Possible fundamental flaw in the language system's design | Tiki Wiki CMS Groupware :: Development
Possible fundamental flaw in the language system's design

Status
Open

Subject
Possible fundamental flaw in the language system's design

Version
18.x
22.x

Category
- Feature request

Feature
i18n (Multilingual, l10n, Babelfish)

Resolution status
New

Submitted by
hman

Lastmod by
hman

Rating
(0)

Description
Please correct me if I misunderstood this: To me the language system just translates one string of source language into another string of target language, is that right? At least that is suggested by the data structure of language.php.

If that is the case, then there is a fundamental flaw in it: The system cannot (by design) take into account that no function f exists that translations a string of source language into target language.

Because language has a context, and different contexts can radically change to correct translation!

One issue I have only seen now that I have shopping cart activated on one of my Tiki Wikis. There is a button "checkout" to log you off the wiki.

And there is a button "checkout" in the shopping cart, but this will NOT log off the wiki, but close an order, meaning the user requests payment.

The first button "Checkout" is at the moment correctly translated with default values into "Abmelden" in German.

But the second buttons "Checkout" on the shop makes no sense being called "Abmelden" as well, because the user pressing this button does not want to log off, he/she wants to pay, so the only correct translation would be "Bezahlen" (pay).

I believe this contextual change of meaning cannot be dealt with by the current language system, it cannot even handle exceptions?

Workaround

Just checking if i can edit this

Importance
5

Easy to solve?
5

Priority
25

Demonstrate Bug (Tiki 19+)

Please demonstrate your bug on show2.tiki.org

Version: trunk ▼

Demonstrate Bug (older Tiki versions)

Ticket ID
7405

Created
Thursday 02 July, 2020 18:00:03 GMT-0000
by hman

LastModif
Saturday 04 July, 2020 10:20:45 GMT-0000

Comments

hman 02 Jul 20 18:06 GMT-0000

Again, cannot edit my own bug report (error "not found").

Jonny Bradley 04 Jul 20 10:14 GMT-0000

Seemed to work...

hman 04 Jul 20 13:18 GMT-0000

Still getting "not found" (tried to change "Easy to Fix" to "Difficult")

Jonny Bradley 04 Jul 20 10:20 GMT-0000

You are correct, the Tiki translation system is not context aware in any way, a single word will be translated the same in different places.
However, as a non-German speaker, i have tried to find how "Abmelden" should be translated as "Checkout" - Google translate thinks "Checkout" should be "Auschecken" and "Abmelden" means "Sign out" (also here) so i would suggest changing that original translation in this case.

Replacing the translation system has been discussed at TikiFests since at least Tiki 6 to my knowledge, but no volunteers have come forward to even design a new system, let alone implement one sadly.

hman 04 Jul 20 12:33 GMT-0000

Changing the original English text to avoid ambiguity is a possible and quick workaround, yes, but may not be feasible in all situations. Some ambiguities might still remain where there is simply a 1:n relation from English to target language remaining. Therefore I would propose a different approach:

Give tran() a second, optional (!) parameter "context". In most cases, do not use it when calling tran(), in which case the existing translations of all existing target languages and almost all of the code of Tiki Wiki may remain totally unchanged. But in cases where ambiguity may be detected in the future (this may be in totally different places than in the combination of English and German, depending on the target language). Occurences so detected in the future might then supply the context variable, enabling two or more translations which can be easily be told apart.

With this method one could also take into account that different scenarios (use cases of a Tiki Wiki installation) dictate different translations. This differentiation could also use the context variable, or even a third, also optional parameter.

The current localization into German is written in informal language, addressing the user in the second person. While in English second person is identical in formal and informal language ("you"), in German it is not. Informal is "Du" and formal is "Sie" and the verb is then used in plural rather than in singular. The imperative "press the button" becomes informal "Drücke die Taste", formal "Drücken Sie die Taste".

While many people in the IT industry tend to talk informal, and in daily life address each other by the given name, such a text might be preferred by a younger audience. Talking to strangers informally without asking and/or using the given name is, from a cultural perspective, considered rude! Actually, when you do this in daily life and address strangers consistently with "Du" and/or given name despite this person preferring to be addressed with "Sie" and surname might even be considered offensive in legal situations...

Therefore I do feel uncomfortable when Tiki Wiki, when set to German, uniformly addresses the users of a Wiki installation informal and the admin cannot change this (except for editing the entire language.php or create numerous exceptions). A parameter to set language level would be a better solution IMHO.

In other target languages it might even be more complicated. While I studied computer sciences, I tried half a semester of Japanese classes, where I learned that different words exist for the same meaning, depending on the social rank of the other partner in a conversation, most of the time defined by the difference in age, and that even applies when talking about something the conversional partner does or shall to in interaction with third parties, and when you don't know, you have to decide upon which version is most polite...

And now for Google "translate":
Google translate is no translation system (trigramm) at all, despite the name. It uses the search engine to find similar result patterns in source and target language to simulate a translation. The result is that if many peoples' text in the target language use a wrong translation, so will Google translate. With results from "funny" to "embarrassing".

I have spent a couple of months working in Brazil, and when I tried to translate newspaper articles

about Sao Paolos Viracopos airport, it did not even recognize Viracopos was a name (capital V...), and so it "translated" that as well. But since Viracopos airport is rather small in comparison to Guarulhos (biggest airport in all of South America), and thus far more portuguese texts contained "Guarulhos" than contained "Viracopos", Google translate "translated" all (each and every) occurrence of Viracopos into Guarulhos when translating portuguese news about Viracopos airport into English.

Edit: Swapping "Checkout" for Shopping Cart with "Pay" might be the easier workaround, since it is used less than the other "Checkout". But "Sign out" for "Checkout" to log off from Tiki Wiki is also possible, yes. But with Shopping Cart, I haven't looked into the code. I feel that most of that code doesn't use tran() at all, because it remains in English even though I have added custom translation for that, and I have manually added the array_merge that is missing on custom.php (See the other bug report for that). I have already seen that the template file for products has hard coded language in it).

Attachments

filename	created	hits	comment	version	filetype
No attachments for this item					

The original document is available at
https://dev.tiki.org/item7405-Possible-fundamental-flaw-in-the-language-system-s-design