

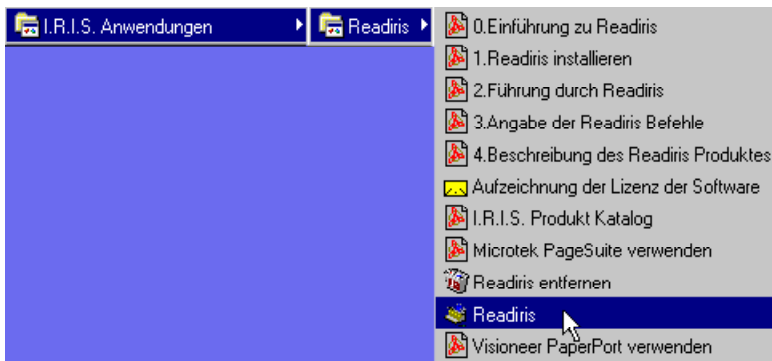
Kapitel 2

KURZE FÜHRUNG

Readiris ist ein OCR-System auf dem neuesten Stand der Technik, das mit einer Vielzahl von fortgeschrittenen Eigenschaften ausgerüstet ist. Dieses Kapitel befaßt sich mit den wichtigsten Eigenschaften und gibt Ihnen viele Tips und Anweisungen, bezüglich der Anwendung von Readiris.

STARTEN DER SOFTWARE

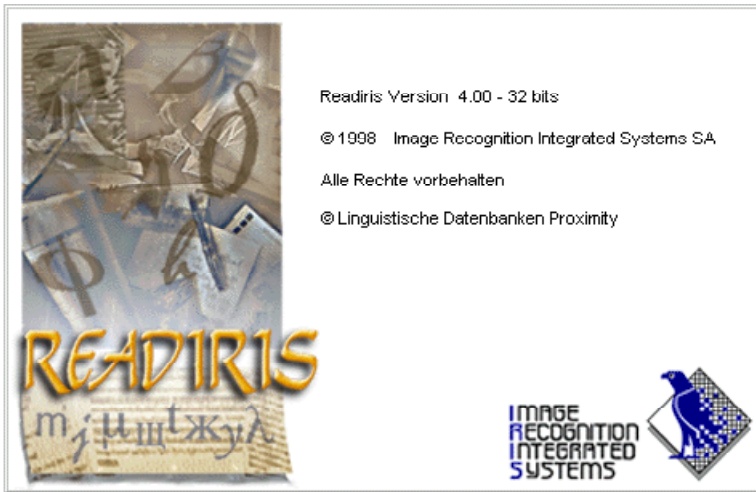
Klicken Sie die Readiris-Taste im Submenü „I.R.I.S. Anwendungen - Readiris“ oder klicken Sie die Verknüpfung in Ihrem Desktop.



Nach einigen Augenblicken erscheint das Readiris-Startfenster sowie das Anwendungsfenster. Das Readiris-Fenster enthält die Versionsnummer sowie



die Copyrights der Readiris Software. Mit einem Mausklick verschwindet das Readiris-Fenster.

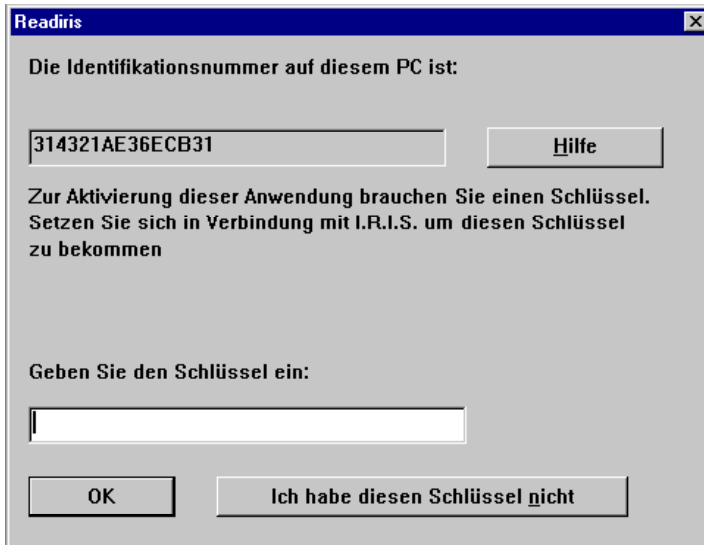


BEIM ERSTEN START

Je nach Software, werden Sie beim ersten Start gefragt, Ihre Lizenz zu registrieren.

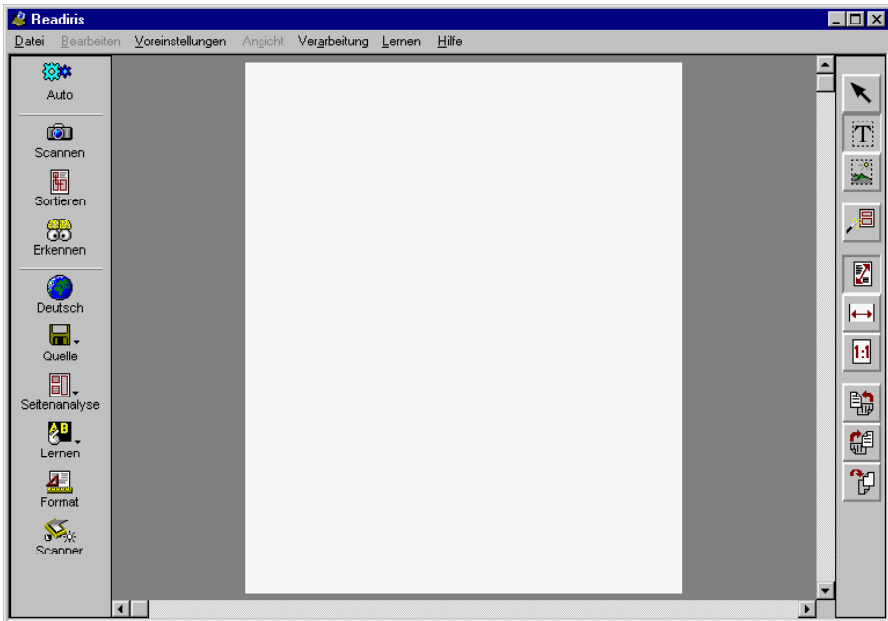
In diesem Fall ist der Gebrauch von Readiris auf einem Monat beschränkt. Bei Registrierung bekommen Sie einen **Softkey** gratis um die Software nach 30 Tagen weiter zu benutzen.

Um diesen Softkey zu generieren benötigen Sie unbedingt die **Identifikationsnummer**; Bitte vergewissern Sie sich, daß diese Nummer verfügbar oder notiert ist wenn Sie Ihre Lizenz registrieren.



ENTDECKEN SIE DAS READIRIS INTERFACE

Das Readiris Anwendungsfenster enthält nicht nur **Befehlsmenüs**, sondern auch zwei Funktionsleisten, die Ihnen den schnellen Zugriff zu den meisten Befehlen ermöglichen. Anfänglich sind die Menüs m.B.a. der Vorschau ausgeblendet: Solange noch keine Bilddateien geöffnet wurden, bleiben sie ausgeblendet.



Das gleiche gilt für die **Bildfunktionsleiste**, die sich rechts in Ihrem Anwendungsfenster befindet. Die Bildfunktionsleiste enthält alle für die Bildvorschau notwendigen Befehle. Die **Hauptfunktionsleiste** links ermöglicht Ihnen den schnellen Zugriff zu fast allen Befehle.

Damit Sie wissen, welcher Befehl mit welcher Funktionstaste übereinstimmt, brauchen Sie nur den Mausschalter kurz über die Taste zu bewegen: Das **QuickInfo** teilt Ihnen die Funktion der Taste mit.



Schließlich gibt es das **Fenster**, in dem die gescannten Bilder angezeigt werden. Zur Erkennung der Bilddateien können diese in das Anzeigefenster gezogen werden („Drag & Drop“).

BEGINNEN MIT EINER ERSTEN EINFÜHRUNG

Die beste und einfachste Art, sich mit Readiris anzuvertrauen ist es zu benutzen. Zu Ihrer Software gehört eine Zahl von **vorgescannten Bildern**, die Ihnen den Start auch ohne Scanner ermöglichen.

Die Funktion „Quelle“ auf der Funktionsleiste bestimmt, ob Sie den Scanner oder das vorgescannte Bild als Bildquelle benutzen werden. Mit Readiris können Sie TIFF Bilder (nicht-komprimiert, Packbits, Gruppe 3 und Gruppe 4 komprimiert), Paintbrush (PCX) Bilder und Windows Bitmaps (BMP) öffnen. Es werden nur Schwarzweißbilder akzeptiert. Diese Möglichkeit ist sehr praktisch, wenn Sie Ihre **Faxe** zur weiteren Bearbeitung in Textdateien konvertieren möchten.

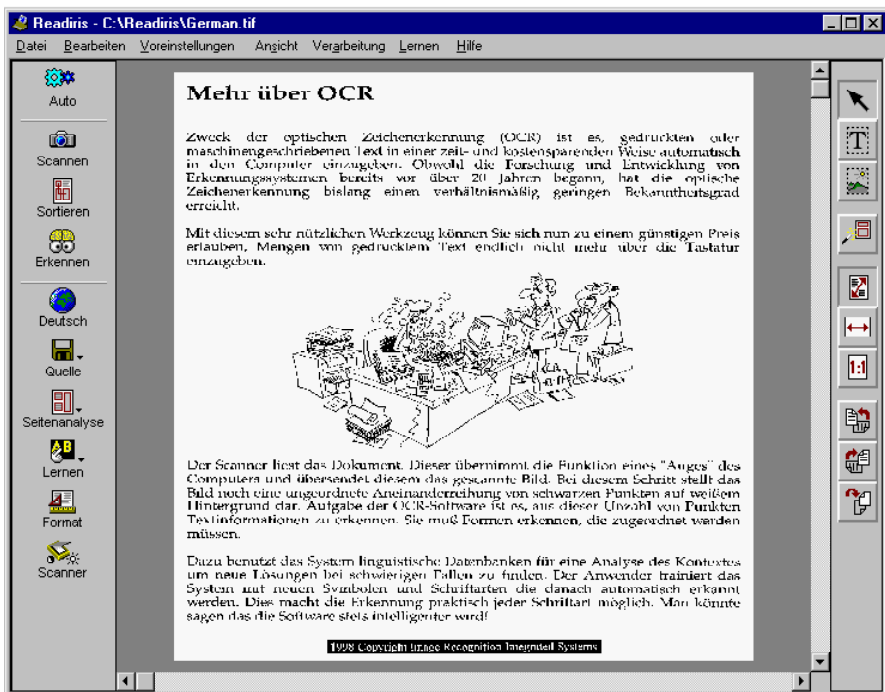
Bevor Sie ein vorgescanntes Bild als Bildquelle öffnen, selektieren Sie die Diskette und nicht den Scanner mit der Schaltfläche „Quelle“ auf der Hauptfunktionsleiste.



Klicken Sie anschließend die Funktion „Scannen“ an. Um das vorgescannte Bild direkt zu öffnen, können Sie auch den Befehl „Öffnen“ im „Datei“ Menü selektieren. Diese Funktion arbeitet sogar wenn Ihr Scanner als Bildquelle läuft.



Selektieren Sie jetzt eine Bilddatei. Wählen Sie die Datei GERMAN.TIF in das Readiris Verzeichnis. Das Bild wird dann von der Festplatte gelesen und nach und nach auf dem Bildfenster angezeigt.



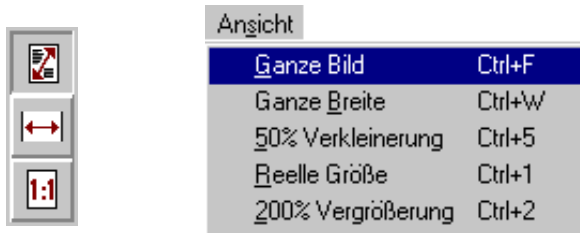
Eine dritte Möglichkeit ist, vorgescannte Bilder mit Hilfe der „Drag & Drop“-Funktion zu öffnen. Ziehen Sie ein Bild vom Windows Explorer zum Readiris Bildfenster. Die Bilder werden sofort geöffnet.

Die Bildfunktionsleiste rechts im Readiris Anwendungsfenster enthält alle notwendigen Befehle für die Bildvorschau: Funktionstasten um die zu bearbeitenden Blöcke zu bestimmen, um das Bild zu drehen, zum Ein- und Auszoomen, usw.

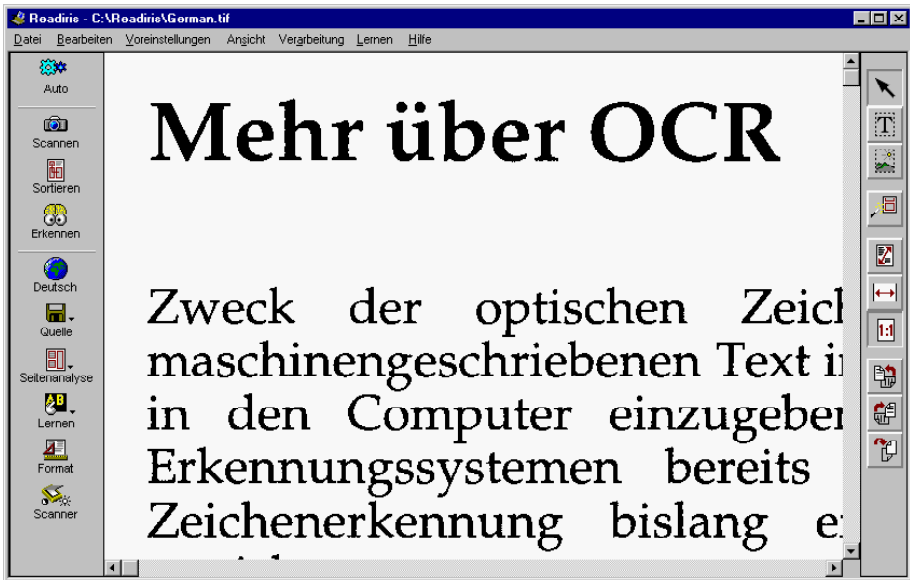
AUF BILDER ZOOMEN

Readiris verfügt über mehrere Befehle um das gescannte Bild zu **zoomen**, z.B. um die Scanqualität zu überprüfen.

Auf der Bildfunktionsleiste befinden sich Funktionstasten, mit denen Sie das gescannte Bild auf die wahre Größe bzw. auf Seitenbreite oder auf Bildschirmgröße einzoomen können. Das Menü „Anzeigen“ enthält noch zwei zusätzliche Zoomstufen: Sie können das Bild auf 50% oder 200% der Originalgröße zoomen. Wenn das Bild auf seine tatsächlichen Größe gezoomt wird, stimmt ein Bildschirmpixel überein mit einem Bildpixel. (Es sind Tastenkombinationen für alle Zoomstufen vorhanden!)



Schließlich können Sie durch Anklicken und Festhalten der *rechten* Taste, die Maus über den zu vergrößernden Teil des gescannten Bildes ziehen. Mit einem zweiten Klick der rechten Maustaste wird wieder ausgezoomt.



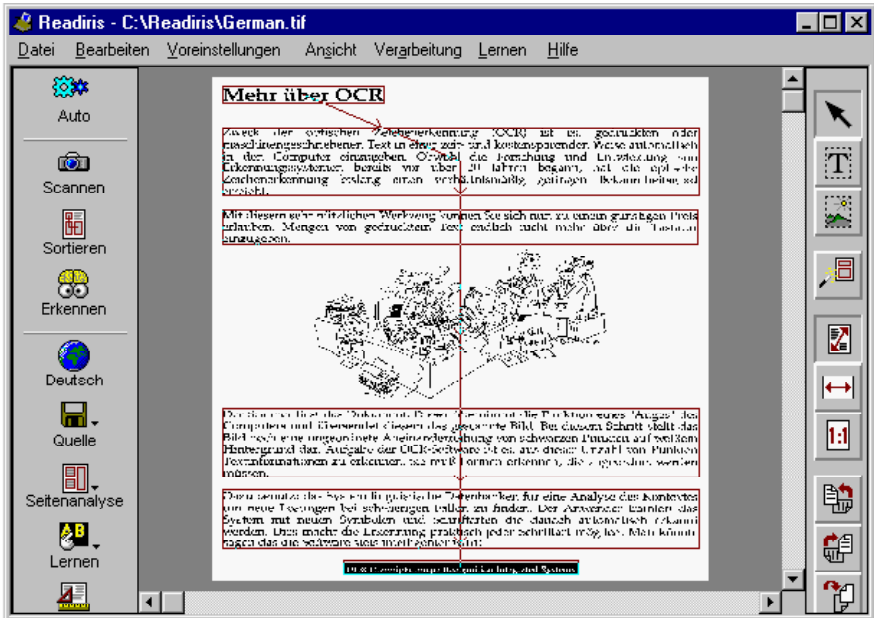
EIN, ZERLEGUNG EINES GESCANNTEN BILDES

Nachdem das Bild gescannt wurde, muß bestimmt werden, welche Teile als Text verarbeitet werden müssen. Hierzu werden Rahmen, sogenannte „Fenster“, um die ausgewählten Bereiche gezogen.

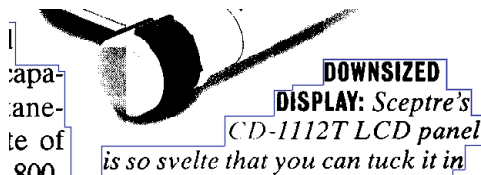
Wenn die Option „Seitenanalyse“ auf der Hauptfunktionsleiste aktiviert wurde, wird diese Handlung automatisch von Readiris ausgeführt.



Die automatische Seitenanalyse ist besonders nützlich, wenn **Texte in Spaltenform** und Dokumente mit einem komplizierten Seitenlayout eingelesen werden müssen.



Die Seitenanalyse ist schnell, toleriert Schrägstellung und arbeitet sehr präzise: „Unregelmäßige“ Schatten werden eingegrenzt und grafische Bereiche mit Fotos, Illustrationen usw. werden ignoriert.



Die Seitenanalyse ermittelt jedoch Zonen mit **weißem Text auf schwarzem Hintergrund**. Solche Eingaben sind kein Problem bei der Erkennung: Weil die Vorschau das gescannte Dokument originalgetreu auf dem Bildschirm anzeigt, ist



Readiris in der Lage das Bild bei Bedarf zu „invertieren“, damit solche Textblöcke erkannt werden. (Ebenfalls können Sie mit Ihrem Scanner ganze Negativbilder erzeugen um weißer Text auf einem schwarzen Hintergrund zu verarbeiten. Sehe weiter.)

Die Zahl der Fenster wird zu jeder Zeit in den QuickInfo's der Textfenster und Grafikfenster-Symbolen angezeigt.



Readiris kann nicht nur Textblöcke finden, es *sortiert* sie auch: die Zonen werden von oben nach unten sortiert und von links nach rechts so daß Dokumenten in Spaltenform richtig verarbeitet werden können. Sie können die Sortierreihenfolge natürlich auch selbst bestimmen - hierzu klicken Sie die Taste „Sortieren“ in der Hauptfunktionsleiste an. Danach klicken Sie das Fenster an. Fenster die *nicht* angeklickt werden, werden nicht erkannt.

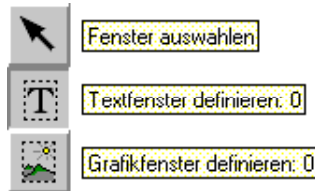


Die Seitenanalyse wird als Standardeinstellung eingestellt. Wenn die Seitenanalyse inaktiviert wurde weil Sie z.B. versehentlich einige Fenster gelöscht haben, können Sie wieder aktiviert werden, indem Sie die Taste „Seite Analysieren“ in der Bildfunktionsleiste anklicken.



ZWEI, MANUELLE FENSTERBESTIMMUNG EINES BILDES

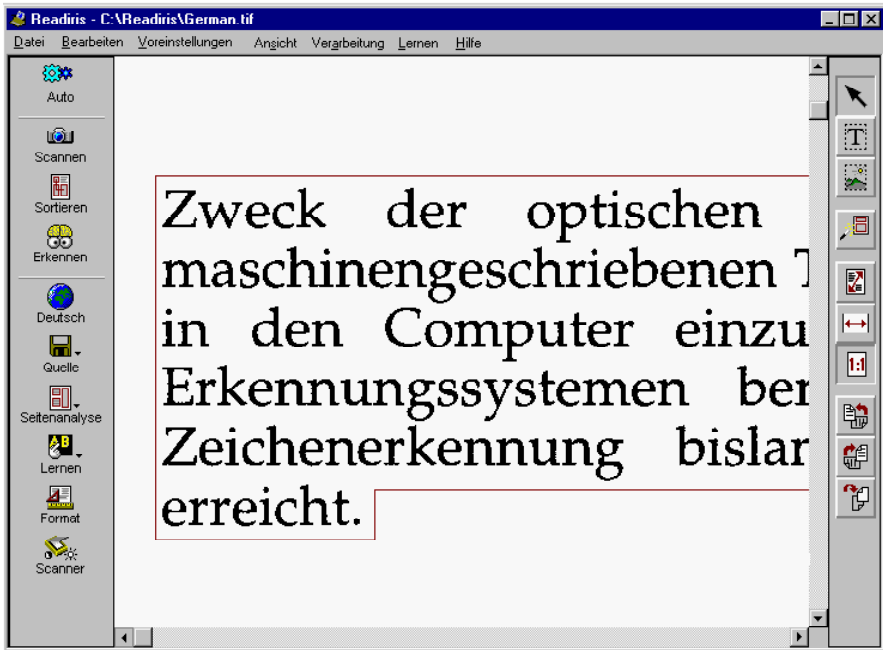
Durch die Seitenanalyse werden die Fenster automatisch bestimmt. Alternativ dazu kann die Fensterbestimmung auch manuell durchgeführt werden mit Hilfe der **Fenstersymbole** von Readiris.



Um ein rechteckiges Fenster zu **erstellen**, wählen Sie das „Textfenster“-Symbol in der Bildfunktionsleiste, bewegen Sie den Cursor in die rechte Ecke oben am Fenster und ziehen Sie das Fenster durch bewegen des Mauszeigers in die untere rechte Ecke und klicken Sie nochmals. (Seiten von weniger als 1 mm. sind nicht erlaubt Sie würden ohnehin kein Zeichen enthalten.)

Die Fenster werden automatisch sortiert: Pfeile zeigen die Sortierreihenfolge an.

Sie können auch „unregelmäßige“ Textblöcke umrahmen, indem Sie **polygonale Fenster** erstellen. Sie können polygonale Fenster erstellen durch rechteckige Fenster zusammenzufügen. Wenn zwei Fenster (derselben Art) einander berühren, macht Readiris automatisch ein polygonales Fenster daraus, als ob Sie ein Haus bauen in dem Sie ein Zimmer nach dem anderen hinzufügen...



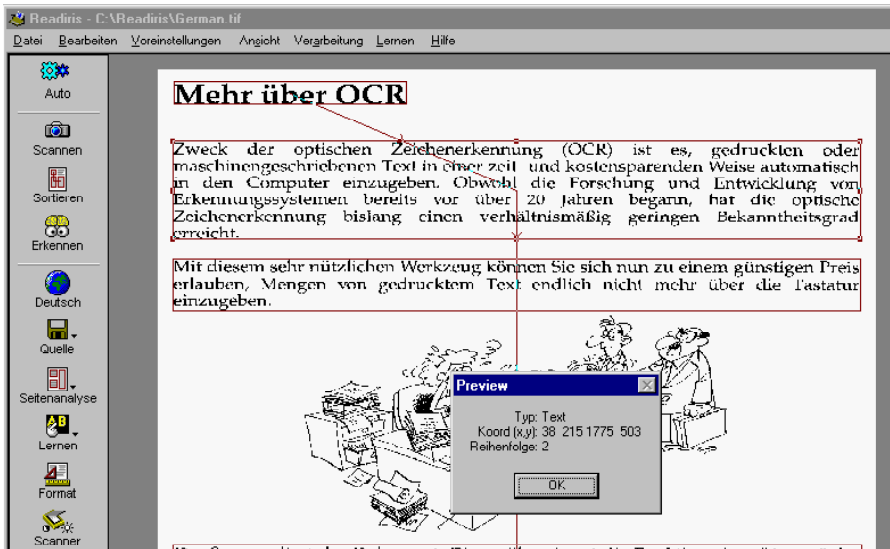
Um Fenster zu verändern, zu verschieben und zu löschen müssen sie zuerst **bestimmt** werden. Selektieren Sie hierzu „Fenster Selektieren“ (das „Pfeil“-Symbol) in der Bildfunktionsleiste und klicken Sie in das Fenster. Es erscheinen jetzt rechteckige Aktivpunkte in jeder Ecke und in der Mitte des Fensters.



Um die **Selektion** eines Fensters zu **annulieren**, klicken Sie mit der linken Maustaste außerhalb des Fensters. Um **zusätzliche Fenster** zu selektieren, halten Sie die Shift-Taste gedrückt während Sie die Fenster anklicken. Um ein **Fenster**

einschließlich des sich wiederum in diesem Fenster (des anderen Typs) befindlichen Fenster zu aktivieren, halten Sie die Ctrl-Taste gedrückt und klicken Sie dabei das Hauptfenster an.

Beachten Sie, daß die Sortierreihenfolge, den Typ und die Koordinaten angezeigt werden können indem die Alt-Taste gedrückt bleibt während Sie selektieren.



Soweit die Fensterauswahl. Um ein Textfenster zu **verändern**, selektieren Sie zuerst das Fenster, bewegen Sie anschließend den Mauszeiger über einen der Aktivpunkte und bewegen Sie die Maus um die Fenstergröße zu verändern.

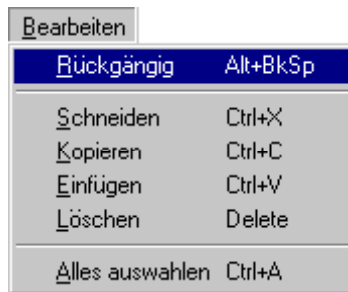
Um ein Fenster zu **verschieben**, selektieren Sie einfach das Fenster und ziehen Sie es zur gewünschten Position.

Um ein Fenster zu **löschen**, selektieren Sie das/die Fenster und selektieren Sie den Befehl „Schneiden“ oder „Löschen“ aus dem Menü „Bearbeiten“. Der



Befehl „Schneiden“ schneidet das Fenster aus und fügt es in ein interner Puffer ein. Der Befehl „Löschen“ löscht das Fenster.

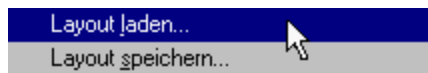
Es können also *alle* bekannten Befehle aus dem Menü „Bearbeiten“ bei den Fenstern angewendet werden: Sie können löschen, ausschneiden, kopieren und einfügen. Der Befehl „Rückgängig“ kann ebenfalls angewendet werden: Wenn Sie versehentlich einige Fenster gelöscht, bewegt, verändert, usw. haben, wird der Befehl rückgängig gemacht mit der Funktion „Rückgängig“.



Für alle Befehle können Tastenkombinationen benutzt werden! Beispielsweise: Damit alle bestehenden Fenster gelöscht werden, können Sie den Befehl „Alles Auswählen“ im Menü „Bearbeiten“ selektieren, oder die Tastenkombination Ctrl+A und den Befehl „Löschen“ aus dem Menü „Bearbeiten“ benutzen oder die Taste Delete. Jetzt kann das Layout wiederhergestellt werden. Um das vorherige Layout wiederherzustellen, selektieren Sie „Rückgängig“ oder benutzen Sie die Tastenkombination Alt+Backspace.

DREI, FENSTERVORLAGEN SPEICHERN

Die Fenstervorlagen können mit vorbestimmten **Fensterstrukturen** mit dem Befehl „Layout Speichern“ im Menü „Datei“ gespeichert werden und in den Speicher geladen werden mit dem Befehl „Layout Laden“.



Wenn Sie Dokumente mit einem ähnlichen Layout lesen müssen, z.B. ein 50-seitiges Protokoll, wo der Briefkopf und die Fußnote aus bestimmten Gründen unsichtbar bleiben sollen, dann kann eine einzige Vorlage benutzt werden um 50 Seiten in Zonen einzuteilen.

Wenn Sie eine solche Vorlage in den Speicher laden wird das Menü „Bearbeiten“ automatisch inaktiviert. Die Layoutvorlage bleibt solange aktiviert bis die Funktion Seitenanalyse auf der Hauptfunktionsleiste wieder aktiviert wird.

MIT READIRIS UM DIE WELT

Vorausgesetzt Sie haben die Fensterbestimmung korrekt durchgeführt, dann sind Sie jetzt fast soweit, daß Sie die Zeichenerkennung aktivieren können. Wir sagen „fast“, weil wir die Sprachenauswahl und die Dokumenteinstellungen noch nicht kontrolliert haben!

Die Spracheneinstellungen finden Sie auf der Hauptfunktionsleiste.



Um die Dokumentensprache zu verändern klicken Sie „Sprache“ an.

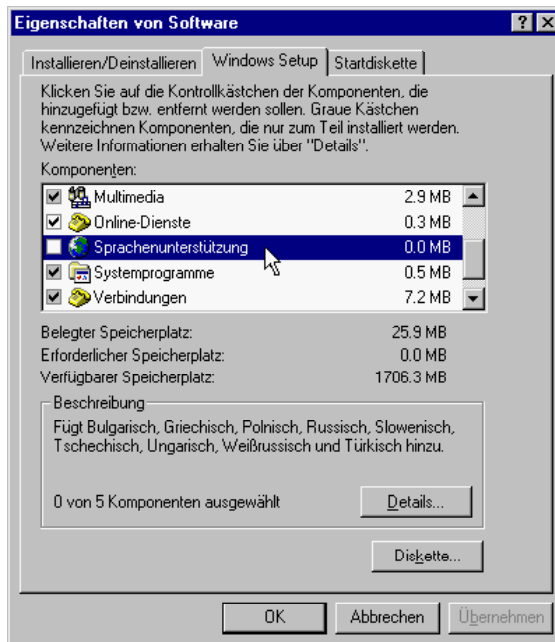


Sie können auch direkt eine Taste auf der Tastatur drücken: Wenn Griechisch oft gewählt wird und Sie möchten jetzt Indonesisch selektieren, drücken Sie dann die „G“-Taste auf Ihrer Tastatur um direkt Zugang zur griechischen Sprache zu bekommen. Für mehrere Sprachen mit dem gleichen Anfangsbuchstaben, drücken Sie den Buchstaben mehrmals bis Sie die gewünschte Option bekommen. Zum Beispiel: Readiris kann Polnisch und Portugiesisch lesen. Wenn Sie das „P“ einmal drücken, selektieren Sie Polnisch, wenn Sie das „P“ ein zweites Mal drücken, selektieren Sie Portugiesisch und beim dritten Mal sind Sie wieder bei Polnisch.

Readiris beschränkt sich nicht nur auf Deutsch: Es werden bis zu 52 **Sprachen** unterstützt! Readiris unterstützt alle amerikanischen und europäischen Sprachen sowie die zentraleuropäischen, griechischen, türkischen und kyrillischen („russischen“) und baltischen.

Merken Sie sich auch, daß die britischen und amerikanischen - oder sagen wir „internationalen“? - Varianten der englischen Sprache ebenfalls differenziert werden.

Merken Sie sich ebenfalls, daß Sie die angepaßte Windows-Konfiguration installieren müssen, damit zentraleuropäische, griechische, türkische und baltische Zeichen angezeigt werden. Vielleicht werden Sie das Windows-Modul „Sprachenunterstützung“ installieren müssen, bevor Ihr Windows-System in der Lage ist mit diesen Sprachen zu arbeiten. Um zu erfahren, ob diese Windows-Modul in Ihrem PC installiert wurde, selektieren Sie das Symbol „Software“ im „Systemsteuerung“. Abhängig von Ihrer Software-Version enthält die Readiris CD-ROM detaillierte Informationen über die Installation dieser Windows-Modul.



Es ist sehr wichtig, die dokumenteneigenen Sprache zu selektieren. Die Software weiß genau welche Zeichen sie zu erkennen hat. Dieses ist natürlich



abhängig von Ihrer Sprachenauswahl. Eine multilinguistische Datenbank hilft bei der Erkennung von „exotischen“ Zeichen so wie ç, ñ, ψ und ø.

Desweiteren benutzt die Software die **linguistischen Datenbanken** sehr ausgiebig, um die Ergebnisse zu kontrollieren. Nehmen wir einmal an das Wort „Präsident“ muß gelesen werden und auf dem Buchstaben „r“ befindet sich einen Tintenfleck, so daß das „r“ wie ein „f“ aussieht. Bei der Kontrolle im deutschen Wörterbuch, wird Readiris automatisch feststellen, daß das Wort „Präsident“ gelesen wurden und daß ein „f“ in diesem Kontext keinen Sinn ergeben würde. Diese „**autodidaktische**“ **Technik** ist natürlich sehr abhängig vom linguistischen Kontext.

Die linguistische Unterstützung kann sehr hilfreich sein bei der Erkennung von „**Zweifelsfällen**“, so wie das „O“, das als „()“ gelesen werden könnte. Ein anderes typisches Beispiel ist „l“ und ‘l’, die oft eine identische Form haben - denken Sie an Texte, die mit einer alten Schreibmaschine geschrieben wurden! Der linguistische Kontext hilft Ihnen zu erkennen, ob es sich um „l“ handelt oder um ‘l’.

Folgendes Beispiel illustriert die Schattenvariationen von ‘l’ und „l“. Die Schatten der ersten Linien sind deutlich abgegrenzt, die der zweiten Linien sind zweifelhaft abgegrenzt. Dieses Problem wird gelöst durch die Sprachwissenschaft.

193 1950s. 1hr
Well, Rossellini

BEI BEDARF VERÄNDERT READIRIS DIE SPRACHEN

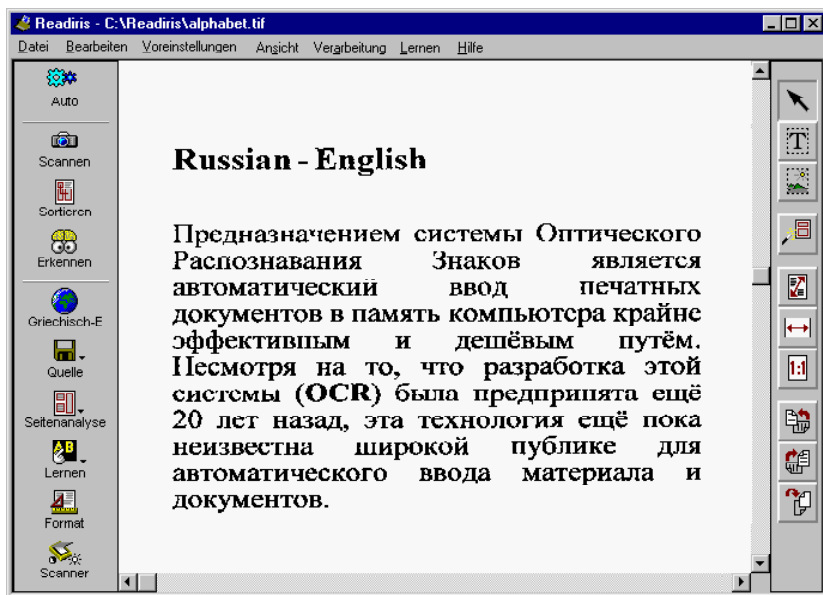
Readiris ist in der Lage, ohne Eingriff des Anwenders, in der Mitte eines Textes von einer Sprache zu einer anderen zu wechseln. Wenn „Westliche“ Worte auftauchen in griechischen und kyrillischen Dokumenten - viele Eigennamen, Markennamen usw. werden nicht umgesetzt aber mit „normalen“, lateinischen Zeichen geschrieben. Readiris kann sofort zum richtigen Alphabet wechseln. Mit

anderen Worten: Sie können ein „**gemischtes Alphabet**“ aktivieren. Die griechischen oder kyrillischen und westlichen Zeichen werden dann gemischt.

Vergewissern Sie sich, daß die „Griechisch-Englische“ oder die damit übereinstimmende kyrillische Einstellung gewählt wurde - z.B. „Weiß-Russisch-Englisch“. Mit anderen Worten: Selektieren Sie nicht nur „Griechisch“ oder „Weiß-Russisch“ als Dokumentensprache in der Hoffnung, daß die westlichen Symbole perfekt erscheinen werden!

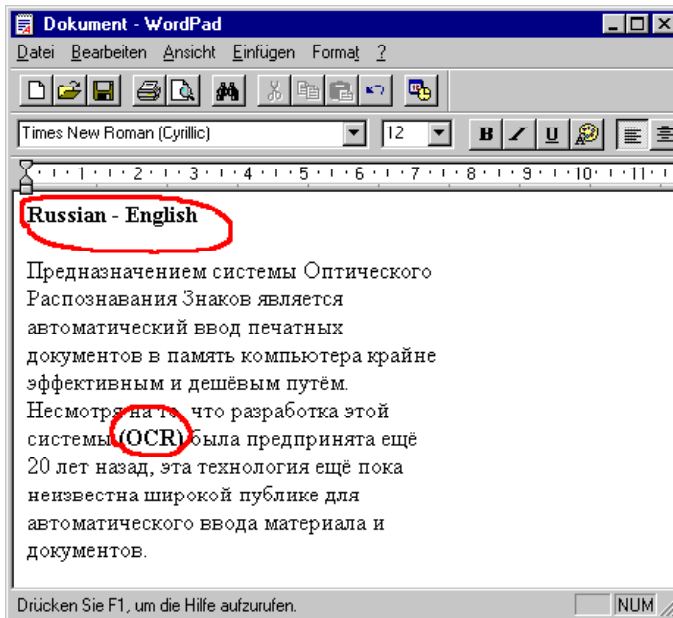


Hier sehen Sie ein Beispiel eines russischen Textes mit einigen englischen Wörtern - öffnen Sie die Bilddatei ALPHABET.TIF wenn Sie dieses selbst einmal ausprobieren möchten.





Das Endergebnis sieht aus wie folgt: Sie werden vielleicht doch noch eine kyrillisches **Zeichentyp** selektieren müssen, damit der russische Text richtig angezeigt wird.



Um **mehrere Sprachen** mit einander auf einer Seite zu **mischen**, wählen Sie die Sprache mit dem größten Zeichensatz. Wenn Sie ein englisches Dokument haben, das ebenfalls die französische Übersetzung enthält, selektieren Sie am besten Französisch als Sprache, damit auch die betonten Buchstaben ç, é und ù erkannt werden können.

BESTIMMEN DER DOKUMENTEIGENSCHAFTEN

Nachdem die Sprache eingestellt wurde, befassen wir uns mit den anderen Eigenschaften des Dokuments. Sie können die Erkennung noch besser gestalten, indem Sie einige Dokumentmerkmale spezifizieren: Typ der Schriftart und Zeichenlaufweite. Was bedeutet dies nun alles?

Zuerst sollten Sie angeben, ob „normal“ gedruckte oder **Matrix** gedruckte Dokumente erkannt werden sollen. Selektieren Sie den Befehl „Typ der Schriftart“ im Menü „Voreinstellungen“.

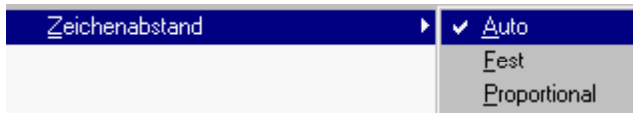


Dieser Parameter muß unbedingt richtig eingestellt werden: „9 Pin“ Dot Matrix Symbole sind aus vielen einzelnen Punkten zusammengesetzt. Hier ist ein hoch spezialisiertes Erkennungssystem erforderlich.

ape-descended life

„Letter Quality“ Dot Matrix-Ausdrucke, auch „25 Pin“ genannt, erfordern normale Einstellungen, so wie die **Druckeinstellungen** Druck, Maschinenschrift, Laserdruck und Inkjet-Druck. „Normal“ ist die Standardeinstellung.

Die **Laufweite** kann mit dem Befehl „Zeichenabstand“ im Menü „Voreinstellungen“ eingestellt werden.



Bei *festen* Schriftarten haben alle Buchstaben die gleiche Breite. Ein „i“ nimmt genau so viel Platz ein wie ein „w“, wie in diesem Satz vorgeführt. Wir denken hier an Dokumente, die mit einer Schreibmaschine geschrieben wurde.



Bei der *Proportionalschrift* ist die Laufweite abhängig von der Form des Zeichens. Zeichen wie *m* und *w* sind breiter und brauchen mehr Platz als die feinen Zeichen so wie *i* oder *j*. Alle Bücher, Zeitschriften und Zeitungen sind in Proportionalschrift gedruckt.

Courier, fester Zeichentyp
Letter Gothic, fester Zeichentyp

Century Schoolbook, proportioneller Zeichentyp
Arial, proportioneller Zeichentyp

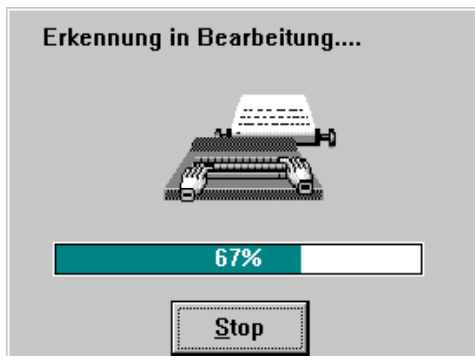
Hier ist es am einfachsten, wenn Sie diese Option in der Standardeinstellung „Auto“ lassen, damit Readiris den Zeichenabstand automatisch erkennen kann.

READIRIS WIRD IMMER INTELLIGENTER!

Wenn die Dokumentensprache und die Dokumenteigenschaften eingestellt sind, können Sie die Schaltfläche „Erkennen“ anklicken.



Das OCR-Programm wird angezeigt. Sie können die Schaltfläche „Stop“ anklicken, um die Texterkennung abzubrechen.



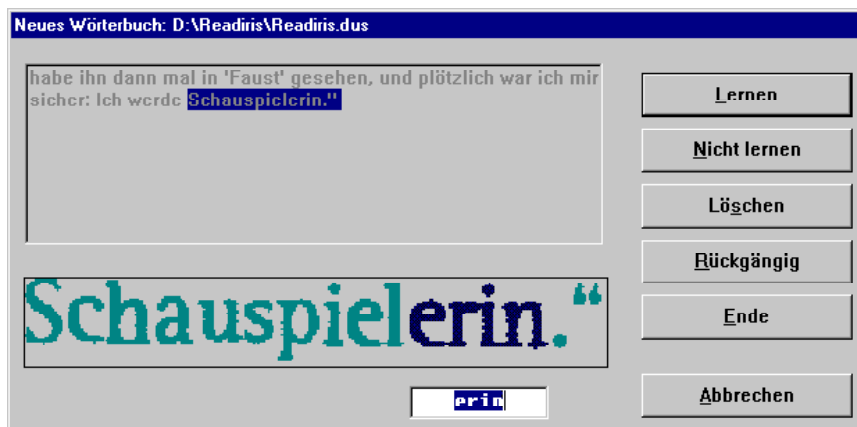
Am Ende jeder Erkennungsphase aktiviert Readiris die interaktive Lernphase. Die Lernfunktion muß aber aktiviert sein mittels der Schaltfläche „Lernen“ auf der Hauptfunktionsleiste. Interaktives Lernen ist standardmäßig aktiviert.



Zeichentraining ist eine Möglichkeit, um die Genauigkeit des Erkennungssystems zu verbessern. Das Training ist eine große Hilfe beim Lesen von Texten mit beschädigten Zeichen wie man in reellen Dokumenten findet, und beim Lesen von stilisierten Schriftarten, welche Readiris nicht optimal erkennen kann.

Das Lernsystem wird auch angewendet um das System mit **Sonderzeichen** vertraut zu machen. Mathematische und wissenschaftliche Zeichen und Dingbats können von Readiris nicht erkannt werden. Einige Beispiele: Readiris kann lernen, das π -Symbol als „pi“ oder das Dingbat „☎“, als „Tel.“ zu erkennen. Die Liste mit den erkannten Zeichen kann aber nicht mit den Symbolen „ π “ und „☎“, aufgestockt werde.)

Der erkannte Text wird schrittweise auf dem Bildschirm wiedergegeben. Das System stoppt, wenn es Zweifel hat ein Zeichen - oder eine Zeichenreihe wie bei „zusammenlaufenden“ Zeichen („Ligaturen“) - richtig zu erkennen. Sie werden immer in ihrem Kontext angezeigt, die Zweifelsfälle werden eingblendet. Nicht erkannte Zeichen werden durch ein Tilde (das “~“-Symbol) dargestellt.



Kontrollieren Sie zuerst, ob das richtige Zeichenwörterbuch und den Wörterbuch-Modus aktiviert wurden. Beide werden immer im Fenster „Lernen“ angezeigt. Wenn nicht, klicken Sie die Schaltfläche „Abbrechen“ an - Das Dokument wird jetzt wieder angezeigt mit dem Fensterlayout. Aktivieren Sie nun das richtige Zeichenwörterbuch oder den richtigen Wörterbuch-Modus und starten Sie OCR neu. (Die Funktion „Zeichenwörterbuch“ wird bald besprochen werden.)

Wenn erforderlich, geben Sie über die Tastatur das richtige Zeichen oder die richtige Zeichengruppe ein und selektieren Sie einen der nachstehenden Befehle.

Lernen

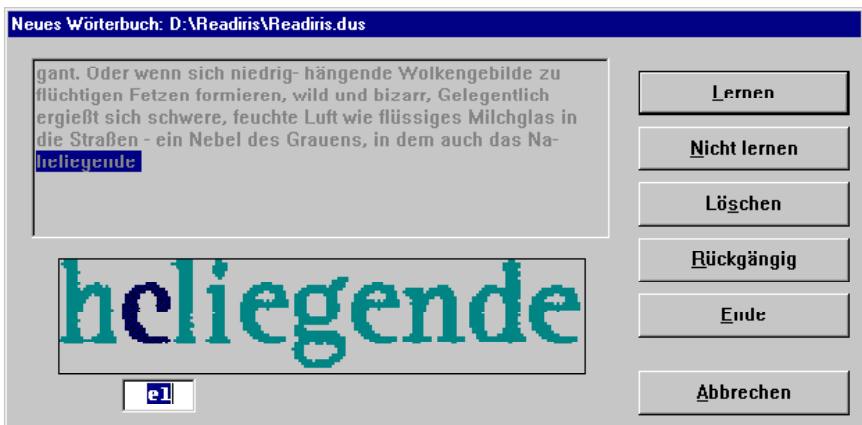
Sie sind mit der vorgeschlagenen Lösung einverstanden oder Sie haben die Lösung korrigiert. Readiris speichert das Zeichen dann im Zeichenwörterbuch als „sicher“ ab. Dieses Zeichen wird dann künftig nie wieder angezeigt werden. Eine Bestätigung Ihrerseits ist nicht mehr erforderlich.

Im oben erwähnten Beispiel hält das System bei einige zusammenlaufenden Zeichen an. Wir klicken „Lernen“, um die Zeichen, die nicht mit andere kunnen verwechselt werden, zu akzeptieren.

Nicht Lernen

Sie sind mit der vorgeschlagenen Lösung einverstanden, oder Sie haben die Lösung korrigiert. Der Unterschied mit der Funktion „Lernen“ liegt darin, daß das erlernte Symbol als „unsicher“ im Zeichen-Wörterbuch abgespeichert wird. Künftig wird das System die gelernte Lösung vorschlagen, sie muß aber immer wieder bestätigt werden.

Diese Schaltfläche wird bei Symbolen die leicht zu verwechseln sind benutzt: Ein schlecht geformtes „e“ kann als „c“ gelesen werden, ein schlechtes „t“ ähnelt einem „r“ usw.



Das oben gezeigte „e“ ist stark beschädigt - eigentlich sieht es mehr aus wie das „c“-Symbol. Jetzt sollten die Schaltfläche „Nicht Lernen“ anklicken, damit es nicht mit dem „c“-Symbol verwechselt wird.

Löschen

Das angezeigte Zeichen verschwindet von der Ausgabedatei. Diese Schaltfläche wird benutzt, um bestimmte „Verunreinigungen“ so wie z.B. Flecke,



Kaffeeflecke, usw. zu eliminieren. Diese könnten als Punkte, Kommas usw. erkannt werden. Es läßt auch jedes ungewollte Symbol verschwinden.

Rückgängig

Widerruft die zuvor erteilten Befehle. Sie können die letzten neun Befehle rückgängig machen.

Ende

Der Lernprozeß wird abgebrochen, aber das OCR-System läuft weiter im Auto-Modus. Alle Entscheidungen werden akzeptiert ohne Bestätigung des Anwenders.

Klicken Sie diese Schaltfläche an wenn Sie feststellen, daß die Erkennung sehr akkurat durchgeführt wurde und Korrekturlesen nicht erforderlich ist.

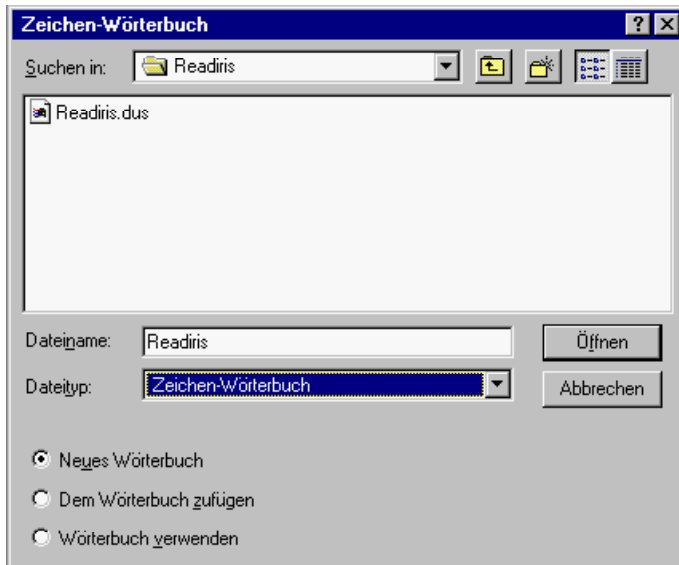
Verwechseln Sie „Ende“ nicht mit „Abbrechen“. Bei „Abbrechen“ gibt es keine Textausgabe und müssen Sie neu anfangen. Bei „Ende“ wurde der Text erstellt, aber noch nicht richtig korrigiert!

DIE ROLLE DER ZEICHENWÖRTERBÜCHER

Die Ergebnisse jeder Lernphase werden übergangsweise im Speicher festgehalten. Sie sollten jedoch in sogenannten Wörterbuch-Dateien gespeichert werden, damit sie auch später noch verwendet werden können.

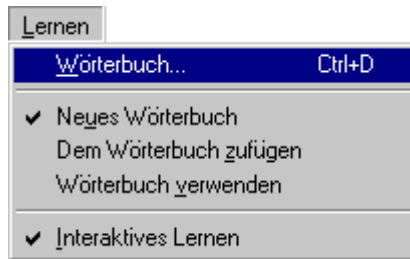
Wenn Sie ähnliche Dokumente erkennen möchten und die Intelligenz Zeichenwörterbücher nutzen wollen, ist es ratsam diese zu speichern. So wird Readiris immer intelligenter!

Wie funktioniert das? Die Funktion der **Zeichenwörterbücher** wird vom Menü „Lernen“ kontrolliert: Es muß ein Wörterbuch mit dem Befehl „Wörterbuch“ gewählt werden und der Handlungsmodus muß bestimmt werden.



Zeichen-Wörterbücher haben eine Kapazität von 500 Zeichen. Für Sonderanwendungen empfehlen wir Ihnen separate Wörterbücher anzulegen, z.B. nach Typ oder nach Dokument. Es ist von großer Bedeutung die Wörterbücher deutlich zu benennen, z.B. BERICHT.DUS, PALATINO.DUS usw. Wörterbücher haben die Standarderweiterung *.DUS. Wenn ein Wörterbuch voll ist, ist es zwecklos noch weiter zu lesen. Die Ergebnisse können dann weder im Speicher, noch im Wörterbuch festgehalten werden.

Der Wörterbuch-Modus kann im Befehl „Wörterbuch“ oder direkt im Menü „Lernen“ eingestellt werden. Es gibt drei Wörterbuch-Modi: neu, zufügen und verwenden.



Wenn Sie „Neues Wörterbuch“ selektieren, geben Sie an; daß die Ergebnisse in einem neuen Wörterbuch gespeichert werden müssen. (Wenn Sie ein bereits existierendes Wörterbuch selektieren, wird der Inhalt gelöscht.)

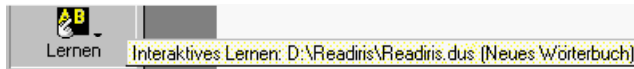
Der Modus „Zufügen“ gibt an, daß die Ergebnisse in einem bereits existierenden Wörterbuch gespeichert werden: Das Erkennungssystem nutzt bereits die zusätzliche Intelligenz des Wörterbücher und Sie fügen neue Zeichen hinzu. Mit anderen Worten: Mit dieser Option wird ein Zeichenwörterbuch schrittweise aufgebaut.

(Wenn Sie für ein neues Wörterbuch einen neuen Dateinamen eingeben und den „Zufügen-Modus“ aktivieren, wird ein leeres Zeichenwörterbuch erstellt, das Sie ergänzen können.)

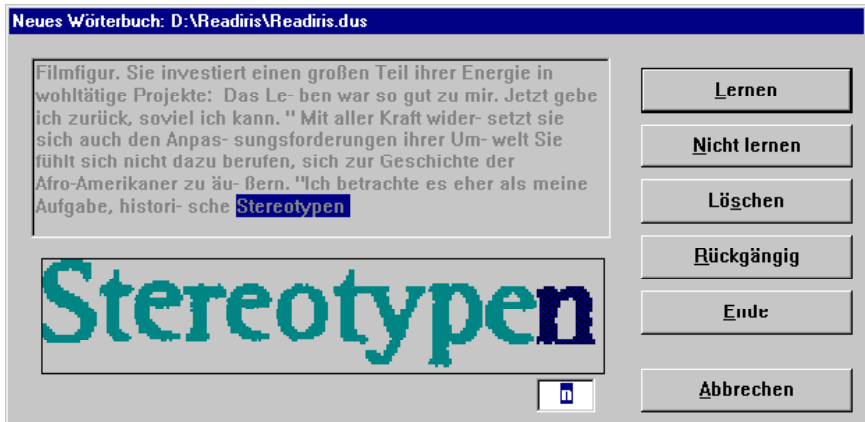
In der letzten Option „Wörterbuch Verwenden“, funktioniert das Wörterbuch nur im Lesen-Modus („Read Only“). Es werden dem Wörterbuch *keine* neuen Zeichen zugefügt. Sie können es aber verwenden.

Wenn Sie nur eine Seite lesen müssen, selektieren Sie den Modus Neu. Bei mehreren Seiten der gleichen Schriftart selektieren Sie den Modus Neu für die erste Seite, den Modus Hinzufügen für einige darauffolgende Seiten und den Modus Verwenden für den Rest des Dokumentes. Wenn Sie ein mehrseitiges Dokument lesen müssen wird beim Lesen der folgenden Seiten das Benutzerwörterbuch in den Zufügen-Modus gebracht.

Das QuickInfo der Schaltfläche „Lernen“ zeigt Ihnen an, welches Wörterbuch aktiv ist und in welchem Modus es arbeitet.

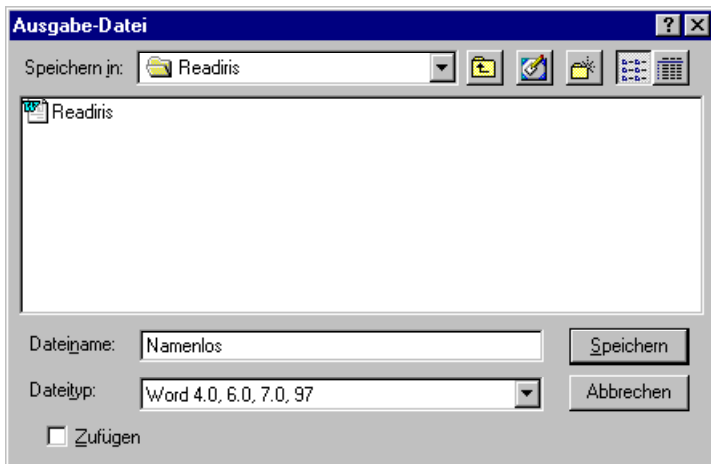


Wenn Sie im interaktiven Lernbereich sind, werden die Wörterbücher und ihre Modi im Fenstertitel angezeigt. Wenn es sich um einen Fehler handelt, sollten Sie „Abbrechen“ klicken.

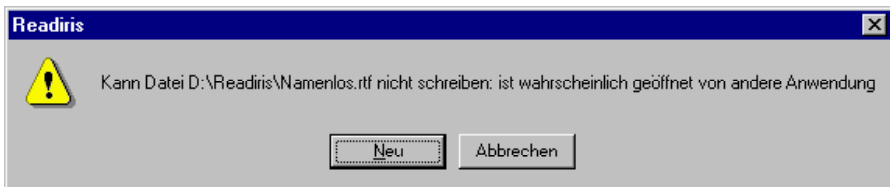


WOHIN MIT DEM TEXT?

Die interaktive Lernphase enthält auch die Zeichenerkennung. Sie werden gebeten, den erkannten Text zu speichern.



Sie können entweder eine neue Datei erstellen und eine existierende Textdatei hinzufügen - hierzu gibt es die Funktion „Zufügen“. Sie könnten z.B. zwei Seiten eines gescannten Textes einem Dokument, das Sie selbst geschrieben haben, zufügen. Sie sollten sich vergewissern, daß diese Datei nicht geöffnet ist, sonst können Sie nichts zufügen!

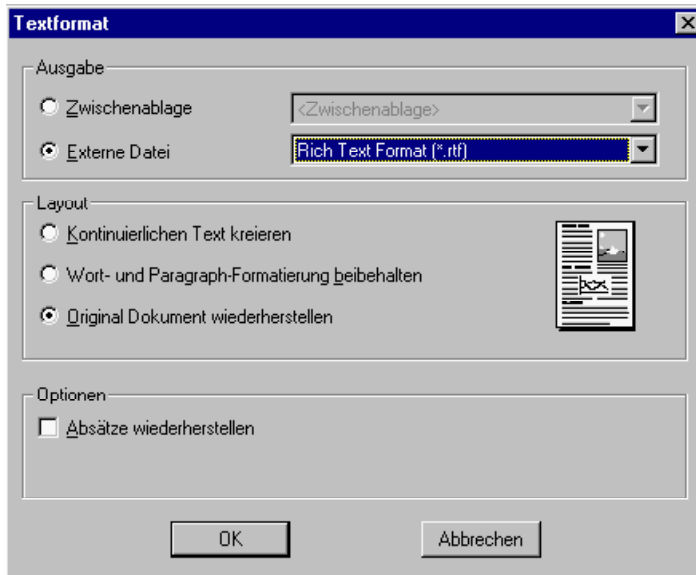


Wenn Sie einen Text speichern, bekommt die Ausgabedatei nicht nur einen Dateinamen, das Textformat wird auch festgelegt und, das Aussehen des Textes wird bestimmt vor die Erkennung.

Das **Textformat** wird über die Funktion „Format“ (oder mit dem Befehl „Textformat“ im Menü „Voreinstellungen“) aktiviert *vor* die Erkennung.



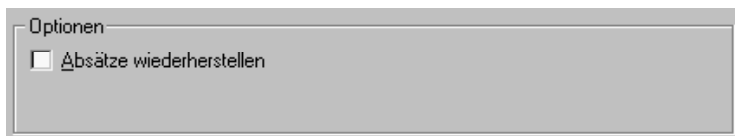
Mit diesem Befehl selektieren Sie ein Format für die Ausgabedatei. Sie können den gescannten Text aber auch sofort in die **Zwischenablage** kopieren. Es ist also nicht unbedingt erforderlich, den erkannten Text in einer neuen Datei zu speichern.



DIE AUSSICHT DER TEXTAUSGABE

Mit einigen Optionen können Sie das Aussehen der Textausgabe beeinflussen.

Der **Textfluß** des Ausgabedokumentes kann direkt durch die Option „Absätze Wiederherstellen“ unter der Funktion „Format“ beeinflusst werden.



Mit dieser Option kann Readiris automatisch die **Absätze ermitteln**: Readiris setzt nur einen festen Zeilenumbruch am Ende eines Absatzes und nicht am Ende einer Zeile. Die Funktion „Absätze Wiederherstellen“ gehört zu den Standardeinstellungen.

Ein Beispiel zur Verdeutlichung. Die ersten drei Zeilen in einem Absatz lauten: „Der neue Präsi-“, „dent stand auf dem Balkon.“ und „Seine Frau kam dazu.“. Die Absatzkontrolle schreibt den Satz wie folgt: „Der neue **Präsident** stand auf dem Balkon. **Seine** Frau kam dazu.“ Die getrennten Silben des Wortes „Präsident“ wurden wieder zusammengefügt. Außerdem wurde einen Leerraum am Ende des Satzes zugefügt. Somit wurde das Aussehen des Textes beeinflusst und ein natürlicher Textfluß erstellt.

Ansonsten hätte Readiris einen feste Zeilenwechselmarke am Ende der Zeile und nicht am Ende des Absatzes gesetzt.

(Die Funktion „Format“ enthält noch mehr Optionen. Mehr darüber später.)

MEHRERE SEITEN ERKENNEN

Wenn Sie „Speichern“ anklicken um den Text zu speichern, werden Sie gefragt, ob Sie noch mehr Seiten scannen möchten. (Wenn Sie den gescannten Text in der Zwischenablage gespeichert haben werden Sie natürlich nicht aufgefordert den Text zu speichern.)



An dieser Stelle wird die Zahl von erkannten Zeichen angezeigt. Selektieren Sie jetzt „Beenden“, um diesen ersten Vorgang zu beenden. Das gescannte Bild erscheint mit den definierten Bereichen wieder auf dem Bildschirm. Es kann jetzt weiter bearbeitet werden und wird solange auf dem Bildschirm angezeigt, bis Sie eine andere Seite lesen möchten.

Sie können jetzt den gescannten Text mit Hilfe Ihres Textverarbeitungsprogramm oder Ihres Textverarbeiters öffnen und ihn in Ihr Desktop-Publishingprogramm oder jedes andere Text-Programm übernehmen. Sie haben jetzt ein Papierdokument in eine Textdatei konvertiert für die weitere Bearbeitung mit einem Text-Programm. Einen Vorgang der 10 bis 25 Mal schneller ist, als das manuelle Abschreiben!

Die Hilfe bei der Verarbeitung einer einzelnen Seite endet hier. Aber wie wird ein mehrseitiger Text bearbeitet?

Eigentlich ist es ganz einfach: Klicken Sie „Nächste Seite“ an und die nächste Seite wird sofort gescannt, wenn Ihr Scanner als Bildquelle eingestellt wurde. Die Ausgabedatei und das Benutzerwörterbuch befinden sich automatisch in Zufügen-Modus, so daß der Text zugefügt werden kann und das interaktive Lernen automatisch fortgesetzt wird.

Sie können jetzt mit der Erkennung mehrerer Seiten fortfahren: Der neue Text wird jedesmal nach dem Speichern der Textdatei zugefügt.

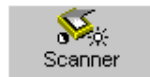


SCANNEREINSTELLUNG

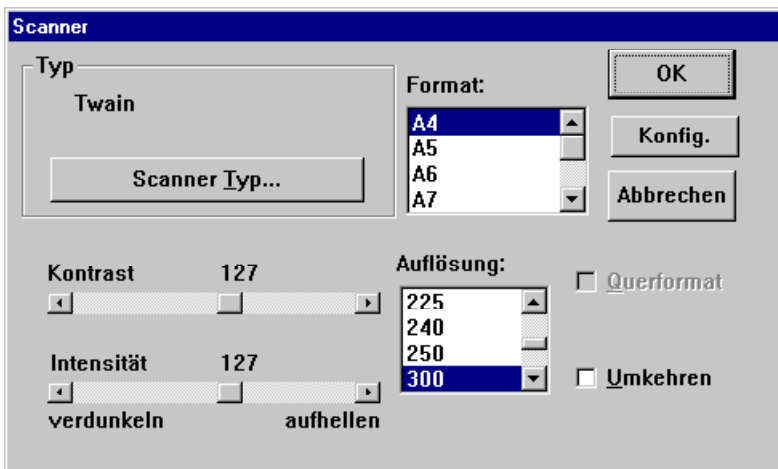
Jetzt werden wir den Scanner einstellen. Wichtige Voraussetzung ist, daß der Scanner und die erforderlichen Treiber richtig installiert wurden.

Wenn Ihr Scanner zusammen mit Ihrer Readiris Software lizenziert wurde brauchen Sie hier wahrscheinlich nichts mehr zu machen. Ihr Scanner ist dann oft bereits unter Readiris installiert.

Klicken Sie auf das Symbol „Scanner“ in der Hauptfunktionsleiste.

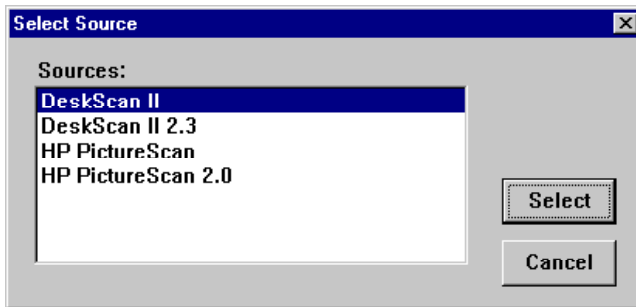


Klicken Sie auf das Symbol „Scannertyp“ um das **Scannermodell** zu bestimmen.



Wenn Sie „<Image>“ (Bild) als „Scanner“ selektieren, dann dienen vorgescannte Bilder immer als Bildquelle. In diesem Fall brauchen Sie das Symbol „Quelle“ nicht mehr zu aktivieren, um die Festplatte als Bildquelle zu aktivieren.

Das Symbol „Konfig.“ ist nur aktiviert wenn der Scanner es erlaubt. Es ermöglicht den Zugriff zu einigen fortschrittlichen Scanparametern: Bei Twain Scannern können Sie die Twain-Quelle selektieren, indem Sie auf das Symbol „Konfig.“ klicken. (Sie können ebenfalls den Befehl „Quelle Auswählen“ im Menü „Datei“ anklicken.)



Ihr Scanner kann ebenfalls mit einer **Scan-Plattform**, wie z.B. Visioneer PaperPort oder HP Document Assistant arbeiten. Abhängig von Ihrer Software-Version gibt es hierzu spezielle Handbücher auf der Readiris CD-ROM.

Nachdem der Scanner gewählt wurde, können Sie im gleichen Fenster die Scanauflösung, das Seitenformat und die Seitenorientierung, die Helligkeit und den Kontrast bestimmen, und andeuten wenn Sie den automatischen Dokumenteingebener Ihres Scanners verwenden möchten. Bei Scannern, die mit Twain kompatibel sind, werden diese Parameter oft innerhalb der Twain-Schnittstelle bestimmt.

Für die normale Schriftart wählen Sie die **Auflösung** von 300 dpi, für eine kleinere Schriftart (unter 10 Punkte) und bei einer schlechten Qualität des Dokumentes wählen Sie eine höhere Auflösung von 400 dpi.

Readiris ist in der Lage, Zeichen mit einer **Größe** von 6 bis 72 Punkten (0,21 bis 2,54 cm) zu lesen.



6 Punkte

72 Punkte

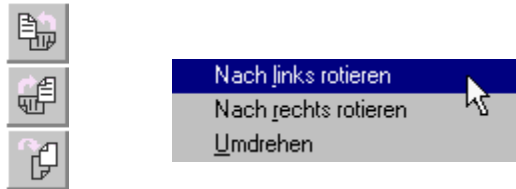
Readiris kann auch **Initialen** („Drop Letters“) erkennen. „Initialen“ sind große Buchstaben, die über mehreren Zeilen gehen - sie können natürlich nie größer sein als 72 Punkte!

Readiris erkennt
 „Initialen“ - das
 sind große
 Buchstaben die über
 mehreren Zeilen gehen.

Durch Aktivierung der Option „Querformat“ geben Sie an, daß die gewählte Seitenorientierung breit („Querformat“) ist anstatt schmal. Die Seitenorientierung kann nur auf reduzierten Seitenformaten angewendet werden: mit einem A4 Flachbettscanner können A5 Texte (halb so groß) gescannt werden im Hoch- und Querformat. Die A4-Seite kann jedoch nur in einer Richtung gescannt werden!



Um die Seitenorientierung zu korrigieren, können Sie die Drehsymbole, die sich auf der Bildsymbolleiste befinden benutzen. (Siehe auch Menü „Ansicht“.)



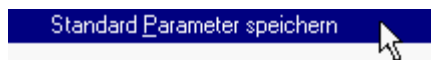
Es gibt drei Drehrichtungen: Nach links, nach rechts und **Drehung** um 180 %.

Stellen Sie die **Helligkeit** und wenn vorhanden, den **Kontrast** ein. Die Kontrasteinstellungen sind nur bei bestimmten Scannern vorhanden. Die ausgewählten Kontrastwerte werden numerisch vorgestellt.

Zuletzt gibt es die Option "Umkehren", die Ihnen die Möglichkeit bietet **Negativbilder** zu erzeugen. Aktivieren Sie diese Option zur Verarbeitung von ganzen Seiten mit weißem Text auf einem schwarzem Hintergrund.

STANDARDEINSTELLUNGEN SPEICHERN

Wenn alle Scanparameter richtig eingestellt wurden, klicken Sie auf den Befehl „Standard Parameter Speichern“ im Menü „Datei“. Die laufenden Einstellungen werden als **Standard Einstellungen** gespeichert. Bei künftigen Aufrufen arbeitet Readiris mit den Einstellungen weiter.

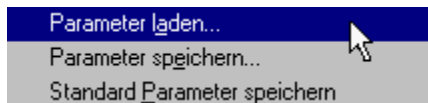


Die Standardeinstellungen enthalten mehr als nur die Scanparameter: Sie können auch bestimmen, ob Sie die interaktive Lernfunktion benutzen werden, welches Benutzerwörterbuch (und Modus) benutzt wird, welche Sprache und welche Schriftart - z.B. Normalschrift, Proportionalschrift - usw. aktiv sind. Kurz, *alle* operative **Einstellungen** werden in den Readiris Standardaufstellungen gespeichert.



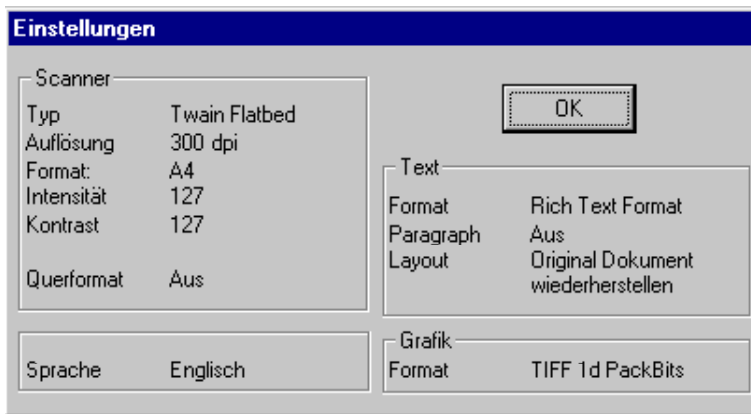
DAS SPEICHERN VON VORDEFINIERTEN EINSTELLUNGEN

Bei jedem Start werden die Standardeinstellungen geladen. Um zu vermeiden, daß Sie die eingestellten Parameter ständig anpassen müssen, können Sie auch unterschiedliche Einstellungen speichern. Die Befehle sind „Parameter Speichern“ und „Parameter Laden“. Sie befinden sich im Menü „Datei“.



Ein Beispiel: Wenn Sie regelmäßig deutsche Dokumente mit einem spezifischen Layout scannen müssen, ist es zu empfehlen, eine Ausgabedatei zu erstellen, die auf diesen speziellen Dokumententyp Anwendung findet. Sie müßten dann „Deutsch“ als Dokumentensprache selektieren. Damit Sie die Fenster nicht immer neu definieren müssen, ist es ratsam, eine Vorlage zu laden. Schalten Sie den Lernmodus aus aber aktivieren Sie ein Benutzerwörterbuch im „Verwenden“-Modus, weil die gleichen Schriftformen systematisch wieder angewendet werden usw.

Um eine Übersicht über die eingestellten Parameter zu bekommen, ist es unnötig die verschiedenen Menübefehle zu überlaufen, Sie brauchen nur den Befehl „Information“ im Menü „Datei“ zu aktivieren.



AUTOMATISCHE SEITENERKENNUNG

Nachdem der Scanner eingestellt wurde, können wir die Seitenerkennung aktivieren. Vergessen Sie nicht, den Scanner mit der Funktion „Quelle“ als Bildquelle zu selektieren.



Anstatt den Scanvorgang schrittweise auszuführen, wird die Seitenerkennung jetzt automatisch durchgeführt. Hierzu klicken Sie auf die Funktion „Auto“ anstatt auf die Funktion „Scannen“.



Automatisches OCR bedeutet, daß die Seite nacheinander gescannt wird, durch die Seitenanalyse oder durch eine Fenstervorlage definiert und erkannt wird ohne interaktives Lernen. Sie brauchen nur die Befehle zu initiieren, den Rest übernimmt Readiris.



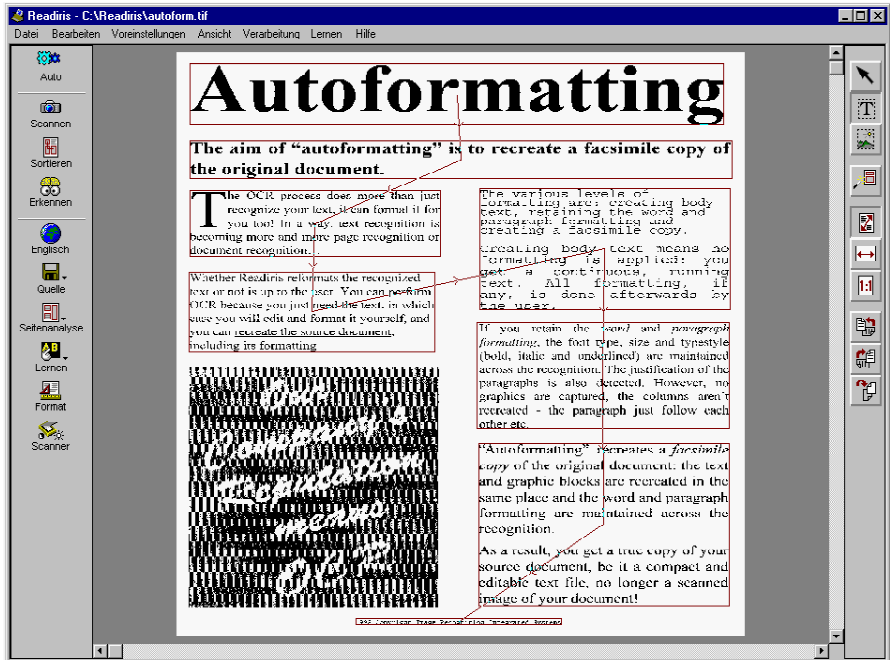
READIRIS ERSTELLT DAS LAYOUT

Automatisches Erkennen darf nicht mit automatischem Formatieren verwechselt werden. Der automatische Erkennungsmodus ermöglicht den automatischen Erkennungsprozeß. „Autoformatierung“ bedeutet, daß Readiris eine **Faxkopie** des gescannten Dokumentes erstellt. Die Wort-, Absatz-, und Seitenformatierung vom Originaldokument werden übernommen.

Ähnliche Schriftzeichen (serif und sans-serif, proportional und fest) werden wie im Originaldokument wiedererkannt. Die Schriftgröße und die Schriftauszeichnung (fett, kursiv und unterstrichen) werden ebenfalls wiedererkannt im Speicher. Der Randausgleich und die Einzüge (links und justiert) von jedem Textblock werden wiederhergestellt. Die Position der Spalten, Text- und Grafikblocks erscheinen wie im Originaldokument. Mit anderen Worten: Readiris ist in der Lage, eine originalgetreue Kopie Ihres Originaldokuments als eine kompakte Textdatei und nicht wie eine Bilddatei zu speichern.

All das bedeutet, daß die Fenstersortierung *nicht* arbeitet, wenn „Autoformatierung“ aktiviert wurde. Die bereits mit der Seitenanalyse durchgeführte Sortierung wird automatisch ignoriert.

Ein Beispiel zur Verdeutlichung. Öffnen Sie das Fenster AUTOFORM.TIF im Readirisverzeichnis.



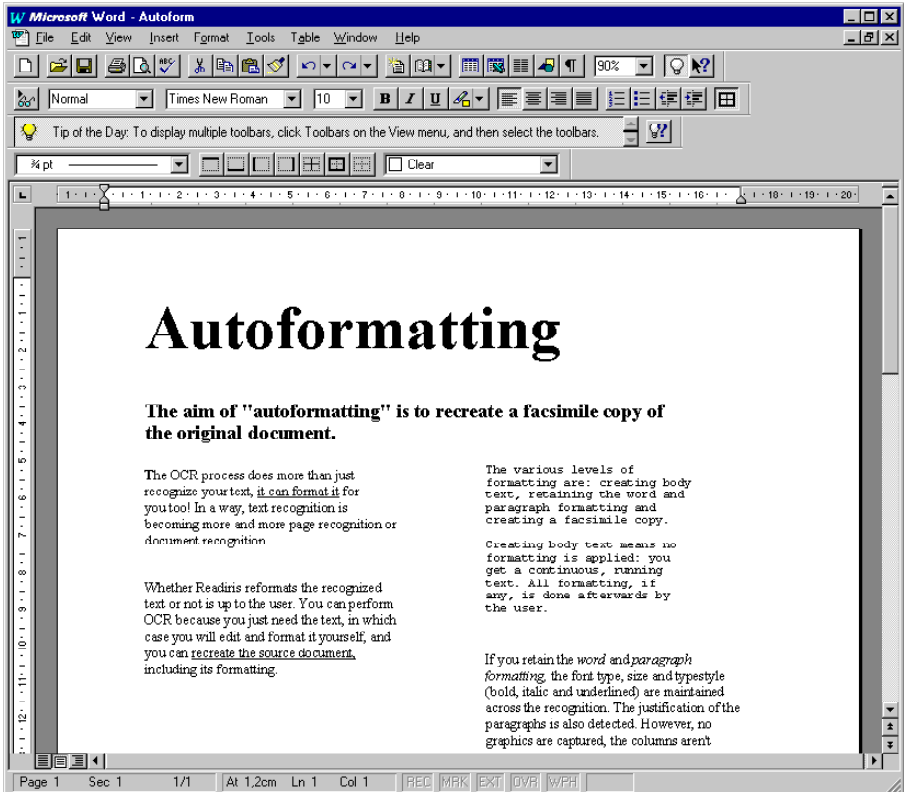
Klicken Sie auf das Symbol „Format“ auf der Hauptfunktionsleiste und selektieren Sie „RTF“ (Rich Text Format) als Ausgabeformat und „Original Dokument Wiederherstellen“ als Layoutoption. (Entscheiden Sie selbst, ob Sie die Option „Absätze Wiederherstellen“ wählen um den Zeilenumbruch zu aktivieren).



Die Wiederherstellung von Layouts ist nur im RTF-Format möglich. Andere, „begrenzte“ Formate so wie ASCII, ANSI, usw. unterstützen *keine* fortgesetzte Formatierung und sollen deshalb nicht im Autoformat-Modus gewählt werden.

Positiv ist, daß das RTF-Format ein sehr verbreitetes Textformat ist und mit jedem Textverarbeiter geöffnet werden kann. (Beachten Sie das WordPad ein beschränkt einsetzbarer Textverarbeiter ist; WordPad ist in der Lage die RTF-Dateien zu öffnen, die Formatierungen wie Block, Einzüge usw. werden ignoriert.)

Wenn der erkannte Text mit einem Textverarbeiter geöffnet wird, sieht er mit überhaupt kein Eingreifen des Verwenders so aus. Um das Dokument in der richtigen Ansicht zu sehen, aktivieren Sie den „WYSIWIG“-Modus, auch „Seitenlayout“ genannt.



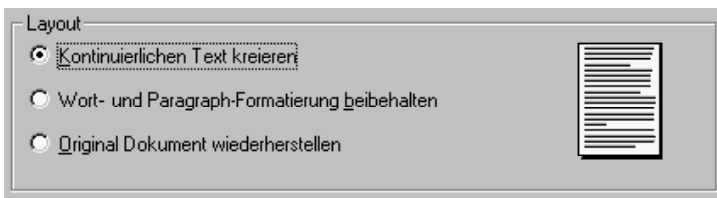
Also, Readiris kann Texte erkennen und formatieren. OCR ist nicht nur eine Texterkennung. OCR wird mehr und mehr eine **Seiten-** oder **Dokumentenerkennung!**



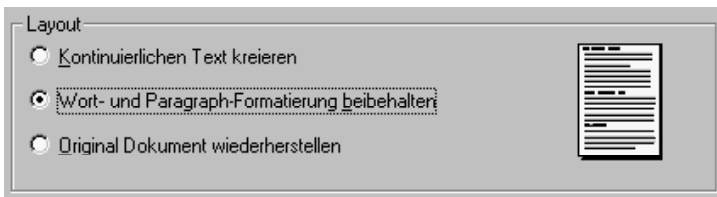
TEXTFORMATIERUNG, TEIL 2

Es gibt auch zwei andere Layoutoptionen: „Kontinuierlichen Text Kreieren“ und „Wort- und Paragraph-Formatierung Beibehalten“.

Sowie das Symbol rechts zeigt, bedeutet „Kontinuierlichen Text Kreieren“ daß ein nicht formatierter, **kontinuierlicheer Text** erstellt wird. Der Text wird ohne Formatierung gespeichert. Diese Option ist nur dann sinnvoll, wenn Sie nur den Text brauchen, aber nicht die Formatierung.



Die Option „Wort- und Paragraph-Formatierung Beibehalten“ ist ein Mittelding. Die **Wortformatierung** - Schriftart (serif - sans serif, proportional - fest), Schriftgröße und Schriftauszeichnung (fett, kursiv und unterstrichen) - werden beibehalten. Das Gleiche gilt für die **Absatzformatierung** - der Randausgleich und die Einzüge (links - justiert) werden ebenfalls beibehalten.



Verwechseln Sie diese Format-Option nicht mit „komplettem“ Autoformatierung: Hier werden die Absätze hintereinander gestellt, es werden keine Spalten oder Textblöcke kopiert. Auch die Grafiken werden ignoriert.

DAS POSITIONIEREN VON GRAFIKEN IM TEXT

In unserem Beispiel war die Grafik nicht im gescannten Text enthalten. Anstelle der Grafik wird ein Leerraum angezeigt. Das ist normal: Weil die Grafiken keinen Text enthalten kann die Seitenanalyse diese auch nicht erkennen. Wir möchten jedoch eine originalgetreue Kopie des Originaldokumentes erstellen. Dazu muß der OCR-Vorgang wiederholt werden, diesmal aber inklusive der Grafik.

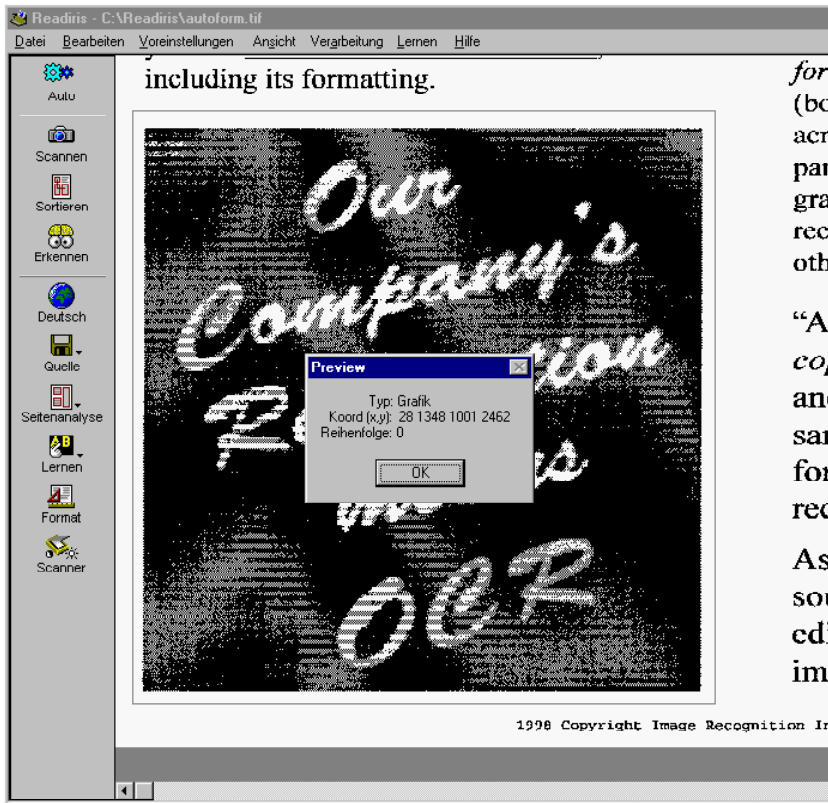
Hierzu muß um die Grafik ein Grafikenfenster gezeichnet werden, indem Sie den gleichen Arbeitsvorgang anwenden wie bei Textfenstern. Selektieren Sie danach das Symbol „Grafikenfenster“.



Das QuickInfo der Grafikensterschaltfläche teilt uns mit, daß wir jetzt ein Grafikenfenster auf unserem Bildschirm haben.



Sie können ebenfalls die Alt-Taste drücken, um ein Fenster zu selektieren. (Im Gegensatz zu den Textfenstern haben die Grafikenfenster keine Sortierreihenfolge.)



Wiederholen Sie nun die Erkennung, indem Sie die Format-Option "Original Dokument Wiederherstellen" selektieren. Sie werden feststellen, daß die Schwarzweißgrafik jetzt mit einbezogen wurde.

GRAFIKEN GETRENNT SPEICHERN

Grafiken können nur im Text gespeichert werden mit „kompletter“ Autoformatierung - dieses setzt voraus, daß Sie im Textformat RTF mit der Layoutoption „Original Dokument Wiederherstellen“ und nicht mit „begrenzte“ Textformate wie ASCII, ANSI usw arbeiten. Readiris kann jedoch Grafiken speichern ohne die Texte gelesen zu haben.

Da die Readiris-Software nur **Schwarzweißbilder** und keine Grautöne oder Farbbilder wiedergibt, hat es wenig Zweck Fotos mit Readiris zu scannen. Sie können jedoch mühelos schwarz-weiße Grafiken speichern.

Wie? Zeichnen Sie ein Grafikfenster um die Illustrierungen, Cartoons usw. Selektieren Sie danach den Befehl „Grafiken Speichern“ im Menü „Datei“.



Geben Sie einen Dateinamen ein und bestimmen Sie das Grafikformat. Selektieren Sie ein Format das von Ihrer Fotoretuschiersoftware unterstützt wird. Die TIFF- und Paintbrush-Formate (PCX) sind vorhanden. (Paintbrush-Dateien können immer mit dem Windows-Programm Paint geöffnet werden.)

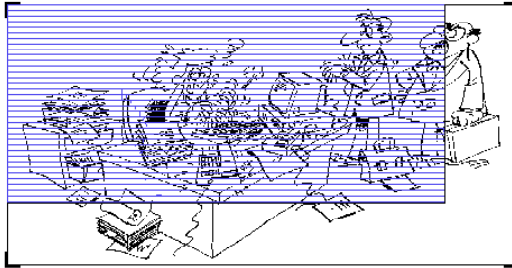


Die Grafiken werden in einer separaten Datei gespeichert. Sie brauchen sich nicht auf einer einzelnen Grafik zu beschränken. Bei der Erstellung von mehreren Grafikfenstern werden diese „gestapelt“ in einer Datei. (Der Befehl „Abschneiden“ im Fotoretuschierprogramm trennt die Dateien wieder von einander.)

Seiten von weniger als 1 mm. sind nicht erlaubt. Bitmap-Formate in der Größe enthalten ohnehin kaum Information. Fenster mit einer unregelmäßigen, nicht-rechteckigen Form bekommen eine rechteckige Form. Readiris bedeckt die größte Fläche, basiert auf den größten Koordinaten. Das Beispiel hier unten zeigt ein

Grafikfenster (hat schwarzen Rand), daß vervollständigt wurde von einem schattierten Bereich.

ur computer automatically.

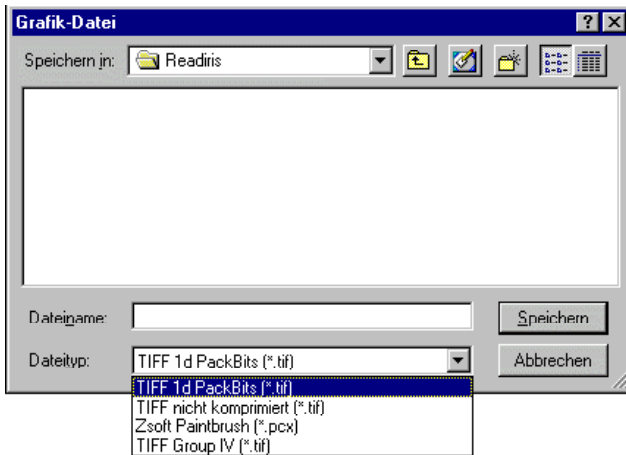


nt is read by your scanner. This device acts as the "eye" of yo

FAXE ERKENNEN UND ZURÜCKGESTELLTE ERKENNUNG

Die Speicherung von Grafiken als Bilddateien bietet Ihnen noch andere Möglichkeiten. Sie können die Bilder als *ganze* Seite speichern und die **zurückgestellte Erkennung** hinterher durchführen. Das gleiche wird mit den vorgescannten Bildern gemacht!

Zeichnen Sie ein Grafikfenster um eine ganze Seite und speichern Sie die Grafik dann im TIFF-Format.



Jetzt können Sie die Diskette als Quelle selektieren und die Bilddatei mit dem Symbol „Scannen“ öffnen. (Wenn Sie den Befehl „Öffnen“ im Menü „Datei“ benutzen, brauchen Sie die Quelldatei nicht auf dem neuesten Stand zu bringen).

Readiris öffnet schwarzweiße TIFF-Bilder (nicht-komprimiert, Packbits, Gruppe 3 und Gruppe 4 komprimiert), Paintbrush (PCX)-Bilder und Windows Bitmaps (BMP). Diese Funktion ist sehr praktisch, wenn Sie Ihre **Faxe** zur weiteren Bearbeitung in Textdateien konvertieren möchten!

Um bessere OCR-Ergebnisse zu erzielen, überzeugen Sie den Faxabsender davon, Faxe in „feiner“ Qualität mit einer Resolution von 200 dpi zu schicken.

Vergessen Sie nicht, daß Sie mit der Editierfunktion „**Drag & Drop**“ Bilddateien öffnen können: Ziehen Sie eine Bilddatei vom Windows Explorer zur Readiris Bildfenster. Die Datei wird sofort geöffnet.

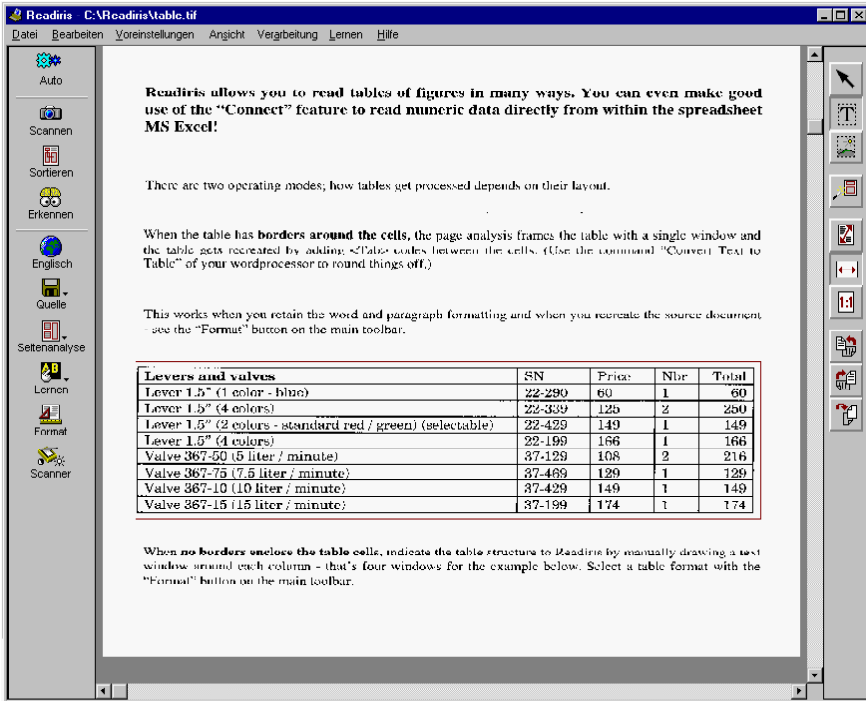
TABELLENERKENNUNG

Bis jetzt haben wir uns mit der Erkennung von Faxen und Texten, sowie mit der Speicherung von Grafiken befaßt. Jetzt werden wir eine Tabelle bearbeiten. Nehmen Sie eine Zeichentabelle und scannen sie oder öffnen Sie die Vorlage TABELLE.TIF im Readirisverzeichnis.

Tatsächlich enthält das Bild TABELLE.TIF zwei Tabellen. Das ist kein Zufall! Die Bearbeitung von Tabellen ist abhängig von Ihrem Layout.

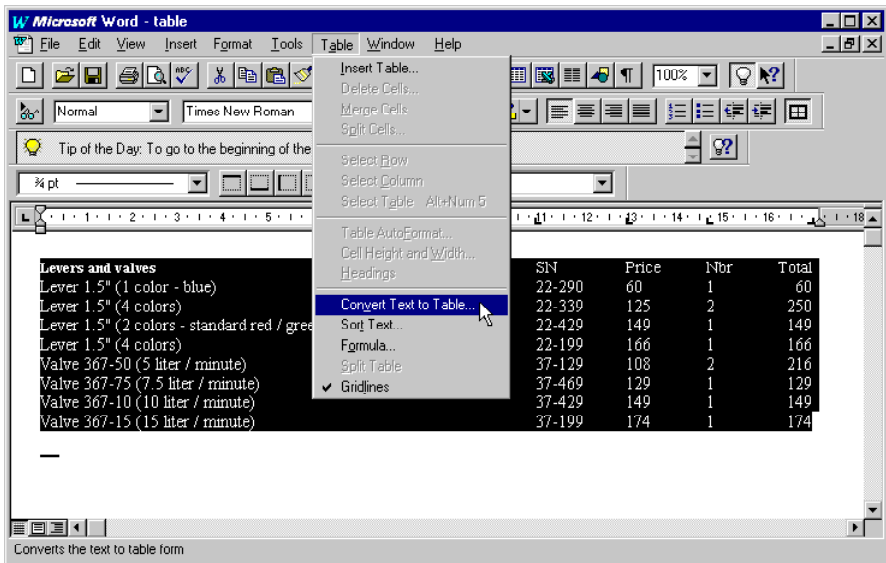
Wenn die Tabellenzellen eingerahmt sind, erstellt die Seitenanalyse ein einziges Fenster für die gesamte Tabelle. Wenn die Tabellenzellen nicht eingerahmt sind, muß der Anwender ein Fenster um jede Spalte ziehen, damit die Ausgabe in den entsprechenden Zellen präzise gestaltet wird.

Hier sind zwei Beispiele. Wir fangen mit der umrahmten Tabelle an. Weil *jede Zelle umrahmt* ist, untersucht die Seitenanalyse die Tabelle mit einem einzelnen (Text)Fenster.

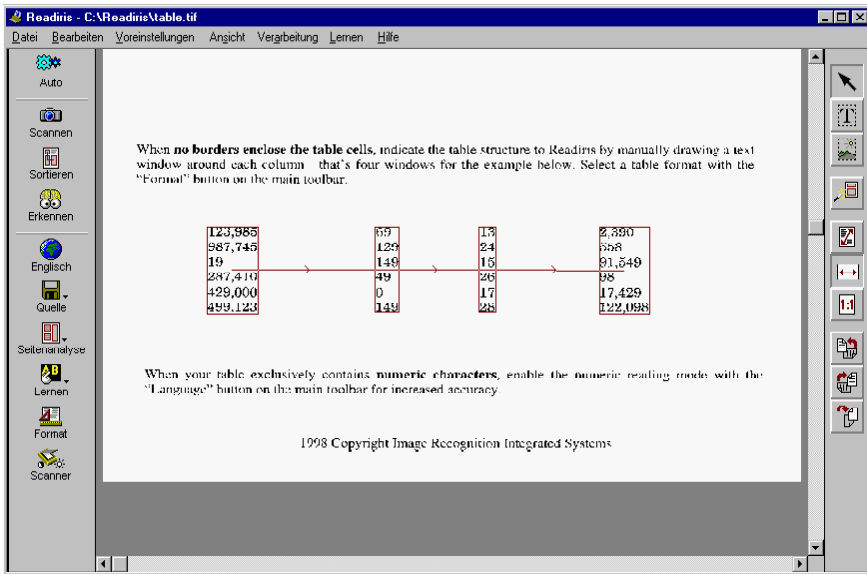


Starten Sie den Erkennungsvorgang mit der Layoutoption „Wort- und Paragraph-Formatierung Beibehalten“ oder „Original Dokument Wiederherstellen“. Die Tabellen werden durch Zufügen von Tabulatoren zwischen den einzelnen Zellen wiederhergestellt.

Öffnen Sie Ihren Textverarbeiter, um sich das Ergebnis anzuschauen. Benutzen Sie den Befehl Ihres Textverarbeiters „Text in Tabelle Umsetzen“ um diese Handlung abzurunden. (Sie könnten den laufenden Text ebenfalls in Ihrer Textdatei abgespeichert haben.)

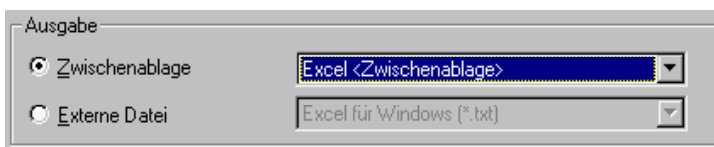


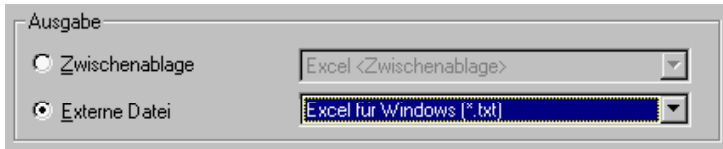
Jetzt das Beispiel mit den *nicht umrahmten* Tabellenzellen. Da die Seitenanalyse die Tabelle nicht als eine Einheit untersucht, zieht der Anwender ein Fenster um jede Spalte, damit die Ausgabe in den entsprechenden Zellen präzise gestaltet wird.



Jetzt müssen Sie mit dem Symbol „Format“ ein Textformat selektieren. Die Tabellen werden durch Zufügen von Tabulatoren zwischen den einzelnen Zellen wiederhergestellt. (Die „Standard“ Textformate reichen *nicht* aus, Sonderformate sind erforderlich um die Reihen und Spalten zu erstellen!)

So wie mit den Textdokumenten können Sie hier auch eine externe Datei erstellen, die in das Spreadsheet importiert werden kann. Sie können die Tabelle auch in die Zwischenablage kopieren und dann direkt in den Spreadsheet einfügen.



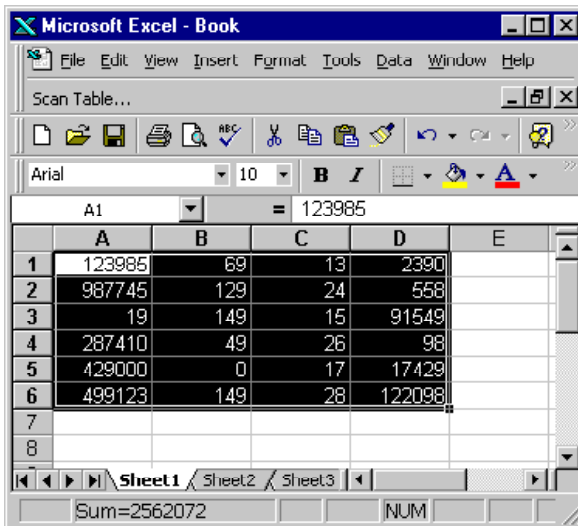


Um eine optimale Scanqualität zu bekommen, sollten Sie die Erkennung auf **numerischen Symbole** mit dem Symbol „Sprache“ begrenzen. (Der numerische Modus ist nicht ausschließlich numerisch, er enthält die Zeichen 0 bis 9, +, *, /, %, , (Komma), . (Punkt) (,), -, =, \$ und £).



Da diese Funktion nur auf Tabellen ohne alphabetischen Text angewendet werden kann - sonst werden die Textabschnitte nicht korrekt erkannt - können wir den numerischen Modus jetzt wohl aktivieren, aber nicht für die erste Tabelle.

Nachdem die Tabelle mit dem Spreadsheet geöffnet wurde, oder in das Spreadsheet eingefügt wurde, sieht sie wie folgt aus: Die typische Tabellenstruktur mit Reihen und Spalten wurde wiederhergestellt und kann die Tabelle sofort weiterbearbeitet werden.

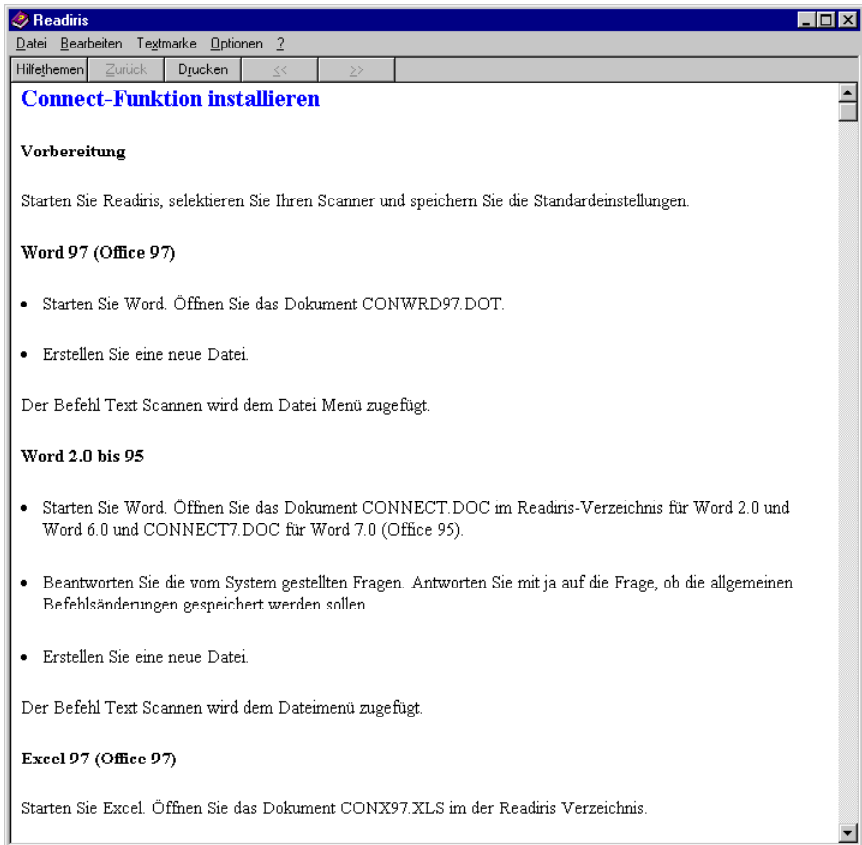


„CONNECT“ BILDET EINE DIREKTE VERBINDUNG

Es gibt noch einen anderen Weg, um Zeichentabellen zu scannen, nämlich direkt über den Spreadsheet.

Die „Connect“-Funktion bietet Ihnen eine direkte Verbindung zwischen Ihrem Scanner und Ihrer Windows-Anwendung. Nach Installation sind sie in der Lage das OCR-System von Ihren favoriten Windows-Anwendungen aus - Textverarbeiter oder Spreadsheet - auf zu starten.

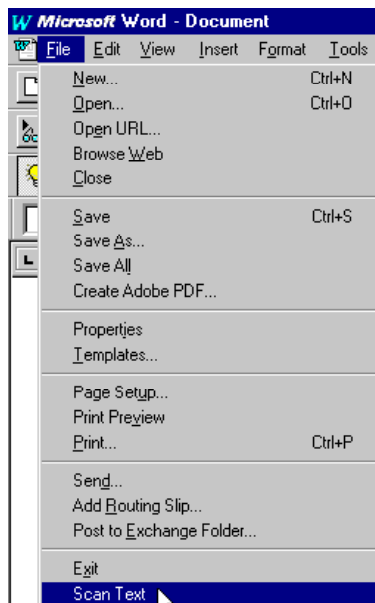
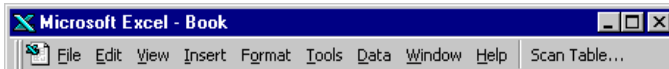
Bei der **Installation** gibt es einige Schritte die manuell ausgeführt werden müssen. Lesen Sie hierzu die Datei CONNREAD.DOC in Ihrem Readirisverzeichnis. (Diese „Lese Mich“-Datei erscheint automatisch am Ende des Installationsprogramms.)



Bei korrekter Installation fügt das „Connect“-Feature Ihrer Anwendung einen neuen Befehl zu. In den unten gezeigten Beispielen wird der Befehl „Text Scannen“



dem Word Menü „Datei“ zugefügt und der Befehl „Tabelle Scannen“ wird der Excel Menüleiste zugefügt.



So funktioniert es: Readiris darf *nicht* im Hintergrund laufen. Um Readiris zu starten, klicken Sie den neuen Befehl in Ihrem Applikationsmenü an. Die „Connect“-Schnittstelle wird Ihnen einige Fragen stellen bevor Readiris startet. Gehen Sie vor wie gewöhnlich und verlassen Sie die „Connect“-Schnittstelle nach Beendigung. Die Leseergebnisse werden der derzeitigen Cursorposition hinzugefügt.

ON-LINE HILFE BEKOMMEN

Hiermit beschließen wir unsere allgemeine Übersicht von Readiris. Es ist möglich, daß die neuesten Informationen noch nicht in diesem Handbuch besprochen wurden. Wir empfehlen Ihnen deshalb, das On-Line-Hilfe-System zu konsultieren wenn Sie zusätzliche Informationen über Readiris brauchen.

Selektieren Sie das Menü „Hilfe“. Der Befehl „Hilfethemen“ oder den Shortcut F1 gestattet Ihnen den Zugriff auf vielen On-Line-Hilfe-Informationen.

