PDFSpeech User Guide

August 5, 2019

Introduction

PDFSpeech is a macOS application published by onekerato.com, available for Macs running OS X 10.7 Lion and later. At its core, PDFSpeech has the ability to read out the text content in a PDF using text-to-speech technology built into macOS. It's purpose is to **help you read more**.

How can PDFSpeech help you read more? Simply put, it reduces the effort involved in reading. Just like watching videos is "easy", reading a book using text-to-speech technology (TTS) is "easy" because you can sit back and "watch" the book unfold. Of course, you still need to pay attention, but what TTS-powered reading gives you for free is **momentum**. And momentum is the secret to achieving your goals.

To make the most of PDFSpeech, load your reading list into the app, and start TTS by **double-clicking** on a word. Momentum will do the rest.

Contents

What's New in PDFSpeech 2

Basics: How to Use PDFSpeech 5

- Choose a Text-to-Speech (TTS) Voice 6
- Bring Your Reading List into PDFSpeech 7
- Get Organized with a PDF Pack 8
- Navigate inside a PDF 9
- Quick Text-to-Speech 10

Configure PDFSpeech to Read Your Content 10

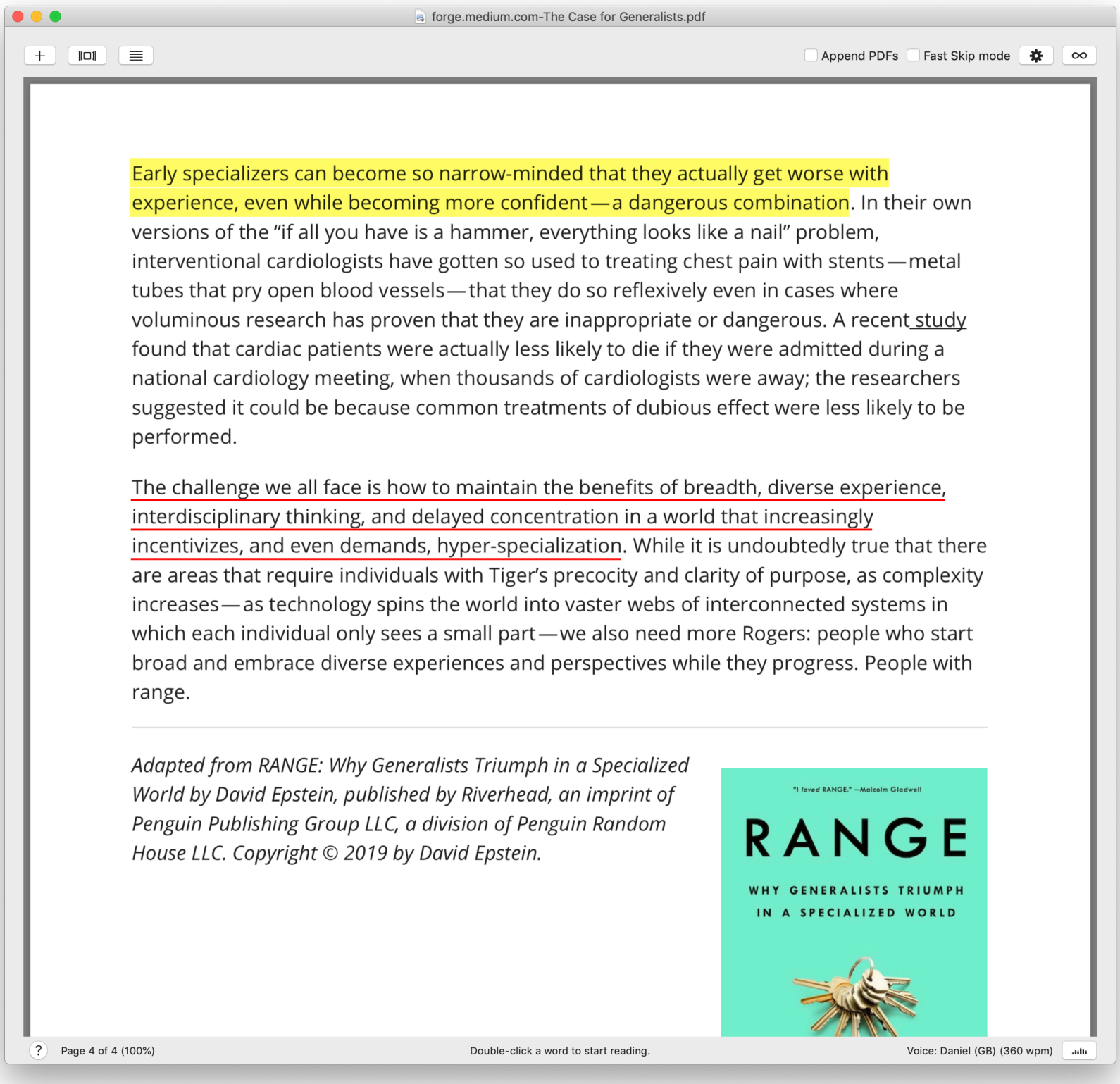
- Skip Headers and Footers 11
- Skip Fixed Width Paragraphs 11
- Fast Skip Mode through Text 11
- Customize Pronunciations 12

Resources 13

What's New in PDFSpeech

What's New in PDFSpeech 1.4

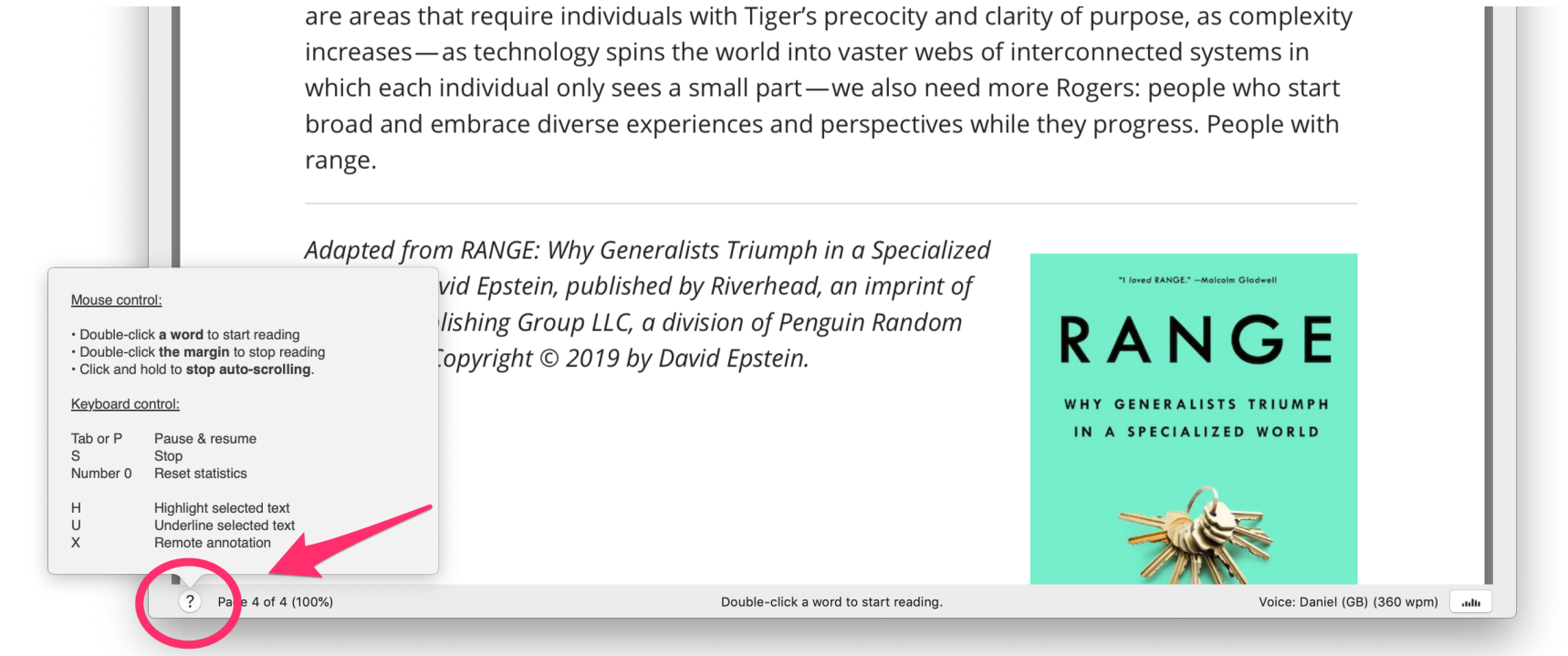
- **Shortcuts**: Use the `Tab` key to pause and resume text-to-speech. The shortcut `p` also works for pause & resume, but the `Tab` key may be easier to tap with the left hand when using the mouse with the right hand to scroll pages and double-clicking to start reading.
- **Annotations**: PDFSpeech is now an "**Editor**" of PDF files. It can now edit PDF files by adding **highlight and underline annotations**, and removing existing annotations.
    - To **add a highlight**, select text and hit `H`.
    - To **add an underline**, select text and hit `U`.
    - To **remove** a highlight or underline annotation, select **all or even a part** of the annotated text, and hit `X`.



The highlighting and underlining features are easy to employ **while using TTS** to read a book. With TTS, sentences are automatically highlighted as they are read, and using the `H` or `U` shortcut key **without having to**

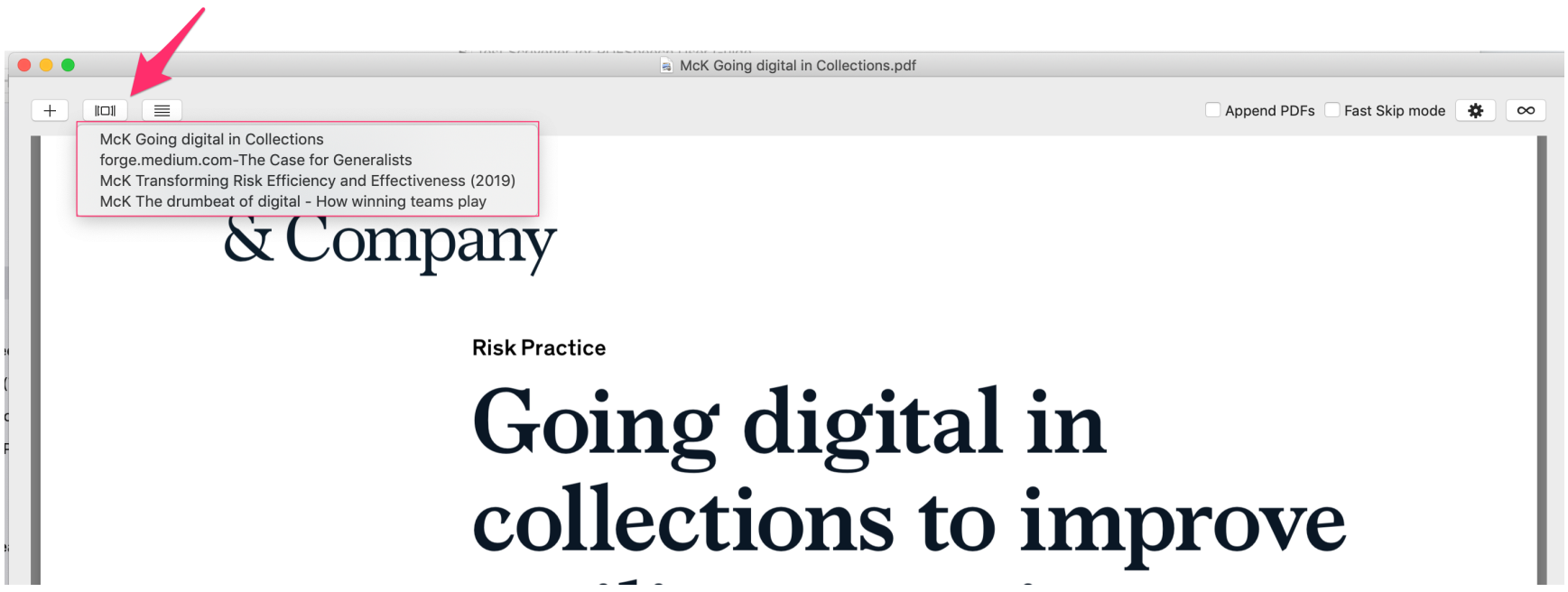
pause TTS makes it fluid to annotate text. The PDF file including annotations is saved automatically.

A handy reference of the **available keyboard shortcuts** is available from the "question" icon button on the bottom left of the PDFSpeech window.

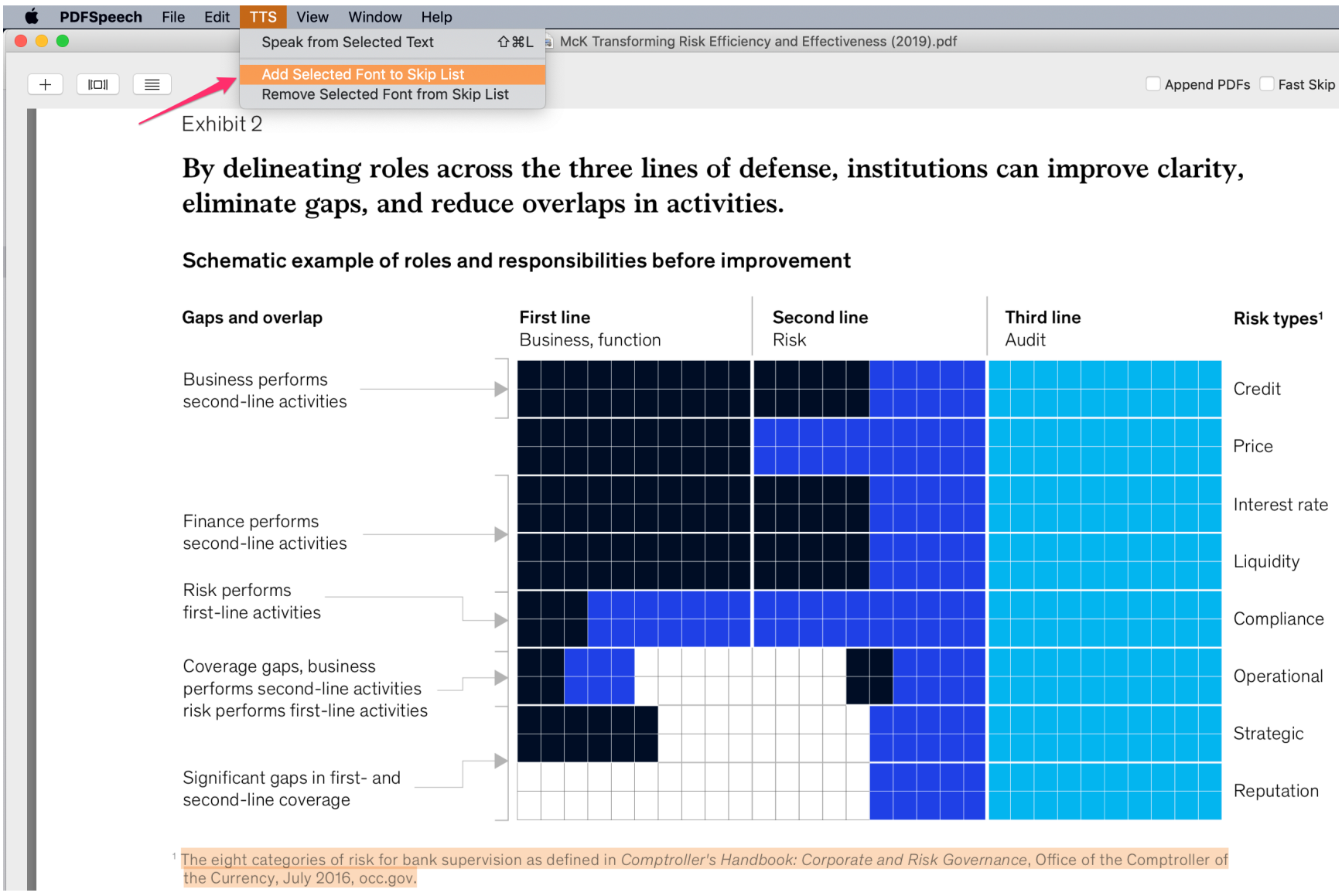


What's New in PDFSpeech 1.3

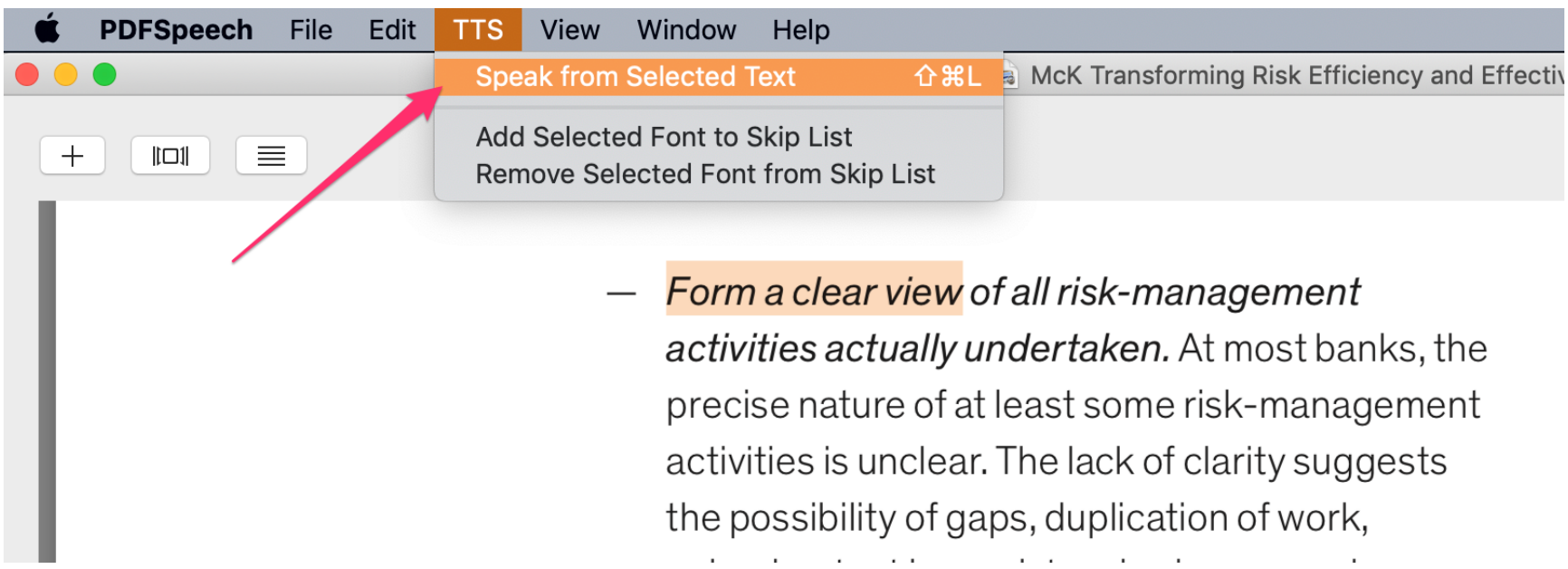
- Load **your PDF reading list** into PDFSpeech a little bit easier by setting up a "**Watched Folder**". A dedicated "**Recents**" button on the top-left shows a **list of the five most recently modified PDFs** in the Watched Folder. By default, the user's **Downloads** folder is the Watched Folder. The folder is only "watched" when you click on the Recents button (shortcut `CMD R`), and thus may take a couple of seconds to populate the list if the folder has many files. Use File > Set Watched Folder... to customize.



- **Skip sentences and paragraphs based on specific fonts**. Select text on the page that you want to skip, then use the new TTS menu > Add Selected Font to Skip List command. Similarly, if you change your mind, select text and use the TTS > Remove Selected Font to Skip List command. PDFSpeech will remember these instructions when you come back to the PDF.

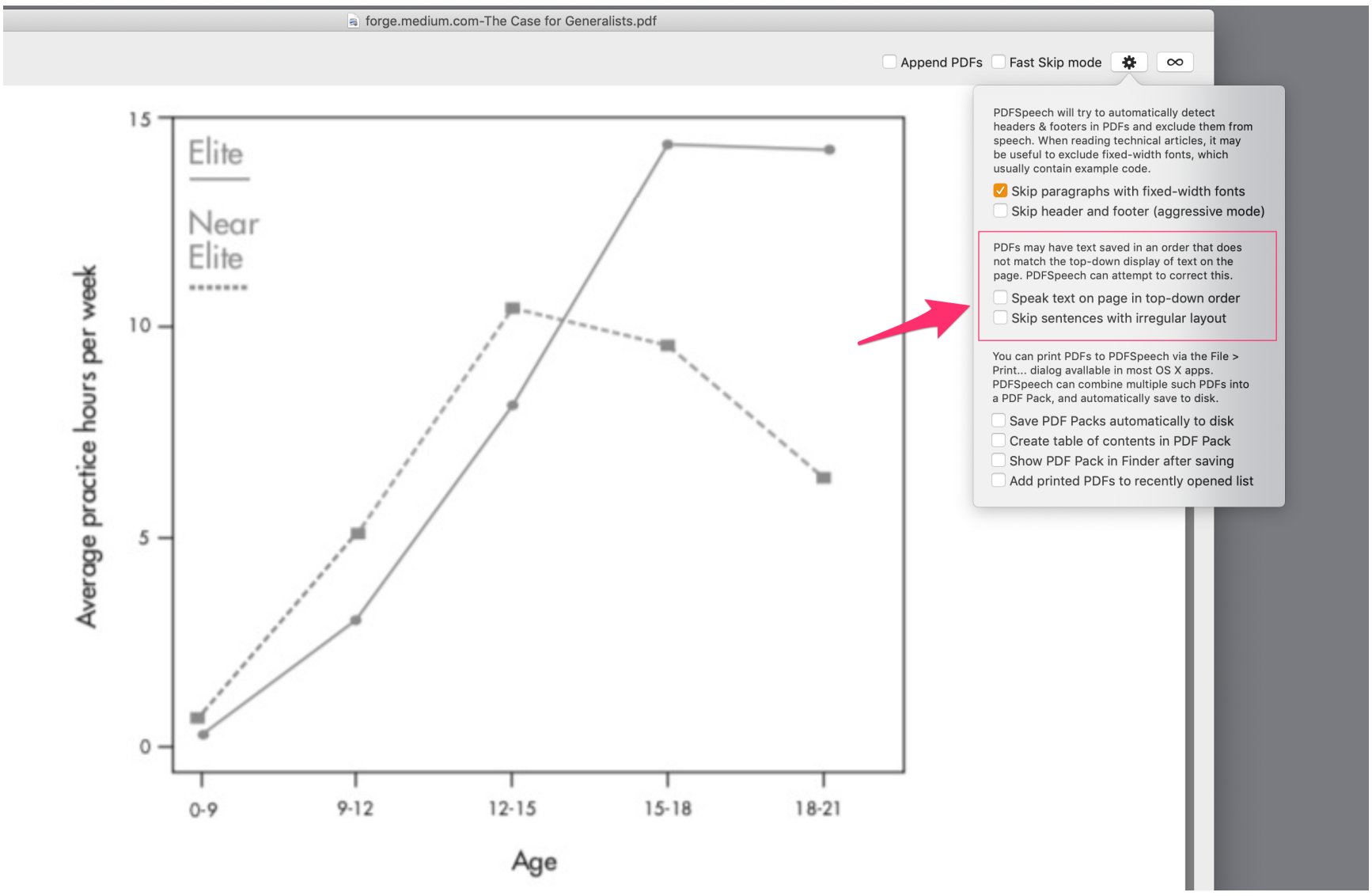


- As always, you can double-click on any word to start reading from that point. But now, you can also select text, and use the command TTS > Speak from Selected Text (shortcut `SHFT CMD L`).



- A **PopClip** extension for PDFSpeech is available at onekerato.com/pdfspeech.html (Help > Go to PDFSpeech website) that allows you to select text in the PDF and invoke the "Speak from Selected Text" using your mouse.

- The **Preferences popover** now provides two additional options to control the text-to-speech engine. Often, a PDF stores text inside a PDF in an order that doesn't match the expected left-to-right, top-to-bottom layout of text on the page. As such, when the PDF engine extracts the text from the PDF, it can be read out in an incorrect order. The new options in PDFSpeech Preferences attempt a workaround by sequencing text to match the visual layout.



- Bug fixed: PDFSpeech's text-to-speech engine no longer gets stuck if a sentence has no pronounceable words.
- The **PDFSpeech User Guide** has been updated with a Contents page, a What's New section, and internal hyperlinks.

What's New in PDFSpeech 1.2

- Ability to select the **text-to-speech voice** and **speaking rate** within the app. Or you can continue to use the system default TTS voice and rate.
- **PDFSpeech** now integrates with **PDFoo**, the app that enables precise linking into any page or point in a PDF. The generated PDFoo URLs open in the PDFoo app, which is included in the PDF Believer Pack available on the Mac App Store.

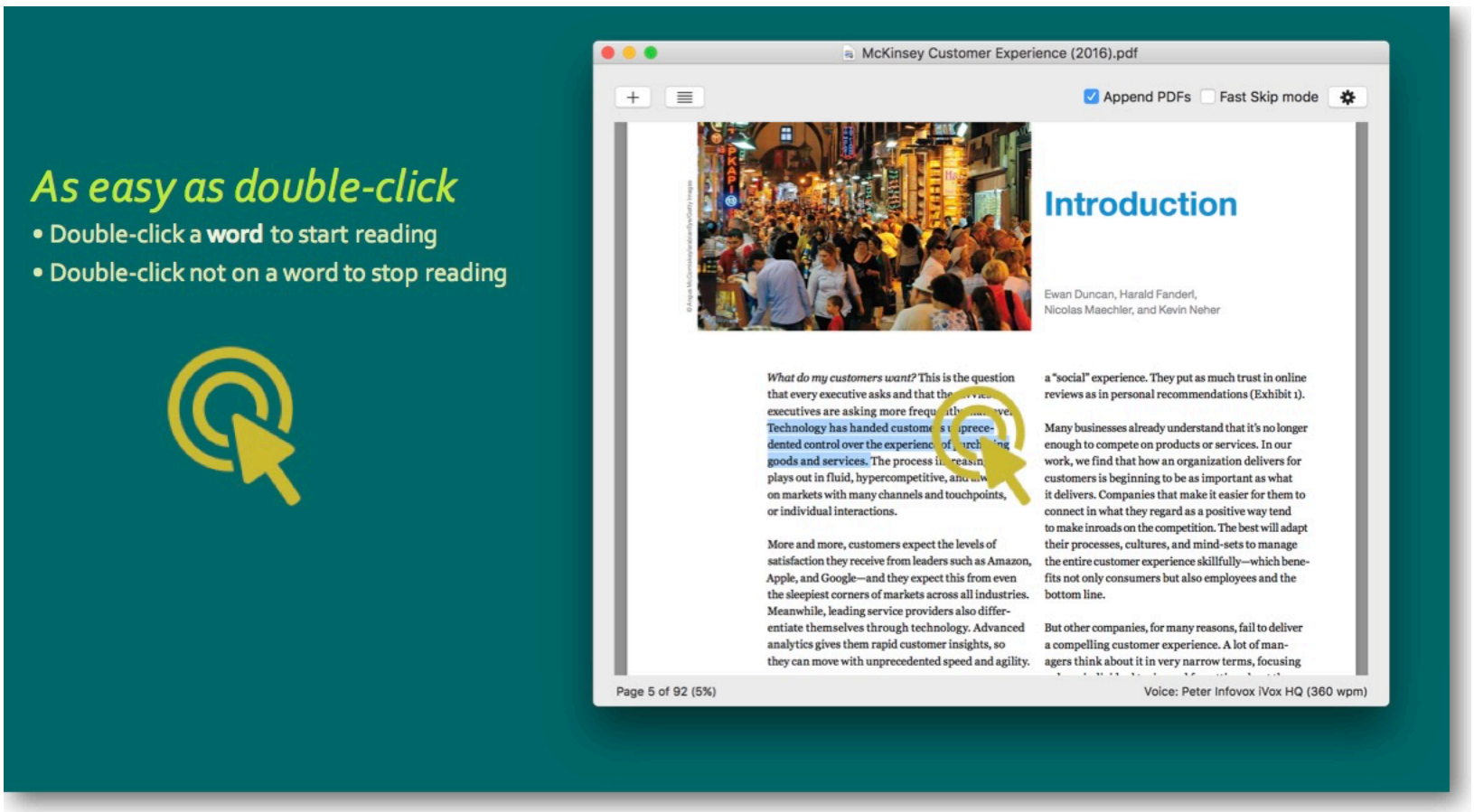
The **PDF Believer Pack** bundles five apps, and is more cost effective than purchasing PDFoo and PDFSpeech separately. However, the bundle itself is only visible on the Mac App Store on macOS Mojave 10.14 and later, though after purchase, the apps can be downloaded and used on prior macOS versions as usual.

Basics: How to Use PDFSpeech

To get started with PDFSpeech, open a PDF and **double-click on any word** in the PDF. Try it now with this user guide. PDFSpeech will begin reading the sentences from that point forward through the document.

To **pause**, hit the `p` key. To resume, hit `p` again.

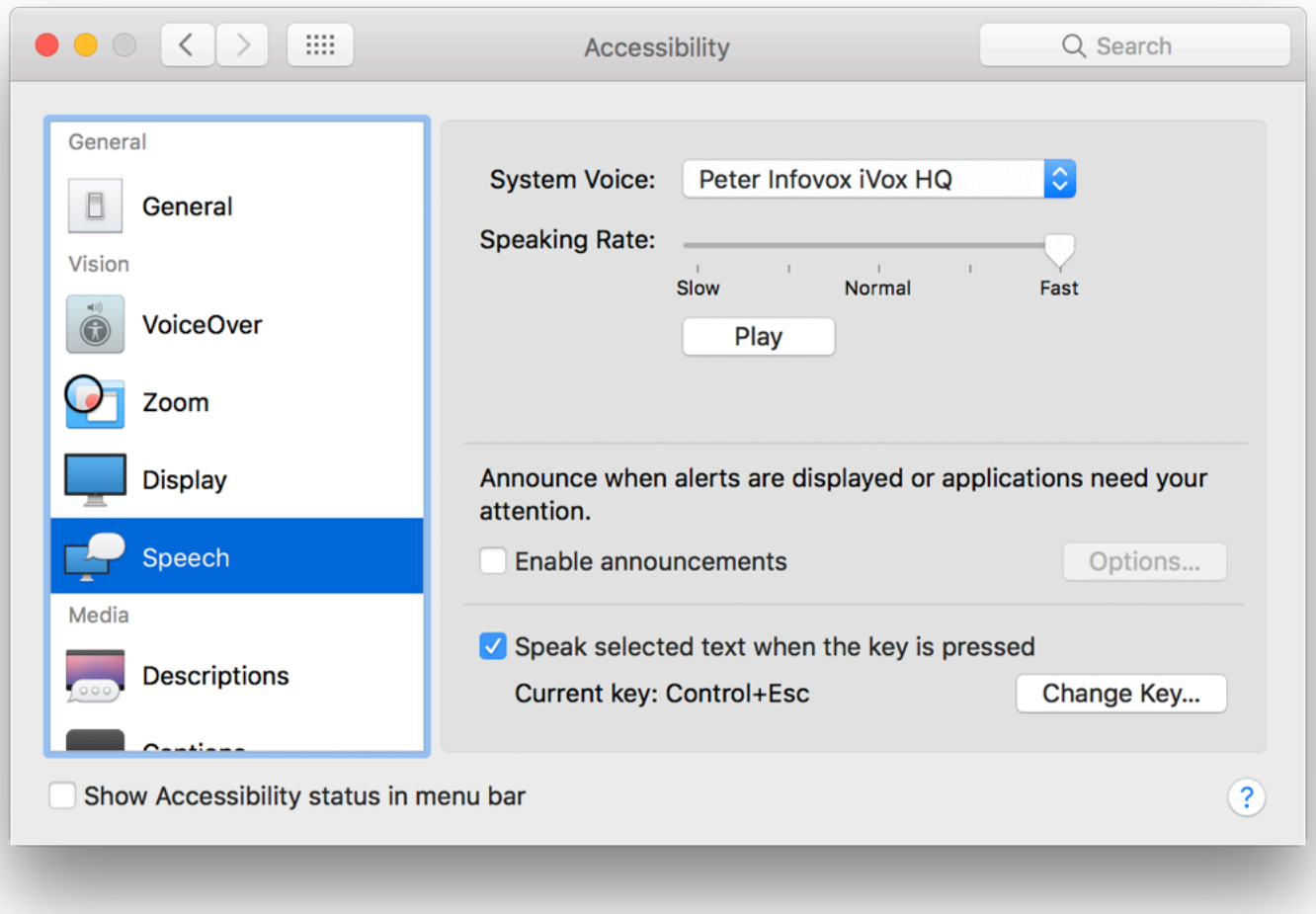
To **stop**, simply double-click anywhere in the margins, or hit the s key.



Choose a Text-to-Speech (TTS) Voice

PDFSpeech uses the text-to-speech settings configured in **macOS System Preferences**. On macOS Sierra, the speech settings are located in the **Accessibility** pane. You can specify the system voice and the speaking rate. If you're wearing a headset, you will be able to listen and follow along even when the speaking rate slider is close to or at the maximum, which corresponds to around **340-360 WPM** (words per minute.) High TTS speeds also help you maintain focus for longer (win-win). At 350 WPM, you can read one long chapter of a business or technical book (say 10,000 words) in 30 mins.

macOS offers several high-quality voices (for free!) that can be used for TTS in any app on your Mac. Two system voices with good clarity for reading technical content, in my opinion, are **Daniel** (UK) and **Samantha** (US). You can also purchase pro-quality Acapela voices for around $15 each from AssistiveWare Infovox iVox, licensed for use on your Mac.



Bring Your Reading List into PDFSpeech

PDFSpeech only works with PDFs, and that may seem like a huge limitation for users who want to bring their diverse reading list into this app. However, there are a few reasons why **PDFSpeech only works with PDFs**:

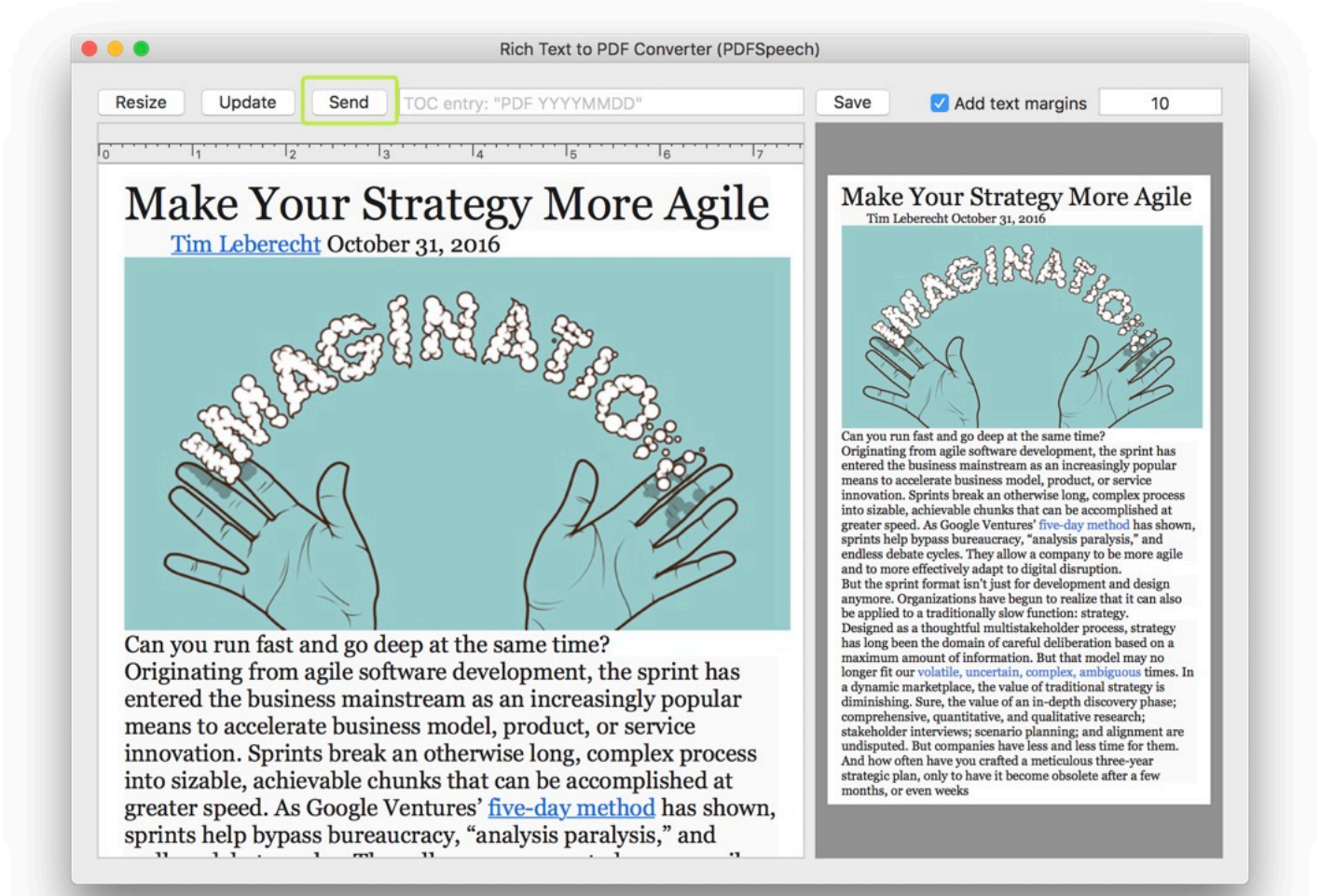
- PDFs are favored by publishers for long-form digital content such as books & magazines because PDFs can faithfully preserve text & images in carefully crafted page layouts, are easy to annotate (for study & reference), and is a mature format that works consistently across all platforms.
- Every macOS app that can print can also make a PDF, thus your reading list can be easily converted to the PDF format.
- macOS makes it convenient for developers to work with PDFs, such as to package multiple PDF documents into a single PDF, and add a table of contents to help with navigating a long PDF.

That said, you will have to bring your reading list into PDFSpeech, by turning it into PDFs where required. Here are four ways to do so in PDFSpeech:

1. **Open a PDF from disk:** Use File > Open… or click the "+" button on the top left of the window. You can also drag PDFs to the PDFSpeech icon in the dock. This is the recommended approach for importing professionally-produced PDFs, which usually contain a table of contents that can be used for navigation in PDFSpeech.
2. **Print to PDFSpeech:** Add PDFSpeech as a target in the Print… dialog, and you can print content from any macOS app such as web content in Safari and Google Chrome, reports in Microsoft Word, or sections of an EPUB from an app like Murasaki. Visit onekerato.com for a handy tip on using custom paper sizes to print a pretty PDF.
3. **PDF from Clipboard Contents:** You can copy text & images in any app, and create a PDF out of it using

File > Insert Clipboard Contents. This menu command comes in two flavors: an interactive window where you can fine-tune the copied text & resize images and the PDF, and a "fast" automatic mode. Note that PDFSpeech always creates a "slim fit" single-page PDF from clipboard contents, which frankly looks quite elegant. If you want to resize the PDF, simply resize the window to change the width of the single page (the height automatically adjusts to fit the content.)

4 **PDF pages from Preview:** Often, you may have already opened a PDF in Preview, the reader app bundled with macOS. You can select one or multiple pages from the thumbnails in the sidebar, copy to clipboard, switch to PDFSpeech and use the menu command File > Insert PDF pages on Clipboard. This is a quick way to tap into PDFSpeech to skim a few pages (or a chapter) with TTS-powered reading. Alternatively, with an AppleScript automation shortcut, you can send the PDF open in Preview directly into PDFSpeech.

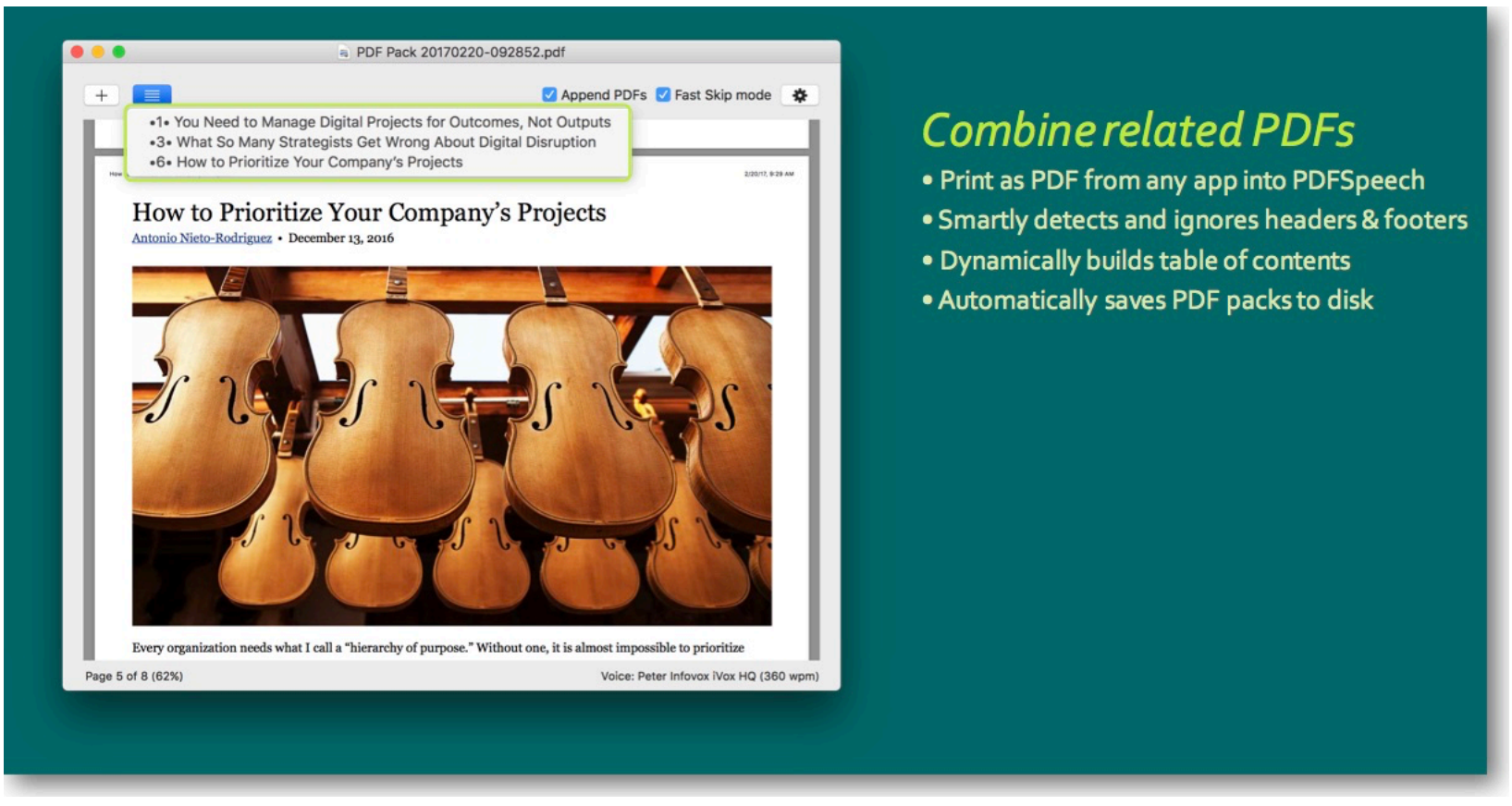


Get Organized with a PDF Pack

While PDFSpeech works great for importing PDFs, it really shines in creating **PDF Packs**, i.e. a collection of PDFs. Conveniently located on the toolbar is an option to "Append PDFs". When this option is checked, any new PDF opened in PDFSpeech will be appended to the existing PDF, to create a PDF Pack. Any existing table of contents in the PDFs will be preserved, and a top-level table of contents will be added with links to the first page of each PDF in the pack. Moreover, these PDF Packs are automatically saved to disk. To open the save location, select File > Show Autosaved PDF Packs Directory. You can also use File > Save As… to save a copy of the current PDF Pack elsewhere.

The ability to create PDF Packs is also present in PDFCombo, a free companion app for macOS, also developed by onekerato.com.

Building a PDF Pack is a great way to **collect all your reading content** in one pack: print from the web, extract pages from existing PDFs, or import text & images from the clipboard. Then you can use PDFSpeech to read through the entire PDF Pack.



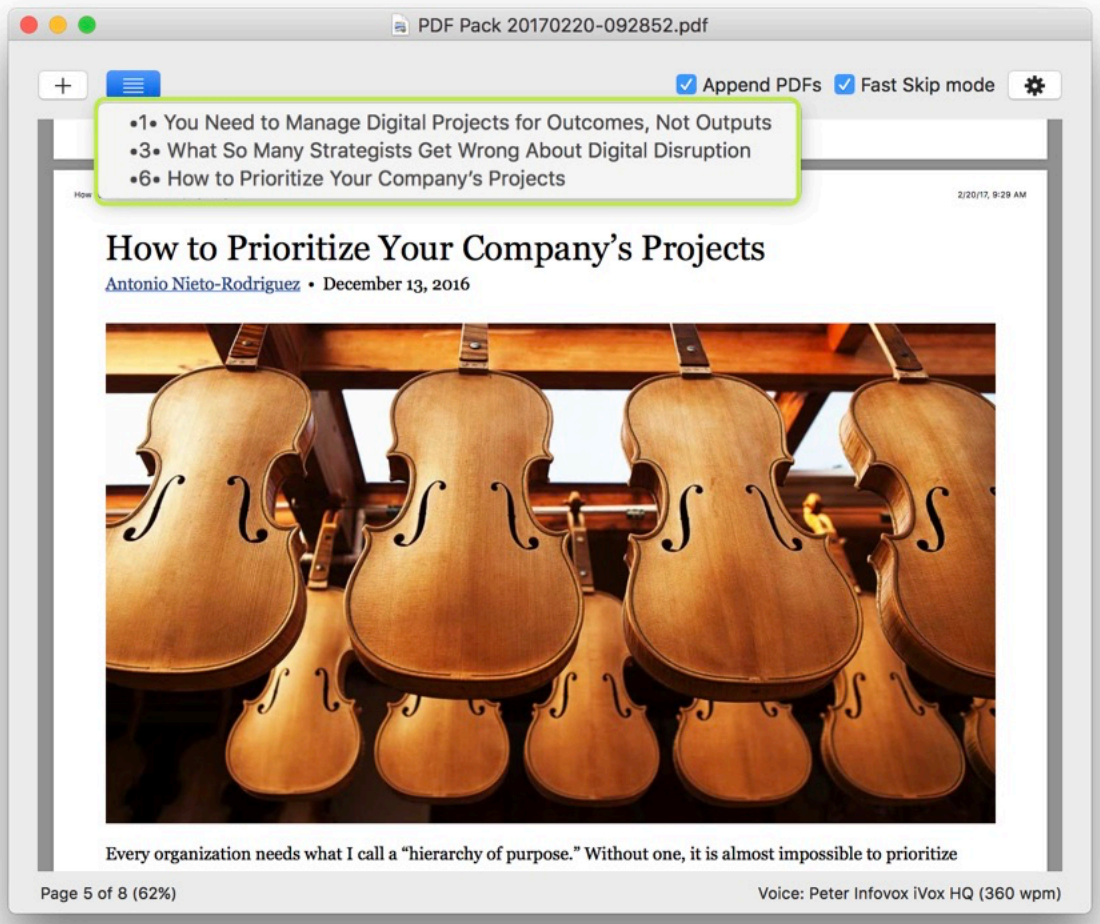
Navigate inside a PDF

In PDFSpeech, the primary means of navigating through the PDF is scrolling through it. When TTS is active, the pages will be scrolled automatically such that the sentence currently spoken by TTS is highlighted and kept visible on screen.

Click and hold the mouse to temporarily disable automatic scrolling - for example, when you want to refer to a figure while the speech continues.

In addition, if the PDF has a **table of contents**, it is shown when clicking on the **outlines button**, and you can then navigate using the outline. For PDF Packs, as PDFs are appended to the pack, the menu updates to show a top-level outline entry for each PDF in the pack. In the screenshot below, three web articles were printed to create the PDF Pack, and the outline allows switching to a specific PDF in the pack. Note also the numbered prefix for each outline entry such as "•3•", which refers to the page number for that PDF in the pack. This allows the reader to quickly estimate the size of each PDF, for example, that the second PDF in this pack is 3 pages long.

A table of contents is especially valuable in long PDFs. If your PDF does not contain a table of contents, or you would like to modify an existing one, you can use PDFOutliner, an OS X app also developed by onekerato.com.



Quick Text-to-Speech

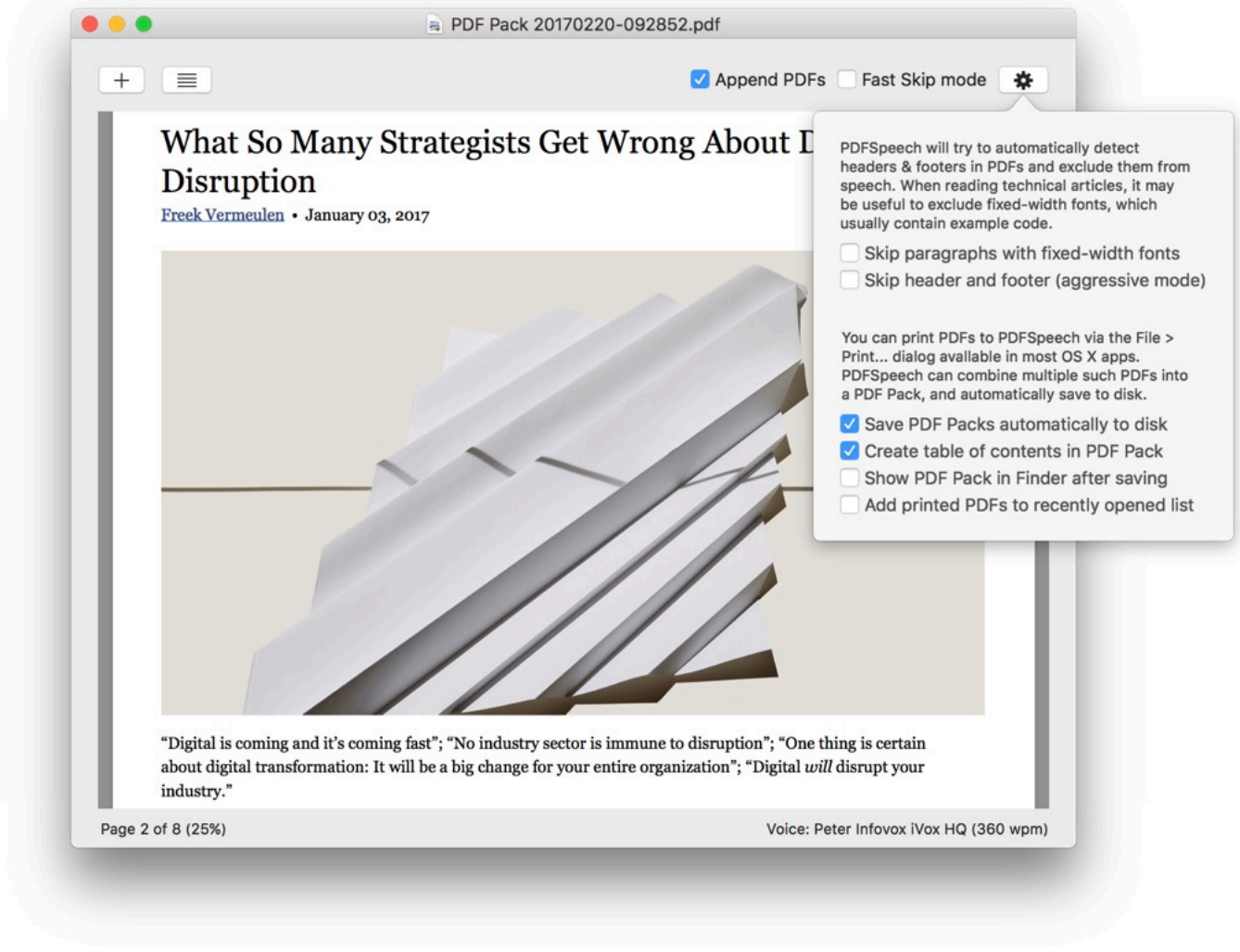
To start text-to-speech, **double-click on any word**. PDFSpeech will start speaking with the sentence containing that word, and then speak subsequent sentences. At any point, you can double-click another word to start the speech from that point forward, which also allows you to go back and read a part again. To stop text-to-speech, double-click anywhere other than a word. Usually, it's easiest to double-click in the whitespace of the page margins.

Next, this guide covers the configuration options available in PDFSpeech that make it easier to listen to the PDF document.

Configure PDFSpeech to Read Your Content

The objective of configuring PDFSpeech is to control what should be read on the page. For example, internally, PDFSpeech tries to detect URLs and not read them out. In web pages printed from Safari's reader view for example, there is usually a long URL in the footer. PDFSpeech, if it detects the URL, will abbreviate speaking it - thereby overriding the default behavior of the TTS engine which is to spell out the entire URL.

In addition, there are a few choices you can make in PDFSpeech to guide the behavior of the TTS engine, as discussed below.



Skip Headers and Footers

Headers and footers are very common in PDF documents, and it gets tedious to listen to the TTS engine read out every header & footer on every page, often breaking the chain of thought as you follow the content from one page to the next. That's why PDFSpeech automatically detects headers & footers and does not send that text to the TTS engine. However, the detection may fail on certain PDFs, and in such cases, you can enable the "**Skip header and footer (aggressive mode)**" option which tries to specifically identify fonts in the document used in headers & footers, and skip any text that uses these fonts. Of course, this algorithm may unfortunately backfire if the font used for headers & footers is identical to that used for body text. You should disable the "aggressive mode" if the TTS engine skips a lot of text that it should not.

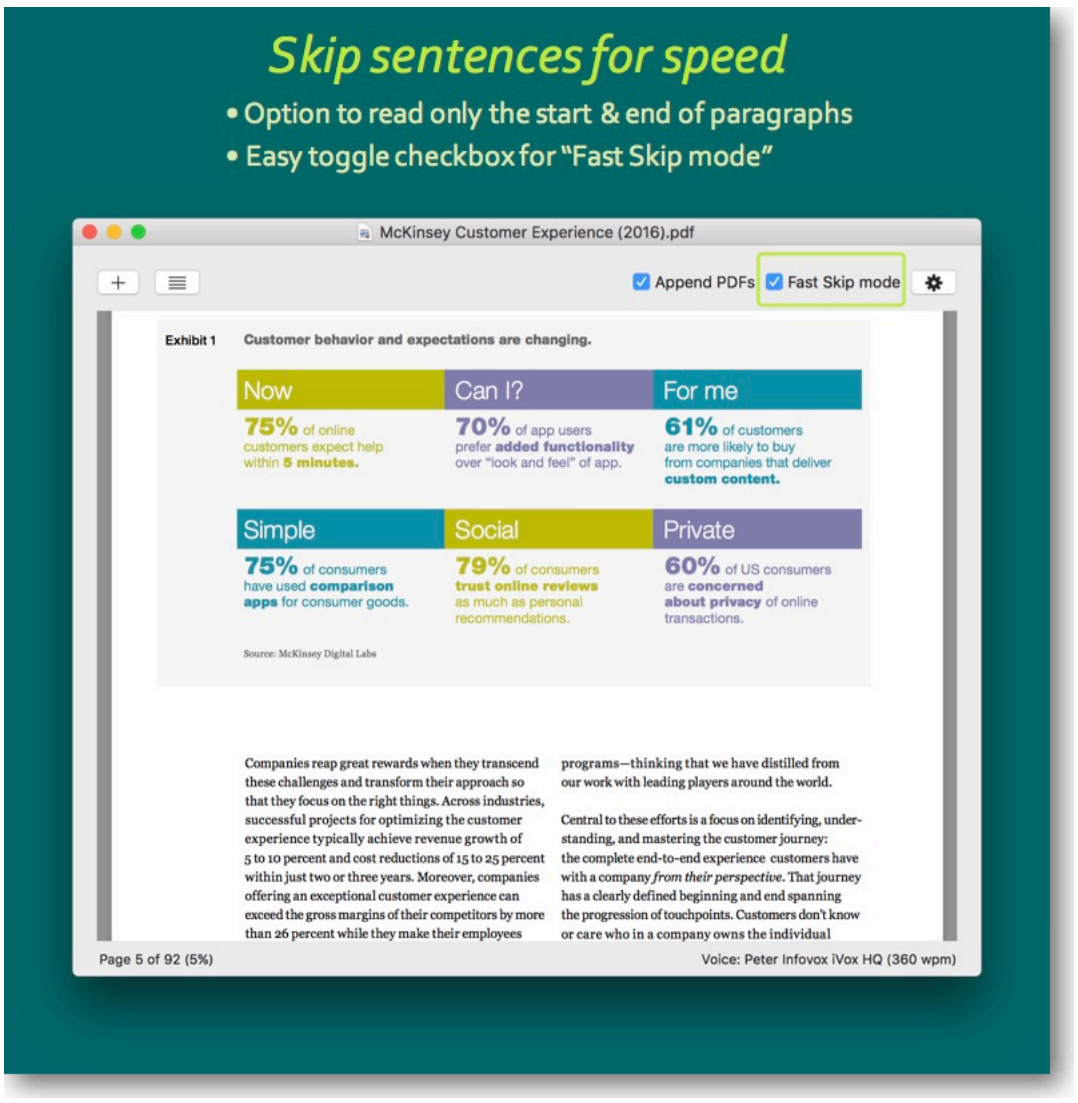
Skip Fixed Width Paragraphs

In many technical books (especially programming books), fixed-width fonts are used to mark code blocks. It may not make sense to pass such code snippets into the TTS engine, which can at best only attempt to spell out unfamiliar programming words that are not part of its vocabulary. The option to "**Skip paragraphs with fixed-width fonts**" when checked tells PDFSpeech to find paragraphs which predominantly contain fixed-width fonts and not pass them to the TTS engine. As noted above, you may find it useful to check this option when reading programming or technical books in PDFSpeech.

Fast Skip Mode through Text

One popular technique for speed reading is to focus on topic sentences and ignore the rest of the paragraph. Usually the topic sentences are at the beginning or end of paragraphs. Hence, PDFSpeech provides the option of "**Fast Skip mode**" wherein only the first two and last two sentences of paragraphs will be read out. Usually, this is sufficient to get a good overview of the content, especially in books that provide detailed explanations. This

option will save you time, and even avoid the frustration that arises with verbose text. Note that paragraphs of four sentences or fewer will be read out in full anyway.

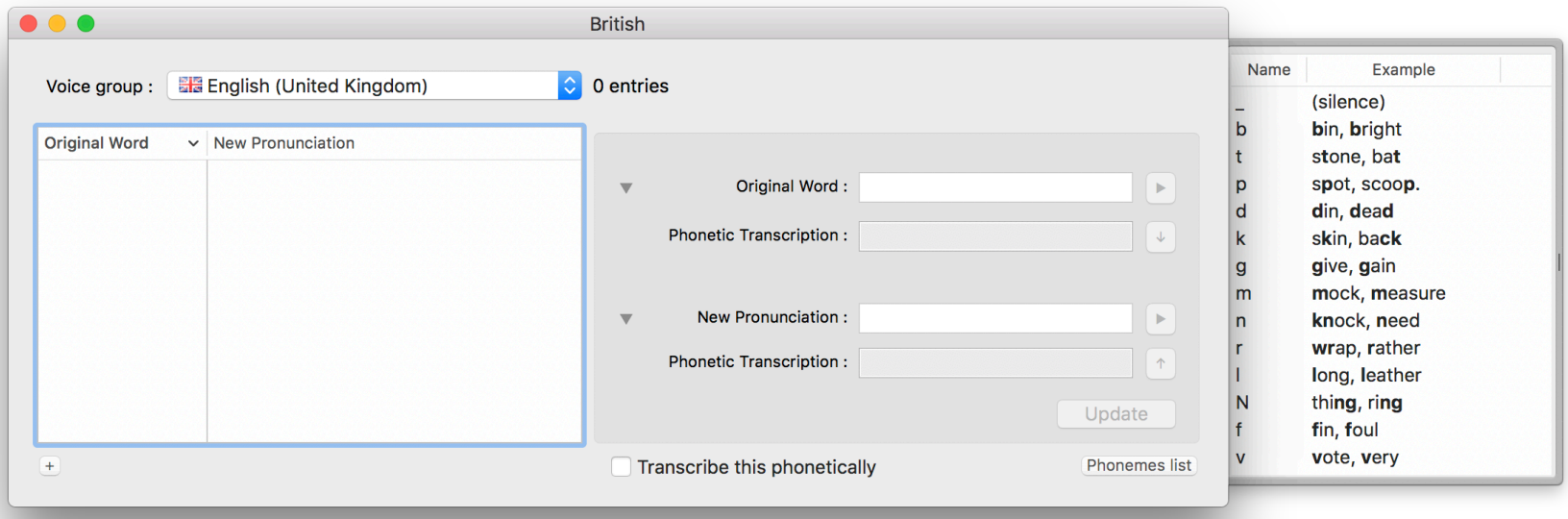


Customize Pronunciations

The TTS engine is equipped with a limited vocabulary, and tries its best to correctly pronounce words that it encounters outside of this vocabulary. For example, the unusual word "PDFSpeech" is by default pronounced by the TTS engine as its component characters "P-D-F-S p.e.e.c.h." Such mispronunciations can quickly get annoying! Therefore, the PDFSpeech application includes a simple yet very helpful word mapping feature, wherein you specify the word it should look for and what to substitute it with. To specify custom pronunciations for words, go to File → **Edit Custom Pronunciations**. Your edits are immediately picked up in text-to-speech.

In the text file that ships with PDFSpeech, you will find word mappings for words like macOS, URLs, onekerato.com and PDFSpeech that make it easier to listen to this user guide. You should add words to this customized pronunciation text file as required to **weed out pronunciation errors**, and thus make it easier for yourself to listen to the text-to-speech voice on a sustained basis.

Note that these word mappings only apply to text that is spoken through the PDFSpeech app, because the words are being intercepted and replaced before they reach the TTS engine. The **InfoVox iVox Voice Manager** third-party app provides similar functionality (via a "pronunciation editor" feature) to customize its own suite of TTS voices, where such customizations would apply across all OS X apps that use these voices.



Resources

PDFSpeech is on the web at onekerato.com/pdfspeech.html.

Visit the product website for tips on how to leverage PDFSpeech and other applications to increase your reading throughput.

For support and feedback, kindly use the menu command, Help > **Contact Customer Support**, or send an email to onekerato@me.com, or you can use the contact form on the website.