

Programme OCR primé :

plus de 70 prix reçus dans le monde entier

ABBYY

SOFTWARE
HOUSE

Leader reconnu :

- Meilleure précision
- Meilleures performances
- Meilleure convivialité

ABBYY

FineReader[®]

Version 6.0

OCR

Manuel Utilisateur

Système de reconnaissance optique de caractères

ABBYY FineReader

Version 6.0

Manuel Utilisateur

Les informations du présent document sont susceptibles de modification sans préavis et n'impliquent aucun engagement de la part d'ABBYY Software House.

Le logiciel décrit dans ce document est fourni aux termes d'un accord de licence. Le logiciel ne peut être utilisé ou copié qu'en stricte conformité avec les termes de l'accord de licence. La copie de ce logiciel sur tout support constitue une infraction aux lois de la Fédération de Russie relatives à la protection juridique des logiciels et des bases de données et aux lois internationales, sauf autorisation expresse dans le cadre de la licence ou des accords de non-divulgateion.

Aucune partie de ce document ne peut en aucun cas être reproduite ni transmise, sous aucune forme ni par aucun moyen, sans l'autorisation expresse écrite d'ABBYY Software House.

© ABBYY Software House, 2002. Tous droits réservés.

ABBYY, BIT Software, FineReader, « fontain image transformation », Lingvo, Scan&Read, Scan&Translate, « one button principle », « Your computer reads by itself », sont des marques déposées d'ABBYY Try&Buy et DOCFLOW sont des appellations commerciales d'ABBYY Software House.

Adobe®, le logo Adobe, Adobe PDF (Portable Document Format) et Adobe Acrobat® sont les marques déposées d'Adobe Systems Incorporated.

Toutes les autres marques sont des appellations commerciales ou des marques déposées de leur détenteur légal. 123015 Moscou, P.O. 72. ABBYY

Table des matières

Chapitre 1

| | |
|--|---|
| Installation et démarrage d'ABBYY FineReader | 3 |
| Configuration matérielle et logicielle requise | 4 |
| Installation de ABBYY FineReader | 4 |
| Installation sur un serveur et une station de travail réseau | 5 |
| Démarrage de ABBYY FineReader | 6 |

Chapitre 2

| | |
|---|----|
| Démarrage rapide | 7 |
| Comment saisir instantanément un document | 8 |
| Fenêtre principale de ABBYY FineReader | 8 |
| Barres d'outils de ABBYY FineReader | 10 |

Chapitre 3

| | |
|--|----|
| Caractéristiques générales d'ABBYY FineReader | 15 |
| A quoi sert un système OCR ? | 16 |
| Nouvelles caractéristiques d'ABBYY FineReader 6.0 | 17 |
| Formats d'enregistrement de document pris en charge par FineReader | 18 |
| Support de formats d'image | 18 |

Chapitre 4

| | |
|---|----|
| Acquisition de l'image | 19 |
| Numérisation | 20 |
| Définition des paramètres de numérisation | 21 |
| Conseils pour le réglage de la luminosité | 22 |
| Numérisation de documents multipages | 22 |
| Ouverture d'images | 24 |
| Numérisation de pages doubles | 24 |
| Ajout d'images de cartes de visite au lot | 25 |
| Travailler avec les images | 25 |
| Numérotation des pages | 27 |
| Options d'image de lot | 28 |

Chapitre 5

| | |
|---|----|
| Analyse de la présentation de la page | 29 |
| Informations générales sur l'analyse de la présentation | 30 |
| Types de blocs | 30 |
| Options d'analyse automatique de la présentation | 31 |
| Tracer et modifier des blocs manuellement | 32 |
| Analyse manuelle de la présentation d'un tableau | 35 |
| Utilisation des modèles de bloc | 35 |

Chapitre 6

| | |
|--|----|
| Reconnaissance | 37 |
| Informations générales sur la reconnaissance | 38 |
| Langue de reconnaissance | 38 |
| Type d'impression du texte source | 39 |
| Autres options de reconnaissance | 40 |
| Reconnaissance en arrière-plan | 40 |
| Reconnaissance avec apprentissage | 41 |
| Apprentissage d'un gabarit utilisateur | 42 |
| Modification d'un gabarit utilisateur | 44 |
| Langues utilisateur et groupes de langues | 44 |
| Création d'une nouvelle langue | 45 |
| Création d'un nouveau groupe de langues | 46 |

Chapitre 7

| | |
|---|----|
| Vérification et modification de texte | 47 |
| Vérification du texte dans FineReader | 48 |
| Options de vérification et de modification de texte | 49 |
| Ajout et suppression de mots dans le dictionnaire utilisateur | 50 |
| Modification de texte dans FineReader | 51 |
| Modification de tableaux | 54 |

Chapitre 8

| | |
|---|----|
| Enregistrement dans des applications et des formats externes | 55 |
| Informations générales sur l'enregistrement du texte reconnu | 56 |
| Options d'enregistrement de texte | 56 |
| Enregistrement du texte reconnu au format RTF et DOC | 58 |
| Enregistrement du texte reconnu au format PDF | 58 |
| Enregistrement du texte reconnu au format HTML | 59 |
| Enregistrement de l'image de page | 60 |

Chapitre 9

| | |
|--|----|
| Travailler avec les lots | 61 |
| Informations générales concernant l'utilisation des lots | 62 |
| Création d'un nouveau lot | 63 |
| Ouverture d'un lot | 63 |
| Ajout d'images à un lot | 63 |
| Numéro de page de lot | 64 |
| Fermeture d'une page de lot ou d'un lot entier | 64 |
| Suppression d'un lot | 65 |
| Recherche en mode texte intégral dans les pages du lot reconnues | 65 |

Chapitre 10

| | |
|---|----|
| Traitement des documents en réseau | 67 |
| Travailler en réseau sur un même lot | 68 |
| Travail de groupe à l'aide des mêmes langues et dictionnaires utilisateur | 69 |
| Travail de groupe à l'aide de dictionnaires personnalisés, pour les langues prenant en charge les dictionnaires | 70 |

Annexe

| | |
|--------------------|----|
| Raccourcis clavier | 71 |
| | 72 |

Bienvenue

Merci d'avoir choisi ABBYY FineReader !

La saisie informatique de texte nous est indispensable. Pendant longtemps, le clavier a constitué le seul moyen de saisie des documents imprimés. Pensez aux longues heures passées à saisir du texte... Si seulement l'ordinateur pouvait lire directement des feuilles de papier !

Les rêves se réalisent parfois. L'application d'OCR (Optical Character Recognition - reconnaissance optique de caractères) FineReader permet à un ordinateur et à un scanner de lire par eux-mêmes les textes imprimés.

Un scanner ne suffit-il pas pour lire le texte ?

Non. Le scanner ne fait que créer une image, une photographie du texte. Une photographie est un ensemble de points noirs et blancs ou, en d'autres termes, un fichier image qui ne peut être modifié dans un traitement de texte tel que MS Word, WordPerfect, Word Pro, etc. Vous avez besoin d'un système d'OCR, qui recherche des symboles parmi les points noirs et blancs, « reconnaît » des lettres dans chaque symbole et transforme l'image en texte modifiable sur un ordinateur. Une telle application produit un fichier exploitable par les traitements de texte et les ordinateurs de bureau.

Je peux donc maintenant saisir automatiquement un document sur mon ordinateur ?

Oui, vous pouvez saisir automatiquement un document sur votre ordinateur sans passer par le clavier.

A vous de jouer !

Manuel de l'utilisateur

Ce guide vous présente les bases d'ABBYY FineReader. Chaque chapitre commence par une brève description de son contenu.

Aide en ligne

L'Aide en ligne de FineReader contient des informations de base et approfondies sur les fonctions, les paramètres et les boîtes de dialogue. L'Aide en ligne est au format HTML. Elle vous permet de trouver les informations facilement et rapidement.

Fichier Lisez-moi

Il contient des informations de dernière minute sur le programme.

Support technique

Si vous avez des questions sur l'utilisation de FineReader, consultez la documentation à votre disposition (guide d'utilisation et fichier d'aide) avant de contacter notre service technique. Vous pouvez également parcourir la section consacrée au support technique sur notre site Web www.abbyy.com. Vous y trouverez peut-être la réponse à vos questions.

Si vous ne parvenez pas à résoudre votre problème, adressez-nous un courrier électronique à l'adresse support@abbyy.com. Les experts de notre support technique auront besoin des informations suivantes pour répondre efficacement à vos questions :

- vos nom et prénom
- le nom de votre société
- votre numéro de téléphone (ou de télécopie ou adresse électronique)
- le numéro de série de votre exemplaire de FineReader
- le numéro de version (voir Aide/A propos de.../Version)
- la description générale du problème et tout message d'erreur complet (le cas échéant).
- votre ordinateur et type de processeur
- la version de votre système d'exploitation Windows

Vous pouvez nous fournir toute autre information que vous considérez importante.

Vous pouvez obtenir certaines de ces informations automatiquement en cliquant sur le bouton **Infos sur le système** dans la boîte de dialogue **A propos d'ABBYY FineReader** (menu **Aide**).

Tous les utilisateurs des versions courantes et précédentes peuvent bénéficier d'une assistance gratuite.

Chapitre 1

Installation et démarrage d'ABBY FineReader

Ce chapitre traite des procédures d'installation de FineReader et de sujets apparentés, notamment la configuration système requise et l'installation sur station de travail et réseau.

La configuration de FineReader est assurée par un programme d'installation spécial, fourni avec le produit de distribution. Utilisez toujours la disquette ou le CD-ROM fourni lors de l'achat de FineReader. L'installation échouera si vous utilisez une copie du programme d'installation.

Contenu de ce chapitre :

- Configuration matérielle et logicielle requise
- Installation de FineReader
- Installation sur un serveur et une station de travail réseau
- Démarrage de FineReader

Configuration matérielle et logicielle requise

La configuration suivante est requise pour utiliser FineReader :

1. PC avec processeur Pentium 200 ou supérieur.
2. Microsoft(r) Windows(r) XP, Microsoft(r) Windows(r) 2000, Windows NT(r) Workstation 4.0 (avec Service Pack 6 ou supérieur) ou Windows 95/98/ME.
3. 64 Mo (Windows XP/2000), 32 Mo (Windows Me/98/NT 4.0), 16 Mo (Windows 95) de mémoire, plus 16 Mo de mémoire pour chaque processeur supplémentaire sur les systèmes multiprocesseurs.
4. Microsoft Internet Explorer 5.0 ou supérieur (Microsoft Internet Explorer 5.5 est fourni avec FineReader).
5. 90 Mo d'espace disque pour l'installation minimale du programme.
6. 70 Mo d'espace disque pour l'utilisation du programme.
7. Scanner 100% compatible TWAIN, caméra numérique ou modem/fax.
8. Lecteur de CD-ROM.
9. Souris ou autre dispositif de pointage.
10. Ecran VGA ou résolution supérieure.

Installation de ABBYY FineReader

Options d'installation

Quand le programme d'installation a contrôlé le système ; tapez votre nom et indiquez le dossier dans lequel vous voulez installer ABBYY FineReader. Le programme d'installation propose alors plusieurs options d'installation. Sélectionnez l'option de votre choix :

- **Typique** (recommandé) - tous les composants sont installés, y compris toutes les langues de reconnaissance, une seule langue d'interface sélectionnée pendant l'installation.
- **Installation personnalisée** - choisissez les composants du produit que vous souhaitez installer (notamment les langues de reconnaissance disponibles).

Remarque : si vous souhaitez utiliser les dictionnaires et les gabarits utilisateur d'une version antérieure de FineReader, ne désinstallez pas celle-ci avant l'installation de la nouvelle version. Une fois l'installation terminée, **les gabarits et les dictionnaires de la version antérieure** seront disponibles.

Installation d'ABBYY FineReader

Si le produit que vous avez acheté contient à la fois un CD-ROM et une disquette, procédez comme suit :

1. Insérez la disquette d'installation dans le lecteur de disquettes.
2. Insérez le CD-ROM dans le lecteur de CD.
3. Cliquez sur le bouton **Démarrer** dans la barre des tâches, puis sélectionnez la commande **Paramètres/Panneau de configuration**.
4. Cliquez deux fois sur l'icône **Ajout/Suppression de programmes**.
5. Sélectionnez l'onglet **Installation/Désinstallation** et cliquez sur le bouton **Installer**.
6. Suivez les instructions d'installation qui s'affichent à l'écran.

Si le produit que vous avez acheté contient seulement un CD-ROM, procédez comme suit :

1. Insérez le CD-ROM dans le lecteur de CD.
2. Cliquez sur le bouton **Démarrer** dans la barre des tâches, puis sélectionnez la commande **Paramètres/Panneau de configuration**.
3. Cliquez deux fois sur l'icône **Ajout/Suppression de programmes**.
4. Sélectionnez l'onglet **Installation/Désinstallation** et cliquez sur le bouton **Installer**.
5. Suivez les instructions d'installation qui s'affichent à l'écran.

Remarque : vous devrez saisir le **code d'installation** pour poursuivre l'installation dans les cas suivants : votre ordinateur ne dispose pas de lecteur de disquettes 3,5", l'installation s'exécute à partir d'une disquette endommagée ou qui n'est pas d'origine, certaines applications installées sur votre ordinateur entrent en conflit avec le programme ABBYY FineReader. Vous obtiendrez le **code d'installation** auprès de la société ABBYY ou de votre revendeur. Le code d'installation est créé à partir de l'**ID du produit** (généré automatiquement par le programme d'installation) et du numéro de série (figurant sur la carte d'enregistrement). Pour obtenir le **code d'installation**, vous devrez remplir un formulaire sur le site www.abbyy.com, envoyer le numéro de série figurant sur votre carte d'enregistrement par courrier électronique, ou tout simplement contacter le support technique par téléphone et lui communiquer les informations requises.

Si un message d'erreur s'affiche, consultez le fichier **Readme.htm** (sur le CD-ROM ABBYY FineReader) pour connaître les instructions à suivre dans les situations d'erreur les plus courantes.

Installation sur un serveur et une station de travail réseau

Installation sur un serveur

Cette rubrique ne concerne que les administrateurs système.

- Si votre progiciel contient à la fois un CD-ROM et une disquette, insérez la disquette d'installation et exécutez le fichier setup.exe du CD-ROM FineReader à l'aide de l'option de ligne de commande /a.
- Si le produit acheté ne contient qu'un CD-ROM, exécutez setup.exe depuis le CD-ROM FineReader à l'aide de l'option de ligne de commande /a.

Licences supplémentaires

Si FineReader est utilisé simultanément par plusieurs personnes, vous devez ajouter des numéros de série une fois l'installation terminée sur le serveur pour assurer un fonctionnement correct.

Pour ce faire :

1. Exécutez le fichier LicSetup.exe situé dans le dossier Fichiers programmes\ABBYY FineReader 6.0 dans lequel le programme ABBYY FineReader 6.0 Corporate Edition a été installé. La boîte de dialogue **Ajouter une licence** s'affiche.
2. Tapez un nouveau numéro de série et cliquez sur le bouton **Ajouter**.

Remarque :

1. Vous ne pouvez pas utiliser d'unités logiques créées par la commande SUBST.
 2. Des composants comme MS Windows NT 4.0 Service Pack 6.0 ou MS Internet Explorer 5.5 permettent d'accéder à certaines fonctionnalités optionnelles de FineReader. Ces composants ne sont pas installés si vous effectuez l'installation sur un serveur. Dans les autres types d'installation de FineReader, ces composants sont installés automatiquement. Pour éviter tout problème lié à l'absence de ces composants, l'administrateur système doit mettre à jour le système avant de procéder à l'installation d'ABBYY FineReader.
 3. Avant l'installation, vérifiez que tous les **utilisateurs** ont accès en lecture et écriture au dossier réseau Utilisateurs (créé automatiquement pendant l'installation et contenant des fichiers temporaires).
-

Installation sur une station de travail réseau

Si le logiciel ABBYY FineReader 6.0 Corporate Edition a été installé sur un serveur de réseau, vous pouvez exécuter le programme d'installation directement à partir de ce serveur.

Pour installer le logiciel ABBYY FineReader 6.0 Corporate Edition sur une station de travail :

- Exécutez le fichier Setup.exe à partir du dossier réseau contenant le logiciel ABBYY FineReader 6.0 Corporate Edition. Suivez les instructions d'installation qui s'affichent à l'écran.

Remarque :

1. Vous devez posséder des droits d'administrateur sur la station de travail où ABBYY FineReader est installé.
 2. Si le message « Impossible de charger FineReader. Il n'y a pas de licence gratuite. » s'affiche, vérifiez le nombre de licences supplémentaires que vous avez ajoutées dans la boîte de dialogue **Ajouter une licence**, ainsi que le nombre d'utilisateurs actuels du logiciel.
 3. Pour que le programme ABBYY FineReader 6.0 fonctionne correctement, l'utilisateur doit avoir l'accès en lecture et en écriture au dossier correspondant.
-

Démarrage de ABBYY FineReader

Pour lancer le programme ABBYY FineReader :

- Sélectionnez la commande **ABBYY FineReader 6.0 Professional (Corporate Edition)** dans le menu **Démarrer/Programmes**.

Remarque : avant de démarrer FineReader, assurez-vous que le scanner est connecté à l'ordinateur et qu'il est mis sous tension. Dans le cas contraire, consultez le manuel d'utilisation du scanner pour connaître les instructions d'installation.

Si vous n'avez pas de scanner, FineReader peut reconnaître vos fichiers d'image (des exemples de fichiers se trouvent dans le dossier **ABBYY FineReader/Démo**).

Chapitre 2

Démarrage rapide

Ce chapitre explique comment saisir un document sans rien connaître sur le fonctionnement du système FineReader.

Vous ferez également la connaissance des fenêtres et des barres d'outils de FineReader. Si vous connaissez déjà FineReader, vous pouvez passer directement à la section *Nouvelles caractéristiques de FineReader 6.0*.

Contenu de ce chapitre :

- Comment saisir instantanément un document
- Fenêtre principale de ABBYY FineReader
- Barres d'outils de ABBYY FineReader

Comment saisir instantanément un document

1. Mettez le scanner sous tension si nécessaire.

Remarque : de nombreux modèles de scanner doivent être mis sous tension avant l'ordinateur.

2. Mettez l'ordinateur sous tension et démarrez FineReader (**Démarrer/Programmes/ABBYY FineReader Professional** ou **Corporate Edition**). La fenêtre principale de FineReader s'affiche.
3. Insérez la page que vous souhaitez lire dans le scanner.
4. Cliquez sur la flèche située à droite du bouton **Numériser&Lire**. Sélectionnez la commande **Assistant Numériser&Lire** dans le menu contextuel.



L'assistant **Numériser&Lire** est un mode spécial de numérisation/lecture. Un échantillon de fichier d'image est proposé dans le dossier **Demo**. Ce dossier est situé dans le dossier où vous avez installé FineReader. FineReader vous guide pas à pas dans la procédure.

5. Suivez les instructions de l'assistant **Numériser&Lire**.

La procédure de saisie du document se déroule en quatre étapes : numérisation, lecture, vérification orthographique et enregistrement du texte reconnu.

Après numérisation, une « photographie » de la page source apparaît dans la fenêtre **Image**. L'application vous demande ensuite de définir les paramètres de reconnaissance. Elle commence ensuite la reconnaissance de la page, dont elle analyse simultanément la présentation. Les parties de l'image déjà traitées sont signalées en bleu.

Le texte reconnu s'affiche dans la fenêtre **Texte**. Vous pouvez vérifier et modifier le texte dans cette fenêtre. Une fois la vérification du document terminée, l'assistant **Numériser&Lire** propose d'envoyer le texte reconnu vers une application donnée, de l'enregistrer dans un fichier ou de traiter d'autres images.

Fenêtre principale de ABBYY FineReader

Avec FineReader, vous travaillez toujours avec un lot de documents donné. Un lot est un dossier contenant des images, des fichiers de texte reconnus et d'autres fichiers d'informations de FineReader. Chaque image numérisée constitue une page de lot séparée. Si un même fichier d'image en contient plusieurs (par exemple avec une image TIFF multipage), chaque image de ce fichier constituera une page de lot séparée.

Lorsque vous démarrez FineReader pour la première fois, le lot par défaut est sélectionné. Vous pouvez utiliser ce lot ou en créer un nouveau. Pour plus de détails, consultez la section *Informations générales sur l'utilisation des lots*.


Fenêtre principale
Barre d'outil standard
Barre d'outil Formatage
Barre d'outils Assistant

Fournit des outils pour toutes les procédures : Numération, reconnaissance, vérification orthographique et enregistrement.

Fenêtre Texte

Affiche le texte reconnu, pour contrôle et vérification orthographique

Fenêtre Image

Affiche l'image numérisée pour l'examiner et tracer des blocs.

Fenêtre Zoom

Affiche l'image agrandie de la ligne de texte en cours de modification ou de la zone d'une image sur laquelle vous travaillez.

Fenêtre Lot

Affiche les pages du lot ouvert, dans l'un des deux modes : Vignettes (comme à présent) ou Détails.

Barre d'outils image

Elle contient des outils pour tracer et modifier les blocs, effectuer un zoom et modifier l'image.

Le menu principal s'affiche en haut de la **fenêtre principale de FineReader**. En dessous de ce menu figurent quatre barres d'outils : Standard, Formatage, Image et Assistant. Vous pouvez afficher ou masquer n'importe quelle barre d'outils.

Pour afficher ou masquer une barre d'outils, sélectionnez la commande **Barre d'outils** dans le menu **Affichage** ou cliquez avec le bouton droit de la souris sur l'une des barres d'outils pour afficher le menu contextuel. La liste des barres d'outils apparaît, les barres actives étant sélectionnées. Cliquez sur le nom de celle que vous souhaitez afficher ou masquer.

La barre d'état s'affiche en bas de la fenêtre principale de FineReader. Cette barre affiche des informations sur l'état de l'application et les opérations en cours, ainsi que des informations succinctes sur les commandes et les boutons que vous avez sélectionnés.

La fenêtre **Lot** est toujours affichée dans la **fenêtre principale**. Il est possible d'afficher trois autres fenêtres : **Image**, **Zoom** et **Texte**.

Les fenêtres **Image**, **Zoom** et **Texte** sont interconnectées : lorsque vous cliquez deux fois sur une zone d'image donnée dans la fenêtre **Image**, la zone correspondante s'affiche dans la fenêtre **Zoom**, et le pointeur, situé dans la fenêtre **Texte** (s'il existe déjà du texte reconnu pour la page), se place à l'endroit où vous avez cliqué.

Vous pouvez réorganiser les fenêtres à l'écran :

- Sélectionnez l'un des éléments ci-dessous : **fenêtre Lot >...; fenêtres Image et Texte >...; fenêtre Zoom >...** dans le menu **Affichage**.

Dispositions conseillées pour les fenêtres à l'écran :

Fenêtre **Lot** sur la gauche ; **Affichage de lot** : **Vignettes** ;
fenêtres **Image**, **Texte** et **Zoom**

Fenêtre **Lot** en haut ; **Affichage de lot** : **Détails** ;
fenêtres **Image**, **Texte** et **Zoom**

Fenêtre **Lot** en haut ; **Affichage de lot** : **Détails** ;
fenêtres **Image** et **Zoom**

Fenêtre **Lot** en haut ; **Affichage de lot** : **Détails** ;
fenêtres **Texte** et **Zoom**

Utile si/quand :

Les pages du lot ne sont pas trop nombreuses.

Les pages du lot sont nombreuses.

Vous effectuez une analyse de la présentation et une reconnaissance.

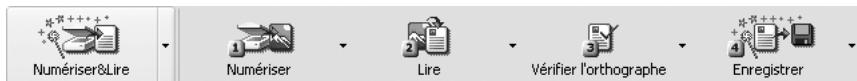
Vous modifiez le texte reconnu.

Pour passer d'une fenêtre une autre :

- Appuyez sur **CTRL+TAB**.
- Appuyez sur **ALT+1** pour activer la fenêtre **Lot**.
- Appuyez sur **ALT+2** pour activer la fenêtre **Image**.
- Appuyez sur **ALT+3** pour activer la fenêtre **Texte**.

Barres d'outils de ABBYY FineReader

FineReader comporte quatre barres d'outils : **Standard**, **Image**, **Formatage** et **Assistant**. La plupart du temps, vous constaterez que les boutons des barres d'outils sont le moyen le plus pratique d'accéder aux fonctions de l'application ; toutefois, les mêmes fonctions sont également accessibles par les menus ou par des raccourcis clavier. Pour connaître la fonction d'un bouton de la barre d'outils, placez le pointeur de la souris sur ce bouton pour afficher la bulle d'aide, ainsi que des informations supplémentaires dans la barre d'état.

Barre d'outils Assistant

Les boutons de la barre d'outils **Assistant** permettent d'exécuter les principales fonctions de FineReader : **Numérisation**, **Lecture**, **Vérification** et **Enregistrement** des résultats de la reconnaissance. Les numéros des boutons indiquent l'ordre dans lequel les différentes actions de saisie du document doivent être effectuées. Vous pouvez exécuter chacune d'elles séparément ou simultanément en cliquant sur le bouton de l'**assistant Numériser&Lire**. L'**assistant Numériser&Lire** effectue automatiquement le cycle complet de traitement du document.

Chaque bouton comporte plusieurs modes. Cliquez sur la flèche située à droite du bouton et sélectionnez un mode dans le menu contextuel. L'icône du bouton affiche toujours le dernier mode sélectionné. Cliquez sur le bouton lui-même pour re-exécuter ce mode.

Numériser&Lire



Assistant Numériser&Lire : exécute un mode spécial de numérisation et de lecture. FineReader vous assiste dans la procédure de traitement du document et indique les actions à effectuer pour obtenir le résultat visé.

Numériser&Lire : lance la numérisation et la lecture d'un document en utilisant les options actives.

Numériser&Lire avec images multiples : numérise et lit plusieurs images à la suite.

Ouvrir&Lire : ouvre et lit les images sélectionnées dans la boîte de dialogue **Ouvrir**.

1 - Numériser



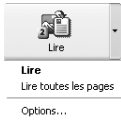
Ouvrir une image : ajoute une ou plusieurs images au lot. Chaque image ajoutée est copiée dans le dossier du lot.

Numériser une image : numérise une image.

Numériser des images multiples : numérise les images jusqu'à ce que vous arrêtez l'opération. Pour ce faire, sélectionnez la commande **Arrêter la numérisation** dans le menu **Fichier**.

Options : affiche la page de l'onglet **Numériser/Ouvrir une image** (boîte de dialogue **Options**), pour vous permettre de configurer les options de numérisation, etc.

2 - Lire



Lire : lit la page de lot ouverte.

Tout lire : lit toutes les pages de lot non reconnues.

Options : affiche la page de l'onglet **Reconnaissance** (boîte de dialogue **Options**) permettant de définir les options de reconnaissance du document.

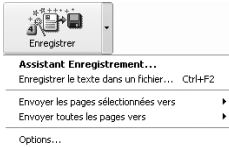
3 - Vérifier l'orthographe



Vérifier l'orthographe : recherche dans le texte les mots mal orthographiés ou incertains (contenant des caractères dont la reconnaissance est incertaine).

Options : affiche la page de l'onglet **Vérification orthographique** (boîte de dialogue **Options**) permettant de définir les options de vérification orthographique.

4 - Enregistrer



Assistant Enregistrement : lance l'**assistant Enregistrement**, qui vous permet de sélectionner l'application dans laquelle le texte doit être enregistré et de définir les options d'enregistrement.

Enregistrer le texte dans un fichier : enregistre le texte reconnu dans un fichier sur disque.

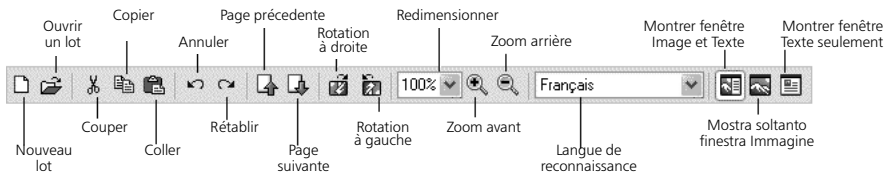
Envoyer les pages sélectionnées vers : cette option permet d'exporter seulement les pages de lot sélectionnées. Dans ce cas sélectionnez les pages à exporter, puis cette option et enfin l'application de destination. FineReader exporte alors les pages vers l'application sélectionnée sans enregistrer le texte sur disque.

Envoyer toutes les pages vers : exporte toutes les pages reconnues vers l'application sélectionnée sans enregistrer le texte sur disque.

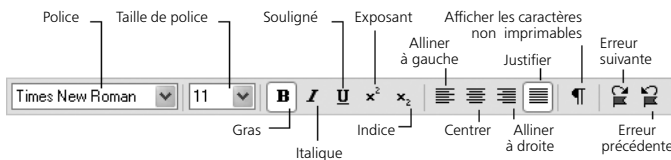
Options : affiche la page de l'onglet **Formatage** (boîte de dialogue **Options**) permettant de définir les options d'enregistrement.

Barre d'outils Standard

Les boutons de la barre d'outils **Standard** donnent accès à des outils de gestion des fichiers et des images (annulation/rétablissement d'une action, défilement des pages de lot, nettoyage et rotation d'image) ainsi qu'à la liste des **Langues de reconnaissance**.

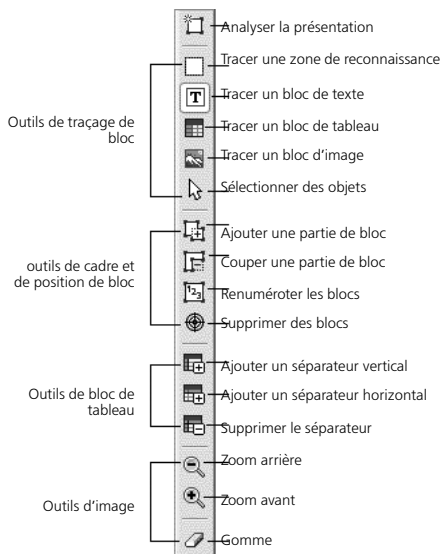


Barre d'outils Formatage



Les boutons de la barre d'outils **Formatage** donnent accès à différents outils de mise en forme du texte. La fenêtre **Texte** permet de modifier le texte et son formatage.

Barre d'outils Image



Les boutons de la barre d'outils **Image** donnent accès à des outils d'analyse de présentation (par exemple, création et édition de blocs), ainsi qu'à des outils de mise à l'échelle et de modification d'image (nettoyage, etc.).

Remarque : vous pouvez utiliser les boutons servant à créer et modifier des blocs dans la fenêtre **Zoom**, ainsi que dans la fenêtre **Image**.

Configuration de la barre d'outils

Remarque : l'affichage de la fenêtre principale de FineReader ou, plus précisément, le nombre de boutons affichés dans les barres d'outils, varie avec la résolution de l'écran. Pour afficher tous les boutons disponibles, augmentez la résolution de votre moniteur. Toutefois, FineReader dispose des mêmes fonctionnalités, même si tous les boutons ne sont pas affichés ; ils ne représentent en effet qu'une partie des fonctions de FineReader et à chaque bouton correspond toujours une commande.

Avec FineReader, vous pouvez configurer les barres d'outils **Standard**, **Image** et **Formatage**. Vous pouvez, par exemple, ajouter ou supprimer les boutons permettant d'accéder à n'importe quelles commandes de l'application.

Chaque commande de menu possède sa propre icône. Pour voir la liste complète des commandes et de leurs boutons respectifs, reportez-vous à la liste **Commandes** de la boîte de dialogue **Paramètres (Outils>menu Configuration)**.

Pour ajouter un bouton à une barre d'outils :

1. Sélectionnez la catégorie requise dans le champ **Catégories**.

Remarque : les commandes de la liste sont classées par élément de menu. Le choix de la catégorie affecte les commandes affichées dans la liste **Commandes**.

2. Sélectionnez la barre d'outils à laquelle vous souhaitez ajouter un bouton dans le champ **Barres d'outils**.
3. Sélectionnez une commande dans la liste **Commandes**, puis cliquez sur le bouton (**>>**).

La commande est alors ajoutée à la liste des commandes de la barre d'outils sélectionnée et s'affiche dans la fenêtre principale de cette barre d'outils.

Pour supprimer un bouton d'une barre d'outils :

- Sélectionnez le bouton que vous souhaitez supprimer dans la liste **Boutons de la barre d'outils** et cliquez sur le bouton (**<<**).

Remarque :

1. L'ordre des boutons de la liste détermine leur ordre sur la barre d'outils. Pour modifier cet ordre, sélectionnez la commande dans la liste des commandes en cours de la barre d'outils, puis cliquez sur **Haut** ou **Bas** pour déplacer la commande dans la liste.
 2. Vous pouvez classer les commandes par groupes : sélectionnez **Séparateur** dans la liste des **Commandes**, puis cliquez sur le bouton **Ajouter**. Un séparateur est alors ajouté à la liste des boutons de la barre d'outils en cours. Vous pouvez déplacer le séparateur selon vos besoins.
 3. Pour restaurer les boutons par défaut dans une barre d'outils, sélectionnez-la dans la liste **Barres d'outils**, puis cliquez sur le bouton **Restaurer**. Pour restaurer les boutons par défaut de toutes les barres d'outils, cliquez sur le bouton **Restaurer tout**.
-

Chapitre 3

Caractéristiques générales d'ABBYY FineReader

FineReader offre des outils permettant la saisie numérique de documents. Il suffit de cliquer une fois sur le bouton Numériser&Lire : l'application s'occupe du reste. Vous pouvez envoyer le texte reconnu vers une application de traitement de texte ou de tableur, l'enregistrer au format RTF/DOC, PDF ou HTML et conserver ainsi la mise en page, ou encore exporter le texte reconnu vers une base de données.

Contenu de ce chapitre :

- A quoi sert un système OCR ?
- Nouvelles caractéristiques d'ABBYY FineReader 6.0
- Formats de documents pris en charge en sortie
- Formats d'image pris en charge

A quoi sert un système OCR ?

Les systèmes OCR (Optical Character Recognition - Reconnaissance optique de caractères) permettent de saisir automatiquement des documents imprimés dans votre ordinateur.

FineReader est un système de reconnaissance optique de textes multi-polices. Cela signifie que le système reconnaît les textes imprimés dans pratiquement n'importe quelle police sans apprentissage préalable. L'une des caractéristiques de FineReader est sa haute précision de reconnaissance et sa faible sensibilité aux défauts d'impression. Ces qualités sont dues à une technologie spéciale de reconnaissance (technologie IPA), basée sur les principes d'Intégralité, de Pertinence et d'Adaptabilité. Cette technologie est intégralement mise en œuvre dans FineReader.

Le traitement de saisie d'un document peut se décomposer en deux étapes :

1. **Numérisation.** Lors de la première étape, le scanner agit comme « l'œil » de votre ordinateur : il regarde l'image et la transfère dans l'ordinateur. L'image acquise n'est rien de plus qu'une image, un ensemble de points noirs, blancs et de couleur, impossible à modifier dans une application de traitement de texte.
2. **Reconnaissance.** Lors de la seconde étape, FineReader effectue le traitement ROC de l'image.

Attardons-nous sur la seconde étape.

Le traitement ROC de l'image par FineReader implique l'analyse du fichier d'image transmis par le scanner (analyse de la présentation) et la reconnaissance de chaque caractère. Les processus d'analyse de la présentation (sélection des zones de reconnaissance, des tableaux, des images, des lignes et des caractères individuels) et de lecture de l'image interagissent de façon étroite : l'analyse de la présentation utilise les informations sur le texte reconnu pour affiner ses résultats.

Comme indiqué précédemment, le processus de reconnaissance de l'image est basé sur les principes d'Intégralité, de Pertinence et d'Adaptabilité.

- **Intégralité** - l'objet de la reconnaissance est décrit en tant qu'entité globale par un ensemble d'éléments de base et les relations qui existent entre eux.
- **Pertinence** - la reconnaissance est le processus consistant à générer des hypothèses et à les vérifier.
- **Adaptabilité** - la capacité d'auto apprentissage du système.

Ces trois principes déterminent le comportement du système. Le système émet une hypothèse concernant un objet à reconnaître (un caractère, une partie de caractère, plusieurs caractères « collés ») puis l'accepte ou la rejette, en essayant de déterminer pas à pas tous les éléments structurels et leurs relations. Ces éléments structurels sont les équivalents informatiques des parties de caractères indispensables à la perception humaine (arcs, cercles, points, etc.). Le programme s'adapte ensuite au texte - il se règle de lui-même - en tirant parti de l'expérience acquise à partir des premiers caractères reconnus avec certitude. La recherche pertinente et l'utilisation des informations du contexte permettent au système de reconnaître même les caractères abîmés ou déformés, de sorte qu'il est presque insensible aux défauts d'impression.

Le résultat final est le texte reconnu, que vous voyez dans la fenêtre **Texte** de FineReader et qui peut être modifié et enregistré dans n'importe quel format réutilisable.

Nouvelles caractéristiques d'ABBYY FineReader 6.0

Caractéristiques générales

- Vous pouvez maintenant ouvrir et lire des fichiers PDF avec FineReader. Le format PDF est un des formats standard en matière de publication de documents sur Internet, d'archivage de documents, etc. Vous pouvez donc ouvrir les fichiers PDF dans FineReader, les lire, les modifier et les enregistrer, au format PDF ou sous tout autre format pris en charge par FineReader.
- Intégration à Windows Explorer. Les fichiers d'image et les lots FineReader peuvent à présent s'afficher dans Windows Explorer.
- Enregistrement des documents reconnus sous le nom de leurs images source.
- Barres d'outils personnalisables.

Traitement de l'image

- Impression des images numérisées et du texte reconnu.
- Outil permettant de séparer des images figurant sur deux pages ou contenant des cartes de visite disposées dans un ordre spécifique.

Reconnaissance

- 177 langues de reconnaissance. Consultez **ABBYY FineReader l'aide**, section *Langues prises en charge*, pour la liste complète des langues prises en charge.
- Algorithme amélioré, pour la reconnaissance de documents de qualité médiocre, intégrant une nouvelle méthode adaptable de binarisation des images (idéal pour les images numérisées en mode Gris ou Couleur), ainsi qu'une nouvelle méthode de suppression de l'arrière-plan.

Enregistrement et modification

- Editeur WYSIWYG à plusieurs colonnes. Les blocs comportant du texte, des tableaux et des images reconnus s'affichent dans ce mode, à leur emplacement d'origine.
- Précision d'enregistrement accrue des documents source dans MS Word : enregistrement des images non rectangulaires, de l'organisation du texte sur plusieurs colonnes et des listes (numérotées et à puces).
- Prise en charge des fichiers PDF multilingues : FineReader enregistre les textes multilingues au format PDF, sans nécessiter l'installation de polices supplémentaires.
- Nouveau mode d'enregistrement au format PDF : Image seulement.
- Outil permettant de sélectionner le taux de compression lors d'un enregistrement au format HTML et PDF.
- Outil permettant de modifier la résolution des images JPEG lors d'un enregistrement au format RTE, DOC et PDF.
- Alignement d'un texte sous forme de tableau lors d'un transfert vers MS Excel ou d'un enregistrement au format XLS.

Caractéristiques professionnelles

- Mode Groupe partagé s'appliquant à l'utilisation des langues et des dictionnaires utilisateur, ainsi qu'aux dictionnaires spécifiques aux langues prédéfinies (version FineReader Corporate seulement).
- Dans le menu **Edition>Recherche avancée** figure un outil permettant d'effectuer une recherche sur le texte ou les termes d'un document, indépendamment de leur format (version FineReader Corporate seulement).
- Application de renseignement de formulaires ABBYY FormFiller (version FineReader Corporate seulement). C'est également une application GRATUITE pour les utilisateurs de ABBYY FineReader Professional enregistrés).

Formats d'enregistrement de document pris en charge par FineReader

FineReader peut enregistrer le résultat de la reconnaissance dans les formats suivants :

- Document Microsoft Word (*.DOC)
- Rich Text Format (*.RTF)
- Adobe Acrobat (*.PDF)
- HTML
- Fichier CSV (Comma Separated Values - valeurs séparées par des virgules) (*.CSV)
- Texte simple (*.TXT). FineReader accepte diverses pages de code (Windows, DOS, Mac, ISO), ainsi que le codage Unicode.
- Feuille de calcul Microsoft Excel (*.XLS)
- DBF

Formats d'image pris en charge

FineReader affiche les formats de fichiers d'image suivants :

PDF: Fichiers au format PDF, version 1.3 ou antérieure

BMP : 2 bits - noir et blanc
4 et 8 bits - palette
16 bits
24 bits - palette et couleurs vraies
32 bits

PCX, 2 bits - noir et blanc

DCX : 4 et 8 bits - niveaux de gris

JPEG : niveaux de gris et couleurs vraies

TIFF : noir et blanc - non condensé, CCITT3, CCITT3FAX, CCITT4, packbits
niveaux de gris - non condensé, packbits, JPEG
couleurs vraies - non condensé, JPEG
palette - non condensé, packbits
multi-image TIFF

PNG : noir et blanc, niveaux de gris, couleur

FineReader enregistre les fichiers d'image dans les formats suivants :

BMP : noir et blanc, niveaux de gris, couleur

PCX : noir et blanc, niveaux de gris

JPEG : niveaux de gris, couleur

TIFF : noir et blanc - non condensé, CCITT3, CCITT3FAX, CCITT4, packbits
niveaux de gris - non condensé, packbits, JPEG
couleurs vraies - non condensé, JPEG

PNG : noir et blanc, niveaux de gris, couleur

Chapitre 4

Acquisition de l'image

La qualité de la reconnaissance dépend dans une large mesure de la qualité de l'image source. Ce chapitre explique comment numériser un document correctement, comment ouvrir et lire une image déjà présente sur votre ordinateur (consultez la liste des formats d'image pris en charge dans la section *Formats d'image pris en charge*), comment traiter les images et les préparer pour la reconnaissance (éliminer la « poussière » de la numérisation), etc.

Contenu de ce chapitre :

- Numérisation
- Définition des paramètres de numérisation
- Conseils pour le réglage de la luminosité
- Numérisation de documents multipages
- Ouverture d'images
- Numérisation de pages doubles
- Ajout d'images de cartes de visite au lot
- Numérotation des pages
- Travailler avec une image
- Options d'image de lot

Numérisation

FineReader « dialogue » avec les scanners par l'intermédiaire de l'interface TWAIN. Cette norme universelle a été adoptée en 1992 pour unifier les normes d'interaction des périphériques d'imagerie (tels que les scanners) et les applications externes. FineReader peut « dialoguer » avec un scanner de deux manières via le gestionnaire TWAIN :

- **Utiliser sa propre interface** : utilisez la boîte de dialogue **Paramètres de scanner** de FineReader pour définir les options de numérisation ; l'option correspondante est **Utiliser l'interface FineReader** ;
- **Utiliser l'interface TWAIN du scanner** : utilisez la boîte de dialogue TWAIN du scanner pour définir les options de numérisation ; l'option correspondante est **Utiliser l'interface de la source TWAIN** ;

Les deux modes présentent des avantages et des inconvénients.

Lorsque vous choisissez l'option **Utiliser l'interface de la source TWAIN**, l'option d'aperçu de l'image devient généralement disponible. L'aperçu permet de définir la zone de numérisation et d'ajuster la luminosité : ces réglages sont faciles car vous pouvez vérifier leur impact sur l'aperçu. Un inconvénient est que les différents scanners utilisent des boîtes de dialogue de gestionnaire TWAIN différentes. Les instructions d'utilisation de la boîte de dialogue TWAIN sont fournies avec le scanner.

Lorsque vous sélectionnez l'option **Utiliser l'interface FineReader**, les fonctionnalités suivantes deviennent disponibles : tout d'abord, vous pouvez numériser des images multiples sur les scanners à chargeur automatique ; ensuite, vous pouvez enregistrer les options de numérisation dans les fichiers de modèle de lot (*.fnt), puis les utiliser avec d'autres lots.

Vous pouvez aisément basculer d'un mode à l'autre :

- Sélectionnez l'onglet **Numériser/Ouvrir une image** dans la boîte de dialogue **Options** (menu **Outils>Options**) et cliquez sur le bouton radio approprié : **Utiliser l'interface de la source TWAIN** ou **Utiliser l'interface FineReader**.

Remarque :

1. L'option **Utiliser l'interface FineReader** peut ne pas être disponible avec certains modèles de scanner.
 2. Si vous souhaitez afficher la boîte de dialogue Paramètres de scanner dans le mode **Utiliser l'interface FineReader**, sélectionnez l'option **Montrer la boîte de dialogue d'options avant de numériser** dans la page de l'onglet **Numériser/Ouvrir une image** (menu **Outils>Options**).
-

Important : consultez la documentation du scanner pour le connecter correctement. Après avoir connecté le scanner à l'ordinateur, n'oubliez pas d'installer un gestionnaire TWAIN ou une application de numérisation.

Pour lancer la numérisation :



Cliquez sur le bouton **1 - Numériser** ou sélectionnez la commande **Numériser** dans le menu **Fichier**. La fenêtre **Image** s'ouvre et affiche une « photographie » de la page numérisée dans la **fenêtre principale de FineReader**.

Si vous souhaitez numériser plusieurs pages, cliquez sur la flèche située à droite du bouton **1 - Numériser** et sélectionnez la commande **Numériser des images multiples** dans le menu contextuel.

Si la numérisation n'est pas lancée immédiatement, l'une des deux boîtes de dialogue suivantes s'affiche :

- La boîte de dialogue **Source TWAIN** du scanner. Vérifiez les options de numérisation et cliquez sur le bouton **OK** pour lancer la numérisation.

- La boîte de dialogue **Paramètres de scanner**. Vérifiez les options de numérisation et cliquez sur le bouton **OK** pour lancer la numérisation.

Conseil :

Si vous souhaitez lancer la reconnaissance dès que les images source sont numérisées, utilisez la commande **Numériser&Lire** ou **Numériser&Lire avec images multiples**.



Cliquez sur la flèche située à droite du bouton **Numériser&Lire** et sélectionnez la commande **Numériser&Lire** ou **Numériser&Lire avec images multiples** dans le menu contextuel. FineReader numérise et lit les images. La fenêtre **Image** affiche une « photographie » de la page numérisée et la fenêtre **Texte** affiche le résultat de la reconnaissance dans la **fenêtre principale de FineReader**. Le texte reconnu peut être transmis à diverses applications externes et enregistré dans divers formats.

Définition des paramètres de numérisation

La qualité de la reconnaissance dépend dans une large mesure de celle de l'image numérisée. Pour régler celle-ci, vous devez définir les principaux paramètres de numérisation : résolution, mode de numérisation et luminosité.

Les principaux paramètres de numérisation sont les suivants :

- **Résolution** : choisissez une résolution de 300 ppp pour le texte normal (taille de police de 10 points ou plus) et de 400 à 600 ppp pour les textes de plus petite taille (9 points ou moins).
- **Mode de numérisation - gris.**
La numérisation en niveaux de gris est le mode le plus performant pour le système de reconnaissance. Si vous numérisez vos images dans ce mode, l'application règle automatiquement la luminosité.
- **Mode de numérisation : noir et blanc.**
Le mode noir et blanc permet d'accélérer la numérisation, mais au risque de perdre certaines informations sur les caractères. Cela peut affecter la qualité de la reconnaissance des documents dont la qualité d'impression est moyenne ou médiocre.
- **Mode de numérisation : couleur.** Si vous numérisez des documents en couleur, qu'il s'agisse d'images, de caractères en couleur ou d'un arrière-plan, il peut être utile de conserver les couleurs dans le document électronique. Dans ce cas, utilisez le mode de numérisation couleur. Dans les autres cas, choisissez le mode gris.
- **Luminosité** : dans la plupart des cas, une luminosité de valeur moyenne (50%) suffit. Certains documents numérisés en noir et blanc peuvent nécessiter un réglage de la luminosité.

Remarque : la numérisation avec une résolution de 400 à 600 ppp (au lieu des 300 ppp par défaut), en niveaux de gris ou en couleur (au lieu du noir et blanc) est plus lente. Sur certains modèles de scanner, la numérisation en 600 ppp prend quatre fois plus de temps qu'en 300 ppp.

Pour définir les paramètres de numérisation :

- Si vous numérisez des images à l'aide de l'**interface TWAIN de FineReader**, sélectionnez la commande **Paramètres de scanner** du menu **Outils**. La boîte de dialogue **Paramètres de scanner apparaît**. Définissez les options de numérisation dans la boîte de dialogue.
- Si vous numérisez des images à l'aide de l'interface de la source TWAIN, la boîte de dialogue TWAIN du scanner s'affiche automatiquement lorsque vous cliquez sur le bouton **1-Numériser**. Définissez les paramètres de numérisation dans la boîte de dialogue. Le nom des options de numérisation peut varier selon le modèle du scanner ; par exemple, le terme

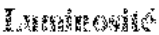

« seuil » peut s'appliquer à la luminosité. Le réglage de la luminosité peut également se présenter sous la forme d'une icône de soleil ou d'un cercle noir et blanc. Ces options sont décrites plus précisément dans la documentation de votre scanner.

Conseils pour le réglage de la luminosité

L'image numérisée doit être lisible. Pour vérifier sa lisibilité, examinez-la dans la fenêtre **Zoom**.

Luminosité - exemple d'image correcte (du point de vue de la ROC)

Si vous constatez que l'image numérisée est médiocre (caractères collés ou fragmentés, consultez le tableau ci-dessous pour savoir comment améliorer la qualité de l'image.

| Votre image a l'aspect suivant : | | Solution: |
|--|--|---|
|  Luminosité | caractères « déchirés » ou très clairs | <ul style="list-style-type: none">● Réduisez la luminosité pour rendre l'image plus sombre.● Numérisez en niveaux de gris (le réglage automatique de la luminosité est utilisé dans ce mode). |
|  Luminosité | caractères déformés, collés ou remplis | <ul style="list-style-type: none">● Augmentez la luminosité pour rendre l'image plus claire.● Numérisez en niveaux de gris (le réglage automatique de la luminosité est utilisé dans ce mode). |

Numérisation de documents multipages

FineReader prend en charge un mode de numérisation spécial, commode pour les documents multipages : Numériser des images multiples. Ce mode permet de numériser autant de pages que vous le souhaitez. Il présente les particularités suivantes :

- Si vous numérisez des images à l'aide de l'**interface TWAIN de FineReader**, le scanner ne s'arrête pas après avoir numérisé une page et passe automatiquement à la suivante.
- Si vous numérisez des images à l'aide de l'**interface de la source TWAIN**, la **boîte de dialogue TWAIN** du scanner ne se referme pas une fois la page numérisée. Vous pouvez insérer la page suivante et la numériser, et ainsi de suite.

Si vous devez numériser un grand nombre de pages, deux possibilités s'offrent à vous : vous pouvez utiliser un chargeur automatique de documents (ADF) ou faire le travail manuellement.

Numérisation avec un chargeur automatique :

1. Si vous utilisez l'**interface de FineReader**, sélectionnez l'option **Utiliser le chargeur automatique de documents** de la boîte de dialogue **Paramètres de scanner** (menu **Outils> Paramètres de scanner**), puis la commande **Fichier>Numériser des images multiples** pour lancer l'opération.

2. Si vous utilisez l'**interface de la source TWAIN**, sélectionnez l'option **Utiliser le chargeur automatique de documents** de la boîte de dialogue TWAIN de votre scanner (cette option peut avoir un nom différent selon le modèle du scanner ; les options de la boîte de dialogue TWAIN sont décrites dans la documentation du scanner), puis sélectionnez la commande **Fichier>Numériser des images multiples** pour lancer la numérisation.

Numérisation sans chargeur automatique :

1. Si vous utilisez l'**interface de FineReader**
 - Sélectionnez la commande **Numériser des images multiples** du menu **Fichier**.

Certaines méthodes permettent de numériser plus facilement des pages multiples sur un scanner à plat sans chargeur automatique. Utilisez l'une des méthodes suivantes :

- Définissez la valeur de la pause, c'est-à-dire le délai séparant la fin de la numérisation d'une page et le début de la numérisation de la page suivante. Sélectionnez l'option **Pause entre les pages**, puis définissez la valeur de la pause (en secondes) dans la boîte de dialogue **Paramètres de scanner** (menu **Outils>Paramètres de scanner**).

Dans ce cas, le scanner ne commence à numériser la page de texte qu'à la fin du délai en secondes que vous avez défini, vous disposez ainsi du temps nécessaire pour insérer la page suivante. Une fois la pause terminée, la numérisation reprend normalement.

- Sélectionnez l'option **Arrêt entre les pages** de la boîte de dialogue **Paramètres de scanner** (menu **Outils>Paramètres de scanner**).

Dans ce cas, chaque fois que le scanner a fini de numériser une page, une boîte de dialogue vous demande si vous souhaitez poursuivre la numérisation. Cliquez sur le bouton **Oui** pour continuer ou **Non** pour mettre fin à l'opération.

Lorsque vous avez fini de numériser les pages, sélectionnez la commande **Arrêter la numérisation** du menu **Fichier**.

2. Si vous utilisez l'**interface de la source TWAIN**
 - Sélectionnez la commande **Numériser des images multiples** du menu **Fichier**. La boîte de dialogue TWAIN de votre scanner s'affiche. Cliquez sur **Numériser** (ce bouton peut également s'appeler **Final** ou autre selon la boîte de dialogue TWAIN de votre scanner) pour lancer la numérisation.

Numérisez la page, insérez une nouvelle page et cliquez sur le bouton **Numériser** de la boîte de dialogue **TWAIN** du scanner et ainsi de suite.

Lorsque vous avez terminé, cliquez sur le bouton **Fermer** (il peut également avoir un autre nom) dans la **boîte de dialogue TWAIN** de votre scanner.

Conseil : pour contrôler davantage la qualité des images numérisées, sélectionnez l'option **Ouvrir l'image pendant la numérisation** dans la page de l'onglet **Numérisation** (menu **Outils>Options**). Si vous activez cette option, chaque page numérisée s'affiche immédiatement dans la fenêtre **Image**. Si l'image vous semble mal numérisée, arrêtez l'opération (sélectionnez la commande **Arrêter la numérisation** du menu **Fichier**) et recommencez.

Ouverture d'images

Vous pouvez reconnaître des images même si vous n'avez pas de scanner (consultez la liste des formats utilisables dans la section *Formats d'image pris en charge*).

Pour ouvrir une image :

- Cliquez sur la flèche située à droite du bouton **1 - Numériser** et sélectionnez la commande **Ouvrir une image** dans le menu contextuel. L'icône du bouton **1 - Numériser** change et le titre devient **Ouvrir**.
- Sélectionnez la commande **Ouvrir une image** dans le menu **Fichier**.
- Dans Windows Explorer : cliquez avec le bouton droit de la souris sur le fichier d'image à ouvrir, puis dans le menu contextuel, sélectionnez l'option **Ouvrir avec FineReader**. Si l'application FineReader est déjà lancée, l'image s'ajoute au lot en cours. Dans le cas contraire, une copie de FineReader est lancée, avant l'ajout de l'image. Cette copie ouvre le dernier lot sur lequel vous avez travaillé.

Sélectionnez une ou plusieurs images dans la boîte de dialogue **Ouvrir**. Les images sélectionnées s'affichent dans la fenêtre **Lot** et la dernière image apparaît dans les fenêtres **Image** et **Zoom** de FineReader. Toutes les images sélectionnées sont copiées dans le dossier de lot. Consultez la section Informations générales sur l'utilisation des lots pour plus d'informations sur l'organisation des lots et l'affichage des images qu'ils contiennent.

Conseil : si vous souhaitez que les images ouvertes soient reconnues immédiatement, utilisez le mode **Ouvrir&Lire** :

1. Sélectionnez la commande **Ouvrir&Lire** dans le menu **Traitement** ou appuyez sur **CTRL+MAJ+D**. La boîte de dialogue **Ouvrir** s'affiche.
 2. Sélectionnez les images à reconnaître.
-

Numérisation de pages doubles

Lorsque vous numérisez un livre, il est plus commode de numériser à la fois les pages de gauche et les pages de droite (des « pages doubles »). Toutefois, la qualité de la reconnaissance est meilleure si la page double est segmentée après numérisation, de sorte que chaque page du lot corresponde à une seule page du livre. La reconnaissance et l'analyse de la présentation sont effectuées séparément pour chaque page, chaque page étant redressée si nécessaire.

Si vous souhaitez segmenter des pages doubles :

- Avant la numérisation, sélectionnez l'option **Segmenter les pages doubles** dans la page de l'onglet **Numériser/Ouvrir une image** (menu **Outils > Options**).

Lorsque cette option est sélectionnée, chaque page double est segmentée en deux pages de lot. Consultez la section Informations générales sur l'utilisation des lots pour plus d'informations sur les lots.

Remarque : si une image double est segmentée de manière incorrecte, désélectionnez la case à cocher **Partager les pages doubles** et numérisez à nouveau la page double ou rajoutez l'image correspondante au lot, puis segmentez manuellement l'image, à l'aide de la boîte de dialogue **Partager l'image** (**Image > Partager l'image**).



Ajout d'images de cartes de visite au lot

Il est pratique, lorsque vous saisissez des cartes de visite, d'en saisir un grand nombre à la fois (capacité maximale du scanner). La qualité de la reconnaissance est optimale (grâce au redressement par exemple) lorsque chaque carte de visite est reconnue sur une page distincte. L'application contient des outils de partage automatique et manuel des images de cartes de visite agencées selon un ordre donné (pour en savoir plus, consultez la section *Utilisation des cartes de visite*).

Pour partager une image :

1. Sélectionnez l'image requise dans la fenêtre **Lot**.
2. Sélectionnez la commande **Partager l'image** du menu **Image**. La boîte de dialogue **Partager l'image** s'affiche.
3. Cliquez sur le bouton **Partager les cartes de visite**.

Remarque :

1. La page divisée est supprimée du lot et remplacée par les éléments qui la composent. Pour en savoir plus, consultez la section Informations générales concernant l'utilisation des lots.
 2. Si l'image n'a pas été divisée correctement, fragmentez-la manuellement en cliquant sur le bouton **Ajouter un séparateur vertical/Ajouter un séparateur horizontal**.
 3. Pour supprimer l'ensemble des séparateurs, cliquez sur le bouton **Supprimer tous les séparateurs**.
 4. Pour déplacer un séparateur, activez le mode **Sélectionner un séparateur** (cliquez sur le bouton ) , puis déplacez le séparateur.
 5. Pour supprimer un séparateur, activez le mode **Sélectionner un séparateur** (cliquez sur le bouton ) , puis déplacez le séparateur hors de l'image.
-

Travailler avec les images

- Nettoyer l'image
- Inverser l'image
- Faire pivoter ou basculer l'image
- Effacer un bloc
- Redimensionner l'image
- Récupérer les informations de l'image
- Imprimer l'image
- Annuler la dernière action

1. Nettoyer l'image

L'image reconnue peut contenir de la « poussière », c'est-à-dire de nombreux points superflus. Ces points en excès sont dus à la numérisation de documents de qualité d'impression médiocre. Les points situés près du contour des caractères peuvent avoir un effet négatif sur la qualité de la reconnaissance.

Il est donc souhaitable de réduire le nombre de ces points. Pour cela :

- Sélectionnez la commande **Nettoyer l'image** dans le menu **Image**.

Si vous souhaitez nettoyer un bloc en particulier :

- Sélectionnez la commande **Nettoyer le bloc** dans le menu **Image**.

Remarque : dans le cas où l'original serait très pâle ou imprimé dans une police très fine, le nettoyage de l'image risque de faire disparaître des points, des virgules ou des parties très fines des caractères, provoquant une dégradation de la qualité de reconnaissance.

Si vous numérisez ou ouvrez des images très « poussiéreuses », sélectionnez l'option **Nettoyer l'image** dans le groupe **Prétraitement de l'image** de la page de l'onglet **Numériser/Ouvrir une image** (menu **Outils > Options**) pour les nettoyer avant que l'application ne les ajoute au lot.

2. Inverser l'image

Certains scanners inversent les images (transforment le noir en blanc et réciproquement) pendant la numérisation.



Vous pouvez utiliser la commande **Inverser l'image** pour rendre au document son apparence standard, c'est-à-dire une police noire sur un fond blanc. Pour ce faire :

- Sélectionnez la commande **Inverser l'image** dans le menu **Image**.

Remarque : si vous numérisez ou ouvrez des images inversées, sélectionnez l'option **Inverser l'image** dans le groupe **Prétraitement de l'image** de la page de l'onglet **Numériser/Ouvrir une image** (menu **Outils > Options**) avant de les ajouter au lot.

3. Faire pivoter ou basculer l'image

La qualité de la reconnaissance dépend de l'orientation de l'image (le texte doit se lire de haut en bas et les lignes doivent être horizontales). Par défaut FineReader détecte automatiquement l'orientation de la page pendant la reconnaissance. Si cette orientation est incorrecte, désélectionnez l'option **Détecter l'orientation d'image (pendant la reconnaissance)** de la page de l'onglet **Numériser/Ouvrir une image** et faites **pivoter l'image** manuellement pour lui donner une orientation standard :


- Cliquez sur le bouton  ou sélectionnez la commande **Rotation à droite** dans le menu **Image** pour faire pivoter l'image de 90° vers la droite.
- Cliquez sur le bouton  ou sélectionnez la commande **Rotation à gauche** dans le menu **Image** pour faire pivoter l'image de 90° vers la gauche.
- Sélectionnez la commande **Rotation à 180°** dans le menu **Image** pour faire pivoter l'image de 180°.

Pour basculer l'image :



- Horizontalement (autour de l'axe vertical) - sélectionnez la commande **Basculement horizontal** dans le menu **Image**.
- Verticalement (autour de l'axe horizontal) - sélectionnez la commande **Basculement vertical** dans le menu **Image**.

4. Effacer un bloc

Si vous ne souhaitez pas reconnaître une zone donnée de l'image ou si de grandes zones de l'image sont « poussiéreuses », vous pouvez les effacer en procédant comme suit :

- Sélectionnez l'outil , puis la zone de l'image que vous souhaitez effacer en maintenant enfoncé le bouton gauche de la souris. Relâchez le bouton pour effacer la zone sélectionnée.

5. Redimensionner l'image

- Sélectionnez l'outil  /  de la barre Image (dans la fenêtre **Image**) et cliquez sur l'image. L'échelle de l'image sera multipliée/divisée par deux.
- Cliquez avec le bouton droit de la souris sur l'image et sélectionnez la commande **Redimensionner**, puis sélectionnez l'échelle de votre choix dans le menu contextuel.

6. Récupérer les informations de l'image

Vous pouvez obtenir les informations suivantes sur l'image : largeur et hauteur en pixels ; résolution verticale et horizontale par pouce (ppp) ; type d'image.

- Cliquez avec le bouton droit de la souris sur l'image et sélectionnez la commande **Propriétés** dans le menu contextuel. Une boîte de dialogue s'ouvre. Sélectionnez l'onglet **Image** dans cette boîte de dialogue.

7. Imprimer l'image

Vous pouvez imprimer l'image ouverte dans la fenêtre **Image**, les images des pages sélectionnées dans la fenêtre **Lot** ou les images de toutes les pages du lot :

- Sélectionnez la commande **Imprimer l'image** dans le menu **Fichier**. La boîte de dialogue **Impression** s'ouvre. Réglez les paramètres d'impression requis (imprimante, nombre de pages à imprimer, nombre de copies, etc.).

8. Annuler la dernière action

- Pour annuler la dernière action, cliquez sur le bouton **Annuler** sur la barre **Standard** .

Conseil : pour annuler l'action **Annuler**, cliquez sur le bouton **Rétablir** de la barre **Standard** .

Numérotation des pages

Un numéro est affecté à chaque page numérisée. Le numéro par défaut est celui de la dernière page du lot majoré d'une unité.

Vous pouvez toutefois définir manuellement le numéro des pages si vous le souhaitez, pour conserver, par exemple, la numérotation initiale ou numériser une série de pages classées par numéro. Pour ce faire :

- Sélectionnez l'option **Demander un numéro de page avant d'ajouter la page au lot** dans la page de l'onglet **Numériser/Ouvrir une image** (menu **Outils>Options**).

Si vous numériser un grand nombre de pages recto verso triées par numéro :

1. Sélectionnez l'option **Demander un numéro de page avant d'ajouter la page au lot** dans la page de l'onglet **Numériser/Ouvrir une image** (menu **Outils>Options**).
2. Spécifiez le numéro de la première page numérisée dans la boîte de dialogue **Numéro de page**, puis sélectionnez l'option **Pages paires et impaires séparément** dans le champ **Numérotation des pages**. Sélectionnez l'ordre de numérotation des pages : croissant ou décroissant. Le mode dépend de la manière dont vous disposez les pages recto verso dans le chargeur automatique de documents, c'est-à-dire si le numéro disposé au-dessus est le plus grand ou le plus petit.

Options d'image de lot

Convertir en noir et blanc les images couleur et en niveaux de gris (**Numériser/Ouvrir une image** menu **Outils>Options**)

Sélectionnez l'option **Convertir en noir et blanc les images couleur et en niveaux de gris** si vous numérisez des images en niveaux de gris avec l'**interface de la source TWAIN** et que les images numérisées ne contiennent ni image, ni police, ni arrière-plan couleur ou si vous ne souhaitez pas conserver les couleurs des images numérisées. Si cette option est sélectionnée, les images numérisées occupent moins d'espace disque.

Chapitre 5

Analyse de la présentation de la page

Avant toute reconnaissance, FineReader doit « savoir » quelles sont les parties de l'image à reconnaître. C'est le rôle que joue la procédure d'analyse de la présentation. Elle trace des blocs de texte, d'image, de tableau et de code barre (FineReader version Corporate Edition uniquement).

Ce chapitre traite des sujets suivants : cas où une analyse de page manuelle peut être utile, types de bloc disponibles, modification des blocs tracés par l'analyse automatique, amélioration de l'analyse de la présentation à l'aide de modèles de bloc.

Contenu de ce chapitre

- Informations générales sur l'analyse de la présentation.
- Types de bloc
- Options d'analyse de la présentation
- Tracé et modification de bloc manuellement
- Analyse manuelle de la présentation d'un tableau
- Utilisation des modèles de bloc

Informations générales sur l'analyse de la présentation

L'analyse de la présentation de la page peut s'effectuer automatiquement ou manuellement. Dans la plupart des cas, FineReader traite lui-même la tâche d'analyse. Lancez l'analyse automatique en cliquant sur le bouton **2 - Lire**. La reconnaissance est effectuée en même temps que l'analyse de la présentation.

Remarque : une procédure autonome d'analyse de la présentation est également disponible (**menu Traitement>Analyser la présentation**). Vous pouvez utiliser cette procédure si nécessaire, mais pensez que la qualité de l'analyse peut être inférieure. En effet, la procédure simultanée d'analyse de la présentation et de reconnaissance met en œuvre les informations acquises pendant la reconnaissance pour seconder l'analyse de la présentation.

Vous pouvez tracer des blocs manuellement dans les cas suivants :

1. Vous voulez reconnaître **une partie** d'une page;
2. L'analyse automatique de la présentation a tracé les blocs incorrectement.

Conseil :

- Dans certains cas, vous pouvez améliorer la qualité de l'analyse automatique et ajustant les options d'analyse. Vérifiez les options en cours (onglet **Reconnaissance**, menu **Outils>Options**).
 - Si l'application trace incorrectement certains blocs, il est souvent plus rapide de modifier les blocs à l'aide des outils de modification de bloc que de supprimer tous les blocs et de les tracer manuellement.
-

Types de blocs

Les blocs sont des zones d'image encadrées. Les blocs indiquent au système quelles zones de l'image reconnaître et dans quel ordre. Les blocs ont également une influence sur la façon dont est conservée la présentation d'origine. Les blocs de types différents comportent des cadres de différentes couleurs. Vous pouvez modifier les couleurs des cadres de bloc dans la page de l'onglet **Affichage** de la boîte de dialogue **Options** (menu **Outils > Options**), dans le groupe **Apparence**. Sélectionnez le type de bloc dans le champ **Élément**, ainsi que la couleur de votre choix dans le champ **Couleur**.

Les types de bloc suivants sont disponibles :

Zone de reconnaissance - ce type de bloc sert à la reconnaissance et à l'analyse automatiques. Après avoir cliqué sur le bouton **2 - Lire**, tous les blocs de ce type seront analysés et reconnus automatiquement.

Texte - ce type de bloc sert aux zones d'image de texte. Il ne doit contenir que du texte formaté dans une seule colonne. Si des images se trouvent à l'intérieur du texte, tracez un bloc séparé pour chacune.

Tableau - ce type de bloc sert aux zones d'image de tableaux ou aux zones de texte structurées en tableaux. Lorsque l'application lit des blocs de ce type, elle trace des séparateurs verticaux et horizontaux dans le bloc, de manière à constituer un tableau. Ce bloc est représenté sous forme de tableau dans le texte en sortie. Vous pouvez tracer et modifier des tableaux manuellement.

Image - ce type de bloc sert aux zones contenant des images. Un bloc de ce type peut contenir une image réelle ou tout autre objet (par exemple une partie de texte) que vous souhaitez afficher sous forme d'image dans le texte reconnu.

Code barre (disponible uniquement dans la version **Corporate Edition**) - ce type de bloc sert aux zones de l'image qui sont des codes barre. Si votre document comporte un code barre imprimé et que

vous ne voulez pas l'afficher comme une image, mais comme une série de lettres et de chiffres dans le texte reconnu, tracez un bloc séparé pour le code barre et définissez le type de bloc sur **code barre**.

Remarque : il existe une option permettant l'analyse et la reconnaissance automatiques des codes barre, mais elle n'est pas définie par défaut. Pour la définir, sélectionnez l'option **Rechercher les codes barre** dans la page de l'onglet **Reconnaissance** (menu **Outils >Options**).

Options d'analyse automatique de la présentation

La procédure d'analyse automatique de la présentation de FineReader trace des blocs des types suivants : blocs de texte, de tableau, d'image et de code barre.

La procédure d'analyse automatique de la présentation (associée à la reconnaissance de texte) démarre lorsque vous cliquez sur le bouton **2 - Lire**. Avant de cliquer sur ce bouton, vous devez définir les principales options d'analyse de la présentation : type de document et options d'analyse de tableau.

Type de document

Dans la plupart des cas, la présentation est détectée automatiquement. La détection automatique s'effectue si la valeur **Détecter la présentation** de la page de l'onglet **Reconnaissance** est définie dans le groupe **Type de document** (menu **Outils >Options**). Cette valeur est définie par défaut.

Dans le cas où vous souhaiteriez définir manuellement le type de document :

- Sélectionnez l'élément de votre choix dans le groupe **Type de document** de la page de l'onglet **Reconnaissance** dans la boîte de dialogue **Options** (menu **Outils >Options**).

Types de documents disponibles :

Détecter la présentation - cette option indique que la présentation est détectée automatiquement. Cette valeur est définie par défaut. Si cette valeur est définie, la reconnaissance de tous les types de textes, y compris les textes sur plusieurs colonnes, les textes en tableau et les images, s'effectue automatiquement.

Colonne unique - cette option indique que le texte est formaté sur une seule colonne. Utilisez cette option dans le cas où l'analyse automatique de la présentation aurait identifié incorrectement le type de texte sur plusieurs colonnes.

Texte simple avec espaces - cette option indique que le texte est formaté sur une seule colonne et qu'il est défini avec une police à espacement fixe d'une seule taille. Dans le texte reconnu, les retraits à gauche seront représentés sous forme d'espaces, chaque ligne représentant un paragraphe séparé et les paragraphes d'origine seront séparés par des lignes vides. Ce type peut être utilisé si, par exemple, vous reconnaissez des listages de programmes en C++ ou d'anciennes sorties informatiques.

Options d'analyse de tableau

Dans la plupart des cas, l'application divise automatiquement le tableau en lignes et en colonnes. Si vous souhaitez ajuster d'autres options de tableau, ouvrez la page de l'onglet **Reconnaissance** dans le groupe **Tableaux**. Réglez ces options si :

- L'analyse automatique de la présentation a tracé incorrectement les lignes et les colonnes du tableau ;
- Le document contient de nombreux tableaux simples de même type (par exemple le tableau ne comporte pas de cellules fusionnées ou chaque cellule ne comporte qu'une ligne de texte).

1. Utilisez l'option **Une ligne de texte par cellule** si vous traitez un tableau sans séparateurs noirs ou un tableau dans lequel certains séparateurs sont noirs et d'autres non, et dans lequel chaque cellule ne comporte qu'une ligne de texte. Par exemple :

| Kilomètres | Miles |
|------------|-------|
| 1 | 0.62 |
| 5 | 3.2 |

- ce tableau ne compte qu'une ligne de texte par cellule

| Phénomène physique | T, degrés Celsius |
|-------------------------------|-------------------|
| Point d'ébullition de l'eau | 100 |
| Point de congélation de l'eau | 0 |

- ce tableau compte plusieurs lignes de texte par cellule

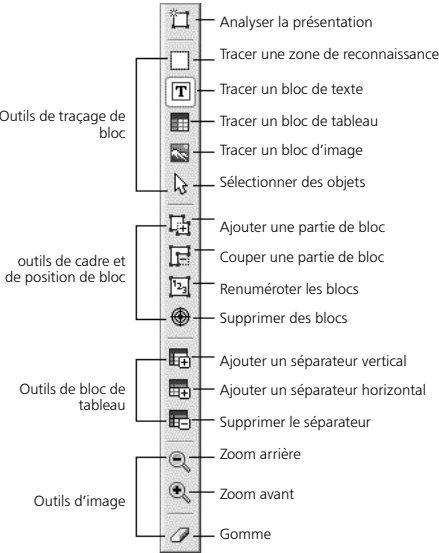
2. Utilisez l'option **Tableau sans cellules fusionnées** si vous traitez un tableau qui ne contient pas de cellules fusionnées. Par exemple :

| Température | |
|----------------|---------------|
| Degrés Celsius | Degrés Kelvin |
| -273 | 0 |
| 100 | 373 |





- la cellule **Température** est une cellule fusionnée

Remarque : ne sélectionnez pas les options **Une ligne de texte par cellule** et **Tableau sans cellules fusionnées** si votre texte comporte des tableaux de structures différentes. Le choix de ces options risque de provoquer des erreurs d'analyse de la présentation et d'avoir une incidence négative sur la qualité de la reconnaissance.

Tracer et modifier des blocs manuellement



Pour créer un nouveau bloc :

1. Sélectionnez l'un des outils suivants :
 -  pour tracer une zone de reconnaissance ;
 -  pour tracer un bloc de texte ;
 -  pour tracer un bloc d'image ;
 -  pour tracer un bloc de tableau .
2. Positionnez la souris au point où vous souhaitez placer un angle de votre bloc. Maintenez enfoncé le bouton gauche de la souris et déplacez le pointeur jusqu'au point où vous voulez placer l'angle opposé du bloc.
3. Relâchez le bouton de la souris.

Un cadre entoure la zone de l'image que vous avez sélectionnée.

Vous pouvez alors changer le type de bloc. Le type de bloc peut être l'un des suivants : **Zone de reconnaissance**, **Texte**, **Tableau**, **Image** ou **Code barre**. Pour modifier le type du bloc :


- Cliquez sur le bloc avec le bouton droit de la souris et sélectionnez l'option **Type de bloc**, puis le type de bloc de votre choix dans le menu contextuel.

Modifier des blocs**Pour déplacer les bords du bloc :**


1. Cliquez sur le bord du bloc et maintenez enfoncé le bouton gauche de la souris. Le pointeur de la souris se transforme en flèche à deux pointes.
2. Déplacez le pointeur dans la direction de votre choix.
3. Relâchez le bouton de la souris.

Remarque : si vous cliquez sur l'angle d'un bloc, vous déplacerez simultanément les bords verticaux et horizontaux du bloc.

Pour ajouter une partie de bloc rectangulaire :

1. Sélectionnez l'outil .
2. Cliquez sur le bloc auquel vous voulez ajouter une partie. Maintenez enfoncé le bouton gauche de la souris et déplacez le pointeur en diagonale. Sélectionnez la zone de l'image que vous voulez ajouter au bloc et relâchez le bouton. Le rectangle que vous avez tracé sera ajouté au bloc.
3. Si nécessaire, déplacez le bord du bloc.


Pour couper une partie de bloc rectangulaire :









1. Sélectionnez l'outil .
2. Cliquez sur le bloc dont vous voulez couper une partie. Maintenez enfoncé le bouton gauche de la souris et déplacez le pointeur en diagonale. Sélectionnez la zone de l'image que vous voulez couper et relâchez le bouton. Le rectangle sélectionné est découpé du bloc.
3. Si nécessaire, déplacez le bord du bloc.

Remarque :




1. Vous pouvez modifier les bordures des blocs en y ajoutant de nouveaux nœuds (points de séparation). Vous pouvez utiliser la souris pour déplacer les segments ainsi créés dans la direction de votre choix. Pour ajouter un nœud, appuyez sur la touche **MAJ**, puis placez le pointeur de la souris à l'endroit où vous voulez l'insérer (le pointeur se transforme alors en croix), puis cliquez sur la bordure. Le nœud se crée alors.
2. FineReader impose certains pré requis concernant la forme des blocs. Il faut notamment veiller à ce que les lignes de texte dans les blocs ne soient pas divisées pour que la reconnaissance s'effectue correctement. Pour satisfaire ces exigences, FineReader corrige automatiquement les bordures des blocs une fois que vous avez ajouté ou coupé une partie de bloc (si vous coupez la partie supérieure ou inférieure d'un bloc ou si vous tentez de couper une partie entre les deux angles du haut ou du bas, l'application réagira comme si vous aviez coupé l'angle de droit du bloc, respectivement haut ou bas), ou interdit certaines opérations lorsque vous essayez de déplacer les segments formant la bordure du bloc.

Pour sélectionner un bloc ou un groupe de blocs :


- Sélectionnez l'outil  et cliquez sur le bloc de votre choix ou appuyez sur le bouton gauche de la souris et dessinez un rectangle contenant tous les blocs que vous voulez sélectionner.

Remarque : vous pouvez sélectionner un ou plusieurs blocs en utilisant les outils habituels de sélection de blocs. Pour sélectionner plusieurs blocs, maintenez enfoncée la touche **MAJ** ou **CTRL** avec l'un des outils choisis : , ,  ou  et déplacez la flèche sur les blocs que vous voulez sélectionner. Pour inverser la sélection (pour ajouter un bloc non sélectionné ou désélectionner un bloc sélectionné), maintenez enfoncée la touche **CTRL** avec l'un des outils choisis : , ,  ou  et déplacez la flèche sur les blocs de votre choix.

Pour déplacer des blocs :


- Maintenez enfoncée la touche **ALT** avec l'un des outils choisis : , ,  ou  et déplacez les blocs.

Pour numérototer des blocs :

1. Sélectionnez l'outil .
2. Cliquez sur les blocs dans l'ordre qui vous convient. Le contenu des blocs s'affichera dans le même ordre dans le texte en sortie.

Remarque : si vous renumérotez des blocs sur une image déjà reconnue, le texte reconnu de la fenêtre **Texte** en mode brouillon sera réorganisé en conséquence.

Pour supprimer un bloc :

- Sélectionnez l'outil  et cliquez sur le bloc que vous souhaitez supprimer.
- Sélectionnez les blocs que vous voulez supprimer et appuyez sur la touche **SUPPR**.

Remarque : si vous supprimez un bloc déjà reconnu, le texte correspondant de la fenêtre **Texte** sera également supprimé.

Pour supprimer tous les blocs de l'image :

- Sélectionnez la commande **Supprimer les blocs et le texte** dans le menu **Lot**.




Remarque : si l'image dans laquelle vous supprimez des blocs a déjà été reconnue, tout le texte reconnu de la fenêtre **Texte** sera également supprimé.

Analyse manuelle de la présentation d'un tableau

Conseil : si la procédure d'analyse de la présentation d'un tableau ne divise pas correctement le tableau en lignes et colonnes, essayez de modifier les résultats de l'analyse automatique, plutôt que de supprimer tous les blocs et de les tracer manuellement. Dans la plupart des cas, la correction des résultats de l'analyse automatique prend moins de temps.

Pour modifier un tableau manuellement :

Utilisez les outils suivants de la barre d'outils **Image** pour modifier le tableau:

-  Ajouter un séparateur vertical
-  Ajouter un séparateur horizontal
-  Supprimer un séparateur

Si la cellule du tableau contient seulement une image, sélectionnez la commande **Traiter la cellule comme une image** dans la boîte de dialogue **Propriétés de bloc** (menu **Affichage>Propriétés**). Si la cellule contient à la fois du texte et des images, tracez un ou plusieurs blocs d'image dans la cellule.

Pour fusionner des cellules ou des lignes de tableau :

- Sélectionnez la commande **Fusionner les cellules du tableau** ou **Fusionner les lignes du tableau** dans le menu **Edition**.

Remarque : vous pouvez segmenter des cellules précédemment fusionnées avec la commande **Segmenter les lignes du tableau** (menu **Edition**). L'option **Fusionner les lignes du tableau** n'affecte pas la division du tableau en colonnes.

Remarque : pour éviter de tracer manuellement des séparateurs horizontaux et verticaux, tracez un bloc de tableau distinct, puis cliquez dessus avec le bouton droit et sélectionnez la commande **Analyser la structure du tableau** dans le menu contextuel. Le système trace tous les séparateurs qu'il estime nécessaires. Vous pouvez modifier le tableau si certains séparateurs sont incorrects.

Utilisation des modèles de bloc

Si vous traiter un grand nombre de documents de présentation identique, comme des formulaires ou des questionnaires, l'analyse de la présentation de chaque page peut prendre beaucoup de temps. Vous pouvez créer un modèle de bloc, contenant un ensemble standard de blocs d'un type particulier correspondant à la présentation de vos pages, puis appliquer ce modèle à toutes les pages similaires que vous devez lire.

Remarque : vous ne pouvez utiliser le modèle créé que pour les documents numérisés avec une résolution identique à celle du document pour lequel vous avez créé le modèle.

Pour créer un modèle de bloc :

1. Ouvrez une image et tracez les block automatiquement ou manuellement.
2. Sélectionnez la commande **Enregistrer les blocs** dans le menu **Image**. La boîte de dialogue **Enregistrer les blocs sous** apparaît. Tapez un nom pour le modèle de bloc.

Pour charger un modèle de bloc :

1. Cliquez sur la **fenêtre Lot** et sélectionnez les pages auxquelles vous souhaitez appliquer le modèle.
2. Sélectionnez la commande **Charger les blocs** dans le menu **Image**. La boîte de dialogue **Ouvrir des blocs** apparaît.
3. Sélectionnez le fichier de modèle approprié.
4. Sélectionnez le bouton radio **Appliquer** à approprié. Le bouton **Toutes les pages** applique le modèle de bloc à toutes les pages du lot, le bouton **Pages sélectionnées** l'applique uniquement aux pages sélectionnées.
5. Cliquez sur le bouton **Ouvrir**.

Chapitre 6

Reconnaissance

L'objectif de la reconnaissance optique de caractères (OCR) est de lire le texte de l'image source et de conserver également sa présentation. Vous devez donc définir les principaux paramètres de la reconnaissance avant de lancer celle-ci : langue de reconnaissance, type d'impression du texte reconnu et type de document. Ce chapitre décrit ces paramètres et traite d'autres questions importantes concernant la reconnaissance, illustre le rôle des différents paramètres de reconnaissance, etc.

Contenu de ce chapitre

- Informations générales sur la reconnaissance
- Langue de reconnaissance
- Type d'impression du texte source
- Autres options de reconnaissance
- Mode de reconnaissance en arrière plan
- Reconnaissance avec apprentissage
 - Apprentissage d'un gabarit utilisateur
 - Modification d'un gabarit utilisateur
- Création d'une nouvelle langue ou d'un nouveau groupe de langues
 - Création d'une langue utilisateur
 - Création d'un groupe de langues

Informations générales sur la reconnaissance

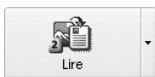
Remarque : vérifiez les options suivantes : langue de reconnaissance, type d'impression du texte reconnu et type de document avant de lancer la reconnaissance.

Vous pouvez :

1. Reconnaître un ou plusieurs blocs tracés sur une image.
2. Reconnaître une page ouverte ou toutes les pages sélectionnées dans la fenêtre Lot.
3. Reconnaître toutes les pages non reconnues du **lot**.
4. Reconnaître toutes les pages en arrière plan. Ce mode vous permet de modifier et de reconnaître les pages en même temps..
5. Reconnaître des pages en mode Apprentissage gabarit utilisateur. Ce mode sert en général à reconnaître les textes composés dans des polices décoratives ou à traiter des volumes importants (plus de cent pages) de documents de qualité d'impression médiocre.
6. Reconnaître le même lot sur plusieurs postes de travail.

Pour lancer la reconnaissance :

- Cliquez sur le bouton **2 - Lire de la barre d'outils Assistant ou**
- Sélectionnez la commande appropriée dans le menu **Traitement** :
 - Lire** - pour reconnaître la page ouverte ou toutes les pages sélectionnées dans la fenêtre Lot.
 - Lire toutes les pages** - pour reconnaître toutes les pages non reconnues du lot.
 - Lire le bloc** - pour reconnaître un ou plusieurs blocs tracés sur une image.
 - Lancer la reconnaissance en arrière plan** - pour lancer la reconnaissance en arrière plan.

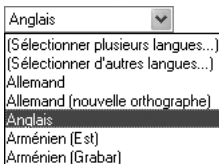


Par défaut, le bouton **2 - Lire** reconnaît l'image ouverte. Pour changer le mode du bouton, cliquez sur la flèche qui se trouve sur sa droite et sélectionnez l'option de votre choix dans le menu contextuel.

Remarque : si vous reconnaissez une page déjà reconnue, la reconnaissance ne portera que sur les blocs modifiés et ajoutés.

Langue de reconnaissance

FineReader reconnaît aussi bien les documents unilingues que multilingues (par exemple anglais-français). Pour définir la langue de reconnaissance du texte, sélectionnez-la dans la liste déroulante de la barre d'outils **Standard**.



Pour reconnaître un document multilingue :

1. Sélectionnez la commande Sélectionner plusieurs langues dans la liste de langues de la barre d'outils **Standard**. La boîte de dialogue **Langue de reconnaissance** apparaît.
2. Sélectionnez les langues appropriées.

Remarque :

1. Si vous utilisez souvent une combinaison de langues déterminée, créez un nouveau groupe de langues et ajoutez-lui les langues appropriées.
2. L'augmentation du nombre de langues de reconnaissance utilisées simultanément peut entraîner une dégradation de la qualité de la reconnaissance. Il est conseillé de limiter ce nombre à deux ou trois.
3. Avant de lire des documents, vérifiez les polices sélectionnées dans la page de l'onglet **Formatage** : ces polices doivent prendre en charge tous les caractères utilisés dans les langues de reconnaissance sélectionnées, sinon le texte ne sera pas affiché correctement (des lettres seront remplacées par le symbole ? ou la lettre j). Consultez ABBYY FineReader l'aide, section *Caractères langues de reconnaissance qui peuvent être représentées incorrectement dans la fenêtre Editeur du Texte*.

Il peut arriver que la langue de reconnaissance requise ne figure pas dans la liste. Il peut y avoir plusieurs raisons à cela :

1. Consultez ABBYY FineReader l'aide, section *Langues prises en charge*, pour la liste complète des langues prises en charge.
2. La langue ne figure pas dans la liste de la barre d'outils **Reconnaissance**. Pour ajouter la langue requise, vous devez sélectionner la commande **Sélectionner d'autres langues** dans la liste de la barre d'outils **Standard**. La boîte de dialogue **Langue de reconnaissance** apparaît. Sélectionnez la langue requise.
3. La langue a été déconnectée pendant une installation personnalisée.

Remarque : veuillez à indiquer le dossier où vous avez installé ABBYY FineReader.

Pour afficher/masquer une langue dans la liste déroulante de la barre d'outils :

- Sélectionnez la langue requise dans la boîte de dialogue **Editeur de langue** (menu **Outils>Editeur de langue**), puis sélectionnez ou désélectionnez l'option **Montrer cette langue dans la liste déroulante de la barre d'outils**.

Conseil : il est même possible de définir une langue de reconnaissance pour un bloc séparé ! Pour ce faire, cliquez avec le bouton droit de la souris sur le bloc concerné et sélectionnez **Propriétés** dans le menu contextuel. La boîte de dialogue **Propriétés** apparaît. Sélectionnez l'onglet **Bloc**, puis choisissez la langue de reconnaissance du bloc dans le champ **Langues** de la page de l'onglet.

Type d'impression du texte source

Dans la plupart des cas, le type d'impression du texte est détecté automatiquement. Pour utiliser la détection automatique du type d'impression, sélectionnez **Détection automatique** dans le groupe **Type d'impression** (menu **Outils>Options**, onglet **Reconnaissance**).

Si vous reconnaissez des sorties d'imprimante matricielle en mode brouillon ou des documents de machine à écrire, vous pouvez améliorer la qualité de la reconnaissance en spécifiant un autre **type d'impression** :

- Sélectionnez l'option **Machine à écrire** pour la reconnaissance de documents dactylographiés.
- Sélectionnez l'option **Imprimante matricielle** pour la reconnaissance de sorties d'imprimante matricielle.

software

Voici un exemple de sortie d'imprimante matricielle. Notez que les lignes des caractères se composent de points distincts.

software

Voici un exemple de document dactylographié. Toutes les lettres ont la même largeur (comparez par exemple les lettres w et a).

Pour modifier le type d'impression :

- Sélectionnez l'option de votre choix dans la page de l'onglet **Reconnaissance** de la boîte de dialogue **Options** (menu **Outils >Options**).

Remarque : lorsque la reconnaissance de documents issus d'une imprimante matricielle ou d'une machine à écrire est terminée, pensez à rétablir le paramètre **Détection automatique** pour reconnaître correctement le texte normal.

Autres options de reconnaissance

Montrer l'image pendant la reconnaissance

La reconnaissance d'un grand nombre de pages est plus rapide si l'image traitée n'est pas affichée. Pour effectuer la reconnaissance sans afficher l'image :

- Désélectionnez l'option **Montrer l'image pendant la reconnaissance** dans la page de l'onglet **Général** (menu **Outils >Options**).

Direction du texte

Si l'application ne reconnaît pas correctement les blocs contenant du texte vertical (bloc de texte ou cellule de tableau) :

- Cliquez sur le bloc de texte vertical à l'aide du bouton droit de la souris, puis sélectionnez la commande **Propriétés** du menu contextuel. La boîte de dialogue **Propriétés de bloc** s'affiche. Sélectionnez l'élément requis dans la liste intitulée **Direction du texte** et relancez la reconnaissance de l'image.



Texte inversé ou basculé

Si l'application ne reconnaît pas correctement le texte inversé ou basculé (bloc de texte, cellule de tableau ou tableau entier) :

- Cliquez sur le bloc approprié à l'aide du bouton droit de la souris, puis sélectionnez la commande **Propriétés** du menu contextuel. La boîte de dialogue **Propriétés de bloc** s'affiche. Sélectionnez l'article **Inversé** ou **Basculé**, puis relancez la reconnaissance de l'image.

Reconnaissance en arrière plan

Le mode de reconnaissance en arrière plan s'avère utile si vous souhaitez modifier les pages déjà reconnues tout en poursuivant la reconnaissance. Pour lancer la reconnaissance en arrière-plan :

- Sélectionnez la commande **Lancer la reconnaissance en arrière-plan** dans le menu **Traitement**. Le signe  apparaît dans la ligne d'état, en bas de la fenêtre principale de Fine Reader. Si le mode d'affichage **Détails** est actif dans la fenêtre **Lot** (pour activer ce mode, cliquez avec le bouton droit de la souris dans la fenêtre **Lot** et sélectionnez **Afficher>Détails** dans le menu contextuel), la page en cours de reconnaissance contient l'icône  dans la colonne **Ouvert par**.

Lorsque le mode de reconnaissance en arrière-plan est actif, la reconnaissance reprend automatiquement quand une page non reconnue est ajoutée au lot.

Remarque : si votre lot contient de nombreuses pages, vous remarquerez que la vitesse de reconnaissance est accrue sur les systèmes multiprocesseurs en **mode Arrière-plan**.

Pour arrêter la **reconnaissance en arrière-plan** :

- Sélectionnez la commande **Arrêter la reconnaissance en arrière-plan** dans le menu **Traitement**.

Remarque : le mode de reconnaissance en arrière-plan utilise les options de reconnaissance en cours.

Reconnaissance avec apprentissage

Comme indiqué précédemment, FineReader lit sans difficulté les textes imprimés dans presque n'importe quelle police et il est pratiquement insensible à la qualité de l'impression. En général, vous pourrez reconnaître n'importe quel texte sans apprentissage préalable. FineReader fournit toutefois des outils d'apprentissage pour les gabarits utilisateur.

Le mode Apprentissage gabarit utilisateur peut vous servir pour :

1. reconnaître des textes imprimés dans des polices décoratives ;
2. reconnaître des textes contenant des caractères spéciaux (par exemple des symboles mathématiques) ;
3. reconnaître des volumes importants (plus de cent pages) de texte de qualité médiocre.

Conseil : n'utilisez le mode **Apprentissage gabarit utilisateur** que dans les cas mentionnés. Dans les autres situations, vous obtiendrez peut-être une très légère amélioration de la qualité de reconnaissance, mais qui ne compensera pas le temps consacré à l'apprentissage.

L'apprentissage de gabarit se déroule comme suit. Dans un premier temps, vous reconnaissez une ou deux pages dans un mode d'apprentissage spécial. Le résultat est appelé gabarit. Ensuite, FineReader utilise ce gabarit pour reconnaître plus facilement le reste du texte.

Il arrive parfois que deux ou trois caractères soient « collés ». Il est souvent impossible de les séparer ; FineReader ne peut pas les inclure dans des cadres contenant un caractère chacun. Si c'est le cas (il n'est pas possible de déplacer le cadre pour qu'il contienne un caractère entier sans englober de fragment des caractères adjacents), vous pouvez apprendre à FineReader à reconnaître la combinaison de caractères « inséparables ». Des exemples de combinaisons inséparables classiques seront ff, fi, fl, etc. Ces combinaisons sont désignées ligatures.

Remarque :

1. Le gabarit ainsi créé n'est utile que pour reconnaître les textes imprimés dans la même police de la même taille et numérisés avec la même résolution que le document sur lequel vous avez effectué l'apprentissage du gabarit.
 2. chaque gabarit est créé pour un lot particulier. Si vous supprimez le lot, vous supprimez le gabarit en même temps. Vous pouvez mettre n'importe quel gabarit à la disposition d'autres lots. Pour ce faire, enregistrez les options de lot dans un fichier au format de modèle de lot.
 3. Si vous devez reconnaître des textes imprimés dans une police différente, n'oubliez pas de déconnecter le gabarit. Cliquez sur le bouton radio **Ne pas utiliser de gabarit utilisateur** de la page de l'onglet **Reconnaissance**, menu **Outils>Options**.
-

Pour effectuer l'apprentissage d'un gabarit utilisateur :

1. Lancez le mode **Apprentissage gabarit utilisateur** : cliquez sur le bouton radio **Apprentissage gabarit utilisateur** de la page de l'onglet **Reconnaissance**, menu **Outils>Options**, groupe **Apprentissage**. Le nom de gabarit par défaut (« Par défaut ») s'affiche sur la ligne d'état.
2. Cliquez sur le bouton **2 - Lire**.
3. Effectuez l'apprentissage du gabarit : reconnaissez une ou deux pages en mode **Apprentissage gabarit utilisateur**.
Les caractères appris sont enregistrés dans le gabarit par défaut. Une fois l'apprentissage du gabarit terminé, FineReader enregistre le gabarit créé (**Default.pat**) dans le dossier de lot actif.
4. Modifiez votre gabarit.
5. Désactivez le mode apprentissage (cliquez sur le bouton radio **Utiliser gabarit utilisateur** de la page de l'onglet **Reconnaissance**).
6. Effectuez la reconnaissance du reste du texte en cliquant sur le bouton **2 - Lire**.



Remarque :

1. Pour créer plusieurs gabarits pour le même lot, ouvrez la boîte de dialogue Editeur de gabarit (cliquez sur le bouton **Editeur de gabarit** de la page de l'onglet **Reconnaissance** ou sélectionnez la commande **Outils>Editeur de gabarit**). Créez un nouveau gabarit (cliquez sur le bouton **Nouveau** dans la boîte de dialogue) et sélectionnez-le (cliquez sur le bouton **Définir comme actif**). L'utilisation d'un gabarit créé est identique à celle du gabarit par défaut ; reportez-vous aux étapes 1-5. N'oubliez pas qu'un seul gabarit peut être actif à la fois.
2. Si vous avez créé plusieurs gabarits pour le même lot, le gabarit actif est le dernier créé. Le nom du gabarit actif est affiché dans la barre d'état. Pour activer un autre gabarit, sélectionnez-le dans la liste de la boîte de dialogue Editeur de gabarit (menu **Outils>Editeur de gabarit**) et cliquez sur le bouton **Définir comme actif**. Cliquez ensuite sur le bouton radio **Utiliser gabarit utilisateur** de la page de l'onglet **Reconnaissance**, menu **Outils>Options**, groupe **Apprentissage gabarit utilisateur**.
3. Si l'option **Utiliser gabarits intégrés** est sélectionnée, FineReader tente de lire le texte à l'aide des gabarits intégrés et ne s'arrête que sur les caractères incertains. Si vous apprenez au système à lire les polices décoratives et/ou non standard (telles que Tibetan), FineReader risque de lire les caractères incorrectement avec les gabarits intégrés. Dans ce cas, désactivez la case à cocher **Utiliser gabarits intégrés** de la page de l'onglet **Reconnaissance** et apprenez au système à reconnaître chacun des caractères.


Apprentissage d'un gabarit utilisateur

1. Assurez-vous que le bouton radio **Apprentissage gabarit utilisateur** de la page de l'onglet **Reconnaissance** (menu **Outils>Options**), groupe **Apprentissage** est activé.
2. Cliquez sur le bouton **2 - Lire**. FineReader lance la reconnaissance. Lorsqu'il rencontre un caractère qu'il ne connaît pas, la boîte de dialogue **Apprentissage de gabarit** apparaît. L'image du caractère y est affichée.

Apprentissage pour reconnaître un caractère :

Le cadre de la fenêtre supérieure de la boîte de dialogue doit contenir **un caractère unique entièrement contenu** dans le cadre. Si celui-ci ne contient qu'une partie de caractère ou plus d'un caractère, cliquez sur les bords du cadre et déplacez-les pour que les conditions ci-dessus soient réunies. Les boutons  et  permettent également de déplacer les bords du cadre (cette option est utile pour l'apprentissage des symboles en italiques ; voir plus loin). Lorsque vous avez terminé, tapez le caractère et cliquez sur le bouton **Apprentissage**.

Remarque :



1. Le système ne peut apprendre à lire que les caractères de l'alphabet de la langue active. Si vous apprenez à FineReader à lire des caractères que vous ne pouvez pas saisir au clavier, vous pouvez utiliser une combinaison de deux caractères pour les signaler ou copier le caractère requis depuis la **Table des caractères** (cliquez sur le bouton  dans la boîte de dialogue Apprentissage de gabarit pour l'ouvrir).
2. Si vous souhaitez apprendre au système à retenir la mise en forme des caractères, sélectionnez **Italique** ou **Gras** dans la boîte de dialogue Apprentissage de gabarit avant de cliquer sur le bouton **Apprentissage**.
3. Veillez à saisir les caractères en majuscules lorsque vous effectuez l'apprentissage d'images de majuscules et en minuscules pour les images de minuscules.

Si vous faites une erreur lors de l'apprentissage, cliquez sur le bouton **Précédent** pour restaurer la position précédente du cadre ; la dernière paire « caractère-image » est alors supprimée du gabarit. Cette annulation est limitée : elle ne s'applique qu'au dernier mot.

Apprentissage de la reconnaissance des ligatures

Une ligature est une combinaison de deux caractères « collés », par exemple, fi, fl, ffi, etc. Ces caractères sont difficiles à séparer car ils sont « collés » à l'impression. Il est préférable de les traiter comme un caractère composé « simple ».

L'apprentissage des ligatures se fait de la même manière que celui des caractères séparés :

1. Tapez la combinaison de caractères requise, puis cliquez sur le bouton **Apprentissage**.
2. Le cadre de la fenêtre supérieure de la boîte de dialogue doit contenir **la ligature en entier**. Déplacez le bord du cadre à l'aide de la souris ou en cliquant sur les boutons  et .

Chaque gabarit peut contenir jusqu'à 1 000 nouveaux caractères. Toutefois, il est déconseillé de créer un trop grand nombre de ligatures, car cela risque de dégrader la qualité de la reconnaissance.

Lors de l'apprentissage de FineReader, tenez compte des limitations suivantes :

1. FineReader ne distingue pas certains caractères habituellement considérés comme différents. Ces images sont reconnues comme un seul et même caractère. Par exemple, les apostrophes simple (') , droite (') et gauche (') sont considérées comme un même caractère dans le gabarit, à savoir l'apostrophe simple. Ainsi, vous ne verrez jamais apparaître les apostrophes droite et gauche dans le texte reconnu, même si vous apprenez au logiciel à les reconnaître.
2. Dans certains cas, une image est reconnue comme un caractère particulier, selon l'environnement.

Modification d'un gabarit utilisateur

Il peut être conseillé de modifier un nouveau gabarit avant de vous en servir pour la reconnaissance. Un gabarit dépourvu de l'apprentissage correct a des effets négatifs sur la qualité de la reconnaissance et il est utile de le vérifier.

Le gabarit ne doit contenir que des caractères ou ligatures entiers. Il faut retirer du gabarit les caractères dont les bords sont endommagés et ceux qui sont incorrectement identifiés.

Pour modifier un gabarit utilisateur :

1. Sélectionnez la commande **Editeur de gabarit** dans le menu **Outils**. La boîte de dialogue **Editeur de gabarit** apparaît.
2. Sélectionnez le gabarit approprié et cliquez sur le bouton **Modifier**. La boîte de dialogue **Editeur de gabarit** apparaît.
3. Sélectionnez un caractère et cliquez sur le bouton **Propriétés** pour modifier le titre du caractère et définir la police correcte : italique, gras, indice ou exposant. Vous pouvez également cliquer sur le bouton **Supprimer** pour retirer du lot les caractères dont l'apprentissage est incorrect.

Langues utilisateur et groupes de langues

En complément des langues et groupes de langues intégrés, vous pouvez créer une langue ou constituer un nouveau groupe (à partir des langues prises en charge par FineReader) et l'utiliser lors de la reconnaissance.

La création d'une langue peut s'imposer dans les cas suivants :

1. Pour vous connecter à un dictionnaire utilisateur.
 - Par exemple, vous devez reconnaître un texte en anglais contenant de nombreuses abréviations. Vous pouvez créer un dictionnaire d'abréviations, créer une nouvelle langue et connecter le dictionnaire à la langue. Vous pouvez ensuite constituer un nouveau groupe de langues constitué de l'anglais (avec le dictionnaire système) et de votre nouvelle langue (avec le dictionnaire d'abréviations) et vous servir de ce groupe pour reconnaître vos textes.
2. Pour reconnaître certains documents spéciaux, par exemple :
 - Des listes d'articles ne comportant que des références de produits. Ces références ne contiennent souvent que des chiffres et certaines lettres. Vous pouvez créer une nouvelle langue ne contenant que les caractères requis et l'utiliser pour reconnaître ces documents.
 - Des documents entièrement en capitales. Vous améliorerez la qualité de la reconnaissance si vous créez une langue interdisant toutes les minuscules.

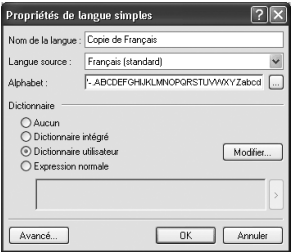
Il est utile de créer un groupe de langues si vous utilisez fréquemment une combinaison de langues déterminée.

Pour créer une langue ou un groupe de langues, ouvrez la boîte de dialogue de l'**Editeur de langue** (menu **Outils**, commande **Editeur de langue**).


Création d'une nouvelle langue

Pour créer une langue de reconnaissance :

- 1. Sélectionnez la commande **Editeur de langue** dans le menu **Outils**.
- 2. Cliquez sur le bouton **Nouveau**. Dans la boîte de dialogue qui s'affiche, sélectionnez le bouton radio **Créer une copie de la langue**, puis une langue source pour la nouvelle langue.
- 3. La boîte de dialogue **Propriétés de langue simples** s'affiche.



Définissez les paramètres suivants pour la nouvelle langue (tous les paramètres sont définis dans la boîte de dialogue **Propriétés de langue simples**) :

- 1. Nom de la nouvelle langue.
- 2. Alphabet de base à utiliser. Ce paramètre est défini dans le champ **Alphabet**. Si nécessaire, modifiez l'alphabet en cliquant sur le bouton .
- 3. Le dictionnaire à utiliser par le système (à la fois pour la reconnaissance et la vérification orthographique). Les options suivantes sont disponibles :
 - **Aucun** (aucun dictionnaire pour la langue)
 - **Intégré** (dictionnaire fourni avec FineReader)
 - **Dictionnaire utilisateur**

Si vous souhaitez ajouter des mots au dictionnaire, ou vous connecter à un ancien dictionnaire utilisateur ou un fichier texte au format Windows (ANSI) ou Unicode (le seul impératif est que les termes soient séparés par des espaces ou autre caractère non alphabétique) cliquez sur le bouton **Modifier le dictionnaire**.

Remarque : la vérification orthographique considère comme corrects les termes du dictionnaire utilisateur s'ils apparaissent comme suit dans le texte : capitales définies dans le dictionnaire, toutes les lettres minuscules, toutes les lettres majuscules, première lettre majuscule. Exemples :

| Capitales définies dans le dictionnaire : | Les formes suivantes seront correctes : |
|---|---|
| abc | abc, Abc, ABC |
| Abc | abc, Abc, ABC |
| ABC | abc, Abc, ABC |
| aBc | aBc, abc, Abc, ABC |

- **Expression normale** (spécifiant la grammaire de la nouvelle langue ; consultez la section Expressions normales).

Remarque :

- 1. Vous pouvez cliquer sur le bouton **Avancé** dans la boîte de dialogue **Propriétés de langue simples** pour définir certaines propriétés avancées de la nouvelle langue, telles que les caractères ignorés dans les mots, les caractères interdits, etc.

2. La nouvelle langue utilisateur créée est enregistrée par défaut dans le dossier de lot. ABBYY FineReader Corporate Edition permet de spécifier un autre dossier. Pour plus d'informations sur le travail de groupe avec des langues utilisateur et leurs dictionnaires, consultez la section Travail de groupe avec les mêmes langues et dictionnaires utilisateur.

Création d'un nouveau groupe de langues

Si vous effectuez fréquemment la reconnaissance de textes dans certaines combinaisons de langues, par exemple l'anglais et l'allemand, vous pouvez créer un groupe de langues associant ces langues. Le groupe sera affiché dans la liste de langues de la barre d'outils **Standard**.

Remarque : vous pouvez spécifier les langues de reconnaissance directement depuis la liste de langues de la barre d'outils **Standard**. Pour cela, sélectionnez l'option **Sélectionner plusieurs langues** dans la liste. La boîte de dialogue **Langue de reconnaissance** apparaît. Sélectionnez les langues souhaitées dans la boîte de dialogue.

Pour créer un groupe de langues de reconnaissance :

1. Sélectionnez la commande **Editeur de langue** dans le menu **Outils** et cliquez sur le bouton **Nouveau**. Une boîte de dialogue apparaît. Sélectionnez la commande **Créer un nouveau groupe de langues** dans la boîte de dialogue.
2. La boîte de dialogue **Propriétés de groupe de langues** apparaît.



Définissez les paramètres suivants de groupe de langues (tous les paramètres sont définis dans la boîte de dialogue **Propriétés de groupe de langues**):

1. Nom du groupe.
2. Langues composant le groupe.

Remarque :

1. Dans certains cas, vous êtes sûr que des caractères ne seront jamais trouvés dans les textes reconnus. Vous pouvez alors spécifier ces caractères (désignés caractères interdits) dans les propriétés du groupe de langues. Cette définition de caractères interdits peut améliorer la vitesse et la qualité de la reconnaissance. Pour spécifier ces caractères, cliquez sur le bouton **Avancé** dans la boîte de dialogue **Propriétés de groupe de langues**. La boîte de dialogue **Propriétés avancées** apparaît. Spécifiez les caractères interdits sur la ligne **Caractères interdits**.
2. Le nouveau groupe de langue utilisateur créé est enregistré par défaut dans le dossier de lot. ABBYY FineReader Corporate Edition permet de spécifier un autre dossier. Pour plus d'informations sur le travail de groupe avec des langues utilisateur et leurs dictionnaires, consultez la section Travail de groupe avec les mêmes langues et dictionnaires utilisateur.

Chapitre 7

Vérification et modification de texte

Quand la reconnaissance est terminée, le texte reconnu apparaît dans la fenêtre Texte. Cette fenêtre constitue l'éditeur intégré de FineReader. Elle vous permet de vérifier le résultat de la reconnaissance et de modifier le texte obtenu.

L'éditeur de FineReader comporte deux caractéristiques spéciales :

1. Un système intégré de vérification orthographique (consultez ABBYY Fine Reader l'aide, section *Langues prises en charge*, pour la liste des langues prise en charge par la vérification..
2. Une aide visuelle commode : l'image source de la ligne de texte que vous vérifiez est affichée dans la fenêtre **Zoom**.

Le système de vérification orthographique contient :

1. Des outils permettant de trouver les mots douteux (contenant des caractères incertains).
2. Des outils détectant les fautes d'orthographe.
3. Des outils pour ajouter des termes inconnus au dictionnaire de FineReader. L'ajout de termes améliore la qualité de la reconnaissance.

Contenu du chapitre :

- Vérification du texte dans FineReader
- Options de vérification
- Ajout et suppression de termes dans le dictionnaire utilisateur
- Modification du texte dans FineReader
- Modification de tableaux

Vérification du texte dans FineReader

Les caractères et les mots reconnus de façon incertaine qui sont absents du dictionnaire sont mis en évidence dans différentes couleurs. Par défaut, le bleu clair indique les caractères incertains et le rose indique les mots absents du dictionnaire. Pour modifier la couleur :

- Sélectionnez l'option **Caractère incertain** (ou **Absent du dictionnaire**) ainsi que la couleur de votre choix dans l'option **Couleur** de la page de l'onglet **Affichage** (menu **Outils>Options**) dans le groupe **Apparence**.

Pour vérifier les résultats de la reconnaissance :

1. Cliquez sur le bouton **3 - Vérification orthographique** de la **Barre d'outils Assistant** (ou sélectionnez la commande **Vérification orthographique** dans le menu **Outils**).
2. La boîte de dialogue **Vérification orthographique** s'ouvre.



3. La boîte de dialogue **Vérification orthographique** comporte trois fenêtres. La fenêtre supérieure est identique à la fenêtre **Zoom** de FineReader, elle affiche l'image originale du mot éventuellement mal orthographié. La fenêtre du centre affiche le mot éventuellement mal orthographié lui-même, tandis que la ligne au-dessus affiche le nom du type possible d'erreur. La fenêtre **Suggestions** du bas propose des suggestions de remplacement (le cas échéant). Les suggestions sont basées sur le dictionnaire sélectionné dans la liste déroulante **Langue du dictionnaire** ; vous pouvez sélectionner la langue de votre choix dans cette liste.



Remarque : vous pouvez agrandir la boîte de dialogue **Vérification orthographique** pour faciliter les opérations de vérification et d'édition. Pour cela, cliquez sur le bord de la boîte de dialogue ; le pointeur de la souris se transforme en flèche à deux pointes. Faites glisser le bord pour agrandir ou réduire la boîte de dialogue.

4. Procédure à suivre si des mots sont mal orthographiés :
 - Cliquez sur le bouton **Ignorer** pour laisser le mot inchangé.
 - Cliquez sur le bouton **Tout ignorer** pour laisser inchangés tous les mots identiques dans le texte.

Remarque : lorsque vous cliquez sur le bouton **Ignorer** ou **Tout ignorer**, le drapeau « incertain » est supprimé du mot : le système considère que le mot ne contient pas de caractères non reconnus ou incertains ; ces caractères ne sont plus mis en évidence. Lorsque vous exportez ces mots au format PDF et que vous activez le mode **Remplacer les mots incertains par des images**, ils ne seront pas remplacés par des images.

- Sélectionnez une suggestion de remplacement et cliquez sur le bouton **Remplacer** ou **Tout remplacer** pour remplacer le mot en cours ou toutes ses occurrences dans le texte. Si la fenêtre **Suggestions** n'affiche aucune suggestion correcte pour le mot, modifiez ce dernier ou les mots avoisinants de la fenêtre, située au centre de la boîte de dialogue (attention : lorsque vous passez en mode de modification, certains boutons changent leurs actions et leurs légendes sont également modifiées). Cliquez alors sur le bouton **Confirmer** (**Tout confirmer**) pour modifier le mot actuel (tous les mots de ce type) dans le texte et pour sélectionner le mot suivant reconnu de manière incertaine.
- Cliquez sur **Ajouter** pour ajouter le mot au dictionnaire. Lorsque le mot a été ajouté, l'application considère que toutes les futures occurrences de ce mot sont correctes dans toutes ses formes.
- Cliquez sur **Options** pour déterminer les paramètres de vérification orthographique.
- Cliquez sur **Fermer** pour refermer la boîte de dialogue.

Se déplacer entre des mots incertains

Pour consulter rapidement les résultats de reconnaissance, vous pouvez utiliser le bouton  et le bouton  pour aller au mot incertain suivant et précédent.

Vous pouvez également utiliser les touches de raccourci **F4 (MAJ F4)** pour naviguer entre des mots incertains.

Options de vérification et de modification de texte

Ces options sont définies dans la page de l'onglet Vérification orthographique (menu **Outils>Options**).

- Niveau d'affichage d'erreur

Remarque : cette option doit être définie avant de lancer la reconnaissance.

- Arrêt sur les caractères incertains
- Arrêt sur les mots absents du dictionnaire
- Arrêt sur les mots composés
- Ignorer les mots contenant des chiffres
- Corriger les espaces avant et après les ponctuations

Niveau d'affichage d'erreur

La liste **Niveau d'affichage d'erreur** vous permet de sélectionner parmi les valeurs suivantes :

- **Aucun** - les erreurs de reconnaissance ne sont pas mises en évidence.
- **Standard** - les caractères non reconnus et incertains sont mis en évidence.
- **Approfondi** - comme Standard, mais les mots absents du dictionnaire sont également mis en évidence.

Remarque : le nombre d'erreurs affiché dans la fenêtre **Texte** changera après la relecture du document.

Arrêt sur les caractères incertains

Si cette option est définie, la vérification orthographique s'arrêtera sur les termes comportant des caractères incertains.

Arrêt sur les mots absents du dictionnaire

Si cette option est définie, la vérification orthographique s'arrêtera sur les termes absents du dictionnaire. Si un mot est absent du dictionnaire, cela peut indiquer qu'il a été lu incorrectement.

Arrêt sur les mots composés

Si cette option est définie, la vérification orthographique s'arrêtera sur les termes absents du dictionnaire et qui peuvent être construits selon les modèles morphologiques disponibles ou à partir des mots du dictionnaire.

Ignorer les mots contenant des chiffres

Si cette option est définie, la vérification orthographique considèrera corrects tous les mots contenant des chiffres, sauf s'ils contiennent des caractères incertains.

Corriger les espaces avant et après les ponctuations

Si cette option est définie, la vérification orthographique ne s'arrêtera pas si elle détecte des espaces incorrects avant ou après les marques de ponctuation, mais corrigera automatiquement ces espaces.

Ajout et suppression de mots dans le dictionnaire utilisateur

Ajouter un mot au dictionnaire utilisateur

L'enrichissement du dictionnaire est une bonne manière d'améliorer la qualité de la reconnaissance. Au cours de la reconnaissance, FineReader vérifie les mots qu'il reconnaît dans le dictionnaire. Il est judicieux d'ajouter les mots inconnus les plus fréquents (par exemple les abréviations, les noms et autres termes fréquemment utilisés dans vos documents).

L'une des caractéristiques spécifiques du système de vérification orthographique de FineReader est qu'il n'ajoute pas simplement le mot sous la forme où il l'a trouvée dans le texte, mais crée le paradigme du mot, c'est-à-dire, l'ensemble de toutes ses formes. Cette fonction permet à FineReader de « connaître » le mot dans sa globalité juste après avoir rencontré la première de ses formes dans le texte, de sorte que par la suite toutes les autres formes peuvent être reconnues en toute confiance.

Pour ajouter un mot au dictionnaire pendant la vérification orthographique :

- Cliquez sur le bouton **Ajouter** dans la boîte de dialogue **Vérification orthographique**.

Définissez les paramètres suivants dans la boîte de dialogue **Forme de base**.

1. **Éléments du discours** (nom, adjectif, verbe, non fléchi).
2. Si le mot s'écrit toujours avec une première lettre capitale, sélectionnez l'option **Nom propre**. Si vous ajoutez une abréviation, sélectionnez l'option **Abréviation**.
3. La forme de base du mot.

Cliquez sur **OK**. La boîte de dialogue **Créer un paradigme** s'ouvre. Dans cette boîte de dialogue, FineReader vous pose des questions concernant les formes du mot. Vos réponses aident FineReader à construire le paradigme du mot que vous ajoutez. Cliquez sur **Oui** ou sur **Non** lorsque FineReader vous pose une question. En cas d'erreur, cliquez sur le bouton **De nouveau** pour répondre correctement à la question. Le paradigme construit s'affiche dans la boîte de dialogue **Paradigme**.

Remarque :

1. Si vous ne souhaitez pas construire le paradigme et préférez toujours ajouter les mots sous forme non fléchiée, sélectionnez l'option **Sauter l'invite de forme de mot** (dictionnaire anglais seulement) de la page de l'onglet **Vérification orthographique** (menu **Outils>Options**).
 2. Vous pouvez également ajouter des mots lorsque vous voyez la liste des mots ajoutés. Pour cela, sélectionnez la commande **Afficher les dictionnaires** dans le menu **Outils**. La boîte de dialogue **Sélection de dictionnaire** s'ouvre. Choisissez la langue dans la boîte de dialogue **Sélection de dictionnaire** et cliquez sur **Afficher**. Le dictionnaire accompagné de la liste des mots ajoutés s'ouvre. Vous pouvez ajouter votre mot en cliquant sur le bouton **Ajouter**.
 3. La construction de paradigmes pour les termes que vous ajoutez dans les dictionnaires utilisateur n'est disponible que pour l'arménien (oriental, occidental, Grabar), l'anglais, l'italien, le français, l'allemand (ancienne et nouvelle orthographe), le russe, l'espagnol et l'ukrainien.
-

Si le mot ajouté existe dans le dictionnaire, le système vous en avertit. Dans ce cas, vous pouvez consulter son paradigme. Si vous pensez que le paradigme existant est incorrect (c'est souvent le cas lorsque vous ajoutez un homonyme appartenant à un autre élément de discours), vous pouvez en construire un autre (bouton **Ajouter** de la boîte de dialogue **Ajouter un terme**).

Conseil :

1. FineReader vous permet d'importer des dictionnaires utilisateur créés dans des versions précédentes (FineReader 3.0, 4.0 et 5.0).
 2. FineReader vous permet d'importer votre dictionnaire utilisateur (*.dic) créé dans Microsoft Word 6.0, 7.0, 8.0 et 9.0.
-

Pour importer un dictionnaire :

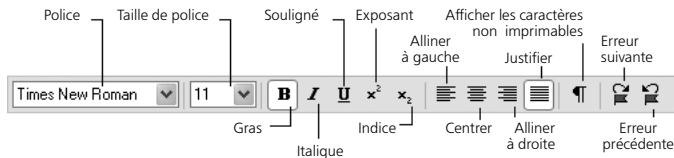
1. Sélectionnez la commande **Afficher les dictionnaires** dans le menu **Outils**, puis sélectionnez la langue du dictionnaire et cliquez sur le bouton **Afficher**.
2. Cliquez sur le bouton **Importer** dans la boîte de dialogue et sélectionnez des fichiers portant les extensions *.pmd, *.txt ou *.dic.

Pour supprimer un mot du dictionnaire :

1. Sélectionnez la commande **Afficher les dictionnaires** dans le menu **Outils**. Sélectionnez la langue et cliquez sur le bouton **OK**. Une boîte de dialogue s'ouvre.
2. Sélectionnez le mot que vous souhaitez supprimer et cliquez sur le bouton **Supprimer**.

Modification de texte dans FineReader

Remarque : si la fenêtre de texte de FineReader n'affiche pas les caractères correctement (du symbole ? ou de la lettre ; s'affichent à la place de tout ou partie des lettres), cela signifie que la police en cours ne prend pas totalement en charge l'alphabet de votre langue de reconnaissance. Sélectionnez une police prenant en charge le jeu de caractères de reconnaissance entier (par exemple Arial Unicode ou Bitstream Cyberbit) dans le groupe **Polices** de la page de l'onglet **Formatage** (menu **Outils>Options**), puis lisez à nouveau le document. Consultez ABBYY FineReader l'aide, section Caractères langues de reconnaissance qui peuvent être représentées incorrectement dans la fenêtre **Editeur du Texte**.



Lorsqu'une page a été lue, son texte s'affiche dans la fenêtre **Texte**. Lorsque vous envoyez le texte vers une application externe, sa mise en page est conservée conformément aux options de préservation de la présentation. Définissez ces options dans la page de l'onglet **Formatage** (menu **Outils > Options**) et dans les boîtes de dialogue correspondant aux différents **formats**.

Les caractères incertains sont mis en évidence. Vous pouvez annuler la mise en évidence de ces mots ; il suffit pour cela de désélectionner l'option **Mettre en évidence les caractères incertains** de la page de l'onglet **Affichage** (menu **Outils > Options**).

L'éditeur FineReader possède deux modes d'affichage des documents : le mode normal (mise en page standard) et le mode brouillon.

Le mode normal affiche les blocs avec le texte, les tableaux et les images reconnus, tels qu'ils apparaissent dans l'image source. Ainsi, ce mode conserve la mise en page standard du document source : colonnes, tableaux, images et lettrines (grandes lettres occupant plusieurs lignes d'un paragraphe). Le bloc dans lequel le pointeur se trouve est activé. Lorsque vous déplacez le pointeur en appuyant sur les touches du clavier, l'ordre de navigation entre les blocs est déterminé par leur numérotation dans l'image source. Si le texte d'un bloc est trop volumineux pour tenir dans ce bloc (par exemple, vous avez modifié le texte, qui prend maintenant trop de place), certaines parties de texte des blocs inactivés risquent de ne pas être visibles ; si tel est le cas, des points rouges apparaissent sur son bord. Lorsqu'un bloc est activé, ses bords s'élargissent pour afficher la totalité du texte contenu dans ce bloc.

Les caractéristiques suivantes du texte ne s'affichent pas dans le mode brouillon : retrait à gauche ; alignement des paragraphes (tous les paragraphes sont alignés à gauche) ; couleurs du texte et du fond. Une police de taille unique (12 pt par défaut) est utilisée dans tout le texte affiché en mode brouillon. La police et ses effets (gras, italique, souligné, exposant et indice) sont conservés dans le mode brouillon.

Vous pouvez basculer du mode brouillon au mode normal et réciproquement en cliquant sur  (pour sélectionner le mode normal) ou sur  (pour sélectionner le mode brouillon) dans la fenêtre **Texte**.

Vous pouvez modifier la taille de la police pour afficher le texte en mode brouillon. Pour cela :

1. Sélectionnez **Options** dans le menu **Outils**.
2. Définissez votre police dans l'option **Taille de la police d'édition brouillon** de la page de l'onglet **Affichage**.

L'éditeur intégré de FineReader contient les fonctions d'édition de texte suivantes :

- Copier, couper, coller
- Rechercher et remplacer
- Style et attributs des polices
- Alignement du texte
- Annuler et rétablir

Copier, couper, coller

1. Avant d'utiliser les commandes copier et couper, sélectionnez le texte.
2. Suivez les instructions ci-dessous, selon ce que vous voulez faire :

Pour copier la sélection :



- Vous pouvez cliquer sur le bouton **Copier** de la barre d'outils **Standard** ou
- sélectionner la commande **Copier** dans le menu **Édition** ou dans le menu contextuel, ou encore
- appuyer sur **CTRL+C**

Pour couper la sélection :



- Vous pouvez cliquer sur le bouton **Couper** de la barre d'outils **Standard** ou
- sélectionner la commande **Couper** dans le menu **Édition** ou dans le menu contextuel, ou encore
- appuyer sur **CTRL+X**

Pour coller le texte copié :



- Vous pouvez cliquer sur le bouton **Coller** de la barre d'outils **Standard** ou
- sélectionner la commande **Coller** dans le menu **Édition** ou dans le menu contextuel, ou encore
- appuyer sur **CTRL+V**

Rechercher et remplacer

Pour rechercher un mot ou une phrase dans le texte que vous modifiez :


1. Vous pouvez :
 - sélectionner la commande **Chercher** dans le menu **Édition** ou
 - appuyer sur **CTRL+F**
2. La boîte de dialogue **Rechercher** s'ouvre. Tapez le mot ou la phrase que vous souhaitez rechercher sur la ligne **Chercher** de la boîte de dialogue et définissez les paramètres de recherche.

Remarque : pour rechercher le même mot avec les mêmes paramètres, appuyez simplement sur la touche **F3**.

Pour chercher et remplacer un mot ou une phrase dans le texte que vous modifiez :


1. Vous pouvez :
 - sélectionner la commande **Remplacer** dans le menu **Édition** ou
 - appuyer sur **CTRL+H**
2. La boîte de dialogue **Remplacer** s'ouvre. Tapez le mot ou la phrase que vous souhaitez rechercher sur la ligne **Chercher** de la boîte de dialogue, puis tapez le mot ou la phrase de remplacement sur la ligne **Remplacer par** et définissez les paramètres de recherche.

Style et attributs des polices

1. Cliquez sur le mot ou sélectionnez le texte dont vous souhaitez modifier la police.
2. Les actions suivantes sont disponibles :
 - Vous pouvez cliquer sur le bouton (par exemple ) sur la barre **Formatage** ou
 - cliquer avec le bouton droit de la souris dans la fenêtre **Texte** et sélectionner la commande **Propriétés de caractère** dans le menu contextuel. La boîte de dialogue **Caractère** s'ouvre. Sélectionnez le type de police de votre choix et définissez ses paramètres dans la boîte de dialogue ou
 - appuyer sur **CTRL+B** pour obtenir des caractères gras, sur **CTRL+I** pour obtenir des caractères en italique ou sur **CTRL+U** pour souligner le mot ou le texte.

Remarque : la couleur d'arrière-plan du bloc et le formatage du texte en petites lettres capitales (toutes les minuscules sont remplacées par des petites lettres capitales) indiqués dans la boîte de dialogue **Polices** ne sont pas affichés dans la fenêtre **Texte**. Ces fonctions sont visibles uniquement lorsque vous avez exporté le texte dans les applications qui prennent en charge ces fonctions.

Alignement du texte

1. Sélectionnez le texte à aligner.
2. Effectuez l'une des actions suivantes :
 - cliquez sur le bouton (par exemple ) dont vous avez besoin sur la barre **Formatage**, ou
 - cliquez dans la fenêtre **Texte** avec le bouton droit de la souris et sélectionnez la commande **Propriétés de caractère** dans le menu contextuel. La boîte de dialogue **Caractère** s'ouvre. Sélectionnez l'élément requis dans le champ Alignement.

Annuler et rétablir

Pour annuler une action :



- Vous pouvez cliquer sur le bouton **Annuler** de la barre d'outils **Standard** ou
- sélectionner la commande **Annuler** dans le menu **Edition**, ou encore
- appuyer sur **CTRL+Z**

Pour rétablir une action annulée :



- Vous pouvez cliquer sur le bouton **Rétablir** de la barre d'outils **Standard** ou
- sélectionner la commande **Rétablir** dans le menu **Edition**, ou encore
- appuyer sur **CTRL+Y**

Modification de tableaux

L'éditeur de tableau contient des outils permettant d'effectuer les actions suivantes :

- Fusionner le contenu de cellules ou de lignes
- Segmenter le contenu de cellules
- Segmenter le contenu de lignes (colonnes)
- Supprimer le contenu de cellules

Pour fusionner le contenu de cellules ou de lignes :



- Maintenez enfoncée la touche **CTRL**, sélectionnez les cellules ou les lignes à fusionner, puis la commande **Fusionner les cellules du tableau** ou **Fusionner les lignes du tableau** dans le menu **Edition**.


Pour segmenter le contenu de cellules :

- Sélectionnez la commande **Segmenter les cellules du tableau** dans le menu **Edition**.

Remarque : cette opération ne peut être appliquée qu'aux cellules précédemment fusionnées.

Pour segmenter le contenu de lignes ou de colonnes :

- Sélectionnez l'outil  ou  dans la barre d'outils de la fenêtre **Image**, puis cliquez sur la ligne (colonne) à segmenter pour ajouter un nouveau séparateur horizontal (vertical).

Conseil : vous pouvez fusionner le contenu de lignes à l'aide de l'outil  ou de la commande **Fusionner les lignes du tableau** (menu **Edition**).

Pour supprimer le contenu de cellules :

- Sélectionnez une ou plusieurs cellules à supprimer dans la fenêtre **Texte** et appuyez sur **SUPPR**.

Chapitre 8

Enregistrement dans des applications et des formats externes

Vous pouvez enregistrer le résultat de la reconnaissance dans un fichier, l'envoyer à une application externe sans sauvegarde, le copier dans le Presse-papiers ou l'envoyer par courrier électronique. Vous pouvez enregistrer toutes les pages ou seulement des pages sélectionnées.

FineReader peut exporter le résultat de la reconnaissance vers les applications suivantes :

Microsoft Word 6.0, 7.0, 97 (8.0), 2000 (9.0) et 2002 (10.0) ; Microsoft Excel 6.0, 7.0, 97 (8.0), 2000 (9.0) et 2002 (10.0) ; Corel WordPerfect 7.0, 8.0, 9.0 et 2002 (10.0) ; Lotus Word Pro 9.5, 97 et Millennium Edition ; StarWriter 4.x et 5.x ; PROMT 98 et toutes les applications acceptant le standard ODMA.

Contenu de ce chapitre :

- Informations générales sur l'enregistrement du texte reconnu
- Options d'enregistrement de texte
- Enregistrement du texte reconnu au format RTF et DOC
- Enregistrement du texte reconnu au format PDF
- Enregistrement du texte reconnu au format HTML
- Enregistrement de l'image de page

Informations générales sur l'enregistrement du texte reconnu

Vous pouvez :

- enregistrer le texte reconnu à l'aide de l'**assistant Enregistrement**,
- enregistrer les pages ouvertes ou sélectionnées dans un fichier ou les exporter vers une application externe,
- enregistrer toutes les pages du lot dans un fichier ou les exporter vers une application externe,
- enregistrer l'image de page.



Cliquez sur le bouton **4 - Enregistrer** pour envoyer les résultats de la reconnaissance à une application ou les enregistrer dans un fichier. L'apparence de l'icône dépend du mode d'enregistrement actif. La légende du bouton **Enregistrer** indique le nom de l'application sélectionnée pour l'exportation.

Pour enregistrer le texte reconnu :

1. Cliquez sur la flèche à droite du bouton **4 - Enregistrer** et sélectionnez la commande appropriée dans le menu contextuel.

Remarque : si vous souhaitez n'enregistrer que certaines pages, sélectionnez-les avant de cliquer sur le bouton **4 - Enregistrer**.

Quand l'exportation est terminée, l'icône du bouton **4 - Enregistrer change**, en fonction de l'opération effectuée : envoi du texte reconnu à une application, envoi par courrier électronique, copie dans le Presse-papiers ou enregistrement dans un fichier. L'icône du bouton **4 - Enregistrer** correspond toujours au dernier mode d'exportation utilisé. Si vous souhaitez exporter d'autres images en utilisant le même mode, il vous suffit donc de cliquer sur l'icône sans passer par le menu contextuel du bouton.

Options d'enregistrement de texte

Les options d'enregistrement sont définies dans la page de l'onglet **Formatage** du menu **Outils>Options**. Certaines options d'enregistrement peuvent être définies dans les boîtes de dialogue **Assistant**

Enregistrement et Enregistrer le texte sous.

- Modes de préservation du formatage et de la présentation du texte
- Conserver les images
- Résolution des images (enregistrement au format RTF, etc.)
- Qualité de résolution JPEG
- Polices à utiliser
- Enregistrer toutes les pages du lot ou seulement les pages sélectionnées
- Modes d'enregistrement du texte reconnu

Modes de préservation du formatage et de la présentation (enregistrement au format RTF, DOC ou HTML)

- **Conserver la présentation complète** : la mise en page du document est conservée intégralement : disposition en paragraphes, police et taille de police, colonnes, direction du texte, couleur du texte, structure des tableaux.
- **Conserver police et taille de police** : la structure des tableaux, la disposition en paragraphes, la police et la taille de police sont conservées.
- **Supprimer tout le formatage** : seules la structure des tableaux et la disposition en paragraphes sont conservées.

Remarque : vous avez la possibilité de définir certaines options avancées selon le format d'exportation choisi. Par exemple, dans le cas des formats **RTF/DOC**, vous pouvez définir la taille de page par défaut et le mode de mise en évidence des caractères incertains ; dans le cas du format **HTML**, vous pouvez définir la résolution de l'image et la page de code etc. Vous pouvez activer ces options dans la boîte de dialogue **Paramètres de format** (menu **Outils>Paramètres de format**). Cette boîte de dialogue comporte un onglet pour chaque format ; cliquez sur l'onglet approprié et activez les options souhaitées.

Conserver les images

Si vous activez cette option, les images sont enregistrées en même temps que le texte reconnu. Cette option est disponible avec les formats d'enregistrement **RTF**, **DOC** et **HTML**.

Résolution des images (enregistrement aux formats RTF/DOC, PDF et HTML)

Vous pouvez diminuer la résolution des images. Les fichiers **HTML**, par exemple, s'affichant généralement dans un navigateur, il n'est pas logique d'enregistrer des images haute résolution dans de tels fichiers. Vous pouvez donc diminuer la résolution des images (et par conséquent la taille des fichiers **HTML**) sans affecter leur qualité visuelle. Pour ce faire, entrez la valeur de résolution requise dans le champ **Réduire la résolution des images** à situer dans la page de l'onglet **Formats>RTF/DOC (PDF, HTML)**.

Remarque : si, dans le champ **Réduire la résolution des images**, vous entrez une valeur de résolution plus élevée que celle de la source, cette valeur est ignorée et les images sont enregistrées en fonction de la résolution de la source.

Qualité de résolution JPEG(enregistrement aux formats PDF et HTML)

Lorsque vous enregistrez du texte aux formats **PDF** et **HTML**, les images sont enregistrées au format **JPEG**. Ce format utilise l'algorithme dit de perte de qualité, afin de comprimer l'image ; en effet, la technologie de compression se base sur la moyenne des groupes de pixels, pour qu'une région entière soit enregistrée comme un nombre unique et non comme un nombre important composé de numéros décrivant chaque pixel. La valeur indiquée dans le champ **Qualité de résolution JPEG** permet de définir la qualité désirée pour l'image (**Outils>Formats**, puis les onglets **PDF** et **HTML**). Vous pouvez indiquer une valeur de 1 à 100 (la valeur par défaut est 50). Plus la valeur indiquée dans ce champ est élevée, plus la qualité de l'image enregistrée est élevée. Cette valeur affecte également la taille de l'image. En effet, plus cette valeur est élevée, plus la taille de l'image au format ***jpg** augmente. Pour que la taille et la qualité de l'image soient parfaitement proportionnées, tâchez d'enregistrer les images avec des valeurs de qualité **JPEG** différentes, puis de les ouvrir dans une application de visualisation d'images. La valeur de la **qualité JPEG** est définie dans la page de l'onglet **Formats>PDF (HTML)**.

Polices à utiliser (disponible si vous effectuez l'enregistrement dans les formats RTF, DOC et HTML).

Par défaut, les polices de la page de l'onglet **Formatage** sont utilisées lorsque vous enregistrez au format **RTF**, **DOC** et **HTML**. Vous pouvez spécifier les polices à utiliser. Vous pouvez changer les polices dans la fenêtre **Texte** ou définir d'autres polices dans la page de l'onglet **Formatage**, groupe **Polices**, puis relancer la lecture du document.

Enregistrer toutes les pages du lot ou seulement les pages sélectionnées

Vous pouvez enregistrer toutes les pages du lot ou n'enregistrer que celles que vous sélectionnez. Si vous souhaitez enregistrer uniquement un certain nombre de pages, sélectionnez-les avant de procéder à l'enregistrement.

Modes d'enregistrement du texte reconnu (disponible si vous enregistrez plusieurs pages de lot à la fois)

- **Créer un fichier séparé pour chaque page** : chaque page de lot est enregistrée dans un fichier séparé. Le numéro de la page de lot est ajouté automatiquement à la fin du nom de fichier.
- **Nommer les fichiers en tant qu'images source** : - utilisez cette option pour enregistrer chaque page dans des fichiers séparés, dont le nom est identique à celui des images source.

Remarque :

1. les pages ne contenant pas d'informations sur l'image source (les pages numérisées par exemple) ne sont pas enregistrées dans ce mode. Un avertissement s'affiche si de telles pages s'intercalent entre les pages enregistrées.
2. Si plusieurs pages d'un lot proches les unes des autres affichent une image identique à l'image source, ou si leurs images source portent le même nom, ces pages sont traitées comme celles d'un fichier TIFF multipages et leur texte est enregistré dans un même fichier. Si les pages, dont les images source portent un nom identique, ne sont pas proches les unes des autres, elles sont traitées comme si leurs fichiers d'images étaient différents et le texte est donc enregistré dans des fichiers différents portant le même nom, suivi d'un index : _1, _2, etc.

-
- **Créer un fichier à chaque page vide** : le lot entier est traité comme une série de groupes de pages, dont la dernière reste blanche. Les pages des différents groupes sont enregistrées dans différents fichiers, dont les noms sont formés à partir d'un nom spécifié par l'utilisateur, suivi d'un index : -1, -2, -3, etc.
 - **Créer un fichier unique pour toutes les pages** : toutes les pages (ou toutes les pages sélectionnées) du lot sont enregistrées dans un même fichier.

Enregistrement du texte reconnu au format RTF et DOC

Les modes de préservation de la présentation sont définis dans la page de l'onglet **Formatage** de la boîte de dialogue **Options** (menu **Outils>Options**).

Remarque : lorsque vous enregistrez le texte au format RTF ou DOC, les polices utilisées sont celles définies dans la page de l'onglet **Formatage** de la boîte de dialogue **Options** (menu **Outils Options**) ou celles que vous avez définies lors de la modification du texte dans la fenêtre **Texte**.

Conseil : si vous préférez modifier le texte reconnu dans Microsoft Word plutôt que dans la fenêtre **Texte** de FineReader, vous pouvez conserver la mise en évidence des caractères incertains. Sélectionnez les options **Avec la couleur de fond et/ou Avec la couleur de texte** dans la page de l'onglet **RTF/DOC**, groupe **Mettre en évidence les caractères incertains**. Dans le fichier enregistré, tous les caractères incertains seront mis en évidence dans la couleur que vous avez définie.

Enregistrement du texte reconnu au format PDF

Options de préservation de la présentation du document :

1. **Texte et images seulement** : seuls le texte reconnu et les images qu'il contient sont enregistrés.
2. **Image seulement** : seule l'image est enregistrée.
3. **Texte par-dessus l'image de page** : la totalité de l'image est enregistrée comme telle. Les zones de texte sont enregistrées comme du texte par-dessus l'image.

4. **Texte sous l'image de page** : la totalité de l'image est enregistrée comme telle. Le texte reconnu est placé au-dessous. Cette option est utile si vous exportez le texte dans des archives : la présentation est totalement préservée et la recherche en mode texte est possible si vous optez pour ce mode d'enregistrement.

Pour définir ces options :

1. Sélectionnez la commande **Paramètres de format** du menu **Outils**. La boîte de dialogue **Paramètres de format** apparaît.
2. Définissez les options appropriées dans la page de l'onglet **PDF** de la boîte de dialogue.

Remarque :

1. Une option spéciale, **Remplacer les mots incertains par des images**, est disponible dans les modes **Texte et images seulement** et **Texte par-dessus l'image de page**. Si vous activez cette option, tous les mots incertains seront remplacés par leur image. Vous pouvez l'activer dans la page de l'onglet **PDF** de la boîte de dialogue **Paramètres de format**.
2. Si vous devez modifier le texte reconnu avant l'exportation au format PDF, veillez à préserver la division en lignes (c'est-à-dire de ne pas supprimer de lignes existantes et de ne pas en rajouter) ; ceci pourrait altérer l'affichage du fichier PDF (les lignes peuvent, par exemple, se chevaucher).
3. Lorsque vous enregistrez des textes utilisant un alphabet non-latin (par exemple, cyrillique, grec, tchèque, etc.), FineReader utilise les polices fournies par la société ParaType (www.paratype.com/shop).
4. Si, lors de l'exportation au format PDF, un message indiquant que certains caractères de votre texte n'existent pas dans les polices standard s'affiche, vous devez sélectionner un mode utilisant les polices de Type 1 et les polices de Type 1 elles-mêmes. Ces polices doivent être disponibles via Adobe Type Manager ou via un installateur de police postscript (sous Windows 2000).
5. Avant d'éditer les fichiers PDF utilisant un alphabet non-latin (par exemple, cyrillique, grec, tchèque, etc.) dans Adobe Acrobat, vous devez modifier la police de la section du texte modifiée avec une police installée sur votre ordinateur.

Enregistrement du texte reconnu au format HTML

Les modes de préservation de la présentation sont définis dans la page de l'onglet **Formatage** de la boîte de dialogue **Options** (menu **Outils>Options**).

Remarque :

Lorsque vous enregistrez le texte au format HTML, les polices utilisées sont celles définies dans la page de l'onglet **Formatage** de la boîte de dialogue **Options** (menu **Outils>Options**) ou celles que vous avez définies lors de la modification du texte dans la fenêtre **Texte**.

Pour conserver les images dans un fichier HTML :

- Activez l'option **Conserver les images** dans la page de l'onglet **Formatage** de la boîte de dialogue **Options** (menu **Outils>Options**).

Remarque : les images sont enregistrées dans des fichiers JPEG (*.jpg) distincts. La résolution et la qualité des images sont définies dans la page de l'onglet **HTML** de la boîte de dialogue **Formats(Outils>Formats)**.

Formats HTML disponibles

1. **Complet (utilise CSS ; nécessite Internet Explorer 4.0 ou supérieur)** : si vous activez cette option, le nouveau format HTML 4 est utilisé. Il prend en charge n'importe quel type

de préservation de présentation (le type de préservation effectivement utilisé dépend des options activées dans la page de l'onglet **Formatage**, groupe **Conserver la présentation**) ; la feuille de style intégrée est utilisée.

2. **Simple (compatible avec tous les navigateurs Internet)** : si vous activez cette option, le format HTML 3 est utilisé. La présentation du document est préservée approximativement : le retrait de la ligne n'est pas conservé et la taille de la police est enregistrée approximativement (le format HTML 3 ne prend en charge qu'un nombre limité de tailles de police ; de ce fait, FineReader choisit la taille de police du format HTML 3 qui correspond à la taille de police effective du texte). Ce format HTML est pris en charge par tous les navigateurs (Netscape Navigator, Internet Explorer 3.0 et supérieur).
3. **Automatique (enregistre deux formats dans un fichier ; sélection automatique selon le type de navigateur)** : les deux formats (**Simple** et **Complet**) sont enregistrés dans le même fichier. Votre navigateur sélectionnera le format pris en charge.

Pour définir le format HTML souhaité :

- Cliquez sur le bouton radio approprié dans la page de l'onglet **HTML** de la boîte de dialogue **Paramètres de format** (menu **Outils>Formats**), groupe **Formats**.

Remarque : l'application détecte automatiquement la page de code. Pour changer de page de code, faites votre choix dans le champ **Page de code** de la page de l'onglet **HTML** de la boîte de dialogue **Paramètres de format**.

Enregistrement de l'image de page

1. Sélectionnez une page de lot.
2. Sélectionnez la commande **Enregistrer l'image sous** dans le menu **Fichier**. La boîte de dialogue **Enregistrer sous** apparaît.
3. Dans le disque et le dossier où vous souhaitez enregistrer le fichier, ainsi que le format du fichier.

Remarque : si vous le souhaitez, vous pouvez n'enregistrer que certaines zones d'image encadrées par des blocs (de tout type). Pour cela, sélectionnez le ou les blocs appropriés, puis, dans la boîte de dialogue **Enregistrer sous**, cliquez sur la case de sélection **Enregistrer seulement les blocs sélectionnés**. Ce mode ne peut être utilisé que lors de l'enregistrement d'une seule image.

4. Spécifiez le nom du fichier.
5. Cliquez sur **OK**.

Remarque : vous pouvez enregistrer plusieurs images dans un même fichier (TIFF multipages). Pour ce faire :

1. Sélectionnez les images souhaitées dans la fenêtre **Lot**.
 2. Sélectionnez la commande **Enregistrer sous** dans le menu **Fichier**. Sélectionnez le format TIFF et l'option **Enregistrer comme fichier d'image multipages**.
-

Remarque : si vous enregistrez des images de plusieurs pages sélectionnées dans la fenêtre **Lot** comme fichiers distincts (vous n'utilisez pas le mode d'enregistrement d'images dans un fichier TIFF multipages), les noms des fichiers ont la syntaxe suivante : le préfixe est le nom du fichier que vous saisissez, le suffixe est le nombre à quatre chiffres de la page.

Chapitre 9

Travailler avec les lots

Le lot est le principal référentiel de données FineReader : les images numérisées, le texte reconnu et les autres données sont conservés dans un lot. La majorité des options de FineReader sont définies dans le cadre d'un lot : numérisation, reconnaissance, options d'enregistrement, etc. Les gabarits utilisateur, les langues utilisateur et les groupes de langues « appartiennent » également aux lots. Quand vous créez un nouveau lot, vous pouvez utiliser les paramètres par défaut, les paramètres du lot en cours ou les paramètres enregistrés dans un fichier *.fbi.

Contenu du chapitre :

- Informations générales sur l'utilisation des lots
- Création d'un nouveau lot
- Ouverture d'un lot
- Ajout d'images à un lot
- Numéro de page de lot
- Fermeture d'un lot
- Suppression d'un lot
- Recherche en mode texte intégral dans les pages de lot reconnues

Informations générales concernant l'utilisation des lots

Lorsque l'application FineReader est utilisée pour la première fois, elle ouvre le lot qui se trouve dans le dossier FineReader. Vous pouvez travailler avec ce lot ou en créer un nouveau. Un lot peut contenir jusqu'à 9 999 pages.

Conseil : il peut être utile d'enregistrer les « mêmes » pages dans un même lot. Par exemple, si vous conservez dans un même lot toutes les pages d'un livre, les textes d'une même langue ou les images possédant la même présentation, votre travail s'en trouvera considérablement simplifié.

La fenêtre **Lot** affiche la liste des pages du lot ouvert. Pour afficher une page, cliquez sur son icône ou cliquez deux fois sur son numéro. Tous les fichiers associés à cette page de lot s'ouvriront dans leur fenêtre respective, c'est-à-dire le fichier texte (si la page a été reconnue) dans la fenêtre **Texte** et le fichier d'image dans la fenêtre **Image**.

Deux modes de visualisation sont disponibles pour afficher les pages dans la fenêtre Lot :

| Affichage de lot | Description |
|------------------|--|
| Vignettes | Dans cette vue, les pages du lot s'affichent sous forme de vignettes. Une vignette est simplement une image très de la page d'origine. Des icônes supplémentaires apparaissent sur les vignettes à mesure que vous traitez l'image, pour vous indiquer les actions particulières effectuées : reconnaissance, enregistrement, etc. Cette vue peut être utile pour ouvrir une page précise d'un lot : vous voyez une petite image de la page dans la fenêtre de lot et pouvez ainsi déterminer facilement celle dont vous avez besoin. Pour ouvrir une image, cliquez simplement sur sa vignette. |
| Détails | Dans cette vue, la fenêtre de lot affiche des informations détaillées concernant chaque page du lot ; vous pouvez trier la liste des pages selon toute caractéristique appropriée. La vue Détails est pratique lorsque vous manipulez des lots de grande taille, car la fenêtre de lot affiche un nombre beaucoup plus grand de pages dans cette vue que dans la vue Vignettes . Cliquez deux fois sur une page pour l'ouvrir. |

Pour choisir le type de vue de la page dans la fenêtre Lot :

- Cliquez dans la fenêtre **Lot** avec le bouton droit de la souris et sélectionnez la commande **Afficher>...** dans le menu contextuel.

Vous pouvez également régler chaque vue de la fenêtre **Lot**, c'est-à-dire choisir les caractéristiques affichées, la façon de trier les pages, etc. Pour cela :

- cliquez dans la fenêtre **Lot** avec le bouton droit de la souris et sélectionnez la commande **Affichage de lot> Personnaliser** dans le menu contextuel. Une boîte de dialogue s'ouvre. Définissez les options de votre choix dans les pages des onglets **Vignettes** et **Détails** de cette boîte de dialogue.

Vous pouvez sélectionner plusieurs pages ou une série de pages :

- **Pour sélectionner plusieurs pages consécutives**, tenez enfoncée la touche **MAJ** et cliquez sur la première et la dernière page du groupe que vous souhaitez sélectionner.
- **Pour sélectionner plusieurs pages**, tenez enfoncée la touche **CTRL** et cliquez sur les pages dont vous avez besoin.
- **Pour sélectionner toutes les pages du lot**, ouvrez la fenêtre **Lot** et choisissez la commande **Tout sélectionner** dans le menu **Edition** ou appuyez sur **CTRL+A**.

Création d'un nouveau lot

Pour créer un nouveau lot :

1. Sélectionnez la commande **Nouveau lot** dans le menu **Fichier**. La boîte de dialogue **Création de nouveau lot** apparaît.
2. Sélectionnez ou créez un dossier pour le nouveau lot dans la boîte de dialogue **Création de nouveau lot**.
3. Sélectionnez le champ **Modèle de lot** et choisissez l'une des valeurs suivantes, selon les paramètres que vous voulez appliquer au nouveau lot : **Paramètres par défaut** - pour appliquer les paramètres par défaut, **Lot en cours** - pour appliquer les paramètres du lot en cours ou **Modèles de lot** (.fnt) - pour appliquer des paramètres enregistrés antérieurement dans un fichier spécial.


Remarque : pour enregistrer des paramètres de lot dans un fichier, cliquez sur le bouton **Enregistrer** de la page de l'onglet **Général** (menu **Outils>Options**). La boîte de dialogue **Enregistrer le modèle de lot sous** apparaît. Spécifiez le nom de fichier. Les options suivantes sont enregistrées : les options définies dans les pages des onglets **Reconnaissance**, **Numériser/Ouvrir une image**, **Formatage** et **Vérification orthographique**, les options définies dans les autres pages d'onglet **Paramètres de format**, chemin vers le dossier contenant les langues utilisateur et leurs dictionnaires. Les langues utilisateur, les groupes de langues et les gabarits utilisateur sont également enregistrés dans ce fichier. Si vous souhaitez rétablir les paramètres par défaut, cliquez simplement sur le bouton **Utiliser paramètres par défaut** dans la page de l'onglet **Général**. Pour charger des paramètres, cliquez sur le bouton **Charger** dans la page de l'onglet **Général** et sélectionnez un fichier de modèle de lot FineReader (*.fnt) contenant les options appropriées.

Ouverture d'un lot

1. Sélectionnez la commande **Ouvrir un lot** dans le menu **Fichier**. La boîte de dialogue **Ouverture de lot** apparaît.
2. Sélectionnez le dossier approprié dans la boîte de dialogue.
Lorsque vous ouvrez un lot, celui qui était ouvert est automatiquement enregistré et fermé.

Au démarrage, FineReader ouvre automatiquement le dernier lot avec lequel vous avez travaillé.

Remarque : vous pouvez ouvrir le lot directement dans Windows Explorer :

- cliquez avec le bouton droit de la souris sur le dossier du lot (représenté par l'icône ) et sélectionnez la commande **Ouvrir avec FineReader** dans le menu contextuel. Une copie de FineReader sera lancée et ouvrira le lot que vous avez sélectionné.

Ajout d'images à un lot

- Sélectionnez la commande **Ouvrir une image** dans le menu **Fichier** ou appuyez sur **CTRL+O**.
- Sélectionnez les images souhaitées dans la boîte de dialogue **Ouvrir une image**. FineReader ajoute l'image au lot ouvert et la recopie dans le dossier de lot.

Remarque : vous pouvez ajouter des images directement depuis Windows Explorer :

1. Dans Windows Explorer, sélectionnez un fichier d'image ou un groupe de fichiers.
2. Cliquez sur la sélection à l'aide du bouton droit de la souris, puis dans le menu contextuel, sélectionnez la commande **Ouvrir avec FineReader**. Si l'application FineReader est déjà lancée, la ou les images sélectionnées viennent s'ajouter au lot en cours. Dans le cas contraire, le dernier lot sur lequel vous avez travaillé s'ouvre dans une copie de FineReader. Cette commande du menu contextuel n'est activée que pour les fichiers au format pris en charge par FineReader 6.0.

Numéro de page de lot

Toutes les pages de lot sont numérotées. Un lot peut contenir jusqu'à 9 999 pages. Le numéro de page est affiché dans le lot. Vous pouvez renuméroter les pages directement dans la fenêtre **Lot** ou par l'intermédiaire de la boîte de dialogue **Renuméroter les pages**.

Pour renuméroter les pages dans la fenêtre Lot :

1. cliquez sur une page de la fenêtre **Lot** ou appuyez sur **F2**.
2. Saisissez ensuite le nouveau numéro de la page.

Une fois le numéro de la page modifié, celle-ci est placée de manière à correspondre au nouveau numéro de la fenêtre **Lot**.

Remarque : cliquer deux fois sur un numéro de page permet d'ouvrir la page correspondante.

Pour renuméroter des pages dans la boîte de dialogue **Renuméroter les pages** :

1. Sélectionnez une ou plusieurs pages.
2. Sélectionnez la commande **Renuméroter les pages** dans le menu **Lot**.
3. Spécifiez le numéro de la première page sélectionnée (celle de plus petit numéro).

Remarque

1. Si vous voulez renuméroter toutes les pages du lot, sélectionnez l'option **Toutes les pages** dans la boîte de dialogue **Renuméroter les pages**.
2. Si vous ne voulez renuméroter qu'une partie du lot :
 - Sélectionnez les pages à renuméroter dans la fenêtre **Lot**.
 - Sélectionnez l'option **Pages sélectionnées** dans la boîte de dialogue **Renuméroter les pages**.
3. Si vous voulez que les pages sélectionnées soient renumérotées séquentiellement, sélectionnez l'option **Numérotation des pages en continu**. Par exemple : les numéros de page étaient 2,5,6. « 1 » a été défini comme premier numéro. Après renumérotation, les numéros de page seront 1,2,3. Si l'option Numérotation des pages en continu n'avait pas été spécifiée, les numéros de page seraient 1,4,5. Ainsi, la première page se voit attribuer le numéro que vous avez saisi et le reste des pages sera renuméroté de manière à conserver les écarts de numéros initiaux.

Remarque : si vous ne renumérotez qu'une partie des pages du lot, une page renumérotée peut se voir attribuer un numéro déjà occupé par une page non concernée par la renumérotation. Le cas échéant, FineReader vous avertira et l'ensemble de l'opération sera annulé.

Fermeture d'une page de lot ou d'un lot entier

Pour fermer une page de lot :

- Sélectionnez la commande **Fermer la page en cours** dans le menu **Lot**.

Pour fermer une page de lot :

- Sélectionnez la commande **Fermer un lot** dans le menu **Fichier**.

Remarque : le lot sera enregistré automatiquement lors de sa fermeture.

Suppression d'un lot

Remarque : la suppression d'un lot implique la suppression de l'intégralité de son contenu, c'est-à-dire toutes ses pages (images et texte), ainsi que tout fichier supplémentaire créé en travaillant avec le lot : gabarits et langues utilisateur, etc. En pratique, vous supprimez tout le contenu du dossier de lot.

Pour supprimer un lot, sélectionnez la commande **Supprimer** le lot dans le menu **Lot**.

Pour supprimer une page de lot :

1. Sélectionnez les pages à supprimer dans la fenêtre **Lot**.
2. Sélectionnez la commande **Supprimer la page** dans le menu **Lot** ou appuyez simplement sur **SUPPR**.

Recherche en mode texte intégral dans les pages du lot reconnues

(prise en charge dans la version FineReader Corporate seulement)

Vous pouvez rechercher des mots sous leurs diverses formes grammaticales dans l'ensemble des pages reconnues. L'élément recherché peut être un mot ou un groupe de mots. Les mots peuvent se présenter sous diverses formes (pour les langues prises en charge par le dictionnaire) ; les mots recherchés peuvent se trouver à n'importe quel endroit du texte et ne pas se suivre.

Pour ce faire :

1. dans le menu **Edition**, sélectionnez la commande **Recherche avancée** ou appuyez sur les touches **ALT+F3**.
2. La fenêtre **Rechercher** s'ouvre en bas de la fenêtre **Zoom**.
3. Dans le champ **Chercher**, tapez le texte à rechercher. Vous pouvez également y coller le contenu du Presse-papiers ou sélectionner une ancienne recherche dans la liste.
4. Cliquez sur le bouton **Chercher**.
Le numéro des pages du lot dans lesquelles figurent TOUS les mots saisis dans le champ **Chercher** s'affiche dans la fenêtre **Résultats de la recherche**. Pour toutes les pages du lot ainsi trouvées, la fenêtre affiche les données dernièrement modifiées, ainsi que la première section de la page contenant les termes recherchés (mis en évidence). Cliquez sur le numéro de la page pour l'ouvrir dans les fenêtres **Image**, **Texte** et **Zoom** ; dans chacune de celles-ci, tous les mots trouvés sont mis en évidence par la couleur du fond.

Remarque : la fonction de recherche avancée ne prend pas en charge les caractères spéciaux (comme les pieds de mouches et les marques de paragraphe).

Chapitre 10

Traitement des documents en réseau

La version Corporate Edition de ABBYY FineReader est spécialement conçue pour la gestion des documents de réseau. Chaque ordinateur impliqué dans la gestion de réseau doit être équipé d'une copie indépendante de FineReader (pour en savoir plus sur l'installation de FineReader en réseau, reportez-vous à la section Installation sur un serveur et une station de travail réseau).

La version Corporate Edition d'ABBYY FineReader vous permet de :

1. Travailler sur un même lot en réseau

Vous améliorerez ainsi les performances de la modification multi-utilisateurs de documents. Vous pouvez également suivre le processus de modification de documents dans son ensemble (identifiant de la personne ayant ouvert la page du lot et nom de son ordinateur, utilisateur ayant numérisé, reconnaissant ou ayant vérifié la page, etc.) Enfin, tout le groupe travaillant sur le même lot peut accéder aux modifications apportées par un utilisateur à une page.

2. Travail de groupe à l'aide des mêmes langues et dictionnaires d'utilisateur

Avec ABBYY FineReader Corporate Edition, plusieurs utilisateurs peuvent travailler avec les mêmes langues et les mêmes dictionnaires. Ils peuvent élargir leurs dictionnaires utilisateurs simultanément, c'est-à-dire pendant la vérification orthographique.

3. Travailler en groupe avec des dictionnaires personnalisés pour les langues supportant un dictionnaire

ABBYY FineReader propose des dictionnaires intégrés pour les langues supportant un dictionnaire. Ces dictionnaires contiennent les mots les plus utilisés. Ils peuvent toutefois ne pas inclure les noms propres, les termes techniques spécifiques, les acronymes, etc. Afin d'améliorer la qualité de la reconnaissance (FineReader vérifie les mots qu'il reconnaît dans le dictionnaire) et d'accélérer la vérification orthographique, vous pouvez ajouter des mots de ce type dans un dictionnaire personnalisé. Avec ABBYY FineReader Corporate Edition, plusieurs utilisateurs peuvent travailler avec le même dictionnaire personnalisé.

Contenu du chapitre :

- Travailler en réseau sur un même lot
- Travail de groupe à l'aide des mêmes langues et dictionnaires utilisateurs
- Travailler en groupe avec des dictionnaires personnalisés pour les langues supportant un dictionnaire

Travailler en réseau sur un même lot

(disponible uniquement dans FineReader version Corporate Edition)

1. Créez/Ouvrez un lot et réglez les options de numérisation et de reconnaissance requises. Lancez FineReader et ouvrez le lot à traiter sur tous les ordinateurs concernés.
2. Lancez la reconnaissance en arrière-plan (**Traitement>Lancer la reconnaissance en arrière-plan**) sur tous les postes concernés.
3. Lancez la numérisation sur le poste de travail équipé d'un scanner à chargeur automatique.

Conseil : si votre scanner à haut débit ne prend pas en charge le format TWAIN, vous pouvez numériser directement les pages dans le dossier de lot de FineReader. Pour ce faire, numérisez les images à l'aide de n'importe quelle application de numérisation fournie avec votre scanner haut débit sur l'ordinateur lié au scanner. Sélectionnez ensuite le dossier de lot FineReader pour l'enregistrement de ces images. Il vous suffit alors de nommer les images numérisées en suivant cette règle : les fichiers doivent porter les noms 0001.tif, 0002.tif, 0003.tif, etc.(selon l'ordre de la numérisation).

4. FineReader détectera et traitera automatiquement les images que vous numérisez.
5. Vous pouvez alors modifier le texte reconnu et l'enregistrer dans un fichier ou l'exporter vers une application de votre choix.

Vous pouvez afficher l'état de chaque page (si la page a été numérisée, reconnue, modifiée ou enregistrée, quel utilisateur du réseau effectue la modification et la numérisation, etc.) dans la fenêtre Lot. Ces informations sont affichées dans les colonnes correspondantes dans l'affichage des Détails des pages du lot. Pour régler l'affichage des Détails des pages :

- cliquez avec le bouton droit de la souris dans la fenêtre **Lot** et sélectionnez **Affichage>Détails** dans le menu **contextuel**.

Vous pouvez personnaliser l'affichage des détails des pages : spécifiez les colonnes qui doivent être affichées dans la fenêtre Lot ou sélectionnez la colonne selon laquelle les pages seront triées. Pour ce faire :

- cliquez avec le bouton droit de la souris dans la fenêtre **Lot** et sélectionnez **Affichage>Personnaliser**. Réglez les options nécessaires dans la page de l'onglet **Détails** de la boîte de dialogue **Paramètres d'affichage du lot**.

Si plusieurs postes traitent un lot, FineReader leur répartit automatiquement le travail. Chaque nouvelle page scannée est « récupérée » par le premier poste de travail disponible sur lequel la reconnaissance en arrière-plan s'effectue. Cette page est alors verrouillée et inaccessible pour tous les autres ordinateurs. Vous pouvez actualiser la liste des pages du lot en appuyant sur **F5** ou en sélectionnant la commande **Actualiser la liste de pages** dans le menu **Lot**. Simultanément, d'autres postes de travail (ou le même) peuvent ouvrir des pages reconnues pour les vérifier, les modifier et les enregistrer. Tout le groupe travaillant sur le même lot peut accéder aux modifications apportées par un utilisateur.

Remarque : si votre lot contient de nombreuses pages, vous remarquerez que la vitesse de reconnaissance est accrue sur les systèmes multiprocesseurs en mode Arrière-plan.

Travail de groupe à l'aide des mêmes langues et dictionnaires utilisateur

(disponible uniquement dans FineReader version Corporate Edition)

Créez un lot et réglez les options de numérisation et de reconnaissance requises.

Tous les dictionnaires et langues utilisateur que vous reliez sont stockés dans un dossier. Il s'agit du dossier de lot par défaut. Avant de créer une langue utilisateur utilisant un dictionnaire utilisateur, vous devez spécifier le dossier dans lequel stocker les langues et dictionnaires utilisateur. Pour spécifier le dossier :

- Cliquez sur le bouton **Modifier** dans la boîte de dialogue **Editeur de langue (Outils> Editeur de langue)**, puis sélectionnez le dossier dans la boîte de dialogue qui s'affiche. Tous les dictionnaires et langues utilisateur que vous reliez sont stockés dans ce dossier.

Une fois l'installation terminée, enregistrez les paramètres du lot dans un fichier de modèle de lot (*.fbt) :

- Sélectionnez **Options>Reconnaissance**, puis cliquez sur le bouton **Enregistrer (Outils> Options)**. Dans la boîte de dialogue **Enregistrer le modèle de lot sous**, sélectionnez le dossier, puis saisissez le nom du fichier.

Pour que plusieurs utilisateurs travaillent avec les langues utilisateur et leurs dictionnaires connexes, ils doivent tous charger les paramètres de lot depuis le fichier .fbt enregistré précédemment lors href="\"../BatchWork/Create.htm\">de la création d'un lot.

Pour ce faire :

Sélectionnez l'élément **Modèle de lot (.fbt)** dans le champ **Modèle**. Dans la boîte de dialogue **Ouvrir un modèle de lot**, sélectionnez le fichier .fbt requis. Les paramètres de lot enregistrés précédemment, ainsi que le chemin d'accès aux langues utilisateur et à leurs dictionnaires sont définis. Par conséquent, tous les utilisateurs auront le même chemin d'accès aux langues utilisateur et à leurs dictionnaires.

L'emploi de ces langues utilisateur pour la reconnaissance et la vérification orthographique vous permet de modifier leurs dictionnaires. Les modifications apportées par un utilisateur sont disponibles à l'ensemble des utilisateurs du dossier. Les langues utilisateur créées dans ce dossier sont disponibles pour tous les utilisateurs ayant chargé ce modèle de lot. La liste des langues utilisateur disponibles, ainsi que leurs propriétés s'affichent dans la boîte de dialogue **Editeur de langue** du groupe **Langues définies par l'utilisateur**.

Lorsqu'un utilisateur ajoute ou supprime un mot du dictionnaire, celui-ci se verrouille. Le dictionnaire est mis à jour lorsqu'un utilisateur clique sur le bouton **Ajouter** de la boîte de dialogue **Vérification orthographique**, ou sur tout autre bouton de la boîte de dialogue **Afficher les dictionnaires**.

Remarque :

1. Si vous utilisez un dossier dans lequel plusieurs dictionnaires de différents utilisateurs sont stockés, ces derniers doivent avoir accès en lecture et en écriture à ce dossier.
2. Lorsqu'une langue utilisateur est utilisée par plusieurs utilisateurs, elle n'est accessible qu'en lecture seulement ; vous ne pouvez donc pas modifier les paramètres d'une langue utilisateur déjà créée. Vous pouvez, cependant, ajouter ou supprimer des entrées dans le dictionnaire utilisateur de cette langue.

Travail de groupe à l'aide de dictionnaires personnalisés, pour les langues prenant en charge les dictionnaires

(disponible uniquement dans FineReader version Corporate Edition)

Créez un lot et réglez les options de numérisation et de reconnaissance requises. Les dictionnaires personnalisés pour les langues principales prédéfinies, (langues pour lesquelles un dictionnaire est disponible), sont enregistrés par défaut dans le dossier où l'application a été installée (pour Windows 2000 : Documents et paramètres\[profil de l'utilisateur]\Données de l'application\ABBYY\FineReader\6.00\UserDictionaries).

Pour permettre à plusieurs utilisateurs d'utiliser les mêmes dictionnaires personnalisés pour des langues prédéfinies, spécifiez un dossier public où l'ensemble des dictionnaires de ce type sont enregistrés. Il peut s'agir d'un dossier local ou réseau. Pour spécifier un dossier :

- Cliquez sur le bouton **Parcourir**, situé dans la page de l'onglet **Vérification orthographique** de la boîte de dialogue **Options** (menu **Outils>Options**). Sélectionnez le dossier dans lequel stocker les dictionnaires utilisateur pour les langues prédéfinies.

Les utilisateurs peuvent ajouter à ces dictionnaires personnalisés. Lorsqu'un utilisateur ajoute ou supprime un mot du dictionnaire, celui-ci se verrouille. Les modifications apportées par un utilisateur sont disponibles à l'ensemble des utilisateurs du dossier. Le dictionnaire est mis à jour lorsqu'un utilisateur clique sur le bouton **Ajouter** de la boîte de dialogue **Vérification orthographique**, ou sur tout autre bouton de la boîte de dialogue **Sélection de dictionnaire**.

Remarque : si plusieurs utilisateurs utilisent le dossier dans lequel les dictionnaires personnalisés sont stockés, tous ces utilisateurs doivent avoir accès en lecture et en écriture au dossier.

Annexe

Raccourcis clavier

Menu Fichier

| Pour : | Appuyez sur : |
|---------------------------------------|---------------|
| Ouvrir une image depuis un fichier | CTRL+O |
| Numériser une image | CTRL+K |
| Numériser des images multiples | CTRL+MAJ+K |
| Arrêter la numérisation | CTRL+T |
| Créer un nouveau lot | CTRL+N |
| Ouvrir un lot | CTRL+P |
| Enregistrer un texte dans un fichier | CTRL+F2 |
| Enregistrer une image dans un fichier | F12 |

Menu Edition

| Pour : | Appuyez sur : |
|--|----------------------|
| Annuler la dernière action | CTRL+Z |
| Rétablir la dernière action annulée | CTRL+Y |
| Couper la sélection vers le Presse-papiers | CTRL+X |
| Copier la sélection vers le Presse-papiers | CTRL+INSER ou CTRL+C |
| Coller le contenu du Presse-papiers | CTRL+V ou MAJ+INSER |
| Supprimer le bloc actif, la sélection ou les pages sélectionnées | SUPPR |
| Sélectionner tout le texte de la fenêtre Texte, toutes les pages du lot ou tous les blocs de l'image ouverte | CTRL+A |
| Chercher le texte spécifié | CTRL+F |
| Chercher l'occurrence suivante du texte recherché | F3 |
| Chercher et remplacer le texte spécifié | CTRL+H |

Menu Affichage

| Pour : | Appuyez sur : |
|--|---------------|
| Agrandir l'image dans la fenêtre Image | CTRL+MAJ+NUM+ |
| Réduit l'image dans la fenêtre Image | CTRL+MAJ+NUM- |
| Zoom avant sur des blocs sélectionnés | CTRL+MAJ+NUM* |
| Propriétés | ALT+ENTREE |

Menu Lot

| Pour : | Appuyez sur : |
|---|----------------|
| Ouvrir la page suivante du lot | ALT+BAS |
| Ouvrir la page précédente du lot | ALT+HAUT |
| Ouvrir une page de numéro spécifié | CTRL+G |
| Fermer la page en cours | CTRL+4 |
| Supprimer le texte reconnu dans la fenêtre Texte | CTRL+MAJ+SUPPR |
| Supprimer tous les blocs dans la fenêtre Image et tout le texte reconnu dans la fenêtre Texte | CTRL+SUPPR |
| Actualiser la liste de pages | F5 |

Menu Traitement

| Pour : | Appuyez sur : |
|---|---------------|
| Numériser et lire une image | CTRL+D |
| Ouvrir et lire une image | CTRL+MAJ+D |
| Lancer l'assistant Numériser&Lire | CTRL+W |
| Analyser la présentation | Ctrl+E |
| Analyser la présentation de toutes les pages du lot | CTRL+MAJ+E |
| Lire les pages actives ou sélectionnées | CTRL+R |
| Lire toutes les pages du lot | CTRL+MAJ+R |
| Lire les blocs actifs ou sélectionnés | CTRL+MAJ+B |

Menu Outils

| Pour : | Appuyez sur : |
|---|---------------|
| Vérifier l'orthographe du texte reconnu | F7 |
| Aller sur le mot incertain de l'erreur précédente | F4 |
| Aller sur le mot incertain de l'erreur suivante | MAJ+F4 |
| Afficher les dictionnaires | CTRL+MAJ+V |
| Traduire un terme avec Lingvo | CTRL+MAJ+T |
| Ouvrir l'Editeur de langue pour créer ou modifier des langues et des groupes de langues | CTRL+MAJ+L |
| Ouvrir l'Editeur de gabarit pour créer ou modifier des langues et des groupes de langues | CTRL+MAJ+A |
| Définir les paramètres de numérisation | CTRL+MAJ+S |
| Ouvrir la boîte de dialogue des paramètres de format pour définir les options d'enregistrement des formats pris en charge en sortie | CTRL+MAJ+X |
| Ouvrir la boîte de dialogue Options | CTRL+MAJ+O |

Menu Fenêtre

| Pour : | Appuyez sur : |
|---------------------------------------|----------------------|
| Ouvrir la fenêtre suivante | CTRL+F6 |
| Ouvrir la fenêtre précédente | CTRL+MAJ+F6 |
| Ouvrir la fenêtre Lot | ALT+1 |
| Ouvrir la fenêtre Image | ALT+2 |
| Ouvrir la fenêtre Texte | ALT+3 |
| Ouvrir la fenêtre Zoom | ALT+4 |
| Passer à la fenêtre Recherche avancée | ALT+5 |
| Ouvrir la fenêtre Recherche avancée | ALT+F3 |

Menu Aide

| Pour : | Appuyez sur : |
|---------------|----------------------|
| Ouvrir l'aide | F1 |

Raccourcis généraux

| Pour : | Appuyez sur : |
|---|--|
| Mettre la sélection en caractères gras | CTRL+B |
| Mettre la sélection en caractères italiques | CTRL+I |
| Mettre la sélection en caractères soulignés | CTRL+U |
| Passer à la cellule de tableau voisine | flèche vers la gauche, la droite, le haut et le bas |