Understanding Students' Use of Code-switching in a Learning by Teaching Technology

Evelyn Yarzebinski[1], Amy Ogan[1], Ma. Mercedes T. Rodrigo[2], Noboru Matsuda[1]

[1]Human-Computer Interaction Institute, Carnegie Mellon University, USA
[2]Department of Information Systems and Computer Science,
Ateneo de Manila University, Philippines
{eey2,aeo,noboru.matsuda}@cs.cmu.edu; mrodrigo@ateneo.edu

Abstract. Personalized learning systems have shown significant learning gains when used in formal classroom teaching. Systems that use pedagogical agents for teaching have become popular, but typically their design does not account for multilingual classrooms. We investigated one such system in classrooms in the Philippines to see if and how students used code-switching when providing explanations of algebra problem solving. We found significant amounts of code-switching and explored cognitive and social factors such as explanation quality and affective valence that serve as evidence for code-switching motivations and effects. These results uncover complex social and cognitive interactions that occur during learning interactions with a virtual peer, and call for more affordances to support multilingual students.

Keywords: teachable agents, personalized learning systems, self-explanations, code-switching

1 Introduction

Personalized learning systems (PLS) [1] are an important part of 21st century classrooms, as they provide teachers with technology-based curricula that have been demonstrated to show significant learning gains across many domains. Used in conjunction with traditional classroom instruction, these systems give students the opportunity to receive individualized feedback, hints, and more as they practice solving problems. Many of these systems incorporate natural language technologies in an effort to bring them ever closer to the gold standard of expert human tutoring [2,3].

Such advanced or intelligent educational technology is still overwhelmingly designed by and for a western, Caucasian, English-speaking audience [4]. Even in American classrooms, however, the concept of a majority demographic in schools continues to lessen. While educators can expect that some students enter the classroom as monolingual native English speakers, more students enter the classroom as members of two or more cultures, as native speakers of multiple English dialects, and even as native speakers of other languages.

While we know that multilingual students perform better when they are able to choose which of their languages to use in a learning situation [5], monolingual PLSs

adfa, p. 1, 2011.

may not have the affordances and capabilities to support students in this way. Some existing intelligent educational technologies have been designed to cater to some of these excluded audiences [6,7], but the majority do not consider that their design choices make the inclusion of students who do not fit normative values difficult. This is especially important because as PLSs are deployed more extensively and even exported to other countries, they take with them their design principles – and limitations.

One illustrative example of these types of multilingual classrooms is the Philippines, in which students speak at least one of over 150 regional languages at home in addition to the common language, Filipino [8]. Since the Philippines is a multilingual country but retains English as a language of instruction, it provides an interesting testbed for the cross-cultural investigation of PLSs, as language translation is not a strict requirement for deployment. In fact, in a systematic review of AIED and ITS papers over the last four years, [9] found that the Philippines was one of the highest producers of ITS research outside of the World Bank's list of high-income nations.

With English often as the sole language of instruction on frequently deployed PLSs [9], students in these contexts may need to choose between their native language of cognition and their language of interaction with a given PLS. Previous research in written tasks involving one's second language has shown that bilingual students chose to use their mother tongue double to triple the amount of time over their second language when performing a math reasoning task, regardless of whether that task had high or low cognitive demands [10]. Despite this evidence from even simple written tasks, PLSs have not typically been designed to address these concerns, even as they become more widely deployed on a global scale. In fact, little prior research has investigated student language use in such systems. It is not known what language students choose to use in a PLS or why. However, this information is a critical component for designers of systems that allow for natural language input, and may even impact those that simply provide content in English.

In this paper, we report on an investigation into the cognitive and social underpinnings of students' interactions with a specific PLS deployed in the Philippines, particularly what language they chose to use and how they positioned themselves towards the pedagogical agent used in the system. We found that students did indeed switch between the language of instruction and their home language, with low prior knowledge students using less English. Student ability was additionally a factor in the affective quality of their language use. We report on these findings below, which together suggest that students may be code-switching to avoid face-threatening situations or to express their frustration with the system – findings that have implications for the design of systems that will be deployed in multi-language contexts.

2 Background and Related Work

The educational technology we deployed in this investigation is SimStudent [11], a machine-learning based system that employs the concept of learning by teaching. In this system, students use an equation-solving interface to "teach" algebra to Stacy, who is an onscreen pedagogical agent introduced to the students as a peer learner.

Students decide what problem Stacy should solve and then, after typing it in the interface, lead her through the process step by step, confirming or rejecting her problem steps, or providing the next step if she has no idea what to do next. Stacy asks a number of contextually-relevant questions in order to elicit deeper thinking from the student, such as questions like: “Why did you choose 3x+5 = 9x-2 for the problem?” and “How did you know to divide 3?”. The student’s goal is to tutor Stacy well enough to pass all 4 sections of a quiz given by the system. This system has previously been successfully used in studies in U.S. classrooms [11].

We believe that SimStudent is a representative PLS for study in multilingual classrooms because it allows for natural language input in the form of self-explanation (SE) and always responds with an “Okay!,” regardless of the language of input or its content (students, however, may believe that the system more deeply parses their input). A system which focuses on learning with a pedagogical agent also creates an interesting potential for deeper social issues to arise in students’ language use, since such systems are intended to simulate human interaction in such a way that students feel they are engaging with a virtual “person” they might care about [12, 13]. Previous research with SimStudent indicates that students do indeed treat Stacy as a social being, using language with the agent that is very similar to that of human friends [14, 15]. SimStudent has also been evaluated previously in the Philippines [16], although students’ use of multiple languages was not explicitly studied.

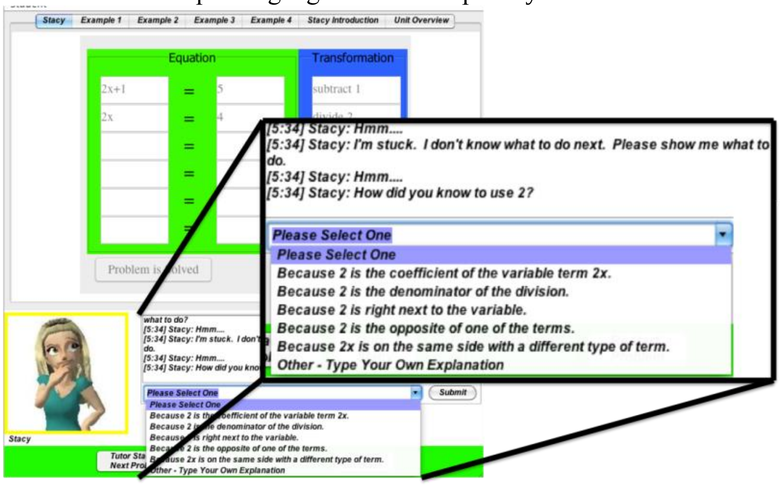

Fig. 1. The SE version of the SimStudent interface with an inset of the feedback module

Alternating between languages within a sentence or conversation is known as code-switching, a common phenomenon that occurs in speakers of multiple dialects or languages as they navigate their multiple linguistic representations of the world around them. Though there is debate in the linguistics field about code-switching vs. code-mixing (as highlighted in [17]), we use the term ‘code-switching’ (CS) for these analyses. There are many ways one can define instances of CS; we operationalize it here to mean the use of at least one word of Filipino in a SE intended for an English-only system. We use the term CS to indicate students are choosing to switch from using the expected input language (English) to a language they were told that Stacy did not understand (Filipino).

In general, there are many complex reasons why someone might choose to use CS. Ritchie & Bhatia [17] discuss four broad motivations: social and interpersonal dynamics, compartmentalizing particular language use only for certain topics (e.g., if words for particular math concepts are not known in one language), indicating additional pragmatic meaning like emphasis or vagueness, and social acceptability of using a particular language in a given context. To further expand on the social dynamics described by [17], CS can also function as a broad label to include people part of an 'in-group' or exclude an 'out-group', based on whether a particular person can understand what is being said [18].

This subtle, but powerful way to demonstrate inclusion and exclusion has the ability to cause friction within groups, especially in situations with speakers of a diverse number of languages. CS in the Philippines is an especially tricky matter. More than 150 languages are spoken around the country, and classroom language policy in the Philippines has historically been the subject of much debate between educational researchers and government leadership. Despite President Arroyo declaring in 2003 that from as young as Grade 3 the medium of instruction for mathematics would be English only, research in Philippines classrooms (as in other multilingual classrooms) has indicated that gains in mathematic problem solving are greater when teaching a student in her native language - whether that language is English or Filipino - than in her second language [19]. A new policy was enacted beginning with the 2012-2013 school year mandating mother-tongue based instruction in grades K-3, but starting in Grade 4, English and Filipino would both be used as mediums of class instruction, with English mandated for math classes [20]. Our dataset was collected prior to this new policy. Based on the language situation in the Philippines, the structure of our technology, and prior literature, we believe there are several viable explanations for why some students may choose to use CS when providing SEs to SimStudent:

Cognitive explanations.

1. *Cognitive load reduction* [18]: students may use Filipino to express math concepts that they are unable to express as well in English.
2. *Knowledge activation* [18]: students may employ CS when using particular math concepts if they were learned in a specific language.

Social explanations.

3. *Rapport building* [18]: if students believe that the system understands their home language, they may use it to build feelings of solidarity or teamwork.
4. *Face threat reduction* [18]: students may answer in a language they believe the system cannot understand, to reduce feelings of shame or inadequacy about being wrong or uncertain in front of a partner.
5. *Frustration with own knowledge* [21]: use of affective language is more strongly linked to one's native rather than second language, so students may use their native tongue to express negativity if they do not feel self-efficacious about their tutoring.
6. *Frustration with partner* [18]: students may blame the system for errors and exclude the agent from the conversation by using their native language.

Alternate explanations posed in the literature for CS, like 'an attempt to be globally-minded,' are unlikely to apply given that the direction of CS in this student population is from the more global, English, to the less common, Filipino. In this paper, we investigate which of the six explanations above have merit.

3 Research Questions

To better understand which explanations have merit, we ask the following research questions regarding students' interactions with Stacy, formulated around factors contained in each of the hypotheses:

RQ1) To what extent do students use CS in a natural language-based PLS?
RQ2) Does a student's prior knowledge affect the likelihood of CS?
RQ3) What content are students expressing via CS?
- *a) Is student CS associated with math content?*
- *b) Is student CS associated with negative affect?*

4 Methodology

The data we use to examine these questions was collected as part of a study run as a class-level randomized study in five classrooms in Manila, Philippines during students' normally scheduled algebra classes. Students were assigned to one of two versions of SimStudent: (1) the Baseline condition, in which students tutor Stacy or (2) the Self Explanation (SE) condition, which additionally prompted the student to provide a SE after some specific actions, such as the student entering a new problem, indicating that one of Stacy's suggested next steps was not correct, or providing the next step if Stacy has trouble. The Baseline condition only allowed students to help Stacy with numerical responses, while in the SE condition students could either choose a SE from a pre-populated menu or elect instead to type in their own SE in a text box. Since we were investigating cognitive and social implications of students' interactions with Stacy, we focused only on the students in the SE condition (who were able to type SEs in natural language). All of the pre-populated menu SEs were only in English, as seen in Fig 1.

Prior to tutoring, students were given a pretest consisting of 5 sections that tested declarative and conceptual knowledge. In this analysis we focus on the conceptual knowledge, because SEs are strongly linked to conceptual learning [22]. After a 3 day intervention, students received an isomorphic posttest [11]. Overall, 131 students took the pretest, posttest, and used SimStudent: 52 Baseline condition students and 78 Self-Explanation condition students. We focus on these 78 students below.

We coded each SE (N = 1810) for the binary categories of CS (Yes or No) and Valence (Non-negative or Negative), and a 6-level category of Quality (labels marking level of mathematic relevance). CS was labeled as Yes for input that included at least one word in Filipino. A Negative label was given to input that e.g., showed frustration or anger or used caps lock (shouting). To decide Quality, 3 coders used a manual that specified whether SEs contained particular content, specifically, (a) a conceptual link; (b) a procedural link; (c) contextually-related, but without having conceptual or procedural links; (d) expression of uncertainty; or (e) a nonsense response or one that was completely unrelated to math. The coding schemes for Valence and Quality are described in full in [16]. Cohen's Kappa was computed for these three code categories: κ(CS)=98.9%, κ(Valence)=68.1%, and κ(Quality)=82.5%. A label of Source noted whether an SE was a pre-existing menu item or was typed in by the student. Source was automatically derived; there is no reported coding scheme or reliability. Some results from this Philippines study were previously published in [16], but the CS data has not been investigated.

5 Results

Below, we describe results from the three research questions which we analyzed using paired t-tests or independent samples t-tests, as appropriate.

RQ1 – To what extent do students use CS in a natural language-based PLS?.

We first investigated whether students used CS in SimStudent. It is important to note that when students asked if Stacy understood Filipino, the native Filipino experimenters recommended that they provide English SEs to Stacy. Per student, the mean number of SEs prompted by the system was 22.9 (SD = 16.0), of which students elected to provide an average of 15.2 (SD = 12.6) SEs in natural language. Seventy-four of the seventy eight students (94.9%) typed at least one SE in natural language, while four students used only menu SEs. Of the seventy-four who used natural language, we found that a full 26% provided at least one CS SE; that is, SEs that included at least one word in Filipino. On average, these students who did employ CS produced 2.25 (SD = 1.5) CS SEs.

RQ2 – Does a student's prior knowledge affect the likelihood of CS?

We investigated which types of students might be more likely to demonstrate CS behaviors. One such factor suggested by our six hypotheses was student ability. We thus performed a pretest median split to differentiate between high prior knowledge (PK) students and low PK students. This enables us to compare successful and struggling students and the different types and amounts of SEs they gave. The median split occurred at a pretest score of 34%, with 39 students both above and below the split.

We compared the conceptual learning gains (defined as (posttest-pretest)/(1-pre)) of these groups to ensure that they were indeed different from each other. Students in the low PK group did have a significantly higher conceptual normalized gain than students in the high PK group: low PK students showed significant conceptual pretest-posttest improvement, but high PK students did not (see Table 1). Given that between-group differences in conceptual learning existed, we could then continue to look at whether they used language differently.

Table 1. Test scores

Measure	Group	Mean (SD)	*t*	Cohen's *d*	*df*
Conceptual Normalized Gain	High PK	-0.05 (0.34)	1.97*	0.50	62[a]
	Low PK	0.08 (0.20)			
Number of CS SEs	High PK	0.54 (1.27)	0.27	0.06	76
	Low PK	0.62 (1.21)			
Number of English-only SEs	High PK	18.10 (11.32)	-2.61*	-0.60	76
	Low PK	11.08 (12.43)			

Note: a = Levene's Test was significant; DF were adjusted. * = $p < 0.05$

Within the low and high PK groups, there was a statistically equivalent number of students who used CS, at 28% and 23% respectively. While similar *proportions* of both groups used CS, we found different *rates* of language use; while students in the low PK group (M=0.62) wrote a statistically equivalent amount of CS SEs as high PK students (M=0.54), low PK students wrote significantly fewer SEs exclusively in English (M=11.08) than high PK students (M=18.10) (see Table 1).

RQ3: What content are students expressing via CS? Is student CS associated with math content or non-math content? Is student CS associated with negative affect?.

Given the lower frequency of English use by low PK students, we investigated what students expressed when they used CS. When considering the mathematical relevance of students' SEs, half (49%) of all CS SEs were classified as being completely unrelated to any mathematical content (Quality code (e)), compared to 38% of all English SEs. For example, students brushed Stacy off without helping:

Stacy: "*What will doing the problem b-2=5 help me learn?*"

S2-(H): "*basta sumagot ka na lang*" [it doesn't matter, just give an answer]

Another third (31.9%) of CS SEs stated that they didn't know the answer to Stacy's question (Quality code (d)), as opposed to only 2.7% of English SEs. Of the remaining CS SEs (19%), the best example of providing part of a valid, mathematical SE (Quality code (a), (b), or (c)) came from a high PK student:

Stacy: "*What will doing divide 1 accomplish?*"

S1-(H): "*kasi nga parehas silang 9 that what i do is to divide it from both side*" [because they are both 9]

We next considered the valence of SEs students provided, as many motivations for CS involve affective behavior. Across all students and all natural language explanations, the mean number of negative SEs was 12.9 (12.0), while the mean number of non-negative SEs was 10.2 (8.8). We then investigated the relative valence of CS SEs, including only students who code-switched (N=20). We found that indeed students provided significantly more negative CS SEs than non-negative (See Table 2).

Comparing by ability, while the low PK group (M=11.31) gave a statistically indistinguishable amount of negative SEs from the high PK group (M=14.51), they gave fewer non-negative SEs (M=7.13) than the high PK group (M=13.36) (See Table 2). For example, S3 – low PK - first provides a seemingly innocuous SE in English, but starts using Filipino in order to insult Stacy:

Stacy: “*Why did you choose this problem?*”

S3-(L): “*because I should teach you first in a easy way. tanga*” [stupid]

while other students stuck completely to Filipino for their insults:

S4-(H): “*wag ka ngang pakialamera masyado ka kcng epal e kaya ang daming nagkakamali*” [stop being nosy. you're such a (expletive), that's why people are making mistakes]

Table 2. Self Explanation Counts

	Groups	Mean (SD)	*t*	Cohen’s *d*	*df*
Number of Negative SEs	High PK Low PK	14.51 (13.90) 11.31 (9.7)	-1.18	-0.27	76
Number of Non-Neg. SEs	High PK Low PK	13.36 (8.55) 7.13 (8.03)	-3.32***	-0.76	76
Number of CS SEs	Positive SEs Negative SEs	0.10 (0.45) 2.15 (1.42)	-6.10***	-2.80	19

Note: ***= $p < 0.001$

We did not compare the relative rates of negative vs. non-negative explanations in low PK students, as the numbers were too low to be meaningful.

6 Discussion and Conclusion

We found that language use is indeed a factor to pay attention to in PLS deployments, as students did code switch while using an English-only system in a classroom with English as the language of instruction. Over a quarter of students provided one or more SEs in Filipino. While identical numbers of high and low PK students chose to employ CS, low PK students wrote fewer SEs in English than high PK students, at first glance suggesting that a cognitive factor may underlie the CS effect. However, CS SEs in general contained less math-related content than English SEs (with 81% containing no math content or coded as “I don’t know”), suggesting that students were not simply *reducing cognitive load* or *activating prior knowledge* in a language they were more comfortable with. Instead, students frequently used Filipino to make a statement of uncertainty (whether real or feigned) – which they very rarely did in English – or chat off-topic. These two findings point towards explanations that include social features.

Given that students were told that Stacy did not speak Filipino and that they should use English with her, our results likely do not indicate that students were working to *build rapport* with the agent. CS SEs also had a more negative valence overall than SEs in English. Although prior work indicates that learners can build rapport with an agent through teasing and impoliteness to learning benefit [15], an informal look at the exact student SEs in our study argues against this interpretation: e.g., despite being told to use English with Stacy, one student called her out in Filipino for not understanding a previous SE in Filipino:

S5-(H): "*the meaning of that is 'let's go'. ang gago ka* " [you're stupid]

Negative valence may indicate students saw Stacy as part of an out-group, with the fact that Stacy appears Caucasian and speaks English reinforcing her status as a cultural "other."

Of the remaining proposed explanations in Section 2, we see some evidence for *face threat reduction*, as admissions of uncertainty occurred mainly in Filipino; *frustration with their own knowledge*, as more negative SEs occurred in Filipino; and *frustration with their partner*, as many negative insults directed at the system and its math abilities (despite being learned directly from the student tutor!) were in Filipino.

None of these explanations are purely social – they all interact with student knowledge, self-efficacy, and meta-cognition. This is particularly salient when combined with the finding from this same dataset that negativity was significantly associated with lower achievement [16]. While our future work will include a wider student sample and deeper analyses of language phenomena, even these first results indicate that systems may well benefit from detecting alternate language use as a signal to trigger an intervention at early stages, even if the language cannot be parsed or understood. Such interventions could provide cognitive support, addressing students who may be lost or frustrated, but more sophisticated systems might equally address social concerns that can augment cognitive support in a more human-like, effective way.

Multilingual classrooms can be difficult places to effectively implement language policy even without technology. Here we have discussed the investigation of a personalized learning system with students who may speak a different language at home than the language of school. In a first look at code-switching in a PLS, we saw that students indeed made use of multiple languages to express themselves when interacting with a virtual agent, with evidence pointing towards underlying social rather than simply cognitive factors. The complex and critical nature of these results calls for further study and deeper consideration of learning system designs that not only support learners' cognitive needs, but social concerns as well.

7 Acknowledgements

This work was supported in part by NSF Grant 1252440 and the Pittsburgh Science of Learning Center, funded by NSF Award SBE-0836012. We thank Regina Ira Antonette M. Geli, Aaron Ong, and Gabriel Jose G. Vitug for helping interpret the data, Rex Bringula, Roselle S. Basa, and Cecilio dela Cruz for data collection and logistical support; Kevin Soo and Joel Chan for their advice; and Gierad Laput for translations.

8 References

1. Schneider-Hufschmidt, M., Malinowski, U., & Kuhme, T. (1993). *Adaptive user interfaces: Principles and practice*. Elsevier Science Inc.. (1993).
2. Biswas, G., Jeong, H., Kinnebrew, J. S., Sulcer, B., and Roscoe, R. D. Measuring Self-Regulated Learning Skills through Social Interactions in a teachable Agent Environment. *Research and Practice in Technology Enhanced Learning*, 5(02), 123-152, (2010).

3. Ritter, S., Anderson, J. R., Koedinger, K. R., and Corbett, A. Cognitive Tutor: Applied research in mathematics education. *Psychonomic bulletin & review*, 14(2), 249-255, (2007)
4. Henderson, L. Theorizing multiple cultures instructional design model for e-learning and e-teaching. *Globalized e-learning cultural challenges*,130-153 (2007)
5. Webb, L. and Webb, P. Introducing discussion into multilingual mathematics classrooms: An issue of code switching?. *Pythagoras: Teaching and learning mathematics in multilingual classrooms*: vol 67, p 26-32 (2008)
6. Pinkard, N.: Rappin' Reader and Say Say Oh Playmate: Using children's childhood songs as literacy scaffolds in computer-based learning environments. *Journal of Educational Computing Research* vol 25, p 17-34 (2001)
7. Gilbert, J.E., Arbuthnot, K., Hood, S., Grant, M.M., West, M.L., McMillian, Y., and Eugene, W.: Teaching algebra using culturally relevant virtual instructors. *International Journal of Virtual Reality* vol 7, p 21–30 (2008)
8. Martin, I. P. Diffusion and directions: English language policy in the Philippines. *English in southeast Asia: features, policy and language in use*, p 189-205 (2012)
9. Nye, B. Intelligent Tutoring Systems by and for the Developing World: A Review of Trends and Approaches for Educational Technology in a Global Context. International Journal of Artificial Intelligence in Education. (in press)
10. Qi, D. S. An inquiry into language-switching in second language composing processes. *Canadian Modern Language Review 54*(3), 413-435, (1998)
11. Matsuda, N., Yarzebinski, E., Keiser, V., Raizada, R., William, W. C., Stylianides, G. J., & Koedinger, K. R. Cognitive anatomy of tutor learning: Lessons learned with SimStudent. Journal of Educational Psychology, 105(4), 1152-1163, (2013).
12. Chase, C. C., Chin, D. B., Oppezzo, M. A., & Schwartz, D. L. Teachable agents and the protégé effect: Increasing the effort towards learning. *Journal of Science Education and Technology*, *18*(4), 334-352 (2009).
13. Lester, J. C., Converse, S. A., Kahler, S. E., Barlow, S. T., Stone, B. A., & Bhogal, R. S. The persona effect: affective impact of animated pedagogical agents. In *Proceedings of the SIGCHI Conference on Human factors in computing systems*, 359-366 (1997)
14. Ogan, A., Finkelstein, S., Mayfield, E., D'Adamo, C., Matsuda, N., & Cassell, J. Oh dear stacy!: social interaction, elaboration, and learning with teachable agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 39-48, (2012)
15. Ogan, A., Finkelstein, S., Walker, E., Carlson, R. and Cassell, J. Rudeness and Rapport: Insults and Learning Gains in Peer Tutoring. *Proceedings of ITS 2012*. 11-21.
16. Rodrigo, M. M. T., Geli, R. Ong, A., Vitug, G., Bringula, R., Basa, R., Dela Cruz, C., Matsuda, N., Exploring the Implications of Tutor Negativity Towards a Synthetic Agent in a Learning-by-Teaching Environment. Philippine Computing Journal, vol 8, 15-20 (2013)
17. Ritchie, W. C., & Bhatia, T. Social and Psychological Factors in Language Mixing. Handbook of Bilingualism and Multilingualism, (2012)
18. Sert, O. The Functions of Code-Switching in ELT Classrooms. *Online TESL Journal*, *11*(8), (2005).
19. Bernardo, A. B. I. Language and mathematical problem solving among bilinguals. Journal of Psychology, 136(3), 283-297, (2002)
20. Philippines Dept. of Ed, http://ceap.org.ph/upload/download/201210/1714521500_1.pdf
21. Pavlenko, A. Emotion and emotion-laden words in the bilingual lexicon. *Bilingualism: Language and cognition*, *11*(02), 147-164 (2008).
22. Chi, M.T.H. Self-explaining expository texts: The dual process of generating inferences and repairing mental models. Advances in Instructional Psychology, vol. 5, 161-238, (2000).