**ORIGINAL ARTICLE**                                                              **Open Access**

CrossMark

# A longitudinal investigation into English speaking self-efficacy in a Japanese language classroom

Paul Leeming

Correspondence:
paulleeming@yahoo.com
Kindai University, Higashiosaka,
Kowakae, 3–4–1, Osaka, Japan

## Abstract

Self-efficacy is the belief in one's own ability to carry out a given task, and has been shown to be a powerful predictor of performance. Although researchers have considered self-efficacy within language learning, it remains a relatively underused and unexplored construct. This longitudinal mixed-method study set out to address this, by developing a questionnaire to measure students' English speaking self-efficacy, which was then given to first-year university students on eight occasions over the course of an academic year. Changes in self-efficacy were modeled using Hierarchical Linear Modeling, and potential predictors of change were assessed. The model showed that students grew in self-efficacy, although there were different rates of growth for individuals. Students were interviewed regarding growth in self-efficacy, and possible reasons for change. Students stated that efficacy increased as they became used to the class, but the importance of context as an influence on self-efficacy also emerged.

**Keywords:** Self-efficacy, Speaking, Longitudinal

## Introduction

Self-efficacy (SE), a well-known construct in general psychology, was introduced by Bandura (1977, 1993, 1997) to describe human agency. SE has been shown to predict task engagement and performance, and has become increasingly used within SLA, often as a variable within different motivational frameworks (Mills, 2014). Mills (2014) explained the importance of SE:

> Students with high self-efficacy to perform academic tasks tend to exhibit lower levels of anxiety, display increased persistence when faced with obstacles, exert greater effort, show more flexible learning strategy use, and display higher levels of intrinsic interest in academic tasks. (p.9).

The characteristics described here are those that every teacher attempts to foster in their students, and therefore developing high-self efficacy is logically a priority in the language classroom. In order to achieve this, teachers and researchers first need to understand the way in which SE changes, and also to what students attribute change. In order to facilitate this understanding, the research described in this paper set out to investigate how speaking self-efficacy changes over the course of an academic year, and

variables that may predict this change. Students views on self-efficacy were also investigated.

Bandura (1997) argues that the strength of SE as a predictor variable comes from its specificity, and with such a great difference in the cognitive demands placed on learners by each of the fours skills, it is necessary to develop measures of SE that are specific to each skill. The research therefore also sought to measure English speaking self-efficacy, which is relatively unexplored in the field of language learning. Language learning is an incredibly challenging task, and so many students fail to achieve success. If teachers are able to foster positive feelings of efficacy, students will persist, will invest more effort, and ultimately reach higher levels of proficiency.

## Background

### Self-efficacy

Bandura (1997) described SE as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p. 3). This is the definition adopted for the current study. SE has been shown to predict performance, and influences the initial decision to undertake an action, the amount of perseverance shown, and the ability to control affective influences during the task.

SE has several sources, the most influential of which is personal experience, described by Bandura (1997) as "enactive mastery experience" (p. 80). If an individual relates to the task in question and has succeeded in completing similar or identical tasks then SE will be high. Vicarious experience is the second greatest source of influence. SE arises when people watch peers whom they deem to be similar to them, successfully perform the task in question. A third source of SE is peer influence through encouragement. The final source of SE discussed by Bandura (1997) is affective and comes from our mental and emotional states. According the Bandura's (1997) theory, there is a cyclic relationship, with positive experiences leading to greater SE, which in turn leads to more commitment to subsequent tasks, and a greater likelihood of success. Therefore in the classroom, if a teacher is able to provide challenging and yet positive learning experiences, SE should experience growth.

### Self-efficacy and SLA

SE has become increasingly studied as a variable within SLA at the start of the twenty-first century (Mills, 2014). A substantial body of work has investigated teachers' feelings of self-efficacy, and how this impacts students' performance (Swanson, 2012, 2013, 2014). Research has also focused on SE within a motivational framework, or within self-regulation and learning strategies (Chang, 2010; Dörnyei & Otto, 1998; Lee, Yu, & Liu, 2017; Wang & Bai, 2016).

Considering SE and performance, Mills, Pajares, and Herron (2006) investigated the relationship between listening and reading SE, anxiety, and reading and listening ability with 93 college students studying French as a foreign language. They found that SE had a positive relationship with reading, but anxiety had no relationship. SE also had a positive relationship with listening ability, but only for female students. The authors suggested that this may be due to men feeling less relaxed when learning foreign languages, but may also be a result of problems with measurement which relied on

self-reported data. The analyses were correlational and therefore causal relationships cannot be determined, but the authors ended the paper highlighting the importance of SE, arguing that knowledge of students' efficacy beliefs helps teachers target those views directly through teaching, and can have a positive impact on students' subsequent performance.

In a later study in a similar context also investigating performance, Mills, Pajares, and Herron (2007) asked 303 college students studying French to express their efficacy beliefs that they could attain certain grades in a French foreign language class, and also their efficacy beliefs related to their ability to effectively study for the course. The researchers were interested in the interaction of efficacy beliefs regarding attainment in the course, and also in the ability to use strategies to study effectively. Results showed that belief in ability to study, rather than belief in ability to achieve a certain level of performance, was a greater predictor of the actual grade achieved by students. Gender was also investigated, and showed that women tended to have higher feelings of SE. The authors concluded by encouraging teachers to foster greater feelings of SE in students due to the benefits for students' performance in language courses.

Research has considered use of strategies and their relation to SE (Graham, 2007; Magogwe and Oliver, 2007), but has particularly been investigated within motivational research. Dörnyei and Otto (1998) included SE as a part of the process model of motivation, believing that SE will influence students' initial decision to begin a given action. Chang (2010) investigated the correlation between motivation, which she measured as SE and autonomy, and group processes, which she described using cohesion and norms. Analysis of the data showed that group processes were weakly related to some aspects of L2 motivation.

Related to motivation, Hsieh and Kang (2010) examined the relationship between attribution and SE in a Korean context. They asked high school students to give general ratings of SE regarding English after receiving their test grades. Results suggested that students high in self-efficacy were more likely to attribute course success to internal factors, and self-efficacy was positively correlated with achievement measures. Among students who were unsuccessful on the test, those with high SE attributed the results more to personal control. The researchers stressed the importance of teachers monitoring students' beliefs, and also directly intervening to change views on the causes of success or failure. Unfortunately, the outcome variable of test performance was self-reported, and therefore some caution is advised when considering the results.

There have been a number of studies that have considered SE from a more longitudinal perspective. Mills (2009) conducted a longitudinal study investigating the effects of project-based learning on the SE of beginner level French students. She measured gains in the five key goals for the course; communication, cultures, connections, comparisons, and communities, and found that students made gains in SE in all the areas measured. Mills (2009) argued that project-based learning is an effective method of increasing the students' efficacy, and that the feedback given throughout the course was an important factor in the increases in students' SE. The author did acknowledge that there was no control group, and therefore it was impossible to state with certainty what was responsible for the increase in SE.

Piniel and Csizér (2015) also used a longitudinal approach to investigate changes in motivation, anxiety, and SE among university students studying academic writing. They

adopted a dynamic systems approach to changes in SE (Dörnyei, MacIntyre, & Henry, 2014; Larsen-Freeman & Cameron, 2008). Using a battery of measures in a mixed method study, the researchers measured motivation, SE and anxiety, six times over the course of a 14-week semester. They used nine items to measure the writing SE of students, and in their analyses considered the data from 21 students. Results for SE showed that there was a linear decline over the 14-week course, although retrospective qualitative data from students suggested that they perceived their SE in writing to have increased. The authors attributed the difference between qualitative and quantitative results to the fact that the qualitative data was retrospective, allowing the students to look back over the entire course and see the improvements that had been made. According to Bandura (1997), positive experiences should lead to greater SE, which in turn should lead to greater effort, leading to yet more positive experiences of learning. It is of interest that this did not seem to be the case in this study of writing SE.

No studies within SLA have used individual difference variables to predict change in SE, but Bandura (1997) claims that the largest influence will be task related proficiency. That is, an individual's ability to perform a specific task. In the language classroom, this would relate to an individual's ability with regards the English skill required for a given task. Aside from ability, personality has also been shown to influence levels of SE, with high anxiety leading to lower levels of SE (Gist & Mitchell, 1992). Particularly in a communicative language classroom, students who are extraverted would be assumed to be less anxious when performing speaking tasks, and therefore should have higher levels of SE. As mentioned above, in language learning contexts gender has also shown to be important, with women generally displaying more SE than their male counterparts.

### Motivation for the current study

Mills (2014) has argued that the two greatest problems with research into SE are that measures used are not specific, or measures of SE are combined with other questions measuring different variables, thus complicating interpretation. Also, there are only a small number of longitudinal studies of SE within SLA (Piniel & Csizér, 2015), and limited research related specifically to speaking SE. There has also been a lack of qualitative data to ascertain how students perceive SE. Bandura (1997) argues that the strength of SE as a predictor comes from its specificity, and therefore research is needed investigating speaking SE. As any language teacher knows, students' feelings of efficacy towards reading in the second language are often vastly different to their feelings of efficacy regarding speaking. If researchers are able to understand how SE changes over the course of an academic year, and any variables that may predict growth, we can attempt to influence students SE, and subsequent performance in the language classroom.

The current study set out to answer the following three questions:

1. How does English speaking SE change over the course of an academic year?
2. What individual difference variables predict changes in SE?
3. To what do students attribute changes in SE?

According to SE theory, positive experiences lead to greater SE and it is therefore hypothesized that students will make gains in speaking SE over the year, as the course is

designed to foster English speaking SE. It is also hypothesized that English ability will be a positive predictor of SE gains, as the development of SE is contingent on positive experiences, and also cyclic, with positive experiences leading to greater SE which in turn leads to positive experiences. Extroverts should also benefit from the social nature of an oral English class, and based on prior research, female participants should experience higher SE. The students' views on SE were unknown prior to the study, and therefore there was no hypothesis for the third research question.

## Method

### Participants

The data reported on in this paper is part of a larger mixed-methods, longitudinal study investigating small group work in the language classroom. The participants ($n$ = 77, 23 female and 54 male) were enrolled in a compulsory first-year oral English course in the science department of a private university in Japan. The first-year students were in three different classes, and were all native speakers of Japanese. Although all students had six years of formal English education, it is generally focused on grammar and reading in order to pass university entrance exams. English speaking experience varied, with some students having experienced oral English classes in high school, a small number of students having lived abroad, and some students attending English conversation school. The average TOEIC score was 390, meaning that that the average student was of upper beginner to lower-intermediate level (ACTFL level 0+ to 1, or A2 in the CEFR), although there was a wide range of abilities in each class, as students were grouped according to major. Participation in the study was optional, and students were given a detailed explanation of their rights regarding subsequent use of any data.

The oral English course ran for one academic year, with the first semester running from April to July, and the second semester from September to January. Each semester was 14 classes in total, with one 90-min class per week. Aside from the oral English class, students also took classes in reading and writing, each for 90 min once a week. The researcher taught all the students in this study for writing and oral class for a total of 3 h per week. The researcher was a native-speaker of English, with over 12 years of teaching experience in the context for this study. The oral English classes followed a Task-Based Language Teaching approach (TBLT) (Willis & Willis, 2007), with students working together in small groups to complete simple tasks that often required some kind of feedback or report to the entire group. All data for the study was gathered in the oral English classes.

Generally, as the participants were quite limited in ability to speak English, the focus was on getting them to maintain simple conversations related to personal experience (e.g. last vacation, family), or to give simple opinions including likes and dislikes. The aim was to provide students with ample opportunity to practice speaking in small groups, and to give explicit feedback regarding gains made.

The research was non-experimental in design and several steps were taken to actively foster speaking SE in this course. Students were given positive experiences of success with a variety of simple speaking tasks. With an emphasis on communication rather than grammatical accuracy, students were told that if communicative objectives were achieved then the task was a success. Students were also given regular positive

feedback. Correction was generally provided to the entire class, and focused on positive examples of language use from students.

Assessment was based on in-class performance, and particularly on mid-term and final speaking tests each semester, which accounted for 40% of the students' grades. The tests were administered in groups of three or four people, with students given a simple topic such as "plans after graduation", or "the best age to get married". Students then had to make conversation for 10 min, without any use of the first language. Tasks were graded to become more difficult as the academic year progressed, moving from personal topics to more abstract topics that required students to give their opinion. Students did not receive a detailed breakdown of their grades, but a final overall grade, approximately one month after the end of each semester.

### Quantitative measures

The SE measure was designed to measure individual students' SE with regard to specific tasks in the oral communication course (See Appendix 1 for the full instrument). It was designed following the guidelines provided by Bandura (2006), using "can do" statements that were designed to be specific to the context. As Bandura (2006) explains, it is the belief in the ability to do something that is central, and this distinguishes the measure from constructs such as Willingness to Communicate (WTC), which typically ask about plans to perform speaking tasks (see Peng & Woodrow, 2010, for an example). A search of the literature revealed several measures of SE in SLA research, but the measures did not relate specifically to English speaking SE (see Rahimi & Abedini, 2009, for an example focusing on listening comprehension). The Motivated Strategies for Learning Questionnaire (MSLQ), originally developed by Pintrich and DeGroot (1990), was subject-specific but could be easily adapted. Although the general questions were largely appropriate for the course, three items were deleted as they asked students to make comparisons with peers, which is criticized by Bandura (1997). One of the other items was adapted as it referred to two aspects of performance in one question. For all the items the wording was changed to simple can-do statements. Items 6 and 7 were added, relating to more specific tasks in the course students were taking.

The SE measure was piloted with 128 students from a previous cohort, and analyses using Winsteps computer software (Linacre & Wright, 2007) showed that the items fit the Rasch model for measurement (Bond & Fox, 2007). All of the questions were related to the ability to perform well in the class with respect to the final grade, and to accomplish specific tasks in the oral communication course. An example item is *I can speak English fluently when taking part in a group discussion*. Students respond to each statement using a six-point Likert-scale ranging from *Not at all true* to *Very true*. Rasch analyses of the eight different administrations of the measure showed that the item reliability (analogous to Cronbach's Alpha) was high, with values ranging from .92 to.98. Results of Rasch Principal Components Analysis (PCA) supported the claim that the instrument was unidimensional.

The instrument was translated into Japanese for use in this study, and back-translated by a Japanese researcher familiar with the research. The website *Survey Monkey* was used to administer the questionnaires, which were given four times each semester, following the timetable outlined in Table 1.

**Table 1** Timetable for the study

| Semester 1 | | Semester 2 | |
| --- | --- | --- | --- |
| Week | Measures | Week | Measures |
| 2 | SE (April), IPIP(Extroversion), Dictation | 2 | SE (September) |
| 5 | SE (May) | 5 | SE (October) |
| 7 | Mid-semester speaking test | 7 | Poster presentation |
| | | 8 | Mid-semester speaking test |
| 10 | SE (June) | 10 | SE (December) |
| 13 | End of semester speaking test | 13 | Poster presentation SE (January) |
| 14 | SE (July) | 14 | End of semester speaking test |
| Interviews: Conducted after each semester with eight students | | | |

*SE* Self efficacy questionnaire

Although all of the students in the school of science were required to take the TOEIC test, many students had been observed to sleep through the test, making interpretation of scores difficult. In order to overcome this problem, an additional measure of English ability was developed. Dictation has been shown to be an accurate and efficient measure of language ability (Cai, 2012; Oller & Streiff, 1975), and was deemed appropriate in that it is primarily a measure of aural skills, with the ability to comprehend, and reproduce spoken English being paramount. Other advantages of dictation were that it could be tailored specifically for the level of the students in the study, and it was relatively easy to administer and grade. The dictation was marked by the author and a research assistant, and any discrepancy in scoring was discussed. The inter-rater reliability was.98. A Rasch analyses was performed to assess dimensionality, and functioning of the items, and the dictation test was found to satisfy the requirements of measurement for the Rasch model. The item reliability was.88.

The extroversion dimension of personality was measured using the IPIP (Donnellan, Oswald, Baird, & Lucas, 2006; Gow, Whiteman, Pattie, & Deary, 2005) which was available in Japanese, and has been used in a Japanese context. This instrument was administered in the second week, again using the online questionnaire format. Rasch analyses were conducted on the extroversion questionnaire and all items displayed adequate fit to the Rasch model for measurement (Linacre, 2007). The item reliability was.85. As with the dictation, scores were converted to logits for use in the subsequent analyses.

### Interviews

Although the study primarily focused on quantitative measures, it was mixed-method and included qualitative data. The research design was a Concurrent Embedded Strategy (Creswell, 2009), where collection of qualitative and quantitative data occur at the same time, and do not impact each other. The central focus in the current study was tracking the changes in SE over a year of English study, but it was hoped that qualitative data would add some insights into student perceptions of change and SE in general. Therefore, in addition to the quantitative measures used, eight students were interviewed at the end of each semester (total of 16 interviews). Participants were selected based on criteria for the larger study mentioned previously. Interviews were semi-structured (McDonough & McDonough, 1997), and were conducted by the

researcher in the students' first language (see Appendix 2 for an outline of the interview procedure and questions relating to SE).

Although the interviews were conducted after the final class in each semester, students had not yet received their grades for the course and there was the danger that this would influence their responses. All interviewees were assured that the interview would have no bearing on their grade for the course, or any subsequent treatment in class. Students seemed to welcome the opportunity to give direct oral feedback to the teacher, which is generally limited in tertiary educational contexts in Japan, and seemed frank and open regarding their views, willing to criticize the way in which groups were organized, and the class was conducted.

The interviews were recorded and subsequently transcribed. Data were analyzed following interpretive analysis, as described by Hatch (2002). The data were revisited several times, and sections that supported the interpretations made were translated into English by the author, for inclusion in this paper.

### Analyses

All of the data from the eight administrations of the SE measures were subject to Rasch analysis to ensure that the questionnaire was unidimensional, and was a valid and reliable measure of SE. (A description of Rasch analysis is beyond the scope of this paper, but see Bond and Fox (2007) for a comprehensive introduction.) All of the values for the SE are given in logits, with high positive values indicating high SE. The data files were stacked for Rasch analysis in order to allow for the direct comparison of data without reducing the reliability (Wright, 2003).

The growth model for SE was constructed using Hierarchical Linear Modeling (HLM) (Raudenbush & Bryk, 2002). In the case of the growth model used in this study, growth is hypothesized to occur within individuals at level 1, and various individual differences variables are added to the model at level 2 to assess their impact on growth. The benefits of HLM over more traditional methods of analyses are discussed in some detail by O'Connell and McCoach (2008). Of particular relevance to this study, HLM allows for the correlation between results in a repeated measures design. Issues of power in HLM are complex, depending on the level of the effects, and fixed or random effects for variables. Tabachnick and Fidell (2007) stated that for sufficient power to discover cross-level effects, we need at least 20 measurements at the second level, with a minimum of three measurements at level 1. The current study had eight measures at the first level (repeated measures), and 78 at the second (people), and therefore meets the criteria suggested. Unequal sample size at each level does not pose a problem for HLM, and missing data is also tolerated.

### Results

#### Descriptive statistics for self-efficacy

The descriptive statistics for SE measures are shown in Table 2. The results are in logits obtained from the Rasch analysis, and show that the measures of SE have normal distributions. Figure 1 shows the development of SE over the academic year and indicates growth. All of the means of the eight administrations were negative, indicating that the students generally had low SE, and found it difficult to strongly endorse any of the items in the questionnaire.

**Table 2** Descriptive Statistics for self-efficacy

|  | Mean (confidence intervals) | SE | SD |
| --- | --- | --- | --- |
| *SE April* | −2.75 (−3.35, −2.16) | .30 | 2.64 |
| *SE May* | −1.49 (−2.10, −.87) | .31 | 2.72 |
| *SE June* | −.75 (−1.43, −.07) | .34 | 3.03 |
| *SE July* | −.47 (−1.17, .22) | .35 | 3.08 |
| *SE September* | −1.06 (−1.55, −.57) | .25 | 2.18 |
| *SE October* | −.70 (−1.23, −.17) | .27 | 2.36 |
| *SE December* | −.09 (−.61, .43) | .26 | 2.29 |
| *SE January* | −.07 (−.49, .63) | .28 | 2.49 |

Fig. 1 shows that initially students had negative feelings of SE, but were able to achieve growth. There was a two-month summer break before the second semester began, and during this time there was a slight fall in SE, as noted by the lower score in September. Growth in the second semester was positive, although a little less steep than in the first semester.

### Growth model for self-efficacy

Research question 1 was interested in how SE changed over the course of the academic year. A two-level growth model was constructed in HLM to test individual differences in growth trajectories, and also variables that may predict changes in SE.

The level 1 model was constructed with SE at times 1–8 as the dependent variable with time as the only predictor in the model.

$$SESTACK_{tij} = \pi_{0j} + \pi_{1j} * (TIME_{ij}) + e_{tij}$$

This model assesses whether there are individual differences in growth over the course of the year. $SESTACK_{tij}$ represents the outcome at time t in SE for student *i*. $\pi_{0j}$



**Fig. 1** Group mean for self-efficacy measures over the academic year
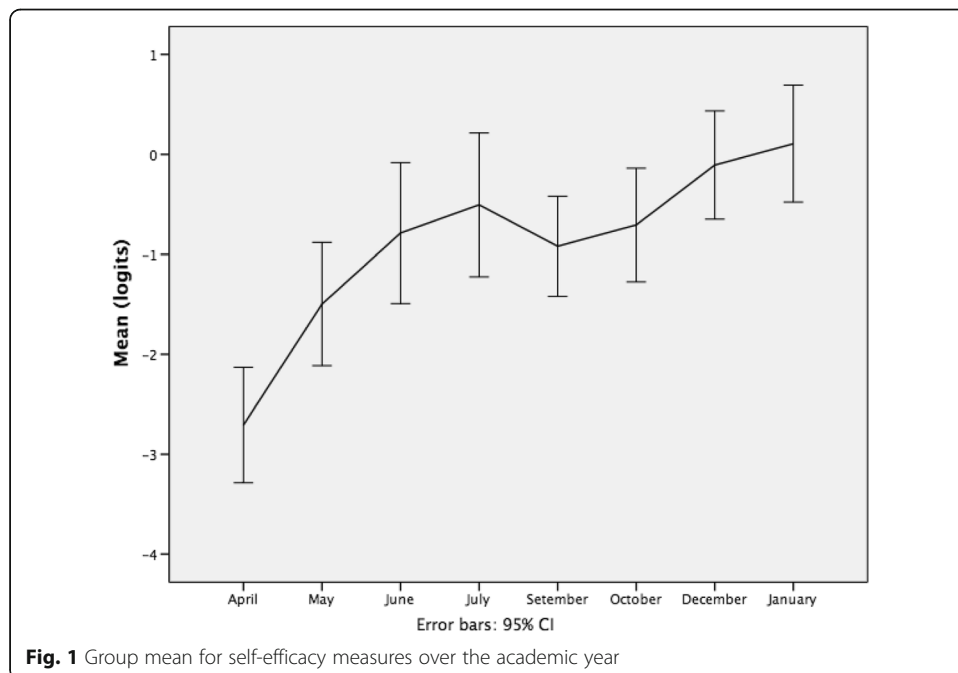
represents the initial intercept for student $i$, while $\pi_{1j}{}^{*}$ ($TIME_{ij}$) represents the growth for student $i$ based on the eight separate time measurements for SE. The final term, $e_{tij}$ represents an error term *at time t* associated with student $i$. At level 2 we have the individual differences language ability (PROF), extroversion (EXT), and gender (GENDER) interacting with both the initial intercept and time in order to ascertain whether these variables have any impact on initial ratings and growth in SE. Level 2 equations incorporate these variables.

$$\pi_{0ij} = \beta_{00j} + \beta_{01j}\,(\text{PROF}) + \beta_{02j}\,(\text{EXT})r_{0ij} + \beta_{03j}\,(\text{GENDER}) + r_{0ij}.$$

$$\pi_{1ij} = \beta_{10j} + \beta_{11j}\,(\text{PROF}) + \beta_{12j}\,(\text{EXT}) + \beta_{13j}\,(\text{GENDER}) + r_{1ij}.$$

The initial state of SE is represented by $\pi_{0ij}$, and is predicted by the variables dictation, $\beta_{01j}$ (PROF), extroversion, $\beta_{02j}$ (EXT), and gender, $\beta_{03j}$ (GENDER). $\pi_{1ij}$ represents the growth rate in SE during the first semester, and this is predicted by the variables dictation, $\beta_{11j}$ (PROF), extroversion, $\beta_{12j}$ (EXT), and gender, $\beta_{13j}$ (GENDER). Gender was dummy coded (0 = female, 1 = male) for the model.

As with all HLM analyses, the first analysis was the unconditional model that seeks to determine whether there is any evidence of growth or change over time. For this analysis there was a growth model at level 1 and no predictors were added at the second level. The results for the unconditional model show are shown in Tables 3 and 4.

The results in Table 3 show that there is statistically significant variation in the initial status of SE, and also that there is significant growth for individuals ($\chi^2 = 177.53$, $p < .01$).

Table 4 shows that the average student began with negative SE ($-1.96$), and gained .30 logits for each occasion of measurement of SE. The growth was also statistically significant.

The next HLM analysis was used to investigate the differences in individual growth by adding individual difference variables to the model at level 2. As mentioned, initial ability in English and extroversion were hypothesized to predict growth in SE and were therefore added to the model. Gender has been shown to lead to significant differences in SE in previous studies (Mills, 2007), and was also selected for inclusion in the model.

The final estimation of level 1 and level 2 variance was significant for both initial SE and growth in SE, with a $\chi^2$ value of 177.53, $p < .01$ for growth. The results for the three level two predictors on both initial status of SE and growth in SE are shown in Table 5.

The results in Table 5 show that extroversion, English ability and gender predict the initial status of SE. All three variables have positive coefficients, suggesting that male students who are more extroverted and proficient than average began with higher ratings of SE. Ability was the only significant variable influencing the growth in SE over the course of the year, and with a negative coefficient of $-.10$, suggests that higher ability students made lower gains in terms of SE.

**Table 3** Estimation of variance components for initial state and growth in self-efficacy

| Random effect | SD | Variance component | df | $\chi^2$ | p |
|---|---|---|---|---|---|
| Initial status, r0 | 2.37 | 5.63 | 76 | 580.82 | < .01 |
| Growth rate, r1 | 0.26 | 0.07 | 76 | 177.53 | < .01 |
| level 1 error, e | 1.43 | 2.03 | | | |

*SE* Self-efficacy

**Table 4** Estimation of Baseline for Initial Status and Growth Rate in SE

| Fixed effect | Coefficient | SE | t-ratio | df | p |
|---|---|---|---|---|---|
| Baseline SE initial state, $\gamma_{00}$ | −1.96 | .29 | −6.74 | 76 | < .01 |
| Baseline SE growth rate, $\gamma_{100}$ | .30 | .04 | 7.87 | 76 | < .01 |

*SE* Self-efficacy

Figure 2 shows the relationship between growth in SE and ability. Although low-ability students began the study with considerably lower feelings of SE with regards to speaking English, by the end of the study they had made greater gains and surpassed the feelings of SE of the more capable students.

### Student views on self-efficacy

Students were asked to reflect on their feelings of SE, with reference to the questions that they had answered eight times over the study. Transcriptions of interviews were examined for themes that emerged regarding student views on SE. Students focused on acclimatization, English ability, and the influence of context, and each of these are addressed below.

#### Acclimatization

Many of the students admitted to having almost no prior experience of oral English classes, and particularly classes taught by a native-speaker of English. As a result of this they expressed initial unease, and were worried as to whether they would be able to understand the teacher and speak English. The students were also in their first year at university and so were not only unfamiliar with this class, but also the general requirements of the university. As time progressed they came to understand the course, the university, and also the teacher, and this resulted in increased feelings of English-speaking SE. This can be seen in the excerpt from Ryo below:

Excerpt 1 Ryo Semester 2 interview

"*It was the first time (to have an oral English class) so I was worried at first but I felt that it was fun and that I would be able to pass. I got used to it and felt that it was fun to speak English.*"

Students were generally nervous, but came to feel that the lessons were fun and that speaking English was a positive experience. Students also stated that they became

**Table 5** Estimation of significant effects of individual differences on initial status growth of self-efficacy

| Fixed effect | Coefficient | SE | t-ratio | df | p |
|---|---|---|---|---|---|
| PROF on initial status, $\gamma_{01}$ | .80 | .11 | 7.06 | 534 | < .01 |
| EXT on initial status, $\gamma_{02}$ | 1.00 | .09 | 10.06 | 534 | < .01 |
| GENDER on initial status, $\gamma_{03}$ | 2.46 | .28 | 8.72 | 534 | < .01 |
| PROF on growth rate, $\gamma_{11}$ | −.11 | .04 | −2.45 | 73 | .02 |
| EXT on growth rate, $\gamma_{12}$ | −.02 | .04 | −0.62 | 73 | .54 |
| GENDER on growth rate, $\gamma_{13}$ | −.11 | .11 | −0.96 | 73 | .34 |

*PROF* = Score on dictation test; *EXT* = Score for extroversion

**Fig. 2** Effect of English ability on individual gains in SE
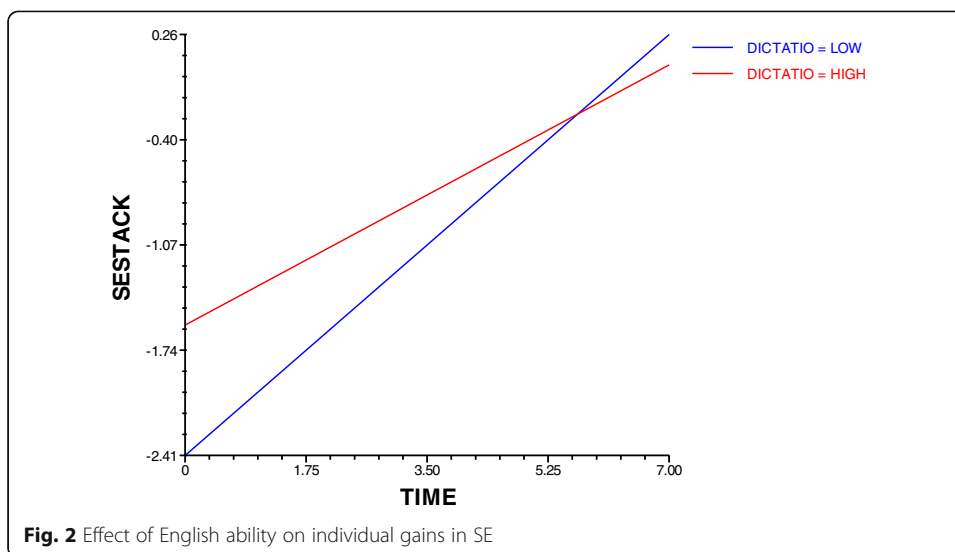
accustomed to the other members of the group, and this gave them greater feelings of efficacy with regards making English conversation.

Although several students mentioned the positive impact on SE from other group members, one student directly mentioned the impact that the teacher had on his own SE.

Excerpt 2 Takuma Semester 1 interview

"*At first it (SE) was low. I thought I would need grammar...that I would have to talk properly. But you said we don't need grammar.*"

Students in this context have studied English grammar for six years in order to pass the often technical and demanding university entrance exams, and therefore have a good understanding of grammar, but little or no opportunity to use English in communicative conversational settings. This means that students often focus on producing grammatically accurate sentences at the expense of fluency. The focus in the class was on developing oral fluency and therefore the teacher regularly reminded students that grammar was not being assessed. This reminder clearly had an impact and meant that students felt more relaxed, leading to greater feelings of confidence and speaking SE.

### Increased English speaking ability

Eight of the students interviewed (half of the total number) mentioned an increase in efficacy speaking English, and three students explicitly stated that their English speaking ability had increased over the course of the semester or academic year. As mentioned previously, students have studied English for a minimum of six years, but generally opportunities to speak English are limited, and therefore given the opportunity to practice students can experience positive growth. Interestingly, the students seemed to differentiate between the ability to speak English, and overall English ability as shown in the excerpt below.

Excerpt 3 Yuki, Semester 1

"*My speaking ability has improved but my English ability hasn't really improved.*"Another student claimed that her English speaking ability had not increased, but it had

become easier to talk in English. In an EFL context such as Japan, students are often judged on their English ability through tests such as TOEIC, which have a listening component but no speaking, or through university entrance exams, which also have reading and listening, but currently do not have speaking ability assessment. As such, students seem to have separated the ability to speak and communicate using English from general English ability, which they consider to be metalinguistic knowledge of grammar and vocabulary. As a result, students do not consider their overall English ability to have improved, despite recognizing that their speaking had improved. Implications are discussed in the next section.

### Contextual influences

A final theme that emerged from the interviews was the influence of specific contextual factors on feelings of SE. Five of the 16 students commented in some detail on specific influences. Two students mentioned factors with only limited relation to the classes themselves. Shigero was interviewed in the second semester, and stated that his SE had declined due to the demands and pressure that he had from his university rowing club. He was a very active member of the club, and was expected to attend practice sessions on an almost daily basis. This seriously limited his time available to study, and he cited this as the main reason for a decline in his SE, as he was aware that his commitment to study had directly suffered as a result of club activities. Another student interviewed in the second semester, Yoichi, stated that attendance had been a problem in the second semester, and being absent from class had negatively impacted his feelings of efficacy. He was worried that his absence from class would leave him struggling to perform the tasks and pass the course. Yoichi had failed in the first semester due to absence from one of the speaking tests.

Other students discussed factors more directly related to the classes. One student noted that if a task had been particularly taxing, then his confidence in his speaking was negatively affected.

Excerpt 4 Fumiyo, Semester 2
"*My self-efficacy changes depending on the (weekly) theme. It depends if I have or don't have the appropriate vocabulary.*".

Task-specific factors could have quite a large influence on how students feel about SE, and is in line with the findings of Yashima, Ikeda, and Nakahira (2016), who discovered that the topic of conversation had a considerable impact on the talk time and participation of Japanese students in an oral English class discussion. Another student mentioned that their feeling changes week by week, and one attributed change to the group context. Shota claimed that he received confidence in speaking from the other members of his group and that had he been in a different group, the result could have been very different.

### Discussion

The current study attempted to address problems with measurement of SE noted by Mills (2014), by creating a questionnaire focused solely on speaking SE, and then using the Rasch Model to analyze the questionnaire data, before conducting subsequent analyses. Rasch analyses of the eight different administrations of the SE measure indicated

that the measure was unidimensional, and conformed to the Rasch model for effective measurement. Results suggest that the measure was suitable, although students did initially have low ratings of SE, suggesting that the items were difficult for students to endorse.

The HLM growth model showed that students grew in SE over the course of the year. This is in line with the findings by Mills (2009), who found that SE grew in her project-based learning course. She attributed this to the course, but in the current study growth may have occurred for several reasons. First, students' initial levels of self-efficacy were low, as indicated by the negative value for the mean of the first administration. This low beginning makes it easy for students to make gains. Furthermore, students in this context have had six years of formal English education that generally is geared towards passing university entrance exams. These exams are focused heavily on reading and grammar, and therefore many students have little experience of oral interaction in English. This means that they are quite nervous when they initially engage in classroom speaking tasks, but due to their reasonably large receptive knowledge of English, once they begin speaking they are able to make large gains in a relatively short period. Students need to feel a sense of mastery (Bandura, 1997), and the course was designed to allow students to gradually build up the conversational skills to take part in a ten-minute group discussion. Students achieved this, and were all given regular positive feedback regarding performance, another recognized source of SE (Bandura, 1997). As all of the students were in their first year at university, there is also the possibility that they may have grown accustomed to the course, and the requirements from the teacher, leading to more positive self-appraisals. Interview data suggested that acclimatization to both the university context and the teacher were factors in leading to increases in SE.

There were significant differences among individuals both in initial rates of SE and growth in SE. Initial differences in SE were predicted by English ability, extroversion, and gender. The fact that English ability predicts initial speaking SE is in line with the theory of SE proposed by Bandura (1997), who claimed that the greatest source of SE is prior success with the same or a similar task. The role of extroversion is a little more difficult to understand, and may be that many of the items on the SE questionnaire relate to interacting with others and working together as a group. Activities such as giving a presentation, and participating in a group discussion are likely to be perceived as easier by extraverted students who enjoy interaction, leading to greater initial ratings of SE. The finding that male students were higher in initial ratings of SE is also in contrast to prior findings (Mills et al., 2007), where females had higher SE. As most groups were male dominated due to the imbalance of gender in the school of science at this university, it may be that female students felt less confident in their ability to actively take part in group discussions, as male students may tend to dominate interactions.

The growth model showed that ability was the only significant predictor of growth, with a negative coefficient showing that students who are higher in English ability struggled to make gains in comparison to students with low ability. Students with high ability began with higher ratings of SE, but by the end of the academic year, their self-ratings of SE were overtaken by the students with lower English ability (see Fig. 1). One possible explanation is that the lower ability students were able to notice their own gains in speaking more, and therefore were able to increase in SE during the course of

the year. Students who were already proficient may not have felt that they made gains, and therefore increases in SE were more limited. This possibility supports Mills (2014) claim that mastery experiences can have both positive and negative influences on SE. For mastery to have a positive influence on SE, students must perceive the task to be challenging and then they can feel a sense of accomplishment upon successful completion. The implication for teachers with mixed-ability classes, such as those in this study, is that they should attempt to provide meaningful experiences of success for students of all levels, and give personal feedback regarding task-performance.

Examination of Fig. 1 reveals that the students made gains over the course of the year, but there was a drop at the start of the second semester, and only by December had students recovered and surpassed SE ratings at the end of the first semester. Students experience a long break in between the first and second semester, with almost two months without classes, and in a largely monolingual context such as Japan, that means that the majority of students had little or no chance to speak English during that time. The implication for teachers is that students should be provided with some opportunity to speak English between semesters. This may be providing special summer courses, or encouraging students to engage in self-study.

Interview data suggested that students generally felt an increase in SE over the course of the semester or academic year, and many attributed this to familiarization with the university, the course, and the group that they were in. Students felt that their speaking ability had increased, but some seemed to distinguish between speaking English and general English ability. In an environment where the focus is heavily on passing tests of grammar and vocabulary, speaking often is given a backwards role, and not emphasized as important. This seems to be reflected in the students' views on speaking. It is important for teachers to emphasize that speaking is central to language ability, and speaking is the skill that is most often used to judge English ability. As Mills (2014) suggests, one role of the teacher is to attempt to change the language learning beliefs of students, and teachers need to make students aware of the central role of speaking in communication, in order to ensure that oral English classes are treated as having value for students.

In line with a dynamic systems theoretical framework (Dörnyei, MacIntyre, & Henry, 2014; King, 2015), the current study highlighted the importance of contextual factors, and showed that language learners are subject to a myriad of influences, including some beyond the classroom. Factors such as club, and attendance can influence how students feel about their SE, as well as more classroom specific variables such as the nature of the task, or the composition of the group. This line of thinking suggests that researchers need to consider context when examining changes in motivation, but also go beyond immediate context to consider influences outside the classroom itself.

There were a number of limitations to the current study. First, convenience sampling was used, and the researcher was the teacher of the three classes, which may have influenced student responses to the questionnaires and interview. Administration of the questionnaire on eight separate occasions also means that students had the opportunity to become familiar with the items, and may have guessed the intentions of the researcher, influencing results. A further limitation is that although the study shows increases in SE, there is no behavioral outcome to show that increased SE has an impact on the classroom practices of students, although research has shown that SE predicts

performance. Future studies should attempt to show how SE directly impacts the way in which students interact in the language classroom. As there was no control group, it is impossible to make strong claims regarding the reasons for changes in SE, and growing accustomed to university may have been a factor. The students were also quite low in ability and SE, and therefore completing simple conversational tasks in English had a large influence on their beliefs. Higher ability students may find it more difficult to sense that their ability is increasing over a short time frame.

## Conclusion

The significant growth in speaking SE over the course of the year is a positive finding for language teachers, and suggests that students can make gains, if given a chance to practice language and experience mastery in the classroom. Mills (2014) summarized research into SE in SLA, stating that that "a multitude of findings from the last decade highlight the critical importance of developing the self-efficacy beliefs of language learners to ensure that learners feel competent and capable in their ability to acquire a FL (p.19)." With the growth in SE in the current study, the students should be more successful in the difficult task of language study.

Despite the limitations, I believe that the current study contributes to our understanding of SE, and how it can change over time. Longitudinal data effectively shows that students are able to make gains in SE over a relatively short period, and that as teachers it is possible to foster feelings of SE within students, potentially leading to greater levels of success with language learning. Future studies should attempt to show how changes in SE relate to gains in language proficiency.

## Appendix 1

### English communication self-efficacy measure

How well do you believe you will perform in this class? Rank yourself on the following statements. 1 (not at all true), 2 (slightly true), 3 (somewhat true), 4 (quite true), 5 (true), 6 (very true):

 i. I can understand what is taught in English Communication Class.
 ii. I can do very well in English Communication Class.
 iii. I can learn the material for English Communication Class.
 iv. I can use my study skills well in English Communication Class.
 v. I can receive a good grade in English Communication Class.
 vi. I can speak English fluently when giving a presentation in front of the class.
 vii. I can speak English fluently when taking part in a group discussion.
 viii. I can give my opinion in English.
 ix. I can enjoy conversation in English.

## Appendix 2

### Interview procedure and questions

- 8 students each semester
- Induvial interviews conducted in Japanese
- Interviews after the final-class

- 20–30 min, recorded

Interview Questions

1. Introduce purpose-consent form. Confidential. No right or wrong answer.
2. Background information-self-introduction. English study to date.
3. Changes in self-efficacy. Ask student to explain changes in self-efficacy over the semester (or year). Ask reasons for change/no change.
4. Show students actual data of self-efficacy changes. Re-elicit reasons for change.
5. Any questions/comments?

**Competing interests**
The author declares that they have no competing interests.

## Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### References
Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215.
Bandura, A. (1993). Perceived self efficacy in cognitive development and functioning. *Educational Psychologist, 28*(2), 117–148.
Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W.H. Freeman.
Bandura, A. (2006). Guide to the construction of self-efficacy scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (Vol. 5, pp. 307–337). Charlotte: IAP.
Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). New Jersey: Routledge.
Cai, H. (2012). Partial dictation as a measure of EFL listening proficiency: Evidence from confirmatory factor analysis. *Language Testing, 30*(2), 177–199.
Chang, L. Y. H. (2010). Group processes and EFL learners' motivation: A study of group dynamics in EFL classrooms. *TESOL Quarterly, 44*(1), 129–154.
Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles: SAGE.
Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment, 18*(2), 192–203.
Dörnyei, Z., & MacIntyre, P. D., & Henry, A. (2014). Motivational Dynamics in Language Learning. Bristol: Multilingual Matters.
Dörnyei, Z., & Otto, I. (1998). Motivation in action: A process model of L2 motivation. *Working Papers in. Applied Linguistics, 4*, 43–69.
Gist, M. E., & Mitchell, T. R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review, 17*(2), 183–211.
Graham, S. (2007). Learner strategies and self efficacy: Making the connection. *Language Learning, 35*(1), 81–93.
Gow, A., Whiteman, M., Pattie, A., & Deary, I. (2005). Goldberg's "IPIP" big-five factor markers: Internal consistency and concurrent validation in Scotland. *Personality and Individual Differences, 39*(2), 317–329.
Hsieh, P. P. H., & Kang, H. S. (2010). Attribution and self efficacy and their interrelationship in the Korean EFL context. *Language Learning, 60*(3), 606–627.
Hatch, J. A. (2002). *Doing qualitative research in education settings*. New York: SUNY Press.
King, J. (2015). The dynamic interplay between context and the language learner. In *Hampshire, UK*. Macmillan: Palgrave.
Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford: Oxford University Press.
Lee, I., Yu, S., & Liu, Y. (2017). Hong Kong secondary students' motivation in EFL writing: A survey study. *TESOL Quarterly*.
Linacre, J. M. (2007). *A user's guide to WINSTEPS: Rasch-model computer program*. Chicago, IL: MESA Press.
Linacre, J. M., & Wright, B. D. (2007). *WINSTEPS: Multiple-choice, rating scale, and partial credit Rasch analysis [computer software]*. Chicago, IL: MESA.
Magogwe, J. M., & Oliver, R. (2007). The relationship between language learning strategies, proficiency, age and self-efficacy beliefs: A study of language learners in Botswana. *System, 35*(3), 338–352.
McDonough, J., & McDonough, S. H. (1997). *Research methods for English language teachers*. London: Arnold.
Mills, N. (2009). A guide du Routard simulation: Increasing self-efficacy in the standards through project-based learning. *Foreign Language Annals, 42*(4), 607–639.
Mills, N. (2014). Self-efficacy in second language acquisition. In S. Mercer & M. Williams (Eds.), *Multiple perspectives on the self in SLA* (pp. 6-19). Bristol: Multilingual Matters.
Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals, 39*(2), 276–295.

Mills, N., Pajares, F., & Herron, C. (2007). Self-efficacy of college intermediate French students: Relation to achievement and motivation. *Language Learning, 57*(3), 417–442.

O'Connell, A. A., & McCoach, D. B. (Eds.). (2008). *Multilevel modeling of educational data*. Charlotte, NC: IAP.

Oller, J., & Streiff, V. (1975). Dictation: A test of grammar-based expectancies. *ELT Journal, 4*(1970), 37–41.

Peng, J. E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language Learning, 60*(4), 834–876.

Piniel, K., & Csizér, K. (2015). Changes in motivation, anxiety and self-efficacy during the course of an academic writing seminar. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 164–194). Bristol, Buffalo, Toronto: Multilingual Matters.

Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33–40.

Rahimi, A., & Abedini, A. (2009). The interface between EFL learners' self-efficacy concerning listening comprehension and listening proficiency. *Novitas-ROYAL, 3*(1), 14–28.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. Thousand Oaks, CA: Sage.

Swanson, P. (2012). The congruence of vocational interests and the workplace environment: Reducing the language teacher shortage. *Language Teaching Research., 16*(4), 519–537.

Swanson, P. (2013). From teacher training through the first year on the job: Changes in foreign language teacher efficacy. *Electronic Journal of Foreign Language Teaching, 10*(1), 5–16.

Swanson, P. (2014). The power of belief: Spanish teachers' sense of efficacy and student performance on the national Spanish exams. *Hispania, 97*(1), 5–20.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson/Allyn & Bacon.

Wang, C., & Bai, B. (2016). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Quarterly*.

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.

Wright, B. D. (2003). Rack and stack: Time 1 vs. time 2 or pre-test vs. post-test. *Rasch Measurement. Transactions, 17*(1), 905–906.