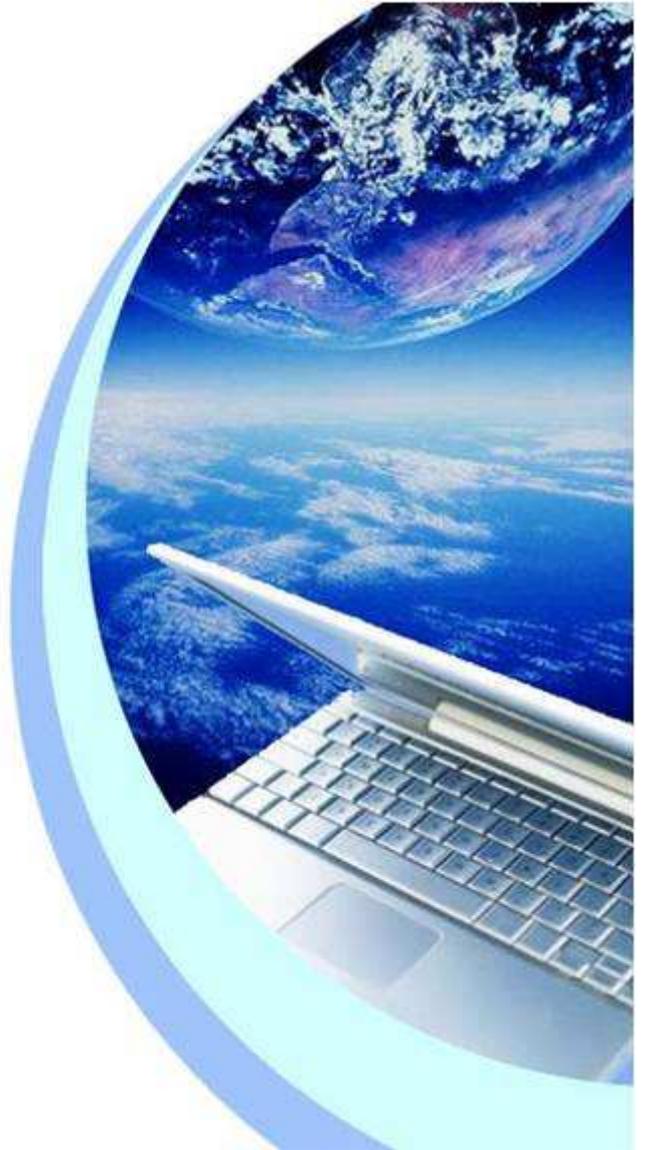




DashesPro™ for Macintosh

Version 7.3



Contents

Introduction	1
Installation	1
Quick reference for QuarkXPress	1
Quick reference for Microsoft Word	1
Hyphenation of files – QuarkXPress tags, RTF, plain text	2
DashesPro™ Menus	2
APPENDIX A – About Hyphenation	6
APPENDIX B – Stylistics	8
APPENDIX C – DashesPro™ Suite: Events and classes for DashesPro™ application	9

Introduction

The DashesPro™ is designed for hyphenation of texts. It works by inserting discretionary hyphens through out the text. It can only be used in conjunction with programs which already support discretionary hyphens. They include Quark XPress and Microsoft Word among others. It can also be used to hyphenate ordinary text files and to unify your hyphenation through applications and platforms.

You can use it to improve hyphenation in your current program and to hyphenate texts in languages not available in your current program.

Installation

1. Launch "Install DashesPro™" application and follow the installation instructions.

Quick reference for QuarkXPress

- 1 Launch DashesPro™.
- 2 Launch QuarkXPress
- 3 Open the file you want to hyphenate
- 4 Make sure that QuarkXPress own automatic hyphenation is turned off
- 5 To hyphenate a selection of text in a story, select the text you want to hyphenate and then choose "Hyphenate QXP" menu command from XPress scripts menu or click on the "Hyphenate QXP" docklet.
- 6 To hyphenate an entire story, select all of the text in the story or simply place an insertion point anywhere in the story text. Then choose "Hyphenate QXP" menu command from XPress scripts menu or click on the "Hyphenate QXP" docklet.
- 7 To hyphenate every story in a Quark document make sure that you don't have either an insertion point or any kind of selection. Then choose "Hyphenate QXP" menu command from XPress scripts menu or click on the "Hyphenate QXP" docklet. To avoid hyphenating taglines and other short display text DashesPro™ will ask you for the minimal length the story needs to be in order to be hyphenated.
- 8 Removing discretionary hyphens from text works exactly the same way except "Remove Hyphens QXP" script is used.

Quick reference for Microsoft Word

DashesPro™ supports hyphenation in the MS Word via clipboard. To hyphenate text:

- 1 Launch DashesPro™
- 2 Launch MS Word
- 3 Open the file you want to hyphenate
- 4 Make sure that MS Word own automatic hyphenation is turned off
- 5 Select the text you want to hyphenate and copy it to clipboard
- 6 Switch to DashesPro™ application
- 7 Select "Hyphenate Clipboard" from the DashesPro™ Actions menu
- 8 Switch back to MS Word and paste the hyphenated text

Please note that in order to retain formatting you have to check "Include formatted text in Clipboard" option in the General section of MS Word preferences.

Hyphenation of files – QuarkXPress tags, RTF, plain text

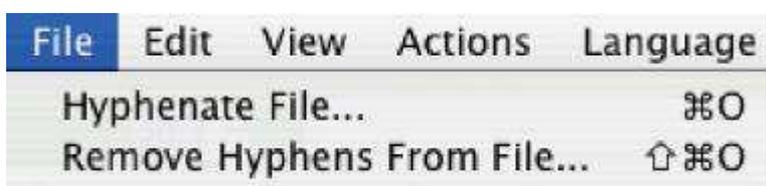
To hyphenate a file, select “Hyphenate File...” under File sub-menu of DashesPro™ menu (see menu structure below). In the dialog that opens choose the type of file you are looking to hyphenate from the popup menu: Quark tag, RTF or text. Select the file you want to hyphenate. DashesPro™ will insert discretionary hyphens throughout the file in accordance with the parameters you have set in the Language Settings and Exceptions dialog boxes. DashesPro™ supports text, RTF, QuarkXPress tag. In RTF, QuarkXPress tag DashesPro™ will keep the text formatting.

DashesPro™ Menus

DashesPro™ menu set consists of DashesPro™ menu, File menu, Edit menu, View menu, Actions menu, Language menu. You can see the description of each point below.

File menu

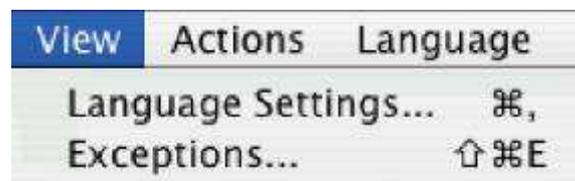
“Hyphenate File...” menu selection hyphenates all text in a file. Selecting this menu item opens “Open File” dialog. In the “Open File” dialog locate the file you are looking to hyphenate, choose from the popup menu the correct file type and click Open button. The file will be hyphenated in accordance to the current language selection and other preferences.



“Remove Hyphens From File...” menu selection removes discretionary hyphens from a file. It works the same way as the “Hyphenate File...” menu.

View menu

“Language settings...” menu selection opens preferences dialog box. This dialog has two areas. The first area, “Language Specification”, manages language specific settings. The second area, “General”, deals with global DashesPro™ properties.



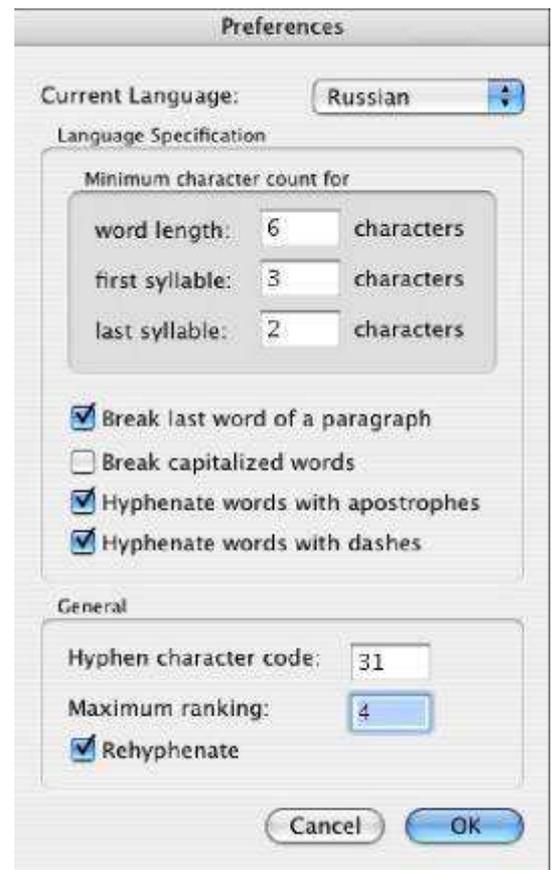
In Language Specifications area you can change the following options:

- Choose the current language by selecting it from the “Current Language” popup menu. Note that the changing language selection in this dialog box changes the current language everywhere else in the program.
- Choose the minimum character count for:
- ‘word length’ - sets the minimum number of characters required for the word to be hyphenated. The program skips the shorter words.
- ‘first syllable’ - sets the minimum number of main characters required before the first hyphen in a word;
- ‘last syllable’ - sets the minimum number of characters required after the last hyphen in a word;
- ‘Break last word of a paragraph’ - if checked, DashesPro™ hyphenates the last words in a paragraph;
- ‘Break capitalized words’ - if checked, DashesPro™ hyphenates capitalized words. Sometimes it is preferable to uncheck this option to prevent hyphenation of proper names;

- ‘Hyphenate words with apostrophes’ - if checked, the words containing apostrophes are hyphenated;
- ‘Hyphenate words with dashes’ -if checked, the words containing dashes, such as mother-in-law, are hyphenated.

In General settings you can change the following options:

- ‘Hyphen Character’ – allow you to enter a decimal ASCII code for discretionary hyphens, that is used when hyphenating plain text files. The supported ASCII code for the discretionary hyphen varies from application to application. Most applications however use ASCII code 31;
- ‘Maximum ranking’ – sets DashesPro™ acceptable hyphenation quality. All hyphenations in DashesPro™ are ranked. Rank shows the quality of each hyphen. You can vary the maximum ranking from 0 to 4, with 0 being the best and 4 the worst;
- ‘Rehyphenate’ – if unchecked, DashesPro™ leaves the words already containing discretionary hyphens untouched.



Note: The rankings are defined as follows:

- 0 – word boundary hyphen as in: white-house
- 1 – prefix or suffix boundary hyphens as in: under-estimate, relation-ship
- 2 – hyphen in the middle of a root as in: estab-lishment, mas-ter
- 3 – hyphen in the middle of a prefix or suffix as in: hy-peractive, un-derestimate
- 4 – hyphen one letter from a word or affix boundary as in: paraphenali-a, o-verreact
pol-ymorphic

“Exceptions...” menu selection opens a dialog that manages hyphenation exceptions that is words that DashesPro™ fails to hyphenate properly. Exception dictionaries are language specific. Before making any changes in the Exceptions dictionary you have to make sure the correct language is selected in the “Current Language” popup menu. The “Font” popup menu allows you to select the most appropriate font for displaying exception words in a particular language.

If you are unhappy with the way Dashes hyphenates a word, you can enter it in the exception dictionary. You can use the exception dictionary either to change the hyphens that Dashes assigns, or to keep Dashes from hyphenating a word altogether.



You may want to do this with certain proper names, for example.

- To add a new word, type it in the edit box and then click 'Add' button.
- To remove a word or words from the list select them and then click "Remove" button.
- You can export the words in the exception list to a text file by clicking on the "Export..." button.

When you add a word to the dictionary, all the required hyphenation points have to be specified. You can optionally specify the hyphenation rank for each hyphen by following it with a number from 0 to 4, for example `bake-0ware`, `va-2ri-1able`, `pol-4y-1mor-2phic`. If you don't specify the hyphenation ranking DashesPro™ assigns the rank of 0.

When you are adding a word to the exception dictionary DashesPro™ will compare its hyphenation against the hyphenation you provide. If the hyphenation you provide agrees with the way DashesPro™ hyphenates the word DashesPro™ issues a dialog letting you know that the word does not need to be in the exception dictionary and does not insert it into the list. This is done to save memory and improve performance.

By default during this process DashesPro™ does not compare hyphen ranks. That is `hel-lo` compares equal to `hel-1lo`, `hel-2lo`, `hel-3lo` or `hel-4lo`. If you want DashesPro™ to also compare the rankings hold down the option key when clicking on the "Add" button.

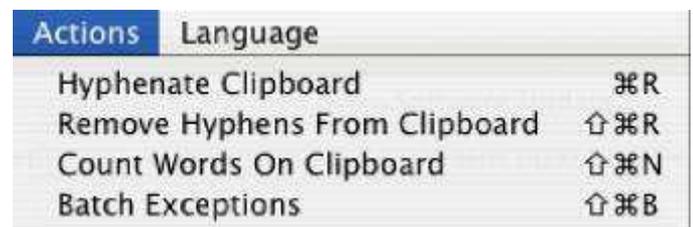
The exception dictionary is not case sensitive. In other words, if you enter a capitalized word

into the exception dictionary, i.e. "Parker", DashesPro™ will recognize the word as an exception when typed in lower-case form in the text, i.e. "parker".

"Show Dashes Window" makes visible the DashesPro™ About window. It can be used to simplify switching between DashesPro™ and other applications.

Actions menu

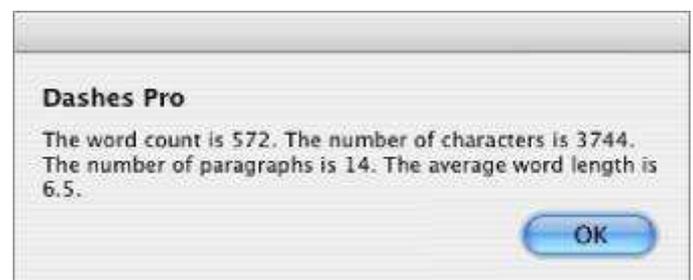
Hyphenate Clipboard menu command hyphenates the text on the Clipboard. DashesPro™ tries to preserve as much formatting as possible. It does this by supporting RTF and styled text formats for text it finds on the clipboard. If your application is able to place RTF on the clipboard then all of the formatting is preserved.



Remove Hyphens From Clipboard menu command removes discretionary hyphens from the text on the Clipboard. Just like the "Hyphenate Clipboard" command it tries to preserve as much formatting as possible.

Count Words on Clipboard menu command shows the number of words on the Clipboard.

Note: a word is a combination of symbols, not separated by white spaces. All the other symbols (such as dots, commas etc.) are counted as the part of the word.



Batch Exceptions. If you already have a list of words which you would like hyphenated in a particular way, and you would like to enter these into the exception dictionary all at once do the following. Open the file of hyphenated words. It must have the following format:

William-son
un-de-ni-abil-ity

In other words, each word you enter must be followed by a carriage return. The hyphens you want entered should appear as hard hyphens in the text. You may also optionally enter hyphen ranks by inserting the appropriate number after the hyphen such as:

William-1son
un-1de-2ni-1abil-1ity

Select all the words you wish to have entered into the exception dictionary and copy them to the clipboard. Switch to DashesPro™ and select “Batch Exceptions” menu item under Actions menu. Your words will be entered into the exception dictionary as you marked them.

The “Batch Exceptions” command only enters the word into the exception dictionary if Dashes hyphenates differently from what you have specified in your word list. For large word lists, this cuts down significantly on the size of the resulting exception dictionary and increases the performance without affecting the accuracy.

By default “Batch Exceptions” command does not compare hyphen ranks. That is, if the difference between the word in your hyphenated list and the hyphenation that DashesPro™ assigns is only in the hyphen ranks, then the DashesPro™ will not enter the word into the exception dictionary. For example, Dashes might hyphenate syn-1chro-2nic-1i-3ty and your batch exception list might contain syn-1chro-2nic-2i-2ty. By default DashesPro™ will not add syn-1chro-2nic-2i-2ty into the exceptions list.

If you wish to override this default and enter such words in the exception dictionary, hold down the option key as you select “Batch Exceptions”. This will allow the word to be entered into the Exception Dictionary even though only the hyphen ranks differ.

Language menu

The **Language** menu allows you to set the current hyphenation language. It also indicates which languages are currently purchased. They are displayed in bold. Languages displayed in plain face are available in the demo mode.



APPENDIX A – About Hyphenation

People have intuitions about where it makes sense to break a word. The principles behind these intuitions can be defined to a large extent, and dictionaries base their hyphenation suggestions on them. The hyphenations suggested by different dictionaries of the same language vary. The variations depend to some extent on the dialect of the dictionary, but most of the variations are due to ambiguities in our

intuitions about where to break a word. We can see this in the fact that all American English dictionaries hyphenate idiosyncratically, as do all British English dictionaries.

Principles of Word Breaking

There are two principles of word breaking that no dictionary in any language that we know of violates.

That is:

- B. In a compound word, break the compound between the component words. In other words, breaks like 'an-teater' are universally unacceptable.
- C. The unit occurring between two hyphens must be a possible syllable in that language. For example, breaks like:

co-mpatibl-e
whi-tehouse
overdr-essed

are also universally unacceptable.

In many languages, the break does not occur on the actual pronounced syllable boundary. For example, if you were to analyze most people's pronunciation of the word 'anteater', we would find that most people pronounce the syllables as 'an/tea/ter'. However, the hyphenation accepted by most dictionaries, namely 'ant-eat-er' still conforms to principle (B), because 'ant', 'eat' and 'er' are all legal pronounceable syllables in English.

In the example above, we can see that there are two forces affecting our intuitions about where to break a word. One is that we want to break words on syllable boundaries. The other is that we want to break words between what linguists call morphemes. Morphemes are meaning bearing elements of the word, such as each of the component words of a compound (prefixes, suffixes and roots). A conflict arises when the morpheme boundary is not the same as the syllable boundary. This conflict is solved differently in different languages and by different dictionaries within the same language. This conflict is the main thing that accounts for the discrepancies among and within dictionaries.

Let us then add a third principle, which is applied consistently in no language that we know of, but which applies sometimes in most dictionaries for most languages.

- D. Break words on prefix and suffix boundaries.

Portuguese is an exception to this. In Portuguese, you never hyphenate on prefix or suffix boundaries unless they coincide with syllable boundaries. The English solution to this conflict of syllable and morpheme boundaries is unnecessarily complex, but certain rules have been established by convention, and must be followed for English.

The last universal rule is:

- E. Avoid stylistically poor hyphens like:
 1. Hyphens too close to a word boundary, such as:

a-mygdaloid
e-vocative
an-tidisestablishmentarianism

2. Hyphens that fall in the middle of a prefix or suffix:

pol-yethylene
Armeni-an

Types of Hyphenation Programs

One basic approach to hyphenation is dictionary-based, and the other is algorithm-based. Any algorithm can be made to match any dictionary by adding an exception dictionary to the algorithm which contains all the words that the dictionary and the algorithm hyphenate differently. Thus the real difference between algorithm and dictionary-based hyphenation lies in what they do with words that do not appear in the dictionary. A dictionary-based program does not hyphenate any words that do not appear in the dictionary. The advantage of this is that you are guaranteed no wrong or unreasonable hyphens. The disadvantage is that even in English, a relatively large percentage of the words that actually appear in text are not in the dictionary.

In languages like German and Scandinavian, this is even more crucial, because of word compounding. The compounds in these languages are the longest words, and therefore those you would most like to hyphenate. However since they are freely generated, no dictionary can contain them all.

The two types of algorithms available are what we call statistically based algorithms and linguistically based algorithms. Statistically based algorithms treat hyphenation purely as an information/dictionary compaction problem. The intention is to use various methods to create the most compact data representation that will enable the reproduction of hyphens present in the original list. Linguistically based algorithms are those, which make extensive use of information we have about how words and syllables are structured. Both methods have advantages. However, we do not know of a statistically based algorithm, which is capable of hyphenating Germanic languages with a high degree of accuracy.

It is trivial to write an algorithm that only finds 40%-50% of the possible hyphens accurately. Such algorithms are available at essentially no cost, and they work very fast. However, they tend to put in hyphens at the ends of words, not in the middle where they are most useful. It is easy to imagine how one might write an algorithm that puts in 1% of the possible hyphens with 100% accuracy. You tell the computer when it finds 'ment', 'ments', 'tion', or 'tions' at the end of a word to put a hyphen in front of it. However, it gets exponentially more difficult to maintain a high degree of accuracy as you put in a higher percentage of hyphens. In fact it is impossible for either dictionaries or algorithms to put in 100% of the possible hyphens correctly, because there are a number of words that are spelled the same, but hyphenated differently, such as 'rec-ord' and 're-cord'.

Dashes is a linguistically-based algorithm. It puts in over 99% of the possible hyphens with over 99.9% accuracy in all of the languages offered.

APPENDIX B – Stylistics

In general, it is recommended that words fewer than 5 characters in length not be hyphenated, and that hyphens not be fewer than 2 letters from the end of the word.

In no language are hyphens one letter from the end of a word stylistically acceptable. In German and Finnish, they are simply considered invalid. In other languages, they are of low stylistic quality. These are hyphens with ranking 4 in Dashes, and they should for almost all purposes be removed.

The following rules hold for rewriting characters:

- a. If a hyphen falls between a 'c' and a 'k' in traditional German, change the 'c' to a 'k':

backen -> bak-ken

- b. In traditional German, as well as in Dutch, Swedish and Norwegian, if three like letters appear in a row due to compounding, one of them is deleted. This consonant is restored when a hyphen is put in.

Schiffahrt → Schiff-fahrt

c. Dutch has a number of rewriting rules:

- 1 Before diminutives ending in '-pje' and '-tje', if the root ends on 'aa', 'oo', or 'uu', a vowel is deleted: autootje → auto-tje, chaltje → chalet-tje, skietje → ski-tje, (cafeetje → cafй-tje)
- 2 The diaeresis is lost on vowels after hyphenation: зеелн → zee-en

In the DashesPro™, the rewriting situations in (a) and (b) are left unhyphenated. In the Dutch examples in (c), the DashesPro™ hyphenates, but does not rewrite.

It is generally not recommended to hyphenate English words with an apostrophe. It is fine to hyphenate French words with an apostrophe.

It is generally not recommended in any language to hyphenate words that are already written with a hard hyphen.

In all languages it is preferable not to hyphenate proper human names. Other proper names may be hyphenated. Use the Exception Dictionary to prevent hyphenation of proper human names.

We have encountered a number of differences of opinion among typographers and other specialists concerning the hyphenation of French. People differ primarily on whether to hyphenate words such as 'instruire' as 'in-struire' or 'ins-truire'. The Canadians in general prefer the latter, but those in France seem to be divided, and feel very strongly one way or the other. For this reason, we have two versions of French, which differ along these lines. Some examples:

French1 French2

in-1strui-3re ins-2trui-3re
 cons-2trui-3re cons-2trui-3re
 conta-2mi-2ner con-1ta-2mi-2ner
 trans-1ac-2tion tran-2sac-2tion

APPENDIX C – DashesPro™ Suite: Events and classes for DashesPro™ application

DashesPro™ supports AppleEvents. It is both scriptable and recordable application. DashesPro™ event suite is described below:

Class application: DashesPro™ application

Properties:

hyphen character integer — ASCII code or character to use as discretionary hyphen **maximum ranking integer** — poorest hyphen quality to use; 0 - best, 4 - worst **rehyphenate boolean** — replace already inserted discretionary hyphens? **word length integer** — number of characters of the shortest word to be hyphenated **first syllable integer** — minimum number of letters before the first hyphen **last syllable integer** — minimum number of letters after the last hyphen **break last word boolean** — hyphenate the last word of a paragraph? **break capitalized word boolean** — hyphenate capitalized words? **language font international text** — font to be used with current hyphenation language **current language international text** — the current hyphenation language **window visible boolean [r/o]** — is DashesPro™ window visible? **frontmost boolean [r/o]** — is DashesPro™ a frontmost application? **name international text [r/o]** — name of DashesPro™ application **version integer [r/o]** — version number of DashesPro™ application

Class text runs: Text with style runs

Properties: **body international text [r/o]** — block of text to be hyphenated **styles a list of integer [r/o]** — list

of offsets into the text at which there is a format change

Class word statistics: Results of “count words” event

Properties: **words integer** [r/o] — number of words **characters**

integer [r/o] — number of characters **paragraphs integer** [r/o]

— number of paragraphs **average length real** [r/o] — average word length

add exceptions: Add words to exceptions dictionary

add exceptions international text — words to add to exceptions dictionary; when this parameter is absent

DashesPro™ adds words on clipboard [**hyphen hierarchy boolean**] — compare hyphen rank?

count words: Report number of words, characters, paragraphs and average word length

count words international text — text to count; when this parameter is absent DashesPro™

counts text on clipboard

Result: **word statistics**

get: Get application property

get reference — application property **Result:**

anything — property data

hyphenate: Hyphenate file

hyphenate alias — text file to hyphenate; when this parameter is absent DashesPro™ hyphenates

text on clipboard [**with tags TEXT/QuarkXPress/RTF**] — file format to process

hyphenate text: Hyphenate string of text

hyphenate text international text — text to hyphenate [**with tags TEXT/QuarkXPress/RTF**] — format to

process **Result:** **international text**

hyphenate text runs: Hyphenate “text runs” object

hyphenate text runs text runs — text runs to hyphenate **Result:** **text**

runs

is word delimiter: Check if the character is in word delimiter set

is word delimiter integer — ASCII code or character to check **Result:** **boolean**

list languages: Return list of all available languages

list languages **Result:** a list of

string

quit: Quit application

quit

remove exception: Remove word from exceptions dictionary

remove exception international text — word to remove from exceptions dictionary

remove hyphens: Remove hyphens from file

remove hyphens alias — text file to remove hyphens from; when this parameter is absent DashesPro™

removes hyphens from text on clipboard [**with tags TEXT/QuarkXPress/RTF**] — file format to process

remove hyphens text: Remove hyphens from string of text

remove hyphens text international text — text to remove hyphens from [**with tags**

TEXT/QuarkXPress/RTF] — format to process **Result:** **international text**

remove hyphens text runs: Remove hyphens from “text runs” object

remove hyphens text runs text runs — text runs to remove hyphens from **Result:** **text**

runs

remove word delimiter: Remove character from word delimiter set

remove word delimiter integer — ASCII code or character to remove

save properties: Make applicaton properties modifications permanent

save properties

set: Set application property

set reference — application property to **anything**

— new value