Visual Input Enhancement via Essay Coding Results in Deaf Learners’ Long-Term Retention of Improved English Grammatical Knowledge

Gerald P. Berent
Ronald R. Kelly
Kathryn L. Schmitz
Patricia Kenney
National Technical Institute for the Deaf, Rochester Institute of Technology

This study explored the efficacy of visual input enhancement, specifically essay enhancement, for facilitating deaf college students’ improvement in English grammatical knowledge. Results documented students’ significant improvement immediately after a 10-week instructional intervention, a replication of recent research. Additionally, the results of delayed assessment documented students’ significant retention of that improvement five and a half months beyond the instructional intervention period. Essay enhancement served to highlight, via a coding procedure, students’ successful and unsuccessful production of discourse-required target grammatical structures. The procedure converted students’ written communicative output into enhanced input for inducing noticing of grammatical form and, through essay revision, establishing form-meaning connections leading to acquisition. With its optimal design characteristics supported by theoretical and empirical research, essay enhancement is a highly effective methodology that can be easily implemented as primary or supplementary English instruction for deaf students. The results of this study hold great promise for facilitating deaf students’ English language and literacy development and have broad implications for second-language research, teaching, and learning.

Input Enhancement

Within the field of second language acquisition (SLA) research and teaching, Sharwood Smith (1991) proposed input enhancement as a theory-based second-language (L2) teaching methodology designed to draw L2 learners’ attention to target language form—lexical items or grammatical morphemes and structures. Drawing learners’ attention to form counters their natural tendency to allocate attention to L2 meaning at the expense of form (VanPatten, 1996). Among other proposals for facilitating L2 form-meaning connections, Long (1991) advocated a focus on form approach, in which teachers draw learners’ attention to specific L2 forms at appropriate discourse junctures. Long carefully distinguished between focus on form (singular) within naturalistic communicative environments and a focus on forms (plural), which refers to the isolation, discussion, and practice of target language grammatical forms as the primary “object” of L2 teaching and learning. An overview of input enhancement and other focus on form methodologies developed since Sharwood Smith’s and Long’s original proposals appeared in Doughty and Williams (1998).

Input enhancement is grounded theoretically in models of SLA, such as Gass (1997), that maintain cognitive-linguistic noticing of L2 input is prerequisite to the further processing of that input. Input noticing permits input intake leading to the establishment of form-meaning associations and the ultimate integration of these associations into the learner’s emerging grammar, in other words, acquisition. Some input enhancement methods were designed for teacher–learner spoken interaction, for example, a teacher’s recast of an...
L2 learner’s utterance to reformulate it to the target-language norm (see Long, 2007). However, many input enhancement methods for use with hearing L2 learners were developed for deployment in the visual modality (see Doughty & Williams, 1998). For example, textual enhancement is visual input enhancement that involves highlighting through underlining, bolding, or a larger font in text in order to induce noticing of a target form or forms during reading.

Textual enhancement has been the most researched L2 focus on form methodology, and assessments of its efficacy have yielded mixed results. In reviewing 21 textual enhancement studies conducted over the past 15 years, Han, Park, and Combs (in press) identified methodological idiosyncrasies inherent in many of the studies and isolated the relevant variables and major issues that have influenced, positively or negatively, the generalizability of their results. These issues are directly relevant to this article and are elaborated further in the Discussion section.

Deaf Learners of English

Just like hearing L2 learners of English, deaf learners of English experience severe constraints on noticing (and subsequent processing of) English language input. In the case of hearing learners, the constraints are associated with later language learning and include cognitive factors related to critical periods for language acquisition, attentional resources, implicit/explicit learning, structural complexity of the L2, influence of the first language (L1), etc. (Gass, 1997). Many of these same variables account for the constraints on deaf learners’ noticing of English language input, precipitated by reduced or no auditory access to spoken English input and the resulting burden on cognitive and perceptual resources for acquiring spoken-language knowledge. In both populations, ultimate attainment of English language knowledge is highly variable and generally falls short of L1 native English competence.1

Given the cognitive, perceptual, and acquisition factors constraining deaf learners’ acquisition of spoken language knowledge, deaf learners typically struggle to attain sufficiently high English skill levels for educational and career success. The challenge and its consequences are apparent in the comparison of deaf learners’ assessed English proficiency levels relative to the English proficiency of their hearing peers (Karchmer & Mitchell, 2003). In view of the typically slow rate of deaf students’ English language and literacy development that has remained constant for decades, frustration over the search for a breakthrough has persisted (Paul, 2000; Qi & Mitchell, 2007).

Despite the pervasive English challenge and the resulting low average levels of attainment, it is important to recognize that, within the deaf population, the range in levels of English attainment is wide. At the high end, many deaf learners, those who have successfully compensated for reduced access to spoken language input, often attain levels of English proficiency no different from their hearing peers. In interviewing such “successful” deaf students at the college level, Toscano, McKee, and Lepoutre (2002) discovered a consistent pattern. All their participants reported growing up in rich communicative environments in which English was accessible via visual and other compensatory means and all family members actively and persistently engaged in communicative interaction. The students reported this kind of family environment whether or not sign language was used in the home. Such retrospective evidence underscores the critical need to facilitate deaf learners’ access to English language input as early and consistently as possible.

Such facilitation of input access is increasing as a result of an increase in early intervention initiatives and improvements in assistive technologies. For example, universal newborn hearing screening (UNHS) has contributed to earlier identification of hearing loss and therefore earlier interventions through which deaf children can access English language input at much earlier ages than in the past. Yoshinaga-Itano (2003a, 2003b) provided an overview of progress in UNHS and evidence of improved English language development (and sign language development) in young deaf children resulting from UNHS. Cochlear implant (CI) use is another contributor to deaf children’s improved access to English language input (see the overview in Spencer & Marschark, 2003). In an assessment of phonological awareness, vocabulary, and reading in deaf children who were CI users, James, Rajput, Brinton, and Goswami (2008) observed progress toward closing the gap between young deaf children and their hearing
peers, especially in the development of phonological awareness. Similarly, Nikolopoulos, Dyar, Archibold, and O’Donoghue (2004) observed a trend toward improved English grammatical knowledge following implantation, especially among deaf children under the age of 4.

In view of such developments, the landscape may be changing for young deaf children’s prospects for attaining higher levels of English language and literacy skills resulting from earlier and better access to English language input. However, as Nikolopoulos et al. (2004) noted, despite progress in the right direction, overall “spoken language grammar acquisition in pre-lingually deaf children using cochlear implants was found to be considerably delayed” (p. 633). Thus, no single advance is likely to be the magic bullet relative to deaf learners’ English language acquisition. Accordingly, the search for highly effective teaching methodologies to facilitate and accelerate deaf learners’ English language and literacy development should remain among the highest of educational priorities. This priority is underscored by research evidence verifying the critical role that strong English language skills play in educational success, especially with respect to college graduation (Cuculick & Kelly, 2003).

**Existing Enhancement Research With Deaf Learners**

In search of more efficacious English teaching strategies, Berent (2005) advocated the development and employment of input enhancement methodologies in teaching deaf learners of English. These methodologies had been shown efficacious in at least some of the studies in the literature on SLA by hearing learners (see Han et al., 2008), and the delivery of many of the methodologies through print exploits the visual modality appropriate for linguistic communication by deaf learners. To assess the efficacy of input enhancement methods in teaching English to deaf learners, Berent et al. (2007) conducted a classroom research study that employed a combination of these methods. Participants in that study comprised three groups of students at the National Technical Institute for the Deaf (NTID), two experimental groups and a control group, taking a 10-week remedial grammar course that met for 1 hr/week. Nine English grammatical structures were the target of instruction in the three groups. The NTID English Program had identified these nine structures for “mastery” in the course objectives. So as not to divert from the goals of the English curriculum, the nine structures were the focus of both course instruction and the associated classroom research.

The two experimental groups received instruction consisting of textual enhancement of selected readings and an essay enhancement coding procedure, which is described further in the Method section. One of the experimental groups received a third enhancement intervention—a **visual dictogloss** (or **visuogloss**), modeled after the oral dictogloss procedure employed in teaching hearing L2 learners of English (see Doughty & Williams, 1998). All these methods are described and illustrated in Berent et al. (2007) and in Berent and Kelly (2008). The control group received “conventional” grammatical instruction (explanation, drill and practice, ASL-English comparison, etc.) but no input enhancement instruction.

Analysis of group performance based on the direct pre- and postassessment of participants’ productive grammatical knowledge revealed significant improvement on the target English structures after 10 weeks by the two experimental groups but no improvement by the control group. The study reported in Berent et al. (2007) was unique in being (presumably) the first empirical study of input enhancement instruction with deaf learners of English and in being one of the only input enhancement studies to explore the interventions employed “long-term” (i.e., over 10 weeks). The results of that study firmly established the efficacy of visual input enhancement in teaching English to deaf students at the college level. However, it did not assess retention of grammatical knowledge after an extended period following the instructional intervention.

Accordingly, the purpose of this article was four-fold: (i) to attempt to replicate the efficacy of input enhancement English instruction with deaf learners at the college level, (ii) to isolate and focus the investigation on just the one input enhancement methodology, the essay enhancement coding procedure, that was the primary enhancement intervention in Berent et al. (2007), (iii) to determine whether any observed improvement in grammatical knowledge is sustained
over an extended time period beyond the instructional intervention phase of the research study, and (iv) to explain the results of the research in the context of the cognitive, perceptual, and acquisitional factors that underlie input enhancement as explored in Han (2005) and Han et al. (2008).

**Method**

**Hypotheses**

This article targeted deaf college students taking the same 10-week remedial English grammar course as the participants in Berent et al. (2007). In light of the positive results obtained in that study, it was hypothesized that students receiving input enhancement instruction would exhibit significant improvement on the target set of nine English grammatical structures relative to a control group of students taking the same course but receiving no input enhancement instruction. More specifically, it was hypothesized that the essay enhancement procedure specifically, the primary intervention provided to the experimental groups in Berent et al. (2007), would account for this improvement after 10 weeks. Secondly, in light of the robustness of the improvement observed in the Berent et al. study, it was hypothesized that a significant portion of the improvement exhibited by the students receiving essay enhancement instruction should be sustained after several months, even in the absence of any follow-up input enhancement instruction, whereas the control group should not exhibit any change in their productive grammatical knowledge. This hypothesis followed the assumption that grammatical improvement represented the acquisition of form-meaning associations firmly integrated into the learners’ mental grammars.

**Participants**

Thirty-four NTID students taking the remedial English grammar course participated in this study. Eighteen of these students received essay enhancement grammatical instruction and comprised the **Enhancement** group. Sixteen of these students received conventional grammatical instruction involving no input enhancement procedures and comprised the **Control** group. Table 1 contains available demographic and test data for the two groups. The mean age of participants in the Enhancement and Control groups was 20.3 and 19.7, respectively. Comparison of group means yielded no significant differences (p > .05) between the two groups in age, Pure Tone Average (PTA) hearing loss, or on any of the English proficiency measures: the California Achievement Tests (Tiegs & Clark, 1957), the NTID Placement Test for Nonfiction Reading Curriculum (Crandall, 1997), and the NTID Writing Test (Albertini et al., 1986).

**Assessment Procedure and Analysis**

The essay coding procedure (Berent, 2005; Berent et al., 2007) was employed both for assessing participants’ productive grammatical knowledge of the nine target grammatical structures and for providing essay enhancement instruction to the Enhancement group. The grammatical structures identified for “mastery” in the remedial English grammar course included seven morphosyntactic verb formations and two clausal structures. These structures had been identified by the NTID English program on the basis of the observed challenges they pose to deaf students at this particular English level and were thus the official focus of the course objectives. The verb formations included

<table>
<thead>
<tr>
<th>Test/characteristic</th>
<th>Group</th>
<th>Enhancement</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td></td>
<td>18</td>
<td>16</td>
</tr>
<tr>
<td>Age</td>
<td>M</td>
<td>20.3</td>
<td>19.7</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.5</td>
<td>1.7</td>
</tr>
<tr>
<td>California Reading</td>
<td>M</td>
<td>8.1</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>1.3</td>
<td>0.7</td>
</tr>
<tr>
<td>NTID Reading Placement</td>
<td>M</td>
<td>98.2</td>
<td>104.5</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>22.4</td>
<td>14.4</td>
</tr>
<tr>
<td>NTID Writing</td>
<td>M</td>
<td>41.3</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>4.7</td>
<td>2.5</td>
</tr>
<tr>
<td>PTA Hearing Loss</td>
<td>M</td>
<td>98.6</td>
<td>97.4</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td>22.5</td>
<td>18.7</td>
</tr>
</tbody>
</table>

aCalifornia Achievement Tests—Junior High Level (Tiegs & Clark, 1957). Mean values represent grade-level equivalent reading scores.
bNTID Placement Test, Nonfiction Reading Curriculum. Percentile scores range from 0 to 200.
cThe NTID Writing Test (Albertini et al., 1986). Possible scores range from 0 to 100.
dPure Tone Average hearing loss measured in the better ear at 500, 1,000, and 2,000 Hz.
the following, with their associated codes shown in parentheses: present tense (PRES), past tense (PST), present tense/progressive aspect (PRES/PROG), past tense/progressive aspect (PST/PROG), infinitive (INF), modal verb plus main verb (MOD), and perfect aspect (PERF). The two clausal structures were that-complements (THAT) and adjective (i.e., relative) clauses (ADJC).

Participants’ productive grammatical knowledge was assessed in short essays that they typed (30–40 min.) under monitored conditions during the first and last weeks of the 10-week course and again five and a half months later. For the First Essay, participants selected one of three personal-experience topics to develop: “My Best Friend,” “My First Job,” or “My Worst Day.” For the Last Essay during the 10th week of the course, each student chose one of the two topics they had not chosen for their First Essay. For the Delayed Essay after five and half months, each student chose his or her remaining topic of the initial three. As in Berent et al. (2007), students’ chosen topics were distributed fairly evenly across the three assessments as well as within and across the two groups. With respect to the essay assessment data gathered by Berent et al., Berent and Kelly (2008) calculated intercorrelations based on frequency of occurrence of grammatical structures produced by the experimental and control groups by essay administration (First, Last) and by essay topic separately. These calculations revealed extremely high correlations indicating that discourse patterns of grammatical structure did not differ either across essay assessments or by topic. Thus, participants’ essays validly reflected change (if any) in grammatical knowledge over time that was not confounded by any differences in discourse structure associated with essay topic or other factors. Because of the demographic similarity of participants in the previous and current studies and use of the identical essay topics, discourse equivalence is assumed to characterize the current essay data as well.

Each essay was coded for the successful production of the nine target structures as required by a student’s own established discourse context. Productive grammatical knowledge was calculated as a percentage based on the number of successful discourse-appropriate productions of the target structures divided by the total number of attempted structures, both successful and unsuccessful. The procedure is demonstrated below in the description of the essay enhancement instructional process. Change in productive grammatical knowledge over time was calculated on the basis of difference scores between the overall percentages successful on the First, Last, and Delayed Essays. Participants’ percentage scores were converted to arcsine values following the procedures recommended by Snedecor and Cochran (1980) for handling a zero proportion and a 100% proportion when \( n \) is less than 50. The percentages were transformed to arcsine values, which were used in the statistical tests—analysis of variance (ANOVA) with repeated measures on the dependent variable.

In addition to scoring by the instructor employing the essay enhancement methodology, two additional raters scored random samples of the participants’ First, Last, and Delayed essays. This scoring was blind: participants’ names were removed from the essays, and essays were not identifiable relative to participants’ group membership or the specific essay category (First, Last, and Delayed). The interrater reliability of essay coding by multiple raters was calculated at .98 using a measure of correlation (Borg & Gall, 1983). Such a high interrater reliability suggests that the coding and scoring process is clear and consistently more objective than subjective. This level of interrater reliability is consistent with the previous research using the coding procedure in Berent et al. (2007).

Essay Enhancement Instructional Methodology

The same essay coding procedure was used in the assessment of participants’ productive grammatical knowledge and in the essay enhancement instructional methodology. Participants in the Control group produced First, Last, and Delayed Essays but did not receive their coded essays and were not aware of the use of a coding procedure for either assessment or instructional purposes. In contrast, after completion of the First Essay, participants in the Enhancement group were instructed in the details and purpose of the essay coding procedure, along with explanation, illustration, and discussion of the nine target grammatical
structures beginning in the second week of the 10-week course. These participants’ First Essays were thus coded for both assessment and instruction and were returned to them after they were coded. Their Last and Delayed Essays were used only for assessment purposes and not returned to them.

Students developed their short-essay topics, spending between 30 and 40 min, on individual laptop computers in a “high-tech” classroom using Microsoft Word software and double-spaced Times New Roman font, size 12. The instructor gathered both the hard copy and electronic MS Word file of each student’s essay. In coding an essay, the instructor studied the student’s established discourse context, for example, reflection on a past experience, developing relationship over time, to determine the time frame and perspective of the particular discourse. This facilitated the instructor’s assessment of the participant’s use of the target grammatical structures. Those structures—and no others—were coded for their successful or unsuccessful production. “Essay enhancement” consisted of the instructor typing structure codes directly in the student’s MS Word file immediately before the relevant grammatical structures; a plus sign preceded the code of a successfully produced structure, and a minus sign preceded the code of an unsuccessfully produced structure. To call attention to students’ successful and unsuccessful structures, codes were typed in Times New Roman bold font, size 14. The bolding and contrast in font size served as visually enhanced input once students received their coded essay files. By this procedure, the student’s own productive output was converted into visually enhanced input via coding. The instructor did not correct any of the student’s unsuccessful productions. The following sentences excerpted from students’ essays illustrate the essay coding procedure relative to the nine target structures.

1. They +PST talked to me –THAT that I +PST had +INF to be on time.
2. The worker +PST told me +THAT that I –MOD must to clean the floor.
3. In my first job where I –PST work at summer camp when I +PST was age 15 years old.
4. He –PRES have many horses –ADJC +PRES are famous for horses show.

5. She +PRES is jealous +THAT that I +PRES have best friend +ADJC who –MOD can hung out and +MOD have more time –INF to conversation with.
6. I never –MOD forget how +PERF I’ve done on my farm experience.
7. I +PRES have another best friend –ADJC that I –PERF know her for a long time.
8. I +PRES/PROG am looking forward –INF to see her again in November for Thanksgiving vacation so we +MOD can spend time together like old times.
10. She +PST/PROG was helping me –INF to studying homework and +INF do projects.
11. After the halftime, we +PST returned to the game and I still –PST/PROG playing bad game.

In addition to the illustration of successful discourse-appropriate productions, the above excerpts illustrate which factors contribute to unsuccessful productions. For example, the modal verb formation in 2 is unsuccessful in view of its syntactic components, must to clean; in 6 –MOD designates the absence of the modal verb will as required by the student’s established discourse, which calls for will forget. An unsuccessful present tense includes a non-English formation, the absence of a present tense verb in a required environment or, in the case of sentence 4, lack of proper agreement (have for has). With respect to that-complements, –THAT in 1 signals that a that-complement is ungrammatical after the main verb talked, unlike the +THAT grammatical production after told in 2. THAT and ADJC designate clausal structures, whether well-formed or ill-formed as –ADJC in 4, which calls for that (or which) are famous for horse show. Other target structures embedded within that-complements or adjective clauses may themselves be successful or unsuccessful (e.g., +PRES are in 4). However, each successful or unsuccessful target production is a discrete entity. More detailed illustration and discussion of longer essay excerpts were provided in Berent et al. (2007) and Berent and Kelly (2008).

Students in the Enhancement group received their coded First Essay files during the second week of the
course and were provided 2 weeks to complete their first revision of the essay, during which their familiarity with the coding system and classroom grammar instruction accrued. They were told to compare their (enhanced) successful productions with their unsuccessful productions, to reflect on these forms and intended meanings, and to convert unsuccessful structures to successful structures as best they could in their first revisions. They were also told not to delete any codes in their first revisions, despite reformulations, so that the instructor could see the locations of original successful and unsuccessful productions. The first revisions were submitted in hard copy on which the instructor acknowledged successful reformulations in ink and circled remaining unsuccessful structures. First revisions were returned to students, who were told to continue reformulation of unsuccessful structures in a second revision, to then remove all original codes from their electronic MS Word files, and to submit their second revisions in hard copy. The instructor verified further successful revision and provided final comments in ink on students’ second revisions of the First Essay. During the remainder of the course, students produced two more short essays, with time for two revisions of Essay 2 and one revision of Essay 3.

Results

Change Between First and Last Essays

As indicated above, participants in the Enhancement and Control groups typed their in-class Final Essays during the 10th week of the 10-week course. The first analysis of the essay data assessed change in productive grammatical knowledge of the nine target structures collectively over the 10-week period. A 2 (Group) × 2 (Essay: First, Last) ANOVA with repeated measures was conducted and yielded an overall significant main effect for Essay, \( F(1, 32) = 27.96, p < .0001 \), indicating an overall significant change between the First Essay and Last Essay means. However, the ANOVA also yielded a significant Group × Essay interaction, \( F(1, 32) = 22.41, p < .0001 \), revealing that the Enhancement group experienced significant improvement in productive grammatical knowledge between the First Essay (61.5% successful) and the Last Essay (85.1%) in contrast to the Control group, which exhibited no significant improvement between the First Essay (73.6%) and the Last Essay (74.9%). These results replicate the findings of Berent et al. (2007), establishing the efficacy of visual input enhancement, in this case, essay enhancement via coding, in teaching English grammar to deaf college students.

Change Between First, Last, and Delayed Essays

Participants in the Enhancement and Control groups had been informed during their course that they would be invited to participate in a follow-up research activity later in the academic year. Students were contacted and invited to participate in this activity approximately five months after the completion of their 10-week remedial grammar course. Of the original 18 and 16 participants in the Enhancement and Control groups, respectively, recruitment resulted in 12 students from each group returning to type the Delayed Essay under the same controlled conditions as during the course. The 24 students typed their Delayed Essays on their remaining topic of the original three topic choices five and a half months after the end of the course. Each student was paid $20 for participating.

The essay data for these 24 students were submitted to a 2 (Group) × 3 (Essay: First, Last, Delayed) ANOVA with repeated measures. There was a significant main effect for Essay, \( F(2, 44) = 13.59, p < .0001 \), indicating significant change across essays, as well as a significant Group × Essay interaction, \( F(2, 44) = 11.26, p < .0001 \). The interaction is plotted in Figure 1. The figure illustrates the significant change exhibited by the 12 Enhancement group participants between their percentage successful on the First Essay (62.3%) and the Last Essay (85.3%) as well as the significant change between the First Essay and the Delayed Essay (73.8%). The performance of the 12 Control group participants revealed no significant improvement across the First (74.9%), Last (76.0%), or Delayed (75.5%) Essays. Thus, the Enhancement group not only improved in their productive grammatical knowledge of the nine target structures over the original 10-week course; their written essays still showed that they also retained this improvement, with
a modest decrease relative to the Last Essay, after five and a half months—nearly half a year.

The remedial English grammar course that all participants took was a co-requisite to their remedial “Academic Writing II” course, which they took simultaneously with the grammar course and which met 4 hr/week. In the writing course, students learn to “organize and develop paragraphs and brief compositions of various discourse types, with particular emphasis on narration, description and process” [catalogue description]. In the intervening period between the Last Essay and the Delayed Essay, of the 24 students completing all three essay assessments, 8 students from the Enhancement group and 10 from the Control group took the developmental “Nonfiction Reading III” course, one student from each group took “Academic Writing IV,” and three Enhancement group students and one Control group student took no English course at all.2 These follow-up courses do not have any formal English grammar focus, and none included any instruction involving input enhancement. Accordingly, the Enhancement group’s retention of productive grammatical knowledge acquired in the grammar course in all likelihood was not influenced by any grammar instruction subsequent to their participation in this study. Furthermore, both groups took essentially the same English courses in the intervening period.

Absence of Regression Effects

Although a regression effect might possibly provide a plausible alternative explanation for the results of the Enhancement Group’s performance on the Last and Delayed essays (as suggested by a reviewer), reasons why this is not the case are as follows. In behavioral research, a subtle source of invalidity is the elusive phenomenon of regression (Hopkins & Glass, 1978). The interpretation of gain scores is often confounded by the phenomenon of regression toward the mean (Borg & Gall, 1983). The regression effect occurs when research participants may have extreme scores (at either the high or low end) on any measure at a given point in time. For purely statistical reasons, the next time participants are tested, there is a natural tendency for their scores to be less extreme—that is, regression toward the true population mean. Those who scored higher on measure X would tend to have on average lower standard scores on Y, and those who scored lower on X would tend to have higher scores on Y (Kenny, 1979). “The regression effect occurs because of errors of measurement in the pretest and posttest and because the tests are correlated with each other” (Borg & Gall, pp. 721–722).

Importantly, this study was designed to minimize any misinterpretation in the gain scores due to potential regression effects. Participants in both the Enhancement and Control groups were similar in age and academic abilities and not selected because of their low or high performance on any single measure (see Table 1 and related commentary). The student participants were all placed in the same level of English instruction. Furthermore, the 24 volunteer participants who contributed to the Delayed Essay analysis were generally not extreme performers on the First and Last Essays. In the Delayed Essay analysis, these 24 students, representing both the Enhancement and Control groups, distributed similarly in their scores on the First and Last Essays. On the First Essay, both groups had 67% within ±1 standard deviation (SD) from the group mean and 33% within ±2 SDs. On the Last Essay, both groups had 75% within ±1 SD of the mean and 25% within ±2 SDs. Therefore, not only were the participants statistically equivalent in age and overall English proficiency but they also distributed similarly about the group means on the First and Last essay measures with no extreme high or low performers.

Furthermore, a repeated measures analysis of variance accounts for the relationship between the
pretests and posttests (related because each participant generates multiple scores that are compared). The most important finding of the repeated measures ANOVA was the significant interaction between groups, indicating a different pattern of performance relative to the First/Last/Delayed Essay scores. Although a regression effect should potentially have equal influence on the posttest scores of both the Enhancement and Control groups because of their statistically similar age and academic ability levels and similar distribution, the Control group performance across all three essays was essentially flat and statistically unchanged. Only the experimental group receiving the essay enhancement instructional methodology showed a statistically significant gain on both the Last and Delayed Essays relative to the First Essay. This is not to suggest that a regression effect did not potentially affect scores in both groups, but the pattern of change by the Enhancement Group is sufficiently robust to strongly suggest that it was due primarily to the instructional methodology. Although there was obviously some degradation in performance between the Last Essay and the Delayed Essay, this decrease could be attributable in part to the hiatus of five and a half months.

Discussion

Improvement and Retention of Grammatical Knowledge

As in Berent et al. (2007), the results of this article documented the efficacy of visual input enhancement in accelerating deaf college students’ acquisition of English grammatical knowledge. It was emphasized in the previous study that “a delayed posttest should certainly be incorporated into the design of future research of this type” (p. 20) in order to determine whether deaf learners’ improvement in English grammatical knowledge resulting from input enhancement instruction is retained over time. The present results established that, indeed, participants who received such instruction not only improved after 10 weeks but also retained that improvement over the long term (five and a half months) with only a modest decrease in assessed grammatical knowledge relative to their 10-week assessment during the instructional intervention period.

The deaf participants in this study, like many if not the majority of deaf individuals, have struggled since childhood to acquire functionally adequate levels of English language knowledge. For deaf college students at this age, further acquisition—and retention—of fundamental morphosyntactic structures of English is unexpected, given their very gradual rate of improvement in English grammatical knowledge otherwise observed (Berent, 1993). Moreover, the retention of this grammatical knowledge in the absence of any input enhancement English instruction after completion of the 10-week grammar course (and most likely any other grammar instruction) is, educationally, a very encouraging result.

The Efficacy of Essay Enhancement

The results of this study lend further support to the efficacy of input enhancement as a theory-based methodology for facilitating English SLA under the assumption that noticing is necessary for further processing of target language input leading to acquisition. Whereas Berent et al. (2007) employed a combination of focus on form methods in teaching their experimental groups, this article isolated one of those methods, essay enhancement, and validated the efficacy of that specific intervention in accelerating deaf college students’ acquisition of English grammatical knowledge and contributing to retention of that knowledge. A specific question is why does the essay enhancement methodology per se have such a positive impact on deaf learners’ English language acquisition?

Expanding on Han (2005), Han et al. (2008) conducted a comprehensive review of 21 SLA research studies of textual enhancement, the most researched input enhancement methodology over the past 15 years. They identified the central conceptual and acquisitional issues inherent in textual enhancement as a teaching strategy, the factors contributing to the positive or negative outcomes observed in existing studies, and the valuable and testable insights gleaned from the textual enhancement literature. The following discussion makes reference to aspects of the conclusions of Han et al.

First and foremost, the essay enhancement procedure inherently satisfies one of the most fundamental
cognitive “prerequisites” for the dual processing of L2 input for both form and meaning, specifically, the sequential processing of input for meaning first and form second (VanPatten, 1990, 1996). Han et al. (2008) found that those textual enhancement studies that employed simultaneous processing for form and meaning generally failed to demonstrate efficacy as opposed to those few studies to date that employed a sequential processing design. With sequential processing, once the input is processed incidentally for meaning, learners’ attentional resources are freed up for the subsequent intentional processing of the input for form. The beauty of the essay enhancement methodology is that it guarantees meaning-based processing up front because, in developing personal-experience essay topics, students ipso facto comprehend their own communicative output. Once that output is converted to visually enhanced input via coding and returned to the students, their attention can be allocated fully to the processing of form.

Another conclusion drawn by Han et al. (2008) is that the noticing of enhanced input is more likely to lead to acquisition if learners have prior knowledge of a target form. In the case of deaf learners of English at the college level, they will have been exposed to English grammatical structures through print and other compensatory means throughout their educational years. Some of this grammatical knowledge will have been acquired, some partially acquired, some available perhaps as receptive rather than productive grammatical knowledge, and so on. With such “prior knowledge,” these deaf learners were well positioned to notice, process, and acquire visually enhanced forms through the essay enhancement procedure.

Sharwood Smith (1991, 1993) emphasized that it is not enough for learners simply to notice enhanced input, they must further have the opportunity to act upon this noticed input. Indeed, Han et al. (2008) noted that many of the textual enhancement studies that they reviewed operated on the assumption that the effects of enhancement should show up immediately in learners’ productions, which was overwhelmingly not the case. A property of the essay enhancement methodology is that it necessitates acting upon noticed enhanced input through a protracted, intense, and intentional process of essay revision. The revision process requires deep reflection, and thus focused attention, on successful and unsuccessful productions to enable the learner to convert unsuccessful productions to successful ones (with no error correction provided). Furthermore, because the essay enhancement procedure was a “long-term” (i.e., 10-week) intervention, it provided participants with multiple opportunities (threecoded essays, five revisions) to notice and act on enhanced input over an extended period of learning.

Another relevant variable is the distinction drawn between simple enhancement and compound enhancement (Han et al., 2008). Simple enhancement involves one enhancement strategy such as textual enhancement via the use of a larger, underlined, or bolded font; compound enhancement involves textual enhancement along with another attention getting strategy such as corrective feedback to the learner. Sharwood Smith (1993) had described textual enhancement as more of an implicit than an explicit device because its underlying purpose might not be apparent to the learner. From that perspective and on the basis of the research they reviewed, Han et al. concluded that compound enhancement is more effective in inducing noticing. Interestingly, they cited Berent et al. (2007) to point out that “compound enhancement is not always necessary” for obtaining successful outcomes.

In fact, though it might outwardly look like simple enhancement, essay enhancement is indeed compound enhancement or what might be called “combined enhancement.” Via the coding procedure, each code inserted into a student’s essay combines and encapsulates complex information. First, insertion of the textually prominent code immediately before the target form substitutes for underlining, bolding, etc., to call attention to the form that follows. Second, unlike underlining, etc., the code carries meaning: it defines the specific form that follows it. Third, the + or – designation actually provides “corrective feedback,” indicating whether the form was produced successfully or unsuccessfully in its discourse environment. Therefore, the fact that essay enhancement is really a type of compound (and not simple) enhancement, and accordingly an explicit device with a clear purpose (cf. Sharwood Smith, 1993), further accounts for its robust effects.4
Implications
A Promising Breakthrough

Empirical classroom research on English teaching and other educational methodologies for deaf students is rare. Given the slow rate of English language and literacy development exhibited on average by deaf learners and the serious negative impact of low English skill levels on their educational and career success, the findings of this empirical study are promising. The extent to which participants in the Enhancement group acquired new grammatical knowledge and retained that knowledge after almost half a year beyond the instructional intervention is considerable relative to the participants who did not receive essay enhancement instruction and showed no change in productive grammatical knowledge across all three assessments.

It must be emphasized, however, that this study was designed to directly assess productive English grammatical knowledge of a set of nine target structures. The results say nothing about the Enhancement or Control group participants’ receptive grammatical knowledge or any other domain of English knowledge that might have increased, decreased, or remained constant over time. Two other fundamental questions that remain unanswered by this study are the extent to which participants improved in productive knowledge of each of the nine specific target grammatical structures separately and the extent to which knowledge of specific target structures was lost over time. Learnability factors and derivational complexity can explain deaf learners’ variable success in acquiring different grammatical structures (Berent, 1996, in press; Berent et al., 2008). Similarly, input enhancement research has explored to some extent how amenable particular target language structures are to intervention (see discussion in Han et al., 2008 and VanPatten, 1996). It was not possible to explore this question in this article because, despite numerical differences in proportions of successful and unsuccessful productions across target structures, there were not sufficiently large numbers of attempted productions of all the structures to satisfy assumptions for statistical comparison. This issue is not only related to the relatively small numbers of participants per group but also to the distribution of the various structures in natural English discourse. For example, participants naturally produced hundreds of past tense forms (e.g., *lived, saw*) for describing past events and circumstances but relatively few perfect aspect formations (e.g., *have lived, has seen*) simply on the basis of discourse properties. Regardless, variation in improvement and retention is an important question for further understanding learnability constraints on deaf learners’ English acquisition and teachability constraints on which structures students might enjoy the greatest improvement and therefore where to allocate instructional and curricular attention.

This study was also not designed to assess the impact of improved English grammatical knowledge on the overall quality or communicative effectiveness of participants’ written English discourse. Although it is reasonable to expect that improvement in grammatical knowledge of the nine target English structures would contribute to the quality of participants’ productive English discourse, there are many other relevant factors that would contribute to discourse quality. For example, accurate knowledge of grammatical structures that were not the focus of the grammar course and general English lexical knowledge would certainly contribute to discourse quality. Equally important, strong organizational and paragraph development skills that were the focus of all the participants’ separate, co-requisite writing course should have had a significant impact on overall discourse quality.

Although the study made no prediction for change across essays other than increased grammatical knowledge of the target structure set by the Enhancement group, the coded excerpts from one participant’s First, Last, and Delayed Essays provided in the Appendix exemplify the type of communicative discourse that participants produced in responding to the three essay topics. Each excerpt in the Appendix is followed by comments on the participant’s successful and unsuccessful productions of the target structures, both within the specific excerpt provided and in the remainder of the essay that is not provided. From these particular samples and commentary, it appears that, after a considerable increase in overall percentage correct between the First and Last Essay, the participant’s decrease in overall percentage correct between the Last Essay and the Delayed Essay is related to a “lapse”
in the successful production of discourse-required simple past tense forms. However, because these are data from only one participant, further study of participant performance from a much larger database should prove highly informative relative to deaf students’ improvement and retention of specific grammatical knowledge over time.

Implementation

Despite the cautions and unanswered questions noted in the discussion above, the results of this classroom research study are promising. Essay enhancement and other input enhancement methodologies may potentially contribute to a long-sought breakthrough in the effort to help deaf students accelerate their pace of English language and literacy development and to attain higher levels of English productive and receptive knowledge. The observed significant improvement and retention of English grammatical knowledge resulting from essay enhancement instruction justifies the incorporation of this methodology into English teaching curricula for deaf students. Essay enhancement and other forms of visual input enhancement deserve ongoing experimentation to determine which students at which educational levels and in which educational settings might derive benefit from this methodology.

Importantly, implementation of essay enhancement instruction does not require a wholesale change in teaching style or strategies. In this article, the essay enhancement instructional methodology was implemented in a 1-hr/week co-requisite course to an English writing course that met 4 hr/week, with significant results. Combined, essay enhancement and related activities occurred only one-fifth of the time available for the participants’ formal English instruction. The key to the enhancement methodology was not correcting the students’ grammatical errors for them but rather focusing their attention on their own productive grammatical forms so they could ponder and reformulate appropriate form-meaning connections through revision. Thus, this methodology, which involves a simple concept and a familiar activity (essay writing), can be applied as an adjunct to any other English teaching approach. Finally, it should be emphasized that, for instructional purposes, essay enhancement need not focus on many different structures simultaneously (e.g., nine structures as in this study). In fact, most existing input enhancement investigations have focused on only one or two target structures at a time. Teachers who implement essay enhancement would therefore be advised to limit their target structures and gauge students’ improvement intuitively based on their experience and observations of their students’ English language growth.

Appendix

Excerpts from One Participant’s First, Last, and Delayed Essays

The three coded essay fragments below, from the same study participant in the Enhancement Group, illustrate the type of English discourse that participants produced in sharing experiences associated with “My Best Friend,” “My Worst Day,” and “My First Job.” This participant’s total essay scores are an example of considerable improvement in successful target productions between the First Essay (52.0%) and Last Essay (87.8%) with a decrease in successful productions on the Delayed Essay (65.0%). Each fragment is followed by comments on its discourse and syntactic properties relative to the target grammatical structures. To protect the participant’s identify, personal identifying information has been replaced with a description of that information (e.g., SCHOOL NAME, RELATIVE, AGE).

First Essay (Total score: 52.0%)

He –PST doesn’t know how +INF to use ASL because he –PST is hard-hearing. I –PST was greeting to him to his new school [SCHOOL NAME] it +INF was very hard time –INF to communication each other so I –PST know how +INF to use my voice –INF to communication with him and –PST teach him how +INF to use ASL. A lot kids –PST picking at him because he –PST don’t know how +INF to use ASL so I –PST helps him +INF to avoid –INF to withdraw from [SCHOOL NAME].

COMMENT: The established discourse prior to the segment shown dictated the use of past tense (e.g., didn’t know rather than doesn’t know) even though present tense might have been acceptable in a different
The fragment reveals lack of past tense marking on several verbs and ill-formed (e.g., helps) past tense formations and some ill-formed infinitives (e.g., to communication). The remainder of the participant’s essay (not shown) contained more (attempted) productions of past tense and infinitives, two successful past tense/progressive aspect formations, one successful and two unsuccessful present tense verb forms, one unsuccessful perfect aspect formation, and one unsuccessful that-complement.

Last Essay (Total score: 87.8%)

My girlfriend +PST thought +THAT I +PST dumped her and I +PST didn’t mean +INF to, I −PST supposed +INF to tell her +THAT I +PST needed a space +INF to relax and +INF forget all the past. However, it +PST was too late for me because I +PST insulted her a lot and I +PST/PROG was hurting her feeling while I +PST called her stupid words. I +PST/PROG was trying +INF to explain her and she +PST started +PST to understand what I +PST/PROG was trying +INF to tell her.

COMMENT: This discourse fragment reveals largely successful production of past tense and past tense/progressive aspect, as well as infinitives. It also contains successfully produced that-complements. The remainder of the essay contained three more successful progressive aspect formations, two successful adjective clauses, and two successful that-complements, with no unsuccessful productions of the last three mentioned structures.

Delayed Essay (Total score: 65.0%)

My [RELATIVE] +PST asked me +INF to help him and he −MOD will pay me +INF to clean it up for [AMOUNT] per hour. My [RELATIVE] and I +PST worked hard through summer for almost [NUMBER] years +INF to open the dirty road. There −PST are a lot log +INF to cut it up and −INF chopped it up for firewood. It +PST was not easy work. We +PST worked under the sun and high humidity and bugs −PST fled around us. I +PST made enough money while I +PST was [AGE] +INF to buy something +ADJC that I really +PST wanted.

COMMENT: The final discourse fragment reveals some unsuccessful instantiations of structures required for a discourse reflecting past time: the modal formation will pay rather than would pay and are instead of were. Infinitive production is mostly successful, and the adjective clause at the end of the segment is well formed.

The remainder of the participant’s essay (not shown) revealed the primary source of the decrease in total percentage correct between the Last Essay and the Delayed Essay. In the 20 remaining discourse environments requiring simple past tense, half of the productions were successful and half unsuccessful. Two of three present tense productions were successful, two of five infinitives, and three of four adjective clauses. In addition, there were two past tense/progressive aspect formations and one modal formation, all successful. For this participant, the diversity of structures attempted in the Final Essay was sustained in the Delayed Essay. The lapse in successful past tense formations over time (though superior to the percentage of past tense formations in the First Essay) appears to have contributed most to the decrease to 65.0% overall success on the Delayed Essay.

Notes

1. Similar constraints on input noticing affecting L2 hearing and deaf learners of English yield remarkable parallels in many domains of their English language development (e.g., Berent, 1983, in press).

2. With reference to catalogue descriptions, in the “Non-fiction Reading III” course, students “learn and practice the reading comprehension skills and English language skills necessary to increase their comprehension of nonfiction reading materials.” In “Academic Writing IV,” students “plan, draft, revise and edit short essays of various discourse types, such as exemplification and process. They learn how to organize and develop their texts for various topics and purposes and how to revise, edit and present them according to the conventions, format and mechanics expected by the discourse community for which they write.”

3. Han et al. (2008) grounded their assessment and analysis of textual enhancement in the context of the vast SLA literature related to theory, processing and acquisition, and L2 teaching and learning. See their article for the specific citations that underlie their conclusions referred to in the this study.

4. Defining essay enhancement via coding as combined enhancement distinguishes it from simple enhancement and complex enhancement as defined by Han et al. (2008). It is not simple enhancement because it is not a single attention-getting device such as underlining. Nor is it the employment of two distinct simple enhancement devices operating in tandem.
Instead, it coalesces two or more of attention-getting functions into one enhancement unit. As such, its capacity to draw learners’ attention to form might be more effective than complex enhancement consisting of two independent devices—certainly an issue to explore further in input enhancement research.

References


Received May 7, 2008; revisions received July 23, 2008; accepted July 23, 2008.