



Internationalization & Pseudo-Localization

Use Lingotek pseudo-localization services to test the localizability of websites, software, and online content such as mobile apps, e-learning, surveys, and video games. Making sure your content is properly internationalized will save time, reduce costs, and ensure high quality localization.

Contents

1) Internationalization	2
1.1 Benefits of Internationalization.....	2
1.2 7 Elements of Localizability	2
2) Pseudo-Localization	4
2.1 Benefits of Pseudo-Localization.....	4
2.2 Pseudo-Localization in Lingotek’s TMS	4
2.3 Checklist of Elements to Inspect with Pseudo-Localized Files.....	5
2.4 Pseudo-Localization Example.....	6
3) Change the Way your Company Internationalizes	6
3.1 Do It Yourself.....	6
3.2 Make Company-wide Change Easier.....	6
3.3 Invite Lingotek to Help with Company Training.....	6
3.4 Additional Resources.....	7

For questions about pseudo-localization, contact Tyson Shelley, Lingotek Language Services Account Director at tshelley@lingotek.com or 801.331.7777.



1) Internationalization

Internationalization is preparation that ensures faster, cheaper, and easier localization.

1.1 Benefits of Internationalization

According to the Localization Industry Standards Association, years of anecdotal evidence show that proper internationalization cuts localization time and costs by 50%.

One way to look at internationalization...

<p>\$1,000 fix in 1 language in internationalization</p> 	<p>OR</p>	<p>\$20,000 fix in 20 languages in localization</p> 	<p>+</p>	<p>\$20,000 fix for each update in 20 languages</p> 
---	------------------	--	-----------------	--

It is better to fix localization problems once in the source language files, so they do not need to be fixed multiple times in multiple localized languages.

1.2 7 Elements of Localizability

The main elements of localizability include the following:

- 1) Regional settings
- 2) Character sets
- 3) Translatable text
- 4) Over-externalization
- 5) Text expansion
- 6) Image text
- 7) Concatenated strings

1) Regional Settings

Can settings for the following be changed automatically according to locale?

- Currency formats (\$, €, £, ¥, etc.)
- Measurement formats (metric)
- Number formats (decimal separators: other countries use commas, where the U.S. would use a decimal point €1,000.00 vs €1.000,00)
- Time/date and calendar formats (other countries list day before month.)
- Address and phone number formats (zip codes)
- Sorting rules (A-to-Z vs Asian character sets)

2) Character Sets

Do various character sets display correctly?

Example Set	Example Characters
European Double byte & Asian Right-to-left	Español, Français, Português, Русский 中文简体, 中文繁體, 日本語 עברית

The following examples show how string concatenation might typically be used and then show a better way to write such strings for internationalization:

Typical string concatenation	Better internationalized strings
1: The files will be deleted 2: The profiles will be deleted 3: one after the other	1: The files will be deleted one after the other 2: The profiles will be deleted one after the other
%s has %d minutes remaining to complete %d questions	User: %s Minutes remaining: %d Questions remaining: %d

2) Pseudo-Localization

Pseudo-localization also goes by the following names:

- Pseudo-translation
- Test translation
- Round-trip test translation
- Translation simulation
- Dummy translation

It is a matter of simulating translation by automatically replacing text with test characters while preserving non-translatable code and simulating expansion or contraction.

2.1 Benefits of Pseudo-Localization

Pseudo-localization enables quick testing of 5 of the 7 elements of localizability:

- 1) Regional settings
- 2) Character sets
- 3) Translatable text
- 4) Over-externalization
- 5) Text expansion
- 6) Image text
- 7) Concatenated strings

Pseudo-Localization helps test at least 5 of the 7 elements of localizability

The results help to ensure a company is reaping the benefits of proper internationalization, including shorter time lines, reduced costs, and more reliable quality.

2.2 Pseudo-Localization in Lingotek's Platform

Lingotek's pseudo-localization includes the following features:

Automatic Character Replacement

Using the same filters that extract and replace text during actual translation, Lingotek automatically replaces all text with test characters.

Multiple Character Sets

Lingotek automatically suggests sets of example characters that match the languages requested for translation. Additional custom characters may also be used.

Simulation of Expansion and Contraction

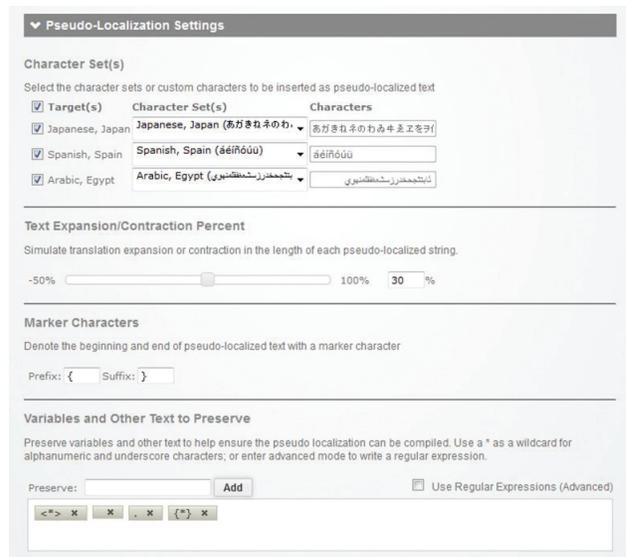
A percentage between -50% and 100% may be selected to simulate expansion or contraction.

Insertion of Marker Characters

Prefix and suffix marker characters may be inserted to easily determine where strings begin and end, thus indicating if expansion is handled adequately.

Preservation of Non-Translatable Code

Variables and other code are preserved to ensure the pseudo-localized files will compile correctly.



The above screenshot displays the pseudo-localization options available within Lingotek's Translation Management System.

2.3 Checklist of Elements to Inspect with Pseudo-localized Files

Pseudo-localized files can help ensure websites and software will be localized quickly and cost-effectively. How localizable is your application or website? Compile the pseudo-localized files (provided by Lingotek) and determine if you can answer “yes” to the following questions. As necessary, adjust the source files and repeat the pseudo-localization process with Lingotek.

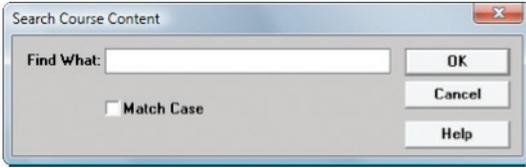
Element	Description	What to look for in the website or application
Over-externalizing	Ensure non-translatable strings are not included with the extracted text for translation.	Did the website or application compile correctly?
Character sets	Ensure that the website or application correctly supports all necessary character sets correctly.	Do all the characters display correctly in the compiled website or application? <ul style="list-style-type: none"> • ASCII characters {XXXXXX} • Extended ASCII characters {óóóóóó} • Double-byte left-to-right characters {剣剣剣剣剣} • Double-byte right-to-left characters {ي ي ي ي ي}
Translatables	Ensure non-translatable strings have been externalized in resource files and / or tagged so that they can be extracted with filters.	Has all the text in the application been replaced by the pseudo-localized characters?
Expansion	Ensure dialog boxes and forms allow for text expansion either with existing sizing or dynamic resizing.	Do the prefix and suffix markers (i.e. “{” and “}”) appear correctly at the beginning and end of each string?
Image text	Identifies text that is embedded into images and must require extra time for localization.	Does no English text remain embedded in images?

Other elements of localizability to examine

Element	Description	What to look for
Concatenations	Ensure there are not too many concatenations that may cause problems for translators.	In the decompiled website or application, do you avoid many variables and / or other code dividing a single string of text in a way that might prevent easy sentence restructuring or word reordering by the translator?
Regional settings	Ensure regional settings (including time and date formats) have not been hard-coded.	Can the following settings be changed for each locale? <ul style="list-style-type: none"> • Currency formats • Sorting rules (e.g. A-Z vs character sets) • Measurements formats (e.g. decimal separators) • Time / date and calendar formats • Address and phone number formats

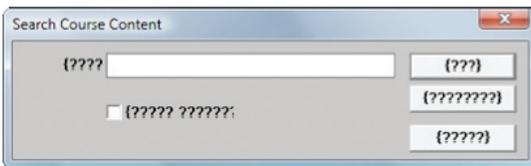
2.4 Example: Pseudo!Localization

Here are two examples of what pseudo-localization might look like for the following screenshot:



The first screenshot shows what the pseudo-localization might look like if the application still needs to be internationalized before translation begins. If adjustments need to be made, pseudo localization should be repeated until the pseudo localized application (or website) looks like the second example, where no problems are identified.

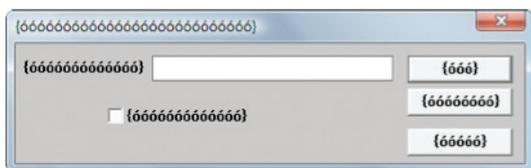
2.4.1 Example: Multiple Internationalization Issues Identified



This example identifies the following potential internationalization problems:

- 1) Some text remains in English, which means one of the following:
 - Not all translatable text is externalized in resource files or tagged for translation.OR
 - Not all translatable text has been separated from images.
- 2) Characters are not displaying properly - in this case, they appear as question marks - so there are apparently issues with encoding and character support.
- 3) Suffix markers are cut off in two strings, indicating that text expansion is not handled properly.

2.4.2 Example: No Internationalization Issues Identified



In this successfully internationalized version, all text is properly replaced with pseudo-translations, all characters display correctly, and prefix and suffix markers show proper handling of text expansion.

3) Change the Way Your Company Internationalizes

Be the hero who makes localization more rapid, cost-effective and problem-free at your company.

3.1 Do It Yourself

Ask your Lingotek representative to create a pseudo-localization for you or use Lingotek's Translation Management System to do it yourself. Then use the checklist provided to identify areas for improvement. Your next localization project will run much more smoothly.

Even if you do not request it separately, pseudo localization is a free, standard part of every Lingotek localization project.

3.2 Make Company!K ide ðH/fbUjcbU]nUjcb Easier

Clients repeatedly tell us they have a difficult time convincing fellow developers, designers, and writers to follow principles of proper internationalization... until they finally show them the pseudo-localization and "a light bulb turns on."

Something about visually identifying the problems finally causes the concepts to "click" with colleagues. People who have been resisting instructions to internationalize a website or something similar finally "get it" and make course corrections very quickly. Try using a pseudo-localization to illustrate opportunities for internationalization and convince your most stubborn colleagues.

3.3 Invite Lingotek to Help with Company Training

Lingotek has shared these free do-it-yourself tips in many conference presentations, webinars, and one-on-one tutorials. Invite us in to explain these principles to your team. We can even pseudo-localize your own content beforehand to display the most relevant examples in the presentation.

3.4 Additional Resources

Check out the following additional resources and reference materials on the topic of internationalization:

1) Internationalization Tips & Guides from Tech Giants

- Microsoft: <http://msdn.microsoft.com/en-us/goglobal/bb892890>
- IBM: <http://www-01.ibm.com/software/globalization/> and <http://userguide.icu-project.org/i18n>
- Oracle: <http://developers.sun.com/global/> and http://www.openoffice.org/specs/collaterals/guides/I18n_in_Software.html
- Apple: <https://developer.apple.com/internationalization/>
- Google: <https://developers.google.com/search?q=internationalization>
- Facebook: <http://developers.facebook.com/docs/internationalization/#best>
- Mozilla: <http://www-archive.mozilla.org/docs/refList/i18n/>
- Adobe: http://blogs.adobe.com/globalization/reference_materials/ and <http://www.adobe.com/cfusion/search/index.cfm?term=internationalization>
- Yahoo!: <http://search.yahoo.com/search?p=internationalization&vs=developer.yahoo.com%2Cdeveloper.yahoo.net> and <http://yuilibrary.com/yui/docs/intl/>

2) Internationalization Tips & Guides from Other Projects

- Debian <http://www.debian.org/doc/manuals/intro-i18n/>
- Ruby <http://guides.rubyonrails.org/i18n.html>
- Perl <http://perldoc.perl.org/perllocale.html>
- PHP <http://php.net/manual/en/book.intl.php>
- Python <http://docs.python.org/library/i18n.html>
- XML <http://www.w3.org/TR/xml-i18n-bp/>

3) Internationalization Books & Tutorials

- A Practical Guide to Localization by Bert Esselink <http://benjamins.com/#catalog/books/liwd.4>
- Going Global with Javascript & Globalize.js by Jukka Korpela http://www.bytelevelbooks.com/books/global_javascript.html
- Java Internationalization by David Czarnecki and Andrew Deitsch <http://www.amazon.com/Java-Internationalization-Series-David-Czarnecki/dp/0596000197>
- Web Internationalization Tutorial by Tex Texin & Yves Savourel <http://www.xencraft.com/training/webstandards.html>

© Copyright 2018 Lingotek, Inc. | No portion of this document may be copied or reproduced without crediting and obtaining express permission from Lingotek



THE TRANSLATION NETWORK

www.lingotek.com

Translation Software:
sales@lingotek.com
(801) 662-0050
(877) 852-4232 opt 2 - Toll free

Translation Services:
services@lingotek.com
(801) 331-7777 opt 1
(877) 852-4231 opt 1 - Toll free

Support:
support@lingotek.com
(801) 331-7777 opt 3
9 am - 6 pm MDT