Straight on through to Universal Grammar: Spatial modifiers in second language acquisition

David Stringer, Beatrix Burghardt, Hyun Kyoung Seo, Yi-Ting Wang
Indiana University

(DRAFT, MAY 2009)

Abstract

There has been considerable progress in L2 research at the syntax-semantics interface addressing how syntax can inform semantics, in terms of interpretive correlates of word order (Slabakova, 2008). This paper provides evidence of the inverse of this phenomenon, with semantics reliably informing learners about syntax. It is argued that there is a functional hierarchy of modifiers in the domain of adpositions which enables the linguistic elaboration of trajectories, but that not all languages lexicalize all types. This study examines whether L2 learners of English are able to overcome the poverty of the stimulus and recruit the relevant functional categories despite their absence in the L1. Modifiers were taught to learners individually, but never in combination. A computer-animated narrative was designed in order to create felicitous contexts for combinations of modifiers, and preference and grammaticality judgment tasks were administered to 131 students from various L1 backgrounds, as well as 20 native speakers. Accuracy scores were remarkably targetlike on binary combinations of modifiers (i) across proficiency levels, (ii) across L1s, and (iii) across the two tasks, revealing that the flow of information is also from semantics to syntax: with the semantics of modifiers in place, the syntactic hierarchy is naturally manifested.

I Introduction

Recent work on the syntax–semantics interface in second language acquisition has made considerable progress in identifying the influence of syntax on semantics, in terms of semantic correlates of word order (for an overview, see Slabakova 2008). This study addresses the question of whether the flow of information is also in the opposite direction, such that semantics can inform syntax. Just as there are robust universal orderings of adjectival modifiers in the nominal domain (Cinque, 1994; Crisma, 1996; Shlonsky, 2004), and of adverbial modifiers in the verbal domain (Alexiadou, 1997; Cinque, 1999), it is here argued that modifiers of the category P (pre/postpositions and particles) appear to stack in a fixed order to the left of the head. However, not all languages lexicalize all types of modifier, which raises the question of whether L2 learners, on gaining awareness of the semantics of modifiers, may reveal knowledge of the syntax of elements in the hierarchy that are not present in the L1, despite lack of instruction and a paucity of evidence in natural input. We examine whether L2 learners of English are able to overcome the poverty of the stimulus and employ functional categories and features not instantiated in the L1 (contra Hawkins and Chan 1991 but consistent with Schwartz and Sprouse 1996). In Section 2, we present an account of the syntax of P-modifiers in

English, in relation to their manifestation in other languages. In Section 3, a description is given of two experiments with novel methodology involving computer animation incorporated into slideshows. The results are summarized in Section 4: in general, accuracy rates were resolutely above chance and at extremely similar levels across the proficiency range, across L1s, and across the two tasks. In Section 5, the implications of these findings are drawn out. Knowledge of the hierarchy of spatial modifiers is in evidence at all stages of development, revealing that lexical semantics can trigger functional projections in L2 syntax. This is commensurate with a view of the language faculty with unidirectional information flow between modules, from lexical semantics to syntax at the lexical interface prior to syntactic derivation, and from syntax to phrasal semantics at the interpretive interface following syntactic derivation. The fact that learners show knowledge of semantics-syntax mappings from the outset, despite the absence of particular mappings in the L1, provides further evidence of the role of Universal Grammar (UG) in second language acquisition.

II The syntax of spatial modifiers in L1 and L2

An emerging consensus in research on the syntax of adpositions is that there is a universal layered structure inside PP, with a higher directional P (PathP), a lower locational P (PlaceP), a locative nominal projection (LocN), and a semantically vacuous PP that assigns case to DP (e.g. Stringer, 2007; den Dikken, 2006; Koopman, 2000; van Riemsdijk, 1990; Svenonius, 2008), as exemplified below in English.

(1) $[_{\text{PathP}}$ from $[_{\text{PlaceP}}$ on $[_{\text{LocN}}$ top $[_{\text{P}}$ of $[_{\text{DP}}$ the table]]]]]

This hierarchy is attested in many languages despite great variability in manifestations of the category P, which may be expressed by means of prepositions, postpositions, particles, or affixes of various types. For example, German circumpositions are best captured by means of the layered PP structure, as shown by van Riemsdijk (1990), and nominal case suffixes in languages such as Lezgian and Inuit are found in precisely the reverse order of prepositions, in line with the Mirror Principle (Baker, 1985), as shown by van Riemsdijk and Huybregts (2007). Languages with less regular paradigms, such as Northern Sami, and languages with mixed systems of adpositions and case suffixes, such as Hungarian, also appear to respect the syntactic hierarchy (Stringer, 2008).

In an extension of this work, we examine how the expression of trajectories may be elaborated by means of modifying elements within spatial PP. Following observations made by Stringer (2005), at least three types of P-modifier may co-occur in a fixed structural hierarchy, as exemplified in (2). That these are indeed P-modifiers with a fixed word order to the left of the head, and not verb particles or 'satellites', is shown by tests of displacement (3a-d).

(2) The fish swam $[_{\text{DEG}}$ {right/straight} $[_{\text{FLOW}}$ {on/back} $[_{\text{TRAJECT}}$ {through/down} $[_{\text{PP}}$ into the cave]]]].

(3) a. The fish swam {*straight through on/*on straight through/*on through straight, etc.} into the cave.
b. It was [straight on through into the cave] that she swam.
c. *It was [through into the cave] that she swam straight on.
d. *It was [into the cave] that she swam straight on through.

The highest modifiers are those of Degree (or intensity), which are well-recognized, and standardly used as a test of prepositional status (Emonds, 1985; Jespersen, 1992 [1924]). Degree modifiers are hard to define precisely, but usually have a sense of 'directly', 'exactly', or 'completely'. *Right* may be used with either directional or locational P, whilst *straight* may only be used with directional P, as shown below.

(4) a. The bird flew {right/straight} into the hole.
b. The bird lived {right/*straight} in the hole.

The motivation for an independent class of Flow modifiers springs from the fact that they have their own particular position, always following Degree modifiers, and always preceding Trajectory modifiers. *On* expresses the continuation of the directional flow, and *back* expresses the reversal of the directional flow. The third class consists of elements normally appearing as lexical P, but functioning in this case as P-modifiers, thus elaborating on simple trajectories. This class of elements includes *up*, *down*, *through*, *over*, and *across*. Their status as P-modifiers can be distinguished from their status as prepositions by means of tests of syntactic distribution as in (3), and by *right*-modification (Ayano, 2001: 79, fn.1).

It must be stressed that the observations offered here concerning P-modifiers only apply to these lexical items on the relevant interpretations. *Straight* is often ambiguous between a Degree modifier reading and a directional adverb reading (in contrast with *diagonally*, *round*, etc.). The following example could mean that Pat ran right home, without getting sidetracked or taking any detours (P-modifier), or that he ran in a straight line (adverbial).

(5) Pat ran straight to the Post Office.

Similarly, *back* can sometimes mean the opposite of *front*, in which case it cannot be construed as a reverse Flow modifier, as shown below.

(6) Harry moved back, because he hated sitting at the front.

In the context of discussion of directional adverbs, Svenonius (2008) considers the co-occurrence of *straight* with a fourth type of modifying element, namely Measure Phrases, observing that the order of measure phrases and directional adverbs is rigid, as in *{twenty centimeters straight/*straight twenty centimeters} below the window*. We find Svenonius' observation to be accurate on the intended interpretation (i.e. the directional adverbial reading: *straight* as against *diagonally*), but Measure Phrases appear to be in strict complementary distribution with *straight* as a Degree modifier, and indeed with any of the P-modifiers in the hierarchy indentified above, and will not be considered here.

The modificational system outlined above can be represented as follows.

(7) [(DEGREE) [(FLOW) [(TRAJECTORY) [P]]]]

Prosody plays a pivotal role in the parsing of phrases with multiple modifiers. Just as in the adverbial hierarchy (Cinque, 1999), the insertion of pauses, shifting of stress or other variance in the intonational contour results in the assignment of different interpretation and different syntactic structure. Consider the sentence in (8a), with variants on the postverbal elements given in (8b-d)

(8) a. Although she was tired after the great migration, the bird flew [right on down] to the lake.
b. {*[right down on] to the lake / right down, onto the lake}
c. {*[down right on] to the lake / down, right onto the lake}

In a context of 'continuation' that renders the Flow modifier meaningful, and with the P-modifiers in a single intonational phrase, the meaning of (8a) is clear, and the word order corresponding to this interpretation is fixed. However, if a pause is inserted after *right down* and *on* forms a prosodic unit with the prepositional head *to*, as in (8b), the syntax changes such that *right* modifies *down*, which is in this case not a modifier but an intransitive head P (following Emonds, 1985), and *on* forms a complex P *onto*. In this case, the meaning changes such that the bird lands on the surface of the lake, which is not entailed by (8a). In the variation in (8c), the degree modifier is not applied to the intransitive P *down* but to the transitive P *onto*, such that emphasis is given not to the swooping downward but to the precision of the landing. Clearly, if such subtleties of interpretation are to be investigated in L2 acquisition, then not only must there be manipulation of syntax, but assiduous control of both context and prosody.

Not all languages lexicalize all types of P-modifier, but when two or more are found, they conform to syntactic predictions. In a pattern suggestive of an implicational hierarchy[1], a language may lexicalize all three (e.g. German, English), the higher two (Estonian, Hungarian), only the highest (French, Spanish), or none at all (Japanese, Korean), as shown below.

(9)	direkt zurück hoch auf den Berg straight back up on the Mountain 'straight back up on the mountain'	DEG-FLOW-TRAJ (German) (**zurück hoch direkt*, etc.)
(10)	otse tagassi teatri-sse right back theatrer-into' 'right back into the theater'	DEG-FLOW (Estonian) (**tagassi otse*)

[1] Thanks to Kamil Ud Deen for this observation.

(11) juste par en dessous le pont DEG (French)
right via LocP underneath the bridge
'right under the bridge'

(12) teburu no (ma-) ue (*ma-) de Ø (Japanese)
table GEN (right-) top (*right-) LocP
'right on top of the table'
(literally: 'on right top of the table')

In languages such as Japanese and Korean, the inventory of adpositions is typically very small, as trajectories are typically expressed by means of verbs or complex PPs with spatial nouns. Such intensity modifiers as exist (e.g. Japanese *ma(n)* 'right', *sugu* 'immediately', *chōdo* 'exactly') attach not to P but to spatial nouns.

This fixed ordering of P-modifiers resembles the syntax of modification by means of adjectives (Cinque, 1994; Crisma, 1996; Shlonsky, 2004) and adverbs (Alexiadou, 1997; Cinque, 1999), in that there is a universal basic order, which I assume with Cinque (1999) to be the product of a functional hierarchy above the lexical head. Such hierarchies have not been the subject of second language research programs to date, although their implications have sometimes been drawn out with respect to other phenomena. In the adjectival domain, for example, Androutsopoulou, Español-Echevarría, and Prévost (2008) examined noun-adjective placement in L2 Spanish by L1 French learners with specific reference to the adjectival hierarchy, although the hierarchy itself was not the target of the investigation. In the adverbial domain, Ionin and Wexler (2001) found that thematic verbs in the L2-English of L1-Russian speakers were more likely to raise past Manner adverbs than Frequency or Epistemic adverbs, a fact which they attribute to the relative position of the adverbs on the functional hierarchy, but the focus of the investigation was verb-raising rather than how the hierarchy itself was revealed in the course of L2 acquisition.

Predictions of patterns of L2 development in this domain vary with theoretical framework. On the assumption that modifiers are hosted in distinct functional projections (Cinque, 1999), the Failed Functional Features hypothesis (Hawkins & Chan, 1997) predicts interminable confusion with regard to the hierarchy for types of modifier that are not instantiated in the L1. In contrast, the Full Transfer/Full Access hypothesis (Schwartz & Sprouse, 1996) predicts an initial facilitation for types manifested in the L1, but rapid convergence on the target grammar once the semantics of individual modifiers has been identified. If knowledge of UG is available in L2 acquisition, the syntax of modification should reveal itself despite the deficiencies in instruction and input, and despite the absence of particular types of modifier in the L1.

III The Aladdin experiments

1 Hypotheses

The linguistics literature to date concerning the syntax of adpositions makes no mention of the syntax of spatial modifiers, and nothing is known of patterns of acquisition in either L1 or L2 acquisition; thus to what degree P-modifiers are learnable in a second language and what role the modificational system of the L1 might play were unknown factors. The general research question we attempted to address was whether learners of English show knowledge of the universal hierarchy of P-modifiers over the course of L2 development. Three contrasting hypotheses were considered.

- H1: The hierarchy will be in evidence from the outset.
- H2: The hierarchy will emerge gradually as learners establish appropriate lexical representations and prosodic constraints.
- H3: Learners who lack L1 analogues will not evince knowledge of the hierarchy.

2 Participants and location

The participants were drawn from six proficiency levels of an Intensive English Program at a large university in the Midwest of the United States. The proficiency levels were derived independently of this project by the battery of placement exams used by the program. Initial placement criteria included composition, reading, vocabulary, grammar, listening comprehension and oral interviews, and promotion in the course involved integrating subsequent sets of test performance scores with previous course grades and current TOEFL scores. A total of 131 students from 6 different levels of proficiency took part in the study, with number of students by level as follows: L2: 4, L3: 15, L4: 29, L5: 45, L6: 32, L7: 6. For purposes of analysis, these were conflated into three general proficiency groups: Lower-intermediate (L2-4: 48), Intermediate (L5: 45) and Advanced (L6-7: 38). Learners came from 17 different L1 backgrounds, listed as follows by numbers of native speakers in the study: Arabic (16), Bambara (2), Chinese (14), French (2), Hungarian (2), Japanese (12), Korean (40), Mongolian (1) , Portuguese (4), Russian (2), Spanish (5), Tajik (2), Thai (1), Tamil (1), Tartar (1), Turkish (27), Vietnamese (2). Three participants were bilingual or multilingual from childhood (1 French/Bambara, 1 French/Bambara/ Russian, 1 Tartar/Russian), and two languages were spoken only by bilinguals (Bambara and Tartar). The main experiments were conducted in a language lab with learners seated in individual booths, but using a main screen and surround speakers. Despite the range of affective factors and processing strategies among students, it was thus possible to synchronize aural and visual stimuli for all participants. The control experiments were conducted with 20 native speakers of English, aged 19-48, all of whom were Americans who had spent most of their lives in the Midwest of the United States, and none of whom were linguists.

3 Materials

a Contextual materials

An original narrative was designed to contextualize PPs and their modifiers. The story was intended, as much as possible, to be interpretable across cultures, and involved characters and scenes which were variations on the well-known Middle-Eastern folk-tale of Aladdin.[2] The narrative runs as follows. In a cave filled with treasure, Aladdin takes a magic lamp from under the nose of a wizard. He then jumps onto a magic carpet and flies up to an opening at the top of a flight of stairs at the side of the cave (*He flies right up out of the cave*). He passes through a tunnel to the outside (*He flies on through to the outside*), where he continues to fly through various spatial environments, each of which provides a plausible context for a targeted combination of prepositions and modifiers. In our version of the tale, he flies everywhere: reducing variability in manner of motion allowed for greater focus on trajectories. For example, in the course of his journey, he flies over some camels, up into the clouds, down to a lake, behind a waterfall, under a rock bridge, across a desert, through a city gate, etc. At one point he drops the lamp, but does not realize until later, and therefore has to fly back through the same environments, enabling further test materials to be utilized. The wizard eventually finds the lamp, but again Aladdin snatches it from under his nose. In the renewed pursuit, the wizard falls into a lake, and Aladdin finally manages to make it back into the desert, where he releases the genie.

The narrative was presented visually. An advantage of using visual stimuli together with linguistic encodings of motion events is that they can force an intended interpretation of a given sentence, despite the polysemy which is rife in adpositional systems. Every P-modifier in the hierarchy that we examine creates potential for ambiguity. That said, whether *right* and *straight* are interpreted as Degree modifiers or as a directional adverbials, whether *on* is interpreted as a Flow modifier or a preposition, and whether *back* signifies return or movement to a posterior location may be largely controlled with appropriate visual stimuli. To embed the visual stimuli within a narrative was necessary in order to provide appropriate context for Flow modifiers, which necessarily express continuation or return with specific reference to prior events.

Previous experimental work on motion events has tended to rely either on two-dimensional images (Berman & Slobin, 1994; Strömqvist & Verhoeven, 2004) or on video (Pourcel, 2002; Hohenstein, Naigles, & Eisenberg, 2004) for elicitation of utterances or judgments. For the current project, novel experimental stimuli were developed, by making use of basic computer animation. The advantages of animation include the incorporation of actual rather than inferred motion in the stimuli (as with video), and the expression of a full range of motion events (as with pictures), without placing actual actors in peril as they fall into lakes, tumble down mountains or fly through tunnels. The animation was achieved in the following way. First, the various characters, objects and background scenes were hand-drawn and colored. Second, the cut-outs and background scenes were scanned as digital images. Third, they were incorporated into Microsoft Powerpoint slides, arranged in layers depending on the desired visibility of objects, and animated to create motion events. So, for example, when

[2] Initial concerns about how learners from different L1 backgrounds might perceive the 'exoticism' of the setting were laid to rest by informal discussion with informants from a variety of L1 backgrounds, and the discovery that such exoticism exists in all of the various versions found internationally. In the oldest Arabic and Syrian folk-tale versions, as well as in early translations into English and French, Aladdin himself is from China (the farthest East imaginable), whilst the sorcerer comes from Morocco (the farthest West imaginable).

the protagonist flies through a gate in a city wall, he disappears as he passes behind the gate, as the latter is the top layer among the objects in the slide.

b Linguistic materials

As discussed previously, prosodic cues are also indispensible to establish the intended interpretation of a string of modifiers. Following a small pilot study, two problems in particular were identified in the formulation of stimuli for preference and grammaticality judgment (GJ) tasks. The first was that the interactions between prosody and parsing were much more complex than we had at first assumed. The main issue was our attempt to ensure that the lexical elements of the modificational hierarchy were in fact parsed as part of a prosodic unit with P, and interpreted as P-modifiers, rather being prosodically aligned with V, and interpreted as verb particles. This might have been simple if there were only two contrastive patterns of prosody. However, there are many ways to form a prosodic unit, which is itself somewhat underdefined in the phonology literature. There may be a pause between prosodic units; there may be a change in pitch such that the first sequence is high and the second low, or vice versa; or both can have internal pitch and stress variation, the shift to the extended PP being marked by sudden rise and gradual fall. The problem was that we could not systematically predict which prosodic pattern would be the most unambiguous delivery of the materials. Sometime a pause was felicitous; sometimes not. Within a given combination type (i.e. Degree-Flow, Flow-Trajectory, etc.), the felicity of particular prosodic patterns varied from item to item, apparently depending at least in part on the choice of lexical items. The second problem was that the response sheets initially had written versions of the carrier sentences, which allowed participants to read as well as listen, either before during or after the oral stimulus, so that the individual's prosodic understanding of the stimulus was an uncontrolled variable. These problems were tackled in the main experimentation by (i) selecting the most appropriate prosody for stimuli on an item-by-item basis, and (ii) removing all written cues from the response sheets. In this way, once all stimuli had been digitally recorded and incorporated into the animated Powerpoint slides as sound files, we were able to ensure the most felicitous prosodic phrasing for the stimuli based on native speaker intuitions, secure the uniform delivery of stimuli in both experiments across many sessions, and restrict linguistic stimuli to oral delivery.

There were 26 slides in total: 3 initial example slides, 2 fillers for narrative coherence, 3 slides targeting onomatopoeia (outside the scope of the current discussion), and 18 test slides targeting the hierarchy of spatial modifiers. The stimuli were formulated as shown in Table 1. The particular combinations used in Experiments 1 and 2 were the same, to facilitate triangulation of the results. For Experiment 2, when categories allowed 6 tokens, 3 were grammatical and 3 ungrammatical, and when categories allowed only 3 tokens, 1 was grammatical and two were ungrammatical. The stimuli were not as balanced as they could have been if created outside the context of a narrative; however, it was decided that the felicitous embedding of multiple modifiers in a coherent narrative was of greater importance, so we strived to balance stimuli within narrative constraints.

Table 1. Targeted combinations in Experiments 1 and 2.

(A) DEG-FLOW (x6)	(a1) straight on, (a2) straight on, (a3) right on, (a4) straight back, (a5) right back, (a6) right back
(B) DEG-TRAJECT (x6)	(b1) right up, (b2) right down, (b3) straight through, (b4) straight down, (b5) right out, (b6) straight out
(C) FLOW-TRAJECT (x3)	(c1) on through, (c2) on down, (c3) back over
(D) DEG-FLOW-TRAJECT (x3)	(d1) right on up, (d2) right back down, (d3) straight back across

In Experiment 1 (the preference task), following oral delivery of two variants of a sentence, learners circled (a) or (b) on their answer sheets, according to which sounded better. In Experiment 2 (the GJ task), positive or negative responses were elicited following pre-recorded delivery of sentence fragments. The order of presentation of targetlike and nontargetlike variants was systematically varied across stimuli in both experiments. Samples of both types of stimuli are given below.

(13) Experiment 1 Sample:
‘He flies ______________ over the camels.’ (**on straight/straight on*)
A B

(14) Experiment 2 Sample:
‘Now where does he go?’ ‘Straight back across the desert.’ (*ok*)
A: good B: bad

The rationale behind the use of sentence fragments was to further control for prosodic reanalysis by subjects. For example, in the context of Aladdin continuing on through a tunnel to the outside, prosody can disambiguate between [*he flies through*] [*on to the outside*], which is grammatical, and *[*he flies* [*through on to the outside*]], which is not. However, prosody can be notoriously difficult in second language acquisition, and an ungrammatical sentence fragment answer such as *[*through on to the outside*], with appropriate prosody, reduces the chance of the P-modifier being reanalyzed as a verbal particle. The rationale behind the forced GJs was to avoid a preponderance of ‘don’t know’ answers given the processing difficulties inherent in many test items and the indeterminate nature of this type of second language judgment. If a participant felt 75% sure of the grammaticality of the sentence, we wanted to ensure that a positive value was recorded, whilst if the opposite were true, we wanted a negative value for the response. We also wanted the initial reactions of students to the stimuli, without the reflective time spent on more elaborate systems of judgment such as Likert scales, which increase the likelihood of prosodic rephrasing during sentence recall. Both learners and controls were told to listen to how the sentence sounded, and to judge it on the way it sounded, without changing the way it was phrased.

In addition to the actual test materials, a set of teaching materials was created, to be used immediately preceding the experiment proper, with the purpose of making clear the meaning of each of the P-modifiers on the intended interpretations in English. Acquisition of the lexical items themselves was not the subject of investigation, but rather their interaction with one another, so pains were taken to ensure that individual lexical meanings were understood and accessible. A handout was also created, which was left on learners' desks throughout the experiments, so that they could quickly recall by means of written and visual aids the meanings of individual items. The items on which they received instruction were the Degree modifiers *right* and *straight*; the Flow modifiers *on* and *back*; the Trajectory modifiers *up*, *down*, *through*, *over*, *across* in prepositional contexts, and the locative nouns *front* and *top*. As with the test materials, the instructional materials were presented in the form of animated Powerpoint slides. The most important aspect of the logic of this part of the experimentation was that students were taught modifiers *in isolation* (i.e. 1 modifier +PP), but they were tested on modifiers *in combination* (i.e. 2 or 3 modifiers + PP).

4 Protocol

The experiments were conducted during a 50-minute period, which began with the teaching session discussed above, followed by the preference judgment task, followed immediately by the GJ task. This order of presentation was maintained because it was considered more likely for the GJ task to affect subsequent preference judgments than the other way round. However, given the plethora of modifiers and combinations in both batteries of stimuli it is unlikely that participants were able to summon up previous judgments during the course of the second task in any case. The response sheets were then collected, and the session was brought to a close with a further instructional session, during which the general principles of the grammar of P-modifiers were explained, so that participants (as students of English) could gain something from the experience.

During the session, one experimenter was responsible for oral delivery of all instructions, and another for the manipulation of visual images and pre-recorded test stimuli. The third and fourth experimenters administered all documents to participants, and performed the crucial function of monitoring the behavior of participants during the session. Notes were taken if participants listened but did not look at the screen, or changed any of their initial answers, or otherwise ignored instructions. Across both experiments, the animations were not run unless all participants were looking at the screen, given the importance of the visual stimuli to interpretation. Thus once a new slide was presented, we checked that all participants had their gaze directed toward the screen, and only following an agreed signal did the experimenter in control of the visual materials start the animation, which was then in sync with the oral stimulus. All linguistic stimuli were repeated once after a four-second pause.

IV Results

1 Results by proficiency level

After applying the same criteria to all subjects (persistent inattention during the experiment, inaccurate responses on two or more of the four fillers, or obviously artificial response patterns), ten subjects were eliminated from Experiment 1 and thirteen from Experiment 2, leaving the numbers of subjects as follows for Experiment 1: Group 1: 42; Group 2: 41; Group 3: 38; and as follows for Experiment 2: Group 1: 41; Group 2: 40; Group 3: 37.

The results in terms of the three proficiency groupings are given in Tables 2 and 3. It had been expected that accuracy on the hierarchy would gradually increase as the lexicon was consolidated and principles of prosody were acquired, in accordance with Hypothesis 2; however, rather than show a strong developmental pattern, the results were remarkably similar across proficiency levels. Twelve 2-tailed comparisons of proportions were conducted for each experiment to calculate differences in accuracy scores on the four types of combination by the three proficiency groups. In Experiment 1, there were no significant differences between Groups 1 and 2, only one between Groups 2 and 3 (Type A: 74-84%, p=0.007), and three between Groups 1 and 3 (Type A: 76-84%, p=0.028; B: 71-81%, p=0.01; D: 64-76%, p=0.042). Overall, these results show a slight (but significant) increase in accuracy on the preference task between Group 1 and Group 3. In Experiment 2, of the twelve comparisons only two were significant: one between Groups 1 and 2 (Type B: 69-77%, p=0.047), and one between Groups 1 and 3 (Type B: 69-79%, p=0.013). Overall, performance on binary combinations of modifiers was extremely similar regardless of proficiency level.

A glance at the descriptive statistics immediately reveals a difference between the relatively accurate performance on Types A and B (Deg-Flow and Deg-Traject), the particularly nontargetlike performance on Type C (Flow-Traject), and the marginally accurate performance on Type D (Deg-Flow-Traject). Such differences require further scrutiny.

Table 2. Exp 1: Preference task accuracy scores by proficiency level, with significance above chance (***p<0.001, **p<0.01, *p<0.05).

	Group means: % accuracy (significance above chance)			
	Group1 (N=42)	Group2 (N=41)	Group3 (N=38)	Controls (N=20
(A) DEG-FLOW	76%***	74%***	84%***	95%***
(B)DEG-TRAJECT	71%***	78%***	81%***	98%***
(C) FLOW-TRAJECT	41%	38%	44%	83%***
(D) DEG-FLOW-TRAJECT	64%**	68%***	76%***	100%***

Table 3. Exp 2: GJ task accuracy scores by proficiency level, with significance above chance (***p<0.001, **p<0.01, *p<0.05).

	Group means: % accuracy (significance above chance)			
	Group1 (N=41)	Group2 (N=40)	Group3 (N=37)	Controls (N=20)
(A) DEG-FLOW	76%***	81%***	80%***	93%***
(B) DEG-TRAJECT	69%***	77%***	79%***	90%***
(C) FLOW-TRAJECT	34%	39%	32%	82%***
(D) DEG-FLOW-TRAJECT	58%	63%**	68%***	98%***

First, let us examine Type C (Flow-Traject) in more detail. Although at first pass it might be thought that the lower reaches of the hierarchy might pose a higher degree of difficulty, analysis by individual stimuli reveals that poor performance on Flow-Traject might be alternatively explained in terms of a lexical effect. In Experiment 1, participants treated items (c1) *on through* and (c2) *on down* very differently from (c3) *back over*. Accuracy rates by group for (c1) were 43%, 20% and 42%; for (c2) they were 12%, 15% and 8%; whilst for (c3) they were 67% (p=0.023), 80% (p<0.001) and 82% (p<0.001). One possible reason for this discrepancy might be that the PPs modified by these combinations were headed by *to*: <u>*on through*</u> *[to the outside]*; <u>*on down*</u> *[to the ground]*; <u>*back over*</u> *[to the waterfall]*. If participants rephrased the first two utterances prosodically as they considered their responses, the resultant forms could be interpretable with *through* or *down* either as verb particles or as P-modifiers, with *on* analyzed not as a modifier at all but as part of the complex preposition *onto*. Such an analysis might be more convincing if the results of GJ task revealed a similar discrepancy. However, in Experiment 2 the accuracy rates by stimulus were as follows: (c1) 32%, 25%, 16%; (c2) 54% 62% 63%; (c3) 15% 28% 19%, with poor performance on all stimuli. It is notable that the controls subjects also had difficulty with (c1) in particular, with scores of 50% in Experiment 1 and 65% in Experiment 2 bringing down the average accuracy for this type. Given the design flaw in juxtaposing *on* and *to* (albeit a legitimate combination in the target language), we are forced to leave more detailed examination of L2 knowledge of Flow-Traject for future work, in which the *to*-PP might be replaced with, for example, an *into*-PP.

Performance on the ternary combinations of Type D was significantly above chance, though at lower levels than the binary combinations of Types A and B. This was partly to be expected given the increase in processing load, but these examples were included to stretch learners as native responses were so robust (indeed, in the actual experiment, the controls attained 100% and 98% accuracy for this type in Experiments 1 and 2, respectively). While the group scores attained significance, again there were item effects. In Experiments 1 and 2, scores for (d1) *right on up* neither reached significance nor showed any improvement across the proficiency levels, standing at 64%, 63% and 63% in Experiment 1, and 49%, 43% and 41% in Experiment 2. In contrast, the other two types showed marked improved across the proficiency range: accuracy scores for (d2) *right back down* were 56% (n.s.), 68% (p=0.019) and 73% (p=0.003) in Experiment 1, and those for (d3) *straight back across* were 68% (p=0.017), 78% (p<0.001) and 89% (p<0.001).

Given the findings discussed above regarding *on* vs. *back*, performance on individual stimuli was checked for item effects. In this regard, differences do emerge on close examination of the data from the lower-intermediate group. A comparison between accuracy rates of Group 1 and Group 3 responses with *right* reveal approximately the same scores; similar comparisons of responses with *straight* usually, although not always, reveal jumps in accuracy rates between Groups 1 and 3. Such differences on the preference task are as follows: (a1) 62% to 79%; (a2) 62% to 79%; (a4) 79% to 92%; (b3) 83% to 92%; (b4) 46% to 54%; (b6) 59% to 81%. It is plausible that the relevant meaning of *straight* is acquired later than *right* in general, and some individuals may have persistent misunderstanding of this item, although Groups 2 and 3 seem to have acquired the relevant meaning and syntax.

One might ask why, although learners performed so well above chance for Types A and B, they did not perform as well as native speakers. We believe the reason for this can be found in the examination of individual results. This is most clearly seen with the control subjects, who all remained for some time after the experiment to give feedback on the experience. While the majority either performed exactly as expected or gave at most one or two unexpected answers out of forty-six, some produced negative results with a degree of consistency. Notably, one subject (NS2), produced 5 out of 6 nonstandard preference judgments for Type A, and in Experiment 2 accepted all variants of Types A and B, resulting in a total of six accurate responses to the grammatical variants and 6 inaccurate responses to the ungrammatical variants. Fortunately, following the experiment (and prior to these results coming to light), she provided an explanation of this by saying, ‘Most of the sentences were fine if you just changed the pronunciation a bit’. Similarly two more subjects (NS12 and NS14) said that once or twice they thought it was ‘OK if you just said it a different way’. These responses show clearly that at least some native speakers disregarded explicit instructions not to change the way sentences sounded before judging them. Perhaps this was to be expected given the fact that nonlinguists are naturally bemused by ungrammatical sentences and that we as humans have a tendency to repair deviant utterances. However, if some native speakers performed in this way, it seems very likely that non-native speakers found it difficult to adhere to the instructions, and prosodically rephrased some of the stimuli, despite our efforts to prevent this by means of restricted visual context, pre-recorded stimuli and explicit instructions. Although this may explain certain responses, the behavior of some individuals is likely to remain mysterious in a cross-sectional study of this type. While one Turkish learner had accuracy scores of 24/24 (100%) on Types A and B over both Experiments, another Turkish learner in the same level scored 8/24 (33%). Numbers of subjects were sufficient to overcome the behavior of outliers, so all subjects were included who met the criteria for inclusion outlined at the beginning of this section.

To summarize the results by proficiency level: The learners were significantly outperformed by the controls in all cases, but they nevertheless showed rates of accuracy that were well above chance for the binary combinations of Types A and B, consistently underperformed on Type C (which contained a design flaw), and showed improvement and eventual accuracy on the ternary combinations of modifiers of Type D.

2 Results by L1

An analysis was also made of the performance of learners from different L1 backgrounds. Following the exclusion of certain subjects as detailed previously, the L1s with the most speakers were, in numerical order: Korean (36), Turkish (25), Arabic (15 and 13, in Experiments 1 and 2, respectively), Chinese (14), and Japanese (12). None of these languages has more than one level of the hierarchy instantiated, so these learners must project functional categories that are absent in the L1. The results by L1 background are given in tables 4 and 5.

Table 4. Exp 1: Preference task accuracy scores by L1, with significance above chance (***p<0.001, **p<0.01, *p<0.05).

	Group means: % accuracy (significance above chance)				
	Korean (N=36)	Turkish (N=25)	Arabic (N=15)	Chinese (N=14)	Japanese (N=10)
(A) DEG-FLOW	78%***	73%***	79%***	77%***	80%***
(B) DEG-TRAJECT	78%***	80%***	82%***	73%***	60%
(C) FLOW-TRAJECT	37%	40%	49%	48%	47%
(D) DEG-FLOW-TRAJECT	72%***	67%**	69%**	71%**	70%*

Table 5. Exp 2: GJ task accuracy scores by L1, with significance above chance (***p<0.001, **p<0.01, *p<0.05).

	Group means: % accuracy (significance above chance)				
	Korean (N=36)	Turkish (N=25)	Arabic (N=13)	Chinese (N=14)	Japanese (N=10)
(A) DEG-FLOW	79%***	78%***	78%***	81%***	73%***
(B) DEG-TRAJECT	82%***	69%***	71%***	76%***	67%**
(C) FLOW-TRAJECT	28%	29%	49%	45%	43%
(D) DEG-FLOW-TRAJECT	65%**	59%	56%	72%**	43%

As with the analysis by proficiency level, the results were remarkably similar across L1 groups. Forty 2-tailed comparisons of proportions were conducted for each experiment to calculate differences in accuracy scores on the four types of combination by the five L1 groups (ten comparisons between languages for each stimulus type). In Experiment 1, only three out of forty comparisons revealed a significant difference, all due to the fact that the Japanese group scored relatively low on Type B (Jap/Kor, 60-78%, p=0.011; Jap/Tur, 60-80%, p=0.006; Jap/Ara, 60-82%, p=0.004). In Experiment 2, only six out of forty comparisons revealed a significant difference, two of which were due to the high score of the Koreans on Type B (Kor/Tur, 82-69%, p=0.005; Kor/Jap, 82-67%, p=0.026), two of which were on the flawed Type C (Kor/Ara, 28-49%, p=0.026; Tur/Ara, 29-49%, p=0.043), and two on the ternary combinations of Type D (Jap/Kor, 43-65%, p=0.038; Jap/Chi, 43-72%, p=0.016). Overall, with a small number of exceptions, performance on all combinations of modifiers was essentially the same regardless of L1 background.

One item effect that surfaced in the analysis by proficiency level was also seen in the analysis by L1. For Type C, the Experiment 1 results show that participants treated items (c1) *on through* and (c2) *on down* very differently from (c3) *back over*. Accuracy rates by L1 for (c1) were 28%, 44%, 40%, 29% and 30%; for (c2) they were 8%, 8%, 20%, 14% and 0%; whilst for (c3) they were 75%, 68%, 87%, 100% and 80%. However, just as in the first analysis, this difference was not apparent in Experiment 2, with poor performance on all stimuli. Given the design flaw previously identified, it remains to be seen whether this lower section of the modificational hierarchy is intrinsically more difficult to acquire or not.

The difference between the Degree modifiers *right* and *straight* found in the analysis by proficiency level was not found here, which was to be expected. More targetlike understanding of *righ*t was observable in the Lower-intermediate Group but the more advanced learners appeared to treat *right* and *straight* in the same way. The analysis by L1 group conflated proficiency levels, so this effect disappeared.

A further question we sought to examine was whether there might have been an interaction between L1 group and proficiency level if, for example, all learners from a single language had been in a single proficiency level. However, as neither the results between L1s nor those between proficiency levels had been significant, such an interaction could not have been possible. For the sake of completion, results from the three largest L1 groups were analysed by proficiency level, revealing that even within languages, performance was remarkably uniform across the levels. To take the largest L1 group as an example, accuracy scores for Korean learners in Experiment 1 for Groups 1 (11 learners), 2 (16 learners), and 3 (9 learners), respectively, were as follows: Type A: 81%, 81%, 82%; Type B: 70%, 89%, 84%; Type C: 36%, 34%, 45%; Type D: 67%, 69%, 85%. As expected, Experiment 2 replicated these patterns: Type A: 73%, 80%, 84%; Type B: 78%, 88%, 87%; Type C: 21%, 29%, 33%; Type D: 45%, 84%, 78%. Comparable results were obtained across the proficiency levels for Turkish and Arabic learners, confirming that the similarity across proficiency levels held within as well as across language groups.

3 Triangulating results across tasks

In comparing the accuracy scores across the two tasks, performance was strikingly uniform, whether analysed in terms of proficiency or L1. As for the three proficiency levels, the results for the four combination types were compared across tasks by means of twelve 2-tailed comparisons of proportions, revealing no significant differences. The lack of task effect was also shown when results were compared by L1. Twenty sets of results were compared, as the five language groups responded to the 4 types of combination, only 2 of which reached significance (Tur, Type B, Exp1-2, 80-69%, p=0.029; Jap, Type D, Exp 1-2, 70-43%, p=0.0.036). The fact that the accuracy scores were so balanced across the two tasks lends credence to the results.

However, close attention to the descriptive statistics suggests a few caveats. The numbers of participants for Arabic, Chinese and Japanese were considerably lower than those for Korean and Turkish, and, as might be expected, reveal more variation. To highlight an extreme example, the accuracy scores of Arabic speakers on Type D were 69% in Experiment 1, with 15 participants, and 56% in Experiment 2, with 13 participants, a difference that would have been significant with a greater number of speakers. However, given the overall tendencies of these learners, it seems plausible that a higher number of speakers would have resulted in a narrower gap. More generally, although Types A, B and C show no task effects at all, Type D contains a pattern which, though not significant, is nevertheless suggestive. Experiment 1 scores were higher than Experiment 2 scores for all three proficiency levels, and for four out of five L1 groups (with the other showing no difference). Moreover, Experiment 1 scores were all significantly higher than chance, for all proficiency groups and all L1 groups, but Experiment 2 scores were less impressive, with performance for several groups being at chance or only weakly significant. Our interpretations of this are as follows: that the learners were able to correctly choose between two alternatives in Experiment 1 arguably does indicate knowledge of the hierarchy despite the extra processing load induced by a ternary combination; that the learners were less targetlike in Experiment 2 may reflect the fact that they were ill-at-ease with such a long concatenation of modifiers.

V Discussion

The results reveal knowledge of the functional hierarchy across general proficiency levels, and irrespective of the L1, despite the fact that the relevant syntax is not taught in the classroom, and despite the scarcity of examples of multiple modifiers in natural input. Of the three hypotheses posited earlier, H3 can be discounted. H2 may be true to a degree, in that clearly the lexical semantics and prosody have to be in place, but the gradual development predicted by this hypothesis was not attested: even very low-level learners performed significantly above chance. It appears that there is direct mapping from lexical semantics to a pre-existing syntax right from the beginning, and that universals of prosody are likely to be in play. More gradual development was observed in performance on ternary combinations of modifiers, presumably due to an improvement in processing skills. With regard to binary combinations of modifiers, H1 is the hypothesis which is best supported by these results: the hierarchy is robustly in evidence from the outset.

This phenomenon permits an interesting perspective on investigations of syntax-semantics correspondences in L2 acquisition. Previously, several lines of research have shown that subtle differences in syntax can lead learners to semantic interpretations that could not possibly be gleaned from instruction or negative evidence, with phenomena such as adjectival restrictions on *wh*-quantifiers (Dekydtspotter & Sprouse, 2001) and *combien* extraction in French (Dekydtspotter, Sprouse & Swanson, 2001). Slabakova (2008) provides an insightful overview of such work, which sheds light on the flow of information at the interface from syntax to phrasal semantics. In an important sense, the current investigation has revealed the reverse of this process, with semantics informing syntax. However, the view adopted here is that there is not a single interface between syntax and semantics with a bidirectional flow of information. Rather, the correspondence is unidirectional, in different directions at two interfaces: from lexical semantics to syntax at the lexical interface, and from syntax to phrasal semantics at the interpretive interface. Slabakova (2006, 2008) is rightly cautious about distinguishing between lexical and phrasal semantics, and the interface with which she is concerned is that of interpretive semantics, following processes of syntactic derivation. If the lexical semantics of modifiers informs the syntax of directional predication, this flow of information from semantics to syntax is at a different point of contact between meaning and grammar, namely, the lexical interface. It is here that the semantics of lexical items helps mould initial syntactic representations (Hale and Keyser, 2002; Jackendoff, 1990; Pinker, 1989). While L2 research in this area has been productive with regard to the selectional properties of verbs (locatives: Bley-Vroman and Joo, 2001, Schwartz, Dekydtspotter, and Sprouse, 2003; datives: Bley-Vroman and Yoshinaga, 1992, Whong-Barr and Schwartz, 2002; psych verbs: White, Brown, Bruhn de Garavito, Chen, Hirakawa, and Montrul, 1999), the current study has charted new ground in its focus on lexical semantics informing the projection of functional hierarchies. The logic was as follows: if the modificational hierarchy is indeed part of the toolkit of UG, and if adult learners have continued access to UG beyond any purported critical period, then once knowledge of the lexical semantics of modifiers is acquired, the syntax of modification should be naturally manifested. This is indeed what was found.

VI Conclusion

This study explored the question of whether L2 learners have prior knowledge of the functional hierarchy of adpositional modifiers, despite the lack of instantiation of parts of the hierarchy in the L1, the lack of formal instruction, and the paucity of combinations of multiple modifiers in natural input. After instruction on the semantics of individual modifiers, a computer-animated narrative was used to contextualize combinations of modifiers, and a preference task and a grammaticality judgment task were administered to L2 learners of English from a variety of L1 backgrounds. The results show a conspicuous awareness of the functional hierarchy from lower-intermediate to advanced learners, irrespective of the L1, with responses triangulated across the two tasks. That learners were able to project the hierarchy on the basis of the lexical semantics of individual modifiers reveals a flow of information from semantics to syntax. However, we suggest that this not due to a bidirectional flow of information at the syntax-semantics interface, but rather that there are two distinct places of intersection between syntax and semantics, each of which exhibits a unidirectional flow of information: the lexical interface, at which lexical semantics informs syntax, and the interpretive interface, at which syntax informs phrasal semantics. That L2 learners so clearly apply semantics-syntax mappings at the lexical interface from individual items to functional categories not instantiated in the L1 constitutes sound evidence for full access to Universal Grammar in second language acquisition.

Acknowledgements

This research was conducted with the collaboration of the Intensive English Program (IEP) and the Center for Language Technology and Instructional Enrichment (CeLTIE), at Indiana University. Many thanks especially to Kathleen Bardovi-Harlig, Roger Crandall, Doreen Ewert, Lisa McCune, Heidi Vellenga, and Tara Zahler. We would also like to thank Kamil Ud Dean, Laurent Dekydtspotter, Bonnie Schwartz and William Snyder for insightful discussion as we framed the research questions and designed the experiments.

References

Alexiadou, A. 1997: Adverb placement: A case study in antisymmetric syntax. Amsterdam: John Benjamins.

Androutsopoulou, A., Español-Echevarría, M. and Prévost, P. 2008: On the acquisition of the prenominal placement of evaluative adjectives in L2 Spanish. In Bruhn de Garavito, J., and Valenzuela, editors, Selected proceedings of the 10th Hispanic Linguistics Symposium. Somerville, MA: Cascadilla, 1-12.

Ayano, S. 2001: The layered internal structure and external syntax of PP. Doctoral dissertation, University of Durham.

Baker, M. 1985: The Mirror Principle and morphosyntactic explanation. Linguistic Inquiry 16, 373-415.

Berman, R.A. and Slobin, D.I., editors, 1994: Relating events in narrative: A crosslinguistic developmental study. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bley-Vroman, R. and Joo, H-R. 2001: The acquisition and interpretation of English locative constructions by native speakers of Korean. Studies in Second Language Acquisition 23, 207-219.

Bley-Vroman, R. and Yoshinaga, N. 1992: Broad and narrow constraints on the English dative alternation: Some fundamental differences between native speakers and foreign language learners. University of Hawai'i Working Papers in ESL 11, 157-199.

Cinque, G. 1994: On the evidence for partial N movement in the Romance DP. In Cinque, G., Koster, J., Pollock, J-Y., Rizzi, L. and Zanuttini, R., editors, Paths toward Universal Grammar. Georgetown: Georgetown University Press, 85110.

Cinque, G. 1999: Adverbs and functional heads: A cross-linguistic perspective. Oxford: Oxford University Press.

Crisma, P. 1996: On the configurational nature of adjectival modification. In Zagona, K., editor, Grammatical theory and Romance language. Amsterdam: John Benjamins, 59-71.

Dekydtspotter, L. and Sprouse, R.A. 2001: Mental design and (second) language epistemology: Adjectival restrictions of *wh*-quantifiers and tense in English-French interlanguage. Second Language Research 17, 1-35.

Dekydtspotter, L., Sprouse, R.A. and Swanson, K.A.B. 2001: Reflexes of the mental architecture in second language acquisition: The interpretation of discontinuous *combien* extractions in English-French interlanguage. Language Acquisition 9, 175-227.

den Dikken, M. 2006: On the functional structure of locative and directional PPs. Ms., CUNY Graduate Center.

Emonds, J.E. 1985: A Unified Theory of Syntactic Categories. Dordrecht: Foris.

Hale, K. and Keyser, S.J. 2002: Prolegomenon to a Theory of Argument Structure. Cambridge, MA: MIT Press.

Hawkins, R., and Chan, C. Y. 1997: The partial availability of Universal Grammar in second language acquisition: The 'failed functional features hypothesis'. Second Language Research 13, 187-226.

Hohenstein, J.M., Naigles, L.R., and Eisenberg, A.R. 2004: Keeping verb acquisition in motion: A comparison of English and Spanish. In Hall, D.G, and Waxman, S.R., editors, Weaving a Lexicon, Cambridge, MA: MIT Press, 569-602.

Ionin, T. and Wexler, K. 2002: Optional verb raising in L2 acquisition: Evidence from L1-Russian learners of English. In Costa, J. and Freitas, M.J., editors, Proceedings of the GALA 2001 Conference on Language Acquisition. Lisbon: Associação Portuguese de Linguistica, 134-141.

Jackendoff, R. 1990: Semantic structures. Cambridge, MA: MIT Press.

Jespersen, O. 1992 [1924]: The Philosophy of Grammar. Chicago: The University of Chicago Press.

Koopman, H. 2000: Prepositions, postpositions, circumpositions and particles: The structure of Dutch PPs. In The syntax of specifiers and heads: Collected essays of Hilda J. Koopman. London: Routledge.

Pourcel, S. 2008: Investigating linguistic relativity: A research methodology. Durham Working Papers in Linguistics 8: 125-138.

Riemsdijk, Henk van. 1990: Functional prepositions. In Pinkster, H. and Genée, I., editors, Unity in Diversity. Dordrecht: Foris, 229-241.

Riemsdijk, H. van and Huybregts, R. 2007: Location and locality. In Karimi, S., Samiian, V., and Wilkins, W., editors, Clausal and phrasal architecture: Syntactic derivation and interpretation: A festschrift for Joseph E. Emonds. Amsterdam: John Benjamins, 339-364.

Schwartz, B.D. and Sprouse, R.A. 1996: L2 cognitive states and the Full Transfer/Full Access model. Second Language Research 12, 40-72.

Schwartz, B.D., Dekydspotter, L. and Sprouse, R.A. 2003: Pouring the fire with gasoline: Questioning conclusions on L2 argument structure. In Liceras, J.M., Zobl, H. and Goodluck, H., editors, Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002). Somerville, MA: Cascadilla Proceedings Project, 248-259.

Shlonsky, U. 2004: The form of Semitic noun phrases. Lingua 114: 1465–1526.
Slabakova, R. 2006: Is there a critical period for semantics? Second Language Research 22, 1-37.
Slabakova, R. 2008: Meaning in the Second Language. Berlin/New York: Mouton de Gruyter.
Stringer, D. (2005). *Paths in First Language Acquisition: Motion through Space in English, French and Japanese*. Doctoral dissertation, University of Durham.
Stringer, D. (2007). Extending the PP hierarchy: The role of bare nominals in spatial predication. In T. Scheffler, J. Tauberer, A. Eliam, and L. Mayol (eds.), *Penn Working Papers in Linguistics (PWPL) Vol.13.1: Proceedings of the 30th Annual Penn Linguistics Colloquium*: 379-392.
Stringer, D. (2008). Spatial affixes in syntax. Paper presented at Berkeley Workshop on Affix Ordering, University of California, Berkeley, October 4.
Strömqvist S. and Verhoeven, L., editors, 2004: Relating events in narrative: Typological and contextual perspectives. Mahwah, NJ: Lawrence Erlbaum.
Svenonius, P. 2008: Projections of P. In Asbury, A., Dotlačil, J., Gehrke, B. and Nouwen, R., editors, Syntax and Semantics of Spatial P. Amsterdam: John Benjamins, 63–84.
White, L., Brown, C., Bruhn de Garavito, J., Chen, D., Hirakawa, M. and Montrul, S. 1999: Psych verbs in second language acquisition. In Klein, E. and Martohardjono, G., editors, The development of second language grammars: A generative approach. Philadelphia: John Benjamins, 171-196.
Whong-Barr, M. and Schwartz, B.D. 2002: Morphological and syntactic transfer in child L2 acquisition of the English dative alternation. Studies in Second Language Acquisition 24, 579-616.