la configuration des fichiers. Vous pouvez définir la liste des bases de données que vous voulez voir explorées normalement. Si vous décidez d'en décharger une ou plusieurs par la suite (un jour ou l'autre)—elles resteront néanmoins dans la destination de saut et seront de nouveau disponibles pour le saut lorsque vous les rechargerez.

Cliquez sur *Afficher le chemin complet* pour le décocher.

En dessous se trouve la case *Avertir avant de fermer une base destinataire de saut*. C'est suffisamment clair. Il est très courant de fermer des fenêtres provisoires et de découvrir qu'on vient de fermer la *dernière* base de données *destinataire de saut*. Cette option vous avertira avant la fermeture du dernier fichier listé dans la destination de Saut. Autrement vous pourriez être déconcerté que Sauter vers ne retrouve pas ce que vous savez pertinemment être là.

Choisissez *Avertir avant de fermer...*

Cliquez sur OK, OK.

Le programme trouvera trois entrées *e*. Cette boîte *Plusieurs Résultats* (correspondances multiples) est la même que celle résultant d'un *Rechercher*, sauf qu'ici elle précise dans quelle base de données la séquence a été trouvée. Comme nous n'avons qu'une base de données actuellement dans notre *Destination de saut*, ce sont les mêmes, mais parfois cette différence peut être intéressante pour choisir.

Remarquez dans le coin droit, en bas de la boîte de dialogue, la petite boîte intitulée *Champ Additionnel*.

Sélectionnez le marqueur **ge** dans *Champ Additionnel*

Sélectionnez l'entrée avec le sens (*question mark*) puis Cliquez sur OK

Shoebox présente alors une fenêtre intitulée SHLEX.DB:3 affichant l'enregistrement.

Fermez la fenêtre.

Maintenant amusons-nous un peu !

Positionnez le pointeur de la souris sur le mot *e* puis faites **w:Clic-droit** ou **m:Option+clic** avec la souris.

Remarquez que la boîte *Plusieurs Résultats* est apparue. La commande **w:Clic-droit** ou **m:Option+clic** exécute un *Sauter vers*, en utilisant la configuration préalablement établie, et en sautant vers le mot près duquel se trouvait le pointeur.

Cliquez sur Annuler et retournez dans la fenêtre dans laquelle vous étiez.

Vous vous souviendrez que nous avons inclus le champ **ge** dans la Destination de Saut. Grâce à cela nous pouvons Sauter vers des mots anglais aussi:

Recherchez le mot *hesy* (choisissez celui qui signifie 'bridewealth')

Déroulez le texte jusqu'à la phrase illustrative.

Maintenant supposons que vous voulez vérifier que le mot *tba* vient en fait de la racine signifiant *go*:

Sauter vers le mot *go* à partir de la traduction anglaise de la ligne suivante. (Faites **w:Clic-droit** ou **m:Option+clic**, le curseur de la souris étant sur ce mot.)

Le programme va rapidement afficher une fenêtre contenant le lexème *-ba* avec le curseur placé à côté du champ **ge go**.

- Fermez la fenêtre Sauter vers.²²

Une autre façon d'utiliser Sauter vers, très utile pour l'interalignement, consiste à désigner une fenêtre existante comme destination de saut (*Fenêtre de saut*). (Nous utiliserons abondamment le saut et les fenêtres de saut plus tard avec l'interalignement.)

- Sélectionnez la deuxième fenêtre, qui est triée par la glose.
 Choisissez Affichage, Fenêtre de saut.
 Sauter vers le mot *go* qui figure toujours dans la première fenêtre. (Faites **w**:Clic-droit ou **m**:Option+clic avec le curseur de la souris sur le mot.)

Remarquez que la deuxième fenêtre se déplace pour montrer le mot. C'est bien plus pratique parce que vous ne perdrez plus de temps à fermer des fenêtres.

Un Saut ne peut conduire qu'à une Fenêtre de saut dont le champ de tri primaire est le même qu'un des champs spécifiés dans la Destination de saut. Ainsi sauter vers un mot qui est dans un champ glose permet de commuter dans la bonne fenêtre parce qu'elle est triée par le champ glose.

- Sélectionnez la première fenêtre.
 Choisissez Affichage, Fenêtre de saut.
 Faites Edition, Sauter vers sur le lexème *e*.
 Dans la boîte de dialogue *Plusieurs Résultats*, sélectionnez celui contenant (*question mark*).

C'est la première fenêtre qui est maintenant mise à jour pour afficher l'enregistrement '*e*' (car c'est la fenêtre triée par le lexème).

- Essayez de sauter vers le mot *question*.

Une boîte apparaît disant que le mot n'a pas été trouvé.

- Cliquez sur Annuler.
 Sélectionnez la phrase (*question mark*) avec les parenthèses.
 Faites **w**:Clic-droit ou **m**:Option+clic pour sauter.

La fenêtre correcte apparaît affichant l'enregistrement contenant la glose (*question mark*). Si du texte est sélectionné, le saut par un Clic-droit se fait vers cette sélection. Si rien n'est sélectionné, c'est le mot sous le curseur qui est utilisé pour le saut.

Rechercher vs Retrouver vs Sauter vers

Ce n'est pas évident mais il y a des différences importantes entre Rechercher, Retrouver, et Sauter vers.

Rechercher ne recherche que dans le champ de tri primaire.	Retrouver peut rechercher dans tous les champs, ou par langue.	Sauter vers peut sauter vers n'importe quel champ spécifié dans le chemin Destination de saut.
Rechercher recherche dans l'ensemble de la base de données filtrée de la fenêtre active.	Retrouver peut être restreint à l'enregistrement en cours ou peut rechercher dans l'ensemble de la base de données filtrée.	Sauter vers peut rechercher dans les autres bases de données chargées. Il <i>n'est pas restreint</i> à la fenêtre active, ni au même type de base de données.

²² Le raccourci Windows pour fermer une fenêtre est Ctrl+F4; sur Mac c'est Commande+W.

Rechercher utilise l'index en cours pour le tri. Le tri prend du temps lorsqu'il est généré, mais par la suite, cela fait gagner du temps.	Retrouver recherche toujours séquentiellement—il parcourt entièrement la base de données, lettre par lettre, à la recherche de la prochaine occurrence de la séquence.	Sauter vers construit un index pour chaque couple base de données-champ dans son chemin. Lorsque ces index sont créés, le Saut est rapide.
Rechercher recherche la séquence en début du champ de tri primaire, ou comme mot entier, si vous le spécifiez.	Retrouver recherche la séquence n'importe où elle pourrait apparaître, suivant qu'on recherche par langue ou par champ. Retrouver peut être configuré pour rechercher un mot complet.	Sauter vers (comme Rechercher) suppose la séquence en début de champ, mais elle ne se limite pas au champ de tri primaire; Sauter vers peut trouver des données en début de n'importe quel champ.

Par exemple, si vous faites un **Rechercher** pour *the*—sans spécifier Champ entier—vous aurez le choix entre *the, them, theme, then, there, etc.*, tous ces mots apparaissant en début de ce champ et commençant par *the*. Si vous faites un **Retrouver** pour *the*—sans spécifier *Mot complet*—en plus des mots trouvés par Rechercher, c.à.d. tous les mots contenant les lettres *the*, et pas seulement ceux commençant par *the*. **Sauter vers** se comporte comme Rechercher pour la correspondance des mots.²³

Filtrer Un filtre est un autre moyen d'effectuer une recherche, mais dans ce cas, il s'agit de retrouver un ensemble d'enregistrements correspondant à un critère particulier. Une autre façon de dire est que Filtrer vous permet de restreindre votre attention à un ensemble particulier d'enregistrements. Les commandes Enregistrement Suivant et Précédent, Premier enregistrement et Enregistrement précédent, ne considèrent que les enregistrements qui répondent au critère.

Filtres élémentaires

Supposons que vous vouliez examiner les entrées qui contiennent la combinaison de voyelles *ai*.

- Sélectionnez la première fenêtre, qui est classée par lexème.
- Choisissez Base de Données, Filtrer.
- Choisissez Modifier Filtres.
- Choisissez Ajouter.
- Spécifiez *ai* comme Nom du filtre en haut de la boîte de dialogue.

Remarquez qu'il y a deux colonnes principales, un peu comme vous l'avez déjà rencontré dans Trier, Destination de saut, et ailleurs. La colonne de gauche s'intitule *Critères de filtrage*. Elle contient une liste des types de filtres et des opérateurs logiques. La colonne de droite est vide, nous n'avons pas encore construit de filtre.

- Choisissez Marq./Séquence dans la colonne de gauche.
- Choisissez Insérer (ou double-cliquez sur Marq./Séquence).

Cela fait apparaître une boîte de dialogue qui vous permet de spécifier le marqueur et la séquence à rechercher pour le filtre Marq./Séquence.

- Choisissez *lx* pour le Marqueur.
- Tapez *ai* pour la *Séquence*.

²³ L'Aide contient ce tableau comparatif sous une présentation plus concise, organisée suivant les trois commandes plutôt que sous la forme de fonctions comparatives. Choisissez Aide, Index. Puis choisissez Rechercher et tapez Jump To.

Avant de sortir de la boîte, remarquez rapidement la section *Classe* sous Séquence. Si vous cliquez sur la petite flèche, une liste des classes²⁴ définies pour l'encodage de langue du marqueur choisi se déroule. Chacune d'elles peut être insérée en cliquant sur le bouton *Insérer dans la Séquence*, en dessous de la liste.

- Cliquez sur OK (trois fois) jusqu'à ce que vous soyez revenu à la petite boîte de dialogue intitulée *Filtre*.
- Choisissez le filtre *ai* dans la liste déroulante.
- Cliquez sur OK.

Le programme va construire l'index approprié. Si l'enregistrement en cours ne satisfait pas au critère du filtre, vous verrez s'afficher le premier enregistrement y satisfaisant. Par la même occasion, constatez que la barre d'état affiche le nombre d'enregistrements satisfaisant aux critères du filtre.

- Choisissez Affichage, Index.

C'est un bon moyen de voir les résultats d'un filtrage. Ici, vous pouvez voir les caractéristiques *ai* des mots dans le champ **lx**.

Remarquez que la case la plus à droite de la barre d'outils contient maintenant *ai*—le nom du filtre actif.

- Déroulez la liste de la case et remontez sur *[pas de filtre]*.
- Choisissez *[pas de filtre]*.

Le filtrage est suspendu, comme vous pouvez le voir tout de suite dans la fenêtre de l'Index.

Filtres et Classes

Maintenant voici un exemple un peu plus élaboré de filtres. Disons que vous vous demandez si le Selaru ne serait pas une langue austronésienne, car généralement ces langues ont un schème syllabique CVCV (consonne-voyelle-consonne-voyelle). Déjà d'un simple coup d'œil, on s'aperçoit que le Selaru connaît des séquences de consonnes (schèmes CC). Essayons de le vérifier.

- Dans la fenêtre 1, choisissez Base de Données, Filtrer.
- Cliquez sur Modifier Filtres. Ceci nous conduit à l'onglet *Filtres* dans la boîte de dialogue *Propriétés du Type de Base de Données*.
- Cliquez sur Ajouter.
- Donnez un nouveau nom de filtre: CC.
- Double-cliquez sur Marq./Séquence.
- Sélectionnez **lx** dans la fenêtre *Marqueur*.

Jusqu'ici ça va. Maintenant nous allons utiliser les *Classes*

- Déroulez la liste sous le mot *Classe*, descendez et cliquez sur cons (la Classe 'consonne').
- Puis cliquez deux fois sur le bouton Insérer dans la Séquence. (Ceci devrait insérer **[cons][cons]** dans la case *Séquence* juste au-dessus de la boîte *Classe*.)
- Vous y êtes, faites maintenant OK, OK, OK (c.à.d. 3 fois).

Ceci devrait vous ramener à la boîte de dialogue *Filtre*.

- Sélectionnez le nouveau filtre CC dans la liste déroulante.
- Cliquez sur OK.

²⁴ Vous pouvez voir et éditer ces Classes en faisant apparaître la boîte de dialogue Propriétés des marqueurs pour ce marqueur et en cliquant sur Propriétés de la langue. Cela vous conduira à la boîte de dialogue Propriétés de l'Encodage de la langue. Choisissez Classes. Vous pouvez aussi vous rendre à la boîte de dialogue Propriétés de l'Encodage de la langue en passant par Projet, Encodages des langues, puis en choisissant la langue en question.

Si le lexème affiché dans la fenêtre 1 ne contient pas la séquence de consonnes (un schème CC), Shoebox passera au prochain enregistrement qui le contiendra.

Cliquez sur le bouton Affichage Index.

Ceci vous fournira la liste de tous les enregistrements qui satisfont au schème CC que vous avez spécifié pour ce filtre. En déroulant la liste vers le bas, on se rend compte que la plupart de ces séquences de consonnes concernent soit *w* soit *y*. On les appelle des *glides*, et elles ont un comportement particulier en Selaru (comme dans beaucoup de langues).

Ajouter une nouvelle Classe

Comme les glides semblent avoir un comportement particulier en Selaru, l'analyse de cette langue serait plus efficace si on créait une Classe juste pour ces glides.

Choisissez Projet, Encodage de langues.

Sélectionnez *vernacular* puis cliquez sur Modifier (ou simplement double-cliquez sur *vernacular*)

Cliquez sur l'onglet *Classes*.

Ceci affiche toutes les *Classes* qui ont été créées pour la langue "vernacular".²⁵ Les *Classes* sont définissables par l'utilisateur, et non imposées par Shoebox, vous pouvez donc modifier celles que vous voulez, comme vous voulez. A chaque Encodage de langue correspond son jeu de *Classes* et c'est ici que nous allons rajouter la Classe pour les "glides".

Cliquez sur Ajouter.

Dans la case *Nom*, tapez : *glide*.

Dans la case *Caractères*, tapez : *W w Y y* (avec un espace entre chaque lettre).

Cliquez sur OK.

Maintenant vous pouvez constater que Shoebox a rajouté la nouvelle Classe *glide* dans la liste des *Classes* pour la langue *vernacular*.

Nous pourrions faire la remarque que la Classe *cons* contient toujours *w* et *y* dans la liste des consonnes. C'est normal (et même bien ici) puisque les glides sont aussi des consonnes. De cette façon, nous pouvons aussi rechercher les glides lorsqu'elles fonctionnent simplement "comme des consonnes".

Filtres complexes

Revenons à nos filtres. Pour voir quel impact ces glides ont sur les séquences de consonnes en Selaru, nous allons nous débarrasser de tous les lexèmes contenant une séquence de consonnes comportant un glide. Ainsi nous pourrions voir combien de lexèmes comporte un schème CC *sans* glides.

Cliquez sur OK puis sur Fermer pour revenir à la base de données.

Puis choisissez Base de Données, Filtrer.

Cliquez sur Modifier Filtres.

Sélectionnez le filtre *CC* puis cliquez sur Copier.

Donnez à la copie du filtre *CC* un nouveau nom : *CC-sans glide*.

Remarquez que ce filtre est déjà près (puisque'il s'agit d'une copie du filtre *CC*) pour restreindre la base de données aux seuls lexèmes contenant un schème consonantique complexe. Nous devons donc ajouter maintenant la contrainte supplémentaire de *ne pas* montrer ceux qui contiennent des glides :

Sélectionnez Et, puis cliquez sur le bouton Insérer (ou double-cliquez sur Et).

Sélectionnez Non et cliquez sur le bouton Insérer (ou double-cliquez sur Non).

Ceci définit la logique de la nouvelle contrainte :

²⁵ Les concepts de *Projets* et d'*encodages de langue* sont abordés plus en détail dans la section Trois concepts , p. 37

- Maintenant sélectionnez Marq./Séquence et cliquez sur le bouton Insérer (ou double-cliquez sur Marq./Séquence).
- Sélectionnez **ix** dans la fenêtre *Marqueur*.
- Déroulez la liste sous le mot *Classe*, descendez et cliquez sur la Classe cons (comme précédemment).
- Cliquez sur le bouton Insérer dans la Séquence *une fois* (pour insérer **[cons]** dans la case séquence).
- Déroulez la liste sous le mot *Classe* à nouveau et cliquez sur la nouvelle Classe glide.
- Cliquez sur le bouton Insérer dans la Séquence (pour faire afficher **[cons][glide]** dans la case *Séquence*).
- Cliquez sur OK.

Nous sommes maintenant revenu à la boîte de dialogue principale du filtre *CC-no glide*. Remarquez comment le filtre complexe traduit : “N’afficher que les enregistrements dont les lexèmes satisfont aux deux contraintes suivantes: le lexème doit contenir un schème CC et il ne doit pas contenir de schème comportant un glide (comme second élément).”

Pour activer le filtre :

- Cliquez sur OK, OK (c.à.d. 2 fois).

Ceci vous fait revenir à la petite boîte de dialogue *Filtre*.

- Sélectionnez le filtre CC-sans glide dans la liste déroulante.
- Cliquez sur OK.

Si la fenêtre 1 est toujours en mode l'Index, vous devriez voir la liste se réduire à 2 lignes. Ceci vous montre simplement la rareté des mots à consonnes complexes sans glide en Selaru. Cet exemple sert juste à illustrer comment les filtres peuvent être utilisés dans l'analyse pour se concentrer sur les données correspondant au problème du moment.

Pour en finir avec les filtres, si vous devez changer de champ de tri primaire, puis utiliser un filtre, commencez par changer l'ordre de tri. Si vous créez le filtre en premier, ceci occasionnera la création d'un index de filtre relatif à l'ordre de tri en cours. Si par la suite vous changez d'ordre de tri, Shoebox devra reconstruire l'index de filtre en plus de reconstruire un index de tri pour toute la base de données (non filtrée).

Quelques autres commandes

Annuler Toutes les modifications

Annuler est une commande très utile pour les individus loin d'être parfaits que nous sommes. Aussi Shoebox dispose de deux types d'annulation: Annuler incrémentiel et Annuler tout.

- Revenez au mode d'affichage normal (c.à.d. pas le mode Index).
- Allez au premier enregistrement de la base de données, et faites quelques modifications dans l'enregistrement. Par exemple, vous pouvez ajouter des gloses. Ajoutez quelques autres marqueurs. Modifiez plusieurs champs.
- Choisissez Edition, Annuler Tout.

Ceci fera apparaître un message vous demandant si vous voulez réellement annuler *tout* ce que vous avez fait dans l'enregistrement.

- Choisissez Oui.

Toutes les modifications que vous avez faites seront annulées. Cela n'annule pas simplement la *dernière* modification, aussi soyez prudent. Pour annuler seulement la dernière modification faite :

- Ajouter du texte et faites quelques modifications à l'enregistrement.
- Choisissez Edition, Annuler.

Ceci annulera la toute dernière chose que vous avez faite.

- Continuez à faire Edition, Annuler pour voir Shoebox revenir en arrière vers l'enregistrement dans son état initial.

Cette annulation incrémentielle est bien utile en édition normale, mais si vous avez fait une bourde dans un enregistrement, l'utilisation de Annuler Tout vous le restituera intégralement.

NOTE IMPORTANTE: Shoebox conserve une trace des modifications apportées à un enregistrement tant que celui-ci est affiché dans une fenêtre. Aussi si vous avez plusieurs fenêtres affichant des données, et que vous cliquez sur une autre fenêtre (et même si vous l'éditez) en revenant dans la fenêtre précédente (qui présente toujours l'enregistrement d'origine), les modifications de la première fenêtre peuvent encore être annulées. **Mais** si dans cette fenêtre vous passez à un autre enregistrement (en utilisant Rechercher, etc.), puis que vous revenez au précédent enregistrement, vous **ne pourrez pas** annuler les précédentes modifications apportées à ce premier enregistrement.

Voici d'autres limitations de la fonction d'annulation :

- Annuler n'annule pas les déplacements de texte—ni à l'intérieur, ni entre des enregistrements.
- Si vous *savegardez* la base de données, vous ne pourrez pas annuler les modifications apportées avant la sauvegarde.
- Si vous modifiez les propriétés d'un marqueur—par exemple, si vous avez changer son nom de champ—il ne s'agit pas d'un changement dans l'enregistrement, et ce changement ne sera pas annulé, mais les modifications apportées avant et après sont annulables.

Quitter et lancer Shoebox

Lorsque vous quittez Shoebox et que vous le relancez, il se souvient de toutes les fenêtres qui étaient ouvertes à la fermeture ainsi que de leur configuration.

- Pour quitter Shoebox, Faites **w:Fichier**, Fermer ou **m:Fichier**, Quitter
- Cliquez sur Oui pour enregistrer les modifications.
- Relancez Shoebox.

Observez que vous avez les mêmes fenêtres, triées de la même façon, et présentant les mêmes enregistrements que lorsque vous avez quitté la programme.

Fin de l'étape

Ici s'achève cette étape de l'apprentissage. Rien de ce qui suit ne s'appuie sur les modifications que vous avez apportées dans cette section.

Créer un nouveau projet et une base de données

Présentation Cette section décrit les concepts de Projet, d'Encodage de langue, et de Type de Base de données, et montre comment les configurer.

L'asterisque * De nombreuses langues nécessitent des caractères spéciaux (romains modifiés ou graphèmes différents) et par conséquent des ordres de tri, polices, claviers particuliers etc., mais d'autres non. Aussi nous avons défini un double cheminement dans cet apprentissage. Ceux qui peuvent se contenter de l'alphabet romain pourront sauter les sections marquées d'un astérisque * en début de titre. Certaines sections comportent un astérisque * au commencement d'une série de paragraphes. Vous pouvez les sauter également. Par contre, pour ceux d'entre vous qui doivent faire face à un alphabet exotique, ces parties vous concernent.

Trois concepts essentiels Il y a trois concepts essentiels à comprendre, même si vous n'avez aucun besoin particulier en caractères spéciaux. Ces trois concepts sont les Projets, les Encodages de langue, et les Types de Base de données. Ces concepts sont assez simples.

Concept 1: Projets Un *projet* c'est principalement la configuration de votre espace de travail sous Shoebox. Un projet mémorise les fichiers chargés dans Shoebox, combien de fenêtres sont ouvertes, où elles s'affichent sur l'écran et comment chacune d'elles est triée et filtrée.

Il y a fort à parier que lorsque vous relancez Shoebox, vous aimeriez retrouver la même configuration que lorsque vous avez quitté le programme—un projet Shoebox sert à cela.

Mieux que cela, vous pouvez définir différents projets pour différentes tâches ; par exemple vous pouvez définir un projet pour un travail lexicographique, un autre pour une recherche anthropologique, un autre pour une analyse grammaticale, et un autre encore pour une analyse de textes ou un travail de traduction ; Passer d'une tâche à l'autre est rendu très simple par l'ouverture du projet correspondant.

Si vous conservez tous vos configurations de projets (pas les données) dans le même répertoire, Shoebox pourra partager des paramètres communs à ces différents projets. Parmi ces paramètres partageables figurent les encodages de langues et les types de bases de données.

Concept 2: Encodage de langue Une langue présente plusieurs caractéristiques, évidemment. L'une d'elles est l'*alphabet* normalement utilisé pour la transcrire. (Si une langue peut être transcrite suivant deux types d'écriture, chacun de ces types constituera un encodage.)

Ce jeu de paramètres liés à la transcription que nous appelons *Encodage de langue* est constitué de :

- l'ordre de tri approprié de cet alphabet
- les associations minuscules et majuscules de ces caractères (si elles existent).
- les regroupements particuliers (appelés *Classes*) de ces caractères qui peuvent être intéressants dans la recherche de données
- une police permettant de représenter ces caractères
- souvent, un clavier spécial pour faciliter l'accès à ces caractères

Shoebox sépare les informations d'encodage des langues des paramètres de Projet, afin que différents projets utilisant les mêmes langues puissent partager ces informations d'encodage des langues sans avoir à les redéfinir dans chaque projet.

Concept 3: Type de Base de Données

Shoebox utilise un concept de *Type de Base de Données*. Comme types génériques de bases de données, nous aurons par exemple les lexiques, les notes culturelles, les notes anthropologiques, ou une collection de textes traditionnels à interaligner.

Un *Type* de base de données comprend:

- Les différents marqueurs figurant dans le fichier base de données de ce type.
- Un nom descriptif pour chacun des marqueurs.
- La position hiérarchique de chacun des marqueurs.
- L'association de chacun des marqueurs à un encodage de langue (Concept 2).
- Des restrictions sur le contenu de certains champs (range set).
- Le type de style associé à un marqueur pour l'exportation vers Word.
- Un modèle (template) des marqueurs pour les nouveaux enregistrements (jeu de marqueurs à insérer).
- Les différents Filtres d'un type de base de données (qui concernent souvent des marqueurs particuliers).
- La destination de Saut.
- Des paramètres de configuration de l'Interalignement ou de l'Adaptation (aide à la traduction) associées à ce type de base de données.
- Les processus d'exportation définis pour les bases de données de ce type.

Comment tout ceci s'articule

Tout fichier ouvert fait partie d'un Projet. Mais un fichier n'appartient pas à un projet particulier, car *plusieurs projets peuvent partager le même fichier* (par exemple, le lexique peut être ouvert par le projet d'analyse grammaticale, comme par le projet de dictionnaire). Chaque fichier de base de données est d'un Type donné ; plusieurs fichiers peuvent être d'un même type (par exemple, plusieurs fichiers texte pourraient être du même type "texte interaligné"). Un Projet peut avoir des fichiers de différents Types chargés au même moment. Chaque marqueur d'un Type de base de données est associé à un Encodage de langue, mais l'Encodage d'une langue n'appartient pas exclusivement à un Type de base de données. Les marqueurs de différents Types de base de données peuvent être associés à un même Encodage de langue, etc.).

Un Projet particulier peut partager certains fichiers avec d'autres Projets, et en comporter d'autres exclusifs. Mais, le Type de base de données d'un fichier ne peut pas être différent d'un Projet à l'autre, de même que l'Encodage de langue associé à un marqueur ne peut pas être différent dans un fichier utilisé dans différents projets.

Le diagramme ci-dessus montre deux projets. Le premier est un projet de Texte interaligné comportant un lexique (divisé en 2 fichiers) et une série de textes à interaligner. Le second est un projet anthropologique comportant une base de données de notes et quelques contes et autres textes. Les textes sont partagés avec le projet interalignement. Les fichiers *texte* sont intégrés dans le projet anthropologique car ils contiennent des contes folkloriques et d'autres textes d'intérêt culturel.

Les fichiers lexiques sont associés au type de base de données *Lexicon*. Ces fichiers lexique contiennent des données en trois langues : la langue nationale, la langue vernaculaire et l'anglais. Les notes culturelles sont associées au type de base de données *Anthropology*. Ce fichier de notes est en anglais et en langue vernaculaire.

Les deux fichiers de projet (Int_text.prj et Anthro.prj) sont conservés dans le même dossier avec les fichiers de configuration TYP et LNG. Les fichiers de données (notes culturelles, lexique et textes) sont conservés séparément dans leur dossier respectif (suivant le choix du

linguiste).²⁶ Disposer ainsi vos fichiers permet à Shoebox de partager les mêmes fichiers de configuration entre les différents projets. Dans cet exemple, les informations sur l'ordre de tri et la sélection de la police de la langue vernaculaire ne sont spécifiés qu'une fois (dans le fichier `V_lang.lng`). Ceci rend la vie plus facile.

Une fois que vos fichiers d'encodage de langue et de type de base de données sont configurés, créer une nouvelle base de données revient tout simplement à faire Fichier, Nouveau et à dire à Shoebox quel type de base de données vous souhaitez créer. Par exemple, si le linguiste utilisant la configuration précédente veut créer une nouvelle base de données de textes interalignés, il n'aura qu'à choisir Fichier, Nouveau et à sélectionner le type de base de données *Text* dans la liste des choix. Shoebox sait ce qui lui reste à faire.

Partir de zéro: pas à pas

Dans cette section vous allez construire un échantillon de base de données lexicales utilisant des données en langue Axininca Campa²⁷. Ce qui suit constitue un parcours des étapes nécessaires au démarrage d'un nouveau projet Shoebox (y compris les étapes concernant le traitement des caractères spéciaux) :

1. Installer les polices nécessaires.
 2. Paramétrer le clavier *Keyman*, si nécessaire.
 3. Lancer Shoebox.
 4. Créer un nouveau Projet.
 5. Créer un Type de base de données.
 6. Configurer vos encodages de langue, ce qui implique de préciser pour chacun :
 - Clavier
 - Police
 - Ordre alphabétique
 - Casse
 - Classes
1. Créer un nouveau fichier de base de données, ce qui implique :
 - Créer le premier enregistrement
 - Associer un encodage de langue à chaque marqueur
 - Créer un modèle d'enregistrement

Cela semble constituer beaucoup d'étapes juste pour démarrer, et ça l'est, mais ceux qui ont des systèmes alphabétiques simples, peuvent sauter les étapes 1, 2, et une grande partie de 6. Finalement, même pour ceux qui ont des besoins complexes, ce n'est pas si lourd qu'il y semble.

Souvenez-vous, les paragraphes et sections précédés par une astérisque * peuvent être sautés par ceux qui ne nécessitent pas de caractères spéciaux.

*Installer la police Axininca

Comme pour la plupart des projets linguistiques, l'Axininca comporte certains caractères spéciaux. Les fichiers de polices de caractères se trouvent dans le dossier POLICES en dessous du dossier SAMPLES. Chaque système a sa manière de charger les polices. Choisissez la méthode appropriée à votre plate-forme.

Sur Mac

Allez dans le dossier POLICES en dessous du dossier SAMPLES. Sélectionner les fichiers de polices pour les données Axininca et amenez-les dans le dossier SYSTEME:POLICES.

²⁶ Il est recommandé de ne pas mélanger dans le dossier des projets, les fichiers de données et les fichiers de configuration. Le mélange des fichiers de données et de configurations sera généralement à éviter sauf si vous devez partager des données avec d'autres, ou contrôler les données d'autres personnes, ou encore si vous êtes en train de développer des exemples pour un cours, etc.

²⁷ Les données Axininca Campa sont utilisées avec l'autorisation de David L. Payne, 1981. *The Phonology and Morphology of Axininca Campa* (The Summer Institute of Linguistics and The University of Texas at Arlington.) Axininca se prononce *aSHINinca*.

(Elles seront également utilisées pour l'apprentissage de l'interalignement. Ensuite, vous pourrez les supprimer.)

Sous Windows Windows 95 et Windows 3.1 ont tous les deux une boîte de dialogue *Ajouter des polices*, mais des routes différentes pour y accéder. Suivez la direction correspondant à votre système, puis allez à la section *...étapes communes* ci-dessous.

... sous Windows 95 Pour installer une police sous Windows 95, dans la barre des tâches,

choisissez Démarrer, Paramètres, Panneau de Configuration.

Vous verrez s'afficher un certain nombre d'icônes.

Double-cliquez sur Polices (un répertoire comportant *A, A, a*).

Vous verrez s'afficher un certain nombre de répertoires.

Cliquez sur Fichier, Installer la nouvelle Police.

Cette boîte de dialogue comporte (probablement) en plus des boutons habituels, une zone vierge pour lister les polices, une boîte de répertoires, et une case à cocher dans le coin inférieur droit de la boîte de répertoire.

Si la case à cocher intitulée *Copier les Polices dans le répertoire Windows* est cochée, décochez-le.

Vous n'avez pas besoin d'une copie permanente de ces polices. Une fois le didacticiel parcouru, vous n'en aurez plus besoin.

... sous Windows 3.1 Voici les étapes pour installer une police sous Windows 3.1:

Allez dans le Gestionnaire de Programmes, puis dans le Groupe Principal, ouvrez le Panneau de Contrôle.

Cliquez sur Polices (icône *ABC*). Vous verrez une boîte de dialogue contenant une liste de polices.

Choisissez Ajoutez.

Cette boîte de dialogue comporte (probablement) en plus des boutons habituels, une zone vierge pour lister les polices, une boîte de répertoires, et une case à cocher dans le coin inférieur droit de la boîte de répertoire.

Si la case à cocher intitulée *Copier les Polices dans le répertoire Windows* est cochée, décochez-le.

Vous n'avez pas besoin d'une copie permanente de ces polices. Une fois le didacticiel parcouru, vous n'en aurez plus besoin.

... étapes communes Les étapes suivantes sont communes aux deux plates-formes Windows :

Naviguez dans le dossier POLICES sous le dossier SAMPLES (qui est lui-même sous le dossier Shoebox).

Vous verrez quatre lignes apparaître dans la Liste des Polices, toutes nommées *Axininca Ft SILDoulos ... (True Type)*.

Cliquez sur Toutes.

Cliquez sur OK.

Fermez tout ce qu'il faut pour revenir à votre espace de travail habituel.

***Keyman (pour Windows seulement)**

NOTE : Cette section peut être abordée sans Keyman, mais si vous utilisez Keyman dans votre configuration de travail habituel, vous apprécierez de savoir comment l'utiliser avec Shoebox.

Si habituellement vous n'utilisez pas Keyman dans votre système, vous pouvez sauter cette étape. Rien ne vous oblige à taper des caractères spéciaux dans ce parcours.

NOTE: La version Windows 95 de Shoebox ne sait pas utiliser Keyman pour commuter automatiquement de clavier lorsque vous placez le curseur dans un champ associé à une autre langue. Si vous devez utiliser Keyman, vous pouvez lancer la version Windows 3.1 de

Shoebox pour accéder à cette fonctionnalité, ou alors vous pouvez continuer à changer de clavier manuellement comme vous le faites avec les autres programmes. La version Win'3.1 de Shoebox tourne un peu moins vite sous Windows 95 que la version Win'95.

- Shoebox vérifie à son lancement, quels sont les claviers disponibles dans Keyman, aussi si vous êtes déjà dans Shoebox, vous devrez commencer par en sortir afin de pouvoir utiliser dans Shoebox un nouveau clavier.
- Vous avez sans doute installé Keyman dans votre groupe de démarrage. Cliquez dessus pour faire apparaître la boîte de dialogue Tavultesoft Keyboard Manager (avec les 5 boutons—Hide, About, Options, Exit, et Help).
- Cliquez sur Options.

Vous verrez une boîte de dialogue comportant un certain nombre d'options.

- Cliquez sur Load (à peu près vers le centre).

Ceci ouvre une boîte de dialogue classique.

- Allez jusqu'au répertoire KEYMAN en dessous du répertoire SAMPLES (lui-même en dessous du répertoire SHOEBOX
- Choisissez *AXININCA.KMN*.
- Cliquez sur OK.

Ceci devrait vous ramener à la boîte de dialogue à 5 boutons.

- Cliquez sur Hide.

Comme Shoebox recherche à son lancement si des fichiers de claviers Keyman sont chargés en mémoire, il faut que Keyman soit lancé avant Shoebox pour que ces claviers soient reconnus par Shoebox.

Démarrer un nouveau projet

Shoebox initialise un projet en créant un fichier de paramètres contenant les informations adéquates.

- Lancez le programme Shoebox.
- Cliquez sur Projet, Nouveau.²⁸

Ceci fait apparaître une boîte de dialogue demandant un nom de fichier pour le projet.

- Allez jusqu'au répertoire SAMP_SET (en dessous du répertoire SAMPLES).
- Tapez le nom du projet, *Axininca*.²⁹
- Cliquez sur Enregistrer (OK sous Windows 3.1).

AXININCA.PRJ est le nom du fichier projet qui sera enregistré dans le répertoire Samp_Set.³⁰

Créer une nouvelle base de données

Pour l'exemple sur lequel nous allons travailler, nous n'avons besoin que d'un petit lexique Axininca Campa-Français.

- Cliquez sur Fichier, Nouveau.

²⁸ Vous n'avez pas besoin de fermer le projet en cours pour en créer un nouveau. Shoebox ne peut avoir qu'un projet ouvert au même moment, aussi choisir Projet, Nouveau obligera Shoebox à fermer le projet en cours avant de créer le nouveau. Il en est de même lorsque vous choisissez Projet, Ouvrir alors qu'un projet est déjà ouvert.

²⁹ Si vous ne mettez pas l'extension .PRJ, Shoebox s'en chargera.

³⁰ Les utilisateurs sous **Windows** peuvent créer un raccourci (icône, sous Windows 3.1) qui servira à ouvrir un projet particulier. Le chemin complet du fichier de projet devra se trouver sur la ligne de cible (ligne de commande dans Win 3.1). pour ouvrir un projet sans créer un raccourci, double-cliquez simplement sur le fichier de projet dans l'explorateur (gestionnaire de fichier). Evidemment vous pouvez toujours utiliser le raccourci (icône) de Shoebox lui-même. Lancer Shoebox sans spécifier un nom de projet ouvre simplement le dernier projet ouvert dans Shoebox. Les utilisateurs sous **Mac** n'ont qu'à double-cliquer sur le fichier de projet pour lancer un projet particulier.

La boîte qui s'ouvre est une boîte de dialogue classique d'ouverture de fichier, avec les choix possibles de lecteur, répertoire et nom de fichier.

- Naviguez dans le dossier AX_DATA (sous le dossier SAMPLES).
- Donnez *AXININCA.LEX* comme nom de fichier.
- Cliquez sur Enregistrer (OK sous Windows 3.1).

Le Type de base de données

Vous voyez apparaître une boîte de dialogue intitulée *Choisir (un) Type de Base de Données*.³¹ Bien qu'une nouvelle base de données soit souvent d'un Type déjà existant, nous allons créer un nouveau Type pour illustrer dans cet exemple le processus initial.

- Cliquez sur Ajouter Nouveau Type.

Ceci fait apparaître la boîte de dialogue *Type de Base de Données* qui liste les différents types disponibles.

- Choisissez Ajouter.

Ceci fait apparaître une boîte de dialogue intitulée *Nouveau Type de Base de Données*. Elle requiert le nom du type de la nouvelle base de données et le nom du marqueur d'enregistrement. Le nom de type peut comporter plusieurs mots.

- Donnez *Lexique Axininca-Français* comme Nom.
- et *lx* comme Marqueur d'Enregistrement.
- Cliquez sur OK.

Ceci nous conduit à une grande boîte de dialogue à onglets intitulée *Propriétés du Type de Base de Données*. Vous êtes à l'onglet *Options*. Le programme nous invite à entrer une description du type de base de données.³²

- Entrez *Créé pour le didacticiel de Shoebox* comme description.

(Le marqueur d'estampillage de date se trouve dans cette boîte de dialogue non pas parce qu'il est si important, mais parce que *Options* est un endroit approprié. Mais vous ne pourrez pas entrer de marqueur de date tout de suite, puisque jusqu'à maintenant seul le marqueur d'enregistrement a été défini, et encore cette information n'est pas complète. Nous y reviendrons plus tard, lorsque vous aurez défini d'autres marqueurs.)

Nous pourrions visiter les autres onglets, mais ce n'est pas nécessaire tout de suite.

- Cliquez sur OK dans la boîte de dialogue *Propriétés du Type de Base de Données*.
- Cliquez sur Fermer dans la boîte *Type de Base de Données*.

Nous sommes revenus dans la petite boîte de dialogue *Choisir Type de Base de Données* où il nous avait été demandé d'entrer un nouveau type de base de données. Vous devriez avoir maintenant *Lexique Axininca-Français* affiché comme type de base de données.

- Cliquez sur OK.

Ceci nous ramène à l'écran principal de Shoebox. Vous y verrez une fenêtre intitulée *AXININCA.LEX*. Elle montre un enregistrement vide comportant un marqueur \lx.

L'Encodage des langues

Le Nom, le Clavier et la Police par défaut

La langue Axininca s'écrit avec un certain nombre de caractères spéciaux. Nous voulons spécifier pour les champs en Axininca, une police appropriée ainsi que d'autres caractéristiques. L'information d'Encodage de langue est tenue séparée de l'information de Type de base de données.

- Cliquez sur Projet, Encodage de langues.

Ceci vous conduit à la boîte de dialogue intitulée *Encodage de langues*. Elle contient une liste des encodages de langue disponibles.

³¹ Voir *Concept 3: Les types de bases de données* au début de cette section.

³² C'est pratique lorsque vous revenez après une longue coupure dans votre travail ou lorsque d'autres personnes travaillent avec vos données.

- Cliquez sur Ajouter.

Ceci fait apparaître une grande boîte de dialogue à onglets intitulée *Propriétés de l'Encodage de langue*. Vous êtes sur l'onglet *Options*.

- Entrez *Axininca Campa* dans Nom.
- Entrez *Créé pour le didacticiel de Shoebox* comme Description.
- Cliquez sur Choisir la Police.

Une grande boîte de dialogue apparaît, avec les options standards de polices. Ce choix de police constituera la police par défaut de cette langue. Ceci signifie que si vous affectez l'encodage de langue Axininca à un champ, c'est cette police qui sera utilisée pour ce champ, à moins que vous ne demandiez expressément d'en utiliser une autre en utilisant l'option Choisir la Police de la boîte de dialogue Marqueurs Propriétés.³³

- *Choisissez *Axininca Ft SILDoulos*.
- Choisissez *Gras, 11pt*.
- *Cliquez sur OK pour revenir à la boîte *Propriétés de l'Encodage de langue*.

*Sur la gauche de la boîte de dialogue de l'onglet *Options*, vous pouvez voir une petite boîte intitulée *Clavier virtuel*.

- *Si vous utilisez Keyman, choisissez *Axininca* comme clavier.

L'ordre de Tri

Pour voir la liste des ordres alphabétiques de cet encodage de langue,

- Choisissez l'onglet Ordres de tri.

Remarquez que le programme a défini un ordre de tri par défaut et lui a donné le nom *Added*.

- Choisissez Modifier.

Une grande boîte de dialogue intitulée *Ordre de tri* apparaît

- Mettez *Axininca Campa* comme nom de tri
- Changez la description en mettant *Créé pour le didacticiel de Shoebox*

NOTE: Si vous n'êtes pas familier de la terminologie utilisée dans cette boîte de dialogue, cliquez sur *Aide*. Ceci vous fournira des exemples sur les différents types de caractères (primaire, secondaire, etc).

Le programme fournit par défaut l'ordre de tri standard de l'Anglais. C'est à peu près correct pour la langue Axininca, mais quelques caractères supplémentaires doivent être rajoutés.

- * Placez votre curseur dans la boîte d'édition de *Caractères de base (même ordre en ligne)*.

* Si vous utilisez Keyman, et la version 16 bits de Shoebox (Win311), vous remarquerez que le clavier Axininca est activé (les autres utilisateurs devront activer eux-même le clavier).

(Les utilisateurs de **Mac** peuvent sauter les trois étapes suivantes.)

- * Placez votre curseur à la fin de la ligne *C c*, et tapez **w**:Entrée
- * *Si vous utilisez Keyman*, tapez les touches suivantes: /C /c —barre oblique suivie de C majuscule, espace, barre oblique suivie de c minuscule—qui devraient s'afficher comme des c et C barrés.
- * *Si vous n'utilisez pas Keyman*, entrez *Alt+0201*, un *espace*, et *Alt+0193*. Deux c et C barrés devraient s'afficher.

³³ Même si les différentes langues dans vos données ne nécessitent pas différentes polices et claviers, en les spécifiant séparément, vous aurez la possibilité d'utiliser certaines fonctionnalités comme la recherche par langue, ou la distinction des champs par des enrichissements de polices différents.

NOTE : Noter que les c barrés, minuscule et majuscule, doivent être séparés par un espace. Les caractères non séparés sont considérés comme des multigraphes.

La langue Axininca comporte un certain nombre de caractères spéciaux. Comme l'objet de cet exercice n'est pas de vous apprendre à saisir de l'Axininca, nous n'ajouterons pas d'autres caractères.

Une remarque à propos du jeu de caractères Axininca. La transcription de l'Axininca utilisée dans cette documentation est une représentation phonologique de niveau profond³⁴ plutôt qu'une orthographe pratique. Elle ne comporte pas de majuscule dans les textes, sauf pour *N*—qui n'est pas identique à *n*. Nous laisserons la majuscule dans l'ordre alphabétique pour représenter un caractère latin modifié.

Pour terminer le travail sur l'ordre alphabétique,

- Supprimez tous les caractères de *Diacritiques triées avant*...
- Supprimez tous les caractères de *Diacritiques triées après*...
- Cliquez sur OK pour revenir à la boîte de dialogue *Propriétés de l'Encodage de langue*

En dessous de la liste des ordres de tri disponibles, se trouve une petite boîte intitulée *Ordre de tri par défaut*. C'est là que vous définirez l'ordre de tri par défaut pour cette langue. Il s'agit de l'ordre de tri qui sera utilisé pour n'importe quel champ Axininca que vous définirez comme champ de tri *secondaire*. Ceci sert également pour l'interalignement. (En choisissant un champ comme champ de tri *primaire*, vous pouvez choisir l'un quelconque des ordres disponibles pour l'Axininca—mais comme nous n'en avons défini qu'un ici, il n'y a évidemment pas de choix possible.) Elle affiche *Axininca Campa*, ce qui est correct.

L'onglet Casse

- Choisissez l'onglet Casse.
- Entrez les C's comme vous l'avez fait dans l'ordre de tri.

Tous les couples de caractères qui sont en rapport majuscule-minuscule doivent être inclus dans la liste descriptive Casse. La disposition des caractères est très semblable à celle de l'ordre alphabétique. Cependant, si nous avons *A Á a á* dans la liste alphabétique ordonnée, (*A* avec et sans accent comme cela existe dans plusieurs langues européennes), les caractères sans accent et avec accent devront, dans Casse, être appariés correctement et placés sur des lignes séparées. En d'autres termes, une ligne comportera *A a* et la suivante *Á á*, car *Á* n'est pas la majuscule de *a*. Si nous avons le cas *A a á* dans l'ordre de tri—avec une majuscule non accentuée servant pour la voyelle avec ou sans accent—alors nous aurions à ce moment là mis les trois caractères sur la même ligne dans Casse.

L'onglet Classes

- Choisissez l'onglet Classes.

Il s'agit des *Classes Utilisateur*.³⁵ Comme nous l'avons vu dans l'Apprentissage, les *Classes* sont des classes de caractères définies par l'utilisateur, utilisée dans les commandes de Filtrage et de Recherche. Les classes typiques dans une langue sont les consonnes, les voyelles, les nasales, les ponctuations et tout autre classification utile au linguiste.

- Choisissez la Classe @. (Il s'agit de l'ancien métacaractère “any letter”.)
- Cliquez sur Modifier.³⁶
- * Entrez les C barrés comme dans l'ordre de tri, mais en ligne, entre les C et les D (avec un espace entre chaque caractère).

³⁴ Payne 6, Note de bas de page.

³⁵ Dans la terminologie de Shoebox pour DOS (SH2) cette liste intègre également les “métacaractères”. Les *Métacaractères* sont des *classes* SH2 dont le contenu ne pouvait pas être changé; La plupart de ces métacaractères ont été remplacés par des classes ordinaires. Les valeurs par défaut de SH2 ont été maintenues afin que les filtres existants puissent continuer à fonctionner.

³⁶ Dans la boîte de dialogue *Définition des Classes*, choisissez Aide pour afficher une liste de Variables par défaut et leur sens général. Remarquez que le nom des Variables peut être changé. Cette section de l'aide décrit également les Variables *Prédéfinies*.

Cliquez sur OK.

Les Classes peuvent être utilisées dans les Filtres et dans la commande de recherche Retrouver. Vous pouvez modifier n'importe laquelle d'entre elles, ou en ajouter d'autres³⁷. Comme elles contiennent des caractères spécifiques, on les considère comme parties intégrantes de la langue.

Normalement il faudrait modifier toutes les Classes qui utilisent ces caractères spécifiques, mais nous nous en passerons ici car vous ne souhaitez probablement pas réellement utiliser ce jeu de caractères et cela serait un peu lassant comme exercice.

Cliquez sur OK dans la boîte de dialogue *Propriétés de l'Encodage de langue*.

Choisissez Fermer dans la boîte de dialogue *Encodage de langues*.

Nous sommes finalement prêt pour entrer le premier enregistrement !

Insérer un enregistrement

Toutes les boîtes de dialogue devraient maintenant avoir été fermées. Vous devriez n'avoir plus qu'une fenêtre vide intitulée AXININCA.LEX. Il devrait y avoir **\lx** dans la zone marqueur.

Faites apparaître les propriétés des marqueurs pour **\lx** (**w**:Clic-droit ou **m**:Option+clic sur le marqueur **\lx**).

Spécifier l'encodage de langue pour le marqueur

Ceci nous amène à la boîte de dialogue Propriétés des Marqueurs.

Spécifiez *lexème* comme Nom de Champ.

Choisissez *Axininca Campa* comme Encodage de langue dans la liste déroulante.

Cliquez sur OK.

Tapez *kitairiki* comme première entrée lexicale.

(Si vous utilisez Keyman, vous remarquerez que le clavier Axininca s'est activé.)³⁸

Tapez **w**:Entrée ou **m**:Retour puis tapez **\ge** suivi d'un espace.

Il vous sera demandé si vous voulez ou non ajouter ce nouveau champ à la liste.

Choisissez Oui.

Ceci vous conduira à nouveau dans la boîte de dialogue Propriétés des Marqueurs (avec **ge** comme marqueur).

Entrez *glose* comme Nom de champ.

Entrez Français comme Encodage de langue.

Cliquez sur OK.

Ceci nous ramène à la fenêtre du lexique Axininca.

Entrez *sanglier* comme glose.

(Si vous utilisez Keyman, vous remarquerez que le clavier Axininca n'est plus activé.)

Nous voulons ajouter deux autres champs avant de définir un modèle pour ce type de base de données. Recommencez le processus précédent pour ajouter un marqueur de champ, un nom de champ, et spécifier l'encodage de langue approprié pour chacun d'eux :

Marqueur: **\nt** Nom du champ: *notes* Encodage de langue: *English*

Marqueur: **\dt** Nom du champ: *date* encodage de langue: *Default*

Maintenant nous voulons que le champ **\dt** soit un champ d'estampillage de date (pour conserver une trace de la dernière date d'édition d'un enregistrement).

³⁷ Remarquez que le contenu de la classe Ponctuation (!), est inclus dans la classe prédéfinie de frontières de mot #. Vous pouvez modifier le contenu de ! mais vous ne devez pas changer son nom.

³⁸ Souvenez-vous que la commutation automatique des claviers ne fonctionne sous Windows 95 que si vous avez lancé la version Windows 3.1 de Shoebox.

- Choisissez Base de données, Propriétés (remarquez les nouveaux marqueurs de champs dans la liste).
- Sélectionnez l'onglet Options.
- Descendez dans la liste jusqu'au marqueur d'estampillage de date **dt date**.
- Cliquez sur OK.

Vous avez maintenant défini un jeu minimal de champs pour démarrer un dictionnaire. Pour votre propre travail vous aurez sans doute besoin de beaucoup plus que cela (ou alors vous utiliserez un jeu prédéfini comme MDF), mais cela sera suffisant pour notre exemple.

Créer un modèle pour les nouveaux enregistrements

Maintenant nous allons créer un modèle pour les nouveaux enregistrements—le jeu de marqueurs à insérer automatiquement chaque fois qu'un nouvel enregistrement sera créé.

- Choisissez Base de Données, Modèle.

Une petite boîte de dialogue apparaîtra disant "Le modèle d'enregistrement sera calqué sur l'enregistrement en cours". Remarquez la case à cocher qui vous donne la possibilité d'inclure également le contenu de champs. Parfois, cela peut être un moyen pratique d'avoir certains champs préalablement remplis, mais la plupart du temps, cela sera inutile. Laissez la case non cochée.

- Cliquez sur OK, ce qui fermera la boîte de dialogue.

Insérer le deuxième enregistrement

Pour insérer un second enregistrement,

- Choisissez Base de Données, Insérer un enregistrement ou bien utilisez le raccourci clavier.

Ceci fera apparaître la boîte de dialogue intitulée *Insérer enregistrement*.

- Tapez *imaa* comme Enregistrement à Insérer.

(Si vous utilisez Keyman, vous remarquerez que le clavier Axininca est de nouveau activé.)

- Cliquez sur OK.

Le programme vous ramène à la fenêtre du lexique, avec *imaa* déjà inséré comme contenu du champ d'enregistrement. Votre curseur est positionné sur la ligne **\ge**. Remarquez que le modèle a inséré pour vous les marqueurs **\ge**, **\nt**, **\dt**.

- Tapez *très* comme glose.

Remarquez que la date d'estampillage n'est pas encore fournie. Elle ne sera pas insérée (ou mise à jour) tant que vous n'aurez pas quitté l'enregistrement.

- Cliquez sur le bouton Enregistrement suivant pour aller à l'enregistrement *kitairiki* (remarquez que sa date d'estampillage est vide également, car nous avons spécifié ce champ estampillage qu'après l'édition de l'enregistrement), aussi...
- Ajoutez une note quelconque dans son champ *notes*, par exemple "*Quel est le lexème pour 'tusk'?*"
- Maintenant cliquez sur le bouton Enregistrement précédent pour revenir à *imaa*. (Son champ date devrait maintenant avoir été mis à jour à la date d'aujourd'hui.)
- Cliquez sur le bouton Enregistrement suivant de nouveau pour voir l'enregistrement *kitairiki*. Sa date d'estampillage devrait également être là maintenant.

Ce sera tout pour cet exercice. Nous espérons que vous aurez maintenant une meilleure compréhension des concepts de Projet, Type de base de données et Encodage de langue. Vous avez également travaillé sur un exemple assez complexe de paramétrage de Shoebox, dont l'installation d'une police, la création d'un nouveau projet, d'un nouveau type de base de données et d'un nouvel encodage de langue. Vous avez aussi appris à créer un modèle d'enregistrement et à définir un estampillage de date. Maintenant que vous avez su faire cela, la prochaine base de données sera du gâteau.

- Sortez de Shoebox. Sauvegardez les modifications.

Fin de l'étape Ici s'achève cette étape de l'apprentissage. Rien de ce qui suit ne s'appuie sur les modifications que vous avez apportées dans cette section.

Texte interaligné

Introduction On appelle texte interaligné, un texte comportant des informations sous chacun de ses mots. On trouve généralement un découpage des mots en morphèmes (segmentation), une glose pour chaque morphème, et parfois la catégorie grammaticale à laquelle ils appartiennent (partie du discours). L'interalignement de texte est une manière efficace de construire un lexique et d'étudier le morphologie d'une langue.

Shoebox peut découper les mots en morphèmes et générer le texte interaligné. Il le fait en recherchant les mots et les affixes dans un lexique.

Principe de base de l'interalignement Notre premier exemple est en Frison³⁹.

- Démarrer Shoebox (s'il n'est pas déjà lancé).
- Démarrer Shoebox (s'il n'est pas déjà lancé).
- Choisissez Projet, Ouvrir
- Naviguer jusqu'au répertoire FRISIAN1 dans SAMPLES
- Ouvrez le projet FRI.PRJ.

Observez que le texte frison dans la fenêtre de gauche comporte les découpages en morphèmes, les gloses et les catégories grammaticales. Par exemple, le premier mot *berne* est composé des morphèmes *bern* et *-e*. Le morphème *bern* a comme glose *bear*, et comme catégorie grammaticale *V* (pour verbe).

Pour voir d'où provient la racine *bern*,

- w**:Clic-droit ou **m**:Option-clic sur *bern* pour Sauter à son entrée lexicale.

Le lexique frison dans la fenêtre de droite présente alors l'entrée *bern*. Dans cette fiche vous pouvez voir la catégorie grammaticale *V* et la glose *bear*. Remarquez que la catégorie grammaticale est étiquetée **\ps** dans le lexique, alors qu'elle est étiquetée **\p** dans le fichier texte. Le marqueur utilisé dans le texte interaligné n'est pas forcément le même que celui du lexique.

Pour voir d'où provient le suffixe *-e*,

- w**:Clic-droit ou **m**:Option-clic sur *-e*

Vous verrez apparaître une boîte intitulée *Correspondances multiples* contenant une liste des suffixes *-e*.⁴⁰

- Choisissez le suffixe comportant la glose *PSTP* puis cliquez sur OK⁴¹

Vous pourrez voir alors l'entrée lexicale *-e* appropriée. Le tiret indique qu'il s'agit d'un suffixes. Le segmenteur de Shoebox découpe *berne* en la racine *bern* et le suffixe *-e*.

Maintenant nous allons interaligner une nouvelle phrase.

- Faites défiler la fenêtre vers le bas pour faire apparaître la seconde proposition de la phrase, qui commence par *sil*

³⁹ Le Frison est parlé dans certaines régions des Pays-Bas et de l'Allemagne. Le texte est tiré du livre *Frisian Reference Grammar* de Pieter Meijes Tiersma.

⁴⁰ Il est possible qu'elle contienne aussi 'e si le tiret est ignoré. Ceci peut être corrigé en modifiant le paramétrage du Saut, mais ce n'est pas important à ce niveau.

⁴¹ On peut sélectionner et cliquer sur OK, ou bien double-cliquer dessus.

- Placez le curseur devant *sil*.

La commande Interaligner se trouve dans le menu Outils. Mais comme elle est très utilisée, il vous sera plus facile d'utiliser le bouton Interalignement dans la barre d'outils (juste à gauche du bouton Index) ou encore d'utiliser le raccourci clavier **w:Alt+I** ou **m:commande+I**.

- Cliquez sur le bouton Interalignement ou utilisez le raccourci **w:Alt+I** ou **m:commande+I**.

Vous verrez Shoebox interaligner automatiquement la ligne. Les gloses sont '*will there his grave be*'.

- Déplacez le curseur dans le champ **f** (traduction libre) et ajoutez: *he will die there as well*.

Remarquez que cette ligne de texte était particulièrement facile à interaligner, mais le processus d'interalignement n'est pas toujours aussi simple, certains morphèmes ont plusieurs sens.

- Déplacez le curseur sur le mot *berne* de la première proposition de la phrase.
- Interalignez (cliquez sur le bouton ou utilisez le raccourci).

Vous verrez apparaître une boîte intitulée *Plusieurs résultats* comportant quatre choix. Elle recouvre le texte aussi vous devrez la déplacer.

- Tirez la boîte *Plusieurs résultats* sur la droite de l'écran. (Elle mémorisera la place où vous l'avez amenée et s'ouvrira là la prochaine fois).

Vous verrez que -e est en surbrillance dans le texte. Cela indique que les quatre possibilités de la boîte sont chacune des gloses possibles pour -e. Remarquez que chaque glose est suivie de sa catégorie grammaticale entre accolades. Ceci permet parfois d'orienter le choix.

- Sélectionnez *PSTP{Tns}* et cliquez sur OK (ou bien double-cliquez sur *PSTP{Tns}*).

Vous verrez l'interalignement de *berne* s'exécuter. Remarquez que le processus s'arrête devant le mot suivant de la ligne. Lorsqu'un texte a déjà été interaligné, l'interalignement ne se refait qu'un mot à la fois.

- Choisissez Projet, Fermer.

Vous verrez une boîte vous demandant si vous voulez enregistrer les modifications apportées au texte.

- Choisissez Oui .

Fin de l'étape

Ici s'achève cette étape de l'apprentissage. Rien de ce qui suit ne s'appuie sur les modifications que vous avez apportées dans cette section.

Paramétrage de l'interalignement

Pour illustrer le paramétrage de l'interalignement, nous allons parcourir les étapes qui ont dû être suivies pour configurer l'exemple précédent. Nous allons ouvrir un projet comportant le lexique et le texte interaligné, mais dont le paramétrage de l'interalignement n'est pas fait.

Ouvrir le projet FRI2.PRJ dans le répertoire SAMPLES\FRISIAN2

Vous verrez le même texte que précédemment, mais sans interalignement.⁴² Remarquez que le bouton Interalignement est grisé. Cela indique que l'interalignement n'est pas opérationnel parce qu'il n'a pas encore été configuré.

Avant de démarrer la configuration de l'interalignement, vérifiez que la fenêtre *Frisampl.txt* est bien celle qui est active, et pas *Friirt.dic* (le lexique). L'interalignement est habituellement configuré comme élément des propriétés de base de données des fichiers *texte* (et pas des fichiers *lexique*).

Pour entrer dans le paramétrage de l'interalignement, cliquez dans la fenêtre *Frisampl.txt* (pour l'activer)

Allez dans le menu Base de données, Propriétés, puis choisissez l'onglet Interalignement.

Remarquez que la liste *De... Vers... Processus* est vide. Pour configurer le processus comme dans l'exemple que nous venons de voir, il nous faut ajouter trois lignes à la ligne de texte. La première ligne affichera les morphèmes, la seconde les gloses, et la dernière ligne les catégories grammaticales. La ligne initiale du texte est marquée par **t**. La ligne de morphème sera marquée par **m** et la ligne de catégorie grammaticale par **p**.

En plus de la spécification des marqueurs à utiliser dans le texte interaligné (t,m,p), Shoebox doit également savoir dans quels champs il devra rechercher les informations dans le lexique pour alimenter les champs d'interalignement du texte. Ces marqueurs sont parfois les mêmes que ceux du texte, mais souvent ce n'est pas le cas. Ici, vous pouvez voir que le lexique comporte le champ **fri** pour le champ morphème, **ps** pour la catégorie grammaticale, et **g** pour la glose.

Chaque ligne d'interalignement est paramétrée par l'ajout d'un processus à la liste Processus. Nous pourrions ajouter ces lignes de processus l'une après l'autre, mais la meilleure façon de paramétrer un interalignement, c'est de choisir le Configuration. Configuration demande un minimum d'information et ajoute les trois processus en une fois.

Choisissez Configuration.

Vous verrez apparaître une boîte affichant les *Configuration : Définition des marqueurs d'Interalignement*. Cette boîte demande les marqueurs à utiliser dans le présent texte interaligné. Comme établi plus haut, notre texte initial est marqué **t**. Nous voulons que la ligne des morphèmes soit marquée **m** et la ligne des catégories grammaticales **p**. Il s'agit là des valeurs par défaut, aussi nous n'avons rien à modifier ici.

Cliquez sur OK.

Vous verrez s'afficher une boîte *Configuration rapide des marqueurs lexicaux*. Cette boîte permet de définir les marqueurs dans lesquels Shoebox devra rechercher l'information. Elle demande également le nom du ou des lexique(s) à utiliser.

Nous voulons qu'il utilise le lexique FRIRT.DIC.

Sélectionnez FRIRT.DIC, puis cliquez sur Insérer dans la liste des *Bases de données disponibles*.

Le marqueur de glose par défaut est **g**, ce qui est bien ce qui nous convient pour ce lexique, mais par contre nous voulons **ps** comme marqueur de catégorie grammaticale, et pas **p**.

⁴² Le texte a été préparé par TEXTPREP.CCT, qui découpe le texte kilométrique en phrases en insérant une ligne de référence, un champ de traduction libre, etc.

- Changez le marqueur de *catégorie grammaticale* en **ps**.

Les marqueurs pour *Forme alternative* et *Forme de base* seront explicités plus loin. Ils peuvent être laissés tel quel pour l'instant.

- Cliquez sur OK.

Ceci termine le Configuration. Vous verrez qu'il y a maintenant trois processus d'interalignement listés. Le premier est un processus de Segmentation, les deux autres des processus de Recherche. Chaque processus a une colonne *De...* et une colonne *Vers...*. Ceci indique le marqueur d'où part le processus et le marqueur qu'il produit. Par exemple, le processus de Segmentation part de **t** et produit **m**. Cela signifie qu'il segmente les mots de la ligne **t** et produit une ligne **m** contenant les morphèmes. Les deux autres processus sont du type Recherche partant de la ligne **m**. Le premier processus lit les morphèmes de la ligne **m** et produit une ligne **g** contenant les gloses. Le deuxième lit les morphèmes de la ligne **m** et produit une ligne **p** contenant les catégories grammaticales.

La liste montre également la structure d'interalignement qui sera produite dans le fichier texte. Le marqueur *De...* du premier processus est le marqueur du texte initial. Le marqueur *Vers...* de chaque processus représente la ligne qui sera ajoutée. Vous pouvez constater dans cet exemple, que trois lignes seront ajoutées. Elles seront étiquetées **m**, **g**, et **p**.

Remarquez que le bouton Configuration n'est plus disponible. Il n'est valide que lorsque la liste des processus est vide, aussi si vous voulez recommencer un Configuration, il vous faudra commencer par supprimer tous les processus.

Nous allons profiter d'être dans cette boîte de dialogue pour aller voir rapidement ce qui se cache dans ces processus. Nous n'aurons rien à modifier, mais vous devez savoir où certaines options peuvent être changées.

- Sélectionnez le processus Segmentation et cliquez sur Modifier.

Vous verrez apparaître une boîte de dialogue avec en haut *Du marqueur* et *Vers le marqueur*. Elle contient également des options qui précisent quoi faire si la segmentation échoue, et des endroits où remplacer les caractères signalant les découpages, les caractères de gloses associées, et le caractère de frontière morphophonologique.

On accède à d'autres détails sur le processus avec le bouton Lexiques.

- Cliquez sur Lexiques...

Vous verrez une boîte de dialogue affichant les *Bases de données disponibles*, les *Champs à explorer* et le *Champ à sortir*. Le processus est paramétré pour consulter la base de donnée FRIRT.DIC, rechercher le morphème en cours dans les marqueurs **fri** et **a** de cette base et ressortir le contenu du marqueur **u** de l'enregistrement trouvé. (Les fonctions de la forme variante **\a** et de la forme sous-jacente **\u** seront discutées plus loin).

- Cliquez sur OK pour retourner à la liste des processus.

- Sélectionnez le troisième processus (le second processus de recherche) et cliquez sur Modifier.

Vous verrez une boîte de dialogue très semblable à celle du processus de Segmentation. Elle présente les zones *Du marqueur*, *Vers le marqueur* et les options pour ce qu'il faut faire si la recherche échoue. Elle permet également de spécifier un caractère séparateur de gloses, avec une option pour s'arrêter à la première. Les processus d'interalignement présentent de nombreuses options parce qu'ils peuvent être utilisés de diverses manières, parmi lesquelles différentes façons de faire de la traduction dialectale.

- Cliquez sur Lexiques...

Vous verrez une boîte de dialogue montrant les *Bases de données disponibles*, les *Champs à explorer* et le *Marqueurs à sortir*. Le processus est paramétré pour consulter la base de donnée FRIRT.DIC, rechercher le morphème en cours dans les marqueurs **fri** de cette base et ressortir le contenu du marqueur **ps** de l'enregistrement trouvé.

Remarquez qu'ici apparaît la différence entre le marqueur **p** du texte et le marqueur **ps** du lexique. Dans la boîte de dialogue principale, le marqueur de *Du marqueur*, c'est **p**. Cela signifie que le texte comportera une ligne marquée **p**. Dans la boîte des Lexiques, le *Champ à sortir* c'est **ps**. Cela signifie que le contenu du champ **ps** du lexique sera recopié dans la ligne du texte interaligné (qui a pour marqueur **p**).

- Cliquez deux fois sur OK pour retourner à la liste des processus.
- Cliquez sur OK à nouveau pour retourner à la fenêtre de texte.

Ceci termine le paramétrage de l'interalignement. Pour vérifier que ce paramétrage fonctionne correctement, nous allons interaligner l'échantillon de données.

- Placez le curseur au début de la première ligne \t.
- Interalignez (par le bouton ou la touche de raccourci).

Vous verrez la boîte de dialogue *Plusieurs résultats* apparaître comme précédemment. Elle apparaît la première fois dans le coin supérieur gauche, mais si vous la déplacez, elle mémorisera son nouvel emplacement par la suite.

- Déplacez la boîte de dialogue dans le coin droit en bas de la fenêtre de Shoebox.
- Double-cliquez sur PSTER{Tns} pour le sélectionner

La boîte de dialogue *Plusieurs résultats* apparaît de nouveau, mais cette fois-ci avec le mot *-e* sur le mot *opgroeid*. Vous devrez peut-être déplacer la boîte de dialogue pour voir le *-e* en surbrillance dans le texte. (Le *-d* final de *opgroeid* est segmenté comme *-e* parce que *-d* est une forme variante de *-e*. Vous irez rechercher *-e* dans le lexique pour le voir.⁴³ Pour le moment :

- Double-cliquez sur PSTP{Tns} pour le sélectionner.

Le reste de la ligne s'interaligne. Ceci montre que l'interalignement fonctionne correctement.

- Pour voir la forme variante *-d* de *-e*, **w**:Clic-droit ou **m**:option-clic sur *-e* dans le texte pour sauter vers l'entrée du lexique.
- Dans la boîte *Plusieurs résultats*, sélectionnez le *-e* du bas avec PSTP comme glose.

Remarquez que l'enregistrement *-e*, qui apparaît dans la fenêtre du lexique FRIRT.DIC, contient le champ \a **-d** comme forme variante (alternate) de *-e*.

- Nous en avons terminé avec ce projet, alors faites Projet, Fermer.

Une boîte de dialogue vous demande si vous voulez enregistrer les modifications apportées au fichier texte.

- Cliquez sur OK.

Fin de l'étape

Ici s'achève cette étape de l'apprentissage. Rien de ce qui suit ne s'appuie sur les modifications que vous avez apportées dans cette section.

⁴³ Une forme variante se retrouve par exemple avec l'adverbe *de*, en français, devant un mot commençant par une voyelle, où il se réalise *d'*.

Un peu plus sur l'interalignement

Pour illustrer un peu plus en détail l'interalignement, nous allons traiter un texte en langue Axininca du Pérou.. (Axininca se prononce ASHINinca.) L'axininca est une langue agglutinante avec une morphophonologie significative, aussi elle illustre pleinement le processus de segmentation. Le texte est tiré de l'appendice du livre *The Phonology and Morphology of Axininca Campa* de David Payne.⁴⁴

La section *Démarrer un nouveau projet et une nouvelle base de données*, p. a défini l'encodage de la langue Axininca et démarrer un lexique Axininca. Elle précisait également comment installer la police Axininca. Si vous avez parcouru cette section et installé les polices Axininca, alors les données de l'exemple apparaîtront correctement. Si vous n'avez pas encore installé les polices Axininca, vous n'êtes pas obligé de le faire maintenant, mais vous constaterez que certains mots axininca apparaîtront différemment dans Shoebox que dans cette section (ce n'est pas important et n'a aucun effet sur la leçon).

L'échantillon d'interalignement comporte un lexique Axininca et des fichiers de configuration provenant d'un répertoire de projets différent de celui précédemment utilisé dans cette section.

- Si vous voyez une boîte de dialogue *Pas de projet ouvert*, choisissez Ouvrir un projet existant et cliquez sur OK.
- Si un projet est déjà ouvert dans Shoebox, choisissez Projet, Ouvrir dans le menu.
- Naviguez dans le répertoire AXINT dans SAMPLES.
- Ouvrez le projet AXINT.PRJ.

Vous devriez voir un texte intitulé WILDBOAR.TXT et un lexique AX.LEX s'ouvrir dans Shoebox. Le texte comporte une traduction libre après chaque phrase, mais il n'est pas interaligné.

Paramétrage de l'interalignement

Cet exemple n'est pas encore configuré par l'interalignement. Le paramétrage de l'interalignement fait partie du type de base de données.

- Vérifiez que c'est la fenêtre du texte qui est activée (et pas celle du lexique) en cliquant n'importe où dans la fenêtre WILDBOAR.TXT.
- Cliquez dans le menu Base de données, Propriétés.

Vous verrez une boîte de dialogue intitulée *Propriétés du type de base de données–Axininca Interlinear Text*.

- Cliquez sur l'onglet Interalignement

Vous verrez une liste vide intitulée *De... Vers... Processus*. Sur la droite il y a un bouton Configuration. Encore une fois, comme il a été dit dans la leçon précédente, c'est la façon la plus simple de démarrer le paramétrage.

- Cliquez sur Configuration.

Vous verrez la boîte de dialogue intitulée *Définition des marqueurs d'interalignement*. Il s'agit des marqueurs qui seront utilisés dans le texte interaligné. La boîte de dialogue présente quatre marqueurs : *Marqueur de texte*, *Segmentation*, *Glose* et *Catégorie grammaticale*. Les marqueurs par défaut sont **t**, **m**, **g** et **p**. Nous utiliserons ces marqueurs, et nous n'aurons donc rien à changer ici.

- Cliquez sur OK.

Vous verrez la boîte de dialogue intitulée *Définition des marqueurs d'interalignement*. Cette boîte demande les marqueurs à rechercher dans le lexique pendant l'interalignement.

Le haut de la boîte de dialogue contient une liste *Base de données disponibles* montrant les bases de données actuellement chargées et une liste vide des *Bases de données disponibles*.

⁴⁴ Nous nous sommes efforcés de générer les mêmes analyses que celles du livre, sauf que nous avons abrégé certaines gloses.

C'est là que nous précisons les noms des bases de données que Shoebox devra consulter. Dans notre exemple, il n'y en a qu'une.⁴⁵

- Sélectionnez *AX.LEX* dans la liste des bases de données.
- Cliquez sur Insérer pour copier *AX.LEX* du côté *Bases de données disponibles*.

La partie basse de la boîte de dialogue comporte quatre marqueurs : *Marqueur de glose*, *Part. du discours* (cat. gram.), *Forme alternative* et *Forme de base*. Les marqueurs par défaut sont **g**, **p**, **a** et **u**. Ce ne sont pas là les marqueurs utilisés pour la glose et la catégorie grammaticale dans le lexique, aussi nous devons les remplacer.

- Remplacer le marqueur de glose par **ge**.
- Remplacer le marqueur de catégorie grammaticale par **ps**.

Les autres marqueurs sont corrects et peuvent être ignorés. Ils seront explicités plus loin.

- Cliquez sur OK.

Comme dans l'exercice précédent, vous verrez la liste des processus d'interalignement montrant les trois processus définis pour vous. La liste des marqueurs *Vers...* montre les lignes d'interalignement qui seront créées par Shoebox lorsque le processus sera lancé. Ces lignes seront **m**, **g**, **p**.

Les noms des processus indiquent ce qu'ils font. Un processus *Segmentation* segmente les mots en morphèmes. Un processus *Glose* recherche les mots ou les morphèmes dans le lexique. Remarquez que les deux processus *Glose* recherchent les morphèmes provenant de la ligne **m**. Le premier retrouve la glose du morphème et le second la catégorie grammaticale du même morphème.

- Cliquez sur OK pour retourner à la fenêtre principale.

Interalignement

- Placez le curseur au début de la première phrase du texte (dans le champ \t).
- Lancez l'interalignement (par le bouton ou la touche de raccourci).

Vous verrez Shoebox interaligner la première phrase. Le mot *kitairiki* était la seule entrée du lexique. Il s'interaligne avec la glose *wild.boar* et la catégorie grammaticale *N*, comme cela figure dans le lexique.⁴⁶ Les autres mots présentent des étoiles en dessous d'eux pour montrer qu'ils n'ont pas été trouvés. (Si *kitairiki* n'est pas correctement interaligné, retournez au début du paramétrage de l'interalignement, supprimez l'ensemble des processus, et relancez le Configuration. Retournez au texte, effacez toutes les lignes d'interalignement mauvaises et relancez le processus.)

Après l'interalignement d'une phrase, le curseur se place en dessous de la dernière ligne interalignée. A ce niveau, si cela n'était déjà fait dans ce texte d'exemple, vous auriez à entrer la traduction libre du texte que vous venez d'interaligner. Pour continuer, il *n'est pas* nécessaire de déplacer le curseur sur la ligne suivante, le processus d'interalignement passe automatiquement à la phrase suivante.

- Cliquez sur le bouton d'interalignement pour sauter en dessous, sur la ligne suivante du texte.

Vous verrez la seconde ligne s'interaligner, ne montrant que des marques d'échec (souvenez vous qu'il n'y a qu'un mot dans le lexique). Le premier mot de la seconde phrase (*ti*) signifie *no*. Pour insérer le mot *ti* dans le lexique,

- Pointez le curseur de la souris sur le mot *ti* et Sauter-Insérer-le (**w:Ctrl+Clic-droit** ou **m:controle+Option+clic**).⁴⁷

⁴⁵ Vous pourriez avoir plus d'une base de données lexicales si, par exemple, vous décidez de séparer les racines des radicaux, ou isoler les noms propres et autres données spéciales de votre base de données principale.

⁴⁶ Il est courant d'utiliser un point pour relier des mots dans une glose à mots multiples comme *wild.boar*. Bien que cela ne soit pas indispensable, cela rend la lecture des gloses à mots multiples du texte plus facile.

⁴⁷ Vous pourriez utiliser un *Sauter vers* normal ici, et ne trouvant pas *ti*, il vous serait proposé de l'insérer.

La fenêtre de lexique se transforme en une nouvelle entrée pour *ti*. Les champs Catégorie grammaticale et Glose sont vides.

- Entrez la catégorie grammaticale *Neg*.
- Entrez la glose *no*.

Le résultat devrait être :

```
\x  ti
\ps Neg
\ge no
```

Maintenant nous pouvons interaligner le mot de nouveau et le voir réussir.

- Placez le curseur devant le mot *ti* dans la fenêtre du WILDBOAR.TXT.
- Interalignez (cliquez sur le bouton).

Vous verrez le mot s'interaligner avec comme glose *no* et comme catégorie grammaticale *Neg*.

Vous venez de faire la séquence normale des opérations d'interalignement :

1. Interaligner une phrase.
2. Pour chaque mot en échec, sélectionner le mot ou sa racine et utiliser le *Sauter-Insérer* pour le placer dans le lexique.⁴⁸
3. Remplir les champs de l'entrée lexicale, et
4. Interaligner le mot de nouveau.

Comme exercice pratique,

- Sauter-Insérez le mot *apaani* dans le lexique avec la catégorie grammaticale *Pron* et la glose *one*.

Le résultat devrait être :

```
\x  apaani
\ps Pron
\ge one
```

- Placez le curseur sur le premier *apaani* de la fenêtre de texte puis interalignez.

Remarquez qu'il y a deux occurrences du mot *apaani* dans une ligne, si bien que la deuxième pourra s'interaligner sans difficulté. (Nous parlerons de la réduplication dans la section intitulée Réduplication, p.)

- Interalignez de nouveau pour traiter le second *apaani*.

Affixes

Le processus de Segmentation sépare les affixes et gère les transformations morphophonologiques des affixes et des racines.

Le mot le plus simple de notre exemple composé de plusieurs morphèmes est le dernier mot de la première phrase *aNtami* qui signifie *in the jungle*. Il comporte un suffixe et se découpe en *aNtami-ki*, glosé *jungle-Loc*.

- Déroulez la fenêtre de texte vers la première phrase.

Nous allons commencer par entrer *-ki* dans le lexique. Pour Sauter-Insérer une partie seulement d'un mot dans le lexique, nous sélectionnons d'abord la partie du mot avant de faire Sauter-Insérer.

- Sélectionnez *ki*.

⁴⁸ Vous pouvez demander à Shoebox de s'arrêter et d'insérer les mots manquants dans le lexique. Pour cela, cochez l'option *Insérer dans le lexique* dans le processus Segmentation. Maintenez également cochée l'option *Afficher une marque d'échec*, ainsi si vous choisissez de sauter un mot et de ne pas l'insérer dans le lexique, Shoebox le marquera pour vous.

- Sauter-Insérez.
- Maintenant, première chose, ajoutez un tiret devant *ki* dans l'entrée lexicale (c.à.d. *-ki*). C'est très important ! Sans ce tiret, *ki* n'est pas un suffixe, et Shoebox le traitera comme une racine.

NOTE: Le tiret d'un affixe est très important et est très facile à oublier, aussi, chaque fois qu'un affixe ne sera pas reconnu, pensez à vérifier que vous n'avez pas oublié son tiret.

- Ajoutez la catégorie grammaticale *Loc* et la glose *in*.

Le résultat devrait être:

```
\lx  -ki
\ps  Loc
\ge  in
```

Pour voir ce qui se passe lorsqu'un affixe est dans le lexique mais que la racine n'y est pas,

- Interalignez *aNtamiki*

Vous verrez le suffixe *-ki* correctement découpé et glosé, et des astérisques à la place où la racine aurait été si elle avait été connue.

Il existe une option dans la paramétrage de l'interalignement qui peut rendre ceci encore plus intéressant, mais ce n'est pas la configuration par défaut de Configuration. Cette option présente le reste du mot après découpage des affixes avec une astérisque devant pour montrer qu'il y a échec. Pour demander cette option:

- Cliquez dans le menu Base de données, Propriétés.
- Cliquez sur l'onglet Interalignement.
- Sélectionnez le processus segmentation et cliquez sur le bouton Modifier.

Vous verrez la boîte de dialogue Segmentation avec un certain nombre d'options.

- Validez la case à cocher intitulée *Afficher la racine présumée* (conservez l'option *Afficher la marque d'échec*).
- Cliquez deux fois sur OK pour revenir à la fenêtre principale.
- Maintenant Interalignez de nouveau *aNtamiki* pour voir la différence que cela fait.

Vous verrez **aNtami -ki* comme découpage. Lorsque la racine supposée est correcte, comme ici, entrer la racine dans le lexique est très simple puisqu'il suffit de faire une Sauter-Insérer sur la racine, sans avoir à la sélectionner. Bien sûr parfois la racine supposée sera erronée, mais cette option reste néanmoins une aide.

- Entrez *aNtami* dans le lexique. Vous pouvez le faire par un Sauter-Insérer sur **aNtami*. NOTE: Comme l'astérisque *n'est pas* une lettre de la langue Axininca, il ne sera pas inclus dans le lexème entré par Sauter-Insérer.
- Entrez la catégorie grammaticale *N* et la glose *jungle*.

Le résultat devrait être:

```
\lx  aNtami
\ps  N
\ge  jungle
```

- Interalignez *aNtamiki* à nouveau pour voir si la segmentation est correcte maintenant.

Morphophonologie

La morphophonologie renvoie aux modifications subies par les sons de certains morphèmes en fonction des morphèmes environnants. Cela se reflète souvent dans l'orthographe. Shoebox dispose d'un puissant système de traitement de la morphophonologie, et l'Axininca a une morphophonologie riche.

Nous allons commencer avec le mot *isaiki*. Il se compose de trois morphèmes: *ir- saik -i*, glosé *3M- be,sit -NF*. (*3M* signifie 3ème personne masculin et *NF* non-futur.) Le préfixe *ir-* présente une modification morphophonologique par laquelle *r* chute devant *s*.

Nous ne pouvons pas sélectionner *ir* et l'insérer dans le lexique par Sauter-Insérer puisque qu'il n'apparaît pas tel quel dans le texte. Nous allons donc insérer *i* dans le lexique, puis lui rajouter le *r*.

- Sélectionnez la lettre *i* et faites Sauter-Insérer.
- Corrigez la nouvelle entrée en rajoutant *r* et le tiret pour faire *ir-*. *N'oubliez surtout pas le tiret !*
- Entrez la catégorie grammaticale *S* (pour sujet) et la glose *3M*.

Vous devriez avoir :

```
\x ir-
\ps S
\ge 3M
```

- Interalignez le mot *isaiki* de nouveau et constatez le résultat.

Vous verrez que *-ki* a été isolé en fin de mot, mais rien n'a été fait en tête du mot. *-ki* est un mauvais découpage qui disparaîtra lorsque la racine sera entrée dans le lexique.

Le préfixe *ir-* n'a pas été trouvé parce que nous n'avons pas encore signalé la transformation morphophonologique qui le change en *i-* devant *s*. Pour cela, nous allons utiliser les champs *Forme alternative* et *Forme de base* dans le lexique. (Vous vous souvenez que le *Configuration* avait retenu les marqueurs *la* et *lu* pour ces champs).

Pour dire au segmenteur que *r* disparaît devant *s*, nous devons écrire que le morphème *ir-* devient *i-* devant *s*.⁴⁹ Pour exprimer cela, nous disons que la forme sous-jacente *ir+s* devient *is*. Pour entrer cette information dans le lexique, en utilisant les champs *Forme alternative*,

- Entrez les deux champs suivants en-dessous de l'entrée *ir-*:⁵⁰

```
\a is-
\u ir+s
```

Encore une fois, ceci peut se lire ainsi : *ir-* devient *i-* devant *s*, ou si vous préférez la séquence de surface *is-* provient de la forme profonde *ir-* plus *s*.

Comme l'objectif principal de cette notation est la segmentation, la forme de surface est placée avant la forme profonde, mais si vous raisonnez en générativiste, cela pourra vous aider de penser à la forme profonde en premier; mais soyez vigilant à respecter l'ordre des champs tel que présenté ici.

NOTE: Un principe important est que la forme de surface doit comporter un tiret pour signifier que c'est un affixe, mais la forme profonde comporte généralement un symbole de frontière de morphème plutôt qu'un tiret. Le symbole de frontière de morphème par défaut est le signe plus (+).

Voici comment le segmenteur de Shoebox opère: lorsqu'il essaye de découper un mot comme *isaiki*, et qu'il constate que *is-* est un préfixe figurant dans le lexique, il regarde la forme sous-jacente (*ir+s*) et remet la partie après la frontière de morphème (le *s*) dans le reste du mot (*aiki* qui devient *saiki*) qu'il essaye de segmenter à nouveau. Le découpage du reste du mot aboutira ou non suivant le contenu le lexique, c.à.d. que si *saiki* existe dans le lexique comme racine, le découpage réussira; autrement, il échouera.

- Interalignez *isaiki* de nouveau.

⁴⁹ Nous avons le projet d'un système de règles génératives qui permettrait de dire que *r* chute devant *s*, mais ce n'est pas encore implémenté dans cette première livraison. La notation utilisée ici provient de SH2.

⁵⁰ Le signe plus dans la police Axinica semble présenter un espace devant et derrière, mais il n'en est rien. Il ne doit pas y avoir d'espace avant ni après le signe plus. Si c'était le cas, la segmentation ne fonctionnerait pas.

Vous verrez le préfixe *ir-* et le suffixe *-ki* isolés, laissant la racine présumée **sai*. Ceci n'est pas tout à fait correct, car la vraie racine est *saik*, mais on s'en rapproche. La transformation morphophonologique de *ir-* a bien été prise en compte.

Il nous manque encore deux entrées lexicales pour achever le découpage de ce mot: la racine et le suffixe correct. Commençons par le suffixe *-i*.

- Sélectionnez la lettre *i* et faites un Sauter-Insérer.
- Rajoutez un tiret devant le *i* de l'entrée. Souvenez vous que c'est très important !
- Entrez la catégorie grammaticale *Tns* et la glose *NF*.

Le résultat devrait être:

```
\lx  -i
\ps  Tns
\ge  NF
```

- Sélectionnez *saik* dans la ligne `\t` et faites un Sauter-Insérer. (Comme *saik* est une racine, ne mettez pas de tiret ici).
- Entrez la catégorie grammaticale *V* et la glose *be,sit* (sans espace entre).

Le résultat devrait être :

```
\lx  saik
\ps  V
\ge  be,sit
```

La virgule est souvent utilisée pour signifier "ou" dans une glose de plusieurs mots. *saik* signifie *be* ou *sit*. Il est assez rare de trouver une glose exactement équivalente, aussi, parfois, une telle technique permet de clarifier le sens.

- Interalignez *isaiki* de nouveau.

Cette fois-ci le découpage est correct. Vous avez maintenant trois ou quatre mots de la première phrase interalignés.

Il y a une chose que nous n'avons pas encore fait, c'est de paramétrer l'encodage de langue et son affectation aux lignes d'interalignement. La seconde ligne, celle du découpage en morphèmes, doit être en langue Axininca.

- w**:Clic-droit ou **m**:option-clic pour faire apparaître les *Propriétés du marqueur*.
- Spécifiez l'*encodage de langue Axininca*.
- Cliquez sur OK.

La seconde ligne apparaît maintenant en Axininca, et en gras. La ligne principale du texte serait mieux mise en évidence si cette ligne de morphèmes n'était pas en gras, nous allons donc la mettre en police normale.

- w**:Clic-droit ou **m**:option-clic sur le marqueur `\m` pour faire apparaître les *Propriétés du Marqueurs*.
- Cliquez dans la case *Utiliser police d'encodage* pour la désactiver
- Cliquez sur le bouton Choisir la police
- Sélectionnez normal au lieu de gras
- Cliquez sur OK, OK pour revenir à la fenêtre principale.

La ligne de morphèmes est maintenant dans un style normal (mais toujours en police Axininca—si cette police avait été installée, voir le début du chapitre).

Une autre chose qui peut aider à la lisibilité du texte interaligné, c'est l'utilisation de styles ou de couleurs différentes pour distinguer chaque ligne. Si vous avez un affichage couleur, nous allons choisir de mettre la glose en rouge, et la catégorie grammaticale en bleue. Si vous avez un affichage noir et blanc, nous mettrons la glose en gras italique et la catégorie grammaticale en italique.

- w**:Clic-droit ou **m**:option-clic \g pour faire apparaître les *Propriétés du marqueur*.
- Cliquez dans la case *Utiliser police d'encodage* pour la désactiver
- Cliquez sur le bouton Choisir la police
- Cliquez sur OK, OK pour revenir à la fenêtre principale.
- Faites de même pour mettre la ligne \p en bleu, ou en italique pour un écran N&B.

Voyez comme les différentes lignes se distinguent bien maintenant.

Penchons nous maintenant sur le premier mot de la phrase. Il présente de la morphophonologie dans le suffixe. Le *c*-barré (qui s'affiche comme un A accentué dans les autres polices) est une épenthèse entre les deux *i*, et le *t* est une épenthèse entre *i* et *a*. Le mot se découpe comme *ir- ii -ai -i -ri* glosé *3M- Nom -PASS -NF -PASS*, et signifie *what is called*.⁵¹

Dans ce qui suit, le plus difficile c'est de ne pas oublier de mettre les tirets dans les entrées affixes du lexique. S'il vous arrive de ne pas obtenir le bon résultat, vérifiez que vous n'avez pas oublié un tiret.

- Entrez le suffixe *-ri* dans le lexique. Ajoutez la catégorie grammaticale *Rel* et la glose *REL*. N'oubliez pas le tiret devant *-ri*.

Vous devriez avoir :

```
\x  -ri
\ps  Rel
\ge  REL
```

- Interalignez le mot *hiitaiçiri* (*hiitaiAiri* dans une autre police). Vous devriez voir le découpage du *-ri* et également du *-i* que nous avons entré précédemment.
- Entrez le suffixe *-ai* dans le lexique. Mettez un tiret devant *-ai*. Ajoutez la catégorie grammaticale *Pass* et la glose *PASS* (pour Passif).

Vous devriez avoir :

```
\x  -ai
\ps  Pass
\ge  PASS
```

- Interalignez de nouveau et constatez que *rien n'a changé*, car l'épenthèse *c*-barré n'a pas été spécifiée.

Pour cela, nous devons entrer une forme variante sous le morphème *-i*. NOTE: Un important principe à retenir est qu'*une règle morphophonologique entre deux morphèmes doit toujours être placée dans l'entrée lexicale de l'affixe le plus extérieur*. Ici, nous n'entrerons pas la règle sous l'affixe *-ai*, mais plutôt sous *-i* qui est le plus extérieur dans le mot.

- Sautez à l'entrée lexicale *-i* (**w**:Clic-droit ou **m**:Option+clic).
- ajoutez en-dessous de l'entrée:

```
\a  -içi
\u  i+i
```

(où ç est le *c*-barré axinica que vous pouvez insérer par copie depuis la fenêtre de texte ou par la combinaison Alt+0193 sur la pavé numérique dans le système Windows; il s'affiche Å dans une autre police).

Ces deux champs signifient que lorsque le suffixe *-i* suit une autre *i*, alors un *c*-barré s'insère entre eux. N'oubliez pas le tiret devant *-içi*.

⁵¹ la glose des affixes signifie troisième personne masculin, passif, non-futur, relatif.

Remarquez que même si cela ressemble à une règle phonologique, ce n'en est pas vraiment une, car la partie à droite du signe plus doit être un suffixe à part entière, pas seulement une lettre. Cela ne peut pas non plus exprimer une règle s'appliquant à une classe, une voyelle par exemple. Nous avons le projet d'ajouter la capacité d'implémenter de vraies règles phonologiques dans une future version de Shoebox.

Notation morpho-phonologique

La même notation des règles morphophonologiques s'applique pour les préfixes comme pour les suffixes, mais dans des directions opposées, ce qui peut amener à des confusions.

Les deux principes de base relatifs à la notation des règles morphophonologiques sont:

1. Les transformations morphophonologiques seront toujours formulées pour le *morphème le plus externe*, et pas le plus interne. Les transformations dans les racines seront toujours formulés dans l'affixe, pas dans les racines. Ce qui est très bien puisque les affixes sont bien moins nombreux que les racines, et le nombre de morphèmes nécessitant des formes variantes pour expliciter des transformations morphophonologiques sera ainsi limité.
2. La partie de la forme sous-jacente du même côté de la frontière de morphème que le tiret est retenue pour la suite du découpage. La partie de l'autre côté de la frontière de morphème est isolée comme étant la forme sous-jacente de l'affixe.

Vérifions que ce que nous avons fait jusque là fonctionne correctement :

- Interalignez le mot *hiitaiétri* de nouveau et observez que *-ai* est découpé maintenant. (Sinon, vérifiez que vous n'avez pas oublié le tiret.)
- Entrez la racine *ii* avec comme catégorie grammaticale *V* et comme glose *Nom*.

Vous devriez avoir :

```
\lx  ii
\ps  V
\ge  Nom
```

Il reste deux transformations à décrire pour que le mot soit correctement segmenté. La première, c'est *t* qui est en épenthèse entre le *-i* final de *ii* et le *a* de *-ai*

- Sauter à l'entrée *-ai* et ajoutez ce qui suit :

```
\a -atai -itai
\u i+ai
```

Ce qui signifie que lorsque le suffixe *-ai* suit un *i*, un *t* s'insère entre eux..

- Interalignez de nouveau.

Remarquez que maintenant le découpage produit **h -i -i -ai -i -ri*. Les deux *-i* internes sont faux, mais ce n'est pas grave car ils disparaîtront lorsque le mot sera entièrement décrit. Lorsque certaines informations ne sont pas trouvées, le programme a tendance à voir plus d'affixes qu'il n'y en a réellement.

La dernière transformation morphophonologique à prendre en compte est la transformation de *ir-* en *h* devant *i*.

- Sauter vers l'entrée *ir-* et ajoutez ce qui suit en bas de l'enregistrement (en-dessous de l'autre forme variante):

```
\a  hi-
\u  ir+i
```

Vous devriez avoir :

\lx **ir-**
 \ps S
 \ge 3M
 \a **is-**
 \u **ir+s**
 \a **hi-**
 \u **ir+i**

Ce qui signifie que le préfixe *ir-* se change en *h* lorsqu'il précède un *i*.

- Interalignez de nouveau et observez que le mot se découpe correctement maintenant :

hiitaiçiri
ir- ii -ai -i -ri
 3M- Nom -PASS -NF -REL
 S- V -Pass -Tns -Rel

Cela peut sembler avoir été un rude travail, mais une fois que les affixes et leurs règles de transformations sont entrés, cela devient très intéressant. Ce qui reste à faire alors, c'est d'entrer les racines de chaque mot nouveau.

Restructuration du texte interaligné

La première phrase devrait être entièrement interalignée maintenant. Vous aurez remarqué que les phrases ont tendance à s'allonger lors de l'interalignement. Ainsi, la première ligne déborde actuellement du bord droit de la fenêtre. Pour remédier à cela nous allons restructurer le texte. Vous définissez la marge en dimensionnant la fenêtre à la largeur que vous voulez et vous spécifiez la longueur de ligne à cette largeur.

- Allez dans le menu *Outils, Enregistrer la longueur de ligne*

Vous verrez apparaître une boîte confirmant que la longueur de ligne sera définie à la largeur de l'affichage en cours.

- Cliquez sur OK
 Placez le curseur dans la première phrase
 Allez dans le menu Outils et faites Restructurer le fichier entier

Remarquez que la phrase et tous les champs interalignés sont maintenant ajustés de façon à être visibles dans la largeur de la fenêtre. La restructuration d'un texte interaligné se fait de la même façon qu'un texte ordinaire. Vous pouvez restructurer entièrement un fichier interaligné en choisissant dans le menu Outils, Restructurer le fichier entier.

Un peu plus sur la morphophonologie:

Le deuxième mot de la deuxième phrase (*haniiçiri*—qui s'affiche *haniiçiri* dans une autre police) est le seul mot ayant échoué à la segmentation. Il doit être découpé selon *ir- anii -i*, glosé *3M- walk -NF*, avec comme catégorie grammaticale *S- V -Tns*. Pour obtenir ce découpage, vous devrez entrer la racine *anii*, ainsi qu'une autre transformation morphophonologique de *ir-* qui dira que *ir-* devient *h* devant *a* (le suffixe *-i* a déjà été traité).

Ceci vous permettra de mettre en pratique ce que vous venez d'apprendre au sujet des transformations morphophonologiques.

- Entrez la racine et la transformation morphophonologique adéquate pour que *haniiçiri* soit correctement découpé.

La racine devrait être :

\lx **anii**
 \ps V
 \ge walk

L'entrée lexicale du suffixe *ir-* devrait être maintenant :

```

\lx ir-
\ps S
\ge 3M
\la is-
\lu ir+s
\la hi-
\lu ir+i
\la ha-
\lu ir+a

```

Interalignez

Vous devriez maintenant voir le texte entièrement interaligné ainsi :⁵²

```

haniçi
ir- anii -i
3M- walk -NF
S- V -Tns

```

Ceci achève l'interalignement des deux lignes du texte Axaninca WILDBOAR. Ce qu'il est intéressant de noter ici, c'est qu'en faisant cet interalignement le lexique est passé d'un mot à douze. Admettons-le, au début ce processus peut paraître long et fastidieux (particulièrement pour les langues à morphophonologie lourde), mais une fois que les affixes principaux auront été bien définis, le processus d'interalignement de texte (et l'accroissement simultané du lexique) gagnera en rapidité.

Faites Projet, Fermer

Vous verrez apparaître une boîte demandant si vous voulez enregistrer les modifications apportées au fichier texte

Choisissez Oui .

Ceci conclut le didacticiel sur l'interalignement. Il restera sans doute de nombreux points qui vous resteront obscurs. Pour plus de détails théoriques sur le processus d'interalignement dans Shoebox, reportez vous à la section *Notes sur l'interalignement*, p.

Traduction dans des langues apparentées.

Cet apprentissage ne s'est pas appesanti sur un volet très important de Shoebox– sa capacité à traiter et adapter des textes dans des dialectes ou des langues apparentées. Cette fonctionnalité est connue sous le nom d'Adaptation en langues apparentées (Adaptation pour simplifier) et se rapproche des logiciels DOS CARLA ou CADA utilisé par certaines équipes sur le terrain.

Une série édifiante de démonstrations des capacités de Shoebox en Adaptation se trouve dans le répertoire ADAPT en dessous de SAMPLES. Ouvrez le projet ADAPT.PRJ dans ce répertoire pour plus de détail sur l'utilisation de ces échantillons.

⁵² Si vous n'obtenez pas ce résultat (c.à.d. si vous obtenez une racine inconnu *aan* ou autre chose), vous avez sans doute entré la forme variante incorrectement. Voici une faute courante dans la spécification de la règle morphophonologique:

```

*\a h-
*\u ir+a

```

Ceci ne marche pas parce que ça enlève le *h* et le remplace par *ir* mais cela ajoute en même temps un *a* à la racine (conduisant à *aan*). vous devez inclure le *a* dans la séquence recherchée (*\a ha-*).

Note: il est possible de laisser tomber le *a* partout:

```

*\a h-
*\u ir-

```

Ceci conduirait à un découpage correct dans cet exemple. Mais cette règle est trop générale car elle autorise Shoebox à enlever n'importe quel *h-* et à le remplacer par *ir-*. Cette règle tendra à produire des ambiguïtés.

Fin de l'étape Ceci conclut l'apprentissage. Le reste de cette documentation décrit de nombreuses fonctionnalités importantes de Shoebox qui n'ont pas été abordées dans cet apprentissage, donne des trucs pour contourner certaines limites de Shoebox, et répond aux questions les plus courantes que nous avons rencontrées.

SECTION 2

RÉFÉRENCES ET QUESTIONS FRÉQUENTES

Principes de base

Les Projets Dans Shoebox, un *Projet* c'est l'espace de travail, c'est-à-dire le nombre et la position des fenêtres ouvertes, quels fichiers elles contiennent, et comment elle sont triées et filtrées.

L'emplacement du fichier de paramétrage du Projet détermine l'emplacement des fichiers d'encodages de langues, et des fichiers de Types de bases de données parce qu'un fichier de Projet demande que les autres fichiers de paramétrage se trouvent dans le même dossier que lui. (Les fichiers de données devraient eux être maintenus ailleurs.)

Différents fichiers de Projet peuvent être dans le même dossier et ainsi partager les mêmes fichiers d'Encodage de langues et de Types de bases de données.

Les utilisateurs individuels tireront un avantage évident à n'avoir qu'un dossier pour tous leurs projets.

Les consultants préféreront sans doute avoir des dossiers différents : un pour chaque personne suivie. Ceci évitera des redéfinitions intempestives de types de bases de données et d'encodages de langues.

Pour changer de Projet alors que Shoebox est déjà ouvert, faites *Projet, Ouvrir* et choisissez un Nouveau Projet.

Vous pouvez créer un nouveau projet en faisant *Projet, Nouveau* et en configurant l'espace de travail à votre guise.

Un projet sera sauvegardé dans chacun des cas suivants:

- En chargeant un nouveau projet (*Projet, Ouvrir*),
- En créant un nouveau projet (*Projet, Nouveau*),
- En le sauvegardant explicitement (*Projet, Enregistrer*),
- En l'enregistrant sous un nouveau nom (*Projet, Enregistrer Sous*)
- En fermant le projet en cours (*Projet, Fermer*)—ce qui vous donne l'opportunité d'en ouvrir un autre, d'en créer un nouveau ou de sortir de Shoebox.
- En quittant le programme (*Fichier, Quitter*).

(comme vous voyez, Shoebox n'a vraiment pas envie de perdre les informations de Projet!)

Shoebox démarrera toujours sur le Projet sur lequel vous avez travaillé la dernière fois. Si vous le voulez, il est possible de faire en sorte que Shoebox démarre toujours avec le projet de votre choix. Voici comment faire (nous supposons que Shoebox n'est pas lancé):

Les utilisateurs **Windows** peuvent ouvrir un projet tout simplement en double-cliquant sur le nom du fichier de projet dans l'Explorateur (Gestionnaire de fichiers, sous Windows 3.1).⁵³ Mais si vous ne voulez pas utiliser l'Explorateur (Gestionnaire de fichiers) pour partir à la recherche d'un fichier de Projet, vous pouvez créer un raccourci (icône) vers celui-ci. Le chemin complet et le nom du fichier de Projet doivent figurées sur la ligne de

⁵³ Ceci marche parce que le programme d'installation de Shoebox associe les fichiers d'extension .PRJ au logiciel Shoebox. Si vous utilisez Windows'95, et que vous avez installé les deux versions de Shoebox, mais que vous décidez d'utiliser la version 16 bits pour sa capacité à commuter les claviers Keyman automatiquement, vous devrez changer l'association. Vous pourrez le faire par l'Explorateur Windows, menu Affichage, Options des dossiers, Types de fichier. Mais il est plus facile de simplement réinstaller la version 16 bits de Shoebox. Ceci n'effacera pas la version 32 bits, par contre le programme d'installation associera les fichiers .PRJ à la version 16 bits.

Cible (ligne de commande, sous Win 3.1) pour que cela marche. En double-cliquant sur ce raccourci (icône) Shoebox chargera le Projet correspondant.

Pour lancer un projet particulier, les utilisateurs **Mac** n'auront qu'à cliquer sur le fichier de projet plutôt que sur l'icône du programme Shoebox.

Les Marqueurs de Champs

Définition des marqueurs

Les marqueurs de champ sont des étiquettes qui repèrent chaque élément logique des données. Un marqueur de champ consiste en un caractère antislash (barre oblique inverse) suivi d'une combinaison de caractères alphabétique, chiffre ou ponctuation.⁵⁴

Le premier marqueur d'un enregistrement, appelé justement marqueur d'enregistrement, représente le point de départ de l'enregistrement et la fin de l'enregistrement précédent. Aussi, le marqueur d'enregistrement ne peut apparaître qu'une fois dans un enregistrement.

Tout autre marqueur peut apparaître autant de fois que nécessaire pour ce type de données et pour cet enregistrement particulier.

Liste des Marqueurs

Une liste des marqueurs est tenue à jour pour chaque base de données. Un fichier appelé fichier de *type de base de données* contient cette information et l'actualise d'une session à l'autre. Ceci signifie qu'elle n'est pas perdue lorsque vous quittez le programme.

Pour voir la liste des marqueurs pour la base de données en cours, choisissez Base de données, Propriétés. Pour voir la liste des marqueurs d'un autre type, faites Projet, Type de Base de Données.

- Les marqueurs en cours d'utilisation sont en Gras; les marqueurs inutilisés ne sont pas en Gras.
- Un marqueur inutilisé est un marqueur qui n'est contenu par aucun enregistrement.
- Seuls les marqueurs inutilisés peuvent être effacés de la liste des marqueurs.

Un marqueur peut être ajouté à la liste s'il n'existe pas déjà dans un enregistrement. Ceci permet la création d'un jeu de marqueurs standard (comme le jeu standard MDF que Shoebox reconnaît). Pour ajouter un nouveau marqueur:

Choisissez Base de données, Propriétés, onglet Marqueurs, Ajouter.

Propriétés des Marqueurs

Le fichier Type de base de données tient à jour des informations sur chaque marqueur. Comme information, il y a le code du marqueur lui-même (par ex. **\x**), le Nom du marqueur, l'Encodage de langue qui lui est affectée, sa place dans la Hiérarchie, ses Valeurs autorisées, s'il est de type Caractère ou Paragraphe, et une éventuelle information sur l'attribution d'une police autre que celle de l'encodage. Cet ensemble d'informations constitue les *Propriétés des Marqueurs*.

On peut accéder aux Propriétés d'un marqueur de deux manières:

- Par un **w**:Clic-droit ou **m**:Option+clic sur le marqueur dans un enregistrement, ou bien
- Choisissez Base de données, Propriétés, onglet Marqueur, sélectionnez le marqueur, et choisissez Modifier.

Nom de Champ

On peut donner un nom descriptif à un champ. C'est le *Nom de champ*.

Un Nom de Champ se définit dans la boîte de dialogue Propriétés des Marqueurs (Clic-droit sur le marqueur).

Pour que Shoebox affiche les Noms de champ dans la zone des marqueurs de la fenêtre d'enregistrement,

Choisissez Affichage, Noms de champ.

Pour ne voir que les marqueurs,

Choisissez Affichage, Noms de marqueurs.

⁵⁴ Certains programmes DOS restreignent la longueur des marqueurs à 4 caractères après l'antislash.

Pour voir les deux,

- Choisissez Affichage, Noms des marqueurs et des champs.

NOTE: Si un champ est copié à partir d'une autre application et que son marqueur n'existe pas déjà dans la liste des marqueurs de la base de données d'accueil, une alerte est donnée. Shoebox entre automatiquement ce nouveau marqueur dans la liste, avec une astérisque (*) comme Nom de champ, ce qui le rendra facile à retrouver en entrant dans la liste des marqueurs. N'oubliez pas d'aller dans la boîte de dialogue des Propriétés de ce marqueur pour corriger ce Nom de champ et vérifier par la même occasion que les autres propriétés (comme l'Encodage de langue, la Hiérarchie, etc.) sont correctes.

Encodage de langue d'un marqueur et Police

Un champ contient des données qui sont évidemment dans une certaine langue. Shoebox assigne automatiquement un Encodage de langue, généralement *Default*, aux nouveaux marqueurs de champ. Ceci se passe soit quand une base de données contenant des marqueurs inconnus est importée la première fois, ou bien chaque fois que vous ajouterez un nouveau marqueur à une base de données existante. Vous pouvez changer la langue attribuée automatiquement dans l'onglet *Marqueurs* de la boîte de dialogue *Propriétés du Type de Base de Données*, en remplaçant l'encodage spécifié dans la case *Encodage de langue par défaut (pour les nouveaux marqueurs)*.

L'Encodage de langue affecté à un marqueur peut être modifié; c'est d'ailleurs la meilleure façon de changer la police d'un champ, dans la mesure où un Encodage de langue renferme non seulement une information sur la Police, mais également l'Ordre de tri et le Clavier virtuel correspondant.

Pour changer l'Encodage de langue affecté à un champ donné, modifiez les *Propriétés des Marqueurs* pour ce marqueur (Clic-droit sur le marqueur). Il n'y a pas de façon simple de modifier les propriétés pour tout un groupe de marqueurs d'une base de données.⁵⁵ Néanmoins, les modifications apportées à un Encodage de langue (Police, Clavier virtuel, et Ordre de tri) se répercuteront sur tous les champs affectés de cet Encodage.

Pour spécifier une police différente ou un enrichissement spécial pour un champ particulier, dans la boîte de dialogue *Propriétés des Marqueurs* de ce champ, décocher la case *Utiliser police d'encodage*, puis cliquez sur le bouton *Choisir la Police*, et spécifier les police, taille, couleur, etc. Les modifications faites pour ce marqueur n'affecteront pas les autres champs.

NOTE: Sous Windows, l'utilisation de polices personnalisées pour des champs particuliers demandent plus de ressources système que l'utilisation des encodages. Dans certains cas, ceci peut être à l'origine d'un message d'erreur du système.

Jeu de Valeurs permises

Il est possible de restreindre le contenu de certains champs à un jeu de valeurs permises (par ex., les catégories grammaticales). Les valeurs de ce jeu seront listées dans la zone de texte *Valeurs permises* des Propriétés des Marqueurs de ce champ. Pour créer un jeu de valeurs permises pour un champ,

- Allez dans la boîte de dialogue *Propriétés des Marqueurs* de ce champ et cliquez sur le bouton *Valeurs permises*.
- Tapez les mots, abréviations, ou codes (ou caractères) dans la zone de texte *Valeurs permises du champ...*, séparés par des espaces.
- Si le jeu de valeurs permises est constitué de caractères ou de chiffres (plutôt que de mots), assurez-vous d'avoir sélectionné *Caractères* dans la boîte d'options *Valeurs permises de type...*
- Cliquez sur OK.

Pour créer un jeu de valeurs permises à partir du contenu d'un champ déjà utilisé,

⁵⁵ Pour les utilisateurs avertis : sortez de Shoebox, éditez le fichier de type de base de données, et recherchez/remplacez les occurrences de l'encodage de langue que vous voulez remplacer, par exemple, pour remplacer l'Encodage 'Default' par 'National', recherchez '**\lng Default**' et remplacez-le par '**\lng National**'. Sauvegardez le texte et redémarrez Shoebox.

- Cliquez sur le bouton Rassembler les valeurs permises dans la boîte de dialogue Valeurs permises.

Shoebox va lire le contenu de chacun de ces champs dans la base de données, et s'en sert pour établir une liste des valeurs permises. Ceci peut vous faire gagner beaucoup de temps et vous aider à mettre le doigt sur des incohérences dans vos données.

Une fois qu'une liste des valeurs permises est établie pour un champ donné, Shoebox ne vous permet pas d'entrer des données invalides dans ce champ. Si vous essayez, Shoebox vous enverra un message disant que le mot ou caractère du champ n'est pas autorisé. Il vous demandera ensuite si vous désirez l'ajouter à la liste. Répondre par Oui est un moyen pratique d'ajouter de nouveaux éléments à une liste—si cela est utilisé avec précaution. Répondre par Non vous permet de revenir en arrière et de corriger une mauvaise orthographe, etc.

Shoebox ne fait de contrôle que pour les champs de l'enregistrement saisi; il ne contrôle pas les anciens enregistrements—ceci vous permet d'afficher d'anciennes données sans être obligé de les corriger sur le moment. Par ailleurs, Shoebox ne vous signalera pas de mauvaises données avant que vous ne quittiez l'enregistrement en cours pour un autre enregistrement.

Hiérarchie La plupart des enregistrements ont une structure interne. Certains marqueurs sont "au-dessus" d'autres marqueurs dans le sens où leurs contenus ont un effet significatif sur les valeurs des marqueurs "en-dessous". Par exemple, tout dans un dictionnaire est "sous" le lexème, c.à.d. que vous ne trouverez pas d'information dans une entrée qui ne relève du lexème. De la même façon, les phrases illustratives sont *sous* un sens particulier—elles illustrent *ce* sens, et pas un autre. Dans Shoebox, cette relation peut être reflétée par l'utilisation du paramètre *Hiérarchie*.

La hiérarchie se paramètre dans la boîte de dialogue *Propriétés des Marqueurs*. Pour un exemple simple d'utilisation de la hiérarchie, voir Hiérarchie des marqueurs à la page 24.

Classes Une *Classe* dans Shoebox est une étiquette générique pour un groupe de caractères. Par exemple, la classe "vowel" comporte par défaut les lettres **A a E e I i O o U u**. Cette étiquette est très puissante lorsqu'elle est utilisée dans un Filtre, ou bien avec la commande Retrouver.⁵⁶

Pour utiliser une Classe dans un filtre ou une commande Retrouver, insérez-la entre crochets, c.à.d. *[vowel]*.

Dans la boîte de dialogue Edition, Retrouver, vous devez saisir la Classe (entre crochets) à la main.

Dans les boîtes de dialogue *Critère de filtre Marq./Séquence* et *Langue./Séquence* vous pouvez choisir une classe définie à partir d'une liste déroulante et l'insérer automatiquement entre crochets en cliquant sur le bouton *Insérer dans la Séquence*. Les Classes prédéfinies sont :

Classes prédéfinies	Description
[#]	<i>frontière de mot</i> —correspond à l'espace ou à un caractère de ponctuation de la Classe <i>[!]</i> de la langue du champ courant, ou au début ou la fin du champ.
[?]	<i>tout caractère</i> —correspond à n'importe quel caractère unique, dont l'espace. (NOTE: Si cette langue contient des unités de tri multigraphes comme <i>gb</i> , celles-ci comptent comme une seule

⁵⁶ La section *Filtres et Classes*, page 33, présente un exemple étape par étape d'utilisation des classes dans un filtre.

	unité quand à la correspondance.)
[...]	<i>sauter des caractères</i> (NOTE: Shoebox sautera autant de caractères que nécessaire pourvu que le reste de la séquence coïncide.)
	<i>un espace</i> dans une séquence de recherche correspond à un ou plusieurs espaces blancs (c.à.d. espace, tabulation, ou passage à la ligne). (NOTE: Les crochets ne sont pas nécessaire ici .)
[-]	<i>frontière de morphème</i> —même caractéristique qu'une frontière de mot, le caractère de découpage morphématique en plus (s'il s'agit du tiret—pour une obscure raison technique, cette Classe ne fonctionne qu'avec le tiret comme frontière de morphème).

NOTE: Les options *Correspondance des caractères* dans la commande Retrouver peuvent affecter significativement l'interprétation des Classes par la commande. La pratique vous l'enseignera mieux qu'un discours.

Les autres Classes sont spécifiques à un Encodage de langue. Ceci parce qu'une Classe contient des caractères qui doivent eux appartenir à un Encodage donné. Pour voir et modifier une classe définie par défaut par Shoebox pour chaque Encodage, ou pour en créer de nouvelle, allez dans la boîte de dialogue *Encodage de langues* (dans le menu Projet), sélectionnez l'Encodage voulu, et cliquez sur Modifier. Ceci vous conduit à la boîte de dialogue *Propriétés de l'Encodage de langue*. Cliquez sur l'onglet Classes pour voir, Modifier, Ajouter, ou Supprimer les classes disponibles.

Impression

Pour imprimer vos données de Shoebox, vous pouvez choisir l'une des deux approches suivantes :

- Vous pouvez obtenir une sortie rapide directement par la commande Imprimer du menu Fichier.
- Si vous voulez avoir plus de contrôle sur la sortie, vous pouvez Exporter vos données dans un fichier en *Format Standard* ou bien dans un fichier au *Format Rich Text* (RTF) et imprimer ce fichier par un traitement de texte comme Microsoft Word. Il s'agit là d'un processus en deux étapes nécessitant un fichier intermédiaire. Voir A propos de l'exportation de fichier, p. 74, pour en savoir plus sur cette méthode.

Impression directe

Les quatre boîtes de dialogue relatives à l'impression directe à partir de Shoebox sont toutes dans le menu Fichier. Il s'agit de Imprimer, Aperçu avant impression, Configuration imprimante et Configuration page.

La boîte de dialogue *Imprimer* permet de préciser le nombre de copies et le choix de l'impression soit de toutes les pages soit d'une sélection. (Il s'agit d'une boîte de dialogue standard, aussi la notion de “pages” ne correspond pas bien à la notion d'enregistrements de Shoebox, mais si vous avez des enregistrements courts et que vous demandez un saut de page après chaque enregistrement—dans Configuration page—alors les numéros de page et d'enregistrements correspondront assez bien.)

La boîte de dialogue *Aperçu avant impression* est l'aperçu d'impression standard qui affiche une page ou deux à la fois et qui est utile pour revoir les sauts de page.

La boîte de dialogue *Configuration imprimante* permet de choisir l'imprimante et d'en préciser certains paramètres qui lui sont spécifiques comme la taille du papier, l'orientation de la page, etc.

La boîte de dialogue *Configuration page* permet de spécifier les en-tête et pied de page, les marges et s'il faut un saut de page entre chaque enregistrement. Elle permet également de choisir entre l'impression de l'enregistrement en cours seulement ou bien de tous les

enregistrements actuellement affichables (si un filtre est actif, seuls ces enregistrements seront imprimés).

- Les marges peuvent être précisées en pouces ou en centimètres.
- Les en-tête et pied de page peuvent contenir du texte et/ou des codes spéciaux permettant l'insertion de données comme les noms du fichier, la date, le marqueur d'enregistrement, etc. Ces codes sont:

&F	nom du fichier (sans le chemin)
&P	numéro de la page courante
&n	numéro de l'enregistrement (NOTE: attention à la casse)
&N	nombre d'enregistrements de la sélection en cours d'impression
&R	marqueur d'enregistrement
&r	contenu du champ d'enregistrement
&K	marqueur de tri
&k	contenu du champ de tri
&D	date courante au format jj/mm/aaaa
&T	heure courante au format 24 heures, hh:mm:ss
&L	nom du filtre courant s'il y en a un
&&	imprime un "&"

Vous devez faire précéder tout code de donnée d'un & (ampersand). Par exemple :

Fichier &F Enregistrement &n/&N

donnera comme en-tête

Fichier SHLEX.DB Enregistrement 8/57

Comme son intitulé l'indique, la validation de la case à cocher *L'enregistrement en cours* provoquera l'impression de l'enregistrement actuellement affiché dans la fenêtre active de Shoebox. Même si vous avez plusieurs bases de données ouvertes dans plusieurs fenêtres, il s'agit ici d'imprimer seulement l'enregistrement de la fenêtre active.

La validation de la case à cocher *Toute la base de données* provoquera l'impression de tous les enregistrements affichables de la base de données courante. Un filtre peut être utilisé pour restreindre l'affichage courant à un ensemble d'enregistrements. Ceci limitera ainsi le nombre d'enregistrements imprimés.

NOTE: Si vous avez trié la base par rapport à un champ présentant plusieurs occurrences pour un même enregistrement, cet enregistrement sera imprimé pour chacune de ces occurrences. En d'autres termes, s'il y a 5 occurrences du champ de tri dans un même enregistrement, cet enregistrement sera présent 5 fois à l'impression.

La case à cocher *Saut (de) page entre enregistrements* obligera chaque enregistrement à commencer en haut d'une nouvelle page. Si cette case n'est pas cochée, les enregistrements seront imprimés les uns à la suite des autres, avec une simple ligne blanche entre eux.

A propos de l'exportation de fichier

La fonction d'exportation de fichier de Shoebox permet de formater et d'imprimer une base de données dans le format RTF (Rich Text format) à travers Word (ou tout autre logiciel comprenant le format RTF). Cependant Fichier/Exporter est bien plus qu'une simple fonction d'impression. Elle permet également d'exporter une base de données dans un fichier en pur Format Standard (SF de la SIL).

Que vous exportiez votre base de données en RTF ou en SF, vous pouvez le faire à travers une table de remplacements de manière à faire des remplacements systématiques dans le fichier de sortie. Vous pouvez également demander d'exporter toute la base de données ou

bien seulement l'enregistrement en cours. Vous pouvez sélectionner les champs à inclure ou à exclure dans le fichier exporté. Pour une exportation en RTF, vous pouvez demander que Shoebox attache automatiquement un *modèle* de document Word au fichier, et ouvre le fichier dans votre traitement de texte (sous Windows seulement). Vous pouvez également créer des configurations d'exportations spéciales et les sauvegarder pour vous en resservir plus tard. Pour accéder à toutes ces fonctionnalités, sélectionnez les commandes Exporter du menu Fichier.

Exportation simple de fichier

Il existe deux processus d'exportation simples de fichier base de données:

- Exportation au Format Standard (SF), ou
- Exportation au format RTF (RTF).

Elles sont présentées plus en détail plus loin.

Quelque soit le type choisi, le processus d'exportation simple sous Shoebox se fait en quatre temps:

1. Choisissez *Fichier/Exporter* dans le menu.
2. Sélectionnez le processus d'exportation que vous voulez (SF, RTF, ou un autre déjà défini) dans la boîte de dialogue *Exporter*, puis cliquez sur OK.
3. Shoebox affiche alors la boîte de dialogue appropriée au type de processus d'exportation choisi. Ici vous pouvez le configurer suivant vos besoins; Shoebox mémoriserà votre configuration pour la prochaine fois. Lorsque c'est fait, cliquez sur OK.
4. Dans la boîte de dialogue *Fichier exporté*, entrez le nom du fichier que Shoebox doit créer, puis cliquez sur OK.

Shoebox exécutera l'exportation comme demandée. S'il s'agit d'un fichier en Format Standard, Shoebox écrit le fichier sur le disque et revient à son écran principal. S'il s'agit d'un fichier en RTF et que vous avez demandé à ce que le traitement de texte soit automatiquement lancé, Shoebox appellera alors le traitement de texte (sous Windows seulement) et lui fera ouvrir le fichier RTF, sinon, Shoebox enregistrera simplement le fichier sur le disque et reviendra à son écran principal.

NOTE: Si vous n'avez pas (ou ne pouvez pas) sélectionné *Ouvrir le document dans le traitement de texte* lorsque vous faites une exportation en RTF, vous devrez lancer votre traitement de texte et ouvrir vous-même le fichier RTF pour pouvoir le voir et l'imprimer.

Enregistrements à inclure

Les options *Enregistrement à inclure* dans la boîte de dialogue *Exporter* vous permettent de déterminer exactement quelle partie de la base de données vous voulez exporter et quels enregistrements sont à sélectionner. Les trois choix qui suivent sont communs aux processus d'exportation SF et RTF :

- *Fichier entier* ne tient pas compte du tri et du filtrage en cours et exporte tous les enregistrements de la base de données dans l'ordre de tri de base du marqueur d'enregistrement.
- *Les enregistrements courants* exporte tous les enregistrements affichables dans la fenêtre sélectionnée, triés et filtrés comme ils s'affichent. Si le champ de tri primaire apparaît plusieurs fois dans un enregistrement, cet enregistrement sera exporté autant de fois.

NOTE : Si la fenêtre en cours est triée suivant un autre champ que celui d'enregistrement, juste avant de vous demander le nom du fichier exporté, vous verrez une boîte de dialogue *Exporter la fenêtre actuelle* qui vous offre deux choix :

- Laisser les champs de chaque enregistrement comme ils apparaissent dans la fenêtre en cours –ne pas réordonner
- Déplacer le champ de tri en cours en début de chaque enregistrement (ainsi le champ de tri de la base exportée sera en évidence). Tous les autres champs de tri (non-pertinents) d'un enregistrement donné ne seront pas exportés. De même les enregistrements ne comportant pas le champ de tri ne seront pas exportés.

Scripts d'exportations multiples

- L'enregistrement en cours n'exporte que l'enregistrement en cours.

La fonction *Fichier/Exporter* de Shoebox vous permet de définir des scripts d'exportations multiples, c.à.d. des scripts d'exportation pour des tâches spécifiques (par ex. un jeu de champs particuliers pour une publication de grammaire). Pour chaque script d'exportation vous pouvez spécifier les champs à inclure, la table de remplacements à utiliser (si nécessaire), et dans le cas d'une exportation en RTF la mise en page et le modèle de document Word à attacher.

Shoebox est livré avec deux scripts prédéfinis : les scripts SF et RTF déjà présentés. Vous pouvez les adapter à votre convenance ou bien redéfinir les vôtres à partir de zéro.

Pour modifier un script déjà défini,

- Sélectionnez le script et cliquez sur Modifier.
- Faites les modifications voulues puis cliquez sur OK.

Vos modifications seront enregistrées pour la prochaine exportation utilisant ce scripts.

Pour créer un nouveau script, vous pouvez soit en *Copier* un déjà défini ou bien en *Ajouter* un nouveau (pour lequel il vous sera donné le choix entre créer un scripts de type SF ou RTF).

Si vous êtes sûr de ne plus avoir besoin d'un de vos scripts d'exportation, *Supprimez-le*.⁵⁷

NOTE: Vous pouvez modifier les scripts SF et RTF d'origine si vous le voulez, mais si vous les supprimez, Shoebox les recréera dès qu'un fichier de ce type de base de données sera ouvert. Ceci afin de toujours pouvoir disposer d'au moins ces deux scripts d'exportation.

Ajouter, copier ou modifier un scripts.

Exportation au Format Standard

Que vous Ajoutiez, Copiez ou Modifiez un script, la boîte de dialogue dépendra du type d'exportation SF ou RTF du script d'exportation :

Le script d'exportation au Format Standard (SF) crée un fichier dont la structure est purement au Format Standard, sans formatage spécial ni champ d'en-tête Shoebox. Ceci afin de faciliter l'importation d'un tel fichier dans n'importe quel autre programme reconnaissant le format SF (ou encore, par exemple, l'ouverture dans Shoebox à travers une table de remplacements).

La boîte de dialogue *Format d'exportation SF* propose les options suivantes pour le script d'exportation SF :

Champs à exporter vous permet de préciser quels champs de la base de données doivent être inclus et lesquels doivent être exclus du fichier de sortie.

La boîte *Table de remplacements* vous permet de spécifier une table CC à appliquer sur l'exportation pour modifier les données "à la volée". Par exemple, si vous exportez un fichier devant être traité dans un programme sous DOS, vous pouvez vouloir rétablir les caractères ASCII par la table de remplacements ANSI-IBM.CCT, ou une table équivalente qui convertie le jeu de caractères Windows (ANSI) en jeu de caractères DOS (ASCII). (Cette table peut avoir à être modifiée pour correspondre à votre propre jeu de caractères.)

Exportation en RTF (Rich Text Format)

Le scripts d'exportation en RTF prend votre fichier de données au format standard et en appliquant les spécifications de paragraphe et style de chaque marqueur (définies dans la boîte *Propriétés des marqueurs*) il convertie la sortie en un fichier au format RTF qui est interprétable par Word et les autres logiciels compatibles RTF.

Pour mieux comprendre le script RTF, faites le parallèle entre votre fichier Shoebox et le fichier document qui en résulte : le fichier base de données est devenu un fichier formaté, le fichier Type de base de données fournit les informations nécessaires à la création du *modèle*

⁵⁷ Si la base de données est de Type MDF, vous aurez également le choix de définir un nouveau script d'exportation MDF.

de document, et les noms de marqueurs (définis dans la boîte *Propriétés des Marqueurs*) deviennent les noms de *style* pour chaque partie de texte du document.⁵⁸

La boîte de dialogue *Format d'exportation RTF* propose les options suivantes pour le script d'exportation :

Champs à exporter vous permet de préciser quels champs de la base de données doivent être inclus et lesquels doivent être exclus du fichier de sortie.

La boîte *Table de remplacements* vous permet de spécifier une table CC à appliquer avec l'exportation pour modifier les données "à la volée". Le processus CC est effectué avant la conversion en RTF, ce qui est intéressant si vous voulez combiner des champs, réordonner des champs, ou faire des remplacements systématiques avant que cela ne devienne un fichier RTF inextricable.

La boîte d'options *Format des styles défini par* vous permet d'utiliser les *Propriétés des Marqueurs* ou bien de spécifier un fichier modèle de document Word (DOT).

TRUC: Si vous configurez un script d'exportation RTF pour la première fois, exportez le fichier en utilisant *Propriétés des Marqueurs*. Ceci entraînera la création par Word d'un fichier dans lequel chacun des noms de marqueurs servira à créer un style dans votre document. En supprimant l'intégralité du texte du document (de façon à ce qu'il soit vierge), puis en l'enregistrant comme *modèle de document* vous aurez un fichier .DOT que vous pourrez utiliser dans l'option d'exportation précédente.

Le bouton *Config. page* vous permet de préciser la taille du papier, les marges et le nombre de colonnes du document exporté.

Autre information d'exportation RTF

Dans l'onglet *Marqueurs* du menu Base de Données, *Propriétés*, la colonne la plus à droite de la liste de marqueurs (avec l'en-tête SFR—pour Style-Font-Range) contient des lettres optionnelles qui donnent des informations sur l'attribut de chaque marqueur. Un marqueur avec :

- P ou C représente un style Paragraphe (à la ligne) ou Caractère (sur la même ligne)
- F précise qu'une police est choisie (c.à.d qu'il n'utilise pas la police de la l'encodage de langue)
- R précise qu'un jeu de valeurs permises est défini.

Ces paramètres peuvent être modifiés pour chaque marqueur dans la boîte de dialogue *Propriétés des Marqueurs*⁵⁹. Les choix de style Paragraphe ou Caractère se trouvent dans le groupe de boutons intitulé *Style à exporter*. 'Paragraph' est la valeur par défaut lors de l'entrée d'un nouveau marqueur.

L'enregistreur de Windows pour créer des macros

Les utilisateurs **Mac** peuvent sauter cette section.

A l'occasion, lorsque plusieurs commandes doivent être répétées, l'Enregistreur de Windows peut être utilisé pour créer des macros. Une exemple détaillé pas-à-pas est présentée plus bas pour créer une macro répondant à la touche de raccourci F3 pour se déplacer sur l'Enregistrement Suivant, et F2 sur le Précédent, modifier une macro, et copier un raccourci de l'Enregistreur dans le *menu démarrage* afin de le rendre disponible lorsque Shoebox est lancé.

NOTE: Une copie de l'Enregistreur à été placée dans le dossier Shoebox, parce qu'il n'est pas inclus dans Windows 95. (Les utilisateurs de Windows 3.1 le trouveront dans le groupe Accessoires.)

⁵⁸ Shoebox convertit également automatiquement les *sous-champs* en leurs styles de caractères propres. Les Sous-champs (*subfields*) sont rapidement abordés dans la réponse à une question au sujet de l'imbrication de polices dans un autre champ, p. 83.

⁵⁹ Sélectionnez un marqueur et faites Modifier dans la boîte de dialogue *Base de données, Propriétés*, onglet *Marqueurs* ou bien **w**:clic-droit ou **m**:option+clic sur le marqueur dans la zone des marqueurs.

- Créer une macro**
- Ajuster la taille et la position de la fenêtre Shoebox pour pouvoir afficher la fenêtre de l'Enregistreur.
 - Lancer le programme *Recorder*.
 - Cliquez sur Macro, un menu se déroulera.
 - Cliquez sur Record... une fenêtre apparaîtra avec des informations à fournir.
 - Commencez par taper *Suivant* comme Nom de macro.
 - Puis, allez à la case *Shortcut key* ; cliquez sur la petite flèche pointant vers le bas, sur la droite de la case, et descendez jusqu'à ce que vous trouviez *F3*.
 - Vérifiez que les cases Ctrl, Shift, et Alt ne sont *pas* cochées.
 - Puis, déplacez-vous dans la zone *Playback* qui est prédéfinie pour *Same Application*.
 - Cliquez sur la petite flèche à droite de la case et choisissez *Any Application*.
 - Descendez dans la case *Record Mouse*.
 - Cliquez sur la petite flèche à droite de la case et choisissez *Ignore mouse*.
 - Cliquez sur le bouton Start dans le coin en haut à droite. La macro est prête à s'enregistrer.
 - Cliquez dans la fenêtre *Shoebox* puis faites Alt+D pour aller dans le menu *Base de Données* puis N pour aller à l'enregistrement suivant.
 - Cliquez sur le bouton du Recorder dans la barre des menu pour arrêter l'enregistrement. (Les utilisateurs Windows 3.1 feront Ctrl+Break ou un double-clic sur l'icône du Recorder sur la bureau pour arrêter l'enregistrement de la macro.)
 - Cliquez sur Save Macro, puis sur OK dans la boîte de dialogue qui apparaîtra.
 - Faites réapparaître le Recorder s'il est caché.

Vous devriez voir la macro listée dans la fenêtre du *Recorder* avec la touche de raccourci d'abord et le nom de la macro à côté.

- Cliquez sur *File* dans la fenêtre *Recorder*.
- Puis cliquez sur *Save As*.
- Tapez *SHW.REC* comme nom de fichier dans la fenêtre qui apparaîtra.
- Revenez à Shoebox.
- Dorénavant, F3 sert à se déplacer à l'enregistrement suivant. Vérifiez-le. *Recorder* doit être en fonctionnement pour qu'une macro fonctionne, aussi :
- Réduisez *Recorder* et continuez à travailler dans Shoebox.

Ajouter une macro Pour ajouter une autre macro, restaurer *Recorder* de nouveau, s'il est réduit.

- Cliquez sur Macro dans le menu de *Recorder*
- Cliquez sur Record...
- Remplissez le Nom de macro avec *Précédent*.
- Choisissez le raccourci clavier *F2*, sans *Ctrl*, *Shift*, ni *Alt*.
- Sélectionnez *Any Application* pour *Playback*.
- Remplacez *Record Mouse* en *Ignore Mouse*.
- Puis cliquez sur Start. La macro est maintenant en enregistrement.
- Dans la fenêtre Shoebox, tapez Alt+P pour aller à l'enregistrement Précédent.
- Arrêtez l'enregistrement en cliquant sur le bouton *Recorder* dans la barre des tâches (Ctrl+Break ou double-clic, sous Win3.1).
- Cliquez sur Save Macro puis OK.

Modifier une macro

- Restaurer la fenêtre *Recorder*, et cliquez sur File, Save pour enregistrer le fichier avec les deux macros.
 - Revenez à Shoebox et vérifiez le bon fonctionnement de la nouvelle macro *F2*.
- L'utilisation de la touche *F3* dans Shoebox fait brièvement activer le menu *Base de Données* chaque fois que Shoebox avance, alors que *F2* déplace à l'enregistrement précédent sans cet effet (sur une machine rapide, cet effet peut passer inaperçu). Supposons que vous voudriez que *F3* fonctionne de la même manière que *F2*. *Recorder* ne vous permettra pas de créer deux macros avec le même raccourci et vous ne pouvez pas éditer la macro existante, aussi vous devrez :
- Restaurer *Recorder*.
 - Supprimer la ligne *F3* en la sélectionnant et en cliquant sur Macro, Delete.
 - Puis cliquez sur Macro, Record... puis redonnez toutes les informations sur la macro: Name: *Suivant*, Shortcut key: *F3*, pas de Ctrl, Shift ni Alt, sélectionnez *Any Application* et *Ignore Mouse* et cliquez sur Start. La macro est enregistrement.
 - Cette fois-ci, dans Shoebox, faites Alt+N au lieu d'utiliser le menu pour aller à l'enregistrement suivant
 - Arrêtez la macro comme précédemment.
 - Cliquez sur Save Macro puis OK.
 - Restaurez *Recorder* s'il n'est pas visible, et cliquez sur File, Save pour enregistrer les modifications sous le même nom de fichier. Maintenant Shoebox utilisera *F2* et *F3* pour aller à l'enregistrement précédant et suivant.

Paramétrer Recorder

Pour que ces macros soient disponibles chaque fois qu'on lance Shoebox, mettez une copie du *Recorder* dans le dossier (ou groupe) de démarrage.

Sous **Windows'95**:

- Localisez le fichier *Recorder.exe* dans l'explorateur, cliquez à droite dessus, puis choisissez Créer un raccourci.
- Maintenant, descendez et retrouvez le fichier tout juste créé (sans doute sous le nom "Raccourci vers *Recorder.exe*"), cliquez à droite dessus et choisissez Couper.
- Allez dans le dossier *C:\Windows\Menu démarrage\Programmes\Démarrage*. Cliquez à droite sur le dossier et choisissez Coller. (Ceci déplacera le raccourci que vous venez de créer dans le dossier Démarrage.)
- Remplacez le nom du raccourci par un nom plus parlant comme *Enregistreur de macros*.
- Cliquez à droite sur le raccourci et choisissez Propriétés.
- Cliquez sur l'onglet Raccourci.
- La case *Cible* devrait être en surbrillance. Utilisez la flèche de droite pour déplacer le curseur à la fin de "...*Recorder.exe*".
- Ajoutez un espace et tapez le nom du fichier de macros que vous souhaitez utiliser, c.à.d. *SHW.REC* (pour celles créées précédemment).
- Cliquez sur OK.

Maintenant, lorsque Windows démarrera, *Recorder* sera activé avec le fichier de macros spécifié, c.à.d. SHW.REC.

NOTE: configuré de cette façon, les macros seront opérationnelles dans *tous* (any) les programmes (avec des résultats imprévisibles), aussi si vous n'utilisez pas Shoebox, fermez *Recorder*. (Vous pouvez facilement relancer *Recorder* en utilisant la commande Démarrer, Programmes, Démarrage et en cliquant sur *Enregistreur de macros*).

Sous **Windows 3.1**:

- Maintenez appuyée la touche Ctrl et glissez l'icône du *Recorder* avec la souris dans le groupe de démarrage pour y créer une copie. Lorsque la copie est au bon endroit, relâchez la touche Ctrl et le bouton de la souris.
- Cliquez sur l'icône *Recorder* dans le groupe de Démarrage pour la mettre en surbrillance.
- Cliquez sur Fichier puis Propriétés dans le Gestionnaire de fichiers, avec l'icône *Recorder* en surbrillance .
- sur la ligne de commande, tapez un espace après *RECORDER.EXE* puis *SHW.REC*.
- Cliquez sur OK.

Maintenant, lorsque Windows démarrera, *Recorder* sera activé avec le fichier de macros spécifié, c.à.d. *SHW.REC*.

NOTE: configuré de cette façon, les macros seront opérationnelles dans *tous* (any) les programmes (avec des résultats imprévisibles), aussi si vous n'utilisez pas Shoebox, fermez *Recorder*. (Vous pouvez facilement relancer *Recorder* en double-cliquant sur son icône dans le groupe de *Démarrage*.)

Astuces Macro

Le *Recorder* n'a pas de fonction de répétition, aussi vous pouvez être amené à devoir répéter autant de fois que nécessaire le raccourci clavier. Dans le choix des raccourcis clavier, essayez de tirer profit de la fonction de répétition des touches en n'utilisant pas les touches Ctrl ou Alt. Par exemple, si vous définissez une macro sur Ctrl+X, vous ne pourrez pas la répéter en maintenant appuyée la touche Ctrl et en répétant X. Vous devrez relâcher la touche Ctrl, puis la rappuyer et taper X de nouveau pour chaque répétition de la macro.

Pour créer une macro compliquée, vous pouvez créer plusieurs macros plus petites, les tester individuellement pour voir si elles fonctionnent correctement. Puis vous créez une macro qui les appelle successivement. Très souvent ces macros simples pourront être légèrement modifiées ou utilisées dans différentes combinaisons pour faire des macros plus compliquées.

Encore une fois, faites attention de quitter le *Recorder* avant d'utiliser d'autres programmes que Shoebox, sinon *Recorder* risque de lancer vos macros intempestivement.

Enregistrer votre Presse-papiers

Le Presse-papiers de Windows est un outil intéressant, mais c'est un endroit très volatil pour conserver des données. Un mauvais geste et le contenu du Presse-papiers est perdu. Par contre vous pouvez enregistrer en permanence le contenu de votre Presse-papiers en utilisant l'utilitaire de Windows *Presse-livres* (*Presse-livres sous Win3.1*). Vous pouvez enregistrer autant d'information de type Presse-papiers que vous voulez avec cet utilitaire puis reprendre la partie qui vous intéresse et la recopier dans le Presse-papiers lorsque vous voudrez la coller dans Shoebox ou tout autre document. Pour vérifier cela, allez faites Démarrer, Programmes, Accessoires (normalement dans le groupe Principal du Gestionnaire de Programmes sous Win3.1).⁶⁰

⁶⁰ Le Presse-Livre peut ne pas être installé sur tous les systèmes Windows. S'il ne l'est pas, lancez la configuration de Windows et sélectionnez-le (dans les outils système de Win'95.)

Questions fréquentes

Démarrer avec Shoebox

Pour le **Mac**, l'application se réserve par défaut 4Mo. Vous devriez augmenter ce chiffre à 8Mo (peut-être plus) si vous constatez beaucoup de transfert avec le disque.

Q: Comment savoir si mes données déjà existantes tiendront?

Pour **Windows**, si vous avez moins de 8 Mo de RAM et beaucoup de données, cela risque d'être trop juste. Voici comment le déterminer:

- Evaluer la place occupée par vos données. Vous aurez probablement plusieurs fichiers chargés en même temps, faites-en la somme. Arrondissez au "K" le plus proche (laissez tomber les trois derniers digits, en arrondissant à la valeur supérieure).
- Multipliez ce total par 2.75..
- Ajoutez 500K à ce nombre pour prendre en compte la place occupée par le programme.

Pour connaître la taille de la mémoire disponible dans votre ordinateur :

- Sous Windows 95, ouvrez Mon Ordinateur; ou la Corbeille; ou Démarrer, Paramètres, Panneau de configuration, chacun d'eux présente une Aide dans le menu et une option A propos de Windows 95. La ligne *Mémoire physique disponible* indique exactement la mémoire dont dispose Windows. La mémoire disponible pour Shoebox dépend des autres applications ouvertes.
- Sous Windows 3.1, dans le *Gestionnaire de programmes*, ouvrez l'Aide, puis A propos du gestionnaire de programmes. Près du bouton de la boîte *A propos* se trouve un nombre intitulé *Mémoire*.

Si la taille de la mémoire disponible est inférieure au nombre que vous avez calculé précédemment, vous ne serez pas en mesure de lancer Shoebox pour Windows avec ces données; essayez de ne charger qu'une base de données, en découpant le fichier en morceau plus petit, ou bien ajoutez de la mémoire RAM.

Nous avons été amenés à modifier une base de données en y enlevant 3000 phrases illustratives. Cela a permis d'en réduire la taille à des proportions acceptables. Les phrases provenaient de textes, et nous avons mis dans le lexique, pour chaque entrée, les références aux phrases appropriées, si bien qu'aucune information n'a été perdue. Une certaine commodité a été perdue, mais les textes ont été chargés comme des bases de données *texte* (pas tous en même temps) et continuaient ainsi à être accessibles lorsque cela s'avérait nécessaire.

Si dans Shoebox 2.0 votre lexique principal était chargé en même temps que le lexique inverse trié suivant un autre champ, il n'est plus nécessaire désormais de dupliquer les données en différents fichiers. Dans Shoebox 3.0, vous pouvez trier suivant n'importe quel champ, et avoir une fenêtre affichant les données triées suivant un champ et une autre fenêtre présentant la **même** base de données triée (et/ou filtrée) suivant un champ différent. Vous pouvez également trier sur plusieurs champs.

Q: N'y a-t-il pas de moyen plus simple de démarrer un nouveau projet ou d'en importer un ancien ? Doit-on vraiment passer par les étapes de Type de Base de données et d'Encodages de Langue ?

A: Il existe bien une autre approche, mais les résultats ne vous permettront pas de tirer profit de toute la puissance que le logiciel met à votre disposition. Néanmoins, cette approche rapide et brouillon peut être progressivement améliorée au fur et à mesure de vos besoins pour exploiter les différentes fonctionnalités disponibles.

Voici les points principaux:

1. Vous ne pouvez pas passer outre le fait que les fichiers d'une même base de données doivent impérativement avoir le même marqueur d'enregistrement, aussi vous devrez forcément gérer différents Types. Cependant, vous n'êtes pas obligé de spécifier autre chose que le marqueur d'enregistrement. Vous n'êtes pas obligé de donner la liste des champs (le programme la constituera automatiquement), ni de donner le nom descriptif de chaque champ, ni d'établir une hiérarchie entre eux.
2. Laissez la langue affectée à chaque marqueur comme *Default*. Cela donnera la même police à toutes les données, le même ordre alphabétique, etc.

Si vous devez spécifier un ordre alphabétique et des relations minuscules/majuscules particuliers, vous pouvez modifier directement la langue par défaut. D'autre part, vous pourriez considérer que *c'est* une langue différente, et en tirer les avantages. Vous pourriez vous soucier d'une police également—il est plus efficace d'affecter une police à un encodage de langue (Encodage de langue, Options) plutôt qu'à chaque champ.

Q: Lorsqu'on démarre un nouveau type de base de données pour un fichier déjà existant, est-ce qu'on doit donner la liste de tous les marqueurs de champs utilisés? Que se passe-t-il si on en oublie ?

A: Non, vous n'avez rien d'autre à connaître que le nom du marqueur d'enregistrement. Le programme constituera automatiquement la liste des marqueurs trouvés en lisant le fichier.

Marqueurs

Q: Qu'est-ce que MDF ?

A: Le Multi-Dictionary Formatter (MDF) est un système assez mature pour développer, exporter et imprimer un dictionnaire multilingue et des glossaires inverses. La capacité de la version d'origine a été rehaussée et il a été intégré dans Shoebox, il est ainsi directement disponible pour tous les utilisateurs de Shoebox.

Mais bien que MDF soit un système très utile pour de nombreuses langues, il a ses limites et n'est pas la solution pour tous les langues ou tous les types de lexiques – ce n'est qu'une méthode. Reportez-vous à la documentation de MDF pour voir comment il marche et s'il peut vous être utile.

Q: Dois-je utiliser les marqueurs MDF ? J'ai déjà mon propre jeu.

A: Le jeu de champs MDF est complètement facultatif. Si vous avez déjà votre jeu de marqueurs lexicaux, vous pouvez l'utiliser.

MDF a été inclus parce que beaucoup l'utilisent et le trouvent utile. Pour plus de détails sur MDF, voir la documentation jointe à Shoebox.

Q: Je veux supprimer de la liste des marqueurs un marqueur qui n'est pas utilisé dans ma base de données, mais Shoebox me l'interdit.

A: Un marqueur qui ne peut être supprimé alors qu'il n'est pas utilisé, est probablement référencé dans une boîte de dialogue ou une autre.

Il peut également avoir été spécifié comme *Champ Additionnel* à afficher, dans une boîte de dialogue de correspondances multiples (Plusieurs Résultats). Si vous ne savez pas comment est configurée la correspondance multiple, essayez de rechercher (Base de Données, Rechercher) juste les 2 premières lettres d'un mot quelconque et vous obtiendrez très probablement des correspondances multiples.

NOTE: Faites attention de ne pas avoir coché la case "*Mot complet*" qui rechercherait exactement le mot tapé.

Ou bien il aura été spécifié comme champ de tri. Essayez *Base de Données*, *Trier* pour vérifier.

Il aura peut-être été spécifié dans la configuration de *l'Interalignement*.

La recherche de l'origine du problème peut devenir exaspérante. Peut-être que la manière la plus simple serait de quitter Shoebox et d'examiner le fichier de Type de cette base de données avec un traitement de texte. Le fichier .TYP contient toutes les références à ce marqueur. Faites une Recherche sur le nom du marqueur qui vous préoccupe (sans le caractère '\'). Le fichier .TYP contient toutes les références à ce marqueur.

Q: Comment repérer tous les enregistrements contenant un marqueur donné ?

A: *Retrouver* peut être utilisé pour cela. Faites *Retrouver* sur ce marqueur, sans spécifier de texte. ([...] sera automatiquement inséré comme texte à rechercher.) Pour cela spécifiez le marqueur dans la case *Rechercher dans la champ*.

Une autre possibilité consiste à faire un filtre *Marq./Séquence Filtre* (Base de données, Propriétés, onglet Filtres) sans aucun texte. Il filtrera tous les enregistrements contenant le marqueur spécifié.

Une autre solution serait de trier (*Base de Données, Trier*) par le champ puis d'afficher l'Index avec ce champ (*Affichage, Champs affichés dans l'Index*). Cela affichera toutes les occurrences de ce marqueur. Les enregistrements ne contenant pas ce marqueur seront étiquetés *pas de champ*. Ceux qui comporte ce champ, mais vide seront étiquetés *vide* (ou ce que vous aurez demandé à Index, Affichage d'afficher pour les champs vides ou inexistant).

Q: Le fichier *TEXTREP.CCT* utilisé pour préparer les fichiers texte produit des champs différents de ceux souhaités. Que faire ?

A: Il y a des commentaires dans le fichier *TEXTREP.CCT* qui permettent de savoir ce qu'il faut changer. Utilisez un traitement de texte pour éditer cette table. La prise en charge des fichiers texte devrait être automatisée dans une version future.

Polices Q: Les diacritiques volantes n'apparaissent pas dans la liste des caractères de l'Encodage de langue. Les diacritiques s'agglutinent quand on les entre dans l'ordre de tri secondaire.

A: Ce qui suit devrait aider les utilisateurs qui ont des problèmes à entrer des diacritiques volantes dans les boîtes d'édition de *Ordre de tri, Casse, et Classes*.

Tout d'abord, le remède le plus rapide semble concerner les caractères devant être précédés par des espaces. Pour certaines polices, un espace n'est pas suffisant, mais deux s'avèrent appropriés.

Un espace blanc est obligatoire, mais avec certaines polices, une diacritique en début de ligne disparaît visuellement, et des diacritiques séparées par un simple espace s'agglutinent en une masse informe. Aussi vous devriez taper au moins deux espaces devant toute diacritique superposable dans Propriétés de l'Encodage de langue, lorsqu'elle doit être utilisée ailleurs que dans un contexte de texte normal.

Q: Comment insérer des données qui sont dans une autre police—par exemple un mot banda dans un champ français ?

A: Il s'agit là d'un problème pour lequel nous n'avons pas encore de solution définitive. Il y a deux solutions possibles à court terme.

Ceux qui travaillent avec des alphabets non-romains préféreront sans doute une solution qui devra être modifiée par la suite, mais qui affichera les caractères de la langue. Cela implique d'interrompre le champ courant pour insérer un nouveau champ dédié à cette langue. Par exemple, un commentaire sur l'étymologie d'un mot fera sans doute référence à d'autres formes. Dans ce cas, le champ `\eth` (ou celui que vous utilisez) devra être interrompu par quelque chose comme `\vw` (pour vernacular word) puis continuer avec un autre `\eth` :

```
\eth xxxxxxxx
\v w yyy
\ecc xxxxxx xxxx ...
```

Les marqueurs `\vw` et `\ecc` doivent tous les deux être déclarés avec le style Caractère dans la boîte de dialogue Propriétés des Marqueurs, de façon à ce qu'à l'exportation par RTF vers Word, ils ne génèrent pas de passage à la ligne.

C'est une solution bancale, mais elle permet de voir les caractères voulus.

Une meilleure alternative, actuellement et pour un certain temps dans Shoebox, pour afficher à l'exportation des polices imbriquées, consiste à utiliser la méthode standard de

codage en-ligne provenant du logiciel Convert-to-Word, dorénavant incorporée dans Shoebox à cet effet. Ces codes en-ligne opèrent comme des styles de caractères en marquant ou entourant le texte imbriqué. Le format est : |fv{...} où 'fv' est un code de style, les ... représentent le texte marqué par le style, et les accolades servant à entourer le texte. (la barre '|' indique à Shoebox qu'il s'agit d'un style de caractère). Shoebox les appelle *sous-champs*.

En fait les sous-champs sont définis exactement comme un marqueur de champ normal (par la boîte de dialogue Propriétés des Marqueurs). Mais pour qu'un sous-champ opère, son option *Style à exporter* doit être définie comme *Style caractère*.

Pour illustrer cela supposons que vous ayez défini un sous-champ pour le 'style' *police vernaculaire* de marqueur **fv** et de nom de champ *Police-vernaculaire*. Pour inclure une phrase en vernaculaire dans un champ discussion utilisant l'encodage de langue Français, cela se présenterait ainsi :

\ue La phrase |fv{ana khwofw matamkwe} 'Je vais te donner une claque' est interprétée comme agressive. Ordinairement, cela se termine par un bain de sang.

En fait, Shoebox ne sait pas afficher à l'écran le contenu de ce champ dans les deux polices, mais il sait l'exporter par RTF vers un traitement de texte de telle sorte qu'à la phrase vernaculaire soit attaché le style *Police-vernaculaire* et qu'au reste des mots du champ soit attaché le style du champ **\ue**. Ceci produira quelque chose comme :

La phrase **ana khwofw matamkwe** 'Je vais te donner une claque' est interprétée comme agressive. Ordinairement, cela se termine par un bain de sang.

(Entre parenthèses, c'est la façon de faire de l'ensemble MDF pour convertir les nombreux champs d'un lexique au format modèle de styles de Word.)

Si vos différentes langues peuvent être transcrites à partir de la même police, alors il n'y a aucun problème pour voir ce que vous insérez entre crochets (comme dans l'exemple précédent). Mais si le champ conteneur utilise une police très différente de celle que vous voulez incorporer, vous aurez peut-être intérêt à aller dans un autre champ utilisant cette police, taper le texte, puis le couper et le coller là où vous en avez besoin. Les mots collés changeront de police et ne s'afficheront donc pas correctement à l'écran dans Shoebox, mais les codes de caractères seront justes et l'option d'exportation RTF les restituera correctement. Néanmoins, la solution précédente peut s'avérer préférable.

Fort heureusement, nous aurons une solution à ce problème dans la nouvelle version. Si cela vous concerne et que vous travaillez avec plusieurs alphabets, faites-le nous savoir et nous vous mettrons sur une liste des personnes devant recevoir une version intermédiaire dès qu'elle sera au point.

Gérer des bases de données

Q: Comment corriger le nom d'un Type de Base de Données ? Comment changer le nom d'un Type de Base de Données ?

A: Si vous avez déjà des bases de données affectées du type que vous voulez corriger, alors vous aurez à vous en préoccuper également. Elle ne seront pas modifiées automatiquement

- Premièrement, Fermer (Fichier) toutes les bases de données du type en question.
- Puis, Choisissez Propriétés, Type de Base de Données
- Dans la liste, sélectionnez le type de base de données que vous souhaitez changer.
- Choisissez Copier.
- Donnez le nom qui vous convient et faites OK
- Retournez à la liste de Types de Base de données en choisissant OK

Vous avez maintenant deux types de BdD qui sont identiques, au nom près

- Sélectionnez le type que vous ne voulez pas et cliquez sur Supprimer.

Questions
diverses

- Maintenant vous êtes prêts à recharger les bases de données. Faites-le. Pour chaque fichier que vous chargerez, vous aurez un avertissement disant que le type de base de données (l'ancien) n'a pas pu être trouvé. Demandez alors à Shoebox de lui attacher le type que vous venez de créer.

Q: Comment changer le Type de Bdd affecté à un fichier de Bdd ?

A: Il y a plusieurs approches possibles. La meilleure (la plus sûre et la moins compliquée) consiste à exporter la Bdd (Fichier, Exporter) dans un fichier au format SF. Puis d'ouvrir ce fichier (Fichier, Ouvrir). Le programme vous informera que ce fichier n'est pas au format Shoebox et vous demandera de lui affecter un Type de Bdd. C'est ici que vous choisirez le type de base de données à lui attacher.

Q: Est-ce que Sauter vers doit toujours ouvrir une nouvelle fenêtre ?

A: Non, et c'est à cela que sert *Fenêtre de saut*. Voir le fichier d'Aide à ce sujet. Voir aussi le didacticiel d'Interalignement de Texte pour trouver des exemples d'utilisation.

Q: Y a-t-il une option de sauvegarde automatique ?

A: Non, ni d'approche alternative qui me vienne à l'esprit. Tachez de sauvegarder tous les 5 nouveaux enregistrements, ou quelque chose comme ça. La commande *Fichier, Enregistrer Tout* sauvegarde toutes les bases de données modifiées ainsi que les fichiers de configurations. La commande *Enregistrer* (et son bouton) sauvegarde la base de données en cours. Nous rajouterons probablement cette fonctionnalité dans le futur.

Q: Comment créer une icône dans Windows pour accéder à un projet Shoebox ?

A: Le programme d'installation de Shoebox devrait avoir associé les fichiers d'extension .PRJ au logiciel Shoebox. En double-cliquant sur un fichier projet sous Mac, ou dans l'explorateur sous Windows 95 (gestionnaire de fichiers sous Windows 3.1) Shoebox devrait se lancer et charger le fichier de projet. Mais si vous êtes un utilisateur de Windows et que vous n'avez pas envie de partir à la recherche de vos fichiers projet, vous pouvez créer des raccourcis vers ces derniers. Voici comment :

Sous **Windows 95**,

- Avec l'explorateur, localisez votre fichier de projet.
- Avec le bouton droit de la souris, glisser-déposez le fichier sur un espace vide du bureau puis relâchez le bouton.
- Cliquez sur Créer un raccourci ici dans le menu qui s'est affiché.
- L'icône de raccourci apparaît alors sur le bureau.
- Cliquez à droite sur l'icône.
- Lorsque le menu apparaît, choisissez Propriétés.
- Choisissez l'onglet Raccourci.

Sur la ligne *Cible*, vérifiez que le chemin complet et le nom du fichier de projet sont bien là, c.à.d.:

"C:\MyShoebox Settings\MONPROJ.PRJ"

Modifiez le chemin dans la case *Démarrer en* en précisant le dossier contenant vos données. Il s'agit du dossier qui se présentera lorsque vous ferez Fichier, Ouvrir dans Shoebox.

- Faites OK pour enregistrer ces modifications.
- Cliquez sur le texte de l'étiquette de l'icône. Il se mettra en surbrillance, mais pas la case le contenant.
- Cliquez de nouveau de sorte que le curseur se positionne dans le texte et que le texte ne soit plus en surbrillance.
- Corrigez le texte comme vous l'entendez pour ce projet, puis faites <Entrée>.

En double-cliquant sur cette icône de raccourci, Shoebox devrait se lancer et charger le projet voulu. Vous pouvez déplacer ce raccourci où bon vous semble.

Sous **Windows 3.1**, tout ce que vous avez à faire, c'est de déplacer le fichier de projet du Gestionnaire de fichiers vers un groupe du Gestionnaire de programmes.

- Ouvrez le Groupe de programme Shoebox (ou le groupe dans lequel vous voulez installer l'icône du Projet).
- Ouvrez le Gestionnaire de fichiers et faites en sorte de voir également le Groupe de programmes dans lequel vous souhaitez mettre l'icône du Projet.
- Recherchez votre fichier de Projet (.PRJ) et tirez-le dans le Groupe de programmes.
- Cliquez sur l'icône du Projet et choisissez Fichier, Propriétés.

Une boîte de dialogue intitulée *Propriétés des programmes* apparaît

- Entrez une *Description* appropriée pour votre projet.

Sur la *ligne de commande*, vérifiez que le chemin et le nom du projet sont corrects. Par exemple:

C:\Myshset\MONPROJ.PRJ

Le *Répertoire de travail* doit spécifier le répertoire contenant les *données*. Il s'agit du répertoire qui se présentera lorsque vous ferez Fichier, Ouvrir dans Shoebox.

- Cliquez sur OK.

En double-cliquant sur cette icône de raccourci, Shoebox devrait se lancer et charger le projet voulu. Vous pouvez déplacer ce raccourci où bon vous semble. (Si cela ne marche pas, c'est que les fichiers d'extension .PRJ n'ont pas été associés au programme SHW.EXE. Vous pouvez corriger cela par le gestionnaire de fichier ou bien en réinstallant Shoebox.)

Notes sur l'Interalignement

Définir un découpage dans le champ de structure sous-jacente

Le champ *forme sous-jacente* (ou *profonde*) peut servir à expliciter le découpage des mots. Par exemple:

```
\lx reruns
\u re- run -s
```

NOTE: Pour qu'un morphème soit reconnu comme un *préfixe*, il doit présenter un espace après le tiret. De la même façon, un *suffixe* doit présenter un espace avant le tiret.

Les morceaux les plus longs ont priorité sur les plus courts

Le segmenteur a tendance à donner la priorité aux séquences les plus longues. Cela signifie que si vous importez une série d'*affixes conjoints* de SH2, les combinaisons d'affixes de votre liste auront la priorité sur des éléments plus courts. Par exemple, avec le mot *linguistically*, si *-ically* est dans le lexique, il aura la priorité sur les séquences *-ic*, *-al* et *-ly* du lexique. De cette façon, *-ically* pourrait avoir des découpages et des gloses forcées autres que les séquences plus courtes. De la même façon, si vous avez *linguistic* dans le lexique, il aura priorité sur la séquence *linguist* et *-ic* du lexique

Cependant, la segmentation se fait de l'extérieur vers l'intérieur, aussi s'il y a différents découpages possibles pour un mot, l'affixe le plus long risque de prendre la priorité sur la racine la plus longue. Par exemple, si le lexique contient les morphèmes cités précédemment, alors l'affixe le plus long, *-ically* se combinera avec la racine la plus courte *linguist* au lieu de la combinaison des affixes courts *-al* et *-ly* avec la racine la plus longue *linguistic*. Le résultat peut être le même, puisque les morphèmes issus de *-ically* peuvent être identiques à ceux produits par les éléments plus courts. Mais s'il y a une différence dans les découpages et les gloses forcées, il peut arriver que vous n'avez pas le résultat escompté. Vous pouvez contourner ce problème en ajoutant dans le lexique des mots entiers ou des combinaisons plus longues de racines et d'affixes, lorsque c'est nécessaire.

Un autre exemple de cette caractéristique est illustrée par le mot *does*, qui peut être découpé en *doe -s* et en *do -s*. Si *-es* existe dans le lexique comme variante de *-s*, alors *-es*, l'affixe le plus long, sera trouvé correct avec la racine *do*, et donnera le résultat *do -s* sans ambiguïté. Cela signifie que vous n'avez pas donné à Shoebox la possibilité de voir que la forme est ici *doe -s*.

Il existe plusieurs façons de contourner ce problème et celle que vous choisirez sera celle qui vous paraîtra la plus sensée suivant votre interprétation et votre manière de travailler. La première façon consiste à déclarer explicitement l'ambiguïté dans une de vos entrées. Par exemple, dans l'entrée *do* vous pourriez ajouter les lignes suivantes:

```
\lx do
\la does
\u do -s
\u doe -s
```

Ceci forcera Shoebox à présenter deux choix possibles de découpage pour *does*, et cela vous montre clairement l'ambiguïté (une forme variante provenant de deux formes sous-jacentes). Mais remarquez que cette méthode fait référence à *doe*, qui n'a aucun rapport avec le lexème *do*. Si cela vous semble aberrant, vous pouvez décomposer cela et mettre les découpages explicites dans chacun des lexèmes *do* et *doe*:

```
\lx do
\l a does
\l u do -s
```

```
\lx doe
\l a does
\l u doe -s
```

Ceci présente l'avantage que chaque découpage est placé dans le lexème dont il relève, mais avec le désavantage que l'ambiguïté n'est plus évidente à voir, c.à.d qu'il n'est pas évident en regardant l'entrée *doe*, de comprendre pourquoi le découpage *doe -s* à besoin d'être explicité ici—le danger étant qu'un jour ou l'autre, dans un moment d'inattention, vous risquez de vous dire “ça n'a rien à faire ici ” et de l'effacer.

Finalement il reste peut-être une meilleure solution, mais cela demande un peu de travail dans un premier temps. Vous créez une autre base de données (du même Type que votre base principale) et vous l'utilisez pour stocker toutes ces ambiguïtés de découpage de “mots entiers”. Voici comment ça marche : Vous créez une entrée dans cette base de données, soit:

```
\lx does
\l u do -s
\l u doe -s
```

Puis vous ajoutez cette base de données dans le chemin de recherche du processus Segmentation.. (Pour cela, allez dans Base de données, Propriétés, Onglet Interalignement. Sélectionnez le processus de Segmentation et cliquez sur Modifier puis sur le bouton Lexiques. Sélectionnez la base de données que vous venez de créer et cliquez sur Insérer. C'est fait, cliquez alors sur OK jusqu'à revenir à Shoebox.) Maintenant lorsque Shoebox interalignera un texte et trouvera *does*, il recherchera ce “lexème” dans votre nouvelle base de données (d'ambiguïtés) et vous proposera le choix entre les deux possibilités. Les avantages de cette méthode est qu'elle laisse votre lexique “vierge” de cette cuisine interne qui n'est nécessaire qu'au bon fonctionnement du segmenteur de Shoebox, et qu'elle crée un espace pour stocker tout ce qui tombe dans ce genre de dilemme. Seulement, lorsque vous sauvegarderez vos données, n'oubliez pas cette base de données d'ambiguïtés.

Réduire les ambiguïtés

Toute ambiguïté oblige le programme à stopper et à demander de faire un choix. Réduire les ambiguïtés diminue le nombre d'interventions manuelles nécessaires.

La première façon de réduire les ambiguïtés consiste à entrer dans le lexique des combinaisons de morphèmes qui soient moins ambiguës que les éléments individuels. Ceci est valable pour les constituants des mots comme pour les mots entiers. Les combinaisons d'affixes peuvent être utilisées pour réduire l'ambiguïté des affixes isolés. Les mots courants qui contiennent des affixes ambiguës pourront faire l'objet d'entrées lexicales à part entière.

Par exemple, si une langue présente des affixes du type: *a-*, *ab-*, *bc-*, et *c-*, des mots comme *abcroot*, peuvent être découpés des deux manières suivantes:

```
ab- c- root
a- bc- root
```

Mais si, dans cette langue, la combinaison *ab- c-* n'est pas signifiante, vous pouvez, pour réduire l'ambiguïté, créer l'entrée suivante:

```
\lx abc-
\l u a- bc-
```

De sorte que chaque fois que Shoebox sera en présence d'une combinaison d'affixes du type *abc-*, il saura que le seul découpage raisonnable est *a- bc-*.

De même que pour l'exemple précédent *does*, vous pourriez décider de conserver ce type d'entrée dans une base de données séparée (peut-être la même que celle des combinaisons mots-entiers).

Dans le futur, nous proposerons un moyen plus rationnel de prendre en compte ce type d'ambiguïtés. (ce qui est une autre bonne raison de conserver de telles entrées dans une base de donnée séparée.)

Valeurs forcées

Une valeur *forcée*, c'est lorsque le contenu ou 'valeur' de la ligne suivante (e.g. la ligne de glose) est explicitement spécifiée ailleurs dans le processus d'interalignement (au lieu d'être trouvée dans un processus normal de recherche). Lorsqu'une valeur forcée est donnée, le segmenteur ne s'interroge pas sur d'autres possibilités, il est "forcé" d'utiliser la valeur spécifiée. Néanmoins, une valeur forcée est contrôlée par le processus suivant qui s'assure qu'elle est valide. Si elle n'est pas trouvée dans le lexique, un message est affichée pour le signaler, et une marque d'échec est présentée à la place de la valeur forcée.

Des combinaisons de morphèmes peuvent utiliser des valeurs forcées pour réduire encore plus les ambiguïtés. Si la forme sous-jacente d'un morphème est ambiguë, alors une valeur forcée peut servir à choisir la valeur correcte pour ce morphème dans le mot en cours de traitement.

Par exemple, le suffixe *-s* en anglais est ambigu car il peut marquer le Pluriel et 3s (suffixe du verbe à de la troisième personne du singulier). Néanmoins dans la combinaison *-tion -s*, il s'agit forcément du Pluriel. Pour réduire le nombre de fois que vous aurez à lever l'ambiguïté sur *-s*, vous pourriez créer l'entrée suivante:

```
\x -tions
\y -tion -s{PL}
```

Ceci signifie que lorsque *-tions* est trouvé, il s'agit du suffixe *-tion* et du suffixe *-s*, lequel *-s* sera toujours glosé Pluriel. Le code *{PL}* est la *valeur forcée*. Les caractères par défaut pour spécifier une valeur forcée sont les accolades, mais elles peuvent être remplacées dans *Délimiteurs de gloses forcées* de la boîte de dialogue Segmentation.

(Dans des verbes comme *rations*, le *-s* marque la 3s, mais comme la racine est *ration* et pas *ra*, Shoebox ne découpera pas suivant le suffixe *-tion*, mais plutôt suivant *ration -s*, et l'ambiguïté du *-s* devra être levée. La solution de glose forcée *-tions* ne permettrait pas non plus de lever l'ambiguïté pour des mots comme *nations*.)

Plusieurs valeurs forcées peuvent être spécifiées, c.à.d par exemple qu'une valeur forcée pour la catégorie grammaticale peut suivre une glose forcée:

```
\x -tions
\y -tion -s{PL}{Nsu}
```

Dans cet exemple, *PL* est utilisé pour la ligne directement en dessous de la ligne à découper, et *Nsu* (suffixe nominal) pour la ligne suivante. Si la ligne de glose se trouvait en dessous de la ligne de catégorie grammaticale dans la configuration de l'interalignement, alors les deux valeurs forcées devraient être interverties.

La hiérarchie des marqueurs peut aider à réduire la nécessité des valeurs forcées multiples. Ceci parce qu'une fois déterminée la première valeur, cela réduit les choix possibles pour la seconde aux seuls choix de la section (de l'entrée) de même niveau hiérarchique. Aussi ne vous souciez des valeurs forcées multiples que si la première pourrait laisser la seconde ambiguë.

La hiérarchie des marqueurs est plus efficace si vous configurez le processus d'interalignement avec la ligne de glose *au-dessus* de la ligne de catégorie grammaticale.— comme le paramétrage express le fait. (Cela peut être l'inverse de ce que vous aurez vu dans d'autres publications.) La raison pour laquelle cela marche mieux est qu'une glose est plus restrictive qu'une catégorie grammaticale, c.à.d par exemple qu'il peut n'y avoir qu'un *PL* dans votre lexique, mais plusieurs formes *Nsu*, si bien qu'une fois que *PL* a été fourni, Shoebox sait qu'il s'agit d'un *Nsu*—ce qui ne serait pas forcément le cas dans l'autre sens.

Réduplication Le processus de reduplication est déterminé dans Shoebox par la combinaison d'une ou plusieurs *classes* (entre crochets⁶¹) dans le champ à explorer (champ **la** dans nos exemples) et la présence des lettres *DUP* (ou *Dup* ou *dup*—avec d'autres lettres si besoin est) dans la forme sous-jacente (champ **\lx** dans nos exemples).

Ce que représente réellement la *reduplication* ici peut être un sujet possible de confusion. Ne voyez pas la reduplication comme une simple schème de lettres répétées. Il s'agit en fait d'un morphème contenant du sens. Par exemple de nombreuses langues utilise la *reduplication totale* pour le pluriel des noms, comme dans *buku-buku* en indonésien, qui signifie 'livres'; la reduplication totale signifie dans cette langue "marque le pluriel".

La reduplication de schème est souvent utilisée pour nominaliser un verbe, par exemple en Selaru la racine du verbe *oban* 'frapper' donne le nominal *oboban* 'bâton', ainsi en Selaru ce type de reduplication de schème signifie "nominalise".

Comme la reduplication est porteuse d'une fonction et d'un sens, il s'agit d'un morphème à part entière et elle doit être intégrée dans le lexique (même si elle ne sera sans doute pas sortie dans le dictionnaire). Si vous le préférez, les formes de reduplication peuvent être maintenues dans un fichier séparé, mais il faudra alors faire savoir au processus d'interalignement où aller chercher ces informations.

Dans les exemples qui suivent, les formes sous-jacentes se trouvent dans le champ **\lx**. C'est ce qui viendra se placer sous la forme redupliquée lors de l'interalignement. Ce peut être ce que vous voulez pourvu que les lettres 'dup' apparaissent quelque part. Le schème recherché (contenant les classes) est lui, dans le champ **la**.

Réduplication totale ou générale Certaines langues redupliquent la *totalité* du mot. C'est ce qu'on appelle la *reduplication totale* ou *générale* qui est désigné par le schème de reduplication totale [...]:

\lx FullDup
la [...]

Le processus de reduplication totale dans Shoebox duplique le mot *précédent*. Par exemple, en Selaru, *wer wer* 'mouillé' (qui dérive de *wer* 'eau') serait analysé par Shoebox en :

wer FullDup

(Retenez que vous pouvez nommer ce processus n'importe comment; FullDup n'est qu'une simple étiquette, c'est [...] qui détermine le processus.)

La reduplication totale peut également être utilisée pour décrire ce qui se passe à *l'intérieur* du mot. Le *préfixe* de réplication totale de la seconde moitié de la racine ou du radical est indiqué par [...]-:

\lx Dup-
la [...]-

Ce processus de réplication aurait été utilisé si l'exemple précédent avait été écrit en un seul mot (ce qui est effectivement le cas en Selaru), comme *werwer* 'mouillé'. Shoebox l'aurait analysé en :

Dup- wer

Le *suffixe* de réplication totale de la première moitié de la racine ou du radical est indiquée par -[...]:

\lx -Dup
la -[...]

Ce schème aurait été utilisé pour *werwer* 'wet' si l'on avait déterminé que la reduplication est en fait un processus de suffixation. (Parfois vous ne pouvez pas vraiment savoir, dans ce cas choisissez un type jusqu'à ce qu'une étude plus approfondie fasse apparaître une évidence plutôt que l'autre.)

⁶¹ Voir *Filtres et Classes*, à la page 33 , pour plus de détails sur les Classes.

Réduplication de schème La *réduplication de schème* fait référence à un processus où un schème (ou une séquence) de consonnes et voyelles constitue la matrice de la forme rédupliquée, tandis que les lettres actualisées sont insérées dans la racine ou le radical par le processus de réduplication lui-même.

Pour définir le processus de réduplication de schème dans Shoebox, on utilise les classes des consonnes et voyelles ([cons] et [vowel] ou [C] et [V]) pour préciser la matrice du schème et le type de lettres qui sont dupliquées. Par défaut Shoebox crée les classes dénommées *cons* et *vowel*, mais dans les exemples ci-dessous nous utiliserons C et V, c'est plus facile à lire.⁶²

Par exemple, de nombreuses langues des Philippines présentent une réduplication du schème CV, donnant des mots comme *bibilih* (qui se découpe en *bi- bilih*) en Tagalog. Ce type de réduplication de schème se définit par :

```
\lx CVdup-
\la [C][V]-
```

Certaines langues présente une réduplication qui se combine avec une lettre fixe comme dans *babilih* (mot hypothétique construit sur le schème *Ca-* et *bilih*). La réduplication de la C suivante avec la voyelle *a* fixe se définit par :⁶³

```
\lx CaDup-
\la [C]a-
```

Shoebox peut aussi gérer la réduplication de la fin d'un mot ou d'une racine. Par exemple la réduplication du dernier schème CV d'un mot se définit par :

```
\lx -CVdup
\la -[C][V]
```

Résumé Seule la *réduplication totale* fonctionne au-delà de la frontière des mots. La réduplication de schème et les préfixes et suffixes de réduplication ne fonctionnent qu'à l'intérieur du mot et doivent être définis dans Shoebox avec un tiret (de préfixe ou de suffixe).

Ceci rend possible le traitement de *flipflip* (par une réduplication de schème ou un préfixe ou suffixe de réduplication), *flipflop* (réduplication de schème), et *flip flip* (réduplication totale). Par contre, Shoebox ne saura pas déchiffrer *flip flop*, car il y a là deux mots (dans cet exemple) ce qui rend inopérant le processus de réduplication de schème. Par contre la réduplication de schème fonctionne au-delà des tirets et autres marques de frontières de morphème, et Shoebox saura traiter *flip-flop* en *Dup-flop* (ou *flip-Dup* suivant la façon dont vous aurez défini le processus). Un affixe de réduplication totale peut également prendre en compte les frontières de morphème : découpage de *flip-flip* en *Dup-flip* (ou *flip-Dup* suivant la façon dont vous aurez défini le processus).

Shoebox recherche la réduplication à tout moment durant le processus de découpage, également avant d'autres découpages, aussi la partie rédupliquée peut sauter des frontières de morphèmes non marquées. Par exemple, le préfixe de réduplication totale ([...]- ci-dessus) peut amener aux découpages suivants :

<i>flipflip</i>	to	<i>Dup-flip</i>
<i>flipflips</i>	to	<i>Dup-flip-s</i>
<i>flipsflips</i>	to	<i>Dup-flip-s</i>
<i>flipflipping</i>	to	<i>Dup-flip-ing</i>
<i>flippingflipping</i>	to	<i>Dup-flip-ing</i>

⁶² Vous pouvez renommer [cons] et [vowel] si vous le désirez.

⁶³ Souvenez-vous que le champ \lx doit contenir quelque chose qui aide à éclairer sur le processus de réduplication en cours, et il doit contenir les lettres 'dup'—quelle que soit la casse.

Le schème de préfixe de reduplication totale [...] n'est trouvé que s'il y a correspondance exacte. Il peut découper *flipflips* parce qu'après le découpage du *s*, il ne voit plus que *flipflip*. Mais il ne découpera pas non plus *fliflip*, parce que le *p* n'est pas isolé comme suffixe. Le résultat du découpage de *flipsflips* est le même que celui de *flipflips* parce que la reduplication trouve d'abord *flips*, puis le *s* est découpé. L'exemple *flipflipping* est possible parce que le découpage isole le *-ing* et la morphophonologie enlève le *p* restant, puis la reduplication peut jouer. Le traitement de *flippingflipping* est identique à celui de *flipsflips*: la reduplication passe en premier, puis le processus de segmentation isole les suffixes.

Infixes

Un *infixe* est un morphème qui apparaît à l'intérieur d'une racine ou d'un radical—et non pas à ses frontières. Souvent certains parlent d'«infixe» pour un préfixe ou un suffixe apparaissant plus près de la racine que d'autres suffixes ou préfixes, mais pour Shoebox, il ne s'agit pas là d'infixes. Par exemple le mot indonésien *mempersilakan* 'inviter' a pour racine *sila*. Le reste se découperait suivant: *mem- per- sila -kan*. Remarquez comment les préfixes *mem-* et *per-* se décortiquent comme des pelures d'oignon. Comme *per-* n'est pas dans *sila*, il n'est pas considéré comme un infixe.

Les véritables infixes dans Shoebox sont repérés par la présence de tiret des deux côtés de la forme. Voici un exemple en tagalog:

```
\x -um-
```

Un infixe peut être trouvé n'importe où dans un mot. Il peuvent être retrouvés avant ou après que d'autres affixes aient été enlevés, mais le programme cherche à séparer les autres affixes avant de rechercher les infixes. Par exemple, si la racine est *bilih* 'acheter':

```
\t bumilih
\m -um- bilih

\t bumibilih
\m -um- Dup- bilih
```

Remarquez dans le second exemple comment l'infixe *-um-* s'est inséré dans la portion *bi-* du radical *bibilih*.

Dans le champ de segmentation, les infixes apparaissent soit au début, soit à la fin de la racine ou radical dans lequel ils auront été trouvés, suivant l'option que vous aurez choisie.

Racines composées

Les Racines composées sont prises en compte lorsque est choisie l'option *Autoriser les racines multiples* dans la boîte de dialogue Segmentation (par défaut, le découpage des racines multiples est activé). Un nombre quelconque de racines peuvent être séparées dans ce processus.

En anglais le mot composé *blackbird* peut être découpé en *black* et *bird*. Dans ce cas, les racines sont affichées, en dessous du mot composé, séparées par un tiret isolé (pour indiquer que ces mots proviennent d'un mot unique). Par exemple le texte interaligné pour *blackbird* donnerait:

```
\t blackbird
\m black - bird
```

Cette possibilité provoquera parfois l'apparition de mots qui ne sont pas dans le lexique. Par exemple, *afar* pourrait se retrouver décomposé en *a* et *far*. Mais ce découpage n'apparaîtra que si *afar* n'est pas dans le lexique. Une fois *afar* entré, étant plus long il prime sur les éléments plus courts.

Les tirets dans le lexique

Certaines personnes mettent des tirets sur des racines de leur lexique. Ils le font pour signifier que ces racines sont des formes réduites exigeant la présence d'un affixe. Par exemple dans l'échantillon de base de données Selaru du projet SHLEX, tous les verbes présentent des tirets à l'initiale. Comme exemple, *-abuk*, qui signifie *call*. Mais ce tiret fait croire à Shoebox qu'il s'agit d'un suffixe. Il ne découpera pas un mot constitué uniquement d'affixes.

Pour résoudre ce problème, ajoutez un champ `\a` à toutes les racines comportant un tiret, avec comme contenu la racine sans tiret. Ce n'est pas une solution idéale, mais ça marche. Vous pouvez trier la base par la catégorie grammaticale pour regrouper les entrées qui nécessitent ce champ supplémentaire, puis utiliser l'enregistreur de macro pour faire une copie du champ et retirer le tiret de la copie.

Le segmenteur reconnaît les préfixes, suffixes et infixes par la présence des tirets appropriés ou d'autres caractères de découpage de morphèmes. Aussi il est indispensable de mettre les tirets aux affixes pour que le processus de segmentation marche.

Si vous avez laissé quelques affixes sans tiret, vous pouvez trier votre base de données par la catégorie grammaticale pour regrouper les affixes. Puis ajoutez un tiret à chaque affixe. Si vous en avez beaucoup, utilisez l'enregistreur de macro pour ajouter le tiret et passer à l'enregistrement suivant.⁶⁴

Caractères séparateurs de morphèmes

Différents caractères séparateurs de morphèmes peuvent être utilisés suivant le types d'affixes. Par exemple, vous pouvez marquer les affixes dérivationnels par un signe égal et les affixes inflexionnels par un tiret. Ceci apporte une distinction nette dans la ligne de segmentation et les lignes en dessous car les caractères séparateurs de morphème se transmettent inchangés de la ligne de segmentation vers les lignes suivantes, la distinction continue ainsi à être marquée.⁶⁵

```
\t  relations
\m  relate =tion -s
\p  V      =nom -PL
```

Ce caractère distinctif peut se perdre dans la morphophonologie. Si un caractère séparateur de morphème n'est pas inclus dans la sortie du processus morphophonologique, ce sera le premier caractère séparateur de la liste qui sera utilisé.

Par exemple:

```
\a  =tion
\u  e+tion
```

Supposons que le *tiret* apparaît avant le *signe égal* dans la liste des caractères séparateurs de morphèmes. Après que cette entrée ajoute un *e* à la fin de la racine, il sort *-tion* au lieu de *=tion*. Pour corriger cela, ajoutez un signe égal à la sortie morphophonologique:

```
\u  e+=tion
```

Un caractère séparateur de morphème explicite est toujours admis dans la forme sous-jacente, mais s'il est présent, il doit apparaître après le séparateur de morphème '+' pour les suffixes et avant le '+' pour les préfixes. Faites attention de ne pas entrer quelque chose comme: *=e+ish*. Le processus de morphophonologie ajouterait un signe égal à la racine, et elle ne serait pas reconnue comme lexème.

Lignes de segmentation multiples

Il est possible de segmenter une seconde fois, par exemple pour isoler les radicaux.

```
\t  employees
\m  employee  -s
\p  n         PL
\m2 employ -ee -s
\p2 n        -suf PL
```

Pour cela, il vaut mieux créer une base de données spéciales pour les radicaux, comportant les gloses et catégories grammaticales suivant les besoins. Cette base de données

⁶⁴ Voir *L'enregistreur de Windows pour créer des macros*, p. 77, pour plus de détail sur l'utilisation de l'enregistreur de macro sous Windows.

⁶⁵ Ces caractères de frontières de morphèmes peuvent être définis dans la boîte de dialogue Segmentation (accessible par le menu *Base de données, Propriétés*, onglet *Interalignement*, puis en sélectionnant une ligne de processus segmentation et en cliquant sur Modifier).

grammaticales sera incluse dans le premier processus de segmentation, mais pas dans le second. (Ou bien vous pourriez essayer de mettre l'information sur les radicaux dans un champ "radical" de vos entrées lexicales principales et d'inclure le marqueur des radicaux dans la première segmentation mais pas dans la seconde.) La première segmentation ne verra que les radicaux et les laissera entier, et la deuxième segmentation ne les verra pas en temps que tel et les découpera.

Mots multiples

Sortie de mots multiples

Les gloses et autres champs insérés durant l'interalignement peuvent contenir plusieurs mots, par exemple en Selaru, la racine *betik* signifie 'aller dehors'. Ceci peut être inséré dans le lexique comme :

```
\lx betik
\ps v
\ge go outside
```

Shoebox lira la locution entière et l'utilisera pour gloser *betik*.⁶⁶ Il ne faudrait pas abuser de cette possibilité de glose en plusieurs mots car cela peut rendre l'interalignement du texte difficile à lire. Certaines conventions préconisent la présence d'un point de séparation comme "aller.dehors". Cette convention fonctionne aussi dans Shoebox.

Correspondance avec des mots multiples

Des mots multiples peuvent être retrouvés par le segmenteur et découpés en morphèmes. Ceci peut être utile pour gloser des locutions n'ayant pas de sens littéral, comme :

```
\lx right away
\ge now
```

Cependant les affixes ne peuvent pas être automatiquement isolés dans des locutions. Aussi la locution entière, avec ses affixes, doit figurée dans l'entrée lexicale ou la forme variante. Sa forme sous-jacente peut figurée découpée, avec ses affixes isolés.

Gloses multiples

Séparateur de gloses

Un caractère comme le point-virgule peut être utilisé comme séparateur de gloses. Le caractère utilisé pour cela est défini dans la boîte de dialogue "Recherche" (lookup)⁶⁷

```
\lx sey
\ps n
\ge house; hut
```

La présence du séparateur de gloses divise le champ glose en des possibilités multiples, qui seront présentés comme des ambiguïtés.

La boîte de dialogue *Recherche* présente une option pour *s'arrêter au séparateur*. Si cette option est validée, seule la partie avant le séparateur est utilisée. Si le séparateur de glose est en début de champ, celui-ci apparaît comme vide. Ceci peut être utile pour invalider certaines gloses lors de l'interalignement tout en les laissant accessibles pour faire un index ou comme commentaire.

Champs de gloses multiples

Si un processus de Recherche trouve plus d'un champ de sortie dans un enregistrement, il les traite comme des ambiguïtés. Cependant, il utilise la hiérarchie des marqueurs pour limiter la profondeur de recherche. Par exemple, si la catégorie grammaticale est au-dessus de la glose dans la hiérarchie des marqueurs, une fois la glose choisie, la catégorie grammaticale à laquelle appartient la glose est automatiquement déduite. Ou encore, dans

⁶⁶ Ceci est un comportement différent de celui de SH2 qui arrêta la lecture au premier espace.

⁶⁷ La boîte de dialogue *Recherche* est accessible par le menu *Base de données, Propriétés*, onglet *Interalignement* en sélectionnant une ligne de processus Glose et en cliquant sur Modifier.

l'interalignement, si la catégorie grammaticale est au-dessus de la ligne pour la glose, une fois la catégorie grammaticale choisie, seules les gloses en-dessous de cette catégorie grammaticale seront prises en considération.

Contrôle orthographique

Le contrôle orthographique lance le premier processus d'interalignement (généralement la segmentation) sur chaque champ ayant l'encodage de langue correspondant à celui du processus. Ceci permet de vérifier que chaque mot peut être formé de morphèmes appartenant au lexique.

Pour que le contrôle orthographique fonctionne, il faut qu'un paramétrage de l'interalignement ait été défini dans le type de base de donnée du fichier à contrôler. Dans le cas contraire, vous pouvez en définir un en lançant le paramétrage express comme indiqué dans le didacticiel d'interalignement. Vous pourrez le tester en lançant l'interalignement sur un échantillon de mots.

Pour lancer le Contrôle orthographique, choisissez Outils, Contrôle de l'orthographe (**w**:Alt+C ou **m**:commande+C). Si un mot échoue au découpage, il est sélectionné et le Contrôle orthographique s'arrête. A cet instant, il peut être ignoré, corrigé ou bien inséré dans le lexique. Pour ignorer le mot, déplacez le curseur en dehors du mot (au moins un espace plus loin) puis relancez le Contrôle orthographique. Pour corriger le mot, tapez le mot correct, ou bien déplacez le curseur dans le mot et corrigez-le. Pour entrer le mot dans le lexique, faites un Sauter-Insérer (**w**:Ctrl+Clic-droit ou **m**:contrôle+Option+clic) sur le mot. Pour n'entrer que la racine du mot dans le lexique, sélectionnez-la et faites un Sauter-Insérer (ou bien faites directement un Sauter-Insérer sur le mot entier, puis corrigez l'entrée dans le lexique, pour avoir seulement la racine).

Si aucun problème d'orthographe n'apparaît, un message s'affiche signalant que le Contrôle orthographique est terminé.

Dans un texte interaligné, seule la première ligne est contrôlée pour l'orthographe, les autres lignes sont ignorées.

Pour contrôler rapidement des enregistrements multiples d'une base de données, utilisez les touches de raccourcis pour passer à l'enregistrement suivant et relancer le Contrôle.

Editer

Vous pouvez corriger des mots dans un texte interaligné. L'alignement sera conservé. Ceci permet d'ajuster des découpages et des gloses à la main ou encore d'entrer des lignes d'annotation qui ne proviennent pas du lexique .

Toutefois, Shoebox se base pour l'alignement des nouvelles annotations, sur la ligne immédiatement en dessus, aussi il n'est pas possible de faire des choses comme ajouter une annotation du niveau mot—comme une glose de mot, ou une information du niveau de la structure d'un mot—sous la ligne de morphème, ni à la main, ni par le lexique. (Vous pouvez entrer ce type d'annotations en dessous de la ligne de texte avant qu'elle ne soit segmentée.)

Si vous souhaitez entrer un champ aligné à la main, vous devrez le faire par la ligne *De* ou *Vers* d'un processus d'interalignement. Cette version ne possède pas de processus d'entrée directe à la main, mais vous pouvez créer un champ aligné à la main en définissant un processus de Recherche avec comme marqueur *Vers* le champ d'entrée à la main, sans préciser de base de données de Recherche.

Coller dans d'autres applications

Bien que le texte interaligné soit affiché dans des polices proportionnelles, si vous copiez et collez du texte interaligné vers une autre application (par ex. Word) vous perdrez l'alignement si vous n'utilisez pas une police à espacement fixe (non-proportionnelle, comme Courier) dans cette application. Ceci parce que la structure profonde de l'interalignement est en fait une structure à espacement fixe. Un programme permettant de gérer correctement l'exportation en espacement proportionnel est en cours de développement.

SECTION 3

APPENDICES

Appendice A: Importer de Shoebox 2.0

Présentation Cette section décrit comment importer vos données et vos fichiers de configuration de Shoebox pour DOS v2.0 (SH2). (Pour les utilisateurs de Shoebox 1.2 -- mille excuses, les paramètres ne sont pas convertibles.) Si vous avez à importer des données qui ne sont pas au format SH2, ou simplement si vous ne voulez pas vous embêter à convertir vos fichiers de configuration, suivez les instructions (plutôt laconiques) de la section suivante intitulée *Comment faire sans les paramètres de SHA.SET ?*

Prenez le temps de revoir la section intitulée Trois concepts, sous Créer un nouveau projet p. 37. Elle précise les notions de Projet, d'Encodage de langue, et de Type de base de données. Vous en aurez besoin.

Cette section va vous apprendre à lancer l'utilitaire de conversion qui va extraire les informations de Type de base de données et d'Encodage de langue de votre fichier SHA.SET (fichier général de configuration de SH2).

Qu'est-ce que cela a à voir avec mon problème ? Dans SH2, tous ces paramètres étaient conservés ensemble, en grande partie dans le fichier SHA.SET. (Le fichier SHB.SET comportait quelques informations sur le projet et aussi sur l'affichage, comme la couleur.) En isolant ces paramètres dans différents fichiers, on favorise leur réutilisation. Ainsi, l'information d'encodage de la langue peut être partagée par des fichiers de différents types, dans différents projets. Plusieurs projets peuvent charger les mêmes BdD, et ainsi utiliser les mêmes informations de type de BdD (et les mêmes informations d'encodage).

Il n'y a que vous qui puissiez déterminer ce qui dans le fichier SHA.SET appartient à un encodage de langues et à un type de BdD donnés. Néanmoins nous vous proposons un outil pour vous faciliter la tâche.

Comment faire sans les paramètres de SHA.SET ? Si vous en concluez que les informations de votre fichier de configuration ne vous sont pas nécessaires, alors traitez les données comme des données non-Shoebox à importer. En bref : Si vous n'avez pas encore de type de BdD approprié, vous devrez en définir un (effort minimal) comme cela est décrit dans le document Créer un nouveau projet, p. 37. Démarrez avec la section Créer une nouvelle base de données à la page 41. Mais plutôt que de choisir de créer un nouveau fichier de base de données, importez vos fichiers de données par Fichier, Ouvrir. Suivez les instructions générales pour ajouter un nouveau Type de base de données et, bien sûr, vous aurez à donner un nom de type, un marqueur d'enregistrement, et une description qui correspondent à vos données.

Puis vous voudrez sans doute paramétrer les fichiers d'Encodage de langues, décrit dans la section immédiatement après celle sur les Types de base de données. (Vous pouvez travailler avec simplement *Default* comme Encodage, mais vous perdrez alors certaines fonctionnalités très intéressantes de ce programme.)

Importer un Projet complet —les étapes Nous allons utiliser la base de données et les fichiers de configuration du didacticiel *Approche orientée Texte* du chapitre 5 du manuel de SH2, en plus du didacticiel *Affixes conjoints* du chapitre 13. Le projet résultant comportait un lexique, une base de données de segmentation, et un fichier de texte interaligné. Le lexique et le texte comportaient tous les deux des données en anglais et en français; la base de données segmentation était intégralement en français (sauf les gloses forcées). La liste des affixes conjoints est également en français.

Les premières étapes sont exécutées par un autre programme que Shoebox 3. Si vous étiez sous Shoebox 3, veuillez en sortir. Vous allez créer de nouveaux encodages de langue et types de base de données, Shoebox ne fera que vérifier ces données au démarrage.

Utilisateurs de Macintosh, étape préliminaire

NOTE: Les utilisateurs de Macintosh devront convertir leur fichiers de données et SHA.SET en des fichiers Mac avant de lancer le processus de conversion. Toutes les tables de conversion et autres processus (comme la reconnaissance des noms de fichiers) suppose que la conversion en Mac a été faite.

Etape 1 : Le programme de conversion de Shoebox 2.

Le programme est assez simple à utiliser.

- Lancer le programme SH2 Converter (qui se trouve dans le dossier SHOEBOX\SH2CONV).

Vous verrez un message d'accueil expliquant l'objet du programme. (Ne tenez pas compte du fait qu'il dise qu'il va convertir votre configuration au format Shoebox 3.0; la version actuelle de Shoebox (v.4) lit sans problème les fichiers de configuration v.3)

Le programme comporte un fichier d'aide pas-à-pas très réussi.

- Cliquez sur Help

Vous verrez le guide pas-à-pas. Si vous voulez avoir un aperçu du programme avant de le lancer, cliquez sur la première étape et lisez. Si vous préférez une aide seulement à la demande, vous pouvez avoir les explications de chaque étape en choisissant le bouton *Help* dans la boîte de dialogue de cette étape. Vous constaterez qu'à chaque étape vous pouvez aller au sommaire (*Guide to...*) ou aux explications de l'étape précédente ou suivante.

De même, si à un moment donné vous avez l'impression d'avoir fait quelque chose de grave et que vous vouliez recommencer, choisissez *Cancel* (qui se trouve dans toutes les boîtes de dialogue), cela vous permettra de sortir sans avoir modifié quoi que ce soit.

- Si vous voulez que le guide pas-à-pas reste visible, cliquez dans la boîte *welcome*. Sinon fermez la boîte d'aide.
- Cliquez sur OK dans la boîte de bienvenue.

Définir le Projet

Une boîte de dialogue *Choose Directories and Project* s'affiche, demandant trois chemins, le dernier étant optionnel.

Le premier demande de localiser le fichier SHA.SET.

- Choisissez le bouton Browse à côté de la boîte *Path of sha.set file*.
- Naviguez jusqu'au répertoire SH2_SET. Vous y verrez les fichiers SHA.SET et SHB.SET dans la liste des fichiers.
- Cliquez sur OK pour accepter SHA.SET et retourner dans la boîte de dialogue *Choose directories and Project*.

Ensuite nous devons choisir un nom pour le projet et le chemin d'accès pour les fichiers de paramétrages que nous allons générer à partir de SHA.SET. Nous allons les mettre dans le répertoire *samp_set* dans lequel nous regroupons tous les fichiers de configuration.⁶⁸

- Choisissez le bouton **Browse** à côté de la boîte *Path and file name....*
- Naviguez dans le répertoire SAMP_SET qui se trouve dans le répertoire SAMPLE
- Entrez le nom de projet TXTFRINT.PRJ (Texte français Interaligné).
- Cliquez sur OK pour revenir à la boîte de dialogue *Choose directories and Project*.

Ensuite, les utilisateurs sous Windows devront spécifier la table de conversion qui convertira les caractères étendus ASCII du DOS en caractères ANSI de Windows. Cela est nécessaire pour convertir les caractères au niveau des séquences de tri et des correspondances majuscules/minuscules.

⁶⁸ Normalement, l'utilisateur courant n'aura besoin que d'un dossier pour les fichiers de configurations, de façon à ce que ses différents projets partagent les informations d'encodage de langue et de type de bases de données. Le consultant, par contre, voudra sans doute maintenir les données et configurations des uns et des autres séparées parce que les encodages de langue des différentes personnes pourraient porter le même nom en ayant des contenus différents.

- Choisissez le bouton Browse à côté de la boîte intitulée *Path for character conversion table*.
- Naviguez jusqu'à IBM_ANSI.CCT dans le répertoire MY SHOEBBOX SETTINGS (MYSHSET sous Windows 3.1).

Si vous n'avez pas de caractères spéciaux (accentués ou autres) à convertir, laissez la case blanche lorsque vous importez vos données. Si vous n'avez pas de table de conversion, contactez votre responsable informatique.

- Cliquez sur OK pour retourner à la boîte de dialogue *Choose directories and Project*.
- Cliquez sur OK.

L'information d'encodage de langue

Une nouvelle boîte de dialogue apparaît vous demandant d'entrer votre encodage de langue. Ce projet comporte de l'anglais et du français. Nous avons déjà un encodage pour l'anglais, mais pour illustrer l'utilisation de multiples encodages, nous définirons les deux.

- Tapez *Francais* puis la touche **w**:Entrée ou **m**:Retour (ou cliquez sur Add).
- Tapez *Anglais* puis **w**:Entrée ou **m**:Retour (ou cliquez sur Add).
- Cliquez sur OK.

La boîte de dialogue suivante s'intitule *Associate Punctuation*. Elle affiche une liste de ponctuation issues du fichier SHA.SET. Cette ponctuation n'est pas spécifiquement française, mais commune aux deux langues.

- Cliquez sur *Français* et sur *Anglais* pour les sélectionner.
- Cliquez sur OK.

Ensuite arrive le premier des ordres de tri. Il s'agit de l'*Ordre de tri par défaut*. Si on fait abstraction des marqueurs \srt que le fichier SHA.SET affiche, vous remarquerez que c'est la même ligne que l'ordre de tri défini dans *Globals, Sort Orders*.

- Cliquez sur *Anglais* pour le sélectionner.
- Cliquez sur OK.

Puis s'affiche un ordre de tri intitulé *Français*.⁶⁹

- Cliquez sur *Français* pour le sélectionner.
- Cliquez sur OK.

Ensuite apparaissent les associations majuscule/minuscule⁷⁰. Dans SH2, les associations maj/min n'ont pas de dénominations, mais comme celle-ci suit immédiatement l'ordre de tri français, il s'agit de l'association relative à cet ordre de tri dans SH2.

- Choisissez *Français* pour le sélectionner.
- Cliquez sur OK.

Après cela, viennent plusieurs Classes, qui s'affichent l'une après l'autre. Les classes sont des regroupements de caractères. Les classes étaient « globales » dans SH2 et servaient pour toutes les langues définies. La première classe regroupe les voyelles françaises. Comme cela englobe également les classes anglaises, nous les attribuerons également à l'anglais. (Mais par la suite, nous l'éditerons pour supprimer les voyelles superflues pour l'anglais).

- Cliquez sur *Français* pour le sélectionner, puis sur *Anglais* également.
- Cliquez sur OK.

La classe dénommée *cons* contient les consonnes, y compris ç et Ç.

⁶⁹ Sous Windows vous verrez sans doute quelques caractères supplémentaires bizarres. Il ne semblent pas corrects parce que la conversion des caractères ASCII n'a pas encore été faite.

⁷⁰ De même, sous Windows, vous ne pourrez pas reconnaître les caractères spéciaux affichés. (Il s'agit des C-cédille et des E-accentués.)

- Choisissez les Encodages de langues *Français* et *Anglais*.
- Cliquez sur OK.

La classe *nasal* ne contient que les voyelles communes aux deux langues

- Choisissez les deux encodages *Français* et *Anglais* pour la classe *nasal*.
- Cliquez sur OK.

Ici s'achèvent les informations qui iront s'inscrire dans les fichiers d'Encodages des langues.

L'information de Type de Base de Données

La boîte de dialogue suivante est relative au Type de Base de Données. Nous avons 3 types de Base de données dans ce projet : un lexique dont le marqueur d'enregistrement est `\le`, une BdD de segmentation dont le marqueur d'enregistrement est `\wd`, et un type texte interaligné dont le marqueur d'enregistrement est `\ref`. Le nom des types de base de données ne se limite pas forcément à un seul mot (le nom des Encodages de langues non plus).

- Entrez *Lexique français*, *BdD Segmentation*, et *Texte interaligné* pour les trois Types de base de données.
- Cliquez sur OK.

La boîte de dialogue suivante affiche le premier d'une série de Filtres⁷¹. Si vous regardez bien, vous pouvez repérer le premier Filtre spécifiant *N* comme "mot" isolé dans le champ **ps** (si vous faites abstraction des caractères bizarroïdes autour des dièses, le filtre est codé `[ps#N#]`).

Le Filtre isolant les noms, et le suivant, les adjectifs, appartiennent manifestement au Lexique puisqu'ils concernent la catégorie grammaticale.

- Choisissez *Lexique Français* puis OK pour chacun de ces deux filtres.

Le filtre suivant est relatif au champ KEY. C'est un peu moins évident puisqu'il existe aussi bien dans le lexique que dans la base de segmentation.

- Choisissez *Lexique français* et *BdD segmentation* pour le filtre *ab*.

Ensuite s'affiche un filtre SEARCH.

- Ne le mettez dans aucune des bases de données. Ce type de filtre est avantageusement remplacé par la capacité de la commande Retrouver à rechercher dans les enregistrements en cours.
- Cliquez sur OK *sans* avoir sélectionné de Type de base de données.

Puis vient un filtre FIELD.

- Choisissez *Lexique français*, puis OK.
- Les filtres DATE et NON-UNIQUE peuvent s'appliquer aussi bien au *Lexique Français* qu'à la base de données *Segmentation*.

Le filtre suivant est de type *Column* (comme vous pouvez le deviner par l'opérateur *With*). Il concerne le texte interaligné.

- Choisissez *Texte Interaligné* puis OK.

Puis vient un filtre se référant au marqueur **mb** et au début de certains mots. Il s'agit d'un filtre spécifique qui n'est pas important ici.

- Sautez-le en cliquant sur OK *sans* spécifier de base de données.

Ensuite vous verrez une boîte de dialogue intitulée *Associate Range Box*. En ignorant les marqueurs `\rng` sur la gauche, vous constaterez qu'il s'agit là de la spécification des *Valeurs*

⁷¹ Comme pour les caractères français, sous **Windows** les chevrons qui sont utilisés autour des *Classes* (jeu de caractères définissables par l'utilisateur) et les *métacaractères* (jeux de caractères prédéfinis) dans les filtres ne semblent pas corrects. Ils ressemblent à gauche à un *O* suscrit avec un *x* à l'intérieur et à droite à un macron (tiret horizontal en haut).

permises (Range set) sous SH2. A la suite du premier `\rng` se trouve le marqueur `\ps`. Il s'agit du champ auquel s'applique ce jeu de valeurs. Ce filtre concerne le lexique Français.

- Choisissez Lexique français puis OK.

Maintenant, le texte suivant apparaît à l'écran :

```
\bx1 Lexicon
\au1 Y
\tp1 \gl
\tp1 \ps
```

En regardant la ligne de titre, vous constaterez qu'il s'intitule Associate Template. Il s'agit du Modèle pour la boîte 1 (`\bx1`) qui s'intitule Lexique. Sautez la ligne `\au1` (information d'autochargement). Le Modèle est constitué des lignes commençant par `\tp1`. Le Modèle par lui-même est alors simple :

```
\gl
\ps
```

L'intitulé de la boîte étant mentionné, on devine quelle Base de données est concernée.

- Choisissez Lexique français.
- Cliquez sur OK.

La boîte de dialogue suivante contient des données similaires :

```
\bx2 Parse
\au2 Y
\tp2 \m
```

Vous n'aurez pas de difficulté à comprendre qu'il s'agit du Modèle pour la boîte 2, intitulée *Segmentation*. Le Modèle consiste lui en l'unique marqueur `\m`.

Vous constaterez autre chose à ce sujet. Le Lexique Français ne fait plus partie de la liste des Base de données! En effet, il ne peut y avoir qu'un modèle par Base de données et comme nous en avons déjà choisi un pour le Lexique, il ne s'affiche plus pour ne pas créer de confusion.⁷²

- Choisissez la base de données Segmentation.
- Cliquez sur OK.

Après une petite pause, un message vous informera que vous avez réussi la mise à jour de votre fichier SHA.SET. Félicitations !

- Cliquez sur OK pour mettre fin au programme.

Etape 2 : Importer les fichiers de données

Cette étape va charger tous les fichiers de données au moment de l'ouverture du Projet. Les fichiers spécifiés dans le projet Shoebox seront chargés automatiquement (indépendamment du fait qu'ils aient pu être autochargés ou non dans SH2).⁷³ Si ces fichiers ont été déplacés de leur répertoire d'origine, le logiciel demandera le nouveau chemin d'accès. De toute façon, le programme demandera le Type de Base de données de chacun des fichiers.

- Lancez Shoebox.
- Choisissez Projet, Ouvrir.
- allez dans le dossier SAMP_SET sous SAMPLES.
- Spécifiez le projet *FrIntTxt*.

⁷² C'est la même chose pour les associations minuscules/majuscules—un par encodage de langue—mais comme il n'y en avait qu'un dans cet exemple l'occasion d'en parler ne s'était pas présentée plus tôt.

⁷³ Si vous ne voulez pas qu'un fichier se charge automatiquement dans le nouveau Shoebox, fermez-le tout simplement. Seuls les fichiers encore ouverts au moment de quitter Shoebox seront automatiquement chargés à la prochaine ouverture.

Shoebox va commencer à charger automatiquement les fichiers. Si vous les avez déplacé du répertoire où ils étaient sous Shoebox 2.0, une boîte de dialogue intitulée *Ouvrir un fichier qui a été déplacé*. Remarquez que le nom du fichier recherché s'affiche dans la case Nom, mais pas son chemin d'accès. Le premier fichier recherché est Lexique.db.

- Naviguer dans le répertoire SAMP_SET sous le répertoire SAMPLES.
- Cliquez sur OK.

Vous êtes alors informé que ce fichier n'est pas au format SH3. Surprise ! Ceci signifie en fait qu'aucun Type de base de données n'a été affecté à ce fichier.

- Choisissez le Type Lexique Français.

Remarquez la case à cocher en bas à droite de la boîte de dialogue *Import*. Elle propose *Enregistrer l'original comme *.ori*. Cela vous donne la possibilité de conserver la version originale de votre fichier dans un fichier du même nom, avec l'extension *.ori*. Par défaut, cette sauvegarde est exécutée. Cela vous garantit une sécurité au cas où quelque chose se passerait mal (bien sûr *rien* de mal ne peut arriver...) Une bien meilleure approche (les deux sont d'ailleurs compatibles) consisterait à faire une sauvegarde de tous vos fichiers de données ailleurs, sur le disque ou sur disquette. C'est généralement une bonne habitude de le faire régulièrement, particulièrement dans les périodes de transition.

Nous allons convertir les caractères spéciaux du fichier Lexique avec la même table de changement que celle utilisée dans le programme de conversion.

- Cliquez sur Parcourir.
- Naviguez dans le dossier MY SHOEBBOX SETTINGS (MYSHSET sous Win 3.1) et sélectionnez le fichier IBM_ANSI.CCT.
- Cliquez sur OK pour revenir à la boîte de dialogue *Importer*
- Cliquez sur OK pour lancer l'importation.

On vous signalera que le marqueur d'enregistrement n'a pas été défini pour cette base de donnée.

- Entrez **le** comme marqueur d'enregistrement.
- Choisissez OK pour lancer l'importation.

Ensuite, c'est la Base de données segmentation qui va être lue. Le programme rencontrera les mêmes problèmes, mais comme vous lui avez déjà dit où trouver le fichier SH2LEX.DB, Shoebox va rechercher le fichier PARSING.DB dans le même dossier. Vous devrez cependant spécifier le Type de base de données. Profitez de ce que la petite fenêtre affiche une portion du fichier pour repérer le nom du marqueur d'enregistrement. Il s'agit de **\wd**.

- Choisissez le type de Base de données Segmentation.
- Remarquez que la Table de remplacements sélectionnée précédemment est encore sélectionnée. Shoebox se souvient qu'une CCT avait été spécifiée lors d'une précédente importation. C'est très pratique, mais lorsque vous n'avez plus besoin de CCT, veillez à décocher la case *Utiliser Table CCT*.
- Cliquez sur OK pour lancer l'importation.

On vous signalera que le marqueur d'enregistrement n'a pas été défini pour cette base de donnée.

- Entrez **wd** comme marqueur d'enregistrement.
- Choisissez OK pour lancer l'importation.

Ensuite, Shoebox va lire le fichier Texte.

- Choisissez le Type de base de données Texte interaligné.
- Entrez **ref** comme marqueur d'enregistrement.

- Choisissez OK pour lancer l'importation.

Vous voyez maintenant les 3 bases de données qui étaient chargées dans SH2. Elles s'affichent dans des fenêtres en cascade, dans l'ordre de chargement sous SH2.

Assigner des langues aux marqueurs

Une étape importante consiste à assigner la langue correcte à chaque marqueur, particulièrement à ceux qui seront utilisés pour le tri. Comme exemple, nous allons affecter une langue au marqueur d'enregistrement du lexique.

- Allez dans la fenêtre SH2LEX.DB.
- Choisissez Affichage, Index (ou cliquez sur le bouton Index).
- Allez à la fin du fichier et remarquez que le « à » est le dernier mot du fichier.
- Revenez à l'affichage normal en cliquant de nouveau sur le bouton Index.
- w:**Clic-droit ou **m:**Option+clic sur le marqueur **le** pour faire apparaître les *Propriétés des Marqueurs*.
- Remplacez la langue du marqueur par *Français*. Cliquez sur OK.
- Refaites Index et constatez que le « à » est maintenant en début de fichier comme il se doit.

Dans vos propres bases de données, le même procédé vous permettra d'attribuer une autre langue que *Default* aux marqueurs concernés. En plus d'établir la séquence de tri correcte, cela permet de définir la police adéquate pour ce champ, de différencier les langues par des couleurs et des types différents, et cela permet également à la commande *Retrouver* de reconnaître les champs d'une langue donnée.

Nettoyage

Vous vous souvenez sans doute que lorsque nous avons attribué les Classes lors du processus de conversion, nous avons validé des Classes contenant des caractères Français pour l'Encodage de langues Anglais. Il est temps maintenant de faire du nettoyage.

- Choisissez Projet, Encodage de langues.
- Sélectionnez Anglais.
- Choisissez Modifier ou double-cliquez sur Anglais.
- Choisissez l'onglet Classes.

Ceci affichera une liste des Classes ainsi qu'une partie de leur contenu. En parcourant la liste, vous verrez que la classe Cons contient les c-cédille et que la classe Vowel contient les voyelles françaises.

- Sélectionnez Cons.
- Choisissez Modifier ou double-cliquez sur Cons.
- Enlevez les C-cédilles.
- Choisissez OK pour revenir à la liste des Classes.

De la même façon, enlevez les voyelles françaises de la classe Vowel.

- Choisissez OK pour revenir à la liste des Encodages de langues.
- Choisissez Fermer pour fermer la boîte de dialogue.

Dans vos propres données, vous aurez peut-être à nettoyer les Classes et aussi les Ordres alphabétiques et Associations majuscules/minuscules et peut-être également quelques Filtres. (Les Filtres font partie du type de base de données.)

Etape 3: Interlignement

Troisième étape,

- Allez dans la fenêtre du texte interligné, SACRIFIC.IT.
- Remarquez que le texte n'est pas aligné. Ceci est dû au fait que le paramétrage de l'interlignement n'a pas encore été fait et les marqueurs ne sont par conséquent pas encore traités comme il se doit.

Paramétrage express

Pour configurer l'Interlignement, nous allons lancer la Configuration rapide.

- Assurez-vous que la fenêtre du texte interaligné soit bien active.
- Choisissez Base de données, Propriétés.
- Choisissez l'onglet Interlignement.
- Choisissez Configuration.

Vous verrez apparaître une boîte intitulée *Définition des marqueurs d'interlignement*. Vous pouvez sans doute encore voir les marqueurs du texte à interaligner sur la gauche de la fenêtre de dialogue.

- Tapez comme marqueurs **tx**, **mb**, **gl**, et **ps**.
- Choisissez OK.

Vous verrez apparaître la boîte de dialogue *Configuration rapide des marqueurs lexicaux*.

- Choisissez Insérer pour mettre le fichier SH2LEX.DB dans *Bases sélectionnées*.
- Tapez **gl** et **ps** comme marqueurs de *Glose* et de *Partie du discours*. Laissez les marqueurs de *Forme alternative* et de *Forme de base* tel quel.
- Choisissez OK.

Vous voyez maintenant l'onglet Interlignement présenter un processus *Segmentation* et deux processus *Glose*.

- Choisissez OK.

Vous devriez voir maintenant le texte interaligné pratiquement partout. Les exceptions sont des endroits comme la glose *father-PL* où il n'y a pas d'espace entre les morphèmes. Shoebox place toujours un espace entre les morphèmes et les utilise comme repère d'alignement. (Il utilise également l'espace pour reconnaître les préfixes et les suffixes dans le processus de *Glose*, ce que SH2 ne faisait pas.) Les endroits qui semblent ne pas être alignés peuvent être laissés tel quel dans le texte interaligné, ou bien ils peuvent être corrigés en interalignant de nouveau le mot. (Ne le faites pas encore.)

Utiliser un corps plus petit dans les champs en majuscules

Vous vous êtes sans doute aperçus également que certains mots de la ligne **ps** sont trop longs pour l'espace qui leur est réservé. Par exemple, *PRONPOSS* est trop long, et *PLSUF* est partiellement caché par le *N* qui le suit. Ceci parce que les lettres en majuscules sont plus larges que la largeur moyenne des lettres en minuscules. Pour corriger cela, nous pourrions augmenter le paramètre d'espacement de l'interlignement, mais il est plus judicieux ici d'utiliser un corps plus petit pour la catégorie grammaticale.

- w**: Cliquez à droite ou faites **m**:Option+clic sur le marqueur **\ps** pour en afficher ses propriétés.
- Décocher la case *Utiliser police d'encodage*.
- Cliquez sur Choisir la Police
- Sélectionnez la taille 10 point.
- Choisissez OK pour retourner aux *Propriétés des Marqueurs*.
- Choisissez OK.

La catégorie grammaticale est encore très lisible sans occuper pourtant trop d'espace.

Paramétrer les encodages de langue des marqueurs d'interlignement

Pour que l'interlignement soit correct, il faut paramétrer les Encodages de langue de chacune des lignes. Ceci parce que l'ordre de tri par défaut pour chaque Encodage est utilisé pour déterminer quels caractères font partie d'un mot et lesquels sont des caractères de ponctuation.

- w**: Cliquez à droite ou faites **m**:Option+clic sur le marqueur de champ **\tx** pour afficher les Propriétés des marqueurs.
- Remplacez son Encodage de langue par *Français*. Cliquez sur OK.
- Remplacez également l'Encodage de langue du marqueur de champ **\mb** par *Français*.

**Convertir les chevrons
dans la base de données
Segmentation**

L'Encodage de langue par *défaut* peut être laissé pour les deux autres champs de cet exemple.

Dans vos propres données, si vous utilisez des *gloses forcées* dans une base de données de *segmentation*, vous devrez les convertir avant d'importer le fichier. Shoebox utilise les accolades pour repérer les gloses forcées, aussi une table de remplacements *CC* dénommée *PARSPREP.CCT* est fournie pour convertir les chevrons en accolades. Toutefois si vous utilisiez déjà les accolades dans vos données, vous devrez choisir autre chose pour les gloses forcées (ou encore, si vous trouvez intéressante l'idée d'utiliser les accolades pour les gloses forcées, vous devrez changer votre façon d'utiliser les accolades dans vos données.) Si vous voulez utiliser autre chose que les accolades dans les gloses forcées, ouvrez le fichier *PARSPREP.CCT* dans le dossier *SHOEBOX\SH2CONV*, et remplacez les accolades par autre chose.

NOTE: la table *PARSPREP.CCT* doit être utilisée avant tout autre table de conversion de caractères, car elle suppose que les chevrons sont encore ceux du code ASCII. Aussi si vous comptez utiliser une autre table de conversion de caractères, vous devrez d'abord lancer le programme *CC* avec la table *PARSPREP.CCT* sur le fichier, ou encore importer le fichier dans Shoebox à travers la table *PARSPREP.CCT* puis faire l'autre conversion de la manière suivante: Exporter le fichier en format *SF* sous le nom *PARSING.SFM*, puis fermer le fichier *PARSING.DB*, copier *PARSING.SFM* sous le nom *PARSING.DB* et le rouvrir dans Shoebox, ce qui fera apparaître la boîte de dialogue d'importation. Vous donnerez alors le nom de votre table de conversion, (par ex. *IBM_ANSI.CCT*).⁷⁴

Cette étape n'est pas nécessaire pour l'exemple que nous traitons ici, car il ne contient pas de gloses forcées dans sa base de données *segmentation*.

**Transformer les affixes
conjointes de SH2 en une
base de données**

Pour utiliser les affixes conjoints *SH2*, nous devons les convertir en base de données. Ceci se fait au moment de l'importation par l'utilisation de la table de remplacements *SHAFF.CCT*. Cette table ajoute des marqueurs et convertit également les chevrons en accolades. Aussi si vous voulez remplacer les accolades par autre chose, éditez le fichier *SHAFF.CCT*, et remplacez les accolades par ce que vous voulez.

- Choisissez Fichier, Ouvrir et ouvrez le fichier *SH.AFF* dans le dossier *SAMPLES\SH2_TEXT*). La boîte de dialogue importation s'ouvre.
- Choisissez le Type de base de données *SHAff* (nous vous l'avons fournie).⁷⁵
- Cliquez sur Parcourir et recherchez la table de remplacements *SHAFF.CCT* dans le dossier *SHOEBOX\SH2CONV* puis faites Ouvrir pour le sélectionner.
- Choisissez OK pour lancer l'importation.

Cet échantillon d'affixes ne contient pas de caractères spéciaux français aussi nous n'aurons pas besoin de conversion de caractères. Si vos propres données nécessitent une conversion de caractères, vous pouvez la faire par le procédé d'exportation en format *SF*, puis d'importation à travers une table de remplacements, comme décrit plus haut.

**Segmentation à
la mode SH2**

Nous allons configurer la *segmentation* dans le style *SH2*, puis nous verrons comment l'adapter au nouveau style de *segmentation* de Shoebox pour voir les différences.

- Activez la fenêtre affichant *SACRIFIC.IT*.
- Choisissez Base de données, Propriétés.
- Choisissez l'onglet Interalignement.
- Choisissez Modifier pour éditer le processus de *segmentation*.
- Choisissez Style *segmentation SH2*.

⁷⁴ Si vous êtes familier du logiciel *CC*, vous pouvez faire cela en fermant le fichier *PARSING.DB*, et en exécutant *CC* sur lui avec un autre nom de fichier de sortie, puis en recopiant ce fichier de sortie sous le nom *PARSING.DB*.

⁷⁵ pour "SH Affixes".

Vous verrez apparaître deux nouveaux boutons *Affixes conjoints* et *Base de segmentation*. Commençons par configurer les affixes conjoints.

- Choisissez *Affixes conjoints*.
- Sélectionnez la base de données SH.AFF comme *Base à explorer*. (Commencez par enlever toutes les autres bases de données.)
- Sélectionnez **aff** comme *Champ à explorer*.
- Sélectionnez **u** comme *Champ à sortir*.
- Choisissez OK pour revenir à la boîte de Segmentation.

Ensuite, nous allons configurer la base de données de segmentation.

- Cliquez sur le bouton *Base de segmentation*.
- Sélectionnez la base de données SH2PARS.DB comme *Base à explorer*. (Commencez par enlever toutes les autres bases de données.)
- Sélectionnez **wd** comme *Champ à explorer*.
- Sélectionnez **m** comme *Champ à sortir*.
- Choisissez OK pour revenir à la boîte de dialogue Segmentation.

Maintenant nous allons définir le symbole de frontière morphophonologique pour qu'il corresponde à celui de SH2.

- Supprimer le signe plus qui est là, et tapez # (dièse) comme caractère délimiteur de frontière morphophonologique.

Puis nous allons remplacer les caractères *Délimiteurs de gloses forcées*. Ceci ne sera pas nécessaire si vous utilisez les accolades comme délimiteurs de gloses forcées. Comme les données que nous allons importées contiennent des parenthèses dans les gloses, nous allons demander des accolades comme délimiteurs de gloses forcées.

- S'il y a des parenthèses comme *délimiteurs de gloses forcées*, supprimez-les et tapez des accolades { et } à la place.
- Choisissez OK pour revenir à l'onglet Interalignement.
- Choisissez OK.

Ajouter des tirets aux affixes dans le Lexique

Dans Shoebox, tous les affixes du lexique doivent comporter des tirets. Si ce n'est pas le cas, il ne seront pas reconnus comme étant des affixes. Le moyen le plus simple pour repérer les affixes dans le lexique, c'est par la catégorie grammaticale. Si nous trions la base de données par la catégorie grammaticale, celles-ci seront regroupées.

- Activez la fenêtre SH2LEX.DB.
- Choisissez Fenêtre, Dupliquer.
- Choisissez Base de Données, Trier.
- Sélectionnez le champ **ps** comme premier champ de tri. Laissez **le** en second.
- Validez la case *Trier premier champ à partir de la fin*.
- Choisissez OK.
- Choisissez Affichage, Champs affichés dans l'Index.
- Sélectionnez les champs **ps**, **le**, et **gl** dans l'ordre.
- Choisissez OK.

Ceci produit un affichage dans lequel il est facile de voir les catégories grammaticales et nous pouvons facilement repérer celles qui semblent être des affixes. Tous les affixes de cet échantillon sont des suffixes et contiennent *SUF* comme catégorie grammaticale.

- Sélectionnez le premier enregistrement *SUF* et sortez du mode d'affichage par Index.
- Ajouter un tiret devant le suffixe dans le champ **le**.

- Repassez en mode affichage par Index.
- Faites ceci pour tous les enregistrements présentant *SUF* comme catégorie grammaticale.
- Une fois terminé, fermez la fenêtre du second lexique dans laquelle vous étiez en train de travailler..

**Paramétrer la
Destination de saut**

Pour de meilleurs résultats, nous devons paramétrer la Destination de saut du texte interaligné pour qu'elle pointe vers les lexiques. Ceci permettra de *Sauter* rapidement vers les mots des lexiques et d'en insérer facilement de nouveaux.

- Revenez à la fenêtre du texte interaligné, SACRIFIC.IT.
- Choisissez Base de données, Propriétés.
- Choisissez l'onglet Destination de saut.
- Définissez la Destination de saut suivante : SH2LEX.DB, champ **le**, SH2PARS.DB, champ **wd** field, et SH.AFF, champ **aff**.
- Choisissez OK.
- Pour chacune des bases de données destination de saut, allez dans sa fenêtre et validez Affichage, Fenêtre de saut.

Pour rendre le *Saut vers* plus efficace, nous allons l'obliger à tenir compte des tirets de la même façon que l'interalignement le fait.

- Retournez à la fenêtre du texte interaligné, et assurez-vous que le curseur est bien en haut de l'enregistrement..
- Choisissez Edition, Sauter vers.
- Sous Correspondance des caractères, choisissez Même ceux qui sont ordin. ignorés.
- Choisissez OK. Si le curseur était bien en haut de l'enregistrement, Shoebox aura sauté sur le mot *sacrifice* dans SH2LEX.DB. Si le saut n'aboutit nulle part faites simplement Annuler.

Vérifier les annotations

Dans vos propres données, il n'y a pas de nécessité d'interaligner de nouveau les mots qui le sont correctement. Vous pouvez travailler sur un nouveau texte ou bien corriger les mots qui ne sont pas correctement découpés. Shoebox ne sait pas encore contrôler les annotations déjà établies. Il ne peut que les remplacer (à la demande), et si un mot est ambigu, il redemande de lever l'ambiguïté. Vous pouvez faire des vérifications par la commande Outils, Contrôle de l'orthographe. Il vérifie que chaque mot de la ligne du haut peut être segmentée. Mais il ne vérifie pas les lignes d'interalignement suivantes.

Néanmoins, pour vérifier que ce texte s'interaligne comme sous SH2, nous allons interaligner chaque mot du texte.

- Retournez dans la fenêtre du texte interaligné, et assurez-vous que le curseur est bien en début d'enregistrement.
- Interalignez (cliquez sur le bouton).

Vous constaterez que les annotations sous le premier mot n'ont pas changées. Le curseur s'arrête au second mot (*nos*).

- Interalignez de nouveau.

Vous verrez apparaître un espace après *PRONPOSS*.

- Interalignez de nouveau.

Vous verrez un espace s'ajouter devant *-PL*, qui sera correctement aligné.

- Continuer à interaligner jusqu'au mot *aux* et constatez que tous les mots présentent les mêmes annotations qu'auparavant.
- Interalignez le mot *aux* et constatez que le morphème *-le* n'a pas de glose.

Ceci illustre une différence importante entre SH2 and SH3.

- Pour voir ce qui cause cette défaillance, Sauter vers *le* (pas *-le*).

Observez que c'est une racine dans le lexique, et qu'elle ne comporte pas de tiret. SH2 ignorait les tirets lorsqu'il recherchait des affixes dans le lexique, alors que Shoebox maintenant tient compte des tirets. Ceci permet à Shoebox de relever moins d'ambiguïtés.

Le mot *aux* doit ici être découpé en deux racines. (SH2 ne le permettait pas mais Shoebox oui.) Pour corriger cela nous allons définir le découpage de *aux* en deux racines au lieu d'une racine et d'un suffixe.

- Sauter vers *aux*.
- Remplacer le tiret dans la ligne `\m` par un *espace*.
- Revenez à la fenêtre du texte interaligné et relancez l'interalignement sur *aux*.

Observez que le découpage est maintenant correct.

- Interalignez le reste des mots.

Ils devraient tous présenter les mêmes annotations qu'auparavant.

Le texte interaligné est sûrement trop large pour la fenêtre dans laquelle il s'affiche, aussi nous allons le restructurer.

- Placez le curseur dans le premier champ `\tx` en début d'enregistrement.
- Choisissez Outils, Restructurer.

Vous verrez le texte en entier se restructurer pour tenir dans la fenêtre.

Le découpage selon Shoebox

Nous en avons terminé avec la démonstration de la configuration de l'interalignement de Shoebox en style SH2. Pour vos propres données, vous pouvez vouloir transformer cette configuration pour utiliser le style natif de segmentation de Shoebox, qui met tout dans le lexique au lieu d'utiliser des bases de données séparées pour les affixes et les découpages.

Nous n'allons pas perdre de temps à le faire avec cet échantillon de projet FRINTXT.PRJ, mais pour vos propres données, vous aurez à suivre les étapes suivantes :

- Désactivez le style de segmentation SH2 afin de permettre la recherche d'affixes multiples.
- Supprimer tous les affixes conjoints qui ne présentent pas de modifications morphophonologiques.
- Déplacez chaque affixe présentant des modifications morphophonologiques dans l'entrée lexicale de sa forme de base (sous-jacente).
- Supprimer de la base de données *segmentation* toutes les entrées qui peuvent être découpées par le lexique.
- Déplacez toutes les autres dans le lexique, sous la racine appropriée, ou comme mot isolé si nécessaire.

Importer vos propres données

Sur combien de "Projets" différents (paramétrés en SH2) travaillez-vous ?

- Faites-en une liste avec leur localisation.

Ils sont sans doute dans différents répertoires à partir desquels vous lanciez le programme SH2. Pour chacune des configurations SH2 que vous avez déjà, vous souhaiterez sans doute créer un nouveau Projet.

- Faites une liste des langues impliquées dans chaque Projet à convertir.

Il y aura sans doute la langue vernaculaire sur laquelle vous travaillez, peut-être une langue nationale et sûrement une autre langue comme le français, l'anglais ou l'espagnol.

Sans doute plusieurs projets partageront certaines de ces langues. Plusieurs d'entre eux partageront peut-être le même lexique ou une autre base de données.

- Examinez chacune de vos configurations SH2.⁷⁶

Traitez-vous une même langue de la même façon dans chacun de vos projets ? Avez-vous des ordres alphabétiques différents ou des classes différentes pour des usages particuliers ?

- Choisissez le “meilleur” traitement pour chaque langue et vérifiez dans quel projet il apparaît.

Si possible, vous n'aurez qu'un seul traitement par langue. Cependant, si deux traitements différents ont des avantages spécifiques, notez-les. Ils pourront être fusionnés après importation.

Quelles langues sont couvertes par l'ordre de tri par défaut ? Y-a-t-il un ordre de tri qui recouvre plusieurs langues ?

Vous devez vous poser la question de l'intérêt de regrouper deux langues qui utiliseraient le même ordre de tri, avec quelques petites modifications n'interférant pas. Il y a des arguments pour cela. Cependant, certaines fonctions de Shoebox font une distinction basée sur la langue. Cette distinction sera-t-elle intéressante pour vous, ou bien ce serait une entrave. L'exemple de recherche dans le parcours des *Fonctionnalités de base* utilisait la distinction sur la langue pour retrouver le mot *a* en Selaru seulement et pas en Anglais. Une police est également affectée à un encodage de langue, bien qu'il soit possible de définir des polices différentes pour des champs particuliers. Généralement dans un dictionnaire bilingue une langue est en gras, et l'autre en normal. La police d'un encodage est pratiquée pour faire cette distinction.

Quelque soit le choix que vous ferez, vous pouvez toujours changer l'encodage de langue d'un champ, vous n'êtes donc pas bloqués par un choix définitif. Cependant une décision sage aujourd'hui sera du temps gagné plus tard.

- Jetez un œil à vos Classes.

Lesquelles sont corrélées à tel encodage de langue ? Lesquelles sont explicitement relatives à une langue donnée ? Lesquelles sont partagées ? Lesquelles portent le même nom tout en ayant des contenus différents pour différents projets ?

- Dressez une liste des différentes bases de données de chaque projet.

Certaines d'entre elles sont partagées ? Lesquelles sont vraiment des données d'un même type ? Par exemple, vous pouvez conserver les noms propres dans une base de données séparée du lexique principal—mais elle n'en constitue pas moins un fichier lexical. D'autres découpages peuvent amener à conserver les affixes, les mots d'emprunt, ou un vocabulaire spécialisé dans des fichiers séparés. Ou encore il est possible qu'en se développant cela devenait trop poussif et vous avez découpé le lexique en sections alphabétiques.

- Examinez vos Filtres.

Certaines personnes ont peut-être passé du temps et fait des efforts importants pour créer des filtres intelligents. Certains seront obsolètes avec le nouveau Shoebox, mais d'autres peuvent restés très utiles.⁷⁷ A quelles bases de données s'appliquent ces filtres ? (De nombreux filtres contiennent des marqueurs, qui les rattachent à une base particulière. Si vous avez un filtre dans lequel vous remplacez régulièrement le marqueur, vous pouvez peut-être penser à l'affecter à toutes vos bases de données.)

L'information d'interalignement appartient au Type de la base de données dans lequel l'interalignement est fait. Aussi, si vous avez interaligné vos phrases illustratives dans votre lexique, l'information d'interalignement devrait s'appliquer au lexique. Néanmoins,

⁷⁶ Pour certains, l'impression du fichier SHA.SET sera une approche raisonnable. Ce fichier peut être vu comme une sorte de résumé de projet. Cependant d'autres le trouveront obscur et difficile à interpréter. Ces derniers préféreront aborder le projet et examiner les paramètres de configuration dans le contexte d'utilisation auquel ils sont accoutumés : Globals, Sort Orders, et autres boîtes du programme SH2.

⁷⁷ C'est peut-être le moment opportun pour lire les commentaires sur Filtres, page 113.

habituellement, l'information d'interalignement s'applique au Type de la base de données Texte.

Utilisez-vous des *caractères spéciaux*—tout ce qui est au-delà de A à Z—dans votre orthographe?

Disposez-vous d'une table de conversion de caractères ?

Si vous avez récupéré des données provenant de Word, vous en avez sûrement une. Sinon, contactez votre responsable informatique pour qu'il vous en procure une. Nous vous avons fourni la table IBM_ANSI.CCT; elle couvre l'ensemble des caractères européens du jeu de caractères ASCII étendu, mais pas les caractères de la section graphique.⁷⁸ IBM_ANSI.CCT et son inverse ANSI_IBM.CCT se trouve dans le dossier SHOEBOX\STD_SET.

Disposez-vous d'une police pour vos caractères spéciaux ?

Disposez-vous d'un clavier virtuel Keyman ?

Après avoir répondu à chacune de ces questions, vous devriez avoir une vision nette de ce qui est nécessaire pour transférer vos données vers Shoebox pour que cela fonctionne aussi bien que sous DOS.

Si cela vous semble une tâche trop lourde pour vous, sachez que vous pouvez toujours repartir de zéro, puis ouvrir vos fichiers de bases de données, créer des nouveaux Types de Base de Données, et définir vos Encodages de langues au fur et à mesure. Repartir de zéro n'est vraiment pas très compliqué. Voir Démarrer un nouveau projet, à la page 37, pour plus de détails. Amusez-vous bien.

Appendice B: version SH2.0, comparaisons

Les différences les plus importantes avec la version DOS de Shoebox

1. Vous pouvez ouvrir plusieurs fenêtres pour la même base de données (comme on pourrait s'y attendre de la part d'un logiciel Windows).
2. Chaque fenêtre peut être filtrée et triée différemment et afficher l'Index à la demande.
3. Vous pouvez trier par rapport à *n'importe quel champ*, et pas seulement par rapport au seul champ d'enregistrement, et vous pouvez trier sur plusieurs champs.
4. Vous pouvez Sauter vers *n'importe quel champ*, et pas seulement vers le champ du marqueur d'enregistrement.
5. Si vous avez trié par rapport à un champ qui présente des occurrences multiples dans un même enregistrement, vous rencontrerez cet enregistrement autant de fois que occurrences, lorsque vous vous déplacerez dans la base de données. (Nous appelons ceci l'*Indexation multiple*. Voir la section sur les *Fonctionnalités de base* ou bien le fichier d'Aide.)

⁷⁸ Si vous souhaitez réexporter vos données vers DOS, la table ANSI_IBM.CCT est là pour ça. Les corrections de correspondances doivent être faites dans les deux tables afin de pouvoir importer de DOS et réexporter vers DOS.

6. Des champs différents peuvent afficher différentes Polices, et le logiciel Keyman peut fonctionner avec Shoebox pour commuter entre des claviers virtuels lorsque le curseur entre dans un champ différent⁷⁹.
7. Une liste des marqueurs relie les marqueurs à l'information d'Encodage de langue de sorte que le programme sait quel ordre de tri utiliser lorsque vous triez par rapport à un champ donné et quelle Police et Clavier virtuel utiliser. Cette liste est entièrement modifiable, mais peut aussi être configurée comme un standard d'utilisation pour d'autres utilisateurs.
8. Vous pouvez Retrouver des données dans n'importe quel enregistrement de la base de données, et pas seulement dans l'enregistrement en cours. Retrouver peut également être restreint à un champ particulier ou à une encodage de langue donné.
9. La fonction d'interalignement peut explorer différents champs des enregistrements du lexique, et pas seulement le marqueur d'enregistrement, de sorte que des allomorphes peuvent être intégrés dans un enregistrement principal au lieu d'être dispersés dans le lexique.
10. Plusieurs affixes simples peuvent être isolés d'un mot—vous n'avez plus à lister toutes les combinaisons possibles dans une liste des affixes conjoints. (Vous pouvez néanmoins continuer à lister ces combinaisons.)
11. Les préfixes peuvent maintenant être distingués des racines, suffixes ou infixes, réduisant ainsi les ambiguïtés.
12. Lorsque vous avez choisi la glose (par exemple) parmi un ensemble d'ambiguïtés, le programme ira chercher la catégorie grammaticale (par exemple) correspondante, vous évitant ainsi d'avoir à choisir dans une autre boîte d'ambiguïtés.

Filtres

Certains filtres courants sous la version DOS de Shoebox, sont maintenant superflus puisque Shoebox sait trier par rapport à d'autres champs que celui du marqueur d'enregistrement, que la commande Retrouver sait faire sa recherche sur toute la base de données, et que la commande Sauter vers sait le faire vers des champs autres que la marqueur d'enregistrement ou le champ de tri primaire.

Le filtrage sur des critères comme *tous les noms* peut se faire en *triant* sur la catégorie grammaticale —et là vous avez tous les adjectifs, les noms, les verbes, etc., regroupés. Cependant si vous voulez tous les noms qui soient aussi des verbes (avec les deux champs de catégorie grammaticale dans la même entrée), alors il vous faudra un Filtre car le Tri ne saura pas faire la distinction.⁸⁰

Le filtre FIND couramment utilisé auparavant est bien remplacé par la commande Retrouver avec l'option *La base entière*, ou encore dans certains cas, par la commande Sauter vers. D'un autre côté, le filtrage crée un index des enregistrements répondant à la requête; Si vous devez utiliser Retrouver sur un ensemble particulier d'enregistrement, et aller et venir parmi eux, vous gagnerez sans doute en vitesse à faire construire un index (en utilisant le filtrage). Si vous ne devez aller là qu'une fois, Retrouver est probablement une meilleure solution.

⁷⁹ Attention, la commutation automatique des claviers fonctionne sous Windows 3.1, mais sous Windows 95 elle ne fonctionne qu'avec la version Windows 3.1 de Shoebox.

⁸⁰ Remarquez l'effet surprenant de la hiérarchie des marqueurs ici. Si la catégorie grammaticale est en dessous de l'entrée lexicale dans la hiérarchie, un filtre créé pour rechercher toutes les entrées qui soient à la fois des noms et des verbes fonctionnera comme prévu si la base de données est triée par rapport à l'entrée lexicale. Mais si la base de données est triée par rapport à la catégorie grammaticale, alors la recherche par filtre est restreinte à ce qui est hiérarchiquement au dessus de cette catégorie grammaticale—"l'enregistrement courant" tel qu'il est vu de la catégorie grammaticale—en particulier, cela n'inclue pas les autres catégories grammaticales!

L'ancien filtre FIELD est maintenant remplacé par le filtre *Marqueur/Séquence*—sans spécifier de texte. Ceci produit néanmoins un résultat différent du tri sur le champ. Si vous trie par rapport à un champ, tous les enregistrements qui ne contiennent pas ce champ ou dans lesquels ce champ est vide, seront listés en début, et la base de données sera triée suivant le contenu de ce champ. Si vous filtrez en utilisant un filtre Marq./Séquence sans texte, les enregistrements dans lesquels le champ n'existe pas ne seront **pas** inclus. Les occurrences vides de ce champ (marqueur sans contenu) apparaîtront suivant l'ordre de tri courant; ce qui veut dire qu'ils apparaîtront pas en début, sauf si vous trie par rapport à ce champ.

L'ancien filtre KEY est maintenant un simple filtre Marqueur/Séquence. Il n'y a pas de traitement spécifique pour le marqueur d'enregistrement.

Shoobox inclue également une option de filtrage sur l'encodage de langue du texte. Ceci permet le filtrage dans une langue donnée. (Les champs se voient affecter un encodage de langue à travers la boîte de dialogue Propriétés des marqueurs, et ceci est enregistré dans le fichier de Type de base de données.)

Shoobox s'est débarrassé des crochets et de la barre verticale utilisés par SH2 pour délimiter les contenus de Filtre. Les *Et*, *Ou* et *Non* liant les critères de filtre ainsi que les parenthèses sont toujours disponibles mais ils sont insérés par une commande spéciale. Ils ne doivent pas être tapés comme séquence dans le filtre.

NOTE: Le filtrage entraîne la création d'un index. Ceci prend un certain temps suivant la vitesse de votre machine, et la taille de votre base de données. Le tri suivant un champ particulier prend également un certain temps. Mais une fois l'index créé, la réponse est rapide. Retrouver ne crée pas d'index, aussi le délai est moindre pour une recherche ponctuelle.

Où est passé ? Cette section répertorie les fonctions disponibles dans les menus SH2 et localise la commande ou l'option correspondante dans Shoobox.

Item du menu DOS		équivalent Shoobox
File		
New		idem
Open		idem
Close		idem
Enregistrer		idem
Save as		idem
Options	box name	pas d'équivalent, le nom du fichier s'affiche
	date stamp marker	Base de données, Propriétés, onglet Options
	indirect date stamp	pas d'équivalent
	autoload	tous les fichiers chargés au moment de la sortie seront chargés la prochaine fois
	database template	Base de Données, Modèle
	sort order	Base de Données, Trier, Ordre de tri du premier champ <i>ou</i> Projet, Encodage de langues, <i>sélectionner encodage</i> , onglet Ordres de tri, spécifier un ordre de tri par défaut pour cet encodage

	key length	fin du champ ou fin de la première ligne, par défaut maximum=100 (plutôt que 10)
Merge		Base de Données, Fusionner des bases
DOS		pas d'équivalent (Win'95, faire Démarrer, Exécuter, Démarrer, Programmes, Commande MS-DOS; Win 3.1, choisir Commande DOS)
Output		Fichier, Exporter
Close		idem
Box		
Box 1-7		fenêtres et bases de données multiples permises
Textfile		pas d'équivalent
Edition		
Edition Key		Le champ "clef" est désormais traité comme les autres champs, édition directe
Margins		Outils, Enregistrer la longueur de ligne et Passage à la ligne automatique—voir l'Aide à ce sujet
Cut		idem, raccourci Ctrl+X, Maj.+Suppr
Copy		idem, Ctrl+C, Ctrl+Inser
Paste		idem, Ctrl+V, Maj.+Inser
Extended Copy		pas d'équivalent direct, Windows utilise le Presse-livre, voir la section <i>Notes diverses</i> au dessus
Extended Cut		pas d'équivalent direct, Windows utilise le Presse-livre, voir la section <i>Enregistrer votre Presse-papiers</i> au dessus , p. 80
Find		idem
Repeat Find		Retrouver le suivant, également dans la barre d'outils (jumelles, Suivant ou Précédent)
Reshape		Outils, Restructurer
Undo All		idem, Shoebox annule également en mode incrémentiel
Jump		idem, également w :Clic-droit ou m :Option+clic
Jump-Insert		idem, également w :Ctrl+Clic-droit ou m :control+Option+clic
Search Path		Edition, Sauter vers, Destination de saut (NOTE: <i>Rechercher</i> ne recherche que dans le champ de tri primaire de la base de données active. <i>Sauter vers</i> ira vers n'importe quel couple base de données-marqueur défini dans la Destination de saut.)

Database

Next		idem, également dans la barre d'outils (flèche pointant à droite)
Previous		idem, également dans la barre d'outils (flèche pointant à gauche)
Top		Base de Données, Premier enregistrement, également dans la barre d'outils (flèche pointant à gauche butant sur une ligne)
Bottom		Base de Données, Premier enregistrement, également dans la barre d'outils (flèche pointant à droite butant sur une ligne)
Find		idem, voir la note au dessus au sujet de Edit, Search path
Browse		Affichage, Index (voir aussi Affichage, Champs affichés dans l'Index)
Filters		Base de Données, Filtrer <i>ou</i> Base de données, Propriétés, onglet Filtres
Tally		Lorsqu'un filtre est actif, la barre d'état signale automatiquement <i>numéro d'enregistrement/nb occurrences filtrées</i> au lieu de <i>numéro d'enregistrement/nb d'enregistrements</i>
Record		
Insert		Base de Données, Insérer un enregistrement
Delete		Base de Données, Supprimer l'enreg.
Move		Base de Données, Déplacer l'enreg.
Copy		Base de Données, Copier l'enreg.
Join		Créer une seconde fenêtre, et faites Couper/Coller
Utility		
Macros		utiliser Macro <i>Recorder</i> (voir <i>L'enregistreur de Windows pour créer des macros</i> , p. 77 ou le fichier d'Aide)
Record macro		utiliser Macro <i>Recorder</i> (voir l'Aide)
Reversal		Faites en sorte que la fenêtre courante soit triée par rapport au champ voulu; utiliser Fichier, Exporter et sélectionner <i>Les enregistrements courants</i>
Inventory	of markers	Base de données, Propriétés, onglet Marqueurs NOTE: liste uniquement, pas de nombre occurrences
	of words	Outils, Liste des mots
Number Text		pas encore implémenté, utiliser TEXTPREP.CCT, voir Texte interaligné, p. 49

Extended Char	utiliser la <i>Table des caractères</i> de Window; si elle est installée, faire Démarrer, Programmes, Accessoires (pour Win 3.1 regarder dans le groupe Accessoires)
Spell Check	Outils, Contrôle de l'orthographe (NOTE: le paramétrage de l'interalignement est nécessaire)

Globals

Interlinear Setup	Base de données, Propriétés, onglet Interalignement
Conjoined Affixes	Désormais dans le lexique ou dans un autre fichier de base de données; voir Texte interaligné, démarrant en page 49
Begin Interlinear	Outils, Interalignement, également dans la barre d'outils, ou encore w :Alt+I ou m :commande+I
Display	Pas d'équivalent, voir le <i>Panneau de configuration</i> de Windows (<i>Bureau</i> sous Win 3.1)
Sort order	Projet, Encodage de langues, <i>choisir l'encodage</i> , onglet Ordres de tri, spécifier l'ordre alphabétique par défaut pour l'encodage
Variables	Projet, Encodage de langues, <i>choisir l'encodage</i> , onglet Classes
Punctuation	Projet, Encodage de langues, <i>choisir l'encodage</i> , onglet Classes
Range Sets	w :Clic-droit ou m :Option+clic sur le marqueur <i>ou</i> Base de données, Propriétés, onglet Marqueurs, <i>choisir le marqueur</i>
Insert Range Member	Edition, Insérer une valeur prédéfinie

Aide

Aide	idem
Keys	pas d'équivalent