

Investigating Syntax Priming in an E-mail Tandem Language Learning Environment

*Christine Appel**
School of Applied Languages
and Intercultural Studies
Dublin City University
Dublin 9
Ireland
appelm@tcd.ie

Carl Vogel
Computational Linguistics Lab
Trinity College
University of Dublin
Dublin 2
Ireland
vogel@tcd.ie

Contents

1	Background	1
1.1	Overview	1
1.2	Tandem E-Mail	1
1.3	Communication Environment	1
1.4	Linguistic Priming	2
2	Experimental Design	3
3	Initial Results and Discussion	5
4	Analyses to Follow	6
	References	6

Abstract

This paper presents the results of an experiment on syntax priming in an e-mail tandem language learning environment. Twenty-two subjects participated in the completion of 6 tasks in pairs. Subject-pairs working together on tasks consisted of a native speaker of Spanish learning English and a native speaker of English learning Spanish. Bilingual asynchronous written communication between pair members took place through a system we had designed for research into learner-language in interaction. The results of this experiment inform the practice of e-mail tandem exchanges addressing the following questions: a) is there evidence of syntax priming in transcripts of e-mail tandem exchanges? If so, can we link syntax priming to second language learning? b) how do we integrate these results into the practice of tandem language learning and the design of systems that support this type of language learning over the Internet?

*The authors both gratefully acknowledge financial support for this research from the Centre for Language and Communication Studies at Trinity College.

1 Background

1.1 Overview

This paper presents the motivations, design and partial analysis of results from an experiment in linguistic priming among language learners for whom communication is mediated via e-mail. In this section we contextualize the research by discussing some of the open questions in computer mediated tandem language learning. We also describe the particular computer environment in which our experiment was conducted. Finally, we discuss relevant aspects of the literature on linguistic priming and their ramifications for open questions in tandem language learning. Section 2 outlines the design of our experiment on priming using the computer mediated environment for tandem language learning and §3 analyzes some of the results. In §4 we discuss the ramifications of these results and describe additional analyses of the data which we are undertaking.

1.2 Tandem E-Mail

Tandem language learning is a traditional technique in which learners pair themselves to learn from each other: each pair has a native speaker of the language the other member is trying to learn. Both speakers speak both languages, so that both learners have the opportunity to encounter native use of the target language, to attempt to use the target language, and to receive feedback on their non-native language production. Recently, there have been efforts to transplant this paradigm to computer mediated settings (Little & Brammerts, 1996), but there is no consensus on whether any particular style of computer mediation is more conducive to language learning than any other, although there is growing consensus that each type facilitates different aspects of language use. One of the intuitions behind adopting the computer mediated setting is that it creates a much larger pool of language learners who can participate than is available if tandem learning relies on face-to-face discussions, and the access to that larger pool of learners outweighs the usual loss of information such as pronunciation, intonation, and visual feedback mechanisms. However, adopting computer mediation does not settle the style of interaction: communication could happen in a synchronous setting, such as a MUD (originally, a Multi-User Dungeon), in which several people can interact simultaneously; or it can happen asynchronously in a forum that functions like a bulletin board, individuals posting notes for all to see, or via individual e-mails. Clearly, another axis of choice is whether the communications are public to a community or exchanged individually between pairs of individuals. Studies have been done to explore the effectiveness of individual strategies (e.g. Appel & Mullen, 2000), but thorough analysis of relative merits has not been undertaken. The current work is a step in that direction. We do not address the issue of synchronous versus asynchronous communication. Rather, we involve the larger literature of linguistic priming in examining the question of pairwise versus community tandem learning.

1.3 Communication Environment

Before addressing the issue of linguistic priming, we provide more information about the computer environment which serves as the setting for the experiment we outline and partially analyze later. The system is an e-mail server designed for tandem e-mail exchanges (Appel & Mullen, 1998). Essentially, it is a purpose-built e-mail server

for language learners. Learners register with the system, providing limited personal information (e.g. age range from one of four categories, occupation, etc.), and are then matched with a partner who is a native speaker of the learner's target language and who reciprocally wants to master the learner's native language. It is a purpose-built feature of the system in that it is not a generalized free e-mail server: the tandem pairs can send e-mails through the system only to each other. Additionally, the system has several research and pedagogically oriented features specific to questions about computer mediated tandem learning. These features include summary statistical extraction of information like the number of messages sent by an individual, the amount of native and target language used by each individual (the latter statistic is heuristic, provided by the automatic language guesser of van Noord (1998) which guesses on a per-line basis), etc. These are pieces of information useful to a teacher whose students are engaged in tandem learning as well as to researchers who want to assess the impact of e-mail in comparison with other modes of computer mediated discourse on tandem learning. From the research perspective it is important as well that one artifact is an automatically constructed corpus of learner language data (cf. Granger, 1998), and in a setting in which the researcher has complete control over the integrity of the data (Hewson, Laurent, & Vogel, 1996). The research could not be appropriately controlled if it depended on learners voluntarily collating and submitting the text of their exchanges to the researchers. Moreover, the task of annotating each individually sent record of exchange is far more complicated than automatically annotating the messages as they are sent, as is enabled by the system. Learners who make use of the system are made aware that their texts are recorded with the following disclaimer: "Joining the ETR site simply implies consent that messages in the exchange may be used anonymously for research purposes."¹

1.4 Linguistic Priming

Priming, the occurrence of an event increasing the likelihood that another event of the same type will occur, has been used in psycholinguistic research to investigate many questions. For example, cross-modal priming between sound and text based activities (e.g. Swinney, 1979) has been brought to bear on the question of modularity in cognitive architecture (see Fodor, 1983). Priming is available at all levels of linguistic representation: phonological, semantic, conceptual and syntactic. Recently, syntactic priming effects have been used to provide evidence for the cognitive reality of phrase structure (Branigan, Pickering, Liversedge, Steward, & Urbach, 1995), a notoriously unverified ontological primitive in linguistics. Branigan et al. (1995) make use of the semantic equivalence between sentences like (2) and (1).

- (1) The child gave a ball of yarn to the cat.
- (2) The child gave a cat a ball of yarn.
- (3) The child's sister had loaned....

Subjects were asked to provide sentence completions² for items like (3). What they found was that subjects presented with the prime involving one of the objects of the verb

¹"ETR" stands for Electronic Tandem Resources.

²Their actual materials provide far richer scenarios; but these concocted examples illustrate the paradigm.

marked with the preposition “to” (1) were significantly more likely to use corresponding explicit dative marking in their completion. Moreover, subjects primed with the sentence involving two unmarked objects (2) were significantly more likely to do the completion with two unmarked objects. Note that the verbs in the prime and in the completion task are similar in expecting two objects, whether the dative argument is marked with a preposition or not. The latter prime is important: because there is no preposition in that case (and because the scenarios were rich enough for completions to involve entities not mentioned explicitly), the priming is not lexical but syntactic, and evidence for syntactic priming is evidence for the underlying reality of syntax. Obviously, priming is one of the fundamental assumptions behind tandem language learning. In the experiment we describe in §2 we use estimates of priming to measure the relative efficacy of the main conditions.

Another major influence on the design of our experiment is research on communication within groups and isolated pairs. In communication oriented to task solutions, it has been shown that actual participation in the communication is essential—non-communicating overhearers are disadvantaged in being able to complete the task (Schober & Clark, 1989). Thus for tandem language learning, whether participants are in a community forum or in isolated pair, one would expect full participation throughout the relevant community to be important to overall success. Garrod and Doherty (1994) examine explicitly the relationship between communities of communicators and isolated pairs, also in task oriented settings. Garrod and Doherty (1994) constructed a situation in which participants without visual contact negotiated each other through mazes. Crucial to success in this task is successful sharing of ways of referring to the overall configuration of maze itself, and locations within it. What they found was a higher degree of convergence in the community situation (in which ultimately each participant had a game with someone who played each other player) than among the isolated pairs. This research is quite relevant to related issues in computer mediated tandem language learning: the question there is whether having a single tandem partner is better than participating in a full tandem community.

2 Experimental Design

Paid participants were either native speakers of Spanish or English, all undergraduate students spending at least an academic year in Dublin, Ireland. We adopted a task oriented communication model—the task was based on tangram-construction, a Chinese puzzle built of geometric shapes cut from a square. In any one game, one individual acted as the instructor and the other as constructor. The instructor was supplied with a design to construct and had to communicate with the constructor about the design. No two pairs had the same puzzle, but considerable pre-experiment study went into balancing the designs across pairs (in terms of figural suggestivity or overall difficulty).

Six games were played overall by each person, three in English and three in Spanish. The analysis provided in §3 focuses on just the English data. All of the native English speakers were located in one computer lab at Trinity College, University of Dublin, and all of the native Spanish speakers were located in a different lab in the same building. None of the English speakers had prior contact with the Spanish speakers. All communication relevant to the experiment happened through the tandem e-mail server which recorded the messages.

Games were limited to twenty minutes, with breaks and refreshments in the intervals. Thus, though e-mail is an inherently asynchronous mode of communication, the style of communication via e-mail induced in this design was relatively synchronous (although still actually asynchronous). We did not record relative success in the task. Firstly, communicative success is about semantic coordination, and our main interest here is in coordination of syntax; secondly, twenty minutes is a short time for the task and none of the constructions actually were perfect replicas (nor did we see value in developing metrics for degrees of success in the task). Participants were allowed to talk during the breaks, but not about the experiment itself.

Subjects were randomly assigned to two conditions: isolated pairs or community. In the isolated pair condition, the same two people played each of the six games. In the community condition each person played a different person in each game, until each person played each other person. The participants were apparently unaware of the two conditions. There were intended to be 12 subjects in each condition, but no-shows on the day of the experiment entailed that we had 6 pairs in the community condition and 5 in the isolated pair condition. Of the latter, one pair is removed from analysis because it was two native English speakers (who nonetheless were not aware they played more than one game together).

Our resulting analysis is based on the text files recorded for each individual over the course of the games. The data records which participant played which other participant in each game of the experiment. In total, 1,431 messages were sent, creating a 20,354 word corpus (773 messages and 11,519 words of English, 658 messages and 8,835 words of Spanish). The data was also tagged for parts of speech for each word using an automatic part of speech tagger (Schmid, 1994). We use this textual data to estimate syntactic coordination between native and non-native speakers in the computer mediated tandem language learning exercise. Using the text files, we generate n-gram frequencies. An n-gram is simply a sequence of items of length n, in the order that they occur in the data. Here we focus on bigrams, and only of the text itself, rather than tagged corpus.

We undertake two comparisons. One is of early games in each condition with late games in the same condition. Here, we are examining the degree of overlap between players within a game. Another is between individuals and the rest of the community in the same condition.

For the first comparison we estimate the degree of priming with the following statistic when comparing two bodies of data: the number of distinct shared n-grams divided by the number of bigrams that could have been shared. As an indication of the number of distinct bigrams that are shared, the numerator in that ratio provides a measure of influence between the two texts. Monitoring distinct bigrams attends to the type/token distinction—one particular two word sequence might account for all of the shared text between two messages, or several distinct bigrams might be at work. The other part of the measure, the denominator, relativizes the number of shared bigrams to the size of the texts. Given a text with n words, there are n-1 bigrams that could be shared. We obtain this ratio for the two texts being compared (distinct values, as they may contain different total numbers of words), and as summary of the comparison use the product of those two ratios. We measure each game this way and examine how the first and final game in English for each of the pairs in the study (thus, for the isolated pair condition, the first and final games involve no change of partner).

We also considered a related statistic in comparing the English corpus for each individual participating in the study with the entire remainder of the corpus in the same condition (isolated versus community). The measure here, for each individual, rather than by game as in the preceding paragraph, was based on the total number of distinct shared bigrams divided by the total number of shared bigrams. When comparing the corpus of an individual to the entire corpus in that condition, minus the individual, we identify the total total number of distinct shared bigrams, but also the number of times those bigrams occurred within the individual's contribution. For an individual, this ratio will tend towards the value 1 when there is considerable variety in the individual's intersection with the remainder. That ratio is an estimate of the probability that the individual will have a varied intersection with the remainder. But also take the same ratio for the remainder. The numerator is the same, but the denominator can be different because in the corpus minus the individual's contribution, the n-grams are likely to occur with different frequencies. Taking the product of those figures yields a joint probability, and thus a measure of varied textual intersection in for the individual with the rest of the condition.

3 Initial Results and Discussion

We used the standard two-tailed T-Test to measure the significance of difference between the conditions. Comparing the similarity metric between the first game and final game of English between the isolated pair and community conditions we found significance in the difference for the community condition ($p = 0.03$; there is a less than 3% chance that the difference occurred by chance). The difference did not come close to significance for the isolated pairs ($p = 0.38$). That is, there was a greater dissimilarity between the first and final games for subjects in the community condition than in the isolated pair condition. The actual measure is of degree of priming, using the estimate described in the previous section. This means that in the isolated pair condition, there was no significant change in degree of overlap between the first game played in English and the final game in English, but the change was significant in the community condition. The direction of the change was towards decreased overlap.

For the comparisons of similarity metrics between the language produced by individuals and the remaining corpus of the respective cohort (the entire English community corpus or the entire English isolated pair corpus), we obtained a significant difference ($p = 0.04$) between the individuals in the community condition compared with the individuals in the isolated pair condition. That is, each individual had greater similarity metrics to the remaining corpus in the community condition in which each member communicated with each other member than individuals in the isolated pair condition had with their overall corpus. Additionally, the members of the community condition had a higher measure of varied intersection—the individuals in that condition exhibited a greater degree of textually varied overlap with the rest of the community.

In sum, in the community condition, the language of individuals was more likely to be like the language of everyone else in that condition than the language of individuals in the isolated pairs to be like that of each of the isolated pairs. Over time, arbitrarily chosen members in the community condition show a tendency to become less affected by their partner. That is, in the community condition, the choice of a partner in any one communication will have less impact on the participant's immediate language, and yet,

the participant will share overall more features with the entire community.

4 Analyses to Follow

These two analyses reached significance using the intended tests using the usual thresholds for significant results ($p < 0.05$). A number of additional analyses on the data are also intended (in addition to examining the same tests for the Spanish corpus). For example, we focussed on bigram measures rather than trigram figures, primarily because one would naturally expect a great deal less sharing of two word sequences than three word sequences, and also because two word sequences are more likely than three word sequences to be syntactic rather than contentful items (“of the” vs. “near the rectangle”). Nonetheless we will carry out the analyses to identify whether they further discriminate the two conditions. Further, we intend to replicate the analyses over the tagged versions of the corpus. Using the part of speech tags for each word rather than the words themselves lets us abstract over the lexicon to identify influences of ways of phrasing things imposed by syntax (sequences of three adjectives are expected to be frequent in this dialog task, but adjective verb adjective triples should be quite rare, if they exist at all). We also intend more to more generally link these findings about mutual influences of partners’ language with long standing issues in second language acquisition.

In these two initial analyses we have found results which suggest that the impact of tandem language exchanges in a computer mediated environment is best facilitated in achieving convergence in ways of saying things (i.e. learning the non-native language) by arranging communications in some configuration which does not restrict communication to isolated pairs. We have identified an empirical methodology for examining the questions associated with computer mediated tandem language learning that seems to be providing substantive answers.

References

- Appel, M. C. & Mullen, T. (1998). A Common Gateway Interface for Tandem Language Learning. In Strotmann, B. (Ed.), *Proceedings of the International Congress on Technology in Teaching*. IATEFL Computer SIG, TESOL-Spain, Universidad Europea de Madrid. 20-22 November 1998.
- Appel, M. C. & Mullen, T. (2000). Pedagogical Considerations for Web-Based Tandem Language Exchange. *Computers and Education*, 34(3-4), 291–308.
- Branigan, H. P., Pickering, M., Liversedge, S., Steward, A., & Urbach, T. (1995). Syntactic Priming: Investigating the Mental Representation of Language. *Journal of Psycholinguistic Research*, 24, 489–506.
- Fodor, J. A. (1983). *The Modularity of Mind*. Cambridge: MIT Press.
- Garrod, S. & Doherty, G. (1994). Conversation, Co-ordination and Convention: an Empirical Investigation of How Groups Establish Linguistic Conventions. *Cognition*, 53, 181–215.
- Granger, S. (Ed.). (1998). *Learner English on Computer*. Austin, TX: Addison Wesley Longman.

- Hewson, C., Laurent, D., & Vogel, C. (1996). Proper Methodologies for Psychological and Sociological Experiments Administered via Internet. *Behavior Research Methods, Instruments, & Computers*, 28(2), 186–191.
- Little, D. & Brammerts, H. (1996). A Guide to Language Learning in Tandem via the Internet. Tech. rep. 46, Center for Language and Communication Studies, Trinity College, University of Dublin. CLCS Occasional Paper.
- Schmid, H. (1994). Probabilistic Part-of-Speech Tagging Using Decision Trees. In *International Conference on New Methods in Language Processing*, pp. 44–49. Manchester, UK.
- Schober, M. F. & Clark, H. H. (1989). Understanding by Addressees and Overhearers. *Cognitive Psychology*, 21, 211–32.
- Swinney, D. A. (1979). Lexical Access During Sentence Comprehension: (Re)Consideration of Context Effects. *Journal of Verbal Learning and Verbal Behaviour*, 18, 645–59.
- van Noord, G. (1998) <http://odur.let.rug.nl/vannoord/TextCat/>. last verified 25 June, 2001.

Biographies

Christine Appel is a Lecturer in the School of Applied Languages and Intercultural Studies at Dublin City University. She has an MPhil in Applied Linguistics from Trinity College, University of Dublin, where she is also a Ph.D. candidate.

Carl Vogel has a Ph.D. in Cognitive Science from the University of Edinburgh and is Lecturer in Computational Linguistics at Trinity College.