

**Système de reconnaissance optique  
de caractères**

# **FineReader Version 4.0**

---

**Guide de l'utilisateur**

**ABBYY (BIT Software)  
Moscou 1999**

**ABBYY (BIT Software)**

**FineReader 4.0**

**Guide de l'utilisateur**

Les informations de ce document peuvent être modifiées sans avertissement préalable et ne représentent aucun engagement de la part d'ABBYY (BIT Software).

Le logiciel décrit dans ce document est fourni sous accord de licence. Le logiciel ne peut être utilisé ou copié qu'en accord avec les termes de cet accord. La copie du logiciel sur tout média, excepté ce qui a été spécifiquement autorisé dans la licence ou les accords de non divulgation, est strictement interdite par les lois et traités internationaux sur les droits d'auteur et la législation en vigueur dans votre pays.

Aucune partie de ce document ne peut être reproduite ou transmise sous quelque forme ou par quelque moyen que ce soit, électronique ou mécanique, sans l'autorisation écrite expresse préalable d'ABBYY.

© ABBYY (BIT Software), 1993-1999. Tous droits réservés.

ABBYY, BIT Software, FineReader, "fontain image transformation", Lingvo, Scan&Read, Scan&Translate, "one button principle", "Your computer reads by itself", "Your computer reads and translates by itself" sont des marques déposées d'ABBYY. Try&Buy, DOCFLOW sont des marques d'ABBYY. Toutes les autres marques sont des marques ou des marques déposées de leurs détenteurs respectifs.

P.O. Box 19, Moscow 105568, Russia, ABBYY

**ATTENTION ! Le LOGICIEL (et tous les documents qui l'accompagnent) vous est transmis aux conditions de l'accord de licence ci-joint. IL NE VOUS APPARTIENT PAS.**

Avant de desceller le contenu du coffret, lisez attentivement l'accord de licence. Le fait d'ouvrir le contenu du coffret signifie votre acceptation sans réserves des termes de l'accord de licence ci-joint.

Si vous n'acceptez pas les termes de cet accord, renvoyez le LOGICIEL scellé dans les sept (7) jours suivant l'achat avec tous les documents qui l'accompagnent à la société qui vous l'a vendu.

**ACCORD DE LICENCE LOGICIELLE**

L'accord de licence logicielle a valeur contractuelle et vous lie (que vous soyez un individu ou une entité légale) à ABBYY.

Le document ci-joint, "Preuve d'achat", fait partie intégrante du présent accord.

Le logiciel (ci-après dénommé "LOGICIEL") distribué par ABBYY est constitué d'une série de programmes, bases de données et documentation protégée par les lois et traités internationaux sur le copyright et les droits d'auteur. Dans le cadre du présent accord, nous entendons par "documentation" tous les documents imprimés et les fichiers de texte qui décrivent le LOGICIEL, et qui en font partie intégrante.

**1. Objet de l'accord de licence logicielle**

1.1 L'objet du présent accord de licence logicielle est l'octroi de la licence non exclusive d'utilisation du LOGICIEL transmis d'ABBYY à l'utilisateur final.

1.2 Tous les éléments cités ci-après réfèrent au LOGICIEL en général et à ses composants en particulier.

**2. Copyright**

2.1 Le copyright exclusif du LOGICIEL est détenu par ABBYY.

2.2 En tant qu'utilisateur final, vous recevez la licence d'utilisation du LOGICIEL si vous acceptez les termes et conditions qui suivent.

2.3 Cette licence vous est octroyée exclusivement à moins d'une autorisation spéciale écrite de la part d'ABBYY.

**3. Conditions d'utilisation du LOGICIEL**

3.1 Il vous est permis d'utiliser le LOGICIEL simultanément sur le nombre d'ordinateurs spécifié dans la colonne "Nombre de licences" du document "Preuve d'achat". Il vous est interdit d'utiliser le LOGICIEL sur un grand nombre d'ordinateurs simultanément. Il vous est interdit de stocker plus de copies du LOGICIEL que vous n'en détenez de licences, qu'il soit installé ou non.

3.2 Il vous est permis d'effectuer une copie du LOGICIEL à des fins de sauvegarde pour le seul cas où l'original viendrait à être endommagé ou égaré. Cette copie de sauvegarde ne doit être utilisée à d'autres fins.

3.3 Il vous est permis de créer vos propres bases de données pour les programmes faisant partie du LOGICIEL si une telle possibilité est mentionnée dans la documentation.

3.4 Si vous recevez le LOGICIEL sur plus d'un support, quel que soit le type ou la taille du support sur lequel vous le recevez, il ne vous est permis d'utiliser que le support correspondant à votre ordinateur. Les deux supports contiennent la même copie du LOGICIEL, et il ne vous est permis d'en utiliser qu'une.

3.5 Il vous est interdit de redistribuer le LOGICIEL. Par redistribution il est entendu l'accès par des tiers à des composants du LOGICIEL copiés par une société quelconque, par réseau ou d'autres manières, toute revente ou tout prêt.

3.6 Les activités suivantes sont autorisées avec des limites :

3.6.1 Il vous est permis d'installer et d'utiliser le LOGICIEL sur un maximum d'un ordinateur à processeur unique. Si le logiciel doit être installé sur réseau, seul un poste de travail peut y accéder à la fois.

3.6.2 Il vous est interdit de décompiler, désassembler et d'appliquer quelque rétrotechnique que ce soit au LOGICIEL, à ses composants, à ses bases de données et à quelque programme que ce soit.

3.6.3 Il vous est interdit d'effectuer quelque changement que ce soit dans le code du programme ou des bases de données à l'exception des changements expressément autorisés par le LOGICIEL lui-même et décrits dans la documentation.

3.6.4 Il vous est interdit d'accorder le copyright de l'utilisation du LOGICIEL ou d'octroyer quelque droit que ce soit, de copie ou autre, à des tiers.

3.6.5 Il vous est strictement interdit d'effectuer quelque action que ce soit qui contreviendrait aux lois et traités internationaux sur le copyright et les droits d'auteurs.

#### **4. Résiliation**

4.1 La présente licence logicielle entre en vigueur après que le LOGICIEL a été installé sur l'ordinateur et le reste pendant toute la durée d'utilisation du LOGICIEL.

4.2 La présente licence sera résiliée si vous ne respectez pas les termes et conditions du présent accord, auquel cas vous devrez détruire toutes les copies du LOGICIEL et tous les documents y afférents, disques, fichiers d'informations, copies d'archives du LOGICIEL et autres sauvegardes.

#### **5. Responsabilité de l'utilisateur**

5.1 Toute utilisation, redistribution ou copie illégales du logiciel constitue une infraction aux lois et traités internationaux et aux lois en vigueur dans votre pays sur le copyright, les droits d'auteur et la propriété intellectuelle.

5.2 Toute violation des termes et conditions du présent accord entraîneront l'annulation par ABBYY du droit d'utiliser le LOGICIEL, des garanties y afférentes, du support technique et des mises à jour gratuites du LOGICIEL.

#### **6. Garantie limitée**

6.1 ABBYY garantit la qualité des données présentes sur le support, la capacité de fonctionnement des programmes constituant le LOGICIEL conformément aux conditions décrites dans la documentation, la conformité des composants logiciels aux spécifications techniques et la qualité typographique de la documentation.

6.2 En ce qui concerne les autres composants, ABBYY déclare que le logiciel est livré « en l'état ». ABBYY ne donne aucune garantie et ne fait aucune déclaration concernant l'emploi ou les résultats découlant de l'emploi du LOGICIEL ou de la documentation quant à leur exactitude, fiabilité, actualité ou autres. En outre, ABBYY ne saurait être tenu responsable de tout dommage direct, indirect, secondaire ou accessoire, y compris les dommages entraînés par la perte de bénéfices, l'interruption des activités ou la perte d'informations et autres, découlant de l'utilisation ou de l'impossibilité d'utiliser le LOGICIEL, y compris les erreurs logicielles ou erreurs d'impression.

6.3 ABBYY ne donne aucune garantie et ne fait aucune déclaration sur le fonctionnement du logiciel avec des logiciels ou des matériels tiers déjà présent sur le système, particulièrement les logiciels et les matériels antérieurs à la date de lancement du LOGICIEL.

6.4 La présente garantie limitée reste valable pendant une période de 60 jours à compter de la date d'achat du LOGICIEL. Pendant cette période, ABBYY renvoie toutes les réclamations à son service d'assurance qualité, au département relatif au LOGICIEL.

*Merci d'avoir choisi FineReader!*

*FineReader est un outil performant qui vous permet de gagner du temps lorsque vous travaillez sur vos documents. Nous travaillons sur ce programme depuis des années et espérons qu'il vous facilitera la vie !*

*L'équipe d'ABBYY (BIT Software)*

---

# Table des matières

<b>Chapitre 1. Démarrage.....</b>	<b>8</b>
Comment reconnaître un texte ou un tableau.....	8
Comment entrer le texte d'un document imprimé dans un ordinateur ? .....	8
Comment reconnaître un fichier image .....	12
Comment reconnaître les formulaires (fonction disponible dans FineReader Professional et Handprint).....	12
<b>Chapitre 2. Installation et configuration .....</b>	<b>20</b>
Configurations logicielle et matérielle .....	20
Installation de FineReader .....	20
Problèmes lors de l'installation .....	21
Démarrage de FineReader .....	21
Utilisation du scanner.....	21
Vous n'avez pas encore acheté de scanner .....	23
Votre scanner ne fonctionne pas avec FineReader .....	23
Instructions à suivre.....	23
<b>Chapitre 3. FineReader 4.0 – Vue d'ensemble .....</b>	<b>24</b>
Fonctions de FineReader 4.0 Standard .....	24
Fonctions de FineReader 4.0 Professional .....	26
Fonctions de FineReader 4.0 Handprint .....	26
<b>Chapitre 4. Utilisation de FineReader 4.0 .....</b>	<b>27</b>
Fenêtre principale.....	27
Barre d'outils principale – Scanner&Lire .....	28
Barre d'outils Standard.....	29
Structure du lot.....	29
Aspect de la fenêtre de traitement .....	30
Numérotation des pages de traitement .....	31
Travail sur les pages de traitement .....	32

Numérisation .....	32
Textes de bonne et de mauvaise qualité .....	32
Numérisation de bonne qualité .....	33
Numérisation de documents de taille importante .....	34
Ajout d'images dans le lot.....	36
Ouverture d'une image couleur ou grise dans FineReader .....	36
Analyse de page (marquage des blocs) .....	37
Spécification de la structure des tableaux.....	38
Edition de blocs dans un tableau .....	39
Reconnaissance .....	40
Langue de reconnaissance .....	40
Type de texte .....	41
Enregistrement de la disposition et de la mise en forme du document .....	42
Exécution de la reconnaissance .....	43
Reconnaissance en tâche de fond .....	43
Vérification et édition des résultats .....	43
Enregistrement des résultats de la reconnaissance .....	45
Enregistrement d'un lot pour une utilisation future .....	46
Apprentissage de nouveaux caractères .....	47
Formation du système .....	47
Langues de FineReader .....	51
Langues de l'utilisateur et groupes de langues (pour FineReader 4.0 Professional et Handprint).....	51
Traitement de lots sur réseau (pour FineReader 4.0 Professional et Handprint) .....	51
Travail avec un lot sur réseau .....	52
Mode Lire&Attendre .....	52
<b>Annexes .....</b>	<b>53</b>
Amélioration de la qualité de la reconnaissance : astuces et conseils.....	53
Principes fondamentaux .....	53
Ajout de termes au dictionnaire .....	53
Apprentissage .....	53
Création de nouvelles langues.....	53
Exemple de reconnaissance du russe ancien .....	54
Formats d'images que FineReader peut charger .....	55

Langues prises en charge.....	56
Utilisation des dictionnaires de l'utilisateur et des motifs de FineReader 3.0 .....	56
Utilisation des touches dans FineReader .....	57
Recognition sample .....	59
Sample form for recognition.....	61



Nous avons tous, à un moment donné, besoin d'entrer des données textuelles dans nos ordinateurs. Que ce soit du texte provenant d'articles de journaux et de magazines, des contrats, des lettres commerciales, des télécopies, des listes de prix, des questionnaires et bien d'autres choses. 90% de tous les documents que nous recevons sont sur papier. Durant des années, il n'existait qu'un seul moyen d'informatiser les documents imprimés : la saisie manuelle. Vous aussi, peut-être, avez passé des heures à taper le texte de documents. Et sans doute avez-vous pensé « Quelle merveille si l'ordinateur pouvait lire lui-même, directement, la feuille de papier » !

Les rêves deviennent parfois réalité. Le programme de reconnaissance optique de caractères FineReader (ROC) permet à votre ordinateur et à votre scanner de lire eux-mêmes les textes imprimés.

*Mais le scanner ne peut-il pas faire le travail tout seul ?*

Non. Le scanner se contente de prendre la photo (le cliché) du texte. Une photographie est un ensemble de points noirs et blancs, ou, en d'autres termes, un fichier *graphique*, et ne peut pas être édité dans les logiciels de traitement de texte tels que MS Word, WordPerfect, Word Pro, etc. Seul un système ROC

(Reconnaissance Optique de Caractères) est capable de repérer les symboles dans cet ensemble de points noirs et blancs, de « reconnaître » une lettre dans chaque symbole et de transformer l'image en texte informatique. Un tel programme crée un fichier compréhensible par les éditeurs de texte et les systèmes informatiques.

*Je peux maintenant entrer directement mes documents dans mon ordinateur ?*

Oui, vous pouvez désormais entrer votre document directement dans votre ordinateur sans avoir à le retaper sur votre clavier.

*Amusez-vous bien !*

## Démarrage

FineReader est un programme ROC destiné à la saisie automatique de documents dans votre ordinateur. Le principe de son fonctionnement est relativement simple : vous insérez une page de texte dans le scanner, ce dernier réalise une image du texte, FineReader reconnaît ensuite l'image et affiche le fichier texte dans l'éditeur de texte. Le texte requis est placé dans votre éditeur de texte entre 30 et 60 secondes plus tard, avec la disposition, les images couleurs et les tableaux de l'original.

## Comment reconnaître un texte ou un tableau

Vous avez installé FineReader dans votre ordinateur, et vous avez connecté un scanner à ce dernier. (Des instructions détaillées sur l'installation du système sont données au chapitre Installation et configuration, p. 20.

Vous pouvez maintenant entrer du texte dans l'ordinateur. (Si FineReader est déjà en cours d'exécution, passez directement à l'étape 5.)

## Comment entrer le texte d'un document imprimé dans un ordinateur ?

1. Allumez le scanner.
2. Allumez l'ordinateur.
3. Lancez FineReader. Pour ce faire, dans le menu *Démarrer*, effectuez la séquence de commandes *Programmes/ABBYY FineReader/FineReader 4.0*.

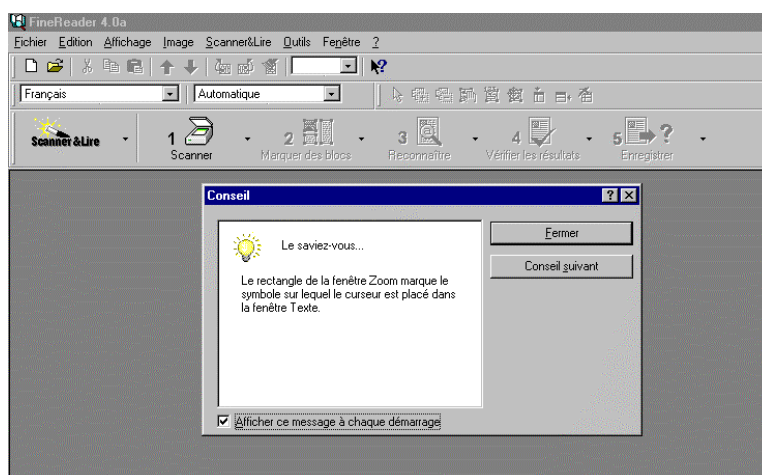


Figure 1. Fenêtre principale de FineReader 4.0.

La fenêtre principale de FineReader (Figure 1, p. 9) et la boîte de dialogue *Conseils* s'affichent. Cette boîte de dialogue contient des informations utiles. Après celle-ci, la boîte de dialogue *Bienvenue* s'affiche.

4. Insérez l'exemple de reconnaissance (il s'agit d'un modèle de formulaire qui se trouve à la fin de ce manuel) dans le scanner.
5. Cliquez sur le bouton *Assistant Scanner&Lire* situé dans l'angle supérieur gauche de l'écran.
6. Suivez les instructions de l'Assistant Scanner&Lire. En quelques minutes, vous allez apprendre à transformer des documents imprimés en fichiers. L'opération d'entrée de documents est très simple et se fait en trois étapes : numérisation, marquage de blocs et reconnaissance.
7. **Si vous utilisez un scanner à plat**, l'opération s'accompagne d'un bruit mécanique et du déplacement d'une source lumineuse sous le capot du scanner. **Si vous utilisez un scanner à entraînement automatique**, la feuille de papier est entraînée de la même façon qu'une télécopie. **Si vous utilisez un scanner à main**, passez-le sur la feuille de papier. Cliquez sur le bouton de démarrage et déplacez le scanner sur la page. Pour terminer la numérisation, appuyez sur n'importe quelle touche du clavier. (Ces conseils sont valables pour la plupart des scanners. Si vous rencontrez des problèmes, consultez le manuel de l'utilisateur de votre scanner.)

*Remarque* : si la boîte de dialogue *Paramètres du scanner* s'affiche, assurez-vous que les paramètres sont correctement définis : résolution – 300 ppp, luminosité – moyenne, mode de numérisation – gris ; cliquez ensuite sur le bouton *Scanner* (ou similaire) de cette boîte de dialogue.



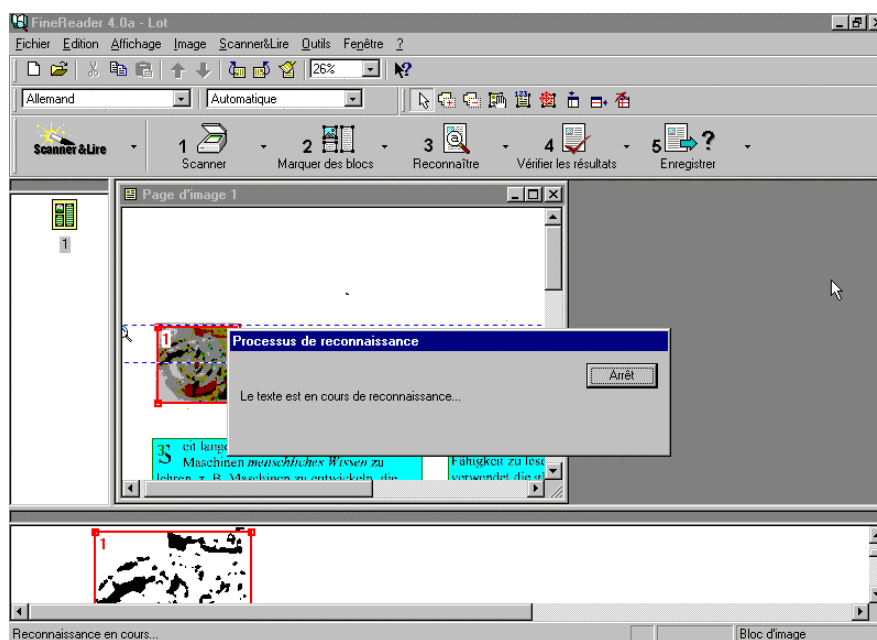


Figure 2. Reconnaissance en cours...

7. Après la numérisation, la fenêtre *Image* contenant un « cliché » de la page s'affiche. Puis, le programme démarre automatiquement l'analyse de l'image en localisant le texte, les images et les tableaux et reconnaît finalement le texte et les tableaux. Au cours de la reconnaissance, FineReader surligne les parties reconnues (Figure 2, p. 10).

**C'EST AUSSI SIMPLE QUE CELA ! LES RESULTATS DU TRAVAIL DE FINE READER PEUVENT MAINTENANT ETRE VISUALISES DANS LA FENETRE TEXTE (Figure 3).**



Figure 3. Fenêtre Texte.

La fenêtre *Texte* est l'éditeur de texte intégré de FineReader. Vous pouvez y éditer le texte reconnu et vérifier les résultats de la reconnaissance (bouton *Vérifier les résultats*).

8. Vous pouvez envoyer le texte reconnu dans l'application MS Word, Excel, etc. Pour ce faire, cliquez sur la flèche située à droite du bouton *Enregistrer* et dans le menu qui s'affiche, sélectionnez la commande *Envoyer à (MS Word, MS Excel, etc.)*.
9. Les résultats de la reconnaissance peuvent également être enregistrés sur le disque. Cliquez sur la flèche située à droite du bouton *Enregistrer* et dans le menu qui s'affiche, sélectionnez la commande *Enregistrer dans un fichier*. Le programme suggère d'enregistrer les résultats de la reconnaissance au format DOC. Ce format vous permet d'enregistrer non seulement le texte, mais aussi les informations relatives à la disposition, aux images et aux tableaux.

## Comment reconnaître un fichier image

1. Allumez l'ordinateur.
2. Lancez FineReader. Pour ce faire, dans le menu *Démarrer*, effectuez la séquence de commandes *Programmes/ABBYY FineReader/FineReader 4.0*.
3. Cliquez ensuite sur la flèche à droite du bouton *Scanner&Lire* et dans le menu correspondant, sélectionnez la commande *Ouvrir&Lire*.
4. Sélectionnez le lecteur, le dossier et le fichier graphique requis dans la boîte de dialogue *Ouverture*. Ce fichier peut porter l'extension TIF, PCX, BMP, DCX ou JPEG (la liste complète des formats qui peuvent être chargés dans

FineReader est fournie dans les Annexes).

Ouvrez, par exemple, le fichier *demo.tif* (situé par défaut dans le dossier *c:\Program Files\ABBYY FineReader\Demo*). Une « photographie » de la page s'affiche dans la fenêtre *Image*. Le programme démarre ensuite la reconnaissance de l'image et, 30 à 90 secondes plus tard, les résultats s'affichent dans la fenêtre *Texte*.

5. Vous pouvez dès lors envoyer les résultats vers une autre application ou les enregistrer sur le disque. Consultez les étapes 8 et 9 de la section précédente.

## Comment reconnaître les formulaires (fonction disponible dans FineReader Professional et Handprint)

Qu'entend-t-on par reconnaissance d'un formulaire ?

Il s'agit de reconnaître des formulaires imprimés standardisés, tels que les questionnaires, bulletins ou divers titres de paiement. Les formulaires standardisés sont des formulaires dont la disposition des champs est semblable.

Quelle est la tâche à réaliser ? Vous avez 50, 100 ou 10000 exemplaires d'un formulaire à votre disposition. Ils contiennent des *champs* d'informations, par exemple, Nom, Prénom, Total (paiement en lettres), Total (paiement en chiffres) et ainsi de suite. Vous devez entrer ces informations dans votre base de données, en lisant la valeur de chaque *champ* de chaque formulaire et en enregistrant celle-ci dans le *champ* correspondant de la base de données. De cette manière, les entrées se trouvant au même endroit sur tous les formulaires sont enregistrées ensemble et en ordre dans la base de données.

Nous expliquons comment entrer des formulaires en utilisant les exemples fournis à la fin de ce manuel (Demande d'emploi). Nous allons les utiliser dans l'exemple qui suit. Ces formulaires sont dactylographiés. En réalité, les formulaires peuvent également être manuscrits. Dans ce cas, utilisez **FineReader Handprint** pour les reconnaître. Un lot de formulaires préparés se trouve dans le dossier *c:\Program Files\ABBYY FineReader\Forms\Demo*, identifié par le nom de modèle **DEMO.FRM**. La description de ce fichier est fournie dans ce chapitre.

### Comment entrer un lot de formulaires ?

Préparez l'original de votre formulaire lisible par une machine. (Trois formulaires vous sont proposés à la fin de ce manuel, dont l'un est déjà rempli.) Nous allons utiliser le formulaire vierge comme original.

1. Effectuez des copies de cet original. N'apportez aucune modification à votre copie, afin que la disposition des champs coïncide exactement avec celle de l'original.
2. Remplissez les formulaires.
3. Exécutez FineReader 4.0 Forms, en effectuant la séquence de commandes *Démarrer, Programmes/ABBYY FineReader 4.0 Forms*.  
La boîte de dialogue *Lot de formulaires* s'affiche (Figure 4).

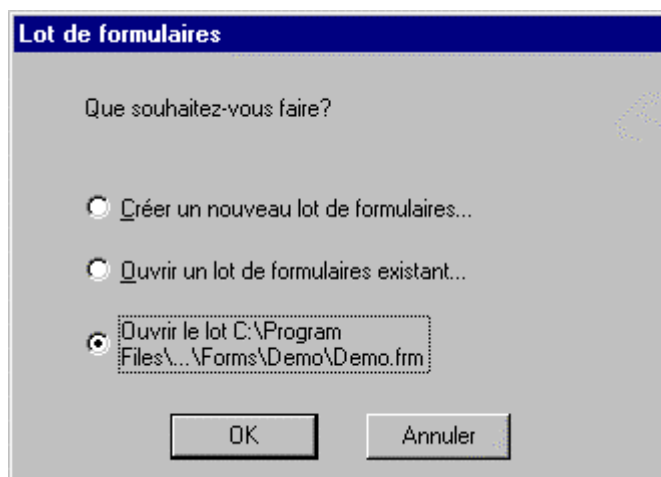


Figure 4. Boîte de dialogue Lot de formulaires.

4. Sélectionnez l'option *Création d'un nouveau lot de formulaires* et cliquez sur *OK*.
5. Dans la boîte de dialogue *Nouveau nom de lot* (Figure 5), spécifiez le nom du lot et le chemin d'accès aux fichiers de ce lot ; cliquez ensuite sur *Suivant*.

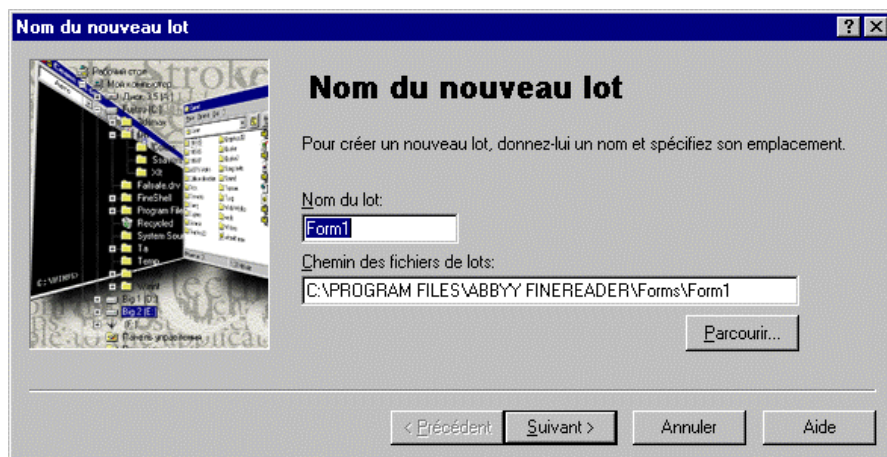


Figure 5. Boîte de dialogue Nouveau nom de lot.

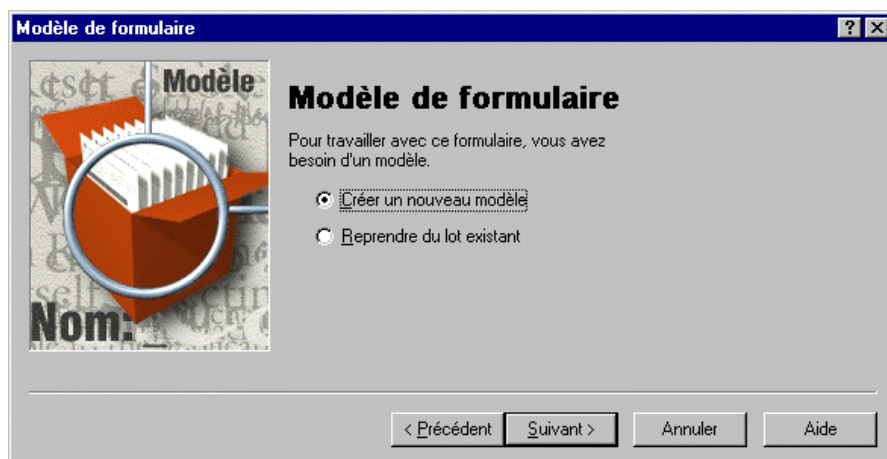


Figure 6. Boîte de dialogue *Modèle de formulaire*.

Dans la boîte de dialogue *Modèle de formulaire* (Figure 6), sélectionnez l'option *Créer un nouveau modèle*. Cliquez sur *Suivant* pour afficher la boîte de dialogue *Image modèle* (Figure 7).

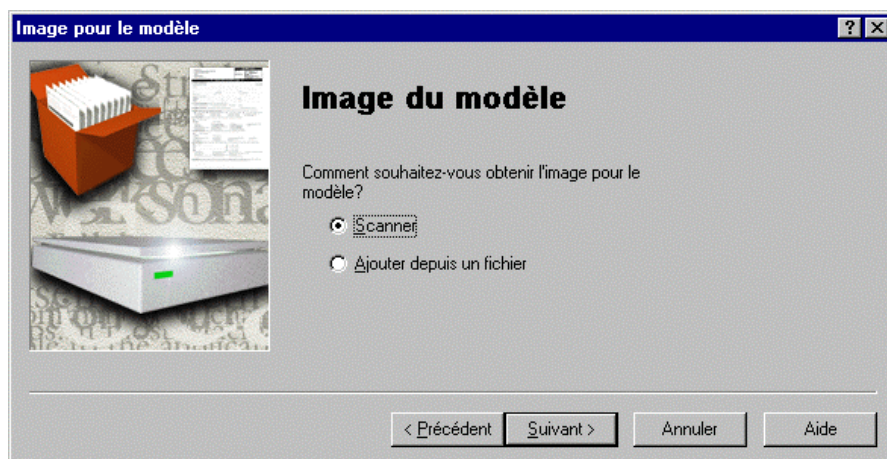


Figure 7. Spécifiez la nature de l'image modèle.

Dans la boîte de dialogue *Image modèle* (Figure 7), indiquez la nature de l'image. Cliquez sur *Scanner*. Nous allons maintenant utiliser le premier des formulaires vierges. Insérez la page dans le scanner et cliquez sur *Suivant*.



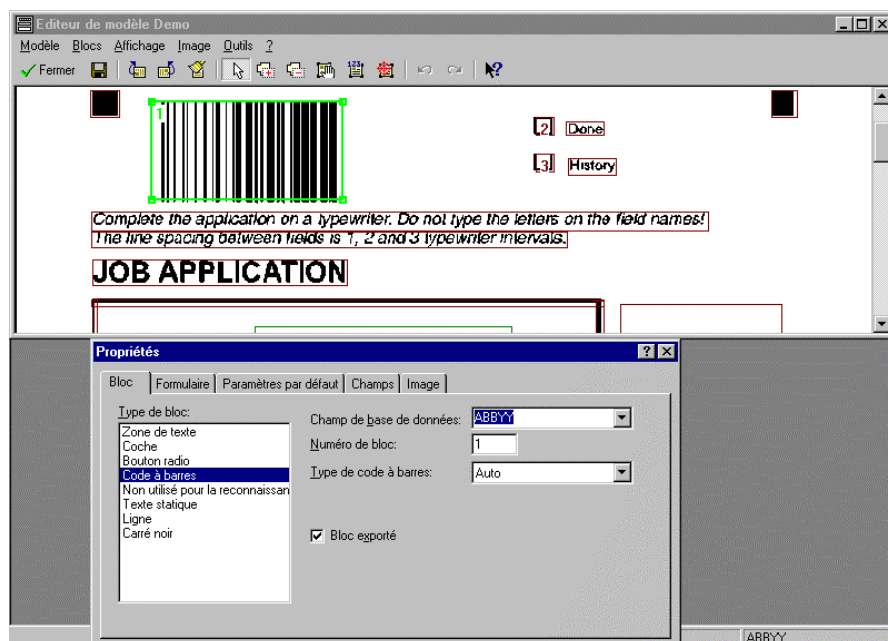


Figure 8. Image du formulaire dans la boîte de dialogue Editeur de modèle.

Après la numérisation, l'image scannée du formulaire original s'affiche dans la boîte de dialogue *Editeur de modèle* (Figure 8). La boîte de dialogue *Propriétés* s'affiche également.

### Création du modèle de formulaire

1. Dans le menu *Modèle*, sélectionnez la commande *Créer un modèle de formulaire automatiquement*.  
Les carrés noirs situés dans les angles du formulaire seront marqués comme blocs de type *Carré noir*. Tous les autres objets du formulaire seront marqués comme *Texte statique*.

**Attention !** N'effacez pas le texte statique ni les carrés noirs, car ils sont utilisés pour le positionnement correct du modèle.

2. Dans la page de l'onglet *Paramètres par défaut* de la boîte de dialogue *Propriétés* (Figure 9, p. 17) modifiez les paramètres suivants : langue – anglais, type de texte – dactylographié, nettoyer le bloc – oui.

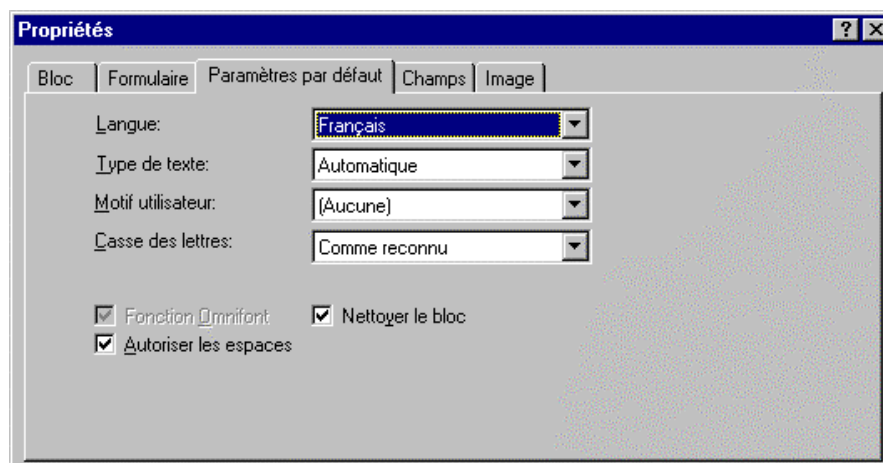


Figure 9. Onglet Paramètres par défaut de la boîte de dialogue Propriétés.

3. Vous devez maintenant « dessiner » les zones de texte qui contiendront le texte devant être reconnu. L'un des champs que vous devez reconnaître dans l'exemple est le champ Nom. Placez le pointeur de la souris sur l'image en haut à gauche du champ Nom. Cliquez faites glisser la souris en bas à droite, jusqu'à ce que le champ entier soit encadré. Relâchez le bouton de la souris. Un bloc (rectangle avec un cadre vert) s'affiche sur l'image. Il s'agit de la zone de texte.  
*Remarque* : tout le contenu des blocs à cadre vert est reconnu. Pour cette raison, vous ne devez pas incorporer le nom du champ ni d'autres lignes d'explication dans un bloc. Dans ce cas, n'incluez pas le titre Nom dans le bloc. Le bloc ne doit contenir que la partie de la page qui contient la valeur du champ (la partie remplie).  
 Si vous devez définir les marges de la zone reconnue plus précisément, dimensionnez le bloc ou déplacez son cadre avec la souris. Si vous faites une erreur lors du marquage des blocs, vous pouvez supprimer le bloc actif en appuyant sur la touche DEL.
4. Vous devez maintenant spécifier le champ de la base de données vers lequel les résultats de la reconnaissance du champ actuel doivent être exportés. Tapez le mot Nom dans la zone de texte *Champ de base de données*.  
*Remarque* : si la base de données existe déjà, sélectionnez la valeur requise dans le champ *Base de données*. Si vous n'avez pas encore créé de base de données, lors de l'attribution des noms aux champs, souvenez-vous des restrictions propres au format que vous projetez d'utiliser pour l'exportation, par exemple, la longueur des noms et les caractères supplémentaires. Vous n'aurez aucun problème avec les formats si vous n'utilisez que les lettres latines et des noms de huit caractères maximum.
5. Renouvelez ces opérations pour tous les champs de *texte* du formulaire : Prénom, Mois de naissance, Ville, Rue, N°, Appartement, Série, Publié le, Date, en donnant à chacun d'entre eux des noms appropriés dans la base de données.  
 Renouvelez ces opérations pour tous les champs contenant des *chiffres* du formulaire : Index, Code postal, Numéro de téléphone, Jour de naissance,

Année de naissance, Date, en spécifiant les champs appropriés dans la base de données. Par contraste avec les zones de texte, désélectionnez la case à cocher *Paramètres par défaut* et choisissez *Chiffres* dans la liste *Langue*.

6. Sélectionnez Masculin (M) et Féminin (F), puis « dessinez » un bloc qui les regroupera comme un seul *groupe bouton radio* Sexe.
7. Dans la partie supérieure du modèle, les cases à cocher (carrés) Terminé et Historique ont été automatiquement marquées comme texte statique. Vous devez maintenant changer leur type en *Cases à cocher*. Pour ce faire, cliquez sur ces blocs l'un après l'autre à l'aide de la souris, en maintenant la touche MAJ enfoncée. Cliquez ensuite à l'aide du bouton **droit** de la souris sur l'un de ces blocs et sélectionnez la commande *Type de bloc* dans le menu contextuel, puis sélectionnez la commande *Case à cocher*.
8. Le modèle électronique de ce formulaire est maintenant prêt. Cliquez sur *Fermer*. La fenêtre de traitement s'affiche.
9. Le travail sur un lot implique un certain nombre d'opérations séquentielles ou parallèles : numérisation, reconnaissance, édition et exportation des résultats de la reconnaissance.

### Numérisation



Préparez les formulaires remplis pour la numérisation (dans notre exemple, deux formulaires sont remplis) et cliquez sur le bouton *Scanner* dans la barre d'outils.

**Attention !** Toutes les images du lot doivent être définies sur la même résolution que le modèle du formulaire.

Pour chaque formulaire scanné, une nouvelle icône s'affiche dans la fenêtre de traitement, et une marque spéciale apparaît dans la colonne *Image*. Ces images ne sont que des représentations de formulaires et ne peuvent pas être éditées.

### Reconnaissance





Après avoir scanné tous les formulaires, vous pouvez démarrer la reconnaissance. Cliquez sur la flèche située à droite du bouton *Reconnaissance* et dans le menu contextuel, sélectionnez la commande *Reconnaître toutes les pages*.

Le programme commence la reconnaissance. Une marque spéciale s'affiche dans la colonne *Texte* à chaque fois qu'un formulaire est reconnu.

### Edition

Pour visualiser et éditer (si nécessaire) les résultats de la reconnaissance, cliquez deux fois sur l'icône du formulaire dans la fenêtre *Traitement*.

A droite de l'écran, vous voyez la fenêtre *Edition* ainsi que le formulaire reconnu, et dans la partie inférieure, l'image agrandie du formulaire.

Pour vous déplacer entre les pages de traitement, utilisez les boutons  et .



### Exportation

Pour exporter les résultats de la reconnaissance vers un fichier, cliquez sur la flèche située à droite du bouton *Enregistrer* et dans le menu, sélectionnez la commande *Exporter vers un fichier*. Sélectionnez le disque et le dossier dans lesquels vous voulez enregistrer les résultats de la reconnaissance et attribuez un nom au fichier. Cliquez ensuite sur le bouton *Enregistrer*.

### **C'EST TOUT ! LE FORMULAIRE RECONNU SE TROUVE DANS LE FICHIER TXT.**

Bien sûr, le format TXT n'est pas le seul format disponible. Vous pouvez enregistrer les résultats de la reconnaissance aux formats DBF, CSV ou XLS. Dans FineReader Handprint, vous pouvez aussi exporter des données via ODBC vers une base de données, telle que MS Access, par exemple.

### Traitement par réseau

Si le réseau comporte plusieurs ordinateurs, vous pouvez effectuer simultanément toutes les procédures décrites ci-dessus sur différents ordinateurs. Pour de plus amples informations sur le traitement sur réseau, consultez la section Traitement des lots sur le réseau, p. 51.

# Installation et configuration

## Configurations logicielle et matérielle

Configuration requise pour installer FineReader 4.0 :

- Un ordinateur équipé d'un processeur Intel 486 ou supérieur (un Pentium 133 ou supérieur est recommandé).
- L'un des systèmes d'exploitation suivants : Windows 95 (la version pan-européenne est recommandée), Windows 98, Windows NT 4.0 (Service Pack 3).
- 16 Mo de RAM pour Windows 95, 98; 32 Mo pour Windows NT.
- 30 Mo d'espace disponible sur le disque dur pour la configuration minimale et environ 60 Mo pour la configuration maximale.
- Un scanner compatible TWAIN ou un périphérique analogue.
- Une souris ou tout autre dispositif de pointage.
- Un lecteur de CD.

## Installation de FineReader

FineReader vous guide pendant l'installation à l'aide d'instructions à l'écran à chaque étape. Pour de meilleurs résultats, n'exécutez aucun autre programme (par exemple, MS Outlook, MS Word, etc.) lors de l'installation.

---

**Attention !** Si vous souhaitez utiliser des dictionnaires et des motifs des versions précédentes du programme, ne désinstallez pas la version antérieure avant d'avoir installé la nouvelle version. Après avoir installé la nouvelle version, vous pouvez utiliser les motifs et les dictionnaires de la version précédente. Pour de plus amples détails consultez l'annexe Utilisation des dictionnaires de l'utilisateur et des motifs de FineReader 3.0

---



1. Insérez le CD-ROM dans le lecteur de CD.
2. Cliquez sur *Démarrer* dans la barre des tâches et sélectionnez la commande *Panneau de configuration*.
3. Cliquez deux fois sur *Ajout/suppression de programmes*.
4. Sélectionnez l'onglet *Installer/désinstaller* et cliquez sur *Installer*.
5. Suivez les instructions à l'écran.

*Remarque :* les étapes 2 à 4 reviennent à exécuter install.exe.

## Problèmes lors de l'installation

En de rares occasions, des problèmes peuvent survenir lors de l'installation à cause de l'incompatibilité entre différents composants logiciels.

Si le programme affiche un message d'erreur, consultez le fichier Readme.hlp qui contient des recommandations utiles dans la plupart des cas.

## Démarrage de FineReader

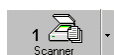
Si vous projetez de numériser, assurez-vous que votre scanner est connecté à votre ordinateur, sous tension\* et en état de marche, avant de démarrer FineReader. Si vous n'avez pas encore connecté le scanner, consultez le Guide de l'utilisateur fourni avec ce dernier pour le connecter.

Si vous n'avez pas de scanner, vous pouvez voir comment FineReader fonctionne dans le fichier demo.tif, que vous trouverez dans le dossier c:\Program Files\ABBYY FineReader\Demo. Pour de plus amples informations sur l'utilisation des fichiers graphiques, consultez la section Comment entrer le texte d'un document imprimé dans un ordinateur ?, p. 8.

### Pour démarrer FineReader :

- Cliquez sur le bouton *Démarrer* dans la barre des tâches et effectuez la séquence de commandes *Programmes/ABBYY FineReader/FineReader 4.0*.

## Utilisation du scanner



---

**Important !** Pour connecter un scanner correctement, lisez son manuel de l'utilisateur. En le connectant, n'oubliez pas d'installer tous les logiciels intégrés (le pilote TWAIN et/ou le programme de numérisation).

---

Pour démarrer la numérisation, effectuez tous les préparatifs et cliquez sur *Démarrer*.

**Votre scanner devrait commencer à fonctionner immédiatement, et 20 secondes plus tard, une « photographie » de la page scannée devrait s'afficher à l'écran.**

Vous pouvez maintenant lire la section Comment entrer le texte d'un document imprimé dans un ordinateur ?, p. 8.

---

\* Si votre scanner possède une alimentation séparée, placez-le sous tension **avant** votre ordinateur.

Si le scanner ne fonctionne pas, la boîte de dialogue des paramètres du scanner s'affiche.

*Remarque* : des scanners différents possèdent des interfaces différentes. Les boîtes de dialogue de gestion du scanner sont des interfaces TWAIN intégrées développées par les constructeurs. C'est pourquoi elles varient d'un scanner à l'autre et qu'elles ne sont pas toujours conviviales et intuitives.

Vérifiez maintenant les paramètres de numérisation. Si vous n'avez jamais utilisé de scanner auparavant, sachez que les scanners possèdent trois paramètres principaux qui influencent la qualité de l'image scannée : la résolution, la luminosité et le mode de numérisation (type d'image).

**Pour définir correctement les paramètres de numérisation, effectuez les réglages suivants :**

- Résolution – sélectionnez 300 ppp.  
*Remarque* : certaines interfaces TWAIN (par exemple, HP PictureScan) ne permettent pas de définir de résolution ROC autre que 300 ppp.
- Luminosité (parfois appelée seuil ou représentée par l'icône « soleil » ou un cercle noir et blanc, etc.) – sélectionnez la valeur moyenne (50%).
- Mode de numérisation – sélectionnez l'option gris (256 couleurs)\*.

Une fois les paramètres définis, cliquez sur *Scanner* (ce bouton peut être appelé différemment, *Fin*, par exemple) pour démarrer votre scanner.

**Si vous utilisez un scanner à plat**, l'opération s'accompagne d'un bruit mécanique et du déplacement d'une source lumineuse sous le capot du scanner.

**Si vous utilisez un scanner à entraînement auto**, la feuille de papier est entraînée de la même façon qu'une télécopie.

**Si vous utilisez un scanner à main**, passez-le sur la feuille de papier. Cliquez sur le bouton de démarrage et déplacez le scanner sur la page. Pour terminer la numérisation, appuyez sur n'importe quelle touche du clavier.

A la fin de la numérisation, la « photographie » de la page scannée s'affiche dans la fenêtre *Image* de FineReader.

Consultez le manuel de l'utilisateur du scanner pour obtenir de plus amples informations.

**Vérification des propriétés de l'image scannée**

Vous pouvez vérifier les propriétés de l'image scannée. Pour ce faire, cliquez à l'aide du bouton **droit** de la souris sur l'image, et dans le menu contextuel qui s'affiche sélectionnez la commande *Propriétés*.

Si vous avez suivi toutes les recommandations de ce manuel, l'image devrait avoir les propriétés suivantes : type d'image – gris (voir ci-dessus), résolution – 300 ppp. Ces valeurs sont les valeurs de reconnaissance les plus communes.

---

\* Pour tous les scanners mis à part les scanners à main. Pour ces derniers – noir et blanc (dessin, ROC, texte)

## **Vous n'avez pas encore acheté de scanner**

Choisissez un appareil compatible TWAIN. TWAIN est un protocole standard reconnu par la plupart des constructeurs de scanners et garantit un fonctionnement normal des scanners avec FineReader.

*Remarque* : malheureusement, vous ne pouvez obtenir des résultats de numérisation satisfaisants avec la plupart des scanners qu'en utilisant l'option d'affichage de la boîte de dialogue TWAIN.

Lorsque cette option est désactivée, certains pilotes TWAIN risquent de ne pas vous permettre de définir la luminosité et/ou la taille du papier manuellement.

Vous pouvez voir la liste des scanners compatibles avec FineReader dans le fichier readme.hlp ou sur le site Web d'ABBYY ([www.abbyy.ru](http://www.abbyy.ru)).

## **Votre scanner ne fonctionne pas avec FineReader**

Si votre scanner ne fonctionne pas avec FineReader, essayez d'effectuer les opérations suivantes :

1. Procurez-vous une nouvelle version du pilote (gestionnaire), auprès d'un revendeur de scanners ou via Internet.

*Remarque* : si vous ne connaissez pas l'adresse du constructeur de votre scanner, contactez l'adresse du comité d'organisation TWAIN : [www.twain.org](http://www.twain.org). Il existe par ailleurs des constructeurs de pilotes TWAIN indépendants pour les modèles les plus courants, par exemple, CFM ([www.cfm.de](http://www.cfm.de)).

2. Rendez-vous sur le site Web d'ABBYY: <http://www.abbyy.ru>. Vous trouverez peut-être le nouveau fichier de configuration de votre scanner.
3. Ecrivez au service de support technique d'ABBYY: [support@abbyy.ru](mailto:support@abbyy.ru).
4. Si vous ne pouvez obtenir de nouvelle version du pilote de votre scanner, contournez le problème de la manière suivante : scannez les images requises à l'aide du logiciel de numérisation, enregistrez-les au format pris en charge par FineReader (TIFF, BMP, JPEG, PCX, DCX\*) et ouvrez ces images dans FineReader.

*Remarque* : vous pouvez assembler un lot automatiquement. Pour ce faire, placez les images dans le dossier de traitement. Les noms des images devant être assemblées dans le lot doivent comprendre le nom du lot ainsi qu'un nombre de quatre chiffres. Pour mettre à jour les pages de la fenêtre de traitement, appuyez sur la touche F5.

## **Instructions à suivre**



Après avoir installé FineReader 4.0, et testé le fonctionnement de votre scanner, rendez-vous au chapitre Démarrage dans lequel vous trouverez des instructions pour parvenir à obtenir rapidement les résultats recherchés.

Vous pouvez apprendre à travailler avec le programme en mode interactif en exécutant l'Assistant Scanner&Lire. L'Assistant vous aide à apprendre les notions fondamentales de la numérisation et de la reconnaissance. Lisez les messages et suivez les instructions de l'Assistant. Bonne chance !

---

\* La liste des formats pris en charge est fournie dans l'Annexe.



# FineReader 4.0 – Vue d'ensemble

FineReader est un logiciel ROC (Reconnaissance Optique des Caractères) polyvalent toutes polices. Ceci signifie que sans apprentissage préalable, il est capable de traiter des textes de presque toutes les qualités, imprimés dans n'importe quelle police de caractères.

FineReader est extrêmement tolérant en ce qui concerne les défauts d'impression. Cette tolérance est le résultat d'une technologie spéciale appelée « fontain image transformation ».

Ce chapitre offre une vue d'ensemble des fonctions de FineReader 4.0 Standard, Professional et Handprint.

## Fonctions de FineReader 4.0 Standard

- FineReader vous permet de saisir des textes 5 à 10 fois plus vite qu'un dactylographe professionnel et avec beaucoup moins de fautes.
- Pour entrer une image dans l'ordinateur, insérez un document dans le scanner et cliquez sur *Scanner&Lire*. Dans un intervalle de 30 à 90 secondes, le texte reconnu s'affiche dans la fenêtre *Edition*.
- L'Assistant Scanner&Lire représente le seul moyen d'obtenir des résultats rapidement, sans entrer dans les détails du programme.
- FineReader permet de reconnaître et d'éditer du texte reconnu sur un seul et même ordinateur simultanément.
- Il prend en charge le déplacement par glissement.
- Il prend en charge le standard MMX.
- Il s'intègre à Lingvo 4.5 et version supérieure.

### Numérisation et marquage des blocs

- Fonctionne avec tous les scanners qui prennent en charge le protocole TWAIN.
- Réglage automatique de la luminosité – pour obtenir de meilleurs résultats lors de la numérisation des parties sombres ou claires de la page.
- Segmentation automatique et manuelle des tableaux.
- Numérisation et enregistrement des images couleurs.

### Numérisation des images

- Détection automatique de l'orientation de la page en cours de reconnaissance et rotation de cette dernière pour la placer dans sa position normale.
- Nettoyage des images.
- Rotation des images de 90, 180 et 270 degrés.
- Inversion automatique des blocs sombres avec du texte clair.

### **Reconnaissance**

- Reconnaissance de textes russes, anglais, français, allemands, espagnols, italiens, suédois et autres, y compris les textes contenant diverses combinaisons de langues reconnues. La liste des langues reconnues est fournie à l'Annexe (p. 56).
- Reconnaissance de texte à multiples colonnes avec des images et des tableaux ; enregistrement de la disposition de la page reconnue.
- Reconnaissance de blocs de texte clair sur fond sombre.
- Reconnaissance des images couleur.
- Formation du programme à la reconnaissance de nouveaux caractères.

### **Edition et exportation**

- Editeur de texte intégré vous permettant de modifier la police et sa taille, ainsi que la mise en forme du texte et des paragraphes, etc.
- Correcteur orthographique intégré vous permettant d'ajouter de nouveaux termes au dictionnaire du système en russe, anglais, allemand, français, ukrainien, espagnol, italien, suédois, néerlandais, danois, norvégien, polonais, finlandais, portugais et tatar\*.
- Création automatique de différents styles de paragraphes de mise en forme et de police différentes.
- Utilisation d'Internet – enregistrement des résultats de la reconnaissance au format HTML.
- Possibilité d'envoyer les résultats de la reconnaissance vers une autre application (par exemple, MS Word 6.0, MS Word 95, MS Word 97, MS Excel 6.0, MS Excel 95, MS Excel 97, Word Pro 97, WordPerfect 7.0, WordPerfect 8.0, Stylus 3.0, PROMT 98) sans les enregistrer sur le disque.

## **Fonctions de FineReader 4.0 Professional**

- Toutes les fonctions de FineReader 4.0 Standard (voir ci-dessus).
- Traitement de documents en réseau afin d'améliorer considérablement la vitesse de traitement de documents importants (lorsque vous possédez des licences pour plusieurs stations de travail).
- Reconnaissance de formulaires sur la base d'un modèle : entrée de quantités importantes d'informations standardisées dans les systèmes d'informations.
- Application des modèles à l'aide de points de référence.
- Concepteur de modèles permettant de reconnaître différents formulaires.
- Reconnaissance de textes en langues non reconnues.

---

\* Le tatar fait l'objet d'une distribution particulière.

- Reconnaissance de codes à barres (Check Code 39, Check Interleaved 25, Code 128, Code 39, EAN 13, EAN 8, Interleaved 25).
- Compensation automatique du désalignement lors de la numérisation pour rétablir la correspondance avec le modèle.

## Fonctions de FineReader 4.0 Handprint

- Toutes les fonctions de FineReader 4.0 Professional (voir ci-dessus).
- Reconnaissance des blocs *manuscripts* en lettres d'imprimerie en russe, anglais, allemand et ukrainien.
- Application automatique de modèles lors de la reconnaissance d'une pile de formulaires non classée dans un lot.
- Exportation d'images dans des bases de données.
- Vérification automatique des résultats de la reconnaissance à l'aide d'expressions régulières, de règles de validation et de consultations de la base de données.
- Exportation vers les bases de données via ODBC.

## Utilisation de FineReader 4.0

Ce chapitre présente les fonctions principales de FineReader 4.0. Pour des informations plus détaillées et des instructions étape par étape, consultez l'aide en ligne.

### Fenêtre principale

Lorsque le programme est en cours d'exécution, un nouveau lot est ouvert par défaut, dans lequel les résultats de la reconnaissance sont placés. Chaque image scannée est représentée sous forme de page séparée dans le lot. La fenêtre principale présente fenêtres de travail suivantes : *Lot*, *Image*, *Texte* et *Zoom* (Figure 10).



Figure 10. Fenêtre principale de FineReader 4.0.

Dans la partie supérieure de la fenêtre se trouvent les barres d'outils. Pour masquer ou afficher ces barres d'outils, utilisez le menu *Affichage > Barre d'outils* ou le menu contextuel. Pour ouvrir le menu contextuel (Figure 11, p. 28), cliquez à l'aide du bouton **droit** de la souris sur l'une des barres d'outils. Les barres d'outils affichées à l'écran sont

marquées d'une coche. La barre d'outils Edition est désactivée par défaut (elle n'est pas affichée à l'écran). Sélectionnez la barre d'outils que vous voulez masquer ou afficher.

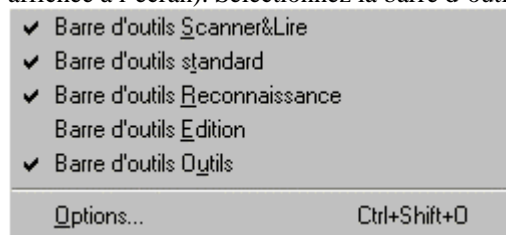


Figure 11. Menu contextuel avec la liste des barres d'outils. Les barres d'outils cochées sont celles qui sont affichées sur l'écran du programme. La barre d'outils Edition est désactivée.

## Barre d'outils principale – Scanner&Lire

La barre d'outils Scanner&Lire (Figure 12) contient des boutons qui réalisent les opérations principales du programme. Les numéros des boutons représentent l'ordre des opérations à effectuer pour obtenir une version électronique du document sur papier : 1 – Numérisation, 2 – Marquage des blocs, 3 – Reconnaissance, 4 – Vérification des résultats et 5 – Exportation. Le bouton le plus à gauche *Scanner/Ouvrir&Lire (Assistant)* peut réaliser la numérisation, le marquage des blocs et la reconnaissance (en mode Scanner&Lire – pour les documents sur papier) ou le marquage de blocs et la reconnaissance (en mode Ouvrir&Lire – pour les images électroniques). Ce bouton peut également fonctionner en mode Assistant Scanner&Lire, qui va vous apprendre à travailler avec le programme ROC.

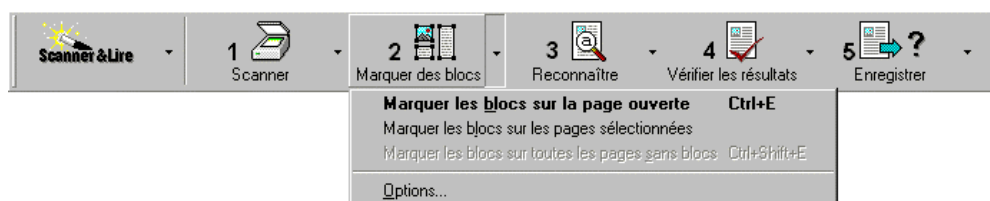


Figure 12. Barre d'outils Scanner&Lire.

Le mode des boutons *Marquer les blocs* et *Reconnaître* dépend de la fenêtre active et aussi du fait que la fenêtre de traitement ait certaines pages sélectionnées ou non.

Si la fenêtre *Image* est active, les boutons traitent la page ouverte.

S'il n'y a aucune image ouverte et qu'au moins une page est sélectionnée dans la fenêtre de traitement, les boutons traitent les images sélectionnées.

*Remarque :* le mode bouton différent du mode par défaut peut être défini manuellement. Cliquez sur la flèche située à droite du bouton pour voir le menu s'afficher. Sur l'image (Figure 12) vous pouvez voir le menu du bouton 2 – *Marquer les blocs*. L'un des éléments du menu est en gras. Cette commande sera exécutée lorsque

vous cliquez sur le bouton. Lorsque vous sélectionnez un autre élément du menu, l'opération correspondante est réalisée.

## Barre d'outils Standard

La barre d'outils Standard (Figure 13) possède les boutons suivants : cinq boutons standard (nouveau, ouvrir, couper, copier, coller), des boutons pour vous déplacer vers le haut et vers le bas dans le lot (vers les pages suivantes et précédentes), des boutons pour faire pivoter et nettoyer les images, ainsi qu'un bouton de mode *Aide*.

Lorsque vous cliquez sur un bouton, vous opérez sur l'image ouverte ou les images sélectionnées dans la fenêtre de traitement. Lorsque vous cliquez sur le bouton *Mode Aide*, le pointeur de la souris fonctionne de la manière suivante : cliquez sur un objet (bouton ou élément de menu) et vous obtenez une aide brève (le pointeur de la souris, dans ce mode, se transforme en une flèche accompagnée d'un point d'interrogation).



Figure 13. Barre d'outils Standard.

## Structure du lot

Lorsque le programme est exécuté, un nouveau lot est ouvert par défaut. Un lot est une sorte de dossier dans lequel les pages scannées et reconnues sont assemblées. Vous pouvez enregistrer des pages séparées du lot ou le lot complet. Dans la fenêtre de traitement, vous pouvez voir la liste des pages du lot en cours. Pour visualiser une page, cliquez deux fois sur son icône ou son numéro. Vous verrez les fichiers présents dans la page : du texte et/ou des images dans les fenêtres *Texte* et/ou *Image*, respectivement.

Les pages du lot sont sélectionnées de la même manière que les fichiers dans l'Explorateur : en cliquant avec la souris (ou à l'aide des touches). Pour de plus amples détails, consultez la section *Sélection des pages*, p. 32.

## Aspect de la fenêtre de traitement

Par défaut, la fenêtre de traitement se présente sous la forme d'une fenêtre « arrimée » au côté gauche de la fenêtre principale, dans laquelle les pages sont représentées sous forme de grandes icônes (Figure 14).

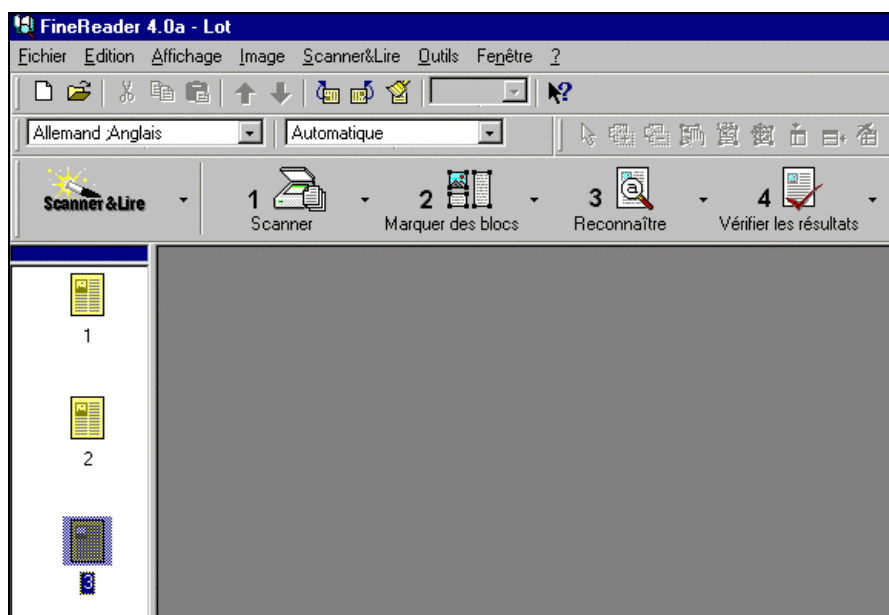


Figure 14. Fenêtre Traitement arrimée contenant de grandes icônes représentant les pages (affichage par défaut).

Si vous voulez déplacer la fenêtre *Traitement*, vous pouvez la désarrimer (cliquez à l'aide du bouton **droit** de la souris dans la fenêtre, puis dans le menu contextuel, désélectionnez la commande *Arrimage*).

La fenêtre *Traitement* peut également afficher des informations sur les opérations effectuées sur telle ou telle page : si elle a été reconnue, éditée, exportée, si une erreur s'est produite lors de la reconnaissance, etc. Pour obtenir des informations sur les pages du traitement, il existe un mode d'affichage spécial des paramètres (Figure 15). Pour obtenir cet affichage, dans le menu contextuel, sélectionnez la commande *Afficher les paramètres de la page*.

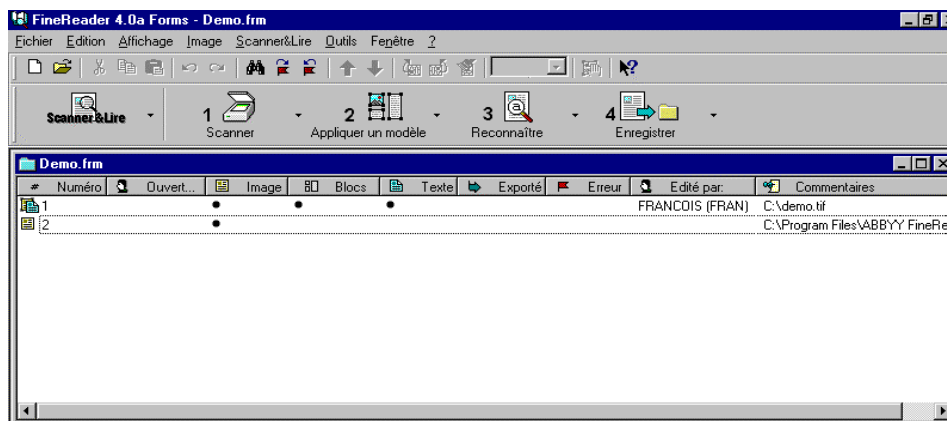


Figure 15. Fenêtre Traitement. Visualisation des paramètres. Les composants présents dans les différentes pages sont indiqués à l'aide de puces (image, blocs, texte, etc.).

### Représentation conventionnelle des pages

Une page de lot peut posséder :

Un fichier *graphique* (image scannée et enregistrée). Ce type de page est indiqué par une icône jaune (si la vue sans paramètres est sélectionnée).

Un fichier *texte* (texte reconnu). Ce type de page est indiqué par une icône blanche (ou bleue).

Une page peut être exclusivement un fichier texte si l'image de la page a été supprimée après la reconnaissance (si la vue sans paramètres est sélectionnée).

Des fichiers *graphiques et textes*. Ce type de page est indiqué par deux icônes qui se chevauchent – jaune et blanche (si la vue sans paramètres est sélectionnée).

### Numérotation des pages de traitement

Un lot peut contenir jusqu'à 9999 pages. Le numéro de la page se trouve à côté de la représentation conventionnelle de la page située dans la fenêtre *Traitement* (voir ci-dessus) et dans l'en-tête de la fenêtre dans laquelle les fichiers de pages sont ouverts, par exemple, Image – 5, Texte – 5.



## Travail sur les pages de traitement

### Sélection des pages

Le lot accepte plusieurs sélections de pages. Cela signifie que vous pouvez sélectionner simultanément plusieurs pages consécutives ou non consécutives.

**Pour sélectionner plusieurs pages consécutives**, cliquez sur la première page à l'aide de la souris. Maintenez la touche MAJ enfoncée, et cliquez sur la dernière page de la sélection.

**Pour sélectionner plusieurs pages non consécutives**, cliquez sur chacune d'entre elles à l'aide de la souris tout en maintenant la touche CTRL enfoncée.

**Pour sélectionner toutes les pages du lot**, activez la fenêtre *Traitement* et dans le menu *Edition*, choisissez *Sélectionner tout...*

*Remarque* : 1. Le résultat de la commande *Sélectionner tout* dépend de la fenêtre activée – *Traitement*, *Image* ou *Texte* : cette commande sélectionne toutes les pages d'un lot, tous les blocs d'une image, ou tout le texte dans l'éditeur de texte.

2. Il est plus pratique de sélectionner plusieurs pages lorsque la fenêtre *Traitement* est affichée avec ses paramètres (menu contextuel, élément *Afficher les propriétés de la page*).

### Travail sur les images

Le mécanisme de fonctionnement des boutons *Marquer les blocs* et *Reconnaître* est décrit dans la section *Barre d'outils principale – Scanner&Lire*, p. 28.

Comme pour la rotation, l'inversion et le nettoyage, ces opérations sont réalisées dans une **fenêtre active**. Si la fenêtre *Traitement* est active, la commande est effectuée sur les pages sélectionnées dans la fenêtre.

### Travail sur le texte

Les opérations effectuées avec les résultats de la reconnaissance (recherche, vérification orthographique, etc.) sont réalisées dans la page active du lot. Si aucune page n'est ouverte, l'opération est réalisée sur la première page de la sélection. Le programme vous invite ensuite à continuer. Pour de plus amples détails, consultez l'aide en ligne.

## Numérisation

L'exactitude de la reconnaissance dépend grandement de la qualité de l'image scannée. Cette section vous aidera à distinguer entre des textes de « bonne » et de « mauvaise » qualité et à apprendre à définir les options du scanner pour obtenir les meilleurs résultats possibles de votre système.

### Textes de bonne et de mauvaise qualité

#### Textes de « bonne qualité »

Les textes de bonne qualité possèdent très peu de caractères désalignés, collés ensemble, tachés ou déformés. Un individu n'a pas à faire d'effort visuel particulier pour lire de tels textes.

Exemples : le Guide que vous êtes en train de lire ; les textes imprimés sur une imprimante laser ; les magazines modernes et les productions couleurs en général.

#### Textes de « mauvaise qualité »

Les textes de mauvaise qualité sont difficiles à lire car ils ont un faible contraste, c'est-à-dire, une différence insuffisante entre les zones de blanc et de noir. De tels textes peuvent soit être trop sombres, soit trop clairs.

Exemples : un texte provenant d'une imprimante par point en mode brouillon, avec un ruban usé ; une copie carbone de machine à écrire ; un texte provenant d'une machine à écrire avec un ruban usé ; tout texte dont les caractères sont déformés, désalignés ou collés ensemble (Figure 16, p. 33).

*Lettres collées ensemble*

*Lettres désalignées*

*Lettres déformées et tachées*

*Figure 16. Exemple de textes de mauvaise qualité. Pour améliorer la qualité de reconnaissance du texte, essayez de régler correctement la luminosité de la numérisation. Augmentez la luminosité (rendez l'image plus claire) pour les lettres déformées, tachées et collées ensemble, et diminuez-la (rendez l'image plus sombre) pour les lettres mal alignées.*

### Numérisation de bonne qualité

Lorsque vous exécutez FineReader pour la première fois, les paramètres de numérisation par défaut sont automatiquement définis. Ces paramètres sont le type image en niveaux de gris (256 couleurs\*), avec un niveau de luminosité moyen et une résolution de 300 ppp. Il se peut que vous ayez parfois à modifier ces paramètres. Par exemple, si vous voulez enregistrer des images couleurs dans le texte en sortie, choisissez alors le mode de numérisation couleur. Pour

---

\* En fait, 16 couleurs suffisent, mais tous les pilotes de scanner ne prennent pas en charge cette fonctionnalité.


scanner les textes imprimés avec des petites polices (8 points et inférieures), réglez la résolution sur 400 à 600 ppp. Dans ces cas particuliers, les para-mètres par défaut ne sont pas appropriés et vous devez les définir manuellement.

### Pour définir les paramètres de numérisation :

1. Dans le menu *Outils*, sélectionnez la commande *Paramètres du scanner...*
2. Dans la boîte de dialogue qui s'affiche, définissez les options nécessaires.

*Remarque* : si dans la page de l'onglet *Numérisation* de la boîte de dialogue *Options (Outils>Options...)*, la commande *Afficher la boîte de dialogue TWAIN* est sélectionnée, vous pouvez utiliser la boîte de dialogue TWAIN de votre scanner pour définir les paramètres de numérisation. Cette boîte de dialogue et ces options sont décrites dans la documentation du scanner. Pour certains modèles de scanners l'option *Afficher la boîte de dialogue TWAIN* est désactivée par défaut. C'est de toute façon vous qui décidez en dernier lieu d'afficher ou non la boîte de dialogue TWAIN. Chaque mode de numérisation (avec ou sans la boîte de dialogue TWAIN) possède ses propres avantages. Ainsi, par exemple, si vous cochez la case *Afficher la boîte de dialogue TWAIN*, d'une manière générale vous pourrez avoir un aperçu de l'image scannée, et définir la hauteur et la largeur, la luminosité et le contraste de la numérisation. Lorsque vous n'utilisez pas la boîte de dialogue TWAIN, vous pouvez définir les options de numérisation dans la boîte de dialogue de FineReader. Lorsque vous scannez sans TWAIN, vous pouvez enregistrer les options de numérisation du lot, scanner en cycle avec une pause définie, etc.

### Nettoyage de l'image

Si vous ne pouvez pas vous débarrasser des « défauts » ou du fond sale de l'image, même en suivant les instructions de numérisation, vous pouvez nettoyer l'image. Pour ce faire, cliquez sur le bouton . Mais n'utilisez pas cette option de manière excessive, car en nettoyant l'image, les points, les virgules et les éléments fins des lettres risquent de disparaître, ce qui entraînerait inévitablement une faible qualité de reconnaissance.

### Numérisation de documents de taille importante

Pour scanner un grand nombre de pages, il est pratique d'utiliser l'alimentation automatique de documents (ADF). Tout particulièrement lorsque les pages du document sont séparées (non reliées). Si vous voulez scanner un livre en utilisant l'ADF, vous devrez nécessairement enlever sa reliure. Pour que le scanner prenne les pages depuis le magasin d'alimentation automatique de documents, dans la boîte de dialogue *Paramètres du scanner (Outils>Paramètres du scanner...)* activez la case à cocher *Utiliser l'alimentation automatique*. Ou, si vous travaillez avec la boîte de dialogue TWAIN, sélectionnez le mode *MultiPage*.

Si pour quelque raison que ce soit, vous ne pouvez pas utiliser l'ADF, vous pouvez scanner de la manière suivante : après la numérisation d'une page, le scanner fait une pause (vous spécifiez la durée de cette pause dans les options), pendant laquelle vous insérez la page suivante dans le scanner. Puis la numérisation se poursuit automatiquement.

Vous ne devez définir que la durée de la pause. Dans la boîte de dialogue *Paramètres du scanner* (*Outils > Paramètres du scanner...*), activez la case à cocher *Pause entre les pages*. Tapez ensuite la valeur choisie (en secondes). Il s'agit du temps qui s'écoulera entre le traitement de deux pages.

### Numérotation des pages

Lorsque vous démarrez la numérisation, le système vous invite à spécifier le mode de numérotation des pages scannées.

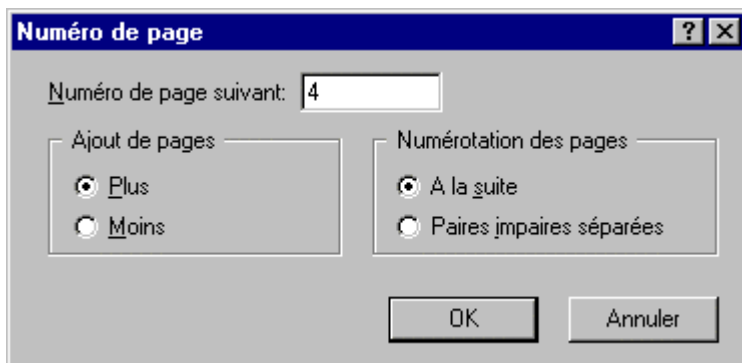



Figure 17. Choisissez le mode de numérotation des pages scannées.

Dans la boîte de dialogue *Numéro de page*, tapez le numéro de la première page scannée. Les pages peuvent être numérotées dans l'ordre ascendant ou descendant. Les numéros de pages peuvent être consécutifs ou non (pages uniquement paires ou impaires). Si vous scannez par exemple une pile de pages recto-verso (le texte est imprimé sur les deux faces) classée par numéro, en utilisant un mécanisme ADF, nous vous conseillons tout d'abord de scanner toutes les pages du même côté (recto ou verso), puis de scanner l'autre côté. Dans ce cas, vous ne sélectionnez que la numérotation paire ou impaire. La sélection d'ordre ascendant ou descendant dépend de la façon dont la pile a été placée dans le magasin ADF, c'est-à-dire si la numérisation commence par le début ou la fin du document.

## Ajout d'images dans le lot

1. Pour ajouter des fichiers images dans le lot, cliquez sur .
  2. Dans la boîte de dialogue *Ouverture*, sélectionnez une ou plusieurs images qui seront ajoutées au lot actuel.  
*Remarque* : dans cette boîte de dialogue, vous pouvez sélectionner simultanément plusieurs fichiers. Maintenez la touche MAJ enfoncée – pour sélectionner plusieurs fichiers consécutifs, ou CTRL – pour sélectionner des fichiers non consécutifs.
- Vous pouvez placer les fichiers dans le lot lors de leur ouverture (cochez la case *Placer dans le lot*, dans la boîte de dialogue *Ouverture*).

## Ouverture d'une image couleur ou grise dans FineReader

### A propos des images couleur et en niveaux de gris

Lorsque FineReader ouvre des images couleur et en niveaux de gris, il crée une copie noir et blanc qui est reconnue. C'est pourquoi, lorsque vous éditez le texte reconnu d'une page couleur ou en niveaux de gris, vous pouvez voir sa copie noir et blanc dans la fenêtre Zoom.

### Comment enregistrer une image noir et blanc, à la place d'une image en niveaux de gris ou couleur

Pour économiser l'espace disque, vous pouvez enregistrer des images couleur ou en niveaux de gris, en tant qu'images noir et blanc. Pour ce faire, cochez la case *Ne conserver que les images noir et blanc dans le lot* (onglet *Outils>Options...>Numérisation*). Lors du chargement ou de l'ajout d'une image couleur ou en niveaux de gris dans le lot, FineReader l'enregistrera alors en noir et blanc.

### Message de résolution incorrecte

Quelquefois, lors du chargement d'une image dont la résolution est inférieure à 100 ppp ou supérieure à 800 ppp, le programme vous suggère de modifier la résolution afin qu'elle soit plus appropriée à la reconnaissance.

Voici les raisons pour lesquelles de telles images sont produites :

- Il existe des formats d'images (par exemple, BMP), dont la résolution peut être lue différemment.
- Format d'image non standard.
- Options de numérisation incorrectes.
- Pour la majorité des scanners à main, la véritable résolution de numérisation des images en niveaux de gris ou couleur n'est pas 300 ou 400 ppp comme spécifié, mais beaucoup moins (75-100 ppp).

Suivant l'origine de l'image, vous pouvez modifier sa résolution ou la laisser inchangée.

### Rappelez-vous qu'une résolution d'image incorrecte peut conduire aux problèmes suivants :

- Le programme ne sera pas en mesure de marquer les blocs de l'image.
- Le programme détectera la taille de la police de manière incorrecte.

## Analyse de page (marquage des blocs)




Avant la reconnaissance, le programme doit « savoir » quelles parties de l'image doivent être reconnues. Pour ce faire, le programme analyse l'image et marque les blocs en tant que texte, images et tableaux. Une page peut être analysée à la fois automatiquement et manuellement. FineReader peut analyser avec succès du texte dont la disposition est mixte. Pour analyser une image, cliquez sur le bouton *Marquer les blocs* dans la barre d'outils *Scanner&Lire*. Pour comprendre le fonctionnement de ce bouton, consultez la section Barre d'outils principale – *Scanner&Lire*, p. 22.



Figure 18. Barre d'outils Outils.

Pour modifier les résultats d'analyse de page automatique, utilisez les outils appropriés (Figure 18).

### Pour marquer et modifier des blocs manuellement

**Pour marquer un bloc**, sélectionnez l'outil  (il est sélectionné par défaut) et cliquez sur l'angle supérieur gauche du bloc que vous voulez marquer. Faites glisser diagonalement jusqu'à ce que le cadre entoure les parties désirées de l'image. Relâchez le bouton de la souris. La partie de l'image sélectionnée sera comprise dans le cadre. C'est un bloc de texte par défaut, donc, il sera reconnu.

Selon l'image entourée, un bloc peut être du type suivant : image, tableau, code à barres\* ou texte dont la reconnaissance est inutile.

**Pour modifier le type d'un bloc**, cliquez sur le bloc à l'aide du bouton **droit** de la souris et, dans le menu qui s'affiche, sélectionnez la commande *Type de bloc*, puis le type de votre choix.

**Pour modifier la taille d'un bloc**, placez le pointeur de la souris sur le cadre d'un bloc, jusqu'à ce qu'il se transforme en une flèche à deux têtes ; cliquez ensuite sur le cadre du bloc et faites-le glisser.

### Marquage automatique de blocs sur une partie de page

Pour marquer des blocs sur une partie de page, sélectionnez-la à l'aide de la souris (comme vous sélectionneriez un nouveau bloc), tout en maintenant la touche CTRL enfoncée. La partie sélectionnée sera analysée automatiquement.

Vous pouvez également marquer le bloc de manière habituelle à l'aide de la souris, puis sélectionner *Analyser le bloc* dans le menu contextuel.

---

\* Uniquement dans FineReader 4.0 Professional et Handprint.

## Spécification de la structure des tableaux

Lorsque le système marque des blocs sur la page, il identifie automatiquement les tableaux et les analyse (il distingue les cellules). Pour la reconnaissance précise d'un nombre important de tableaux du même type, vous pouvez définir leurs caractéristiques :

### **Vous pouvez spécifier les caractéristiques suivantes :**

- Si le tableau est « régulier » ou non.
- Si le tableau est linéaire ou non.
- Si la division des cellules coïncide avec les lignes noires de l'image originale.

Le tableau B (voir ci-dessous) est « régulier ». Les lignes noires d'un tel tableau traversent **tout** le tableau (par contraste avec le tableau A – qui est « irrégulier »).

Le tableau A (voir ci-dessous) est linéaire, car il comporte des cellules contenant *une ligne* de texte chacune. Le tableau B est non linéaire, car il comporte des cellules contenant plus d'une ligne de texte.

Température	
Celsius	Kelvin
-273	0
100	373

*Tableau A. Linéaire, irrégulier*

	Celsius
Point d'ébullition	100
Point de congélation	0

*Tableau B. Non linéaire, régulier.*

Les tableaux linéaires peuvent ne pas posséder de lignes noires – verticales ou horizontales (Tableau C).

Kilomètres	Miles
1	0.62
5	3.1
10	6.2

*Tableau C. Tableau linéaire sans lignes noires.*

### **Pour spécifier le type de structure d'un tableau :**

1. Dans le menu *Outils*, sélectionnez la commande *Options...*
2. Dans la page de l'onglet *Analyse de page* de la boîte de dialogue *Options*, activez les cases à cocher requises dans le groupe *Tableaux*.

**Important :**

1. La structure du tableau devrait être spécifiée dans les options du système avant la reconnaissance.
2. Une spécification incorrecte de la structure d'un tableau peut affecter la qualité de la segmentation.
3. Si vous voulez réaliser la saisie d'un très grand nombre de tableaux de différentes structures, ne sélectionnez pas d'élément dans le groupe *Tableaux* de la page de l'onglet *Analyse de page* (boîte de dialogue *Options*).

**Edition de blocs dans un tableau****Quels tableaux devraient être analysés manuellement ?**

Si un tableau ne possède pas de lignes noires horizontales, mais possède plus d'une ligne de texte par cellule (il n'est pas « linéaire ») et des intervalles égaux entre ses lignes (Figure 19, p. 39), il est préférable de l'analyser manuellement.

Placez un tel tableau dans un bloc séparé et marquez-le comme un *Tableau* ; puis ajoutez des lignes verticales et horizontales à l'aide des boutons de la barre d'outils *Outils* (Figure 18, p. 37).

1.	Jim Johnson	1990.
2.	Mary Lay	1996.
3.	Boris Namin	1997.

*Figure 19. Tableau nécessitant une analyse manuelle. Le tableau n'est pas linéaire et ne possède pas de lignes horizontales.*

Vous pouvez également analyser le bloc de tableau automatiquement (cliquez à l'aide du bouton **droit** de la souris sur le bloc comportant un tableau et, dans le menu qui s'affiche, sélectionnez la commande *Analyser le tableau*), puis modifiez le résultat manuellement.

**Pour analyser un tableau manuellement :**

1. Activez la fenêtre *Image*, en cliquant sur celle-ci à l'aide de la souris.
2. Si le tableau situé dans la fenêtre est trop petit, augmentez l'échelle. Pour ce faire, cliquez à l'aide du bouton **droit** de la souris dans la fenêtre *Image* et dans le menu qui s'affiche, sélectionnez la commande *Echelle*, puis l'élément visé.
3. Si le tableau n'est pas marqué, marquez-le à l'aide de la souris (consultez la section *Pour marquer et modifier des blocs manuellement*, p. 37).
4. Puis utilisez les outils suivants :



*Ajouter une ligne verticale*



*Ajouter une ligne horizontale*




*Supprimer une ligne*



*Remarque :* 1. Pour supprimer une ligne dans le tableau, vous pouvez aussi la déplacer jusqu'à ce qu'elle coïncide avec le cadre du tableau.

2. Pour modifier la direction de la ligne que vous ajoutez au tableau (de la verticale à l'horizontale et vice versa), appuyez sur la touche MAJ.

5. Lorsque l'ajout ou la suppression des lignes est terminé, cliquez sur  ou appuyez sur la touche ECHAP pour que le pointeur de la souris revienne en mode d'affichage normal.
6. Vous pouvez également fusionner ou fractionner les cellules d'un tableau (*Edition>Fusionner les cellules d'un tableau* et *Fractionner les cellules d'un tableau*, respectivement).

*Remarque :* pour sélectionner plusieurs cellules, utilisez la souris (ou la flèche droite-gauche) et la touche MAJ.

## Reconnaissance

Le but de la reconnaissance est de transformer l'image en entrée (scannée) en texte. En d'autres termes, elle remplace l'image de chaque caractère entré par son code informatique. Avant de commencer la reconnaissance, assurez-vous d'avoir défini les paramètres de reconnaissance appropriés dans la barre d'outils Reconnaissance (Figure 20) : langue de reconnaissance et type de texte.



Figure 20. Barre d'outils Reconnaissance.

### Langue de reconnaissance

- Pour reconnaître des textes anglais, espagnols, italiens, allemands, russes, anglo-russes, ukrainiens, français, suédois, danois et néerlandais, polonais, finnois et norvégiens, mais aussi des textes contenant des chiffres arabes, sélectionnez la langue dans la liste de la barre d'outils Reconnaissance.

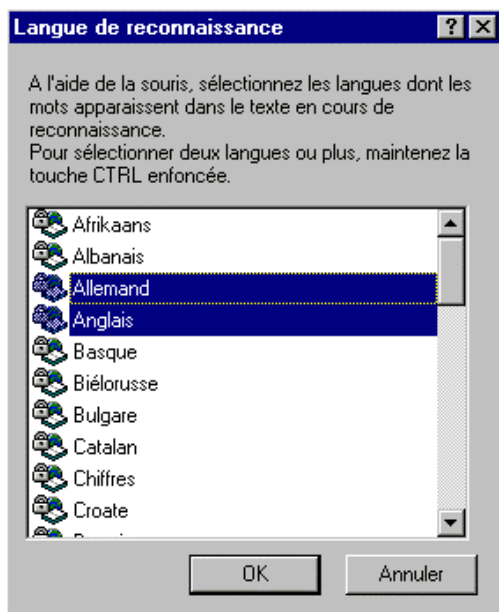


Figure 21. Boîte de dialogue *Langue de reconnaissance*. Sélection des langues de reconnaissance de textes anglo-allemands.

- Pour reconnaître d'autres langues et combinaisons de langues, sélectionnez la valeur *Autre...* et continuez avec l'élément suivant.  
La boîte de dialogue *Langue de reconnaissance* (Figure 21) s'affiche, dans laquelle vous pouvez sélectionner plusieurs langues, dont les mots sont susceptibles d'apparaître dans votre texte.

Après avoir cliqué sur *OK* dans cette boîte de dialogue, la combinaison des langues que vous avez choisies s'affiche dans la barre d'outils *Reconnaissance* (Figure 22).

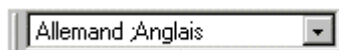


Figure 22. Liste des langues de reconnaissance. Combinaison de langues pour la reconnaissance de textes anglo-allemands.

## Type de texte

Les textes imprimés sur une imprimante matricielle en mode brouillon ou sur une machine à écrire ont des caractéristiques spécifiques (voir les images ci-dessous) par comparaison aux caractères typographiques standard. Les symboles imprimés sur une imprimante matricielle sont

formés de points séparés (Figure 24). Les symboles typographiés (Figure 25), ont une épaisseur de trait pratiquement constante.

Pour la majorité des textes, le type de texte est détecté automatiquement : ligne *Automatique* est sélectionnée dans la liste *Type de texte* (Figure 23). Comme pour les textes dactylographiés et ceux d'imprimante matricielle, pour augmenter la qualité de reconnaissance de tels textes, sélectionnez une valeur spéciale dans la liste *Type de texte* – *Machine à écrire* ou *Imprimante matricielle*, respectivement.

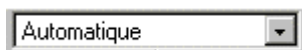


Figure 23. Ligne *Automatique* sélectionnée dans la liste *Type de texte*.

## hardware

Figure 24. Fragment de page, imprimé sur une imprimante matricielle en mode brouillon. Vous pouvez voir que les lettres sont constituées de points séparés.

## direction

Figure 25. Fragment de page dactylographiée.

### Enregistrement de la disposition et de la mise en forme du document

Avant la reconnaissance, vous devez spécifier la manière dont vous souhaitez enregistrer les résultats de reconnaissance : c'est-à-dire, l'enregistrement ou non de la mise en forme des polices, de la disposition du texte et des images. Les options de mise en forme du texte et de disposition sont spécifiées dans la page de l'onglet *Mise en forme* de la boîte de dialogue *Options* (*Outils>Options...*).

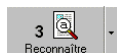
- Si vous n'avez besoin que du texte lui-même, et que sa disposition ou sa mise en forme vous importent peu (par exemple, si vous voulez l'insérer dans un autre texte), sélectionnez l'option *Supprimer la mise en forme*.
- Si vous souhaitez enregistrer les interruptions ou les décrochements des lignes, sélectionnez l'option *Utiliser les espaces pour afficher la mise en forme*.
- Si les décrochements de lignes ne vous intéressent pas, mais que vous voulez enregistrer la mise en forme des polices, sélectionnez l'option *Enregistrer la forme des lettres et la police*.
- Si vous voulez enregistrer la disposition et la mise en forme des polices dans son entier, sélectionnez l'option *Enregistrer la mise en forme avec encadrés*. Cette option vous permettra d'obtenir un résultat proche de l'original.
- Les cadres peuvent cependant compliquer l'édition (voir ci-dessus). C'est pourquoi, si vous projetez d'éditer le résultat dans votre éditeur de texte, sélectionnez le mode d'enregistrement le plus adéquat pour l'édition *Enregistrer la forme des lettres, la police et les colonnes*. Les cadres seront le plus souvent utilisés pour les images.

### Enregistrement des images dans un texte reconnu

Pour enregistrer des images dans un texte reconnu, dans la page de l'onglet *Mise en forme* de la boîte de dialogue *Options* (*Outils>Options...*) activez les cases à cocher *Enregistrer les images*.

*Remarque* : vous pouvez également enregistrer des images dans des fichiers séparés. Pour ce faire, activez le bloc d'image sur l'image et dans le menu *Fichier*, sélectionnez la commande *Enregistrer la copie de l'image*. Dans la boîte de dialogue qui s'affiche, sélectionnez l'option *Enregistrer le bloc actif uniquement*.

## Exécution de la reconnaissance



Pour démarrer la reconnaissance, cliquez sur le bouton *Reconnaître* dans la barre d'outils *Scanner&Lire*.

Si une image est ouverte, l'opération par défaut du bouton est la reconnaissance de l'image ouverte. Si une ou plusieurs pages sont sélectionnées dans la fenêtre *Traitement*, le programme commence par la reconnaissance des pages sélectionnées.

Vous pouvez modifier le mode du bouton. Pour ce faire, cliquez sur la flèche située à droite du bouton et dans le menu qui s'affiche, sélectionnez l'élément de votre choix.

*Remarque* : lors de la reconnaissance, le système affiche une boîte de dialogue qui vous avertit que certains paramètres sont incorrectement définis. Les messages sont formulés de la façon suivante : « Luminosité trop faible (ou trop élevée) », « Type de texte incorrect », etc. Ces messages vous aident à vérifier les paramètres, mais vous pouvez les ignorer ou ne pas les afficher si vous le souhaitez (*Outils>Options...>Reconnaissance>Afficher les conseils en cours de reconnaissance*).

## Reconnaissance en tâche de fond

FineReader vous permet d'exécuter la reconnaissance et d'éditer les pages reconnues simultanément, ceci accélère considérablement le traitement des documents sur le même ordinateur. Démarrez la reconnaissance en tâche de fond (*Scanner&Lire>Reconnaître en tâche de fond*).

Dans la barre d'états, l'indicateur s'affiche. Si les paramètres de la page sont affichés dans le lot (menu contextuel, élément *Afficher les propriétés de la page*), la ligne *Reconnaissance en tâche de fond* s'affiche près de la page en cours de reconnaissance. Vous pouvez aussi voir les puces qui s'affichent en vis à vis des pages traitées dans les colonnes *Blocs*, puis – *Texte*.

Tandis que la reconnaissance est en progrès, vous pouvez ouvrir et modifier les pages reconnues.

## Vérification et édition des résultats

La barre d'outils *Edition* (Figure 26, p. 43) vous permet d'éditer le texte reconnu.

La barre d'outils *Edition* est désactivée par défaut. Pour l'afficher à l'écran, cliquez à l'aide du bouton **droit** de la souris sur l'une des barres d'outils et sélectionnez la commande *Edition* dans le menu contextuel.

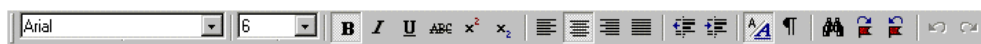


Figure 26. Barre d'outils *Edition*.

### Vérification des résultats de reconnaissance

L'une des fonctions utiles de l'éditeur de texte de FineReader est le système de vérification de l'orthographe intégré pour les textes dans les langues suivantes et leurs combinaisons : russe, anglais, allemand, français, ukrainien, espagnol, italien, néerlandais, danois, suédois, polonais, norvégien, portugais, tatar\*.

Le texte original peut contenir des erreurs d'orthographe et des erreurs d'impression qui sont transférées dans le texte reconnu. En plus des erreurs du texte original, le texte reconnu peut contenir des mots incertains et non reconnus (des mots contenant des caractères incertains et non reconnus). Pour que le correcteur d'orthographe s'arrête sur de tels mots, cochez l'option correspondante (*Outils>Options...>onglet Vérification des résultats*).

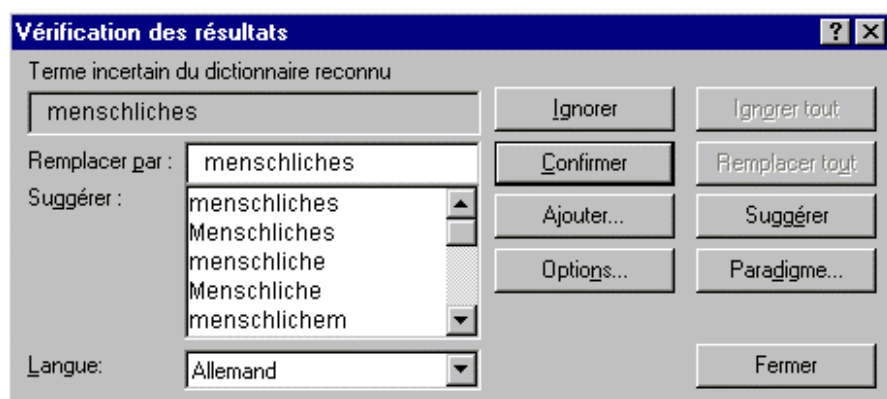
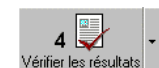


Figure 27. Boîte de dialogue Vérification des résultats.

### Pour vérifier les résultats de reconnaissance :



1. Cliquez sur le bouton *Vérifier les résultats*.  
La boîte de dialogue *Vérification des résultats* (Figure 27) affiche une erreur possible et la fenêtre *Zoom* affiche l'image originale.
2. Sélectionnez l'une des options suivantes pour le mot :
  - Cliquez sur *Ignorer* pour que le mot reste tel quel.
  - Cliquez sur *Ignorer tout* pour ignorer toutes les instances du mot.
  - Cliquez sur *Ajouter* pour ajouter le mot trouvé au dictionnaire.

\* Le tatar fait l'objet d'une distribution spéciale.




- Cliquez sur *Paradigme...* pour visualiser le paradigme (somme de toutes les formes du mot) <sup>1</sup> du mot sélectionné dans la liste *Suggérer*.
- Cliquez sur *Suggérer...* pour obtenir des variantes de remplacement pour le mot édité dans la case *Remplacer par*.
- Cliquez sur *Options...* pour définir les options de vérification<sup>2</sup>.
- Cliquez sur *Fermer* pour refermer la boîte de dialogue.

### Recherche des erreurs

Pour vérifier et modifier les résultats de reconnaissance, vous pouvez utiliser les boutons  et  pour vous déplacer vers les mots suivants et précédents, respectivement. Vous pouvez également utiliser les touches directes : F4 et MAJ+F4 respectivement.

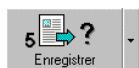
### Visualisation des résultats de reconnaissance

Les résultats de reconnaissance par défaut sont affichés en mode brouillon, tout le texte est affiché avec une seule taille de police. Pour afficher la police à sa taille réelle, cliquez sur . Vous pouvez modifier la taille de la police de l’affichage en mode brouillon (*Outils>Options>Général>Taille de la police en mode brouillon*).

## Enregistrement des résultats de la reconnaissance

Vous pouvez enregistrer les résultats de reconnaissance dans un fichier, les envoyer vers une application (MS Word, MS Excel, etc.) sans l’enregistrer sur le disque, les copier dans le Presse-papiers ou les envoyer par e-mail. Vous pouvez choisir d’exporter toutes les pages reconnues ou uniquement celles qui sont sélectionnées.

### Pour exporter les résultats de reconnaissance :



1. Cliquez sur la flèche située à droite du bouton *5-Enregistrer* et dans le menu qui s’affiche, sélectionnez la commande *Assistant Exportation...*
2. Dans la boîte de dialogue affichée, (Figure 28) sélectionnez l’emplacement dans lequel vous voulez exporter les résultats de reconnaissance.  
Si vous ne voulez exporter que les pages sélectionnées dans la fenêtre *Traitement*, activez le bouton radio en position *N’enregistrer que les pages sélectionnées*.
3. Cliquez sur *Terminé*.

<sup>1</sup> Pour le russe, l’anglais, l’allemand, le français, l’ukrainien, l’espagnol et l’italien.

<sup>2</sup> Pour définir les options de vérification de l’orthographe, sélectionnez l’onglet *Vérification des résultats* de la boîte de dialogue *Options* (*Outils>Options...*).

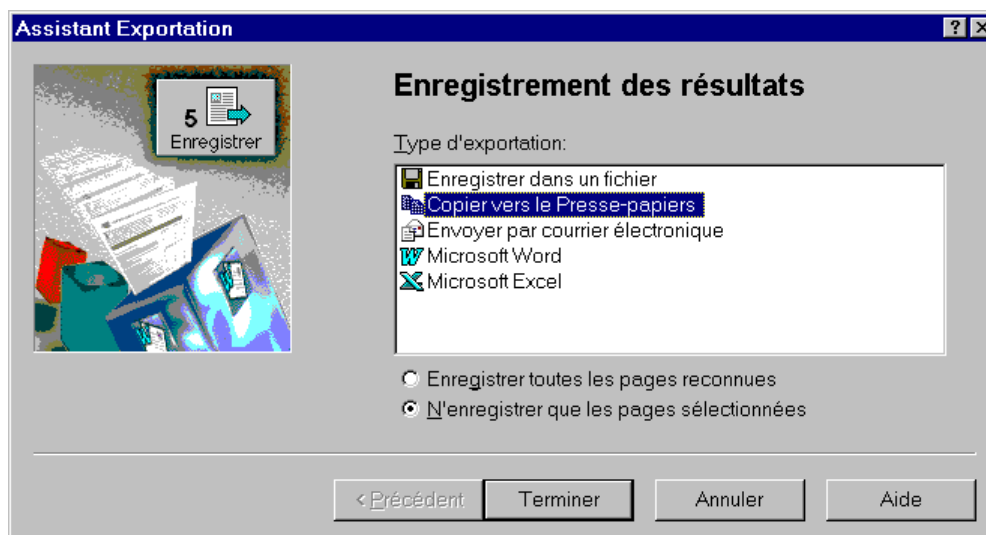


Figure 28. Assistant Exportation.

*Remarque :* si lors de votre travail sur un nouveau lot (sans nom), ce travail était annulé (par exemple, l'alimentation a été coupée), la prochaine fois que vous exécuterez FineReader, il s'ouvrira sur ce même lot. Vous ne perdez que les informations de la page sur laquelle vous travailliez. Les informations sur les autres pages seront récupérées.

## Enregistrement d'un lot pour une utilisation future

Si vous projetez de continuer à travailler avec le lot dans la prochaine session, vous pouvez l'enregistrer (*Fichier > Enregistrer le lot...*). Par défaut, un dossier spécial est créé dans le dossier c:\Program Files\ABBYY FineReader\Packets pour chaque lot. Le titre du dossier est le titre du lot. Les fichiers suivants sont stockés dans le dossier du lot :

- Fichier de lot (BatchName.pac).
- Fichier image (BatchName\_number.TIF) et/ou texte (BatchName\_number.FRF) pour chaque page.
- Fichier avec des blocs (BatchName\_number.BOX) (uniquement pour les pages contenant un fichier image).
- Fichiers possédant des options de numérisation du lot (BatchName.SCN).
- Fichier d'options du lot (BatchName.opt).
- Le dossier Formulaires possède également un fichier de base de données de langues textlang.dat (pour FineReader Professional et Handprint).

## Apprentissage de nouveaux caractères

FineReader peut reconnaître des textes de n'importe quelle qualité, imprimés avec n'importe quelle police. Les textes de bonne et de moyenne qualité, et les polices standard peuvent être reconnues sans formation préliminaire. Les textes de mauvaise qualité et les polices inhabituelles peuvent être reconnus après un apprentissage spécial.

Pour que le système reconnaisse tout le texte correctement, formez-le sur un échantillon de quelques pages. Un motif sera établi, et le système peut utiliser ce dernier pour reconnaître le reste du texte.

L'apprentissage se fait lors de la reconnaissance dans un mode spécial. Le résultat de l'apprentissage est un appariement de l'image du symbole et de son nom, configuré dans la mémoire ROC. La somme de toutes les paires, créées lors de l'apprentissage, est appelé le « motif ». Ce motif est utilisé pour la reconnaissance du reste du texte.

Le motif utilisateur ne peut être utilisé que pour reconnaître du texte de la même taille que la police sur laquelle il a été formé.

## Formation du système

Pour reconnaître des textes de mauvaise qualité ou des polices inhabituelles, effectuez les opérations suivantes :

1. Créez un motif (nommez-le).
2. Définissez le mode *Lire&Apprendre* (cochez l'élément correspondant dans la page de l'onglet *Reconnaissance* de la boîte de dialogue *Options*).
3. Lors de l'apprentissage au programme de polices inhabituelles ou différentes des polices cyrilliques ou latines (par exemple, grecques, irlandaises), désactivez la fonction *Omnifont*.
4. Faites reconnaître au programme 1 à 2 pages en guise de formation.  
*Remarque* : si vous souhaitez reconnaître des textes dont la langue n'est pas prise en charge dans FineReader, imprimez n'importe quel texte avec les caractères typographiques (ni gras, ni italiques) qui serviront à la formation.
5. Editez le motif (*Outils>Editeur de motif...*).
6. Annulez le mode *Lire&Apprendre*.
7. Démarrez la reconnaissance du reste du texte avec le motif associé.  
*Remarque* : lorsque vous passez à la reconnaissance de textes imprimés à l'aide d'autres polices, n'oubliez pas de désactiver le motif (sélectionnez *Aucun* dans la liste *Motif utilisateur actuel* dans la page de l'onglet *Reconnaissance* de la boîte de dialogue *Options*).

Nous allons décrire ces opérations en détail plus tard.

### Création et association d'un motif

1. Choisissez *Options* dans le menu *Outils*.
2. Dans l'onglet *Reconnaissance* de la boîte de dialogue *Options* (*Outils>Options*), cliquez sur le bouton *Motifs* dans le groupe *Apprentissage*.
3. Dans la boîte de dialogue *Motifs*, cliquez sur *Nouveau*.
4. Nommez le motif.



5. Avant de commencer une reconnaissance avec apprentissage, assurez-vous que le motif créé est sélectionné dans la case *Motif utilisateur actuel* et que la case à cocher *Lire&Apprendre* est activée.

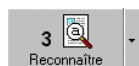
*Remarque :* vous ne devez désactiver l'option omnifont que lorsque vous formez le programme à utiliser des polices inhabituelles ou différentes des polices cyrilliques ou latines (par exemple, grecques, irlandaises).



Figure 29. Boîte de dialogue Apprentissage manuel.

#### Pour démarrer la reconnaissance avec apprentissage :

1. Vérifiez si la case à cocher *Lire&Apprendre* est sélectionnée dans la page de l'onglet *Reconnaissance* de la boîte de dialogue *Options (Outils>Options...)*.
2. Cliquez sur le bouton *Reconnaître*.  
Le programme démarre la reconnaissance et affiche l'image d'une lettre non reconnue dans la boîte de dialogue *Apprentissage de motif* (Figure 29).
3. Nommez la lettre et cliquez sur *Former*.  
*Remarque :* lors de l'apprentissage d'un motif, assurez-vous que les images des lettres majuscules sont nommées avec des lettres majuscules et celles des lettres minuscules – avec des lettres minuscules.



4. Le rectangle de sélection situé dans la partie supérieure de la boîte de dialogue doit contenir **un caractère entier**. S'il contient une partie de lettre ou plus d'un caractère, utilisez les boutons *Droite* et *Gauche* pour déplacer le rectangle afin qu'il encadre une lettre entière.
5. Si vous avez fait une erreur lors de l'apprentissage, vous pouvez cliquer sur le bouton *Précédent*, le rectangle de sélection retournera à sa position précédente, et la dernière paire « image – caractère » formée sera supprimée du motif.  
*Remarque* : le bouton *Précédent* fonctionne à l'intérieur d'un mot.
6. Vous pouvez modifier le motif après qu'il ait été formé. Consultez la section **Modification d'un motif**.  
*Remarque* : lors de l'apprentissage d'un motif, rappelez-vous les restrictions suivantes : le système de reconnaissance ne fait pas la distinction entre certains caractères et connecte leurs images à un seul et même caractère. Ainsi, par exemple, les images de l'apostrophe droite ('), ouvrante (') et fermante (') sont stockées dans le motif en tant qu'images d'une apostrophe droite. C'est pourquoi vous ne verrez jamais d'apostrophe ouvrante ou fermante dans un texte reconnu, bien que vous ayez tapé ces caractères lors de la reconnaissance avec apprentissage.

#### **Pour annuler le mode Lire&Apprendre :**

- Dans la page de l'onglet *Reconnaissance* de la boîte de dialogue *Options* (*Outils>Options...*), désélectionnez la case à cocher *Lire&Apprendre*.

#### **Modification d'un motif**

Avant d'exécuter la reconnaissance en utilisant le nouveau motif, vous devez le vérifier et le modifier si nécessaire. Ceci afin de minimiser les erreurs de reconnaissance, causées par l'apprentissage incorrect d'un motif par le système. Votre motif ne doit contenir que des symboles entiers ou des ligatures (combinaisons de deux ou trois lettres, qui ne peuvent pas être dissociées pendant l'apprentissage, et sont donc considérées comme un seul caractère).

#### **Pour modifier un motif :**

1. Dans le menu *Outils*, sélectionnez la commande *Motifs*.
2. Sélectionnez le motif dans la liste, et cliquez sur *Modifier* pour afficher la liste des caractères (Figure 30).
3. Supprimez les paires incorrectes et refermez la boîte de dialogue.

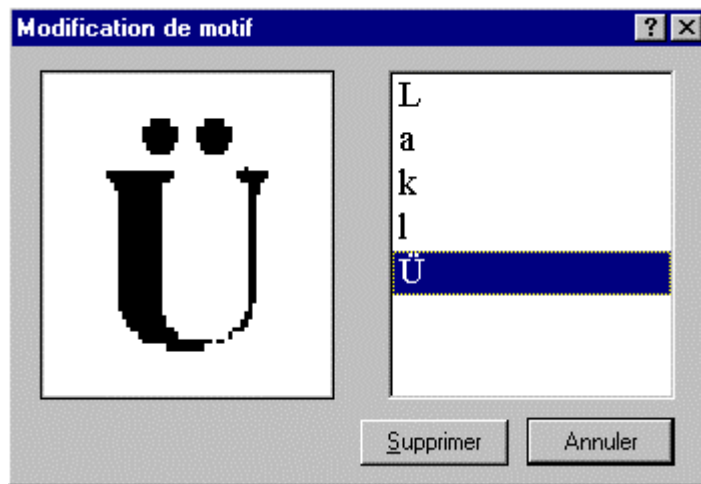


Figure 30. Boîte de dialogue Modification de motif.

### Apprentissage des ligatures

Certains groupes de deux ou trois caractères du texte peuvent être collés ensemble à cause de leur forme. Si pendant l'apprentissage, vous ne pouvez pas déplacer le rectangle de sélection afin qu'il encadre un symbole entier et aucune partie des symboles adjacents, vous pouvez éduquer cette combinaison de caractères comme étant un unique symbole. De telles combinaisons de deux ou trois symboles qui ne peuvent pas être dissociées sont appelées ligatures. Par exemple, ff, ffi, ffl, etc. Lors de l'apprentissage du système, vous pouvez sélectionner la ligature nécessaire dans la liste *Symbole*. Si vous ne trouvez pas la ligature nécessaire dans la liste, créez-en une nouvelle. Un motif ne peut contenir plus de 100 nouveaux symboles. Ne créez pas trop de nouvelles ligatures, car cela peut affecter la qualité de la reconnaissance.

*Remarque :* si vous apprenez au programme à reconnaître de nouveaux caractères qui ne peuvent pas être tapés à partir du clavier, utilisez une combinaison de lettres double pour les nommer ou copiez ces caractères depuis le tableau des caractères. Exemple, une lettre a accentuée peut être apprise en tant que #a. Répondez *Oui*, lorsque le programme vous demande si vous voulez créer une nouvelle ligature.

## Langues de FineReader

Pour reconnaître des textes en russe, anglais, allemand, français, ukrainien, espagnol, italien et toute autre langue de la liste des langues reconnues, sélectionnez la valeur correcte dans la liste des langues de la barre d'outils Reconnaissance. Si vous ne trouvez pas la langue de votre choix dans la liste, sélectionnez la valeur *Autre...* et créez une nouvelle combinaison (pour de plus amples détails, consultez la section Langue de reconnaissance, p. 40). La liste des langues reconnues se trouve dans l'Annexe, p. 56.

### Langues de l'utilisateur et groupes de langues (pour FineReader 4.0 Professional et Handprint)

FineReader reconnaît 53 langues, nombre suffisant pour la plupart des utilisateurs (la liste des langues reconnues se trouve dans l'Annexe, p. 56). Mais dans certains cas, il se peut que vous deviez créer vous-même de nouvelles langues.

#### A quel moment créer une nouvelle langue :

- Pour attacher un fichier texte ou une expression régulière à un dictionnaire. (Utile lors de la création de nouvelles langues de reconnaissance des champs de formulaires.)
- Pour reconnaître des textes dont les langues ne sont pas reconnues par le programme.
- Lorsque vous utilisez souvent telle ou telle combinaison de langues, par exemple, anglais-espagnol.

Vous pouvez attacher un dictionnaire à une nouvelle langue. Il doit se présenter sous forme de fichier texte en code ANSI. Dans FineReader Handprint, vous pouvez aussi attacher une colonne de base de données en tant que dictionnaire.

Pour créer une nouvelle langue ou un groupe de langues, ouvrez la boîte de dialogue *Editeur de langue* (*Outils> Editeur de langue...*).

Un exemple de création d'une nouvelle langue est donné en Annexe, p. 54.

### Traitement de lots sur réseau (pour FineReader 4.0 Professional et Handprint)

Vous pouvez traiter un lot sur plusieurs ordinateurs **simultanément**. Cela augmente considérablement la vitesse de saisie de documents importants. Chaque ordinateur impliqué dans le traitement du lot doit posséder une copie installée du programme.

L'entrée de texte dans l'ordinateur implique la numérisation, l'analyse de page, la reconnaissance, l'édition et l'exportation. Lors du traitement simultané d'un seul lot sur plusieurs ordinateurs, vous pouvez répartir le travail afin qu'une opération ou combinaison d'opérations puisse être réalisée sur chaque ordinateur. Par exemple, un ordinateur, équipé d'un scanner et d'un mécanisme ADF, peut être utilisé pour la numérisation, tandis qu'un autre ordinateur (ou plusieurs) effectue la reconnaissance des pages scannées et/ou modifie et exporte les résultats de la reconnaissance.

Si la reconnaissance du lot est effectuée simultanément sur plusieurs ordinateurs, ils se partagent le travail automatiquement : tant que le lot comprend des pages non reconnues, les ordinateurs ne restent pas inactifs.

## Travail avec un lot sur réseau

1. Créez un nouveau lot (à moins qu'il n'existe déjà) et enregistrez-le dans un dossier qui soit accessible à la lecture et à l'écriture par tous les ordinateurs impliqués dans le traitement du document.
2. Fermez le lot pour enregistrer ses options.
3. Exécutez FineReader et ouvrez le lot sur les ordinateurs qui travailleront dessus.
4. Démarrez la numérisation sur l'ordinateur équipé d'un scanner et d'un mécanisme ADF.
5. Démarrez la reconnaissance en mode *Lire&Attendre* sur l'ordinateur dédié à la reconnaissance (*Scanner&Lire > Reconnaître toutes les pages non reconnues et attendre*), puis sur les ordinateurs restants.
6. Les premières pages étant reconnues, vous pouvez commencer à les éditer sur n'importe quel ordinateur libre, ou sur ceux pour lesquels la reconnaissance s'exécute en tâche de fond, sans attendre les pages restantes.
7. Les pages reconnues (toutes ou celles sélectionnées) peuvent être exportées vers un fichier ou envoyées vers une autre application.

*Remarque* : afin de surveiller le traitement du lot, cliquez à l'aide du bouton **droit** de la souris sur la fenêtre de traitement et dans le menu contextuel affiché, sélectionnez la commande *Afficher les paramètres de la page*. Cette visualisation du lot facilitera le travail en réseau effectué sur le lot (Figure 15, p. 31).

## Mode Lire&Attendre

En mode *Lire&Attendre*, la reconnaissance se poursuit automatiquement lorsque de nouvelles pages non reconnues s'affichent dans le lot.

### Pour exécuter la reconnaissance en mode Lire&Attendre :

- Cliquez sur la flèche située à droite du bouton *Scanner&Lire* et dans le menu qui s'affiche, sélectionnez la commande *Reconnaître toutes les pages non reconnues et attendre*.

# Annexes

## Amélioration de la qualité de la reconnaissance : astuces et conseils

### Principes fondamentaux

Une des raisons pour lesquelles un texte comporte de nombreux caractères non reconnus peut être la mauvaise qualité de l'image source. Définissez la luminosité correcte, le type de résolution du texte et numérisez l'image à nouveau. Si l'image, comporte des défauts, nettoyez-la.

Une mauvaise reconnaissance de texte peut être due à des paramètres de reconnaissance incorrects, par exemple, la sélection d'une langue incorrecte ou d'un trop grand nombre de langues. Les erreurs de reconnaissance systématiques dans chaque ligne peuvent, par exemple, être causées lors de la reconnaissance d'un texte anglais avec la langue anglo-russe.

### Ajout de termes au dictionnaire

Le système utilise un dictionnaire pendant la reconnaissance. L'ajout des termes et des abréviations fréquemment utilisés au dictionnaire améliore la précision de la reconnaissance. Vous pouvez ajouter des mots au dictionnaire du correcteur orthographique (*Outils>Afficher les dictionnaires...*) ou vous pouvez créer une nouvelle langue et lui attacher des termes en tant que fichier texte (en code ANSI) ou colonne de base de données.

Pour de plus amples informations, consultez la section Langues, p. 56 et l'aide en ligne.

### Apprentissage

Si vous devez reconnaître des textes de taille importante de très mauvaise qualité ou avec des polices inhabituelles, vous pouvez former le système sur un échantillon de 2 à 3 pages. Le système reconnaîtra ensuite le reste du texte précisément. L'apprentissage se fait dans un mode spécial lors de la reconnaissance. Le résultat de l'apprentissage est un motif utilisateur, appliqué lors de la reconnaissance du texte. Pour de plus amples détails, consultez la section Apprentissage de nouveaux caractères, p. 47.

### Création de nouvelles langues

Pour accélérer la reconnaissance et améliorer sa précision, vous pouvez créer de nouvelles langues et leur attacher des dictionnaires. Vous pouvez utiliser un fichier texte en code ANSI en tant que dictionnaire.

Par exemple, pour reconnaître une colonne de tableau, dans laquelle une gamme très limitée de mots est reproduite, vous pouvez créer une nouvelle langue et lui attacher un dictionnaire, contenant tous les mots pouvant être rencontrés dans cette colonne.

## Exemple de reconnaissance du russe ancien

Pour reconnaître des textes en russe ancien, effectuez les opérations suivantes :

1. Créez une copie de la langue russe.
2. Définissez la langue créée comme langue de reconnaissance.
3. Créez un nouveau motif (vide), connectez-le et commencez la reconnaissance avec apprentissage.
4. Reconnaissez 1 ou 2 pages de russe ancien en mode *Lire&Apprendre*.
5. Visualisez et éditez le motif.
6. Désactivez le mode *Lire&Apprendre* (mais ne désactivez pas le motif formé).
7. Exécutez la reconnaissance.
8. Dans le texte reconnu, remplacez les ligatures contenant le caractère « \* » avec les caractères correspondants (voir ci-dessous).

Nous expliquerons plus tard comment créer et attacher une nouvelle langue en détail.

### Créer une nouvelle langue :

1. Dans le menu *Outils*, sélectionnez la commande *Editeur de langue*.
2. Dans la boîte de dialogue affichée, cliquez sur *Nouveau*.
3. Activez le bouton radio en position *Modifier la copie de la langue existante*.
4. Dans la liste qui s'affiche, sélectionnez *Russe* et cliquez sur *Suivant*.
5. Tapez le nom de la nouvelle langue : *Russeancien*.
6. Des caractères russes s'affichent dans la boîte de dialogue. Ajoutez maintenant le caractère « \* » (ou tout autre caractère existant dans la page de code cyrillique que vous pouvez taper à partir du clavier). Ce caractère sera utilisé pour créer des ligatures. Pour que le système ignore le caractère « \* » lors de la vérification de l'orthographe, marquez-le afin qu'il soit ignoré (cliquez sur le bouton situé à droite de la case *Ignoré* pour le sélectionner dans le tableau).

### Créer un nouveau motif :

1. Dans le menu *Outils*, sélectionnez l'élément *Editeur de motif*.
2. Dans la boîte de dialogue *Motifs*, cliquez sur *Nouveau*.
3. Nommez le nouveau motif dans la boîte de dialogue *Création de motif*, par exemple, *Russeancien*.

### Connexion du motif et activation du mode Lire&Apprendre

1. Dans la page de l'onglet *Reconnaissance* de la boîte de dialogue *Options* (*Outils>Options...*) dans le groupe *Apprentissage*, sélectionnez le motif que vous venez de créer et que vous voulez utiliser pour l'apprentissage, à partir de la liste *Motif utilisateur actuel*.

2. Cochez les cases *Lire&Apprendre* et *Fonction Omnifont*<sup>1</sup>.

**Pour spécifier la langue de reconnaissance :**

1. Dans la liste *Langue de reconnaissance*, de la barre d'outils Reconnaissance, sélectionnez la valeur *Autre*.
2. Dans la boîte de dialogue qui s'affiche, sélectionnez l'option russe et russeancien<sup>2</sup>.  
Puis démarrez la reconnaissance.

**Reconnaissance avec apprentissage**

Lors de la reconnaissance avec apprentissage (mode *Lire&Apprendre*), le système affiche la boîte de dialogue *Apprentissage de motif*. Dans la partie supérieure de la boîte de dialogue, vous apercevez un fragment de ligne de texte en cours de reconnaissance, pour lequel le caractère non reconnu est encadré par un rectangle (Figure 29, p. 48). Tapez le caractère à l'aide du clavier et spécifiez sa langue (Russeancien). Si le caractère « i » est compris dans le rectangle, tapez alors les caractères « \*й » dans la case *Symbole* ; si le caractère « v » est compris dans le rectangle, tapez la ligature « \*и » ; pour le caractère « Θ », la ligature « \*φ » ; pour le caractère « Ъ », la ligature « \*е ».

Lorsque l'apprentissage est terminé, le motif est créé. Avant de l'utiliser pour la reconnaissance, visualisez et modifiez-le le cas échéant (*Outils>Editeur de motif...*) (Figure 30, p. 50).

Après avoir utilisé le motif pour l'apprentissage sur une ou deux pages du texte, reconnaissez le reste du texte, et connectez le motif créé et modifié. Puis spécifiez que vous allez effectuer la reconnaissance sans l'apprentissage en utilisant le motif créé, et démarrez la reconnaissance en spécifiant la combinaison de langue russe et russeancien dans la barre d'outils.

**Remplacer dans l'éditeur**

Après que le texte a été reconnu, dans l'éditeur de texte qui possède des caractères spéciaux (par exemple, Θ, v etc.), effectuez le remplacement contextuel des ligatures par les caractères, c'est-à-dire, remplacez \*φ par Θ, \*и par v et ainsi de suite.

**Formats d'images que FineReader peut charger**

**BMP:** n/b; gris, couleur  
**PCX, DCX:** n/b; gris, couleur (16, 256 couleurs)

---

<sup>1</sup> Omnifont n'est désélectionné que dans des cas rares – pour former le programme à reconnaître des lettres de certaines langues exotiques.

<sup>2</sup> La langue russe est connectée afin de permettre l'utilisation du dictionnaire qui améliore la qualité de la reconnaissance et la vérification de l'orthographe.



<b>BMP:</b>	n/b; gris, couleur
<b>JPEG:</b>	gris, couleur
<b>TIFF:</b>	n/b – non condensé, CCITT3, CCITT3FAX, CCITT4, Packbits; gris – non condensé, Packbits, JPEG; couleur – non condensé, Packbits, JPEG
<b>PNG</b>	n/b; gris, couleur

## Langues prises en charge

Afrikaans, Albanais, **Allemand**, **Anglais**, Basque, Biélorusse, Breton, **Bulgare**, Catalan, Chechen, Crimean Tatar, Croate, **Danois**, **Espagnol**, **Estonien**, Fijian, **Finnois**, **Français**, Gaélique, **Grec**, Hawaïien, Hongrois, Islandais, **Italien**, Kabardien, Latin, Letton, Lithuanien, Macédonien, Malais, Maori, Moldavian, **Norvégien (Bokmel, Nynorsk)**, Néerlandais (**Belgique**, **Standard**), Ossetic, **Polonais**, **Portuguais (Brésil, Standard)**, Rhaeto-Romanic, Roumain, **Russe**, Samoan, Serbe (Cyrillique), Slovaque, Slovenian, **Suédois**, Swahili, Tagal, Tchèque, **Turc**, **Ukrainien**.

## Utilisation des dictionnaires de l'utilisateur et des motifs de FineReader 3.0

### Utilisation des dictionnaires de l'utilisateur

Une fois la nouvelle version installée, convertissez les anciens dictionnaires au nouveau format :

1. Dans le menu *Outils*, sélectionnez la commande *Options*.
2. Dans la page de l'onglet *Vérification des résultats* de la boîte de dialogue *Options*, cliquez sur *Ajouter au dictionnaire...*
3. Dans la boîte de dialogue qui s'affiche, spécifiez le chemin d'accès au dossier de FineReader (le dossier dans lequel le fichier Fine32.exe de l'ancienne version est situé).  
La liste des langues avec dictionnaires de l'utilisateur s'affiche dans la partie supérieure de la boîte de dialogue.
4. Sélectionnez les langues de votre choix et cliquez sur *Ajouter*.

### Utilisation des motifs utilisateur

Une fois la nouvelle version installée, copiez les fichiers des motifs (avec une extension .ptn) depuis le dossier de la version précédente du programme vers le dossier c:\Program Files\ABBYY FineReader.

## Utilisation des touches dans FineReader

	Pour :	Appuyez sur :
<b>Menu Fichier</b>	Créer un nouveau lot	CTRL+N
	Ouvrir un lot ou une image, ajouter une image au lot	CTRL+O
<b>Menu Edition</b>	Exécuter l'Assistant Exportation	CTRL+D
	Enregistrer dans un fichier	ALT+SHIFT+S
	Envoyer dans MS Word	ALT+SHIFT+W
	Envoyer dans MS Excel	ALT+SHIFT+E
	Envoyer dans Word Pro	ALT+SHIFT+D
	Envoyer dans WordPerfect	ALT+SHIFT+P
	Envoyer dans l'e-mail	ALT+SHIFT+M
	Envoyer dans le Presse-papiers	ALT+SHIFT+C
	Annuler une action	ALT+RET.ARR ou CTRL+Z
	Répéter une action	ALT+MAJ+RET.ARR ou CTRL+Y
<b>Menu Affichage</b>	Couper la sélection et la placer dans le Presse-papiers	CTRL+X
	Copier la sélection dans le Presse-papiers	CTRL+INS ou CTRL+C
	Insérer le contenu du Presse-papiers dans le texte	CTRL+V ou MAJ+INS
	Supprimer les blocs actifs, une sélection, des pages sélectionnées	DEL
	Supprimer des pages...	ALT+DEL
	Sélectionner tout le texte dans l'éditeur, ou toutes les pages du lot, ou tous les blocs de l'image en cours	CTRL+A
	Transformer le bloc en Texte	CTRL+1
	Transformer le bloc en Tableau	CTRL+2
	Transformer le bloc en Image	CTRL+3
	Fusionner les cellules du tableau	CTRL+ENTREE
	Diviser les cellules du tableau	CTRL+MAJ+ENTREE
	Rechercher le fragment voulu dans le texte reconnu	ALT+F3 ou CTRL+F
	Répéter la recherche	F3
	Rechercher le fragment nécessaire dans le texte reconnu et le remplacer par un autre fragment	CTRL+H
	Ouvrir la boîte de dialogue Polices	CTRL+MAJ+F
	Ouvrir la boîte de dialogue Mise en forme	ALT+SHIFT+F
	Ouvrir la boîte de dialogue pour définir la langue de sélection	CTRL+L
	Sélection en gras	CTRL+G
	Sélection en italique	CTRL+I
	Formater la sélection avec un souligné continu	CTRL+S
<b>Menu Image</b>	Afficher l'image entière	CTRL+MAJ+1
	Mettre l'image à l'échelle selon sa largeur	CTRL+MAJ+2
	Mettre l'image à l'échelle selon sa hauteur	CTRL+MAJ+3
	Zoom avant	CTRL+ HAUT
	Zoom arrière	CTRL+ BAS
	Ouvrir la page suivante	CTRL+ Num +
	Ouvrir la page précédente	CTRL+ Num -
	Ouvrir la page numéro...	CTRL+G
	Mettre la liste de la page à jour	F5
	Afficher les propriétés de l'image active (bloc)	ALT+ENTREE
<b>Menu Image</b>	Sélectionner un outil pour créer un bloc de texte	ALT+1
	Sélectionner un outil pour ajouter une partie de bloc	ALT+2
	Sélectionner un outil pour supprimer une partie de bloc	ALT+3
	Sélectionner un outil pour déplacer des blocs	ALT+4

	<b>Pour :</b>	<b>Appuyez sur :</b>
	Sélectionner un outil pour renuméroter des blocs	ALT+5
	Sélectionner un outil pour supprimer des blocs	ALT+6
	Sélectionner un outil pour ajouter des lignes verticales à un bloc de tableau	ALT+SHIFT+1
	Sélectionner un outil pour ajouter des lignes horizontales à un bloc de tableau	ALT+SHIFT+2
	Sélectionner un outil pour supprimer des lignes d'un bloc de tableau	ALT+SHIFT+3
	Exécuter la numérisation avec reconnaissance	F9
	Exécuter la numérisation et la reconnaissance de plusieurs pages	MAJ+F9
	Ouvrir et reconnaître l'image	ALT+SHIFT+O
	Exécuter l'Assistant Scanner&Lire	CTRL+J
	Exécuter la numérisation	CTRL+K
	Exécuter la numérisation de plusieurs pages	CTRL+MAJ+K
	Arrêter la numérisation	ALT+SHIFT+K
	Marquer les blocs sur la page ouverte	CTRL+E
	Marquer les blocs sur toutes les pages sans blocs	CTRL+MAJ+E
	Reconnaître la page ouverte	CTRL+R
	Reconnaître toutes les pages non reconnues	CTRL+MAJ+R
	Reconnaître toutes les pages non reconnues en tâche de fond	ALT+SHIFT+R
	Vérifier les résultats	F7
<b>Menu Outils</b>	Se déplacer vers le mot non identifié suivant	F4
	Se déplacer vers le mot non identifié précédent	MAJ+F4
	Supprimer les marques d'erreur	CTRL+Q
	Afficher les dictionnaires	CTRL+MAJ+D
	Traduire le mot ou la combinaison de mots sélectionné	CTRL+MAJ+T
	Ouvrir la boîte de dialogue Editeur de langue	CTRL+MAJ+L
	Ouvrir la boîte de dialogue Editeur de motif	CTRL+MAJ+A
	Ouvrir la boîte de dialogue Paramètres du scanner	CTRL+MAJ+S
	Ouvrir la boîte de dialogue Options	CTRL+MAJ+O
	Organiser les fenêtres	CTRL+W
<b>Menu Aide</b>	Obtenir l'aide contextuelle de l'objet de l'interface sélectionné (élément de menu ou de boîte de dialogue)	MAJ+F1
	Ouvrir la page de traitement sélectionnée	ENTREE
<b>Autres commandes</b>	Activer la fenêtre Traitement	ALT+0
	Activer la fenêtre Image	ALT+8
	Activer la fenêtre Texte	ALT+9
	Se déplacer d'une cellule de tableau à une autre	ALT+ GAUCHE, DROITE, BAS ou HAUT



## PUISSANCE INCROYABLE & PRECISION ETONNANTE

Depuis longtemps, l'homme rêve d'enseigner *la connaissance humaine* aux machines, de créer par exemple, une machine capable de lire. Lors des 50 dernières années, ce rêve a commencé à devenir réalité. La reconnaissance est l'une des

applications de l'IA<sup>1</sup> les **plus** prometteuses. Il s'agit d'une solution proche de la lecture humaine qui utilise les principes de la reconnaissance des systèmes vivants. Connue sous le nom de technologie de la perception réfléchie intégrale et sélective.

### Principales différences entre Standard, Professional et Handprint.

	Standard	Professional	Handprint
Types de texte reconnu	Imprimé		Imprimé, manuscrit
Reconnaissance de codes à barres		Oui	Oui
Reconnaissance de textes en colonnes, y compris les graphiques et les tableaux. Enregistre le document reconnu au format RTF	Oui	Oui	Oui
Utilisation avec Internet : enregistre le document reconnu au format HTML	Oui	Oui	Oui
<b>Langues prises en charge</b>			
Langues reconnues	40 langues ! La liste des langues prises en charge se trouve sur le côté du coffret.		
Correcteur orthographique intégré	Oui	Oui	Oui
Reconnaissance de documents multilingues	Oui	Oui	Oui
Extensible pour inclure de nouvelles langues		Oui	Oui
Interaction avec d'autres applications			
<b>Reconnaissance de tableaux</b>			
Reconnaissance des tableaux, enregistrement du résultat aux formats RTF, CSV, XLS et DBF	Oui	Oui	Oui
Segmentation des tableaux manuelle et automatique	Oui	Oui	Oui
Post-édition des tableaux reconnus	Oui	Oui	Oui
<b>Reconnaissance de formulaires</b>			
Reconnaissance de formulaires et enregistrement au format DBF		Oui	Oui
Conception des modèles de formulaires		Oui	Oui
Placement automatique des formulaires en utilisant les points de référence		Oui	Oui
Exportation vers des bases de données par ODBC 32			Oui
Sélection automatique du modèle de formulaire			Oui
<b>Traitement par lot des documents et des fonctions avancées</b>			
Traitement distribué des documents via un réseau		Oui	Oui
Intégration à d'autres applications par API, OLE Automation, Active X.	Disponible dans SDK <sup>2</sup>		

<sup>1</sup> - IA - intelligence artificielle ; <sup>2</sup> - SDK - kit de développement de logiciel



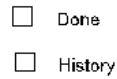
- ☐ Done  
☐ History

Complete the application on a typewriter. Do not type the letters on the field names!  
The line spacing between fields is 1, 2 and 3 typewriter intervals.

## JOB APPLICATION

1. Last Name										
First Name										
2. Date of Birth	Day	Month (in Words)		Year						
3. Sex (tick the appropriate square with "x")										
4. The address where you live now										
Post or Zip Code	Town/City									
Street										
House Number	Apartment		Area Code		Phone					
5. Telephone Number										
6. Passport										
Series	Number		Day		Month (in Words)		Year			
	N									
Issued										
7. Application Date	Day	Month (in Words)		Year						
Signature										

Place  
photo  
here



## JOB APPLICATION



- ☒ Done  
☒ History

Complete the application on a typewriter. Do not type the letters on the field names!  
The line spacing between fields is 1, 2 and 3 typewriter intervals.

## JOB APPLICATION

1. Last Name	Savinova		
First Name	Olga		
2. Date of Birth	Day 4	Month (in Words) April	Year 1964
3. Sex (tick the appropriate square with 'x')	M	F	X
4. The address where you live now			
Post or Zip Code	Town/City		
14097	Moscow		
Street			
High			
House Number	Apartment	Area Code	Phone
67	51	7095	0087657
5. Telephone Number			
6. Passport			
Series	Number	Day	Month (in Words)
BBCAI	N 70098	11	October
Issued	Year 1984		
7. Application Date	Day 14	Month (in Words) January	Year 1995
Signature			

