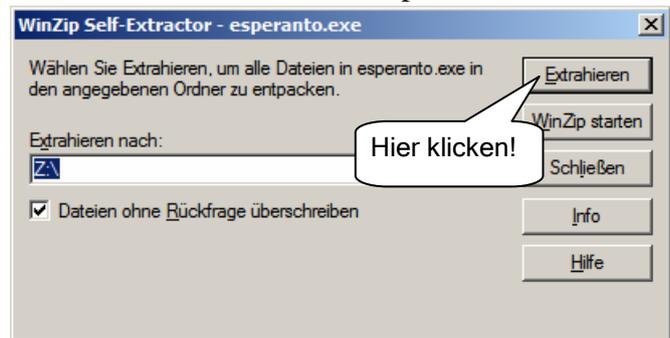


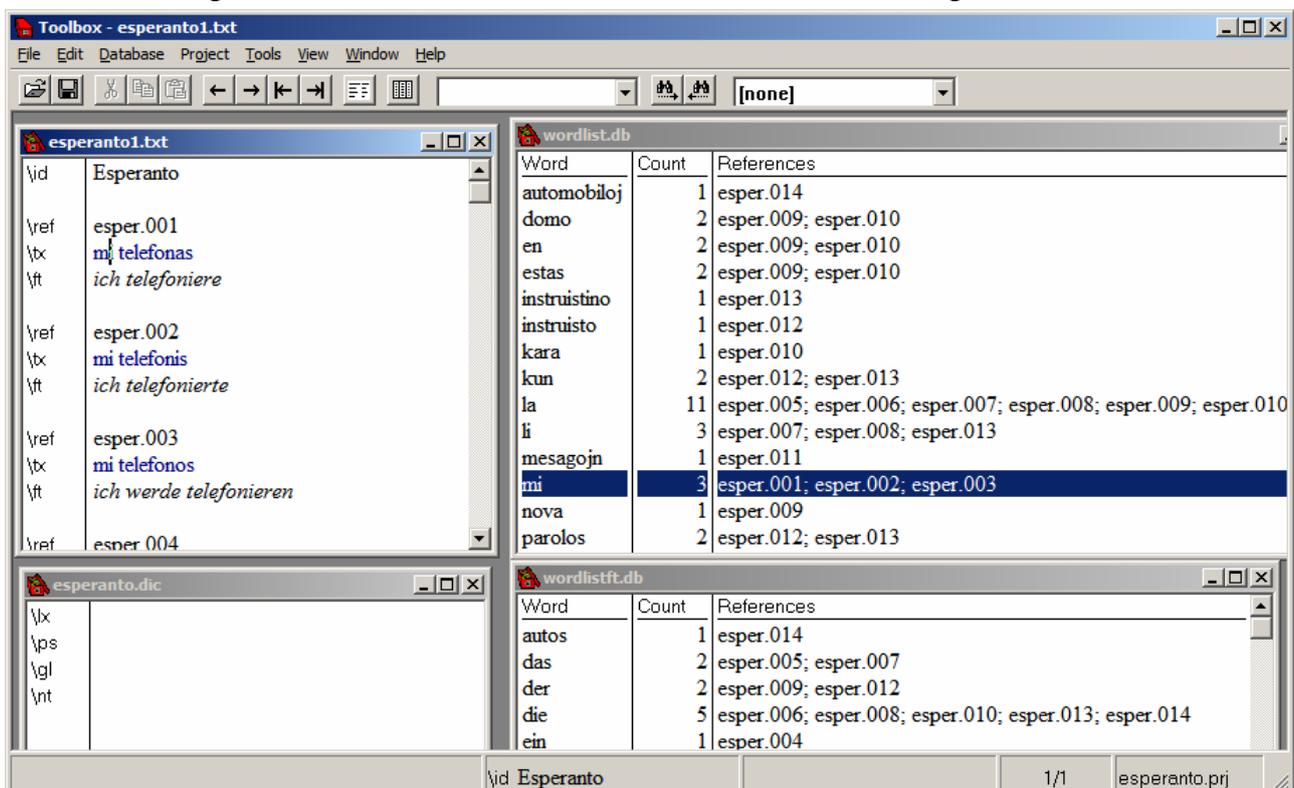
Esperanto-Aufgabe: Anleitung zur Analyse

Für diese Aufgabe wurde eine spezielle Toolbox-Konfiguration vorbereitet, die alle wesentlichen Einstellungen enthält. Sie können sich also unmittelbar der Aufgabe der morphologischen Analyse zuwenden.

Schritt 1. Im Public-Verzeichnis (Laufwerksbuchstabe P:\) auf dem Server befindet sich im Verzeichnis COMPUTERWERKZEUGE >> SHOEBOX >> ESPERANTO eine Datei mit dem Namen ESPERANTO.EXE. Es handelt sich dabei um eine selbstentpackende ZIP-Datei. Führen Sie dieses "Programm" mit einem Doppelklick aus. Es öffnet sich das nebenstehende Fenster. Klicken Sie auf den Schalter **Extrahieren**. Die Dateien sollten dann mit der richtigen Verzeichnisstruktur auf Ihre Festplatte kopiert werden. Möglicherweise werden dabei bereits im Esperanto-Verzeichnis vorhandene Dateien überschrieben, so insbesondere die Esperanto-Sprachkodierung.



Schritt 2. Gehen Sie in Ihr Esperanto-Verzeichnis (Z:\Computerwerkzeuge\Shoebox\Espesanto) und klicken Sie doppelt auf die Datei ESPERANTO.PRJ, um das Projekt zu öffnen. Das Programm Toolbox öffnet sich mit vier Fenstern wie nachfolgend zu sehen ist.



Das Fenster mit der Titelleiste *esperanto1.txt* zeigt den zu bearbeitenden Text, darunter findet sich das noch leere Wörterbuch *esperanto.dic*, rechts oben sehen Sie eine alphabetische Liste aller im Text vorkommenden Esperanto-Formen mit Belegstellen (*wordlist.db*), die von *Toolbox* automatisch erstellt wurde, und darunter eine entsprechende Liste *wordlistft.db* mit den deutschen Entsprechungen, die auf die Übersetzungen im \ft-Datenfeld des Textes zurückgeht. Wenn Sie mit der rechten Maustaste auf ein Esperanto-Wort im Text klicken, werden Sie feststellen, dass das Programm zur Esperanto-Wortliste springt und die entsprechende Form anzeigt. Entsprechendes

gilt für die Wörter in der deutschen Übersetzung. Wenn Sie z.B. auf *ich* im ersten Satz (esper.001) klicken, wird dieses in der Wortliste angezeigt.

Diese Wortlisten können insbesondere bei umfangreichem Datenmaterial sehr hilfreich sein, wenn es darum geht, äquivalente Formen zu vergleichen. Wenn Sie beispielsweise auf das Wort *telefonas* im gleichen Satz klicken, erhalten Sie nacheinander stehend eine ganze Liste von Wörtern mit der Wurzel *telefon*. Wir werden zu einem späteren Zeitpunkt die damit gegebenen Möglichkeiten weiter eingehen.

Achtung – Neuerung:

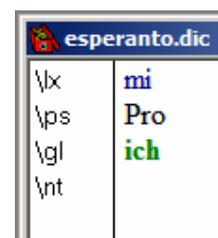
Sie werden feststellen, dass der zu bearbeitende Esperanto-Text sich anders verhält, als Sie es bisher gewohnt sind. Dies liegt darin, dass der Datensatz (*record*) nicht mehr der Satz ist, sondern der ganze Text. Wenn Sie also auf einen der Pfeilschalter  klicken, kommen Sie nicht weiter, weil es – bis jetzt – nur einen Datensatz gibt.

Schritt 3. Sie können jetzt mit der Arbeit beginnen, indem Sie den ersten Satz (esper.001) interlinearisieren, indem Sie entweder auf den Schalter  klicken, oder – was auf die Dauer schneller ist – die Tastenkombination **Alt-I** drücken.

Das Ergebnis sehen Sie rechts: Da das Wörterbuch noch leer ist, kann die Analyse nur fehlschlagen. Das Programm ist aber so eingestellt, dass dabei nicht einfach – wie bei den Feldern `\gl` und `\ps` – drei Sterne angezeigt werden, sondern es versucht, einen brauchbaren Vorschlag zu machen. Dabei ist es umso erfolgreicher, je umfangreicher das Wörterbuch bereits ist. Klicken Sie jetzt mit der rechten Maustaste auf das rostfarbene ***mi**. **N.B.** gehen Sie **nicht** vom Textfeld `\tx`



aus, weil Sie sonst in der Esperanto-Wortliste landen! Wenn Sie vom Morphemfeld `\mb` ausgehen, wird versucht die Form im Wörterbuch zu finden. Da dies misslingt, öffnet sich das nebenstehende Fenster. Klicken Sie auf den Schalter **Insert**. Im Wörterbuch wird ein Eintrag für das Lexem *mi* gemacht, für das Sie die Eingaben für die Wortart und die Kurzdefinition (Glosse) nachtragen müssen. Aus den Vorarbeiten (Minigrammatik, Besprechung der Hausaufgabe) wissen Sie, dass *mi* dem dt. *ich* entspricht. Es ist ein Pronomen. Tragen Sie für *Part of Speech* **Pro** ein und als Glosse **ich**.



Wichtig!!

Sie können sich die Arbeit enorm erleichtern, wenn Sie nach einem Eintrag im Wörterbuch nicht durch einen Mausklick zum Text zurückkehren sondern mit der Tastenkombination **Strg-R**. Wenn Sie dies tun, wird nach dem Zurückspringen in das Textfenster automatisch die Interlinearisierung erneuert. Dies funktioniert allerdings nur, wenn sich auch vom Textfenster aus in das Wörterbuch "gesprungen" sind.

Schritt 4. Drücken Sie also die Tastenkombination **Strg-R**. Das bringt Sie zum Text zurück. Sie werden feststellen, dass die Glosse und die Wortart dort jetzt nachgetragen sind. Sollte dies nicht der Fall sein, müssen die Interlinearisierung noch einmal mit **Alt-I** vornehmen.

Schritt 5. Als nächstes wenden wir uns der Form ***telefonas** zu. Sie wissen aus den Vorarbeiten, die Sie – hoffentlich – gemacht haben, dass dieses Wort aus der Wurzel **telefon** und

dem Suffix **-as** besteht. Klicken Sie jetzt mit der rechten Maustaste im Morphemfeld (**\mb**) auf ***telefonas** und machen Sie einen Eintrag für **telefon** (Sie müssen einfach das **as** weglöschen). Verwenden Sie als *Part of Speech* **NW** (für *nominale Wurzel*) und als Glosse *Telefon*. Kehren Sie dann wieder mit **Strg-R** zum Textfenster zurück. Nachdem jetzt ein Eintrag für die Wurzel **telefon** vorhanden ist, wird sofort als Zerlegung *telefon-as* angeboten.

\ref	esper.001
\tx	mi telefonas
\mb	mi telefon - *as
\gl	ich Telefon - ***
\ps	Pro NW - ***
\ft	

Schritt 6. Machen Sie als nächstes einen Eintrag für die Endung **-as**. Vergessen Sie dabei keinesfalls den Bindestrich vor dem **a**, weil der Eintrag sonst nicht als Suffix erkannt wird. Da es sich bei **-as** um ein Tempusmorphem handelt, wollen wir als *Part of Speech* **Tmp** eintragen und als Glosse **Präs** (für **Präsens**). Rückkehr wie gewohnt mit **Strg-R**.

Schritt 7. Interlinearisieren Sie als nächstes den Satz **esper.002** (*mi telefonis*). Sie werden feststellen, dass wir schon ziemliche Fortschritte gemacht haben. Hier muss nur noch ein Eintrag für das Suffix **-is** gemacht werden ([Tmp: Prät]). Der nächste Satz (**esper.003**) geht analog. Etwas Neues liefert uns erste wieder Satz **esper.004**:

\ref	esper.002
\tx	mi telefonis
\mb	mi telefon - *is
\gl	ich Telefon - ***
\ps	Pro NW - ***
\ft	
\ft	ich telefonierte

\ref	esper.004
\tx	telefono sonoras
\mb	telefon - *o *sonor -as
\gl	Telefon - *** -Präs
\ps	NW - *** -Tmp
\ft	ein Telefon klingelt

Hier scheitert das Programm an ***o** und ***sonor**. Wir wissen, dass **-o** das Kennzeichen für Nomina (**Nom**) ist. Diese Kennzeichnung wollen wir mit dem Symbol **KM** (für *kategoriale Markierung*) anzeigen: [KM:Nom], wobei **KM** als *Part of Speech* und **Nom** als Glosse einzutragen ist. Das verbleibende ***sonor** ist eine verbale Wurzel (**VW**).

esperanto1.txt	
\ref	esper.004
\tx	telefono sonoras
\mb	telefon -o sonor -as
\gl	Telefon -Nom klingel -Präs
\ps	NW -KM VW -Tmp
\ft	ein Telefon klingelt

Die Interlinearisierung von Satz **esper.005** erfolgt nahezu problemlos, da fast alle Bestandteile bereits im Wörterbuch vorhanden sind. Was fehlt, ist der bestimmte Artikel **la**. *Part of Speech* sei **Art** (für Artikel) und die Glosse **Def** (für *definit*, da es ja der bestimmte Artikel ist).

\ref	esper.006
\tx	la telefonoj sonoras
\mb	la telefon - *oj sonor -as
\gl	Def Telefon - *** klingel -Präs
\ps	Art NW - *** VW -Tmp
\ft	
\ft	die Telefone klingeln

In Satz **esper.006** gibt es Probleme mit *telefonoj* 'Telefone'. Das Suffix **-o** müsste eigentlich bekannt sein. Es funktioniert trotzdem nicht, weil es keinen Eintrag für die Pluralmarkierung **-j** gibt. Klicken Sie mit der rechten Maustaste auf ***oj** und machen Sie einen Eintrag für **-j**, indem Sie das **o** durch den Bindestrich ersetzen. Als *Part of Speech* tragen Sie **Num** ein (für *Numerus*) und als Glosse **Pl** (für *Plural*). Nach Rückkehr mit **Strg-R** ist das Problem gelöst.

\ref	esper.007
\tx	li purigis la telefonon
\mb	*li *purig -is la telefon - *on
\gl	*** *** -Prät Def Telefon - ***
\ps	*** *** -Tmp Art NW - ***
\ft	 er reinigte das Telefon

Bei Satz **esper.007** gibt es eine Reihe von Problemen. Es fehlt ein Eintrag für das Personalpronomen **li** 'er' ([Pro: er]) sowie für das Verb **purig** 'reinigen'. Man könnte es sich einfach machen und **purig** als Wurzel auffassen. Es ist aber abgeleitet aus der Adjektivwurzel **pur** 'rein' und einem Derivationsuffix (DS) **-ig** 'machen, bewirken', hier also 'rein machen': [DS:Kaus] (Kaus von *Kausativ*). Bei *telefonon* besteht das Problem darin, dass das Suffix **-n** zur Markierung des Objekts ([Kas:Obj]) noch nicht bekannt ist.

In Satz **esper.009** fehlt die nominale Wurzel **patr** 'Vater', die Verbwurzel **est** 'sei', die Präposition **en** 'in', die Adjektivwurzel (AW) **nov** 'neu', die Kategorialmarkierung **-a** für das Adjektiv (**Adj**) sowie die NW **dom** 'Haus'.

In Satz **esper.010** bedarf *patrino* einer Sonderbehandlung. Das Wort enthält das Derivationsuffix **-in** in der Bedeutung 'weibliches Lebewesen' (**Fem** von *female* 'weiblich', nicht *feminin*).

In Satz **esper.011** müssen zunächst zwei Kodierungsfehler beseitigt werden. Das erste Wort muss *sxi* lauten, nicht *si*, das letzte Wort *mesagxojn* und nicht *mesagojn*. Wir machen Einträge für das Pronomen **sxi** 'sie', die Verbwurzel **send** 'senden' und die Nomenwurzel **mesagx** 'Nachricht'.

In Satz **esper.012** müssen wir die Basis *instruist* analysieren. Sie besteht aus der Verbalwurzel **instru** 'lehren' und dem Derivationsuffix **-ist** 'jemand der berufsmäßig etwas tut, "Profi"'. Weiterhin sind zu ergänzen die Verbalwurzel **parol** 'sprech', die Kategorialmarkierung **-e** für das Adverb (**Adv**), sowie die Präposition **kun** 'mit'.

In Satz **esper.014** fehlt die Nomenwurzel **auxtomobil** 'Auto', die Adjektivwurzel **rapid** 'schnell' sowie die Verbwurzel **vetur** 'fahr'.