What Do Students Feel About Learning Programming Using Both English And Their Native Language?

Adalbert Gerald Soosai Raj
Interdisciplinary program in Computer Sciences and Education
University of Wisconsin-Madison
Madison, Wisconsin, USA
Email: gerald@cs.wisc.edu

Kasama Ketsuriyonk
Department of Computer Sciences
University of Wisconsin-Madison
Madison, Wisconsin, USA
Email: ketsuriyonk@cs.wisc.edu

Jignesh M. Patel
Department of Computer Sciences
University of Wisconsin-Madison
Madison, Wisconsin, USA
Email: jignesh@cs.wisc.edu

Richard Halverson
Department of Educational Leadership and Policy Analysis
University of Wisconsin-Madison
Madison, Wisconsin, USA
Email: rich.halverson@wisc.edu

***Abstract*—Programming is taught in India using English as the medium of instruction to students whose native language is not English. This places a high cognitive load on students who learn programming for the first time and who are not very proficient in English. Our study aims at finding out what the students feel if their native language is used along with English for teaching programming. As a part of our study, we taught linked list, a basic concept in programming, to two groups of undergraduate students for a week in Tamil Nadu, India. We used English to teach one group of students and English and Tamil (the native language in Tamil Nadu) to teach the other group. Our intervention consisted of 3 lectures and 1 live-coding session. We collected qualitative data by means of an open-ended feedback from the students. The analysis of this feedback shows that students have expressed positive sentiments about our bilingual teaching methodologies.**

I. Introduction

Students in India learn their subjects throughout their K-12 either in English or their *native/first language* (i.e. the language that a person has spoken from earliest childhood). This choice of language depends on whether the student studies in an English-medium school or a native-language-medium school during their K-12 [15]. Although there are two mediums of instruction during K-12, almost all the STEM subjects in undergraduate education are taught only in English [16]. Computer Science (CS), one of the STEM subjects, is also taught in Indian colleges in English. The main reason for teaching CS in English is because software companies in India use English as their language of communication as most of them are affiliated with U.S. based companies. Therefore students who aren't very comfortable in English (e.g. students who studied in a Tamil-medium school throughout their K-12 and students who aren't proficient with English even though they may have studied in an English medium school) find it difficult to understand programming concepts since the subject is already new to them and they are also forced to learn it in a language that they are not comfortable with [4], [7]. This increases the cognitive load on these students so much that they end up failing their programming courses, and eventually develop inferiority complexes about their programming abilities [10], [11].

Our research started with the following question: Can we reduce the cognitive load faced by these students by teaching them CS using both English and their native language? To answer this question and to better understand the impact of the medium of instruction to teach CS, we conducted an experiment where we taught programming using both *Tamil and English* to a group of undergraduate students (experimental group) whose native language was Tamil. We also taught programming to another set of students (control group) *only in English* even though their native language was also Tamil. We used *English along with Tamil* to teach the experimental group since we believe that even though Tamil may help students to better understand programming concepts, English is needed for them to communicate with other programmers around the world. Moreover since the documentation of the programming languages (e.g. K&R C [9]) and the online programming forums (e.g. Stack Overflow) are in English, it is very important for these students to learn English along with programming.

To teach CS using 2 languages, we use a technique known as *code-switching* [17], *code-mixing* [18], or *translanguaging* [19], [20] in which the instructor alternates between the 2 languages while communicating with the students. Code-switching, especially from the secondary to the primary/native language helps the students to focus, clarify and reinforce lesson materials that leads to better understanding of the

2475-1057/17 $31.00 © 2017 IEEE
DOI 10.1109/LaTiCE.2017.8

subject and makes the secondary language (e.g. English) more accessible to these students [17].

Our research aims at addressing the following questions:

1) What is the sentiment among students when their native language is used along with English for teaching programming?
2) Is there any difference in students' sentiments when programming is taught using only English and when it is taught using both English and their native language?

II. Related Work

Pal and Iyer analyzed the effects of medium of instruction on acquiring programming abilities among students in North India whose native language is mostly Hindi [1], [2], [3]. Programming was taught to two groups of first-year undergraduate students from Hindi-medium background (i.e. students who studied in Hindi-medium schools throughout their K-12) using English-only, and Hindi-only medium of instruction. The programming abilities of the two groups of students were tested and compared, and it was found that the students from Hindi-medium background learnt programming better when they were taught in Hindi when compared to when they were taught in English. Our work differs from this work in various aspects. We don't use Tamil-only to teach programming but instead we use a combination of Tamil and English to teach programming since we recognize the benefits of using both languages. We don't target students who are from a Tamil-medium background but instead we target all students whose native language is Tamil, even though they may have studied in an English-medium school throughout their K-12.

Lau and Yuen studied the impact of medium of instruction on teaching and learning of computer programming [5]. They conducted their study among two groups of K-12 students from Chinese-medium and English-medium schools in Hong Kong. They taught bubble sort using either C or Pascal to these students, and report that Chinese-medium students appear to understand programming concepts better than their English-medium counterparts. In our study, the students were not separated into two groups based on their medium of instruction during K-12. This is because CS is taught only using English as the medium of instruction in undergraduate colleges in India, irrespective of the students' prior medium of instruction.

Boulet studied the role of language in teaching and learning of Mathematics [6]. The author addresses some specific issues pertaining to languages, that students use to define mathematical terms, to read and interpret mathematical notations, and to describe mathematical processes. The teacher's role to foster productive mathematical discourse in the classroom using their language as a tool is also highlighted. In our study we made sure to use all the technical terms in Computer Science as it is in English. We do not try to translate these terms into Tamil as most programmers across the world use English to communicate about programming. For example, the terms like *linked lists*, *arrays*, *pointers*, *structures*, etc, are used as they are used in English. Tamil is used only to explain these programming concepts.

Probyn interviewed some teachers in South Africa who use English along with Xhosa, an official South African language, to teach Science [7], [8]. The study shows that the language of learning and teaching frequently creates a barrier to learning when it is not the native language of the learners. The benefits of code-switching between the two languages for increased comprehension among students is also highlighted. We consider our study to be an extension to this study where we try to find what the students feel about learning programming using a combination of English and their native language.

Fennema-Bloom studied the value of naturally occuring code-switching during bilingual content instruction in Mandarin/English among non-traditional immigrant high school students [17]. The author found that code-switching is a valuable pedagogic tool used by bilingual teachers to make content more comprehensible. Our work targets on finding the sentiments among students when code-switching is for teaching programming.

III. Methodology

In this section we explain the methodology that we used to conduct the experiment and to collect the data. Details about the experimental procedure including lectures, live-coding session, and students' feedback are explained below. We begin by describing the participants in our research.

A. Participants

The experiments were conducted in a well reputed Engineering college in Tamil Nadu. Two groups of first-year students enrolled in a data structures course were selected for the study. One group was treated as the control group and the other group was treated as the experimental group. The total number of students in the experimental group was 51, and the total number of students in the control group was 52. There was only one student in the experimental group who was from a Tamil-medium background during K-12 and there were two such students in the control group. Since these two groups were chosen at random from two 2 groups of students studying data structures, we had no control over the students' medium of instruction during their K-12.

B. Experimental Procedure

The following activities were performed with both the control group and the experimental group as a part of our intervention. There were three in-class lectures, and a live-coding session. The programming (coding) was done in C [9](A high level programming language). We collected an open-ended feedback from the students in both groups to understand what they felt about our intervention.

1) Three lectures on linked lists: Three classroom-based lectures, each of 50 mins duration, were taught to both groups. The basics of linked lists were taught in those three lectures. Topics discussed were: declaring a node structure for the linked list, adding a node at the beginning of the linked list, deleting a node from a linked list, calculating the length of a

given linked list, printing a linked list, etc. The same topics were taught to both the experimental group and the control group.

Memory diagrams were drawn during the lecture to help students visualize what happens in the computer's memory (stack and heap) when each line of code is executed. Memory diagrams were used in both the groups.

The main differences between the lectures for the 2 groups were the following: The lectures were taught *only in English* for the students in the control group. In addition to that, the students in the control group were required to communicate with the instructor and their classmates during the lecture only in English.

On the other hand, the lectures were taught using both *English and Tamil* in the experimental group, and the students were free to communicate in any of those two languages, whichever they felt more comfortable with. The instructor used both English and Tamil nearly equally (i.e. 50%-50%) while teaching the experimental group. We note that more than 90% of the students in the experimental group asked their questions during the lecture in Tamil. The instructor answered the questions in either English or Tamil depending on which language was used for asking the question.

The instructor used English for introducing a topic, explaining the syntax, and for explaining some technical terms (e.g. self-referential structures). Tamil was used whenever the instructor felt that the particular topic needed a detailed explanation in order to help the students understand the idea in a better way (e.g. How to change the head of a linked list when we add an element at the beginning?). Tamil was used only for oral explanations, and all content written on the chalk-board during the lectures were only in English.

A sample video of the instructor using both *English and Tamil* to teach linked lists to the experimental group can be found here [1].

2) Live coding session: Following the three classroom-based lectures on linked lists, a live-coding session was conducted for about 90 minutes. The instructor projected his laptop on a screen, and wrote C code for the following linked list functions from scratch:

1) Adding a node at the beginning of the linked list.
2) Printing all the elements in the linked list.
3) Deleting all the nodes from the linked list.

The instructor was thinking aloud throughout the live-coding session. He showed the students how he would go about writing the code for these three functions. He also showed them some common sources of errors while writing code for linked lists. The content taught during the live-coding session was the same for both the control group and the experimental group.

A sample video of the live-coding session that was conducted in the experimental group can be found here [2].

[1]http://bit.do/cs_in_tamil_and_english
[2]http://bit.do/live_coding

3) Open-ended feedback: We collected open-ended feedback from the students in both groups to understand their reactions to our intervention. Almost all the students in both the groups have expressed positive emotions about our intervention. Many students have specifically mentioned about the usefulness of Tamil, live-coding session, and memory diagrams for teaching programming. We have presented a detailed analysis of the students' feedback in the next section. The actual feedback from the students in the experimental group and the control group can be found in experimental group's feedback [3] and control group's feedback [4] respectively.

IV. Analysis of Feedback

We analyzed the qualitative data that we collected from the students. i.e. the open ended feedback that each student gave at the end of our one week of classes. We used frequency count, bigram analysis, Vader Sentiment analysis and thematic analysis to understand the sentiment among the students in both the groups and we have presented our results in this section. Since our study mainly focuses on comparing the different sentiments among the 2 groups, we decided to perform sentiment analysis rather than using an approach like grounded theory.

A. Frequency Count

We performed an analysis to count the frequency of words in each student's feedback. The feedback from the experimental group and the control group were analyzed separately. We were interested in the words that occured at least 10 times in either group's feedback. The frequency count of the most frequently occured words in the feedback are shown in Figure 1. The words with an asterisk at the end means that the word had occurred in multiple forms. For example, the word teach* means that the word had occurred in one of the following forms: teach, teaching, taught, etc.

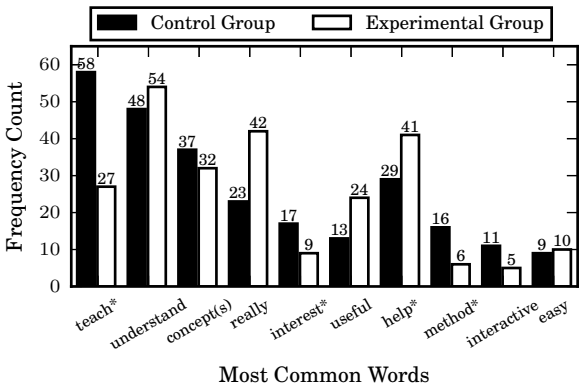

Fig. 1. Frequency count comparison of the most common words in the feedback from both the groups.

[3]http://bit.do/feedback_experimental
[4]http://bit.do/feedback_control

1) Trends in English-Tamil group: The words *understand*, *useful*, *helpful*, and *easy* have occurred more frequently in the experimental group than in the control group as shown in Figure 1. This trend is an indication that the students in the experimental group may have felt more comfortable within the classroom when compared to the students in the control group since the former were taught Computer Science using *Tamil (their native language) and English* while the latter were taught only using English.

2) Trends in English-only group: From Figure 1, we can see that the words *teach*, *method*, *interest*, *concept(s)*, and *interactive* have occured more frequently in the control group than they have occured in the experimental group. The reason for this trend may be because of the fact that even though the students in both the groups really liked the way we taught them Computer Science using visual representations and interactive discussions, the students in the control group have mentioned more about our interactive way of teaching CS since that was the main aspect that differed significantly when compared to their usual CS classes. Moreover, since the usage of the language within the classroom was the same when compared to their usual CS classes, many students in the control group had mainly commented on our teaching style rather than their level of comfort with respect to the language used within the classroom.

3) Expression of strong emotions: We found that the students in the experimental group have expressed strong emotions about our intervention. For example, many students in the experimental group had written something like *really useful*, *really helpful*, *really interesting*, *really understand*, *really awesome*, etc. The word *really* has appeared 42 times in the feedback from the students in the experimental group but has appeared only 26 times in the feedback from the students in the control group. This shows that the students in the experimental group have expressed stronger emotions when compared to the students in the control group. One possible reason for this is because these students haven't attended a CS class before where the instructor uses their native language to help them understand the programming concepts. Therefore, when we used their native language to explain some difficult concepts in linked lists, they got very excited and have expressed extremely positive sentiments.

B. Bigram Frequency

1) Raw Bigram Frequency: Table I and Table II show the raw bigrams from the control group's feedback and the experimental group's feedback respectively. The raw bigrams that are shown in these 2 tables have separate entries for the terms *understand concept* and *understand concepts*. Also note that the terms *live coding* and *coding session* means the same thing. These raw bigrams didn't give us much useful information about any trends that might emerge for the feedback from these 2 groups and hence we decided to consolidate these raw bigrams as shown below in the section on *Consolidated Bigram Analysis*.

Word 1	Word 2	Bigram frequency
linked	list	23
live	coding	14
teaching	method	9
coding	session	7
easy	understand	7
linked	lists	7
understand	concept	6
understand	linked	5
data	structures	5
understand	concepts	5

TABLE I
TOP-10 RAW BIGRAMS FROM THE FEEDBACK GIVEN BY STUDENTS IN THE CONTROL GROUP

Word 1	Word 2	Bigram frequency
linked	list	27
live	coding	15
linked	lists	15
coding	session	11
understand	concept	8
really	useful	7
helpful	understand	7
really	nice	7
understand	concepts	7
really	helpful	6

TABLE II
TOP-10 RAW BIGRAMS FROM THE FEEDBACK GIVEN BY STUDENTS IN THE EXPERIMENTAL GROUP

2) Consolidated Bigram Analysis: The raw bigrams that had the same root words (e.g. understand concept and understand concepts), that conveyed the same ideas (e.g. live coding and coding session), that expressed the same type of feelings (e.g. really useful and really helpful) were grouped together to form consolidated bigrams. The most frequently occured consolidated bigrams along with their frequencies are shown in Figure 2. The bigrams with an asterisk (*) in them means that they represent the collection of all bigrams that were similar. For example, the bigram *really* * represents all the bigrams such as *really useful*, *really helpful*, *really interactive*, *really awesome*, etc.

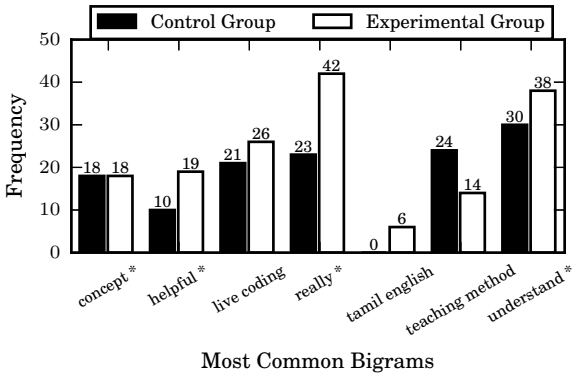

Fig. 2. Frequency comparison of the most common bigrams in the feedback from both the groups.

The following are the trends that we observed in the frequency comparison of the most commonly occurred consolidated bigrams in Figure 2 and our reasoning behind these observed trends.

- **Live coding session:**
  The bigram analysis of the students' feedback showed us that the live coding session had a huge impact on the students. The total number of times the 2 bigrams *"live coding"* and *"coding session"* occurred in both the groups' feedback combined is 47. Only one student's feedback contained two occurrences of the bigram *"live coding"* and all other occurrences of these bigrams were found in different students' feedback. This means that nearly *50% of the students* have explicity mentioned about the *usefulness of the live coding session* that was conducted as a part of our intervention. We manually read all the students' feedback that contained these 2 bigrams *"live coding"* and *"coding session"* and verified that all the occurrences were positive in nature.
- **Strong positive emotions:**
  As we already mentioned in the previous section on frequency count, the students in the experimental group displayed *strong positive feelings* about their experiences with our intervention. This is even more clearly evident from Figure 2 where we can observe that the bigram *"really *"* was found 42 times in the feedback from the students in experimental group but was found only 23 times in the feedback from the students in the control group. We read the feedback to make sure that every occurrence of the bigram *"really *"* had a positive word following "really". Examples of some words that followed the word "really" are: useful, helpful, interactive, nice, good, awesome, fantastic, etc. This means that *more than 80% of the students in the experimental group* have expressed *strong positive feelings* about our intervention while *only 44% of students in the control group* have expressed similar feelings.
- **Tamil and English:**
  The students' feedback from the experimental group had 6 occurrences of the bigram *"Tamil (and) English"* after removing the stop words like "and". There was no occurrence of this bigram in the students' feedback from the control group since those students didn't experience any change with respect to the language used within the classroom. i.e. Usually they are taught in English and we too taught them in English-only throughout our intervention. We manually read through all the feedback in the experimental group and found that 8 students have specifically mentioned that they were *greatly benefited by the idea of using Tamil along with English for teaching* linked lists as it made them *feel more comfortable within the classroom*.
  There was one student for whom the usage of Tamil within the classroom didn't help since he was a non-native Tamil speaker. He was from a state called Orissa in India where people speak a language named Oriya. Before we started our intervention we gave this student an option to attend the same lectures in the control group, where only English was used but he opted to stay in this class since he was already getting used to Tamil since it is the language that is spoken almost everywhere outside the classroom and one of his friends volunteered to translate the parts in Tamil that he didn't understand.
- **Other trends:**
  The bigram *"understand *"* which represents bigrams like *understand concept*, *understand programming*, etc, and the bigram *helpful* * which represents bigrams like *helpful (to) understand*, *really helpful* have occured more frequently in the experimental group than in the control group. This shows that the lectures that were taught in both Tamil and English to the students in the experimental group may have been more helpful than the lectures that were taught only in English to the students in the control group.
  The only bigram that has occured more frequently in the control group than in the experimental group is the bigram *"teaching method"*. As stated previously, the difference in the method of teaching Computer Science using memory diagrams and interactive sessions has had a major impact on the students in the control group since that was the only major change when compared to their usual classes. i.e. There was no change with respect to the language used within the classroom for these students.

In the next section, we use a sentiment analysis tool in Python [14] to quantitatively analyze the amount of positive and negative sentiments among the students about our intervention.

C. Sentiment Analysis

The feedback from all the students from both the groups were overwhelmingly positive. This shows that both the groups were positively affected by our intervention. We wanted a way to measure the students' feedback in a quantitative way to find if there was any difference in the amount of positivity expressed in the feedbacks from the 2 groups. Therefore, we used a tool named Vader [12] in Pythons's Natural Language Toolkit (NLTK) [13] to quantitatively analyze the amount of positive and negative sentiments among the students in both the groups. Using Vader, we measured the sentiment of each student's feedback individually and the sentiment of each group's feedback as a whole. The results from our sentiment analysis is presented below.

1) Interpretation of Vader results: For each sentence, Vader reports the following 4 values:

1) Negative
2) Neutral
3) Positive
4) Compound

Negative represents the negative sentiment value of the feedback, *Neutral* represents the neutral sentiment value of the

feedback, and *Positive* represents the positive sentiment value of the feedback. The sum of the values of Negative, Neutral and Positive will always be equal to 1. *Compound* represents the overall sentiment of the feedback. The compound value is a number between -1.0 to +1.0 and its interpretation is described as follows:

- -1.0 (Extremely Negative)
- -0.5 (Negative)
- 0 (Neutral)
- 0.5 (Positive)
- 1.0 (Extremely Positive)

For example, the sentence *"VADER is smart, handsome, and funny."* has the following sentiment values: neg: 0.0, neu: 0.254, pos: 0.746, compound: 0.8316. This means that this sentence is positive with a 75% probability and neutral with a 25% probability. The probability for this sentence to be negative is 0%. The compound value of 0.8316 tells us that this sentence is very positive since the compound value is in the range between positive (0.5) and extremely positive (1.0).

The sentiment values for the sentence *"A really bad, horrible book."* are: neg: 0.791, neu: 0.209, pos: 0.0, compound: -0.8211. These values inform us that this sentence is negative with 79% probability and neutral with nearly 21% probability. There is a 0% probability that this sentence may express a positive sentiment. The compound value of -8.211 classifies this sentence to have a sentiment between negative (-0.5) and extremely negative (-1.0).

2) Sentiment analysis of individual students' feedback: The compound value of individual students' feedback from both the control group and the experimental group are plotted in Figure 3 and Figure 4 respectively.

The key observations from these plots of individual students' compound values are as follows:

1) Around 94% of the students' feedback in the experimental group and 91% of the students' feedback in the control group had a compound value greater than 0.5. This means that the students in both the groups have really liked the way we taught them CS. Especially, the students in the experimental group have given more positive feedback and this may be due to the influence of using the native language to teach CS.
2) There is no student's feedback that is either neutral (compound value = 0), negative (compound value = -0.5), or extremely negative (compound value = -1.0). This means that our intervention didn't have any negative impact on any of the students in both the groups.
3) The mean of the compound values of the students' feedback from the control group and experimental group are 0.83 and 0.86 respectively. This highlights the fact that students in the experimental group have given slightly more positive feedback when compared to the students in the control group.

These observations highlight the fact that both the groups have given very positive feedback about our intervention and the experimental group's feedback is slightly more positive than the control group's feedback due to the use of the native language (Tamil). The number of students in Figure 3 and Figure 4 is more than those reported in Section III-A since our classes gained popularity among students from other batches and some of them attended our classes even though they weren't officially enrolled.

3) Sentiment of feedback about usage of Tamil: We measured the sentiment of the individual student's feedback from the experimental group where students had explicitly mentioned the word *Tamil*. The average value of the compound, negative, neutral, and positive sentiments were calculated for these feedback. The results from this analysis are shown in the Figure 5. It can be observed that the compound value is 0.919 which means that the *overall sentiment among students for using Tamil (along with English)* is *extremely positive*. We can also see that the average negative sentiment value of these feedback where Tamil was mentioned is extremely low (with a value of 0.017). This analysis shows us that students welcomed the idea of using Tamil along with English for teaching programming.

D. Thematic Analysis

We performed a thematic analysis of students' feedback to find the things the students liked the most in our intervention and the reasons for why they liked them. Most of the students' feedback specifically mentioned about the usefulness of Tamil to teach programming, effectiveness of the live-coding session, and how helpful the memory diagrams were for understanding programming concepts. Based on these main themes from the students' feedback we group the feedback into the following categories:

1) Usefulness of Tamil: Tamil was used to explain difficult concepts in the experimental group. All the students' comments about the usage of Tamil to explain programming were very positive. Selected comments include: *"Your class made me more attentive in the class. It is very comfortable, when you interact with us in both Tamil and English."; "The lecture was really very useful and it was easy to understand since the mixture of English and Tamil language helps us to learn better."; "The usage of both the languages Tamil and English actually kept us engaging."; "It is very helpful for us to understand the concept. I really felt comfortable in learning this topic that was taught to us in both the languages."; "This class was just rocking, we feel comfortable when we are taught in both the languages. As far from my part linked list is the toughest portion in C programming, but you made us understand easily. There was full freedom throughout the class (with respect to using the language of our choice)."*

From these students' feedback we found that using 2 languages for teaching programming made them *more attentive*, *feel more comfortable*, and helped them to *understand the concepts* in a better way.

2) Live-coding session: The live-coding session that we conducted was received well by the students. Selected students' comments include: *"It is the first time I have been in a class of interactive coding session where you learn while*

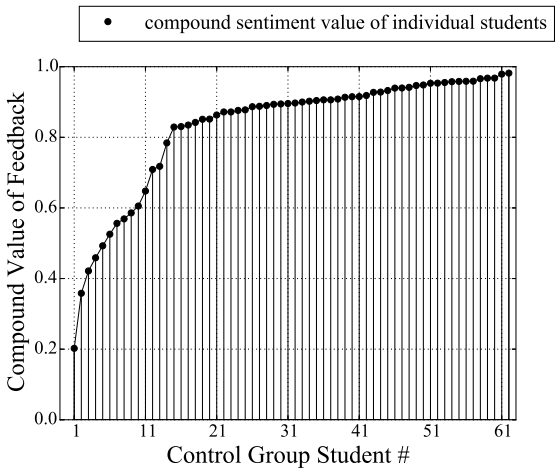

Fig. 3. Compound value of individual students' feedback in the control group

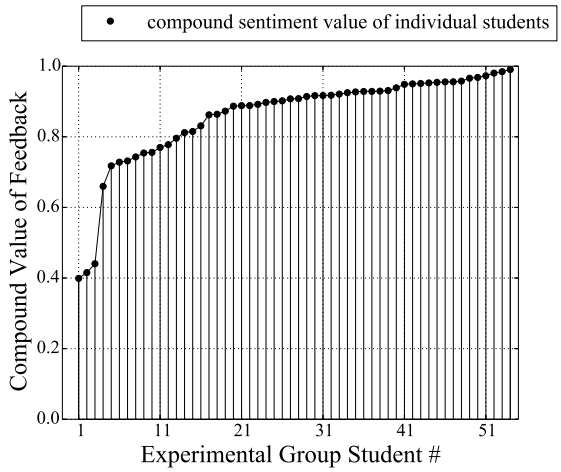

Fig. 4. Compound value of individual students' feedback in the experimental group

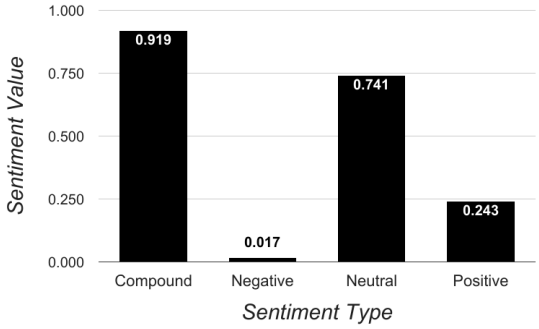

Fig. 5. Sentiment of feedback about Tamil and English

you code. This way of approach has definitely helped me (a beginner in data structures) understand linked list easily."; "It was very helpful to understand what actually happens while executing each line of code. The live-coding session helped to understand the possible bugs that may occur while coding."; "I actually found the live coding program very helpful, it was like very easy to understand other than just teaching us the theoretical concepts of programming during class hours."; "The live coding session was really useful. In that session, I came to know about the errors what we do while coding."

Based on the students' comments on live-coding session, we found that the live-coding session has helped the students to *learn about the possible bugs* that they may commit while writing code for linked lists and has also helped them *better understand the concept* of linked lists.

3) Memory Diagrams: While teaching linked lists, the instructor drew diagrams to show the students what happens in the stack and the heap portions of the computer's memory. This teaching methodology was highly appreciated by the students. Sample comments are as follows: *"The idea of visualising what happens in a computer memory helped me to understand the concept pretty easily."; "Teaching complicated concepts with the help of memory diagram helped me to understand very clearly."; "Attending the lecture what I felt very useful was the diagrammatic representation of memory cells, which avoided straining my brain a lot."; "Your teaching method like explaining how the compiler read the source code and how the works were done in memory was really wonderful, and make me to understand deeply the concepts."*

The memory diagrams that were drawn during the lectures have helped students to gain a deeper understanding of linked lists and have helped them in their learning.

V. Discussion

From our analysis of students' feedback we found that students in both the groups have expressed positive sentiments about our intervention. The students in the experiemental group have expressed strong positive emotions when compared to the students in the control group and we belive that the reason being using their native language along with English for teaching programming. Also, the overall sentiment among students for the usage of Tamil for teaching programming is extremely positive. The things that students liked the most in our intervention are code-switching between Tamil and English, live-coding session and memory diagrams.

A. Instructor's Experience

Some of the things that the instructor observed while conducting our experiment which weren't captured from the data we collected are discussed below.

The *code switching* from English to Tamil was very natural for the instructor but the code switching from Tamil to English was a little difficult. This may be due to the fact that the native

language of the instructor is Tamil and so switching from Tamil to English may have been slightly difficult. It may be interesting to study how the students feel when code switching from Tamil to English and from English to Tamil while they learn programming in a bilingual classroom.

Based on the observations by the instructor, more than 90% of the questions asked by the students from the experimental group within the classroom was in Tamil. This was surprising to the instructor since the students are usually not used to asking questions in Tamil within the classroom. It may be interesting to measure and analyze the language used by the students within the classroom when they are free to use either English or Tamil.

Only one student in our experimental group did his/her schooling till K-12 in a Tamil-medium school. All the other students in the experimental group have studied in an English-medium school. Even though most of the students in the experimental group have studied in an English medium school before their undergraduation, it is interesting to see how much they liked our way of using Tamil along with English to teach programming. This makes us wonder what may the effect of using Tamil in a class where there are mostly students who did their schooling in a Tamil-medium school till their K-12. We think that it would be very interesting to study this and we are planning to conduct such experiments as a part of our future work.

B. Limitations

India is a country with many regional languages. Therefore our teaching methodology of using the native language along with English would work only in regional colleges where almost every student speaks the native language. Even in these colleges, our methodology poses some challenges to students from other states who may not speak the native language of the state that they are attending college in. This number is usually pretty low since regional colleges are mostly attended by students who belong to the same state. In our study, there was only one student in the experimental group who wasn't very comfortable with Tamil. Even though the number of such students are very low, we acknowledge that this is a limitation of the teaching methodology used in our study.

VI. Conclusion

In this work, programming was taught to students using their native language (Tamil) along with English. Students have expressed positive sentiments about using both Tamil and English for teaching programming. The students in the experimental group have expressed strong positive sentiments than the control group about our intervention. We attribute this increased positivity in sentiments to the usage of the native language within the classroom. Even though this work was done specifically in India, the results of our work are applicable for teaching CS more effectively in many other countries, where English is not the native language.

Acknowledgment

Our sincere thanks to Prof. Saswati Mukerjee and Prof. Annie Bennet for helping us with conducting our experiments in India. We thank Prof. Michael Swift and Prof. Matthew Berland for their reviews and feedback on our work. Our heartfelt thanks to Devin Wixon for her suggestions about our experimental methodology.

References

[1] Pal, Yogendra, and Sridhar Iyer. *"Effect of medium of instruction on programming ability acquired through screencast."* Learning and Teaching in Computing and Engineering (LaTiCE). IEEE, 2015.

[2] Pal, Yogendra, and Sridhar Iyer. "*Comparison of English versus Hindi Medium Students for Programming Abilities Acquired through Video-Based Instruction.*" Technology for Education (T4E). IEEE, 2012.

[3] Pal, Yogendra, and Sridhar Iyer. "*Classroom Versus Screencast for Native Language Learners: Effect of Medium of Instruction on Knowledge of Programming*." Proceedings of the 2015 ACM Conference on Innovation and Technology in Computer Science Education. ACM, 2015.

[4] Pal, Yogendra. *A Framework for Scaffolding to Teach Programming to Vernacular Medium Learners*. Diss. INDIAN INSTITUTE OF TECHNOLOGY BOMBAY, 2016.

[5] Lau, Wilfred WF, and Allan HK Yuen. "*The impact of the medium of instruction: The case of teaching and learning of computer programming*." Education and Information Technologies 16.2 (2011): 183-201.

[6] Boulet, Genevieve. "*How does language impact the learning of mathematics? Let me count the ways.*" Journal of Teaching and Learning 5.1 (2007).

[7] Probyn, Margaret. "*Teachers voices: Teachers reflections on learning and teaching through the medium of English as an additional language in South Africa.*" International Journal of Bilingual Education and Bilingualism 4.4 (2001): 249-266.

[8] Probyn, M. J. "*Learning science through two languages in South Africa*." International Symposium on Bilingualism. 2005.

[9] Kernighan, Brian W., and Dennis M. Ritchie. *The C programming language*. 2006.

[10] T Murugavel, *The Problems of Non-English Medium Engineering Students and Possible Solutions*, http://worldlitonline.net/the-problems-of-non-english.pdf, 2011. Accessed: March 16, 2017.

[11] Vasudha Venugopal and Aravind Kumar, *Language barrier puts students from Tamil medium in a fix*, http://www.thehindu.com/features/education/issues/language-barrier-puts-students-from-tamil-medium-in-a-fix/article3336855.ece, 2013. Accessed: March 16, 2017.

[12] Hutto, Clayton J., and Eric Gilbert. *"Vader: A parsimonious rule-based model for sentiment analysis of social media text*." Eighth International AAAI Conference on Weblogs and Social Media. 2014.

[13] Bird, Steven. *"NLTK: the natural language toolkit*." Proceedings of the COLING/ACL on Interactive presentation sessions. Association for Computational Linguistics, 2006.

[14] Van Rossum, Guido. "*Python Programming Language*." USENIX Annual Technical Conference. Vol. 41. 2007.

[15] Annamalai, E. "*Medium of power: The question of English in education in India.*" Medium of Instruction PoliciesWhich Agenda? Whose Agenda (2004): 177-94.

[16] Jayaram, N. "*The language question in higher education: trends and issues.*" Higher Education 26.1 (1993): 93-114.

[17] FENNEMA-BLOOM, JENNIFER R. "*Code-scaffolding: A pedagogic code-switching technique for bilingual content instruction.*" Journal of Education (2009): 27-35.

[18] Berthele, Raphael. "*The influence of code-mixing and speaker information on perception and assessment of foreign language proficiency: An experimental study.*" International journal of bilingualism 16.4 (2012): 453-466.

[19] Garca, Ofelia, and Claire E. Sylvan. "*Pedagogies and practices in multilingual classrooms: Singularities in pluralities.*" The Modern Language Journal 95.3 (2011): 385-400.

[20] Mazak, Catherine M., and Claudia Herbas-Donoso. "*Translanguaging practices at a bilingual university: A case study of a science classroom.*" International journal of bilingual education and bilingualism 18.6 (2015): 698-714.