THE PEDAGOGICAL MEDIATION OF A DEVELOPMENTAL LEARNER CORPUS FOR CLASSROOM-BASED LANGUAGE INSTRUCTION

Julie A. Belz
Indiana University-Purdue University Indianapolis
Nina Vyatkina
University of Kansas

Although corpora have been used in language teaching for some time, few empirical studies explore their impact on learning outcomes. We provide a microgenetic account of learners' responses to corpus-driven instructional units for German modal particles and pronominal *da*-compounds. The units are based on developmental corpus data produced by native speakers during interactions with the very learners for whom the units are designed. Thus, we address the issue of authentication in corpus-driven language pedagogy. Finally, we illustrate how an ethnographically supplemented developmental learner corpus may contribute to second language acquisition research via dense documentation of micro-changes in learners' language use over time.

INTRODUCTION

This article explores the pedagogically mediated application of a learner corpus in language teaching and in the developmental analysis of second language acquisition (SLA). We distinguish between *longitudinal* and *developmental* analyses with regard to the density of observation of learner performance over time. While 'longitudinal' may refer to analyses in which waves of data are elicited at more distant intervals (e.g., at the beginning and end of a semester), we reserve the term 'developmental' for those analyses in which learner performance is documented at close intervals or at all points of production. We refer to the learner corpus under study as a *Developmental Learner Corpus* (DLC)[1] because it contains the complete record of all native speaker (NS) and non-native speaker (NNS) interactions during two-month telecollaborative partnerships (see Belz, 2002; Belz & Thorne, 2006; O'Dowd, 2007) The interlocutors are English NSs who are learning German and German NSs who are learning English. Because participants use both languages, our data are bilingual and consist of native (L1) English, second-language (L2) English, L1 German, and L2 German.

We begin with a review of the application of corpora in language teaching to date. Next, we describe *in vivo* classroom applications of our DLC wherein the very learners who produce the corpus data revisit it in the form of pedagogically mediated teaching materials. Thus, the learners in our experiments encounter their own productions (as well as the productions of their NS keypals and their US classmates) as a subsequent source of didacticized input (Swain & Lapkin, 1995). We further illustrate how the dense performance records archived in our DLC facilitate the close, corpus-driven[2] tracking of micro-changes in learner language use over time.

First, we report on a "developmental pedagogical intervention" (Belz & Vyatkina, 2005, p. 20) for four German modal particles (MPs): *ja, denn, doch,* and *mal*.[3] We chose these four MPs in particular because the NSs in our corpus used them most frequently. The MPs in general were chosen for examination because they are important carriers of interpersonal meaning. As Weydt (1969) puts it, "[t]he German listener expects a particle. If it is absent, the sentence acquires a specific stylistic value: without a particle it sounds choppy, harsh, unfriendly, its utterance is apodictic, abrupt, blatantly noncommittal" (p. 20; our translation). Despite the communicative significance of the MPs, early *in vivo* contrastive analysis revealed a marked mismatch between learner and NS uses. Previous studies have shown that the MPs are difficult to master and that there are few robust materials for their instruction (Götze, 1993; Möllering,

2004). Our MP intervention is characterized by focused instruction based on corpus-enabled contrastive and developmental analyses.

Next, we describe a developmental pedagogical intervention for the German *da*-compounds, which function as cohesive devices and contribute to textual coherence. They were chosen for instruction because early corpus-driven contrastive analysis revealed a mismatch between learner and NS uses. Previous studies have shown that learners under-use them with respect to frequency, range, and syntax; and they are often inadequately addressed in teaching materials (Belz, 2005). Again, the compound intervention was designed in response to contrastive and developmental analyses of learner performance that were enabled *in vivo* by our DLC and the materials were used with the same learners who participated in the interactions on which they are based.

Finally, we illustrate how a richly annotated DLC can be utilized to address issues of SLA – in conjunction with ethnographic data and telecollaborative pedagogy – by facilitating developmental analyses of individual pathways of L2 use during specific learning events. In previous studies, we have exemplified the development of entire cohorts of learners in our experimental courses (e.g., Belz, 2005; Belz & Kinginger, 2003; Belz & Vyatkina, 2005; Vyatkina, 2007; Vyatkina & Belz, 2006). One salient finding emerging from these studies is the idiosyncratic nature of learners' development toward NS norms and the sensitivity of that development to sociocultural factors, including history, agency, and structure (Belz, 2002; Layder, 1993). In order to illustrate the ways in which we have utilized our DLC to mediate the elucidation of individual developmental pathways and their sensitivity to learners' lives and (instructional) experiences, two focal learners, Christie and Jeremy (pseudonyms), are examined. Such dense examinations not only offer insight into SLA at the level of individual micro-changes in performance and awareness over time, but they also exemplify how a pedagogically mediated DLC may afford a tight synergy between classroom-based research and praxis.

CORPORA AND LANGUAGE INSTRUCTION

Johns (1986) introduced the term *data-driven learning* to describe the use of NS concordances in language teaching (Sinclair, 2003). To date, corpus-based teaching proposals primarily have focused on vocabulary (Coxhead, 2002; Römer, 2004) and grammar (Mindt, 2002), although a few studies target pragmatic and discourse features (Belz, 2008; Möllering, 2001, 2004). In most corpus-driven teaching proposals, learners examine teacher-prepared corpus excerpts, but are unable to explore the corpus on their own. In contrast, Bernardini (2002, 2004) advocates discovery learning wherein learners browse a NS corpus in order to discover patterns of use, establish meanings, and facilitate reading, writing, and translation (see also Aston, 2002, p. 10; Leech, 1997; St. John, 2001).

Seidlhofer (2002) suggests *learning-driven data* in addition to Johns' *data-driven learning*. In this method, learner productions are systematically archived for comparison with NS corpora in order to give learners an idea of "where they are…situated in their L2 learning context, and where they eventually (may) want to [go]" (Seidlhofer, 2002, p. 215). Granger (1998) refers to the comparison of NS and learner corpus data as *contrastive interlanguage analysis* (CIA).

Although learner corpora are "a fairly recent phenomenon" (Nesselhauf, 2004, p. 125; Pravec, 2002), many researchers already have employed CIA (Granger & Tribble, 1998; Granger, Hung, & Petch-Tyson, 2002). Notably, the research team associated with the *International Corpus of Learner English* has suggested corpus-driven teaching materials in the areas of lexicography (Rundell & Granger, 2007) and phraseology (De Cock, 2004; Nesselhauf, 2005). Nevertheless, "it is less clear how – and when – [such suggestions] can best be carried over into effective pedagogical practice" (Lee & Swales, 2006, p. 57); indeed, many corpus-driven materials are still awaiting empirical testing of their influence on learning outcomes.

There are a number of limitations connected with corpus-driven language pedagogy, including the small number of languages and text types represented in existing corpora and the lack of developmental corpora. Another major issue is the controversy over the authenticity of corpus data (see Seidlhofer, 2003, pp. 77-123). On the one hand, proponents of corpus-driven language pedagogy suggest that a key advantage to this approach is the *genuine* nature of NS corpus data in contrast to "concocted" (McCarthy & Carter, 1995, p. 370) textbook examples. On the other hand, Widdowson (1991, 2000) questions the alleged authenticity of corpus samples by claiming that authenticity is a function of text reception as well as text production. If learners are unable to authenticate or create an appropriate interpretive context for NS corpus samples, then they may fare no better than "concocted" texts as mediators of L2 learning. Prodromou (1996) supports Widdowson by pointing out that corpus language is "extremely context and culture-bound" and therefore "inappropriate to the needs and interests" of learners (p. 372). Although Bernardini (2002) advocates discovery learning, she nevertheless warns that it relies heavily on learners' "unflagging curiosity and interest" which may quickly wane if learners become lost in the vast "unknown land" of large corpora (p. 167). In a similar vein, Thurston and Candlin (1998) warn that ''[o]ver-exposure to concordance lines can conceivably tire students'' (p. 278). Indeed, Möllering (2004) reports that learners found working with concordances "slightly overwhelming" (p. 245).

Although this controversy addresses the use of NS corpora, it is equally applicable to learner corpora. The latter are not more relevant to the interests of a particular group of learners than the former and therefore do not necessarily resolve the issue of authentication.

THE STUDY

The pedagogical interventions were implemented in the context of *telecollaboration*, a language learning configuration in which distally located learners use Internet communication tools for social interaction, dialogue, and debate with NS age peers (Belz & Thorne, 2006; O'Dowd, 2007). In the focal course, English-speaking learners of German at a large public university in the United States engaged in genuine interactions with German-speaking learners of English at a German teachers' college for eight weeks during Fall 2005. These participants were divided into four transatlantic groups, each of which had both American and German members. Learners chatted and emailed with their other group members during class time. Some participants also chatted on their own time. All correspondence was archived in a web-based teleconferencing program as they were produced. Each group had their own electronic folder in which all their correspondence was maintained, but all participants had access to all folders.

Emails and chats were entered regularly into a locally designed DLC in association with 20 learner and task variables while the course under study progressed. When the focal learners were not chatting/emailing during class time, they explored the correspondence of other groups in both the folders and as DLC excerpts under the guidance of their instructor; therefore, all learners developed a level of familiarity with the telecollaborative correspondence during the semester. As a result, the learners in this study functioned as both participants and observers with respect to the DLC data (Gavioli & Aston, 2001). As participants, they engaged in the interactions that subsequently formed the DLC. As observers, they retrospectively examined these data and critically analyzed them under the guidance of the teacher.

During non-telecollaborative class sessions, corpus-driven pedagogical interventions were delivered that were designed in response to the learners' (emerging) use of the focal features as archived in the DLC. In general, the interventions were informed by the principles of form-focused instruction, which "involves attempts to intervene directly in the process of interlanguage construction by drawing learners' attention to or providing opportunities for them to practice specific language features" (Ellis, Basturkmen, & Loewen, 2001, p. 407). In addition, however, our pedagogical approach combines the focus on specific linguistic forms with meaning-focused and task-based instruction because the focal forms are taught in the context of meaningful interactions (Kasper & Rose, 2002) and practiced by students in communicative tasks (Ellis, 2006) in the form of their on-going telecollaborative correspondence.

Moreover, the very need for the pedagogical interventions in this study was ascertained on the basis of students' documented performance during meaningful language use, i.e., an under-use of the focal features by the learners in comparison with the NSs. Positive evaluative feedback constituted a part of the intervention in that learners were encouraged to practice the focal features after each interventional stage. When they did so, their uses were pointed out and praised *in plenum* in the form of DLC excerpts.

The ethnographic component of the data set includes biographical and technological surveys, post-telecollaboration interviews, formative and summative individual learner portfolios, and field notes based on exhaustive participant observation by both authors. The DLC was used to perform a microgenetic analysis of individual learners' use of the focal features in relation to the pedagogical interventions, time, and the performance of their classmates and NS keypals. Microgenesis entails tracing the history (i.e., genesis) of processes over relatively short periods of time in order to understand "how the human mind functions as a consequence of its formation in cultural activity" (Lantolf & Thorne, 2006, p. 57). The dense documentation of learner performance in the DLC allows us to examine the process of SLA in blow-by-blow fashion for any searchable linguistic feature. The ethnographic data allow us to situate each micro-change in the developmental process within the cultural activity of its origin.

THE MODAL PARTICLE INTERVENTION

The four-stage MP intervention (Figure 1) follows Möllering and Nunan (1995; see Vyatkina & Johnson, 2007, for examples of handouts used in the MP intervention); important procedural differences capitalize on the developmental nature of our DLC.

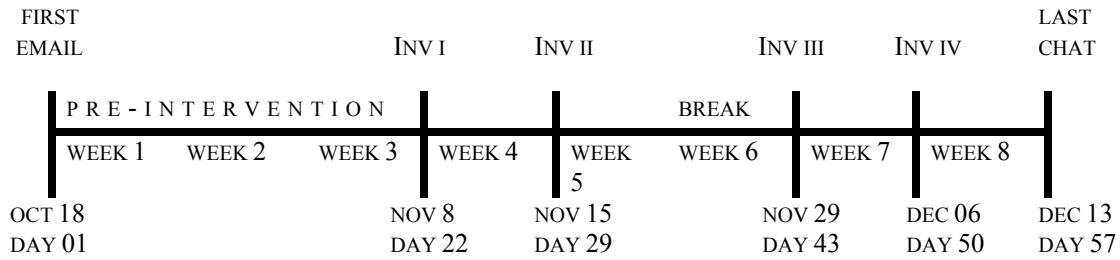

Figure 1. Modal particle intervention timeline.

Pre-Intervention Stage

The pre-intervention data enabled the establishment of a control and a comparative baseline (learner and NS productions, respectively) for the pedagogical experiment. Table 1 below reports NS and learner uses of the focal MPs prior to any intervention for the entire 2005 cohort. Only 2 learners used only 2 MPs one time each, while the NSs used 51 MPs. This result reinforces Möllering and Nunan's (1995) finding that mere exposure to MP-rich texts "without focus on the particles [themselves] seems to have no influence on performance" (p. 58).

Table 1. Pre-Intervention NS and Learner MP Uses.

Category	All NSs	All Learners
Total German Words Pre-Intervention	11,887	9,588
Total MPs Pre-Intervention	51	2
Rate per 1000 German Words Pre-Intervention	4.29	0.20

Intervention Stage I

The objectives of instruction module 1 were the collection of pre-test metapragmatic awareness data and awareness-raising focused instruction based on an enhanced condition (Robinson, 1997, p. 224), which

involved highlighting the focal features in the instructional materials without giving explanations about their use. This 40-minute module was administered at the start of Week 4.

First, we defined "pragmatics" and "pragmatic competence." Next, learners filled out two metapragmatic awareness questionnaires. The first questionnaire presents two short dialogues in German (Weydt, Harden, Hentschel, & Rösler, 1983). Dialogue A contains MPs, while dialogue B does not. The learners underlined the words that were different in the two dialogues and ranked each dialogue with respect to its perceived expressive force. The second questionnaire was similar, but contained examples from the DLC (taken from all NSs in the given cohort). We bolded the MPs, projected the examples on a large screen, and told the learners that the bolded words carried expressive force. Finally, we gave the learners handouts with the DLC excerpts and advised them to observe the NSs' use of the bolded words until the next instruction module in one week. They chatted with their keypals immediately after instruction module 1.

The learners identified most of the MPs in dialogue A and in the DLC excerpts. Similar to NSs, they perceive dialogue A to be more natural, warmer, friendlier, more fluent, and more authentic than dialogue B. Despite these NS-like intuitions, however, the performance data show that only the most proficient learner, Christie, used the MP *ja* once during stage I (Table 2 below; see also Belz & Vyatkina, 2005). These results suggest that enhanced instruction alone does not trigger immediate performance.

Table 2. NS and Learner MP Uses in Stage I.

Category	All NSs	All Learners
Total German Words after Intervention I	5,025	2,552
Total MPs after Intervention I	38	1
Rate per 1000 German Words after Intervention I	7.56	0.39

Intervention Stage II

The objectives of instruction module 2 were the collection of additional metapragmatic awareness data and explicit form-focused instruction, which involves directing the learners' attention to both form and meaning while explaining the underlying rules of MP use (Robinson, 1997). This 40-minute module was administered at the start of Week 5.

First, the learners answered questions related to meta-linguistic knowledge and meta-pragmatic awareness based on the DLC excerpts used in module 1. They were asked to name the word category of the MPs, list other words belonging to this category, select the functions that these words can fulfill from a given list, and list any words that they have used for the same functions.

Only one learner, Saul, accurately used the term "modal particle"[4]. Six out of seven thought that MPs make communication more natural, while five wrote that MPs help create a friendly atmosphere and indicate colloquial/informal communication. These findings suggest that learners' 'feel' for the pragmatic force of the MPs does not automatically translate into meta-linguistic knowledge and productive ability.

After introducing the term "modal particle", we lectured briefly on the meanings, functions, syntactic restrictions, and homonyms of the four focal MPs. To illustrate, one homonym of the MP *ja* is the answering particle *ja* 'yes', while one homonym of the MP *denn* is the causal conjunction *denn* 'because'. We gave the learners a summary of this information on a handout. Next, we showed the learners additional NS MP uses and their homonyms taken from the DLC. Then we examined Christie and Saul's pre-intervention uses. Finally, we advised the learners to observe their keypals' use of the MPs in pending correspondence and to use them themselves. The learners chatted with their keypals immediately after instruction module 2.

In this chat, a third learner, Jeremy, began using MPs. To illustrate, in turn (3) of example (1), Jeremy uses *ja* in a formulaic pattern 'This + is + *ja* + attribute' (Möllering, 2004, p. 234) for intensifying a positive appraisal. In turn (5), he uses the same MP in a free construction to mark reference to shared knowledge. Both uses are syntactically accurate and pragmatically appropriate. Although Jeremy's German partners also use the MP *ja* later in the same chat, Jeremy's uses precede their uses and thus cannot be attributed to direct imitation.

Example (1)

(1) **Jeremy:** Ich bin gut. 'I am good.'

(2) **Carine:** Mir geht's auch gut! 'I am also good!'

(3) **Jeremy:** Oh das ist ***ja*** gut. 'Oh this is **ja** good.'

[...]

(4) **Alma-Lora:** Genau, wir können auch die deutschen fragen, was sie über amerikanischen Patriotismus denken und umgekehrt. 'Exactly, we can also ask the Germans what they think about American patriotism and the other way around.'

(5**) Jeremy:** Ja, und wir konnen ***ja*** die Amerikanische Leute uber Deutsches Patriotismus. 'Yes, and we can **ja** [ask] American people about German patriotism.'

[...]

(November 15)

In a later email, Jeremy uses *ja* six times (once in a formulaic pattern and five times in free constructions) and *denn* once. All uses are pragmatically appropriate, although the syntax of *ja* is twice inaccurate. In comparison to the NSs, Jeremy overuses *ja* by a margin of 7. This initial overuse may indicate Jeremy's desire to experiment actively with a new feature.

Christie does not use any MPs in this stage. Our participant observation helps account for this finding. Christie and her American partner participated collaboratively in a chat during stage II, but the partner, a much less proficient learner, typed. Additionally, there are only 69 German words in this chat. Further, Christie wrote three emails between instruction modules 2 and 3, but two of these emails contain mostly project entries (argumentative, monologic writing), where MPs are not likely to occur. Therefore, the absence of MPs at this stage may be attributed to her lack of general correspondence in German. Learner and NS MP uses for stage II are summarized in Table 3.

Table 3. NS and Learner MP Uses in Stage II.

Category	All NSs	All Learners
Total German Words after Intervention II	2,947	3,605
Total MPs after Intervention II	26	9
Rate per 1000 German Words after Intervention II	8.82	2.50

Intervention Stage III

The primary purpose of instruction module 3 was fine-tuned instruction on the meaning and syntax of the MPs. It was administered at the start of Week 7 and lasted for 40 minutes. First, we showed the learners new NS MP uses since instruction module 2. Importantly, we highlighted Jeremy's new uses and offered recommendations for future use.

Next, we gave learners four worksheets as a homework assignment. Each worksheet contained concordance lines of NS MP uses from the entire 2005 cohort to date plus questions that were designed to stimulate their discovery of patterns in the data. As in previous stages, the learners chatted with their

keypals immediately after instruction module 3. In the ensuing week, Jeremy used *ja* once in an appraisal pattern as seen below.

Example (2)

> **Jeremy:** du hast ***ja*** besser Antworten 'you have **ja** better answers'
>
> (November 28)

Christie used *mal* and *ja* once each in the chat immediately following instruction module 3. Both MPs occur in fixed patterns as exemplified on handouts: an appraisal pattern with *ja* and an informal singular imperative with *mal*. In a later email, Christie uses *ja* twice and *denn* once as seen in example (3).

Example (3)

> An Thanksgiving habe ich bei meinem Onkel gegessen. Meine ganze Familie war dabei und wir haben alle ***ja*** viel gegessen! [...] wann habt ihr winterferien? wie lange habt ihr ***denn*** frei? wir haben ab 17.12 Ferien, bis 9.1. So ca. 3 Wochen haben wir frei. Hoffentlich arbeite ich zwischendurch und verdiene ein bisschen Geld! Ich brauche ***ja*** Geld um Buecher fuer die Universitaet zu kaufen.
>
> 'On Thanksgiving, I had dinner at my uncle's. My whole family was there and we all ate **ja** a lot! [...] when do you have a winter break? How long are you **denn** free? We have a break from 12/17 to 1/9. So, we are free for about 3 weeks. Hopefully, I will work during this time and make a bit of money! I need **ja** money to buy books for he university.'
>
> (November 29)

Notably, Christie uses *ja* in free constructions for the first time. The asynchronous nature of the email may have given Christie more time for monitoring her production and using new MP functions. All five of Christie's MP uses during this stage are appropriate and accurate. All learner and NS MP uses during stage III are summarized in Table 4.

Table 4. NS and Learner MP Uses in Stage III

Category	All NSs	All Learners
Total German Words after Intervention III	883	1,332
Total MPs after Intervention III	9	6
Rate per 1000 German Words after Intervention III	10.19	4.5

Intervention Stage IV

The fourth and final instruction module was administered at the start of Week 8. The primary purpose of this 30-minute unit was fine-tuned instruction on MP use. The session began with a concordancing task. We gave learners a handout containing a summary of demonstrated MP patterns and examples of conventionalized MP expressions. Then we showed them more DLC examples of emergent learner MP use taken from the whole cohort. We discussed accurate and inaccurate uses and related these uses to NS patterns encountered during the concordancing task. Finally, we encouraged learners to use MPs in their own writing. A chat session followed instruction module 4. The learners communicated with their keypals for one more week until the close of the American semester.

During this week, the learners' use of the MPs increased sharply, and quantitatively surpassed NS use in the same week as seen in Table 5 below. NS MP uses also increase during this stage. One explanation for the NS increase may lie in the nature of the task. At the beginning of the partnership, keypals were required to introduce themselves and their institutional contexts to their partners. Such descriptions entail a monologic narrative register in which fewer interpersonal features (e.g., MPs) occur. During the post-intervention period, the learners were primarily completing the joint web-project that involved many

discussions carried out in an interactive, dialogic register. Interpersonal linguistic features cluster in such text types (Biber, 1988).

Table 5. NS and Learner MP Uses in Stage IV.

Category	All NSs	All Learners
Total German Words after Intervention IV	3,317	3,704
Total MPs after Intervention IV	44	55
Rate of MP Uses per 1000 German Words after Intervention IV	13.26	14.85

For example, Jeremy uses MPs 9 times in stage IV. His last chat is particularly rich in MPs: *ja* (5x), *doch* (1x), and *denn* (1x). All uses are pragmatically appropriate and structurally accurate except for the use of *denn* in a declarative sentence. *ja* occurs three times in the formulaic construction *Das ist ja gut* 'That is *ja* good'. Although this phrase is pragmatically appropriate, it is not NS-like with regard to collocational patterns. NSs frequently use *ja* and *doch* in positive appraisal patterns with more emotionally loaded attributes such as *super, cool,* and *echt lustig* 'really funny'. Additionally, NSs often use *ja* with inverted word order (e.g., in questions), whereas Jeremy uses only direct word order (e.g., in statements). Finally, Jeremy overuses the appraisal pattern with *ja* in comparison to NSs.

Christie produces 8 focal MPs in chat: *ja* (5x), *doch* (1x), *denn* (1x), and *mal* (1x). All uses are pragmatically appropriate and syntactically accurate except for a word order mistake with *denn*. She also appropriately uses an additional MP (*eigentlich*). *ja* is used twice in free constructions and three times in the appraisal pattern. In contrast to Jeremy, Christie uses several variations of the fixed pattern *Das + ist + ja + attribute* as seen in example (4). The verb *sein* is used in the past tense in (4a) in the subjunctive mood in (4b); additional constituents are added and the word order is reversed in (4c). This variation provides evidence of Christie's developing ability to use *ja* not only idiomatically but also creatively.

Example (4)

a) das war ***ja*** ein bisschen weit! 'this was **ja** a bit far*!'*

b) das waere ***ja*** cool! *'*this would be **ja** cool!*'*

c) heute ist es hier ***ja*** ganz schoen kalt! 'today is it here **ja** really very cold!*'*

(December 13)

Post-intervention: Cumulative Portfolios

Christie confirms in her cumulative, written course portfolio (which functions as a source of post-intervention meta-pragmatic awareness data) that she was not aware of the MPs before the course and did not notice them in her partners' pre-intervention writing. She noticed how frequently her partners used them "very soon" after the intervention and began using more of them in her own writing.[5] Furthermore, Christie includes an email excerpt from a personal German friend in which she accurately highlights 11 of 12 MPs. This application of her MP knowledge in a new context is an important piece of evidence for the development of Christie's meta-pragmatic awareness.

Jeremy reports in his portfolio that he previously noticed MPs, but did not know their meaning. He includes several emails from his keypals in which he highlighted sentences containing MPs and demonstrates his increased awareness by commenting on one particular use of *mal* by his German partner: "Du solltest wirklich *mal* nach Deutschland kommen" 'You should really *mal* come to Germany'. Jeremy explains that *mal* makes the sentence friendlier and less direct, transforming a command into an invitation. Finally, Jeremy aptly terms the MPs "small words with big meaning" that "make a big difference in a German sentence."

THE *DA*-COMPOUND INTERVENTION

Semantically, the German *da*-compounds function as cohesive ties that link topics across phrases, clauses, larger units of text, and even turns-at-talk. Syntactically, they occur as both anaphoric and cataphoric proforms that index noun phrases as well as larger units of language. To illustrate, in example (5) *dazu* is used anaphorically to refer to the fact that the addressee swims, while, in example (6), *dadurch* is used cataphorically to refer to the fact that the speaker and his companions stayed away from areas where there were many tourists. (The *da*-compounds are underlined, while their referents are double underlined.)

Example (5)

> Ah, du schwimmst. Wie oft machst du das denn? Machst du noch andere Sportarten? Wie bist du ***dazu*** gekommen?
>
> 'Oh, you swim. How often do you do that? Do you do other sports? How did you come **to that** [i.e., how did you get involved in that]?

Example (6)

> ***Dadurch***, dass wir uns von den touristischen Gebieten weitgehend ferngehalten haben, haben wir sehr viele Einheimische kennen gelernt und sind in Familien eingeladen worden zum Essen.
>
> '**By virtue of the fact** that we kept away from the tourist areas we got to know a lot of locals and were invited to eat in the homes of families.'

Belz (2005) ascertained that learners under-use the *da*-compounds by a margin of 0.7 to 1 in comparison to NSs. She also found that NS uses indexed a unit larger than a noun phrase 65% of the time, while learner uses indexed a larger unit only 25% of the time. 21% of NS uses were in cataphoric constructions, while only 8% of learner uses fell in this category. Finally, just 12% of NS uses occurred in verb-and-preposition phrases (e.g., *Ich freue mich darauf!* 'I'm looking forward to it!'), whereas 47% of learner uses occurred in these environments.

Similar patterns emerged in an early contrastive interlanguage analysis of the current data set. Therefore, we determined that the *da*-compounds were an appropriate feature for intervention among this particular group of learners.

Pre-Intervention Stage

The timeline for the *da*-compound intervention is given in Figure 2. The pre-intervention phase lasted 31 days, when instruction module 1 was delivered. The length of each intervention depended on the components of the overall syllabus and instructional schedules at both institutions, as well as the time required for data entry, feature analysis, and subsequent materials development.

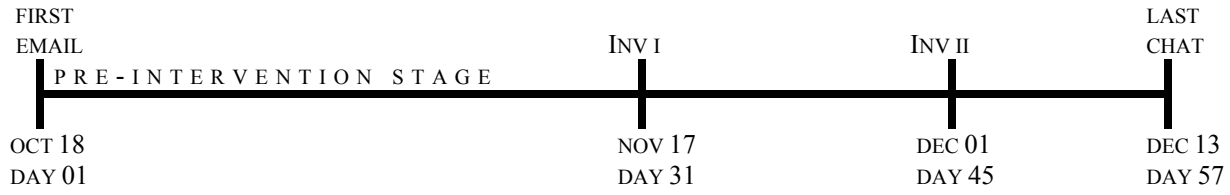

Figure 2. *da*-Compound intervention timeline.

Table 6 lists aggregate, pre-intervention NS and learner uses of the compounds. Results show a normalized rate of 5.57 *da*-compounds per 1000 German words for the NSs and 1.71 for the learners. As with the MPs, the two most proficient learners in the group, Christie and Saul, used the most *da*-compounds during the pre-intervention stage. Christie used a total of five different *da*-compounds seven

times, while Saul used six compounds 11 times. In the following subsections, we again provide a detailed report on the progression of our two focal learners, Christie and Jeremy.

Table 6. Pre-Intervention NS and Learner *da*-compound Uses

Category	All NSs	All Learners
Total German Words Pre-Intervention	19,736	15,161
Total *da*-compounds Pre-Intervention	106	26
Rate per 1000 German Words Pre-Intervention	5.57	1.71

Intervention Stage I

Instruction module 1 consisted of four parts and lasted for 40 minutes. First, we explained that *da*-compounds are similar to personal pronouns in that they can index full noun phrases as well as larger stretches of text. In part 2, we elicited awareness data via the following seven prompts: 1) Name as many *da*-compounds as you can; 2) Do you think that you have used *da*-compounds with your keypals? If so, which ones? 3) Do you think that your keypals have used *da*-compounds? If so, which ones? 4) How do you think your rate of *da*-compound use compares to your keypals' rate of use? 5) Which *da*-compounds have your keypals used most frequently? least frequently? 7) Do you know any phrases in which *da*-compounds occur?

In response to the first prompt, Christie listed ten of 18 total possible *da*-compounds or 56%. She thought she used *damit, danach,* and *davon*, when, in fact, she had only used *davon*; however, she had also used *dafür, dazu, dazwischen,* and *darüber*. Syntactically, her uses represent anaphoric reference to either the object of a prepositional phrase (1x), a noun phrase (4x), or a larger unit (2x). In response to question 3, Christie accurately reported that the keypals had used *damit, danach, davon, dazu,* and *dabei*, but she failed to mention that they also had used 11 additional compounds, including the two most frequent, *darauf* and *dazu*. In response to question 6, Christie listed *daraus* and *darin*. *daraus* was indeed one of two compounds that was not used by the NSs and *darin* was used only once. In response to question 7, Christie wrote the following three sentences: 1) *Ich bin dagegen/dafür* 'I am against/for that'; 2) *Damit kann ich leben* 'I can live with that'; and 3) *Darauf bin ich stolz* 'I am proud of that'.

Jeremy, in contrast, listed only *dafür, damit,* and *davon* (17%). He also inaccurately listed *dass* (a subordinating conjunction) and *da* (a subordinating conjunction and adverb of place). In response to question 2, Jeremy stated that he had not used any *da*-compounds; in fact, he had once used *dazu* but inaccurately. Jeremy thought that the keypals had used many more compounds than he had, but the only one that he could remember was *damit*. He was not sure which compounds were used the least and he could not list any phrases containing them.

During part 3, we showed the learners a table in which all NS and learner uses of the *da*-compounds during the first three weeks of the partnership were compared.[6] The following points were emphasized: 1) the Germans used *da*-compounds 7 times as frequently as the Americans did; 2) the compounds that are used most frequently by the Germans are not the compounds that are used most frequently by the Americans; and 3) five learners each used at least one compound for a combined total of 11 uses (7 accurate) in the pre-intervention stage, while two learners used no compounds. The goal of part 3 was to raise learners' awareness of their own performance in comparison to both their NS keypals and their US classmates in order to allow them to direct their attention to those segments of the pending instruction that would be most relevant to their individual situation.

In part 4, we gave the learners a handout that contained each NS *da*-compound use up to that point as well as their own 11 uses. We told them to read the handout thoroughly before the next module two weeks later. In addition, they were asked to note any differences between NS and NNS uses, to look for *da*-

compounds in their partners' future emails/chats, and to use the compounds in their own correspondence, where appropriate.

Following instruction module 1, the learners met for just one more class session (the Tuesday before Thanksgiving). Consequently, there was little correspondence between modules 1 and 2. Further, the learners were not able to chat with their keypals following the individual modules of the *da*-compound intervention as they could for the MPs, because of scheduling.

Nevertheless, all learners produced a total of 14 *da*-compounds between modules 1 and 2 (Table 7). The learners' rate of *da*-compound use per 1000 German words increased threefold from 1.71 during the pre-intervention stage to 5.12 during stage I, but still lagged behind German rates by a margin of about 1.7 to 1 (see below for an explanation of the NSs' increased rate of use).

Table 7. NS and Learner Compound Uses in the Pre-Intervention Stage and in Stage I.

Category	All NSs	All Learners
Total German Words Pre-Intervention	19,736	15,161
Total *da*-compounds Pre-Intervention	106	26
Rate per 1000 German Words Pre-Intervention	5.57	1.71
Total German Words After Intervention I up to Intervention II	4,109	2,734
Total *da*-compounds After Intervention I up to Intervention II	35	14
Rate per 1000 German Words After Intervention I up to Intervention II	8.52	5.12

While Jeremy did not produce any compounds between modules 1 and 2, Christie used five different compounds (*darüber, darauf, dabei, dazu,* and *darum*) one time each, three of which (*darauf, dabei,* and *darum*) she did not use in the pre-intervention stage. From a structural perspective, Christie still uses *da*-compounds only in anaphoric constructions in stage 1.

Intervention Stage II

Instruction module 2 was delivered on Day 45. It consisted of three parts and lasted 45 minutes. The main purpose of part 1 was to introduce the learners to the range of structures in which the *da*-compounds occur. First, we introduced learners to the concept of cataphoric reference. Next, we showed the learners examples in which *da*-compounds anaphorically index noun phrases and larger pieces of text. After that, we showed learners examples in which the *da*-compounds cataphorically index a subordinating *dass*-clause, an infinitive phrase, or a subordinating *wenn*-clause. The given examples were drawn from the NS keypals' telecollaborative discourse archived in the DLC.

Part 2 was a whole-group recognition task during which we gave the learners a handout containing 10 *da*-compound examples. First, the learners underlined the piece of text to which each *da*-compound refers. Then, they stated whether or not the given *da*-compound occurred in an anaphoric or cataphoric construction.

Finally, we gave the learners a one-page summary of what they had learned about the *da*-compounds, including summary statements about *da*-compound use based on the results of the empirical analysis of the corpus data. We also offered suggestions for how the learners could improve with respect to *da*-compound use (see Appendix for the full summary sheet).

Table 8. NSs and Learners Use of *da*-Compounds after Intervention II.

Category	All NSs	All Learners
Total German Words After Intervention II to Semester Close	8,713	4,629
Total *da*-compounds After Intervention II to Semester Close	64	36
Rate per 1000 German Words After Intervention II to Semester Close	7.34	7.78

After instruction module 2, the learners' rate of compound use increased from 1.71 in the pre-intervention stage to 7.78 (Table 8), which is actually greater that the NSs' rate in the same period.

Unlike stage I, Jeremy uses three *da*-compounds in stage II: *dafür* (2x) and *dazu* (1x). Each compound is used in an anaphoric construction and all uses are accurate with the exception of word order in one case. In his portfolio, Jeremy states that his German keypals used many *da*-compounds, but that he initially had "no idea" what they meant or how important they were. He then gives several examples from his NS partner's emails and correctly labels them as anaphoric. Jeremy explains that he made a first conscious effort to use *dafür* and *dazu* in his final email on Day 50. Nevertheless, Jeremy characterizes the *da*-compounds as difficult and attributes this difficulty to the fact that the compounds do not exist in English *per se*. Lastly, Jeremy explains that he has re-read his keypals' email and now recognizes many uses of *dafür, damit, dazu*, and *davon* in them. He no longer finds the emails confusing because he better understands what the *da*-compounds mean and how they stitch pieces of text together.

In contrast to Jeremy, Christie uses just one *da*-compound (*darauf*) in an anaphoric construction after instruction module 2. Thus, her telecollaborative performance data alone would seem to indicate that she does not develop much after this stage; however, her portfolio data present a different picture. In her final portfolio, Christie included an entry on the *da*-compounds which consists of a reflective commentary, a copy of the notes she took during instruction module 2, and a copy of an email she wrote to her German friend after instruction module 2.

In her commentary, Christie states that she did not know much about *da*-compounds before discussing them in class. She further explains that she did not know that there were anaphoric and cataphoric compounds and that she did not really notice them in her partners' emails. Christie remarks that she tried to use more *da*-compounds after the intervention. To illustrate, she notes that she used six *da*-compounds in the email to her German friend discussed above. In fact, Christie does indeed accurately use five different compounds six times. What is more important is the fact that three of these uses are cataphoric as seen in examples (7) to (9):

Example (7)

> Freue mich schon ***darauf***, eine Mail von dir zu bekommen!
>
> 'I am looking forward **to** getting an email from you!'

Example (8)

> Meine Mutter meinte, dass eure Paket schon angekommen ist. Habt ihr was ***dagegen***, wenn ich es vor Weihnachten auspacke?
>
> 'My mother says that your package has already arrived. Do you have anything **against it**, if I open it before Christmas?'

Example (9)

> Hoffentlich bekommst du bald mein Paket! Freue mich schon ***darauf***, wie dir deine Geschenke gefallen.
>
> 'Hopefully you will get my package soon! I am looking forward **to** [finding out about] how you like your presents.'

Examples (7) to (9), demonstrate considerable development not only because they are Christie's first attested uses of cataphoric *da*-compounds, but also because of the range of structures in which they occur. In example (7), Christie uses *darauf* to refer ahead to the entire subsequent infinitive clause. This use closely mimics the example given during part 1 of instruction module 2. In example (8), Christie appropriately uses *dagegen* to cataphorically index the subsequent *wenn*-clause. This use marks her as an

advanced speaker because such constructions are numerically under-represented, even for NSs. In her final example, Christie appropriately uses *darauf* to cataphorically index the content of a subsequent subordinate clause introduced by the conjunction *wie*. This configuration is syntactically and semantically parallel to the cataphoric reference of the *wenn*-clause in example (8). Christie's use in example (9) again marks her as an advanced user of German because she creatively varied a given model to produce an accurate structure that was not exemplified in the instructional materials.

Table 9 offers an aggregate comparison of all NS and learners' *da*-compound uses during the pre-intervention and entire post-intervention stage.

Table 9. Comparison of Pre- and Post-intervention NS and Learner *da*-Compound Uses

Category	All NSs	All Learners
Total German Words Pre-Intervention	19,736	15,161
Total *da*-compounds Pre-Intervention	106	26
Rate per 1000 German Words Pre-Intervention	5.57	1.71
Total German Words After Intervention I to Semester Close	12,822	7,363
Total *da*-compounds After Intervention I to Semester Close	98	49
Rate per 1000 German Words After Intervention I to Semester Close	7.64	6.65

On the whole, these data show that both NSs' and learners' rate of *da*-compound use increased from the pre- to post-intervention stage. One explanation for the increase in NS uses might be the changing nature of the task near the end of the telecollaborative partnership. In the final weeks, the participants were required to construct collaboratively a website that contained a jointly authored, bilingual essay. In the last weeks the learners often included excerpts of this essay in their emails for correction and discussion. We conjecture that one might find a higher rate of *da*-compounds in more formal written genres. Another reason why the rate of *da*-compound use increased is to be found in the intervention itself. Indeed, one might argue that learners were able to meet the demand for increased compound use in the written essay at the end of the semester precisely because they were exposed to the compound intervention prior to the completion of this task.

DISCUSSION AND CONCLUSIONS

Learner Corpus Research

Our learner corpus differs from the majority of learner corpora with respect to language and text type. The data consist of conversational discourse in both English and German in contrast to argumentative essays in English only (e.g., Granger, Dagneaux, & Meunier, 2002). Furthermore, our corpus does not require an external NS corpus for contrastive interlanguage analysis because it is an integrated bilingual corpus, which results in a built-in NS comparative baseline. This configuration is an advantage in contrastive interlanguage analysis because learner productions are not separated in time and space from the baseline. In addition, our corpus data are supplemented by ethnographic data, which afford a degree of "pragmatic convergence" (Tao & McCarthy, 2001, p. 658) between the analyst and the analyzed data that is not possible in the case of large NS corpora such as the *British National Corpus*. In fact, Flowerdew (2005) notes that "recent studies of small, specialised corpora bear witness to the increasingly important role that the ethnographic dimension now plays in corpus analysis" (p. 329).

The value of ethnographic data in the interpretation of quantitative corpus data is evident in the case of Saul, the only learner in our study to accurately name the modal particles in stage I. Our field notes indicate that Saul noticed the modal particle *ja* in his keypals' writing in the pre-intervention stage and discussed their uses with the instructor. Moreover, our biographical survey data reveal that Saul is a highly motivated Classics major who has previously learned Latin and Greek and is therefore well

acquainted with meta-linguistic terminology. These ethnographic data help explain his single MP use during the pre-intervention stage.[7] Finally, metapragmatic awareness data in the form of post-intervention portfolio entries allowed us to establish when and under what circumstances learners noticed (Schmidt, 2001) the focal features in the course of the interventions. As a result, we employed all "three sources of fact about language" (Cook, 1998, p. 59), observation, introspection, and elicitation.

Corpus-driven Language Pedagogy

The initial corpus-driven pedagogical interventions were designed in response to the needs of specific learners as established via "empirical pedagogical enquiry" (Widdowson, 1991, p. 23) in the form of contrastive interlanguage analysis of their own pre-intervention productions archived in our DLC. Because all learner and NS interactions were entered into the DLC on a daily basis as the telecollaborative partnership unfolded, each subsequent instruction module was based on ongoing assessment of the students' performance since the previous module. This type of dynamic assessment facilitated the production and modification of corpus-driven teaching materials according to individual developmental pathways at each particular point in time. For example, when we ascertained that learners overused MPs in formulaic patterns of appraisal, handouts containing concordance lines for each attested NS pattern were prepared and distributed to learners. Learners were thereby shown that most NS MP uses occurred in free constructions. Following these explanations, advanced learners like Christie experimented with more creative MP uses, whereas less advanced learners like Jeremy continued using more formulaic patterns but occasionally also tried out free constructions. Because the interventions under study were sensitive to particular learners' emerging performance as well as other ethnographic data, they differ drastically from corpus-driven instructional proposals based on isolated excerpts from external corpora.

The Authenticity and Authentication of Corpus Samples

Many researchers have responded favorably to the application of NS corpora in language pedagogy because corpora are seen as an unprecedented source of authentic language samples. Widdowson (1991), however, points out that corpus excerpts may be genuine in the sense that they are produced by NSs in real-life contexts, but language learners do not perceive them as authentic unless they are able to interpret and understand them in a personally meaningful way.

Our corpus-driven teaching materials side-step these issues, we argue, because they are comprised of NS-NNS interactions in which the learners themselves have participated. We maintain that such materials are particularly salient because previous research on telecollaboration has indicated that learners may 'try harder' to speak accurately and appropriately because they may wish to maintain positive face in front of NS age peers with whom they have built personal relationships (e.g., Belz & Kinginger, 2003). In the course of the pedagogical interventions, learners were involved in an iterative cycle of observation and subsequent use of the focal features in keypal correspondence. Thus, they did not run the risk of "produc[ing] corpus-attested but contextually inappropriate language" (Cook, 1998, p. 60). As a result, learners engaged with the corpus data via observation as well as participation (Gavioli & Aston, 2001); we argue that these dual roles enabled our participants to authenticate the corpus excerpts with which they worked during the form-focused pedagogical interventions.

Second Language Acquisition Research

The great majority of learner corpus research to date has employed a cross-sectional design wherein a synchronic slice of learner productions at a single moment in time is compared to a synchronic slice of NS productions in order to ascertain learners' mis-, over-, or under-use of particular L2 features (e.g., Granger, 1998; Granger, Hung, & Petch-Tyson, 2002). In contrast, this study addresses "the developmental course of instructed SLA" (Mellow, Reeder, & Forster, 1996, p. 326) because it is based on the full contingent of learner productions over a two-month period. Although 'longitudinal' analyses

have attracted more attention in recent publications (Ortega & Byrnes, 2008; Ortega & Iberri-Shea, 2005), they are still under-represented in SLA research in general and in corpus linguistics in particular.

A richly documented, ethnographically supplemented DLC facilitates the ecological tracking of micro-changes in learner performance over time. For example, our microgenetic analysis of Jeremy revealed that his first MP uses cannot be explained as imitations of his keypals because his productions preceded theirs. In this case, instruction appears to have triggered the noticing of features. This assumption is confirmed by numerous post-course portfolio entries (see also Vyatkina & Belz, 2006) such as the one by Christie who remarks that she began noticing MPs in NS writing "very soon" after the instruction. The pedagogical intervention thus functioned as an impetus for the learners to start working as explorers of the NS writing. We find much evidence of such "discovery learning" (Bernardini, 2002, p. 166) in post-course portfolio entries such as Christie and Jeremy's where learners included excerpts from their partners' and their own writing with MPs highlighted and accompanied by meta-commentaries. Moreover, our findings show that all learners developed in the general direction of the NS MP use, but never simply 'copied and pasted' NS phrases into their writing. Close comparisons of chat and email exchanges show that learner use always exhibits variations of patterns used by NSs (Vyatkina, 2007). This finding allows us to infer that learners "did not simply look for a model to imitate, […] but rather for patterns that [they] could adapt to [their] own text" (Gavioli & Aston, 2001, p. 242).

NOTES

1. The developmental learner corpus under study is *The Telecollaborative Learner Corpus of English and German (Telekorp*). The 1.5 million-token corpus consists of NS-NNS interactions between approximately 200 Americans and Germans collected during six different telecollaborative partnerships from 2000-2005. Thus, the corpus contains one cohort of learners for each of the six data collection cycles. The interactions comprise many different communicative situations, including requests, compliments, complaints, disagreement, flirting, comparison, explication, decision-making, negotiation of tasks, and information sharing.

2. Tognini-Bonelli (2001) contrasts *corpus-driven* and *corpus-based* approaches to the description of language use. Corpus-driven accounts derive from the empirical analysis of large amounts of corpus data, while corpus-based accounts rely on corpus data in order to test pre-formulated theories of language use.

3. Because there are no corresponding MPs in English, it is difficult to provide an accurate translation for them; nevertheless, the following glosses are representative of their expressive force in German. *ja* indicates: a) an intensification of positive appraisal; b) evidence of shared knowledge; c) presupposed agreement with the partner. *denn*: a) refers to the preceding context of the utterance or the situation; b) renders a question more natural and friendly; c) emphasizes the interest in the hearer's response. *doch* signifies: a) shared knowledge, a slight disagreement and a wish to overcome it; b) awareness of a possible objection on the part of the hearer. *mal*: a) lends an element of temporariness to an action; b) makes commands unobtrusive and friendly; c) lowers the commitment to fulfill a promise.

4. These responses are consistent with the responses given by the participants in Belz & Vyatkina (2005, p. 35). Lexical categories suggested by the learners in both studies included adverb, adjective, article, conjunction, and expletive.

5. English translations are presented here for space considerations.

6. Although the pre-intervention phase actually covered four weeks, learners were shown results for the first three weeks of their interaction only in part three of instruction module 1 because the researchers/instructors needed time to prepare the instructional materials.

7. Saul began his study of German in an intensive course in summer 2005, two months prior to the course under study.

APPENDIX: Summary of *da*-Compounds

Definition

da-compounds are pronouns like *personal* pronouns such as *er, sie, es, wir.* Pronouns are used to address people or to refer to nouns that have already been mentioned in a text. Pronouns substitute for nouns. *da*-compounds can refer to nouns but they also refer to larger pieces of a text such as clauses, sentences, or even bigger pieces such as scenes or ideas.

Direction of Reference

da-compounds can refer *back* to things that have already been mentioned or they can refer *forward* to things that have not yet been mentioned. When *da*-compounds refer back, they are called *anaphoric.* When *da*-compounds refer forward, they are called *cataphoric.*

Anaphoric Reference

da-compounds refer anaphorically to the following things:

1. nouns
2. objects of prepositions
3. larger pieces of text such as clauses, whole sentences, paragraphs

Cataphoric Reference

da-compounds refer cataphorically to the following things

1. *dass*-clauses
2. infinitive clauses
3. other subordinate clauses such as *wenn*-clauses

da-compounds can be used both anaphorically and cataphorically in conjunction with verb-and-preposition combinations such as *sich freuen auf, sich interessiern für.*

Corpus-Driven Analysis of Native Speaker vs. Learner Use of *da*-Compounds

1. NSs frequently use *da*-compounds. Learners rarely use *da*-compounds.
2. The most common *da*-compounds used by NSs are *dazu, davon, dafür, darauf, damit*. Learners tend to use *darüber* the most. They use *dafür* and *dazu* rarely.
3. NSs use *da*-compounds cataphorically about 25% of the time. Learners almost never use *da*-compounds cataphorically.
4. NSs use *da*-compounds to refer to bigger pieces of text more than 50% of the time. Learners most frequently use *da*-compounds to refer to small pieces of text (nouns) or in conjunction with verb-and-prepositions combinations.

How Can Learners Improve?

1. Use *da*-compounds more often. Use a greater variety of *da*-compounds.
2. Use the *da*-compounds to refer to larger pieces of text.
3. Use the *da*-compounds cataphorically.

ACKNOWLEDGMENTS

The authors wish to acknowledge in alphabetical order Nathaniel Carney, Carolin Fuchs, Lisa Hundley, Jonathon Reinhardt, Emily F. Rine, Nathan Shrefler, and Mickey Wheeler for data entry. Special recognition goes to Lisa Hundley for her assistance with an early draft of this article.

ABOUT THE AUTHORS

Julie A. Belz (Ph.D., *University of California at Berkeley*) is Associate Professor of English/Applied Linguistics and World Languages and Cultures at *Indiana University-Purdue University*, Director of the TESOL Graduate Certificate Program, and Director of Research in the Division for Health Communication at the *Indiana Center for Intercultural Communication.* She specializes in Internet-mediated language learning, intercultural communication, second language acquisition, pragmatics, learner corpus research, discourse analysis, learner identity, multilingual writing, and health discourse. She is the co-editor of *Internet-Mediated Intercultural Foreign Language Education* (Heinle & Heinle, 2006) and guest editor of a special issue of *Language Learning & Technology* on telecollaboration. Julie serves on the editorial boards of *Applied Linguistics* and *The American Association for University Supervisors and Coordinators of Foreign Language Programs* (AAUSC).

E-mail: jbelz@iupui.edu

Nina Vyatkina (Ph.D., *The Pennsylvania State University*) is Assistant Professor of German/Applied Linguistics, Director of the German Language Program, and Co-Director of the Center for German Business Language at the *University of Kansas*. Her research interests include longitudinal learner corpus analysis, computer-assisted language learning, and pedagogical applications of developmental L2 research.

E-mail: vyatkina@ku.edu

REFERENCES

Aston G. (2002). The learner as corpus designer. In B. Kettemann, & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 9–25). Amsterdam: Rodopi.

Belz, J. A. (2002). Social dimensions of telecollaborative foreign language study. *Language Learning & Technology, 6*(1), 60–81.

Belz, J. A. (2005). Corpus-driven characterizations of pronominal *da*-compound use by learners and native speakers of German. *Die Unterrichtspraxis/Teaching German, 38*(1), 43-59.

Belz, J. A. (2008). The role of computer mediation in the instruction and development of L2 pragmatic competence. *Annual Review of Applied Linguistics*, 27, 45-75.

Belz, J. A., & Kinginger, C. (2003). Discourse options and the development of pragmatic competence by classroom learners of German: The case of address forms. *Language Learning*, *53*(4), 591-647.

Belz, J. A., & Thorne, S. L. (Eds). (2006). *Computer-mediated intercultural foreign language education.* Boston, MA: Heinle & Heinle.

Belz, J. A., & Vyatkina, N. (2005). Learner corpus analysis and the development of L2 pragmatic competence in networked inter-cultural language study: The case of German modal particles. *Canadian Modern Language Review/Revue canadienne des langues vivantes, 62*(1), 17-48.

Bernardini, S. (2002). Exploring new directions for discovery learning. In B. Kettemann, & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 165-181). Amsterdam: Rodopi.

Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. Sinclair (Ed.), *How to use corpora in language teaching* (pp. 15-36). Amsterdam: John Benjamins.

Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.

Cook, G. (1998). The uses of reality: A reply to Ronald Carter. *ELT Journal*, 52, 57-63.

Coxhead, A. (2002). The academic wordlist: A corpus-based word list for academic purposes. In B. Kettemann, & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 73-89). Amsterdam: Rodopi.

De Cock, S. (2004). Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures, 2*, 225-246.

Ellis, R. (2006). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly*, *40*(1), 83-107.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, *35*(3), 407-432.

Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: Countering criticisms against corpus-based methodologies. *English for Specific Purposes, 24,* 321-332.

Gavioli, L., & Aston, G. (2001). Enriching reality: Language corpora in language pedagogy. *ELT Journal, 55*(3), 238-246.

Götze, L. (1993). Modalpartikeln aus der Sicht der Sprachlehr- und Sprachlernforschung. [Modal particles from the perspective of language teaching and learning research.] *Deutsch als Fremdsprache*, *4*, 227-232.

Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In S. Granger (Ed.), *Learner English on computer* (pp. 3-18). London: Longman.

Granger S., Dagneaux, E., & Meunier F. (2002) The International Corpus of Learner English. Handbook and CD-ROM. Louvain-la-Neuve: Presses Universitaires de Louvain (162 pp.).

Granger, S., Hung, J., & Petch-Tyson, S. (Eds.). (2002). *Computer learner corpora, second language acquisition and foreign language teaching.* Amsterdam: John Benjamins.

Granger S., & Tribble, C. (1998). Learner corpus data in the foreign language classroom: form-focused instruction and data-driven learning. In S. Granger (Ed.), *Learner English on Computer* (pp. 199-209). London: Longman.

Johns, T. (1986). Microconcord: A language learner's research tool. *System*, *14*(2), 151-162.

Kasper, G., & Rose, K. (2002). *Pragmatic development in a second language*. Malden, MA: Blackwell.

Lantolf, J. P., & Thorne, S. (2006). *Sociocultural theory and the genesis of second language development.* Oxford: Oxford University Press.

Layder, D. (1993). *New strategies in social research.* Cambridge, England: Polity Press.

Lee, D., & Swales, J. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-compiled corpora. *English for Specific Purposes, 25*(1), 56-75.

Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, A. M. McEnery, & G. Knowles (Eds.), *Teaching and language corpora* (pp. 1-23). London: Longman.

McCarthy, M., & Carter, R. (1995). Spoken grammar: What is it and how do we teach it? *ELT Journal*, *49*(3), 207-218.

Mellow, D., Reeder, K., & Forster, E. (1996). Using time-series research designs to investigate the effects of instruction on SLA. *Studies in Second Language Acquisition, 18*, 325-350.

Mindt, D. (2002). A corpus-based grammar for ELT. In B. Kettemann, & G. Marko (Eds.), *Teaching and learning by doing corpus analysis* (pp. 91-105). Amsterdam: Rodopi.

Möllering, M. (2001). Teaching German modal particles: A corpus-based approach. *Language Learning & Technology*, *5*(3), 130-151.

Möllering, M. (2004). *The acquisition of German modal particles.* Bern: Lang.

Möllering, M., & Nunan, D. (1995). Pragmatics in interlanguage: German modal particles. *Applied Language Learning*, *6*(1-2), 41-64.

Nesselhauf, N. (2004). Learner corpora and their potential for language teaching. In Sinclair, J. (Ed.), *How to use corpora in language teaching* (pp. 125-152). Amsterdam: John Benjamins.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins.

O'Dowd, R. (Ed.) (2007). *On-line intercultural exchange: A practical introduction for foreign language teachers.* Clevedon, UK: Multilingual Matters.

Ortega, L., & Byrnes, H. (Eds.). (2008). *The longitudinal study of advanced L2 capacities*. Mahwah, NJ: Lawrence Erlbaum.

Ortega, L., & Iberri-Shea, G. (2005). Longitudinal research in second language acquisition: Recent trends and future directions. *Annual Review of Applied Linguistics, 25,* 26-45.

Pravec, N. (2002). Survey of learner corpora. *ICAME Journal*, 26, 81-114.

Prodromou, L. (1996). Correspondence. *ELT Journal, 50*(4), 371-373.

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition,* 19, 223-47.

Römer, U. (2004). A corpus-driven approach to modal auxiliaries and their didactics. In Sinclair, J. (Ed.), *How to use corpora in language teaching* (pp. 185-199). Amsterdam: John Benjamins.

Rundell, M., & Granger, S. (2007). From corpora to confidence. *English Teaching Professional, 50*, 15-18.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Seidlhofer, B. (2002). Pedagogy and local learner corpora. In Granger, S., Hung, J., & Petch-Tyson, S. (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 213-234). Amsterdam: John Benjamins.

Seidlhofer, B. (2003). *Controversies in applied linguistics*. Oxford: Oxford University Press.

Sinclair, J. (2003). *Reading concordances: An introduction.* New York: Longman.

St. John, E. (2001). A case for using a parallel corpus and concordancer for beginners of a foreign language. *Language Learning & Technology, 5*(3), 185-203.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, *16*(3), 371-391.

Tao, H., & McCarthy, M. J. (2001). Understanding non-restrictive *which*-clauses in spoken English, which is not an easy thing. *Language Sciences*, *2*, 651-677.

Thurstun, J., & Candlin, C. N. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes,* 17, 267–280.

Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Amsterdam: John Benjamins.

Vyatkina, N. (2007). Development of second language pragmatic competence: the data-driven teaching of German modal particles based on a learner corpus. The Pennsylvania State University, Unpublished doctoral dissertation.

Vyatkina, N., & Belz, J. A. (2006). A learner corpus-driven intervention for the development of L2 pragmatic competence. In K. Bardovi-Harlig, C. Félix-Brasdefer, & A. S. Omar (Eds.), *Pragmatics and language learning* (pp. 315-357). University of Hawai'i at Manoa, National Foreign Language Resource Center.

Vyatkina, N., & Johnson, K. E. (2007). *Teaching German modal particles: A corpus-based approach*. University Park, PA: CALPER Publications. Retrieved September 25, 2008, from http://calper.la.psu.edu/publications.php

Weydt, H. (1969). *Abtönungspartikel*. Bad Homburg v.d.H., Berlin, Zürich: Verlag Gehlen.

Weydt, H., Harden, T., Hentschel, H., & Rösler, D. (1983). *Kleine deutsche Partikellehre.* Ernst Klett: Stuttgart.

Widdowson, H. G. (1991). *Aspects of language teaching*. Oxford: Oxford University Press.

Widdowson, H. G. (2000). On the limitations of linguistics applied. *Applied Linguistics*, *21*(1), 3-25.