

**Système de reconnaissance optique de caractères**

# **ABBYY FineReader**

**Version 7.0**

**Guide de l'utilisateur**

Les informations du présent document sont susceptibles de modifications 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-divulgaration.

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.

© 2003 ABBYY Software Ltd. 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 Moscow, P.O. 72. ABBYY

# Table des matières

<b>Bienvenue</b> .....	7
Chapitre 1	
<b>Installation et démarrage d'ABBYY FineReader</b> .....	9
Configuration matérielle et logicielle requise .....	10
Installation de ABBYY FineReader .....	10
Installation sur un serveur et une station de travail réseau .....	11
Démarrage de ABBYY FineReader .....	12
A propos de l'activation d'ABBYY FineReader .....	12
Chapitre 2	
<b>Démarrage rapide</b> .....	15
Comment saisir instantanément un document .....	16
Fenêtre principale de ABBYY FineReader .....	16
Barres d'outils de ABBYY FineReader .....	18
Chapitre 3	
<b>Caractéristiques générales d'ABBYY FineReader</b> .....	25
A quoi sert un système OCR ? .....	26
Nouvelles caractéristiques d'ABBYY FineReader 7.0 .....	27
Formats d'enregistrement de documents pris en charge par FineReader .....	30
Support de formats d'image .....	30
Chapitre 4	
<b>Acquisition de l'image</b> .....	33
Numérisation .....	34
Définition des paramètres de numérisation .....	35
Conseils pour le réglage de la luminosité .....	36
Numérisation de documents multipages .....	37
Ouverture d'images .....	39
Acquisition d'images depuis le répertoire de veille .....	39
Numérisation de pages doubles .....	40
Ajout d'images de cartes de visite au lot .....	40
Numérotation des pages .....	41
Travailler avec les images .....	42
Options d'image de lot .....	44

## Chapitre 5

<b>Analyse de la présentation de la page</b> .....	45
Informations générales sur l'analyse de la présentation .....	46
Types de blocs .....	46
Options d'analyse automatique de la présentation .....	47
Tracer et modifier des blocs manuellement .....	49
Analyse manuelle de la présentation d'un tableau .....	52
Utilisation des modèles de bloc .....	53

## Chapitre 6

<b>Reconnaissance</b> .....	55
Informations générales sur la reconnaissance .....	56
Langue de reconnaissance .....	56
Type d'impression du texte source .....	58
Autres options de reconnaissance .....	59
Reconnaissance en arrière-plan .....	59
Reconnaissance avec apprentissage .....	60
Apprentissage d'un gabarit utilisateur .....	62
Modification d'un gabarit utilisateur .....	64
Langues utilisateur et groupes de langues .....	64
Création d'une nouvelle langue .....	65
Création d'un nouveau groupe de langues .....	66

## Chapitre 7

<b>Vérification et modification de texte</b> .....	69
Vérification du texte dans FineReader .....	70
Options de vérification et de modification de texte .....	72
Ajout et suppression de mots dans le dictionnaire utilisateur .....	73
Modification de texte dans FineReader .....	74
Modification de tableaux .....	78

## Chapitre 8

<b>Enregistrement dans des applications et des formats externes</b> .....	81
Informations générales sur l'enregistrement du texte reconnu .....	82
Options d'enregistrement de texte .....	82
Enregistrement du texte reconnu au format RTF, DOC et Word XML .....	85
Enregistrement du texte reconnu au format PDF .....	86
Enregistrement du texte reconnu au format HTML .....	87
Enregistrement de texte reconnu au format PowerPoint .....	88
Enregistrement de l'image de page .....	88

## Chapitre 9

<b>Traitement des documents en réseau</b> .....	91
Travailler sur un même lot en réseau .....	93
Travail de groupe à l'aide des mêmes langues et dictionnaires utilisateur .....	94
Travail de groupe à l'aide de dictionnaires personnalisés, pour les langues prenant en charge les dictionnaires .....	95

## Chapitre 10

<b>Travailler avec les lots</b> .....	97
Informations générales concernant l'utilisation des lots .....	98
Création d'un nouveau lot .....	99
Ouverture d'un lot .....	99
Ajout d'images à un lot .....	99
Numéro de page de lot .....	100
Enregistrement d'un lot .....	101
Fermeture d'une page de lot ou d'un lot entier .....	101
Suppression d'un lot .....	101
Paramètres du lot .....	102
Recherche en mode texte intégral dans les pages du lot reconnu .....	102

## Annexe

<b>Raccourcis clavier et Glossaire</b> .....	103
Raccourcis clavier .....	104
Glossaire .....	106



# BIENVENUE

Merci d'avoir choisi ABBYY FineReader.

Nous avons tous besoin, de façon plus ou moins occasionnelle, de traiter sur ordinateur des documents texte, qu'il s'agisse d'articles de journaux ou de magazines, de contrats, de lettres commerciales, de fax, de listes de prix ou de questionnaires. Pendant des années, la seule façon d'y parvenir était de saisir à l'aide du clavier le contenu de ces documents papier. Rappelez-vous les longues heures que vous avez passées à saisir à l'ordinateur le texte d'un document imprimé... Ne serait-il pas formidable qu'un ordinateur puisse lire lui-même le texte d'une feuille de papier ?

Certains rêves peuvent devenir réalité ! Et bien ceci est désormais possible grâce à FineReader, un programme qui permet à votre ordinateur et à votre scanner d'effectuer eux-mêmes une opération unique : lire du texte imprimé.

ABBYY FineReader est un système OCR (Reconnaissance Optique des Caractères) qui convertit les documents imprimés et PDF en documents au format modifiable : Word, Excel, HTML, PowerPoint sont quelques exemples de formats pris en charge par le logiciel. Aucune modification n'est effectuée dans la disposition du document d'origine. Des millions de personnes utilisent déjà des systèmes OCR, ce qui leur évite d'avoir à saisir manuellement le contenu des documents papier. Ces systèmes sont de plus en plus présents dans les entreprises partout dans le monde.

ABBYY FineReader est très facile à utiliser : numériser et reconnaître un document n'est plus qu'une question de minutes. Il fournit également plusieurs paramètres et options supplémentaires destinés aux utilisateurs avancés qui trouveront ainsi la possibilité de personnaliser l'application et de voir tous leurs besoins entièrement satisfaits.

N'hésitez pas à vous reporter au présent Guide de l'utilisateur : il vous aidera à obtenir d'excellents résultats, de façon à la fois rapide et facile, ainsi qu'à tirer le maximum de profit de ce puissant programme OCR.

Amusez-vous bien !

## **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 permettra 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](http://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](mailto: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**)
- description générale de votre problème et message d'erreur complet (éventuellement).
- 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 **À propos d'ABBYY FineReader** (menu **Aide**).



## Chapitre 1

# **Installation et démarrage d'ABBYY 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 le CD-ROM fourni avec votre lot de logiciels.

### **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
- A propos de l'activation d'ABBYY FineReader

## Configuration matérielle et logicielle requise

### La configuration suivante est requise pour utiliser FineReader :

1. PC avec Intel® Pentium®/Celeron®/Itanium®/Xeon™, AMD K6/Athlon™/Duron™ ou processeur compatible. La vitesse du processeur doit être de 200 MHz ou supérieure
2. Microsoft® Windows® XP, Microsoft® Windows® 2000, Windows NT® 4.0 (avec Service Pack 6 ou supérieur) ou Windows ME/98 (la prise en charge de la langue correspondante est requise pour les interfaces localisées).
3. 64 Mo (Windows XP/2000/NT 4.0), 32 Mo (Windows Me/98), plus 16 Mo de mémoire pour chaque processeur supplémentaire sur les systèmes multiprocesseurs
4. 150 Mo d'espace disque disponible pour une installation typique minimale
5. 70 Mo d'espace disque pour utiliser le programme
6. scanner 100% compatible TWAIN, caméra numérique ou modem/fax
7. Une carte et un écran vidéo (résolution min. 800 × 600)
8. Un clavier et une souris ou tout autre dispositif de pointage

---

**Remarque :** Microsoft Internet Explorer 4.0 ou supérieur est requis pour effectuer des recherches dans des pages reconnues et pour lire des nouvelles sur le canal ABBYY Community news (uniquement pour ABBYY FineReader 7.0 Professional Edition).

---

## Installation de ABBYY FineReader

Le programme d'installation vous guide tout le long de l'installation d'ABBYY FineReader. Veuillez fermer toutes les autres applications avant d'installer ABBYY FineReader.

### Pour installer ABBYY FineReader :

1. Insérez le CD-ROM dans le lecteur de CD. Le programme d'installation démarre automatiquement
2. Suivez les instructions du programme d'installation.

Si le programme d'installation ne démarre pas automatiquement :

1. Cliquez sur le bouton **Démarrer** dans la barre des tâches, puis sélectionnez la commande **Paramètres/Panneau de configuration**.
2. Cliquez deux fois sur l'icône **Ajout/Suppression de programmes**.
3. Sélectionnez l'onglet **Installation/Désinstallation** et cliquez sur le bouton **Installer**.
4. Suivez les instructions du programme d'installation.

## Option d'installation

Au cours de l'installation, vous êtes invité à sélectionner l'une des deux options d'installation suivantes :

- **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** – vous pouvez choisir un nombre quelconque de composants présents dans le produit (notamment les langues de reconnaissance disponibles).

Si un message d'erreur s'affiche, reportez-vous au fichier Readme.htm, se trouvant sur le CD d'installation d'ABBYY FineReader pour obtenir de l'aide.

---

**Remarque :** si vous souhaitez utiliser les dictionnaires et les motifs utilisateur d'une version précédente de FineReader, ne désinstallez pas cette version avant d'installer la nouvelle. Tout **motif et dictionnaire** utilisateur existant pourra alors être utilisé dans la version plus récente.

---

## Installation sur un serveur et une station de travail réseau

### Installation sur un serveur réseau

Cette rubrique ne concerne que les administrateurs système.

L'installation se décompose en deux étapes. Le programme est d'abord installé sur le serveur et peut être installé depuis le serveur sur des stations de travail, via l'une des quatre méthodes suivantes :

- via Active Directory
- via Microsoft System Management Server (SMS)
- via la ligne de commande
- manuellement en mode interactif

Pour installer ABBYY FineReader 7.0 Corporate Edition sur le serveur :

1. Insérez le CD-ROM dans le lecteur de CD.
2. Exécutez le fichier setup.exe depuis le CD-ROM de FineReader avec l'option de ligne de commande /a.

Pour plus d'informations sur l'installation d'ABBYY FineReader sur des stations de travail, l'utilisation de License Manager et l'utilisation du programme sur un réseau local, reportez-vous au "Guide de l'administrateur système" qui se trouve dans le dossier **Administrator's Guide** sur le serveur où FineReader est installé.

## Démarrage de ABBYY FineReader

### Pour démarrer ABBYY FineReader :

- Sélectionnez la commande **ABBYY FineReader 7.0 Professional Edition (Corporate Edition)** dans le menu **Démarrer/Programmes**.

---

**Remarque:** avant de démarrer FineReader, assurez-vous, que le scanner est sous tension connecté à l'ordinateur. Si ce n'est pas le cas, consultez le manuel d'utilisation du scanner pour des instructions d'installation.

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

---

## À propos de l'activation d'ABBYY FineReader

ABBYY est sensible aux problèmes de piratage de logiciels et de droits sur la propriété individuelle. Le piratage de logiciels nuit aussi bien aux fabricants de logiciels qu'aux utilisateurs finaux. En effet, lorsque les utilisateurs achètent une copie légale d'un logiciel, ils s'offrent la garantie que le produit qu'ils ont acquis est un produit de qualité et qu'aucune modification ou altération n'est introduite par une tierce personne, ce qui n'est pas le cas pour les produits illégaux qui ne sont jamais sécurisés.

Une technologie de protection spéciale est utilisée pour ABBYY FineReader 7.0, afin de protéger les produits ABBYY contre la copie et la distribution illégale. Cette technologie empêche toute utilisation non autorisée des produits ABBYY par des personnes n'ayant signé aucun Accord de licence auprès du propriétaire du copyright du logiciel.

Si l'Accord de licence ne permet d'installer la copie du logiciel que sur un seul PC, toute installation supplémentaire de la même copie sur d'autres PC constitue une enfreinte à l'Accord de licence, ainsi qu'à la loi sur le copyright de la Fédération russe et à la loi sur le copyright international. La technologie d'activation contrôle le nombre de copies et empêche l'installation d'une copie dont la licence est accordée pour un nombre illimité de stations de travail. Par ailleurs, le nombre d'installations et d'activations sur un même PC ne fait l'objet d'aucune limitation que ce soit.

Selon la version du produit et le territoire où il est distribué, les limitations ci-dessous peuvent s'appliquer au niveau de la fonctionnalité. Si vous avez acheté une version de ce programme dans un des états membres de l'Union Européenne ou un des pays suivants (Norvège, Suisse, Lichtenstein), une activation sera nécessaire pour profiter des fonctions suivantes :

- Sauvegarde et impression des textes contenant des caractères cyrilliques (par exemple en russe) (FineReader 7.0 Professional Edition)
- Utilisation de FineReader 7.0 Corporate Edition

Si vous avez des questions concernant la version du logiciel que vous avez acheté et l'activation, veuillez contacter ABBYY Europe sur [sales@abbyy.com](mailto:sales@abbyy.com).

## **Comment activer le programme ?**

L'activation est très rapide et est facilitée par l'intervention d'un **Assistant d'activation** intégré au programme. Cet **Assistant d'activation** possède une interface conviviale et permet d'envoyer les informations d'activation nécessaires à ABBYY. Il s'utilise pour entrer le code d'activation (Professional Edition) ou le fichier d'activation (Corporate Edition) envoyé par ABBYY.

Les informations d'activation sont envoyées sous forme de code (**ID d'installation**) qui est généré à partir des informations relatives à l'ordinateur sur lequel le programme est installé. Aucune information personnelle sur l'utilisateur ou l'ordinateur n'est utilisée pour générer ce code, qui ne peut pas être utilisé à des fins d'identification de l'utilisateur.

### **Méthodes d'activation :**

- **Via Internet**

L'activation est effectuée automatiquement et ne prend que quelques secondes. Pour ce type d'activation, vous devez posséder une connexion Internet.

- **Via un courrier électronique**

L'utilisateur doit envoyer un courrier électronique généré par le programme et contenant les informations requises pour l'activation. Pour recevoir un message de réponse automatique rapidement, veillez à ne modifier ni l'objet ni le corps du message.

- **Par téléphone ou par fax** (Professional Edition uniquement)

L'utilisateur doit contacter le magasin ou le revendeur ABBYY le plus proche par téléphone et communiquer l'ID d'installation et le numéro de série de son produit à l'opérateur. Dans la plupart des pays, il est également possible d'envoyer un fax pour activer le programme. Le message contenant des informations requises pour l'activation est généré automatiquement et il ne reste plus qu'à l'utilisateur à l'imprimer et à l'envoyer par fax au magasin ou revendeur ABBYY le plus proche. Il reçoit alors le code d'activation nécessaire à l'activation du programme.

Si vous choisissez l'activation via Internet, l'ensemble du processus est effectué de façon automatique. Si vous optez pour une activation par téléphone ou par fax ou via un courrier électronique, vous devez entrer les informations du Fichier d'activation envoyé par ABBYY dans les champs correspondants de l'Assistant d'activation.

Une fois l'activation terminée, le programme peut être utilisé sur l'ordinateur actuel sans aucune restriction.

ABBYY FineReader 7.0 peut être réinstallé sur le même ordinateur un nombre de fois illimité sans qu'il soit nécessaire de l'activer à chaque fois. Toutefois, si vous procédez à des mises à niveau majeures, au formatage de votre disque dur ou à la réinstallation du système d'exploitation, vous serez peut-être amené à effectuer une activation supplémentaire.

## **Activation de la politique de confidentialité**

Pour activer FineReader en mode version complète, vous devez activer le programme. ABBYY vous demande d'activer le produit afin de vérifier si le produit que vous installez est un produit ABBYY authentique. Lorsque vous activez votre produit, aucune information personnelle n'est transmise à ABBYY. L'activation est entièrement anonyme.

Au moment de l'activation, l'assistant d'activation de FineReader crée un ID d'installation unique qui représente la configuration de votre PC lors de l'activation. Cet ID d'installation ne contient aucune information personnelle, aucune information sur les logiciels ou les données sur votre ordinateur, pas plus qu'il n'en contient sur le modèle ou le fabricant de votre ordinateur. Il identifie l'ordinateur uniquement dans le but de l'activation. Outre l'ID d'installation, l'assistant envoie le nom, le numéro de série, de la version et la langue de l'interface de votre copie de FineReader au serveur d'activation d'ABBYY. Ces informations ne sont utilisées que dans le cadre de la sélection de la langue appropriée et du contenu du message de réponse, afin de vous informer des résultats de la procédure d'activation et elles ne sont utilisées à aucune autre fin.

L'assistant d'activation de FineReader gère la détection et la tolérance de votre PC aux modifications. Les mises à niveau mineures ne requièrent pas de réactivation. Si vous procédez à la révision complète de votre PC ou à la réinstallation de votre système d'exploitation, vous pouvez être amené à activer une nouvelle fois votre produit.

## Chapitre 2

# Démarrage rapide

Ce chapitre vous explique comment saisir un document sans rien connaître du 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 7.0.

### **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 Edition** ou **Corporate Edition**). La fenêtre principale de FineReader apparaît.
3. Insérez dans le scanner la page que vous souhaitez lire.
4. Cliquez sur la flèche 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 **Démo**. Ce dossier est situé dans le dossier où vous avez installé FineReader. FineReader vous guide pas à pas dans la procédure.

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 "photo-graphie" 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. Quand vous avez terminé la vérification du document, l'assistant **Numériser&Lire** vous propose d'envoyer le texte reconnu à une application spécifiée, 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 textes 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.

Par défaut, un nouveau lot est ouvert au démarrage de FineReader. Vous pouvez continuer à travailler avec ce nouveau lot ou en ouvrir un autre. Pour plus de détails, consultez la section Informations générales sur l'utilisation des lots.







Le menu principal apparaît 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 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.

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

- Cliquez sur  ou  sur la barre d'outils **Standard**, ou
- Cliquez à l'aide du bouton droit de la souris sur la fenêtre **Lot** et sélectionnez l'élément **Afficher>...** dans le menu local.

La barre d'état apparaît 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.

Le reste de la fenêtre principale est occupée par les fenêtres **Lot**, **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**.

<b>Dispositions conseillées pour les fenêtres à l'écran :</b>	<b>Utile si/quand :</b>
Fenêtre <b>Lot</b> sur la gauche ; <b>Affichage de lot : Vignettes</b> ; fenêtres <b>Image</b> , <b>Texte</b> et <b>Zoom</b>	Les pages du lot ne sont pas trop nombreuses.
Fenêtre <b>Lot</b> en haut ; <b>Affichage de lot : Détails</b> ; fenêtres <b>Image</b> , <b>Texte</b> et <b>Zoom</b>	Les pages du lot sont nombreuses.
Fenêtre <b>Lot</b> en haut ; <b>Affichage de lot : Détails</b> ; fenêtres <b>Image</b> et <b>Zoom</b>	Vous effectuez une analyse de la présentation et une reconnaissance.
Fenêtre <b>Lot</b> en haut ; <b>Affichage de lot : Détails</b> ; fenêtres <b>Texte</b> et <b>Zoom</b>	Vous modifiez le texte reconnu.

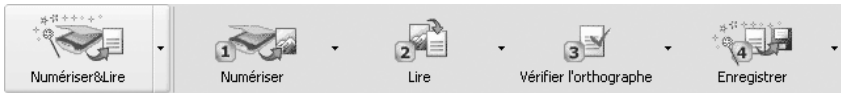
Pour passer d'une fenêtre à l'autre :

- 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 savoir à quoi sert un bouton d'une barre d'outils, placez le pointeur de la souris au-dessus. La bulle d'aide du bouton apparaît et la barre d'état affiche des explications.

## 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 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 réexécuter 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 vous indique les actions à mener pour obtenir tel ou tel résultat.

**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.

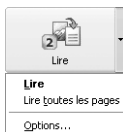
**Numériser des images multiples** : numérise les images jusqu'à ce que vous arrêtez l'opération. Pour ce faire, sélectionnez la commande

**Arrêter la numérisation** dans le menu **Fichier**.

**Répertoire de veille** (Corporate Edition uniquement) : lance le contrôle de dossiers (toutes les images ajoutées au dossier spécifié seront automatiquement ouvertes dans la fenêtre FineReader). Pour désactiver le contrôle de dossiers, sélectionnez **Désactiver le répertoire de veille** dans le menu **Fichier**.

**Options** : ouvre 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.

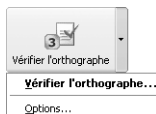
**Tout lire** : lit toutes les pages du lot non reconnues.

**Options** : ouvre l'onglet **Reconnaissance** (boîte de dialogue

**Options**) où vous pouvez 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** : ouvre l'onglet **Vérification orthographique** (boîte de dialogue **Options**) où vous pouvez définir les options de la 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.

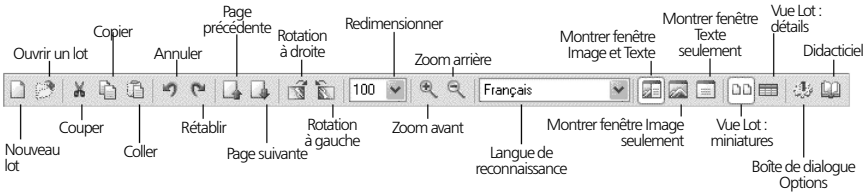
**Envoyer les pages sélectionnées vers** : cette option permet d'exporter seulement les pages de lot sélectionnées. Dans ce cas sélectionnez les pages à exporter puis cette option, et enfin l'application vers laquelle vous souhaitez exporter les pages. FineReader exporte alors les pages vers l'application 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.

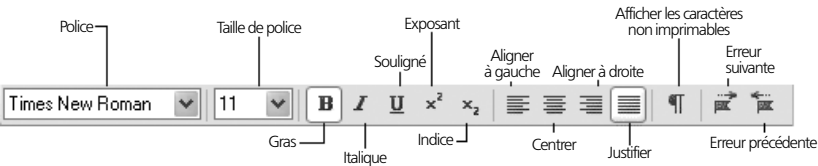
**Options** : ouvre l'onglet **Formatage** (boîte de dialogue **Options**) où vous pouvez définir les options d'enregistrement.

## 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.

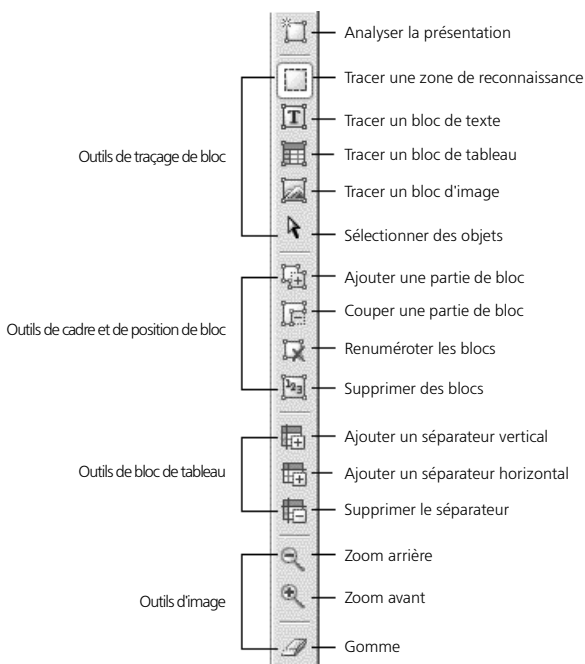


## Barre d'outils Formatage



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

## Barre d'outils Image



---

**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 selon la résolution de l'écran. Pour afficher tous les boutons disponibles, vous devez augmenter 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 nouveau bouton à une barre d'outils :

1. Sélectionnez la catégorie requise dans le champ **Catégories**.  
**Remarque :** les commandes de la liste sont classées par commandes de menu. Le choix de 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 sélectionnée 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 dans 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 faire monter (ou descendre) 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 donnée, sélectionnez cette dernière dans la liste **Barres d'outils**, puis cliquez sur le bouton **Restaurer**. Pour restaurer les boutons par défaut dans toutes les barres d'outils, cliquez sur le bouton **Restaurer tout**.
-



## Chapitre 3

# **Caractéristiques générales d'ABBYY FineReader**

FineReader vous offre des outils assurant la saisie numérique de documents. Il vous suffit de cliquer une fois sur le bouton Numériser&Lire : l'application s'occupe du reste et vous évite de perdre du temps en complications. Vous pouvez transmettre le texte reconnu à une application de traitement de texte ou de tableur, l'enregistrer au format RTF/DOC/Word XML, 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 7.0
- 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) vous permettent de récupérer 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 la récupération 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 OCR de l'image.

### **Attardons-nous sur la seconde étape.**

Le traitement OCR de l'image de 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 interviennent 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 interrelations. 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 distordus, 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 7.0**

### **Précision de la reconnaissance**

- La précision de la reconnaissance a été améliorée de 25 %. Le programme peut désormais analyser et reconnaître des documents à la mise en page complexe, notamment des documents comportant du texte sur un arrière-plan raster ou en couleur, des documents avec des tableaux complexes (y compris des tableaux avec des séparateurs blancs et des tableaux avec des cellules en couleur).
- Des dictionnaires anglais et allemand spécialisés comprenant la terminologie légale et médicale fréquemment utilisée ont été ajoutés. Ils permettent d'améliorer une précision de la reconnaissance de textes spécialisés dans le domaine légal et médical inégale.

### **Prise en charge XML et intégration à Microsoft Office**

- FineReader inclut désormais un nouveau format d'enregistrement : Microsoft Word XML. Les utilisateurs de Microsoft Office 2003 pourront désormais profiter de tous les avantages fournis par le format XML.
- L'intégration à Microsoft Word 2003 permet d'allier la puissance des deux applications de traitement de textes reconnus. Vous pouvez désormais vérifier et modifier les résultats de la reconnaissance via les outils communs de Word Microsoft, tout en comparant les résultats exportés avec l'image d'origine : la fenêtre Zoom de FineReader s'affiche désormais dans Microsoft Word.
- De nombreuses caractéristiques ont été ajoutées à FineReader 7.0 afin de faciliter votre travail. Lors de la création d'un document Word, vous pouvez lancer FineReader, reconnaître un élément de texte et l'insérer juste en dessous du curseur dans le document Word en cours. Ceci signifie que vous pouvez collecter des informations provenant de divers documents papier ou PDF et les regrouper dans un seul document électronique. Les résultats de la reconnaissance peuvent désormais être envoyés en pièce jointe à un courrier électronique, dans tout format d'enregistrement pris en charge.

## **Reconnaissance améliorée des documents PDF**

- La qualité de la reconnaissance des documents PDF a été considérablement améliorée. La plupart des PDF contiennent du texte en arrière-plan. FineReader 7.0 peut extraire ce texte et l'utiliser lors de la vérification des résultats de la reconnaissance, améliorant ainsi la qualité de la reconnaissance.
- Vous pouvez désormais modifier des documents PDF reconnus dans l'Éditeur de FineReader : les changements effectués seront enregistrés dans l'un des formats d'enregistrement de documents PDF pris en charge.
- Les documents PDF créés dans FineReader sont optimisés en vue de leur publication sur le Web : vos lecteurs pourront voir les premières pages d'un document, en attendant son téléchargement complet.

## **Nouvelles options d'enregistrement**

- Un nouveau format d'enregistrement, **Microsoft PowerPoint**, vous permet de créer rapidement des présentations ou de modifier des présentations existantes.
- Lors de l'enregistrement des résultats dans **Microsoft Word**, la taille du fichier .doc sera réduite par rapport aux versions précédentes du fichier. La rétention de la mise en forme des documents contenant des séparateurs divers a été améliorée et de nouvelles options d'enregistrement d'images ont été ajoutées.
- La rétention des éléments de mise en forme complexe a été améliorée au format **HTML** (par exemple, le texte flottant encadrant des images non-rectangulaires). La taille des fichiers HTML qui en résultent est désormais réduite, ce qui est important pour les documents publiés sur Internet.

## **Facilité d'utilisation**

- **L'interface du programme**, désormais plus conviviale et intuitive, a été mise à jour. Les paramètres avancés sont à présent plus accessibles. Les barres d'outils d'édition ont été déplacées dans la fenêtre dans laquelle les résultats de la reconnaissance sont affichés. De nouveaux outils de personnalisation vous permettent d'ajuster les fenêtres de FineReader ; vous pouvez par exemple régler indépendamment les paramètres de zoom de chaque fenêtre.
- Grâce au **Didacticiel** mis à jour, l'utilisateur débutant pourra immédiatement commencer à utiliser le programme, tandis que l'utilisateur confirmé trouvera des conseils pratiques pour régler le programme afin d'en optimiser la qualité et la productivité.

### **Caractéristiques avancées**

**FineReader Professional Edition** inclut désormais les caractéristiques suivantes qui étaient auparavant accessibles par les utilisateurs de **FineReader Corporate Edition** :

- Reconnaissance améliorée des **codes barre** et prise en charge de la reconnaissance pour les codes barre PDF-417 2D.
- **Outil de fractionnement des images** qui permet de fractionner l'image en plusieurs zones et d'enregistrer chaque zone en tant que page distincte. Ce mode est très pratique pour la reconnaissance de cartes de visite numérisées sur une page, de livres et d'impressions de présentations PowerPoint.
- **Recherche avec prise en charge morphologique**. Tout lot créé dans ABBYY FineReader peut être utilisé en tant que base de données de petite taille avec des fonctions de recherche sur le texte entier. Vous pouvez rechercher des mots dans toutes les formes grammaticales possibles (cette caractéristique est valide pour les 34 langues prenant en charge des dictionnaires).
- **Prise en charge de la technologie Hyper-Threading d'Intel**. Cette technologie permet d'accroître considérablement la productivité lors de la reconnaissance de grandes quantités de documents.

### **Caractéristiques avancées supplémentaires :**

- **Numérisation duplex**. Si vous numérisez un document recto verso avec un scanner duplex, vous obtenez deux pages différentes, chacune correspondant à une face du document. Cette option peut être désactivée si vous n'avez pas besoin de numériser en mode duplex.
- Les fichiers image **JPEG 2000** peuvent être ouverts et enregistrés.

### **Capacités réseau de FineReader Corporate Edition**

- **Les méthodes d'installation principales depuis le serveur sur des stations de travail sont prises en charge**. FineReader Corporate Edition prend en charge toutes les méthodes principales d'installation réseau : via Active Directory, via Microsoft Systems Management Server ou via une ligne de commande.
- **Prise en charge de périphériques multifonctions, y compris les périphériques multifonctions en réseau**. Les périphériques multifonctions (MFP) qui allient les fonctions d'un scanner, d'une imprimante, d'une photocopieuse et d'un fax deviennent de plus en plus populaires. FineReader gère l'utilisation de tels périphériques, les deux connectés à une station de travail et installés sur un réseau. Des paramètres de programme spécifiques permettent aux utilisateurs d'ouvrir automatiquement des images n'importe où depuis le réseau ou depuis un serveur FTP et les reconnaître.

- **Modèles de licences d'entreprise multiples.** Outre le programme de licence concomitante, d'autres schémas de licence ont été mis en place. Vous avez la possibilité de choisir le schéma de licence qui vous convient le mieux.
- **License Manager est un nouvel utilitaire pour la gestion de licences dans un environnement réseau.** License Manager vous permet de contrôler l'utilisation de **FineReader Corporate Edition** sur des stations de travail, de réserver des licences pour des types de stations de travail spécifiques et d'ajouter des licences.

Pour plus d'informations sur l'installation d'ABBYY FineReader sur des stations de travail, l'utilisation de License Manager et l'utilisation du programme sur un réseau local, reportez-vous au "Guide de l'administrateur système" qui se trouve dans le dossier **Administrator's Guide** sur le serveur où FineReader est installé.

## **Formats d'enregistrement de documents 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)
- Document Microsoft Word XML (\*.XML) (Microsoft Word 2003 uniquement)
- Adobe Acrobat Format (\*.PDF)
- **HTML**
- Format Microsoft PowerPoint (\*.PPT)
- Fichier CSV (Comma Separated Values – valeurs séparées par virgule) (\*.CSV)
- Texte simple (\*.TXT). FineReader accepte diverses pages de code (Windows, DOS, Mac, ISO), ainsi que le codage Unicode.
- Feuille Microsoft Excel (\*.XLS)
- **DBF**

## **Support de formats d'image**

**FineReader ouvre les fichiers d'image dans les formats suivants :**

**PDF:**

Fichiers dans le format PDF, version 1,3 ou ultérieure.

**BMP :**

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

**PCX, DCX :**

- 2 bits – noir et blanc
- 4- et 8-bit – Palette
- 24-bit – TrueColor

**JPEG :**

- niveaux de gris et couleurs vraies

**JPEG 2000 :**

- 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
- 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

**JPEG 2000 :**

- 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 vous 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 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
- Acquisition d'images depuis le répertoire de veille
- Numérisation de pages doubles
- Ajout d'images de cartes de visite au lot
- Travailler avec une image
- Numérotation des pages
- 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 un gestionnaire TWAIN :

- **en utilisant sa propre interface** : dans ce cas, vous utilisez la boîte de dialogue **Paramètres de scanner** de FineReader pour définir les options de numérisation ; l’option correspondante est **Utiliser l’interface FineReader** ;
- **en utilisant l’interface TWAIN du scanner** : dans ce cas, vous 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.**

Quand vous choisissez l’option **Utiliser l’interface de la source TWAIN**, l’option d’aperçu de l’image devient généralement disponible. L’aperçu vous 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 commuter 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 voulez voir 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** sur 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 et/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** apparaît et montre une “photographie” de la page numérisée dans la **fenêtre principale** de **FineReader**.

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

Si la numérisation n'a pas démarré directement, l'une des deux boîtes de dialogue suivantes apparaît :

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

**Conseil :** Si vous souhaitez lancer la reconnaissance dès que les images sources sont numérisées, vous pouvez utiliser la commande **Numériser&Lire** ou **Numériser&Lire avec images multiples**.



Cliquez sur la flèche 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 de qualité d'impression 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 dans une résolution de 400 à 600 ppp (au lieu des 300 ppp par défaut), en niveaux de gris ou en couleur (au lieu de 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 :**

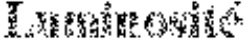

- Si vous numérisez vos images au moyen de l'interface TWAIN de **FineReader**, sélectionnez la commande **Paramètres de scanner** du menu **Outils**. La boîte de dialogue **Paramètres de scanner apparaît**. Définissez les options de numérisation dans la boîte de dialogue.
- Si vous numérisez les images au moyen de l'interface de la source TWAIN, la boîte de dialogue TWAIN de votre scanner apparaît 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 de scanner ; par exemple, le terme "seuil" peut s'appliquer à la luminosité. Le réglage de 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. Vous vérifier sa lisibilité, examinez-la dans la fenêtre **Zoom**.

## **Luminosité** – exemple d'une bonne image (du point de vue de l'OCR)

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 cette allure :	Remède possible :
 <p>caractères "déchirés" ou très clairs</p>	<ul style="list-style-type: none"> <li>● Réduisez la luminosité pour rendre l'image plus sombre.</li> <li>● Essayez de numériser en niveaux de gris (le réglage automatique de la luminosité est utilisé dans ce mode).</li> </ul>
 <p>caractères déformés, collés ou remplis</p>	<ul style="list-style-type: none"> <li>● Augmentez la luminosité pour rendre l'image plus claire.</li> <li>● Essayez de numériser en niveaux de gris (le réglage automatique de la luminosité est utilisé dans ce mode).</li> </ul>

## Numérisation de documents multipages

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

- Si vous numériser vos images au moyen 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 numériser vos images au moyen de **l'interface de la source TWAIN**, la **boîte de dialogue TWAIN** du scanner n'est pas refermée quand le scanner a fini de numériser une page. 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 vous permettent de numériser plus facilement des pages multiples sur un scanner à plat sans chargeur automatique. Deux alternatives s'offrent à vous :

- Définissez la valeur de la pause, c'est-à-dire le délai séparant le moment où le scanner finit de numériser une page et celui où il commence à numériser la 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'au bout du délai en secondes que vous avez défini ; ainsi, vous disposez de suffisamment de temps pour insérer la page suivante. Une fois la pause terminée, la numérisation se poursuit 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 apparaît. 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-en une nouvelle et cliquez sur le bouton **Numériser** de la boîte de dialogue **TWAIN** de votre 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 :** vous avez également la possibilité de contrôler plus directement la qualité des images numérisées. Pour ce faire, sélectionnez l'option **Ouvrir l'image pendant la numérisation** sur l'onglet **Numérisation** (menu **Outils>Options**). Si vous activez cette option, chaque page est ouverte dans la fenêtre **Image** immédiatement après la numérisation. 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 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.
- Dans Microsoft Outlook ou Windows Explorer : cliquez sur le fichier image que vous souhaitez ouvrir et faites-le glisser dans la fenêtre minimisée de FineReader. L'image est ajoutée au lot en cours et ouverte dans la fenêtre **Image**.

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

---

**Conseil** : si vous voulez 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** apparaît.
2. Sélectionnez les images à reconnaître.

---

## Acquisition d'images depuis le répertoire de veille

### (Corporate Edition uniquement)



Si votre société utilise des périphériques multifonctions en réseau alliant les fonctions d'un scanner, d'une imprimante et d'une photocopieuse, vous pouvez configurer FineReader de sorte à recevoir automatiquement les images provenant de tels périphériques.

L'une des méthodes qui peut être utilisée est le contrôle de dossiers. FineReader gère le contrôle d'un dossier spécifique sur un disque local, un réseau ou un serveur FTP. Dans ce

mode, le programme ouvre automatiquement toutes les images numérisées, fax et documents PDF ajoutés au répertoire de veille. Une fois l'image ajoutée au lot FineReader, elle est supprimée du dossier contrôlé.

Pour activer le répertoire de veille :

- Sélectionnez l'option **Répertoire de veille...** dans le menu **Fichier**, ou
- Sélectionnez l'option **Répertoire de veille** dans le menu local de la barre d'outils **Numériser et Lire**.

Lorsque le mode Répertoire de veille est activé, l'icône du bouton **Numériser** change d'apparence, la boîte de dialogue **Répertoire de veille** s'ouvre et l'icône  s'affiche dans la barre d'état. En cas d'erreur, l'icône change d'aspect et apparaît sous la forme : . Cliquez deux fois sur cette icône pour afficher le message d'erreur correspondant.

Pour désactiver le mode Répertoire de veille :

- Sélectionnez l'option **Désactiver le répertoire de veille...** dans le menu **Fichier**, ou
- Cliquez sur le bouton **Arrêter** dans la barre d'outils **Numériser et Lire**.

## 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 reconnaissance est meilleure si la page double est segmentée après numérisation, de sorte que chaque page du lot correspond à une seule page du livre. La reconnaissance et l'analyse de la présentation sont effectuées pour chaque page séparément, chaque page étant redressée si nécessaire.

Si vous souhaitez segmenter des pages doubles :

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

Si cette option est définie, 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 en deux de manière incorrecte, désélectionnez la case à cocher **Partager les pages doubles**, numérisez à nouveau la page double ou rajoutez l'image correspondante au lot, puis essayez de segmenter 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 (autant que le scanner peut en contenir). La qualité de la reconnaissance est optimale (grâce





au redressement, par exemple) lorsque chaque carte de visite est numérisée sur une page à part. Le système présente des outils de partage automatique et manuel des images contenant des 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 est incorrectement divisée dans les cartes de visite, essayez de la fragmenter 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.

## **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, par exemple pour conserver la numérotation initiale ou numériser une série de pages classées par numéro. Pour cela :

- Sélectionnez l'option **Demander un numéro de page avant d'ajouter la page au lot** sur 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** sur 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.

## 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

### 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 est 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 l'onglet **Numériser/Ouvrir une image** (menu **Outils > Options**) pour les nettoyer avant que l'application ne les ajoute au lot.

---

### 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 une apparence standard, c'est-à-dire avec une police noire sur fond blanc. Pour cela :

- 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 l'onglet **Numériser/Ouvrir une image** (menu **Outils > Options**) avant de les ajouter au lot.

---

### Faire pivoter ou basculer l'image

La qualité de 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 FineReader a mal détecté l'orientation de la page, désélectionnez l'option **Détecter l'orientation d'image (pendant la reconnaissance)** 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**.

### 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. Pour cela :

- Sélectionnez l'outil  et sélectionnez la zone de l'image que vous voulez effacer en maintenant enfoncé le bouton gauche de la souris. Relâchez le bouton pour effacer la zone sélectionnée.

### 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.

### 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.

### 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.).

### 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  .

---

## 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 vos images en niveaux de gris avec **l'interface de la source TWAIN** et que les images numérisées ne contiennent ni image, police ni arrière-plan couleur ou si vous ne souhaitez pas conserver les couleurs sur les images numérisées. Si vous définissez cette option, 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.

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 en 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 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** – 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 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 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 souhaitez définir manuellement le type de document :

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

### Types de document 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 a 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 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



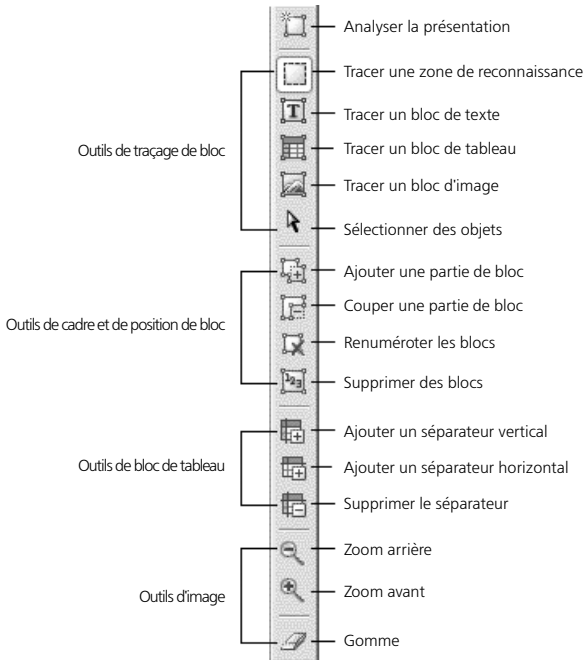
2. Utilisez l'option **Tableau sans cellules fusionnées** si vous traitez un tableau qui ne contient pas de cellules fusionnées. Par exemple :

<b>Température</b>	
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 prérequis 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 renuméroter 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 traitez 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 blocs 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 :

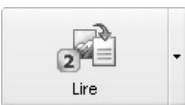
- Cliquez sur le bouton **2 – Lire de la barre d'outils Assistant ou**
- Sélectionnez la commande appropriée dans le menu **Traitement** :

**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.

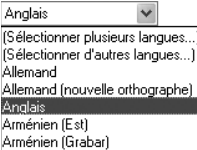
---

## Langue de reconnaissance

FineReader reconnaît aussi bien les documents unilingues que multilingues (par exemple anglais–français). Lors de la reconnaissance de documents en anglais ou en allemand, vous pouvez également utiliser ces langues avec les dictionnaires spécialisés correspondants. Outre



les dictionnaires spécialisés dans le domaine médical et légal, ces langues comprennent aussi des dictionnaires FineReader généraux. 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 sur 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 “?” ou “i”). 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. La langue n'est pas prise en charge par FineReader. Consultez la liste complète des langues de reconnaissance dans ABBYY FineReader l'aide, section “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. Pour ajouter d'autres langues, exécutez le programme d'installation de FineReader en mode installation des composants (**Installation personnalisée**), désactivez tous les éléments à l'exception de **Langues de reconnaissance** et cliquez sur le bouton **Détails**. La liste des langues prises en charge par FineReader est affichée. Sélectionnez les langues requises.

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

---

### **Pour montrer/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 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 "w" et "a").

### **Pour modifier le type d'impression :**

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

**Remarque :** quand vous avez terminé la reconnaissance de documents issus d'une imprimante matricielle ou d'une machine à écrire, 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** sur 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 vous sera 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 FineReader. 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 comporte 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 utilisateurs.

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 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 l'onglet **Reconnaissance**, menu **Outils>Options**, groupe **Apprentissage**. Le nom de gabarit par défaut ("Default") 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.ptn**) dans le dossier de lot actif.
4. Modifiez votre gabarit.
5. Désactivez le mode apprentissage (cliquez sur le bouton radio **Utiliser gabarit utilisateur** 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 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 l'onglet **Reconnaissance**, menu **Outils>Options**, groupe **Apprentissage gabarit utilisateur**.
3. Si l'option **Utiliser gabarits intégrés** est active, FineReader tente de lire le texte au moyen 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 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 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 vous trompez pendant l'apprentissage, vous pouvez cliquer sur le bouton **Précédent** pour restaurer la position précédente du cadre ; la dernière paire “caractère–image” est alors supprimée du gabarit. Cette opération d'annulation est limitée : elle n'est applicable 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**. Vous pouvez déplacer le bord du cadre au moyen 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 fait pas la différence entre 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 dans le gabarit comme un même caractère, 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 nouvelle 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 nouvelle 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.

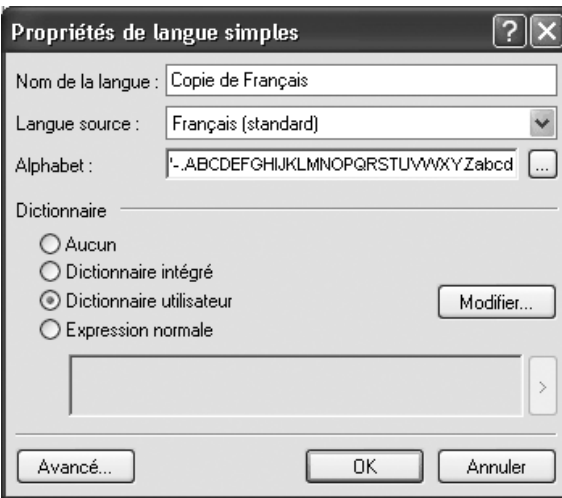


Pour créer une langue ou un groupe de langues, ouvrez la boîte de dialogue de l' **Editeur de langue** (menu **Outils**, commande **Editeur de langue**).


## Création d'une nouvelle langue

### Pour créer une nouvelle 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 apparaît, sélectionnez le bouton radio **Créer une copie de la langue** et sélectionnez une langue source pour la nouvelle.
3. La boîte de dialogue **Propriétés de langue simples** apparaît.



### 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). Vous avez les possibilités suivantes :
  - **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 sûrs 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 (voir la liste des langues prise en charge par la vérification dans ABBYY FineReader l'aide, la section "Langues prises en charge").
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 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 vous offre 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 dans cette liste la langue de votre choix.

---

**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. Ce que vous pouvez faire avec les mots éventuellement mal orthographiés Les possibilités sont les suivantes :
- 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 aller sur 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 fenêtre de 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 sur 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 fléchie, sélectionnez l'option **Sauter l'invite de forme de mot** (dictionnaire anglais seulement) 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 :**

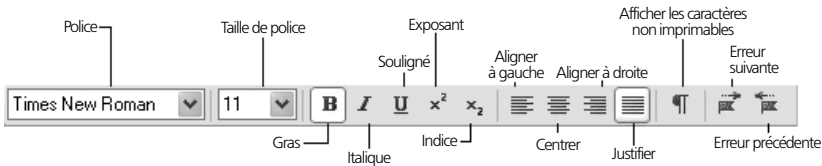
1. Sélectionnez la commande **Afficher les dictionnaires** dans le menu **Outils**. Sélectionnez la langue et cliquez sur le bouton **OK**. Une boîte de dialogue s'ouvre.
2. Sélectionnez le mot que vous souhaitez supprimer et cliquez sur le bouton **Supprimer**.

## **Modification de texte dans FineReader**

---

**Remarque :** si la fenêtre de texte de FineReader n'affiche pas les caractères correctement (des signes “?” ou “ı” 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 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 votre texte vers une application externe, sa mise en page est conservée conformément aux options de préservation de la présentation. Définissez ces options dans 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 vous suffit pour cela de désélectionner l'option **Mettre en évidence les caractères incertains** de l'onglet **Affichage** (menu **Outils > Options**).

L'éditeur FineReader possède deux modes d'affichage pour les 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 passer du mode brouillon au mode normal et réciproquement en cliquant sur  (pour passer en mode normal) ou sur  (pour passer en 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** dans l'onglet **Affichage**.

### **L'éditeur intégré de FineReader vous offre 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 **Edition** 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 **Edition** 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 **Edition** 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 :**

vous pouvez :

- sélectionner la commande **Chercher** dans le menu **Edition** ou
- appuyer sur **CTRL+F**

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 **Edition** 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 à votre disposition :
  - 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

vous pouvez :

### – 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 qui vous permettent de :

- 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 et sélectionnez 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 **Édition**).

---

**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

Les résultats de la reconnaissance peuvent être enregistrés dans un fichier, exportés vers une application externe sans être enregistrés sur le disque, copiés dans le Presse-papiers ou envoyés par courrier électronique dans tout type de format d'enregistrement pris en charge. Il est possible d'enregistrer toutes les pages ou uniquement les 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), 2002 (10.0) et 2003 (11.0) ;  
Microsoft Excel 6.0, 7.0, 97 (8.0), 2000 (9.0), 2002 (10.0) et 2003 (11.0) ;  
Microsoft PowerPoint 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, DOC et Word XML
- Enregistrement du texte reconnu au format PDF
- Enregistrement du texte reconnu au format HTML
- Enregistrement de texte reconnu au format PowerPoint
- 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 sur 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, Word XML, PPT 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/**Word XML**, 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 ; cliquer 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, **Word XML**, **PPT** et HTML.

### **Résolution des images (enregistrement aux formats RTF/DOC/Word XML, PDF, PPT et HTML)**

Vous pouvez réduire 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 réduire la résolution des images (et donc la taille des fichiers HTML), sans que la qualité visuelle des images en soit affectée. Pour ce faire, entrez la valeur de résolution requise dans le champ Réduire la résolution des images à, situé dans la page de l'onglet Formats>RTF/DOC/**Word XML** (PDF, **PPT**, HTML).

**Remarque** : si, dans le champ Réduire la résolution des images à, 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 RTF, DOC, Word XML, PDF, PPT et HTML)**

Lorsque vous enregistrez du texte aux formats PDF, PPT et HTML, les images sont enregistrées au format JPEG.

Lors de l'enregistrement au format **RTF**, **DOC** et **Word XML**, vous pouvez sélectionner le format JPEG pour enregistrer les images.

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**, **RTF**, **DOC**, **Word XML**, **PPT** 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 (PPT, RTF/DOC/Word XML, HTML)**.

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

Par défaut, les polices de l'onglet Formatage sont utilisées lorsque vous enregistrez au format **RTF**, **DOC**, **Word XML**, **PPT** et **HTML**. Vous pouvez spécifier les polices à utiliser. Vous pouvez changer les polices dans la fenêtre **Texte** ou définir d'autres polices sur 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 multipage, 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 nouveau 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, DOC et Word XML

---

**Important** : l'enregistrement au format Word XML est possible uniquement avec Microsoft Word 2003.

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

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

---

### Conseils :

- 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 sur l'onglet RTF/DOC/Word XML, 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.
- Lors de l'enregistrement des résultats dans Word XML, l'image reconnue peut être affichée dans la fenêtre Zoom intégrée à Microsoft Word. Dans cette fenêtre s'affiche l'image agrandie de la ligne ou de la portion du document. Par défaut, l'option **Activer la fenêtre Zoom de FineReader dans Microsoft Word 2003** est cochée dans l'onglet **RTF/DOC/Word XML**.

## 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 sur 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 sur l'onglet **PDF** de la boîte de dialogue **Paramètres de format**.
  2. 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](http://www.paratype.com/shop)).
  3. 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). Pour plus de détails sur les polices de Type 1, consultez ABBYY FineReader l'aide, section "Utilisation des polices de Type 1 lors de l'exportation au format PDF".
  4. 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 sur 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 sur 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** sur 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 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 sur 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 qu'il prend en charge.

**Remarque :** si votre navigateur ne prend pas complètement en charge HTML4/CSS (par exemple Microsoft Internet Explorer 3.0 ou version précédente, Netscape 4.x, etc.), il est recommandé d'utiliser le mode d'enregistrement **Simple**.

---

**Pour définir le format HTML souhaité :**

- Cliquez sur le bouton radio approprié sur 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 l'onglet **HTML** de la boîte de dialogue **Paramètres de format**.

---

## **Enregistrement de texte reconnu au format PowerPoint**

Les modes de rétention de mise en page peuvent être définis dans l'onglet **Formatage** de la boîte de dialogue **Options** (menu **Outils>Options**).

---

**Remarque :**

- Lors de l'enregistrement de texte au format PPT, les polices utilisées sont celles définies dans l'onglet **Formatage** de la boîte de dialogue **Options** (menu **Outils>Options**) ou celles définies au cours de la modification de texte dans la fenêtre **Texte**.
- 

**Important :** lorsque vous enregistrez les résultats au format PPT, les résultats sont enregistrés en tant que fichiers HTML particuliers comprenant différents éléments de la présentation. Pour enregistrer la présentation dans un seul fichier, celle-ci doit être à nouveau enregistrée dans PowerPoint (sélectionnez **Enregistrer sous** dans le menu **Fichier**, puis spécifiez **PPT** comme format d'enregistrement).

---

## **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 multipage). 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 multipage**.

---

**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 multipage), 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

# 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 plus d'informations sur l'installation d'ABBYY FineReader sur des stations de travail, l'utilisation de License Manager et l'utilisation du programme sur un réseau local, reportez-vous au "Guide de l'administrateur système" qui se trouve dans le dossier **Administrator's Guide** sur le serveur où FineReader est installé.

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-utilisateur 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, reconnu 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 sur un même lot en réseau


(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). FineReader détectera et traitera automatiquement les images que vous numérisez.

4. 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 sur  dans la barre d'outils **Standard**, ou
- 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 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 lorsque 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\7.00\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. Lorsque 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 lorsque 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.

---





## Chapitre 10

# 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 \*.fbt.

### 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
- Enregistrement d'un lot
- Fermeture d'un lot
- Suppression d'un lot
- Paramètres du lot
- Recherche en mode texte intégral dans les pages de lot reconnues

## Informations générales concernant l'utilisation des lots

Par défaut, un nouveau lot est créé au démarrage de FineReader. 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 possibilités sont à votre disposition 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 réduite 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 <b>Détails</b> 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 <b>Vignettes</b> . Cliquez deux fois sur une page pour l'ouvrir.

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

- Cliquez sur  ou sur  dans la barre d'outils **Standard**, ou
- 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 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 :**

- Sélectionnez la commande **Nouveau lot** dans le menu **Fichier** ou
- Cliquez sur le bouton **Nouveau lot** (  ).

## **Ouverture d'un lot**


Par défaut, un nouveau lot est créé au démarrage de FineReader.

---


**Remarque :** si vous souhaitez qu'à son démarrage, FineReader ouvre le dernier lot que vous avez utilisé, cochez l'option **Ouvrir le dernier lot au démarrage** dans l'onglet **Général** de la boîte de dialogue **Options (Outils>Options)**.

---

Pour ouvrir un autre lot :

1. Sélectionnez la commande **Ouvrir un lot** dans le menu **Fichier** ou cliquez sur le bouton **Ouvrir un lot** (  ). 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. Si vous avez travaillé sur un nouveau lot et que vous pensez avoir besoin de ce fichier ultérieurement, assurez-vous d'enregistrer ce lot avant de quitter le programme.

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 FineReader est actuellement exécuté sur votre ordinateur, les fichiers sélectionnés sont ajoutés au lot actuel. Dans le cas contraire, une copie de FineReader est lancée et les fichiers sélectionnés sont ajoutés à un nouveau lot.  
Cette commande du menu contextuel n'est activée que pour les fichiers au format pris en charge par FineReader 7.0.

## 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 **Renumeroter 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 Renumeroter les pages :

1. Sélectionnez une ou plusieurs pages.
2. Sélectionnez la commande **Renumeroter 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 **Renumeroter 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 **Renumeroter 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.

---

## Enregistrement d'un lot

### Pour enregistrer un lot :

- Sélectionnez **Enregistrer le lot sous** dans le menu **Fichier**.
- Dans la boîte de dialogue **Enregistrer le lot sous** qui s'affiche, spécifiez le nom du lot et l'emplacement dans lequel il sera enregistré.

## 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**.

## 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**.

## Paramètres du lot

### Pour enregistrer les paramètres du lot dans un fichier :

- Cliquez sur le bouton **Enregistrer** de l'onglet **Général** (**Outils>Options**). La boîte de dialogue **Enregistrer le modèle de lot sous** s'ouvre.
- Entrez un nom de fichier.

Les paramètres suivants sont enregistrés : les paramètres des onglets **Reconnaissance**, **Numériser/Ouvrir une image**, **Formatage** et **Vérification orthographique**, ainsi que les paramètres de l'onglet **Paramètres de mise en forme**. Les langues utilisateur, les groupes de langues utilisateur et les motifs utilisateur sont également enregistrés dans un fichier. Si vous souhaitez appliquer ce modèle à tous les lots, cochez la case **Appliquer ce modèle aux nouveaux lots** dans la boîte de dialogue **Enregistrer le modèle de lot sous**.

### Pour renvoyer les paramètres par défaut :

- Cliquez sur le bouton **Défaut** dans l'onglet **Général**.

### Pour charger les paramètres :

- Cliquez sur le bouton **Charger** dans l'onglet **Général**, puis sélectionnez le fichier **Modèle de lot FineReader (\*.fbt)** contenant les paramètres de votre choix.

## Recherche en mode texte intégral dans les pages du lot reconnues

---

**Remarque :** pour utiliser cette option, Explorer 4.0 ou ultérieur est requis.

---

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).

---

Annexe

## **Raccourcis clavier et Glossaire**

### **Contenu du chapitre**

- Raccourcis clavier
- Glossaire

## Raccourcis clavier

Menu	Pour :	Raccourcis clavier
<b>Fichier</b>	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
<b>Edition</b>	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 <b>Texte</b> , 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
<b>Affichage</b>	Agrandir l'image dans la fenêtre Image	CTRL+MAJ+NUM+
	Réduire l'image dans la fenêtre Image	CTRL+MAJ+NUM-
	Zoom avant sur des blocs sélectionnés	CTRL+MAJ+NUM*
	Propriétés	ALT+ENTREE
<b>Lot</b>	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



<b>Menu</b>	<b>Pour :</b>	<b>Raccourcis clavier</b>
<b>Traitement</b>	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
<b>Outils</b>	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
<b>Fenêtre</b>	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
Aide	Ouvrir l'aide	F1
<b>Raccourcis généraux</b>		
	Mettre la sélection en <b>caractères gras</b>	CTRL+B
	Mettre la sélection en <i>caractères italiques</i>	CTRL+I
	Mettre la sélection en <u>caractères soulignés</u>	CTRL+U
	Passer à la cellule de tableau voisine	flèche vers la gauche, la droite, le haut et le bas

## Glossaire

### A

---

#### **Abréviation :**

forme concise d'un mot ou d'une phrase servant à représenter le tout, par exemple, MS-DOS (pour MicroSoft Disk Operating System), ONU (pour Organisation des Nations Unies), etc.

#### **Activation :**

processus par lequel l'utilisateur reçoit un code spécial de la part d'ABBYY, qui lui permet d'utiliser sa copie du logiciel en mode version complète sur un ordinateur spécifique.

**ADF** (Chargeur automatique de documents) : dispositif du scanner servant au chargement automatique des documents. Un scanner doté d'un chargeur automatique de document vous permet de numériser autant de pages que vous le souhaitez sans aucune intervention manuelle. FineReader prend également en charge la numérisation d'images multiples.

**Analyse de la présentation** (tracé des blocs) :

processus d'analyse de la mise en page et de la répartition des différentes zones de l'image en blocs, en fonction de la mise en page. Les blocs peuvent être de types différents. L'analyse de la présentation peut s'effectuer soit automatiquement dans le cadre d'une procédure simultanée de reconnaissance et d'analyse de la mise en page (en cliquant sur le bouton **2 - Lire**), soit manuellement.

#### **Apprentissage :**

création des paires "image de caractère : caractère". Consultez la section Reconnaissance avec apprentissage pour plus d'informations.

#### **Assistant Numériser&Lire :**

exécute un mode Numériser&Lire spécial. FineReader vous assiste dans la procédure de traitement du document et vous indique les actions à mener pour obtenir tel ou tel résultat.

### B

---

#### **Bloc :**

zone encadrée de l'image.

#### **Bloc actif :**

bloc sur lequel vous effectuez des opérations (par exemple suppression, changement de type, etc.). Le cadre du bloc actif est en gras et ses coins sont repérés par des "carrés".

### C

---

#### **Caractères ignorés :**

tout caractère qui n'est pas une lettre susceptible de se trouver dans un mot, par exemple caractère syllabique ou accent. Ces caractères sont ignorés au cours d'une vérification orthographique.

#### **Caractères incertains :**

caractères reconnus avec un certain degré d'incertitude. Si FineReader n'est pas tout à fait sûr d'avoir reconnu certains caractères, il les marque comme caractères incertains.

#### **Caractères interdits :**

dans certains cas, vous êtes sûr que certains caractères ne se trouveront jamais dans les textes que vous reconnaissez. Vous pouvez alors définir l'ensemble de ces caractères (appelés caractères interdits) dans les propriétés de groupe de langues. La définition de ces caractères peut améliorer de façon sensible la vitesse et la qualité de la reconnaissance. Pour indiquer le jeu de caractères interdits, cliquez sur le bouton **Avancée** dans la boîte de dialogue **Propriétés avancées de groupe de langues**. La boîte de dialogue **Propriétés avancées de groupe de langues** s'ouvre.

Indiquez le jeu de caractères interdits sur la ligne **Caractères interdits**.

**Code d'activation :**

code envoyé par ABBYY à chaque utilisateur de FineReader Professional Edition lors de la procédure d'activation. Le code d'activation est nécessaire à l'activation de FineReader sur l'ordinateur sur lequel l'ID d'installation a été généré.

---

**E****Effets de police :**

variations autour du dessin d'une police (gras, italique, souligné, barré, exposant, indice, petites majuscules).

---

**F****Fichier d'activation :**

fichier envoyé par ABBYY à chaque utilisateur de FineReader Corporate Edition lors de la procédure d'activation. Il contient les informations nécessaires à l'activation du logiciel sur le serveur ou sur un ordinateur autonome le cas échéant. Depuis le serveur, le produit est activé sur les stations de travail.

**Forme de base :**

forme sous laquelle les mots apparaissent dans le dictionnaire.

---

**G****Gabarit :**

ensemble de paires "image de caractère : caractère" créé pendant l'apprentissage de gabarit. Un gabarit est utilisé comme source d'informations supplémentaires au cours de la reconnaissance.

**Gestionnaire :**

programme qui contrôle un périphérique (par exemple un scanner, un écran, etc.).

---

**I****Image inversée :**

image dont les caractères sont blancs sur fond noir.

---

**L****License Manager :**

utilitaire utilisé pour la gestion des licences de FineReader et l'activation d'ABBYY FineReader 7.0 Corporate Edition.

**Ligature :**

combinaison de deux ou plusieurs caractères "collés", par exemple, fi, fl, ffi etc. Ces caractères sont difficiles à distinguer car ils sont généralement "collés" à l'impression. Il est plus facile de les traiter comme un caractère composé "unique".

**Lot :**

dossier contenant des fichiers d'image, des fichiers de texte reconnus ainsi que des fichiers d'informations FineReader. Un lot peut contenir jusqu'à 9 999 pages. 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.

**Luminosité :**

paramètre de numérisation permettant de régler le contraste entre les zones noires et blanches de l'image. Le réglage correct de la luminosité améliore la qualité de la reconnaissance.

## M

### **Mise en page :**

elle comprend : la façon dont le texte, les tableaux et les images sont disposés sur une page, la façon dont le texte est découpé en paragraphes, la police et la taille de la police du texte, le nombre de colonnes de texte, la couleur des caractères et du fond, l'orientation du texte.

### **Modèle de bloc :**

disposition particulière de blocs qui peut être utilisée pour reconnaître les pages de présentation identique. Il est possible d'enregistrer un modèle de bloc dans un fichier spécial.

### **Mot composé :**

mot formé de deux ou plusieurs unités lexicales (sens général) ; mot introuvable dans le dictionnaire, mais pouvant être formé de mots du dictionnaire (sens FineReader).

### **Mots incertains :**

mots contenant un ou plusieurs caractères incertains.

## N

### **Nettoyer l'image :**

suppression de l'excès de petits points noirs de l'image.

### **Numériser&Lire :**

bouton principal de FineReader. Cliquez sur ce bouton pour que FineReader numérise et reconnaisse votre(vos) image(s).

## O

### **Ouvrir&Lire :**

traite un fichier d'image : ouvre le fichier, analyse la présentation de la page et reconnaît celle-ci.

## P

### **Page de code :**

tableau définissant les relations entre les codes de caractère et les caractères eux-mêmes. FineReader affiche la page de code sous la forme d'un jeu de caractères dans lequel vous pouvez choisir.

### **Paradigme :**

ensemble des formes grammaticales d'un mot.

### **Police à espacement fixe :**

tous les caractères d'une police à espacement fixe possèdent la même largeur. Courier New est un exemple de police à espacement fixe. Sélectionnez l'option **Machine à écrire** du groupe **Type d'impression** (onglet

**Reconnaissance**) pour augmenter la qualité de la reconnaissance des documents définie pour les polices à espacement fixe.

### **ppp (points par pouce) :**

unité de mesure de la résolution, c'est-à-dire du nombre de points par pouce.

## R

### **Reconnaissance en arrière-plan :**

mode spécial de reconnaissance. Vous pouvez modifier et enregistrer les pages déjà reconnues pendant que FineReader effectue la reconnaissance d'autres pages.

### **Réglage automatique de la luminosité :**

réglage de la luminosité effectué automatiquement par le scanner ou par FineReader. Pendant le réglage automatique, la luminosité est définie séparément pour chaque zone de l'image.

### **Répertoire de veille :**

dossier spécial contrôlé par FineReader. Toute nouvelle image ajoutée à ce dossier est automatiquement ouverte dans la fenêtre de FineReader .

**Résolution :**

paramètre de numérisation qui détermine le nombre de ppp utilisés lors de la numérisation. Une résolution de 300 ppp convient aux textes imprimés dans des polices de corps 10 pt ou supérieur, une résolution de 400 à 600 ppp est préférable pour les textes imprimés dans des polices de plus petite taille (9 pt ou moins).

---

**S****Scanner :**

périphérique permettant de récupérer des images dans un ordinateur.

**Système multi-polices :**

système de reconnaissance optique qui reconnaît les jeux de caractères de toutes les polices et de toutes les tailles sans apprentissage.

---

**T****Tiret optionnel :**

tiret (¬) indiquant où précisément doit être coupé un mot ou une combinaison de mots à la fin d'une ligne (par exemple "autotracté" sera divisé en "auto-tracté"). FineReader remplace tous les tirets trouvés dans les mots du dictionnaire par des tirets optionnels.

**TWAIN, boîte de dialogue TWAIN :**

boîte de dialogue du scanner.

**Type d'image :**

paramètre de numérisation qui détermine si une image doit être numérisée en noir et blanc, en niveaux de gris ou en couleur.

**Type d'impression du texte source :**

paramètre indiquant comment a été imprimé le texte source (sur une imprimante laser ou équivalente, sur une imprimante matricielle en mode brouillon ou sur une machine à écrire). Choisissez le mode **Auto** pour les textes imprimés sur une imprimante laser, le mode **Machine à écrire** pour les textes imprimés sur une machine à écrire et le

mode **Imprimante matricielle** pour les textes imprimés sur une imprimante matricielle en mode brouillon.

---

**U****Unicode :**

norme développée par le Consortium Unicode (Unicode, Inc.). Cette norme est un système de codage international sur 16 bits pour le traitement des textes écrits dans les principales langues du monde. Cette norme peut s'étendre facilement. La norme Unicode détermine le codage des caractères ainsi que les propriétés et les procédures utilisées dans le traitement des textes écrits dans une langue donnée.