Principles for Simplifying Translation of Marathi Terms in Mobile Phones

Nikhil Welankar
Tech Mahindra Ltd
Mumbai
nikwel@gmail.com

Anirudha Joshi
IDC, IIT Bombay
Mumbai 400076
anirudha@iitb.ac.in

Kirti Kanitkar
IDC, IIT Bombay
Mumbai 400076
cutekirti@gmail.com

As technology products reach a larger portion of the world's population the importance of localization increases. In the last five years, mobile phones in India have reached many “first-time” users of information and communication technologies. These users have limited exposure to English and are comfortable in communicating in another Indian language. This preference is visible for the language spoken at home, outside home, in their choice of newspapers, and TV channels. However, we observed that many users prefer to have their mobile phones in English. One reason for this could be that there is much scope for improvement in the translation of the interfaces in Indian languages – Marathi terms used in current phones are not easily understood by users. We evaluated mobile phone models that are currently available in the Indian market to identify the problems with Marathi localisation. From these problems, we derived seven principles that can simplify the translation of terms: Use conversational terms, minimize use of formal, Sanskritised terms. Retain familiar English terms by transliterating them. Avoid transliterating English terms if users are not familiar with them. Avoid abbreviations. Translate commands as verbs or nouns + verbs, not as only nouns – a common mistake. Localisation causes grammatical mistakes – translate whole sentences, not just individual words. ‘Cool commands’ and metaphors are lost in translation – recreate new, culturally appropriate metaphors and use them consistently. Using these principles, we created a set of equivalent terms that we believe are simpler to use. We conducted three experiments with the help of three successively higher fidelity prototypes to compare the new set with the old one. Our findings show that users preferred and performed better with the set of terms derived from the principles above compared to the original terms. Users also expressed their willingness to use Marathi in their mobile phones if simpler and conversational terms are provided.

Translation, localisation, Indian languages, mobile phones, user interfaces.

1. INTRODUCTION

As technology products reach a larger portion of the world's population the importance of localization increases. The mobile phone penetration in India has been growing at a phenomenal pace. The current tele-density in India has crossed 44.87 % (Telecom Regulatory Authority of India, 2009). We compared these numbers with education data from the most recent census (Census of India, 2001) (figure 1). While doing so, we assumed that people that are more educated are likely to buy phones before people that are less educated. This implies that most new users of mobile phones in the year 2009 and in the near future will have had an education of less than primary school. In the last five years, mobile phones in India have reached many first-time users of information and

Figure 1. Education in India (Census of India, 2001) superimposed on mobile phone penetration (Telecom Regulatory Authority of India, 2009). The red bands on top of

communication technologies. These users have limited exposure to English and may be comfortable in communicating in one or two non-English Indian languages.

Effective translation of the user interface into the local languages is an important first step in the localization of a product, especially in the context of emerging markets. An otherwise well-designed product can be rendered unusable because of poor translation, as users may not understand the terms of the user interface. The problems can be particularly acute in the Indian context where for many years people tend to believe that any technology-based product can be only used in English.

Our aim was to study obstacles faced by users who are not familiar with English while using Marathi interfaces in phones and to derive principles for simpler, more suitable and more acceptable Marathi translations.

Section 2 describes some prior work related to localising interfaces. The section 3 presents the findings from a survey that we conducted to identify patterns of language use on mobile phones in India. In section 4, we present our analysis of the problems with the current localisation and principles that we believe ought to be used to solve the problems. We also present a list of 30 common terms used in mobile phones, their current translations, and translations that we derived using our principles. We evaluated these terms using three prototypes with users – we present these findings in section 5. Section 6 presents our conclusions and possibilities for the future.

2. PRIOR WORK

Literature on user interface localisation typically talks about issues such as text input and keystroke related analysis, space management (leaving about 30% extra space for characters from another language), localising all parts of the user interface including help system and tool tips, providing grammatically correct complete or near-complete sentences and issues related to managing localisation projects (O'Reilly & Associates, Inc., 2002), (Rohae 2004), (Min 2005). Some recommendations identify best practices for localization. Some of these are technical (such as use of Unicode text, isolating text from code, avoid text in bitmaps and icons, etc.) while others are cultural and political issues (such as avoiding use of slangs, ethnocentric material or controversial maps) (Microsoft Inc., 2009). There is some literature about localisation for mobile phones for Indian needs (Rajeshkannan et al 2008), but they primarily focus on localising predictive text input

methods for Tamil mobile user interfaces. Many authors have talked about the socio-cultural importance of having localised user interfaces e.g. (Marcus et al, 2000). However, there seems to be no attempts to identify principles that one could use while translating and localising the user interfaces, particularly the text used in commands and labels.

In a review about user interfaces in Indian languages on mobile phones, (Katre, 2006) identifies several wide-ranging problems related to localisation. He identifies issues such as input and poor legibility of display typefaces and translation issues such as “language drop out” (not translating some parts of the interface), ad-hoc use of English words transliterated in Hindi and Marathi. He calls for constituting a national committee for standardisation of Hindi terms for use in mobile phones.

Our work is a continuation of these thoughts. We agree that standardisation and consistency of the terms used in mobile phones will help improve the usability and growth of mobile phone user interfaces. We feel it is important to identify the principles of localisation before such standardisation can happen.

3. LANGUAGE USAGE SURVEY

We conducted a survey to compare the usage of English and Indian languages on mobile phones and in other contexts. We collected data about age, gender, education, mother tongue, duration of mobile phone use, current language on the phone, whether the user ever used a non-English language and whether Marathi or Hindi were available on their phones. We also collected data about languages that users speak at home, outside home (including work), the languages in which they preferred to read newspapers, watch news, or entertainment on TV, and the languages they follow while watching cricket broadcasts.

We collected data from 40 users (21 females, 19 males) from different locations in Maharashtra (34 from urban, 6 rural). We ensured that we had 8 users in each of these education categories: Graduate +, 11th standard to diploma, 8th to 10th standards, 5th to 7th standards and 4th standard or less. Age ranged from 19-62 (average 31.4, standard deviation 9.2). They had been using a mobile phone from 0.5 years to 10 years (average 3.4 years, standard deviation 2.2).

People overwhelmingly preferred using English on their mobile phones. 32/40 users were currently using their phones in English, 7/40 used it in Marathi and 1/40 used it in Hindi. The break-up of number of English users across education

categories was as follows: Graduate + (8/8), 11th standard to diploma (8/8), 8th to 10th standards (6/8), 5th to 7th standards (5/8) and 4th standard or less (5/8).

For further analysis, we encoded education as an ordinal variable (graduate + = 5 to 4th standard or less = 1), current language on phone (English = 1 and otherwise = 0), gender (female = 0, male = 1) and location (rural = 0, urban = 1). Significant Spearman's rho correlations emerged between location and current phone language ($\rho = 0.490$, $p = 0.001$) and between education and currently phone in English ($\rho = 0.398$, $p = 0.011$). Educated people and people living in urban areas are more likely to use their phones in English.

Correlations of current language on phone with age, gender and phone since years were insignificant, implying an even distribution across these categories. Current language on phone does not correlate with age, gender, or number of years of phone usage.

Among the users who have been using their phones in English, only 1/32 user reported having *changed to English* 3 years ago after having used it in an Indian language for 1 year. In contrast, 5/8 users with phones currently in non-English Indian languages reported having *changed to their current language* after a substantial period of use in another language.

There seems to be an 'initial barrier' to use the phone in an Indian language and people seem to prefer to stick to an Indian language once they overcome this initial barrier. The reasons for these barriers could be ignorance of availability of the option, non-availability of the language or usability issues arising out of localisation.

When asked whether a non-English Indian language is available on your phone, 8/40 users responded that they did not know. Of the remaining, 24/32 users reported that an Indian language was available on their phones and 20/32 explicitly reported that Marathi was available on their phones. Yet, as many as 27/40 reported that they had never tried a non-English language on their phones.

In contrast with the languages used on the phone, people heavily preferred non-English Indian languages in other media. 33/40 reported getting at least one newspaper at home, of which only 3 reported getting at least one English newspaper – all of these belonged to the graduate+ education group. 35/40 people reported watching news on TV, of which only 4 included English as one of the languages they watched TV news in, again all in the graduate+ education category. English is less popular for entertainment on TV. 32/40 people reported watching entertainment programmes on TV, of which only 3 reported watching English entertainment programmes – again, all from the graduate + education category.

standard to diploma. None of the lower education groups reported using English either inside homes or outside.

Similarly, people preferred non-English Indian languages in other media. 33/40 reported getting at least one newspaper at home, of which only 3 reported getting at least one English newspaper – all of these belonged to the graduate+ education group. 35/40 people reported watching news on TV, of which only 4 included English as one of the languages they watched TV news in, again all in the graduate+ education category. English is less popular for entertainment on TV. 32/40 people reported watching entertainment programmes on TV, of which only 3 reported watching English entertainment programmes – again, all from the graduate + education category.

The only category that English seems to be acceptable is watching cricket on TV. 22/40 people mentioned they follow cricket on TV (18/19 males and 4/21 females). Of these, 12/22 reported they could follow the English commentary (of these 9 belonged to either graduate+ or 11th to diploma education categories). 10/22 reported that they could follow the commentary only if it was in Hindi.

4. LOCALISATION PROBLEMS IN MARATHI AND PRINCIPLES OF MORE USABLE LOCALISATIONS

We studied existing terms from the mobile phones available in the market and shortlisted 300 frequently used terms. We then had brainstorming sessions with speakers and experts of Marathi to identify problems with current terms and explore more understandable and conversational terms for these terms. We identified seven common problems with localisation in Marathi and principles of making them more usable. These are listed below with examples.

Problem: Some terms are literal translations in a very formal, Sanskritised Marathi. Such language is not used in conversational Marathi and users are not familiar with these words.

Example: keypad active = कळफलक सक्रीय, Options = विकल्प

Solution: keypad active = कुलुप काढले, Options = पर्याय

Principle: Use conversational terms. Minimize use of formal, Sanskritised terms.

Problem: Some terms are translated from English to Marathi although Marathi speaking users are more familiar with transliterations of either original or alternative English.

Examples: Cancel = रद्द करा, Contacts (or phonebook) = संपर्क, Edit number = संपर्क संपादन

Solution: Cancel = कॅसल, Contacts = डायरी, Edit number = /3eeeeee

Principle: Retain familiar English terms by transliterating them.

Problem: Some terms are transliterations of English terms, but users are not familiar with them in day-to-day conversations in Marathi.

Example: Unlock = अनलॉक

Solution: Unlock = कुलुप काढा

Principle: Avoid transliterating English terms if users are not familiar with them.

Problem: Abbreviations are used in Marathi terms which users were not able to interpret properly.

Example: Speaker = ध्व.क्षे.

Solution: Speaker = स्पीकर

Principle: In particular, avoid abbreviation, as abbreviations are never used in Marathi.

Problem: The commands were translated as nouns instead of as verbs.

Example: Call (which implies 'make a call') = कॉल, Edit number = संपर्क संपादन

Solution: Call = फोन करा, Edit number = नंबर बदला

Principle: Translate commands as verbs or nouns + verbs, not as only nouns – a common mistake.

Problem: Localisation causes grammatical mistakes in computing displays, particularly when a part of the sentence is to be generated programmatically. This happens because the location of a word in a language is not changed according to the needs of Marathi grammar.

Example: Calling Nikhil = कॉल होत आहे निखिल

Solution: Calling Nikhil = निखिल ला फोन होत आहे or निखिल यांना फोन होत आहे

Principle: Translate whole sentences, not just individual words.

Problem: Metaphors and 'cool' commands are lost in literary translations. Some terms are translated in the wrong sense and used inconsistently.

Example: Go to (implies 'a quick way to go to frequently used places') = जा (literary translation of "Go to"), Answer (implies answer the phone / respond to the incoming call) = उत्तर (literal translation in the sense of "answer a question"), Call Register (a record of all incoming and outgoing calls) = नोंद (literal translation of "a note"), Clear (means remove or erase in this context) = स्पष्ट (wrongly translated in the sense of "uncluttered" or "not unclear" rather than "delete")

Solution: Go = शॉर्टकट, Answer = फोन घ्या, Register = आल्यागेल्या फोनच्या नोंदी.

We extended the metaphor of the lock from the command "Unlock" (कुलुप काढा) to the status message "Keypad active" (कुलुप काढले). This extension was not necessary in English, but brought in consistency in the metaphor in Marathi.

Principle: Create new, culturally appropriate metaphors and use them consistently.

5. EVALUATION

We wanted to evaluate if the terms derived using the principles presented above would lead to better usability in the prototype. Due to logistical constraints, it was not possible to evaluate all 300 terms in our experiment. We identified 5 common tasks and shortlisted 35 terms associated with those tasks for the purpose of evaluation. The tasks were unlocking the phone, making a call (short and long method), receiving a call, adding new name to phone book (short and long method) and checking a missed call. Table 1 lists these 35 frequently used terms, their current translations and recommendations using our principles. We also wanted to investigate the right level of prototyping that might be desirable for evaluating such a product. Given that we were only evaluating textual terms, we were wondering if it would be sufficient to test the products with a low fidelity prototype.

First, we made a 2-D 'paper-only' prototype with a printed image of a phone and cut out a window. Interface screens using the terms were printed on paper and were inserted behind this window (figure 2).

Figure 2: Paper prototype.

Usability tests were conducted with 8 users (4 males, 4 females) at Bhivpuri Road and Shelu – villages located 100 km from Mumbai. These users had been using their phones in English. None of them was aware of a Marathi language option on their phones.

Table 1: 35 commonly used terms on mobile phones, their current Marathi translations, and our suggestions based on the principles identified above.

Current English terms	Current Marathi translations (Set A)	Suggested Marathi translations (Set B)
Unlock	अनलॉक	कुलुप काढा
Keypad active	कळफलक सक्रीय	कुलुप काढले
Go to	जा	शॉर्टकट
Options	विकल्प	पर्याय
Save	जतन करा	जमा करा
Erase	पुसा	खोडा
Message	संदेश	निरोप
Call	कॉल	फोन करा
Back	मागे	वाहेर
End	समाप्त	बंद करा
Calling Nikhil	कॉल होत आहे निखिल	निखिल ला फोन होत आहे
Speaker phone	ध्व.क्षे.	स्पीकर
Call duration: 3:20	कॉल सारांश ३:२०	तुम्ही ३ मिनीटे २० सेकंद बोललात
Nikhil Calling	निखिल कॉल चालू	निखिल यांचा फोन आला आहे
Answer	उत्तर	फोन घ्या
Mute	मौन	आवाज थांबवा
Contacts	संपर्क	डायरी
Log	नोंद	आल्यागेल्या फोनच्या नोंदी
Setting	सेटींग	व्यवस्था
Details	तपशिल	माहिती वघा
Add new contact	संपर्कात नवीन भर	नवीन नंबर लिहा
Edit contact	संपर्क संपादन	नंबर बदला
Delete contact	संपर्क काढून टाका	नंबर काढून टाका
Search contacts	संपर्क शोधा	नंबर शोधा
Move contact	संपर्क हलवा	नंबर हलवा
New contact	नवीन संपर्क	नवीन नंबर
First name	प्रथम नाम	नाव
Surname	आडनाव	आडनाव (no change)
Phone number	फोनक्रमांक	फोन नंबर
Next	पुढील	पुढे
Save	जतन	डायरीत जमा करा
Cancel	रद्द करा	कॅसल
Number saved	संपर्क फोन मेमरीत जतन	नंबर डायरीत लिहीला
Call	कॉल करा	फोन करा
Missed call	सुटलेला कॉल	सुटलेला फोन
View	सूची	पहा

Figure 3: Thermacole based 3D prototype.

Users were asked to perform the five tasks one after another using both sets of words (sets A and B from Table 1). Half the users were asked to use set A first and the rest were asked to use set B first. Their performance was measured in terms of time they took to perform each task and number of errors they made. Each time the user chose a wrong option than intended was counted as an error. At the end of the task, they were also asked to give 'marks' to each set on a scale of 0-10. Qualitative feedback about inclination to use their phone in Marathi was collected after the test. We also encouraged users to suggest their own ideas, terms for the existing terms.

Set B outperformed set A (Table 2). We found that users made more errors in set A (3.1 errors per user across 5 tasks) and fewer errors in set B (0.6 errors per user across 5 tasks). They also gave more marks to set B (7.9/10) than set A (4.9/10). The difference is statistically significant ($p < 0.01$). Users performed somewhat faster in Set B as compared to set A, though the difference was not statistically significant.

Users preferred set B as compared to set A because they found the terms familiar and simpler. Users expressed their willingness to use Marathi in mobile phones on regular basis if the interface uses a familiar and simple language.

A limitation of the paper prototype was that users could not relate the paper prototype to the usage of an actual mobile phone. Therefore, we created a physical model in which we could slide paper strips into a 'screen' (Figure 3). The printout was of a slightly bigger size when compared to the actual mobile phone to enable easy operation. The two sets of terms A and B from Table 1 were re-evaluated with the help of this 3-D prototype with 8 users in the method described above. The test

results (table 2) showed that the difference in errors was still significant at $p < 0.05$, though the difference in marks was not significant.

From a methodological perspective, we observed that although the prototype gave a 3-D user experience, users were still unable to relate it with their actual mobile phones. At times, some users tried that task with their own mobile phones before performing it with the prototype. The moderator of the experiment tried to control the timed tasks as accurately as possible (e.g. in unlocking the keypad), yet even the smallest of delays created a time lag and a sense of loss of control for the users. The effect of timed events (e.g. message "keypad active") and its interactions with the users could not be tested effectively, on this and earlier prototype.

Therefore, we created two Flash Lite prototypes and deployed them on Nokia 6300 (Figure 4). One used the terms from set A, the other used terms from set B. We made sure that the timed events in the prototypes were as similar to real mobile phones as possible.

We repeated the test with 8 users. This time users could actually use a real mobile phone for performing tasks. Again, set B did better than set A in errors, time, and marks. The difference between the two sets is statistically significant for errors and marks ($p < 0.01$) (Table 2).

Overall, by combining results from all the three tests, we could see that set B did better than set A on errors and marks ($n = 24$, $p < 0.01$). The difference between the times taken for tasks was not statistically significant.

Table 2: Findings of usability tests comparing terms in set A and set B using three different prototypes.

Prototype	Average no. of errors		Time taken (s)		Marks (0-10)	
	A	B	A	B	A	B
Paper (n = 8)	**3.1	**0.6	57	46	**4.9	**7.9
3-D (n = 8)	*2.1	*0.6	52	44	6.8	8.0
Flash (n = 8)	**10.9	**5.2	68	59	**6.6	**9.1
Overall (n = 24)	**5.4	**2.2	59	50	**6.1	**8.3

* Difference is significant at $p < 0.05$.

** Difference is significant at $p < 0.01$.

6. CONCLUSIONS AND FUTURE WORK

Our survey established a glaring disparity in language use in Maharashtra – we prefer to speak in Indian languages at home and outside, predominantly read Indian language newspapers, and watch TV entertainment and news in Indian languages. However, we prefer to use our phones (and follow cricket) in English. Predictably, urban and highly educated users exclusively used English on their phones (though the same users frequently used Indian languages for other things). Non-English Indian language usage on phones was observable only among those with lowest education and those from rural areas.

In about 25% of the cases, users have no choice but to use English, as their phones do not support other Indian languages. This could happen in very old phones but also in the latest, high-end phones. Another 25% of English usage could be attributed to ignorance – some users do not know whether their phone supports another Indian language. Yet, an important reason for English usage for a large number of users could be the poor usability of the Indian language interfaces, in particular poor translations.

We must acknowledge the limitations of our survey – we only asked factual questions and did not investigate the reasons why users prefer one language on mobile phones and other in other contexts. Deep qualitative interviews and ethnographic studies could explain the reasons better and could be a matter of future investigations.

We identified seven types of problems with localisation of Marathi terms and seven corresponding principles for better, more usable localisation. Using these principles, we derived a set of 300 alternative Marathi terms for phone UIs that we believe would be better. We evaluated 35 of the more commonly used terms and presented findings in this paper. The other terms are available on request.

We found that less educated users from rural areas performed fewer errors and preferred to use our terms in comparison with current terms. The difference in speed of use was not significant, though the set using our terms did somewhat better. We also found that they are not satisfied with current terms and would prefer to use Marathi on their mobile phones if the terms are simpler and easier to understand.

In our evaluations, we restricted ourselves to quantifiable parameters such as speed, errors, and preference ratings. Future work using think-aloud techniques could investigate the reasons why some

translations work better than others, probe the perceptions, thought processes and confusions caused by specific terms. This will enable further refinement of the principles and help them become more exhaustive.

Our study was restricted to the state of Maharashtra and to Marathi user interfaces, though we feel that many problems, principles and solutions could be applied in other non-English languages in India. The need for simplification of terms is required in all languages.

From a methodological perspective, we found interesting effects of fidelity of prototypes. As the fidelity of the prototype increased, the performance of users (speed and errors) deteriorated but their preference ratings improved. We believe that performance deterioration happened because the higher fidelity prototypes were directly handled by the users, while the lower fidelity prototypes were handled through the interventions of the moderators. On the other hand, higher fidelity prototypes provided greater context and possibly greater confidence to users, resulting in higher preference ratings. Though relative performance of the two sets did not change due to the fidelity of the prototype, the significance of findings varied. Since this was not the focus of our research, we did not investigate the reasons for this further, though it could certainly be a matter of future investigation.

Figure 4: Flash Lite based prototype deployed on a Nokia 6300 phone.

7. ACKNOWLEDGMENTS

We thank all the participants in our usability tests. We also thank Nokia for sponsoring research reported in this paper.

8. REFERENCES

- Census of India. (2001) *Educational Level by Age and Sex for Population age and above*.
http://censusindia.gov.in/Census_And_You/literacy_and_level_of_education.aspx (retrieved February 21, 2010).
- Katre, D. (2006), Cross-cultural Usability Issues of Bilingual (Hindi & English) Mobile Phones. Indo-Danish HCI Research Symposium, Guwahati, 2006.
- Marcus, A. Gould, E.W. (2000) Crosscurrents: Cultural Dimensions and Global Web User-Interface Design. *Interactions*, volume 7, issue 4, pages 32-46.
- Microsoft Inc. (2009) Best Practices for Globalization and Localization MSDN Visual Studio Developer Centre. [http://msdn.microsoft.com/en-us/library/aa291552\(VS.71\).aspx](http://msdn.microsoft.com/en-us/library/aa291552(VS.71).aspx). (retrieved December 15, 2009).
- Min, L. (2005) Chinese character entry for mobile phones: a longitudinal investigation. *Interacting with Computers*, Volume 17, Issue 2, Pages 121-146.
- O'Reilly & Associates, Inc. (2009) Creating Applications With Mozilla.
<http://www.csie.ntu.edu.tw/~piaip/docs/CreateMozilla/mozilla-chp-11.html>. (retrieved December 15, 2009).
- Rajeshkannan, R, Nareshkumar, M., Ganesan, R. and Balakrishnan, R. (2008) Language Localization for Mobile Phones.
<http://tifac.velammal.org/CoMPC/articles/11.pdf> (retrieved February 21, 2010)
- Rohae, M. (2004) Keystroke-level analysis of Korean text entry methods on mobile phones. *International Journal of Human-Computer Studies*, Volume 60, Issues 5-6, Pages 545-563.
- Telecom Regulatory Authority of India. (2009) *Telecom Subscription Data as on 31st October 2009*. <http://www.trai.gov.in/> (retrieved December 1, 2009).