Ein Mini-Wörterbuch und eine Mini-Interlinearisierung in Shoebox

Im Folgenden geht es darum, zur Einführung in die Arbeitsweise des Shoebox-Programmes ein englisches Mini-Wörterbuch anzulegen und es auf die Interlinearisierung eines kleinen englischen Textes anzuwenden.

Vorarbeit:

Legen Sie auf Ihrem Verzeichnis auf dem Server (username\$ auf "FB10pdc" (Z):) im Ordner Computerwerkzeuge einen Ordner namens Shoebox an und darin einen weiteren Ordner namens Englisch:



Mini-Wörterbuch



> New und speichern ein neues Lexikon unter dem Namen minilex.lex in dem neu erstellten Verzeichnis ab. Achten Sie darauf, wirklich die Kennung .lex einzugeben. Daraufhin öffnet sich das folgende Fenster:



Select Database Type	×	
Please select the database type for:	ОК	
test.iex	Cancel	•
	<u>H</u> elp	
Add New Type		

Klicken Sie auf den Schalter Add New Type... Es öffnet sich das folgende Fenster:

Database Types		×	Κ
<u>N</u> ame	File Name	Cl <u>o</u> se	E
		<u>A</u> dd	E
		<u>С</u> ору	Ν
		<u>M</u> odify	Si
		<u>D</u> elete	(f
		<u>H</u> elp	
Settings folder: Z:\Sh	oebox\Test		

Klicken Sie auf den Schalter Add...

Es öffnet sich ein weiteres Fenster mit zwei Eingabefeldern. Tragen Sie in das erste Feld als Namen für die neue Database Minilex ein. Tragen Sie in das zweite Feld als Datensatzmarkierer Ix (für Lexem) ein:

New Database Type	>
Name:	OK
Minilex	Cancel
Becord Marker:	Help
x 💌	2.4
	•

Klicken Sie auf OK. Es öffnet sich das folgende Fenster, in dem Sie auf der Karte Options eine Kurzbeschreibung der Database eingeben können.

Database Type P	roperties - Test 🛛 🗙	Klicken Sie anschließend auf OK.
Markers Filters	Jump Path Options Interlinear	Klicken Sie in den beiden darauf-
Type <u>N</u> ame:	Test	folgenden Fenstern jeweils auf Close
Description:	Database für das Minilexikon	bzw. OK.
		Daraufhin öffnet sich das Lexikonfenster
		mit einem Datensatzmarkierer (\lx) im
	<u> </u>	Datensatzmarkiererfeld und der Eingabe-
<u>R</u> ecord Marker:	lx .	aufforderung im rechten Feld:
Date <u>S</u> tamp Marker:	[none]	🖌 minilex.lex
	Multi-Dictionary Formatter	\lx
	OK Abbrechen Hilfe	

Geben Sie hinter die Eingabeaufforderung auf der rechten Seite das Wort *the* ein. Gehen Sie mit Enter in eine neue Zeile und betätigen den Backslash (\). Geben Sie als neuen Datensatzmarkierer ps ein und bestätigen Sie, dass sie

Marker Propert	ies	<u>×</u>
<u>M</u> arker: (without \)	ps	ОК
Field <u>N</u> ame:	Part of Speech	Cancel
Under what in the Hierarchy:	lx 💌	<u>H</u> elp
Marker for f <u>o</u> llowing field:	[none]	🔽 Use Language Fon
Language Encoding:	Default	Choose <u>F</u> ont
Description:	Language Properties	<u>R</u> ange Set
		Style to Export
		C Character
		eragraph
		No Word Wrap

diesen aufgenommen haben möchten. Darauf erscheint das folgende Fenster, in dem Sie den Field Name als Part of Speech angeben

'Part of Speech' (Redeteil) entspricht 'lexikalischer Kategorie' und es wäre eigentlich angenehmer, hier so etwas wie z.B. \c oder \cat zu verwenden, da wir diese Bezeichnungen auch in PC-PATR verwenden. Im Shoebox Tutorial aber wird mit \ps gearbeitet und um das Ganze nicht unnötig zu verkomplizieren behalten wir diese Bezeichnung bei.

_ 🗆 ×

Sie gelangen zurück in das Lexikonfenster. Geben Sie als lexikalische Kategorie Det für Determinator ein:

Vervollständigen Sie den Lexikoneintrag für 'the', so dass er die folgende Information umfasst

😭 min	ilex.lex	
\lx	the	
∖ps	Det	

(gl steht für 'Gloss')und enthält im vorliegenden Beispiellexikon entweder 'grammatische' Information wie Definitheit beim Artikel oder eine deutsche Übersetzung bei Substantiven und Verben. In dem Fenster Marker Properties können Sie neben der Feldbezeichnung auch die Reihenfolge der Marker festlegen.

🙈 minilex.lex

the

Det

Def

١x

\ps

\gl

Erstellen Sie jetzt aus diesem Lexikoneintrag ein Template, sozusagen ein Muster für die Lexikoneinträge im Mini-Lexikon.

Dazu gehen Sie auf Database > Template, es erscheint das folgende Fenster:

•••••••••••••••••••••••••••••••••••	
Set Template 🛛 🗙	
The record template will be set to the fields contained in the current record.	
Cancel	
Include field contents	

Klicken Sie auf OK.

Gehen Sie nun mit der Schreibmarke vor das 'the', klicken Sie auf Database und Insert New Record oder drücken Sie die Tasten Strg + N. Geben Sie als Lexem in das erscheinende Fenster das Wort boy ein. Wie Sie feststellen werden, ist das Lexikonfenster für das neue Lexem nach dem Muster des alten Lexems aufgebaut, entspricht mithin dem 'Template'. Sie müssen nur die Angaben zu Part of Speech und Gloss vervollständigen.

Erstellen Sie auf der Grundlage dieser Arbeitsschritte das folgende Mini-Lexikon:

Lexem	the	boy	dress	girl	in	love	red	- <i>S</i>	-5
Part of Speech	Det	N	N	N	Р	V	А	Suff	Suff
Gloss	Def	junge	kleid	mädchen	in	lieb-	rot	3.Pers.Sg	Pl

Wenn Sie das zweite -s-Suffix eingeben, erkennt Shoebox, dass bereits ein Eintrag mit denselben Daten im Lexemdatenfeld vorhanden ist. Es erscheint ein Fenster namens 'Multiple

Markers'. Klicken Sie hier auf 'Insert' um auch das zweite Suffix in das Lexikon aufnehmen zu können.

Überprüfen Sie über die Browse-View Funktion, dass Sie alle Einträge aufgenommen haben. Das Fenster soll mithin wie nebenstehend aussehen.

Mini-Interlinearisierung

Starten Sie über Start > Programme > Zubehör > Editor den Windows-Editor, geben Sie den folgenden Text ein, speichern diesen unter der Bezeichnung minilex.txt in dem Ordner Shoebox\Test auf Ihrem Verzeichnis auf dem Server ab und verlassen Sie den Editor wieder:

The girls love the boy. The boy loves the girl in the red dress. The cat under the car loves the boys. The boy kicked the ball into the window.

Gehen Sie in Shoebox auf File > Open und wählen Sie die Datei minilex.txt aus. Darauf erscheint das folgende Fenster:

Import		×
This file is not in Shoebox 3 / 4 file form Z:\Shoebox\Englisch\minilex.txt Please select an appropriate database type, or add a new one:	vat:	OK Cancel <u>H</u> elp
Add a new Database Type Make Consistent Changes	The girls love the b	xoy. The boy I 🛌
Use Change Table: None Browse	▼ Backup original	es *.0RI

Klicken Sie auf Add a new Database Type und im nächsten Fenster auf Add....

Es öffnet sich (erneut) das folgende Fenster mit zwei Eingabefeldern. Tragen Sie in das erste Feld als Namen für die neue Database Minilex Interlineartext ein. Tragen Sie in das zweite Feld als Datensatzmarkierer ref ein und klicken Sie auf OK:

New Database Type	Þ	K
Name:	OK	
Minilex Interlineartext	Cancel	
Becord Marker:	Help	

Geben Sie als Kurzbeschreibung in das nächste Fenster auf der Karte Options 'Interlineartext für das Minilex-Projekt' ein.

In einem nächsten Schritt muss für diese Database ein Bezug hergestellt werden zwischen dem Minilexikon und dem Minilex-Text.

Klicken Sie dafür auf die Karte Interlinear und den Schalter Quick Setup.

Es erscheint das Fenster Quick Setup Interlinear Text Marker:

Übernehmen Sie die vorgebenen Werte, d.h. klicken Sie auf OK.

Im nächsten Fenster (Quick Setup Lexikon Markers) müssen a) minilex.lex in

Quick Setup Inter	inear Text Markers	×
Tevt Marker		OK
Morpheme Breaks	m	Cancel
Gloss	9	<u>H</u> elp
Part of Speech	P	

die Dateiliste übernommen und b) die Tags für den Interlineartext mit den Tags aus dem Lexikon verbunden werden. In unserem Minibeispiel sind dafür nur zwei Änderungen notwendig, es muss das 'p' dem 'ps' und 'g' dem 'gl' zugeordnet werden:

Quick Setup Lexicon Mark	ers		×
Available Databases minilex.lex	Insert ->	Databases to <u>S</u> earch minilex.lex	OK Cancel <u>H</u> elp
Show Whole Path Gloss Marker Part of Speech Alternate form Underlying form	gi pd a u		

Klicken Sie auf OK. Sie kommen zurück auf das Fenster Database Type Properties. Wählen Sie darin die Karte Jump Path und aktivieren Sie den Schalter First:

Database Type P	roperties - Test Inte	erlineartext	×
Markers Filters	Jump Path Options	Interlinear	
<u>A</u> vailable Databases		Databases In <u>P</u> ath	Field to Search
minilex.lex	<u> </u>	minilex.lex	k
	Last >	(
	Insert >		
E-U-C	< <u>R</u> emove		
Field to Search	▼ < <u>C</u> lear		
□ Show <u>W</u> hole F	Path		
Notify me befo	re closing a database r	eferenced by a jump p	bath
	OK Abbreche	n	Hilfe

Klicken Sie auf anschließend auf die Karte Markers. Hier müssen Sie noch die Feldnamen für die Markierungen des Interlineartextes eingeben Klicken Sie dazu auf den entsprechenden Marker, wählen Sie Modify und geben die Bezeichnung ein:

Database	Patabase Type Properties - Test Interlineartext 🛛 🔀					
Markers	Filters Jump Path	Options I	nterlinea	ar]		
Mar <u>k</u> er	Field Name	Language	Under	SFR		
\g	Glosse	Default	ref	Р	Add	
\m	Morphem	Default	ref	Р		
١p	Part of Speech	Default	ref	Р	<u>С</u> ору	
\ref	Referenz	Default		Р	Modifu	
M	Text	Default	ref	Р		
					Delete	
					Note: Cancel	
					doesn't undo	
Default I	Language Encoding:			_	Modify, or	
(for new	ly added markers) 🗌	Derauit		<u> </u>	Delete.	
	ОК	Abbrechen			Hilfe	

Klicken Sie anschließend auf OK und im Database Type Fenster auf Close. Sie gelangen wieder auf das Fenster Import.

Import		×
This file is not in Shoebox 3 / 4 file form	at	OK
Please select an appropriate database tune, or add a new one:		Cancel
Database Type	The side laws the b	Help
Add a new Database Type	The gris love the t	oy. The boy 1
Make Consistent Changes		
None		
Browse	•	⊻ ►
Remove hanging indent spaces	Backup <u>o</u> riginal	as *.ORI

Hier müssen Sie das Kästchen Use Change Table aktivieren. Sie werden aufgefordert, dem Programm den Pfad für die entsprechende Datei mitzuteilen. Es handelt sich dabei um einen Filter, dessen Aufgabe ist u.a. darin besteht, den zu interlinearisierenden Text in Datensätze zu zerlegen, wobei das Trennkriterium die das Satzende markierenden Satzzeichen sind (Punkt, Fragezeichen, Ausrufungszeichen).

Diese Filter-Datei heißt textprep.cct und befindet sich (hoffentlich) im Ordner D:\Programme\ShoeboxSettings.

Wählen Sie die Datei aus und klicken Sie auf Öffnen. Sie gelangen wieder auf das Fenster Import. Klicken Sie auf OK. Daraufhin wird der Text importiert und das folgende Fenster öffnet sich:

🚷 mini	lex.txt
\ref	001
//	The girls love the boy.
A	

Gehen Sie mit der Schreibmarke vor das erste Wort des Satzes und klicken Sie auf den Interlinearisierungsbutton. Sie werden aufgefordert, anzugeben, ob es sich bei dem Suffix -s um 3. Pers. Sg oder um die Pluralmarkierung handelt. Wenn Sie das getan

haben, wird der Satz durchinterlinearisiert und Sie können unter f noch eine freie Übersetzung eingeben:

🚯 mir	nilex.txt					_ 🗆 ×
\ref	001			.1		
\t \m	the oirl	- 9	love	the the	boy. boy	
\g	Def mädchen	-P1	lieb-	Def	junge	
\p	Det N	-Suff	V	Det	N	
A	"Die Mädchen lieben den Jungen"					

Verfahren Sie genau so mit dem nächsten Datensatz (= der zweite Satz). Beim dritten Satz werden Sie feststellen, dass nicht alle Wörter im Lexikon enthalten sind. Shoebox kommt genau so weit wie im folgenden Screenshot:

🚡 Shoebox - [testtext.txt]					
🚯 <u>F</u> ile	<u>E</u> dit <u>D</u> atabase Pr <u>i</u>	<u>oj</u> ect <u>T</u> ools ⊻iew ⊻	∕indow <u>H</u> elp		_ 8 ×
¢ 🛛	X @ @ ←	→ ┝ → 票			▼ ∰,∰ [nc
∖ref	003				
\t	The cat under	the car loves		the	
\m	the *** ***	the *** love	-S	the	
١g	Def *** ***	Def *** lieb-	-3.Pers.Sg	Def	
\p	Det *** ***	Det *** V	-Suff	Det	
\t \m \g \p	boys boy -s junge -3.Pers.S N -Suff	Šg			
For Hel	p, press F1	\ref 003			3/3 te: //

Klicken Sie in solchen Fällen mit der RECHTEN Maustaste auf das nicht erkannte Wort und ergänzen Sie das Lexikon. Wenn Sie anschließend vor das nicht erkannte Wort gehen und den Interlinearisierungsbutton erneut klicken, wird auch dieses bearbeitet.