

- Herschensohn, J. 1998. Minimally raising the verb issue. In *Proceedings of the 22nd annual Boston University conference on language development*, ed. A. Greenhill, M. Hughes, H. Littlefield, and H. Walsh, 325–336. Somerville: Cascadia Press.
- Krashen, S., and T. Terrell. 1983. *The natural approach: Language acquisition in the classroom*. Oxford: Pergamon.
- Lardiere, D. 2009. Further thoughts on parameters and features in second language acquisition: A reply to peer comments. *Second Language Research* 25(3): 409–422.
- Larsen-Freeman, D. 1991. Teaching grammar. In *Teaching English as a second or foreign language*, 3rd ed, ed. M. Celcia-Murcia, 279–296. Boston: Heinle & Heinle.
- Larsen-Freeman, D. 2003. *Teaching language. From grammar to grammaring*. Boston: Heinle & Heinle.
- Lightfoot, D., and N. Hornstein. 1994. Verb movement: An introduction. In *Verb movement*, ed. D. Lightfoot and N. Hornstein, 1–18. Cambridge: Cambridge University Press.
- MacWhinney, B. 1997. Implicit and explicit processes – Commentary. *Studies in Second Language Acquisition* 9(2): 277–281.
- Norris, J., and L. Ortega. 2000. Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning* 50(3): 417–528.
- Pollock, J.-Y. 1989. Verb movement, UG and the structure of IP. *Linguistic Inquiry* 20: 365–424.
- Robertson, D., and A. Sorace. 1999. Losing the V2 constraint. In *The development of second language grammars: A generative approach*, ed. E. Klein and G. Martohardjono, 317–361. Amsterdam: John Benjamins.
- Schwartz, B., and M. Gubala-Ryzak. 1992. Learnability and grammar reorganization in L2A: Against negative evidence causing the unlearning of verb movement. *Second Language Research* 8: 1–38.
- Schwartz, B., and R. Sprouse. 1994. Word order and nominative case in nonnative language acquisition: A longitudinal study of L1 Turkish-L2 German interlanguage. In *Language acquisition studies in generative grammar*, ed. T. Hoekstra and B. Schwartz, 317–368. Amsterdam: John Benjamins.
- Schwartz, B., and R. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research* 12(1): 40–72.
- Spada, N., and Y. Tomita. 2010. Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning* 60(2): 263–308.
- Swan, M. 2006. Does grammar teaching work? *Modern English Teacher* 15(2): 5–13.
- Trahey, M., and L. White. 1993. Positive evidence and preemption in the second language classroom. *Second Language Research* 12: 111–139.
- Westergaard, Marit. 2003. Unlearning V2: Transfer, markedness and the importance of input cues in the acquisition of word order in English by Norwegian children. In *EUROSLA yearbook*, vol. 3, ed. S.H. Foster-Cohen and S. Pekarek Doehler, 77–101. Amsterdam: John Benjamins.
- White, L. 1990/1991. The verb-movement parameter in second language acquisition. *Language Acquisition* 1: 337–360.
- White, L. 1991. Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research* 7: 133–161.
- White, Lydia. 1992. Long and short verb movement in second language acquisition. *Canadian Journal of Linguistics* 37: 273–286.
- White, Lydia. 2003. *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press.
- White, L., N. Spada, P. Lightbrown, and L. Ranta. 1991. Input enhancement and L2 question formation. *Applied Linguistics* 12(4): 416–432.
- Yuan, B. 2001. The status of thematic verbs in the second language acquisition of Chinese: Against inevitability of thematic-verb raising in second language acquisition. *Second Language Research* 17: 248–272.
- Zobl, H. 1995. Converging evidence for the ‘acquisition-learning’ distinction. *Applied Linguistics* 16(1): 35–56.

Chapter 5

Modifying the Teaching of Modifiers: A Lesson from Universal Grammar

David Stringer

5.1 Introduction

One area of grammar that has received relatively scant attention in research on pedagogy involves the word order of modifiers, that is to say, modifiers of nouns (adjectives), verbs (adverbs), and prepositions (P-modifiers). While it is generally acknowledged that command of such elements is a fundamental part of grammatical knowledge, little is known about how learners develop systems of modifiers over the course of acquisition, and the efficacy of existing teaching materials is open to question. In this chapter, I consider the pedagogical implications of recent findings of second language research on language universals in the syntax of modifiers and in doing so illustrate how theoretically oriented, formal research can have practical implications for syllabus design and materials development.

The syntax in question is illustrated in the following example, in which the alternative orders of adjectives (**yellow lovely*), adverbs (**completely soon*), and P-modifiers (**back right*) are all clearly ungrammatical to native speakers.

1. The lovely yellow bird soon completely vanished right back into the trees.

Native judgments are also fairly robust even when there are three or more modifiers together, as exemplified below.

2. She bought a *beautiful old red wooden* box.
3. He *probably no longer completely* believes her.
4. I ran *straight on through* into the room.

Learners who achieve high levels of proficiency – Interagency Language Roundtable (ILR) level 4/Common European Framework of Reference (CEFR) level C2 – are

D. Stringer (✉)
Department of Second Language Studies, Indiana University, Memorial Hall 310,
Bloomington, IN 47405, USA
e-mail: ds6@indiana.edu

expected not to make errors with such orderings, yet it is not clear how this knowledge is to be acquired.¹ Adjective order is included as a topic in almost all North American English-as-a-second-language (ESL) textbook series, but the treatment is problematic in several ways: the materials are generally introduced in one-off topic-specific lessons, either at beginner, intermediate, or advanced levels, and are never systematically recycled; there is no evidence that students make use of the “rules” introduced; and teachers express doubts as to the usefulness of explicit instruction in this domain. The learnability problem is even greater for adverbs and P-modifiers, as multiple instances of these elements are essentially absent from current teaching materials, with no more than accidental occurrence.

While research on the acquisition of adjective order can shed light on the efficacy of existing teaching materials at all levels of instruction, related research on adverbs and P-modifiers can inform the development of materials at more advanced levels. Over the last decade or so, there has been a growing interest in discovering areas of grammar not covered in traditional syllabi for the purpose of improving instruction at higher levels of proficiency, especially in nonuniversity governmental institutions, for the training of military personnel, diplomats, and foreign correspondents. In the United States, institutions such as the Defense Language Institute, the Office of Naval Research, and the Foreign Service Institute have been concerned with the problem of raising proficiency from ILR level 3, superior, to level 4, distinguished. (The status of level 5 remains controversial.) A small number of university-sponsored centers have also focused on the development of near-native speakers, such as the Center for the Advancement of Distinguished Language Proficiency at San Diego State and the Center for Language Study at Yale. However, the vast majority of university-run English language programs offer no such instruction, the most “advanced” students graduating with a TOEFL PBT proficiency of 500–550 (approximately ILR level 2/2+). Strings of modifiers can be shibboleths in the native/nonnative distinction and, as such, must be addressed in the teaching process in some way. Although some of the implications of the studies discussed here may be for high-level instruction, participants in the experiments were drawn from a range of lower levels of proficiency (beginner to advanced in standard programs, ILR levels 1–2+), in order to more strictly control for degree of exposure to the target language while holding experimental instruction constant.

Learners need to overcome much more than the lack of available teaching materials, as they also face a learnability problem of a different nature: the need for constraints on the kinds of interlanguage grammars they generate. For example, in the case of modifiers, the margin for error increases dramatically with the number of elements combined, following the “*n* factorial” (in mathematical notation: *n*!). In a phrase such as *beautiful old red wooden box*, the 4 modifiers have 24 possible means of combination ($1 \times 2 \times 3 \times 4 = 24$), only one of which is correct. In a sentence such as *Jack fell straight back down from right up on top of the hill*, the 8 elements

¹The ILR and the CEFR are the standard governmental language proficiency scales used in the United States and Europe, respectively.

preceding the PP phrase *of the hill* have 40,320 possible combinations, again only one of which is correct.² Yet native speakers are unswervingly accurate in production, and in comprehension can process such strings in milliseconds. It seems unlikely that first language (L1) or second language (L2) learners of English could converge on the appropriate grammar in such cases without some kind of innate knowledge, if not of the order itself, then of more abstract underlying principles of combination. From the outset, research on Universal Grammar has been particularly concerned with this type of learnability problem (Chomsky 1957: 13). As we shall see, considering the multiplicity of logically possible alternatives, the fact that variation in the syntax of modifiers is so limited across languages bolsters the claim that language learning is guided by language universals. Formal research on this topic can reveal what such universals might be, thus enabling us to distinguish what learners already know and the specific knowledge they must acquire to master the syntax of modifiers in a particular language.

L2 research from the perspective of Universal Grammar has the potential to inform the teaching of modifiers with regard to two general sets of questions, the first concerning the role of language universals and the second concerning the role of L1 knowledge in processes of L2 development. The two studies discussed in this chapter address these issues in turn. The first study was developed from the assumption that comparative research on the syntax of P-modifiers is on the right track, such that there are aspects of grammar in this domain that do not have to be taught. However, as regards implications for pedagogy, the precise nature of the universals is important. If the syntactic structures involved exist independently of lexical items, as “templates” or “constructions,” then frequency of exposure to such constructions should facilitate acquisition. If, however, the grammatically relevant aspects of word meaning are sufficient to determine the syntactic environment, according to universal mapping principles, then classroom teaching can focus on contextualized vocabulary: the correct word order should be naturally manifested once the lexical semantics is in place, without specific instruction on multiple modifiers.

The second study, *inter alia*, addressed the question of the influence of the L1 in the acquisition of English adjective order. In early approaches to “language transfer,” some argued that a priori contrastive analyses of languages could make predictions about where learner problems were likely to be found, and teaching materials could be designed following such analyses (Lado 1957). Others maintained that as a priori analyses sometimes predict errors that do not occur and predict ease of acquisition in areas where problems do surface, it is best to wait to see the kinds of errors that learners make and then amend teaching materials following an a posteriori error analysis (Gradman 1971; for further discussion, see Schachter 1974; Gass 2013). In designing the study, both approaches were pursued: predictions were made based on a comparative syntactic analysis of the relevant languages, and learner data was experimentally solicited in the form of different kinds of judgment

²Abbreviations used for syntactic categories are as follows: N (noun), V (verb), Adj (adjective), Adv (adverb), and P (preposition/postposition/particle); phrasal projections are abbreviated NP, VP, AdjP, AdvP, and PP.

tasks, to identify areas of instructional need based on actual learner responses. General questions concerning transfer include the following. In what ways does modifier order vary across languages? What predictions might there be for crosslinguistic influence? Is there evidence that learners transfer aspects of syntax from their L1s, leading to different paths of development, such that different L1 groups would benefit from different teaching materials?

The answers to these questions, of course, may differ according to the precise grammatical domain; for example, in the world's languages, there appears to be much more variation in the syntax of adjectives than P-modifiers. They may even differ within a syntactic category; for example, certain combinations of adjectives may follow a fixed, universal order, while others may be subject to language-specific preferences. Nevertheless, formally oriented studies of syntactic representation, crosslinguistic variation, and second language acquisition can provide the foundations for syllabus and materials design in this area. Ideally, teachers faced with the task of instructing their class on the use of modifiers, or confronted with spontaneous questions in class, should be able to understand what learners already know and what they need to learn, what differences may exist among different L1 groups of learners, and whether teaching should focus on syntactic hierarchies (as is currently the case with published materials on adjective order) or whether the syntax will appear naturally once the meanings of the modifiers are in place.

In Section 5.2, a review is given of the type of pedagogical materials already available, and the efficacy of such materials is questioned. Recent linguistic insights into the nature of syntactic hierarchies of modifiers are briefly summarized, and an outline is given of the second language research project, currently in progress, whose findings to date are discussed in Sections 5.4 and 5.5. These sections provide summary reports of experiments investigating the L2 acquisition of P-modifiers and adjectives, respectively, with some clear, immediate implications for second language pedagogy, as well as some results which require further investigation. Finally, these findings are discussed in the more general context of how research on language universals can support language teaching.

5.2 Hierarchies of Modifiers: Beyond the Textbook

In order to ascertain whether or not ESL learners in the United States are exposed to instruction on word order of modifiers, a survey was conducted of pedagogical materials in the ESL library of Indiana University, one of the more established centers of English language teaching in the United States, which contained mainstream textbooks from the last four decades of instruction (1970–2010). In particular, all grammar series were consulted, with specific attention paid to those currently used by the students who were to participate in the experiments discussed below. Of the three hierarchies of modifiers, only adjective order has ever received any attention in standard American ESL textbooks. Moreover, instruction has never been systematic: in no textbook series was there any recycling of materials. In a single lesson,

learners are presented with the word order of adjectives, and it is hoped that from that lesson on they will consciously remember that *size* come before *depth* or that *age* comes before *material* or that *weight* comes before *nationality*. Strikingly, although these word order facts are often thought of by linguists as quite complicated, and thus appropriate for higher-level instruction, such lessons may appear at any stage of proficiency, depending on the series. For example, Azar and Hagen (2006a: 410–413) introduce the hierarchy [opinion>size>age>color>nationality>material] at the beginner level; Thewlis (2007: 150–152) introduces the hierarchy [evaluation/opinion>appearance (usually size>shape>condition)>age>color>origin (geographical>material)] at the intermediate level; and Maurer (2000: 138–145) introduces the hierarchy [opinions/qualities>size/height/length>age/temperature>shapes>colors>nationalities/social classes/origins>materials] at the “advanced” level (approximately ILR level 2). The textbooks and accompanying workbooks encourage students to memorize the order of categories of adjectives and then provide practice through exercises typically involving preference tasks, error correction, and unscrambling. Examples of each are given below.

Preference task: Azar and Hagen (2006b: 216)

Directions: Choose the correct completions.

We work in ____ office building.

A. a large old B. an old large

Error correction: Thewlis (2007: 152)

Are these sentences correct or incorrect? If they are incorrect, identify the problem and correct it.

I bought a green, old, pretty vase at the flea market.

Unscrambling: adapted from Maurer (2000: 144)

Unscramble the sentences in the conversation.

BILL: This is (party / office / formal / a) isn't it? What if I wear (tie / my / silk / new)?

NANCY: That's fine, but don't wear (shirt / purple / ugly / that / denim) with it. People will think you don't have (clothes / any / suitable / dress-up).

Anecdotal reports and online blogs reveal uncertainty on the part of language teachers about the efficacy of such materials. A representative comment from an ESL weblog runs as follows.

5. “If explicitly teaching grammar and syntax is largely ineffective, then explicitly teaching English adjective order must be nearly at the top of the list of ineffectual classroom activities. It just feels like one of those things that simply, but not easily, needs to be “picked up””

(Retrieved November 12, 2010 from <http://eshweb.net/blog/?p=287>)

As yet, there is no research available that speaks to learning outcomes following this type of instruction, and no treatments have been implemented that allow

Table 5.1 Three categories of spatial modifiers

1 Degree	2 Flow	3 Trajectory	PP
Right	On	Over	Into the valley
Straight	Back	Through	
		Across	
		Up	
		Down	

down, through, over, and across. Their status as modifiers can be distinguished from their status as prepositions by means of tests of syntactic distribution as in (9). Not all languages lexicalize all types of modifier, but when two or more are found, they conform to syntactic predictions: German and English lexicalize all three; Estonian and Hungarian only have the higher two; French and Spanish only have the highest; and Japanese and Korean have none at all, expressing such functions elsewhere in the grammar (see Stringer et al. 2011, for further discussion).

For this set of experiments, the focus was not on transfer but on testing for knowledge of the universal hierarchy even when there are no multiple P-modifiers in the L1. Given the scope for possible errors, the lack of instruction, and the relative rarity of multiple modifiers in the input, accuracy in the absence of the possibility of L1 transfer would suggest access to Universal Grammar on the part of adult L2 learners. In addition, the experiments sought to shed light on the question of whether target-like accuracy in the word order of prepositional modifiers depends on (i) frequency of exposure to the hierarchy itself or (ii) acquisition of the lexical semantics of the individual modifiers, such that the hierarchy is naturally manifested. If (i), then pedagogical materials might be developed that include selected instances of multiple modifiers; if (ii), then explicit teaching of the hierarchy is unnecessary, and advanced instruction should focus on contextualized vocabulary rather than syntax.

5.3.1 Experiment I: Aladdin Preference Task

The participants were all enrolled in a university Intensive English Program. The six proficiency levels tested were derived independently of the 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 121 students successfully took part in the first experiment, after ten subjects were eliminated according to preestablished criteria. For purposes of analysis, participants were grouped as follows: lower intermediate ($n=42$), intermediate ($n=41$), and advanced ($n=38$). Learners came from 17 different first language (L1) backgrounds. While all language backgrounds were represented in the general analysis by

proficiency level, a comparative analysis was also made of the performance of learners from five L1 groups with no evidence for multiple modifiers in the native language: Korean, Turkish, Arabic, Chinese, and Japanese. None of these languages has more than one level of the hierarchy instantiated: Korean, Japanese, and Turkish have no P-modifiers, and Arabic and Chinese have at most one type (for discussion, see Stringer et al. 2011). In short, these learners must project a syntactic hierarchy that is absent in the L1. The experiments were conducted in a language lab with a communal main screen and surround speakers, so that it was possible to synchronize aural and visual stimuli for all participants. A control experiment was conducted with 20 native speakers of English, aged 19–48, all of whom had spent most of their lives in the Midwest of the United States.

An original animated slideshow was designed to contextualize PPs and their modifiers, which took the form of a narrative variation on the story of Aladdin. PowerPoint slides of each scene were created by scanning hand-drawn images, arranging them in layers depending on the desired visibility of objects, and then animating the slides. The embedding of 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. 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 complete animation is available for download from the author's professional webpage (<http://www.indiana.edu/~dsls/faculty/stringer.shtml>), and the linguistic materials subject to experimental manipulation are reproduced in Appendix I.

There were 6 examples of degree–flow, 6 of degree–trajectory, 3 of flow–trajectory, and 3 of degree–flow–trajectory: the stimuli were balanced as well as possible within narrative constraints. Prosody plays a pivotal role in the parsing of phrases with multiple modifiers: the insertion of pauses, shifting of stress, or other variance in the intonational contour results in the assignment of a different syntactic structure with a different semantic interpretation. The most appropriate prosody for stimuli was selected on an item-by-item basis, based on native-speaker judgments, and sentences were embedded as sound files in the slides. In addition, participants' response sheets contained no written cues to reduce the risk of prosodic rephrasing during the experiment.

In advance of the experimentation, the vocabulary to be used was presented to the students for the purpose of making clear the meaning of each of the 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. 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). 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*, and *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.

Experiment I was a 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. All linguistic stimuli were repeated once after a four-second pause. An example stimulus from Experiment I is given below.

10. Experiment I sample:

"He flies _____ over the camels." (*on straight/straight on)
 A B

Participants were told to listen to how the sentences sounded and to judge them immediately on the way they sounded, without considering other pronunciations. The order of presentation of target-like and non-target-like variants was systematically varied across stimuli.

Once the results were tabulated, a mixed design ANOVA was conducted with stimulus type as the within-subject factor and proficiency group and L1 as between-subject factors. The stimulus types were as previously described. The proficiency levels included the three learner groups and the native controls. L1 was assessed in terms of five L1 populations: Korean (36), Turkish (25), Arabic (15), Chinese (14), and Japanese (12), as well as the controls. The results for all four proficiency groups and the five L1 populations are given in Figs. 5.1 and 5.2. The p -values displayed indicate significance above chance and are unadjusted from t -statistics using estimated means and standard errors from the repeated measures ANOVA. The possibility of false-positive increases due to multiple comparisons was controlled by using Benjamini and Hochberg's (1995) method for false discovery rate. The results of the native-English controls, who served as both a proficiency group and a language group, are reported only once, in Fig. 5.1.

In brief, a main effect of stimulus type was found, but this was due only to the results for Type C. The results for Types A, B, and D did not reveal any significant differences. A main effect of proficiency was also found, due to two comparisons: Group 4 (the native controls) was significantly different from all other groups, and a small but significant difference was also found between Proficiency Groups 1 and 3, $t(78)=2.878$, $p=.031$, $\eta^2=.071$. With respect to L1 background, no significant differences were found between learner groups. There was no interaction between L1 background and proficiency level, and performance was remarkably uniform across the levels within each language.

A glance at the descriptive statistics immediately reveals a difference between the relatively accurate performance on Types A and B at all proficiency levels (A, 76 %, 74 %, 84 %; B, 71 %, 78 %, 81 %), the particularly non-target-like performance on Type C at all proficiency levels (41 %, 38 %, 44 %), and performance on Type D, which showed significant accuracy at all levels but improvement with general proficiency (64 %, 68 %, 76 %). The same pattern may be observed in the analysis by L1 group, with accurate performance on Types A and B by Korean, Turkish, Arabic, Chinese, and Japanese learners (A, 78 %, 73 %, 79 %, 77 %, 80 %; B, 78 %, 80 %, 82 %, 73 %, 60 %), poor performance on Type C (37 %, 40 %, 49 %, 48 %, 47 %),

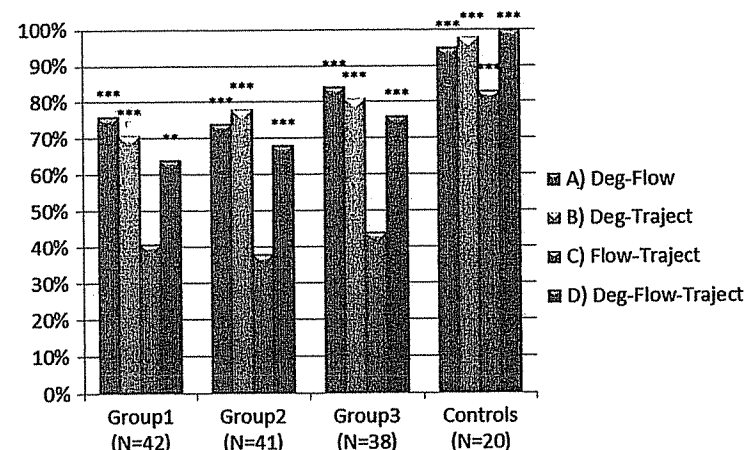


Fig. 5.1 Preference task accuracy scores by proficiency level, with significance above chance (***) $p < 0.001$, ** $p < 0.01$, * $p < 0.05$)

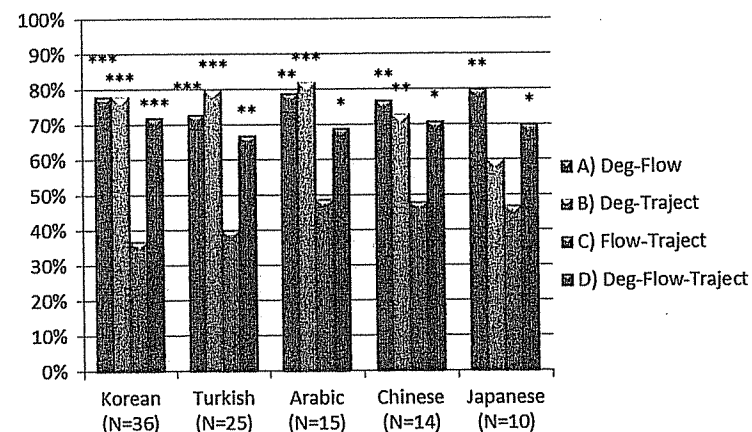


Fig. 5.2 Preference task accuracy scores by L1, with significance above chance (***) $p < 0.001$, ** $p < 0.01$, * $p < 0.05$)

but with no improvement on the more complex ternary combinations of Type D as proficiencies were collapsed in the L1 data (72 %, 67 %, 69 %, 71 %, 70 %).

The high accuracy rates on Types A and B in evidence at all proficiency levels are particularly striking. In comparison, the generally weak performance on Type C calls out for further scrutiny. Although it might appear that the lower reaches of the hierarchy pose a higher degree of difficulty, analysis by individual stimuli reveals that poor performance on Type C might be alternatively explained in

terms of a lexical effect. In Experiment I, participants treated items (c1) *on through* and (c2) *on down* very differently from (c3) *back over*. Accuracy rates for the former were not significantly above chance, but scores on (c3) in Proficiency Groups 1, 2, and 3 were, respectively, 67 % ($p = .023$), 80 % ($p < .001$), and 82 % ($p < .001$). One possible reason for this discrepancy might be that the PPs modified by these combinations were headed by *to*: *on through [to the outside]*, *on down [to the ground]*, and *back over [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*. This issue is revisited in discussion of Experiment II. Performance on the ternary combinations of Type D was significantly above chance, though showing an increase in accuracy with proficiency. This was to be expected given the increase in processing load. These examples were included to stretch learners, as native responses were so robust: the controls attained 100 % accuracy for this type.

To summarize the results, 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. There was no L1 effect.

5.3.2 Experiment II: Aladdin Grammaticality Judgment Task

Experiment II was conducted with the same participants to obtain binary judgments of grammaticality, rather than preference judgments, and to control for any possible task effects. Following the same criteria for exclusion, 13 participants were eliminated, leaving a total of 118 students. Again, results were analyzed in terms of three general proficiency groups: lower intermediate (levels 2–4, $n = 41$), intermediate (level 5, $n = 40$), and advanced (levels 6–7, $n = 37$). As before, learners came from 17 different L1 backgrounds, five of which were tested for L1 effects: Korean, Turkish, Arabic, Chinese, and Japanese.

The Aladdin animation was run again, but this time with different embedded sound files. For each slide, a male voice asked a question about the narrative, and a female voice answered by means of a sentence fragment, which participants judged as good or bad. The order of presentation of target-like and non-target-like variants was systematically varied across stimuli. An example stimulus from Experiment II is given below.

11. Experiment II 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, prosody can disambiguate between [*he flies back*] [*right into the desert*], which is grammatical, and *[*he flies [back right into the desert]*], which is not. An ungrammatical sentence fragment answer such as *[*back right into the desert*] provides a clear contrast to the grammatical [*right back into the desert*] and reduces the chance of the modifier being reanalyzed as a verb particle.

Again, a mixed design ANOVA was conducted with stimulus type as the within-subject factor and proficiency group and L1 as between-subject factors. As before, the analysis by proficiency included the three learner groups and the native controls, and L1 was assessed in terms of the five largest L1 populations: Korean (36), Turkish (25), Arabic (13, as compared to 15 in Experiment I), Chinese (14), and Japanese (12). The results for all four proficiency levels and the five L1 populations are given in Figs. 5.3 and 5.4.

As in Experiment I, a main effect of stimulus type was found, with Type C significantly different from the other three types. However, in Experiment II, significant differences were also found for Type D vs. Type A and Type D vs. Type B, both $p < .001$. There was no effect of proficiency level. With respect to language background, no significant differences were found between learner groups. Unlike in Experiment I, an interaction of stimulus type and language group was found, $F(12.286, 292.409) = 2.092$, $p = .017$, due to the poor performance of the smallest groups (Arabic and Japanese) on Types C and D. As before, there was no interaction between L1 group and proficiency, and performance was consistent across the levels.

Again, the descriptive statistics clearly indicate the difference between the relatively accurate performance on Types A and B at all proficiency levels (A: 76 %, 81 %, 80 %; B: 69 %, 77 %, 79 %) in comparison with the other two types. Type C stimuli produced a notably non-target-like performance at all proficiency levels (34 %, 39 %, 32 %), and Type D again showed improvement with general proficiency, although the accuracy levels were lower at each proficiency level than in Experiment I (58 %, 63 %, 68 %). In the previous experiment, the analysis of Type C results revealed considerably lower rates of accuracy for items (c1) *on through* and (c2) *on down* as compared to (c3) *back over*, and it was hypothesized that the first two might have been phonologically rephrased by participants, so that the displaced *on* could merge with the following preposition *to*, resulting in *onto*. However, in Experiment II, the results did not reveal the same discrepancy, with poor performance on all stimuli. It is notable that the control subjects also had difficulty with (c1) in particular, with scores of 50 % in Experiment I and 65 % in Experiment II, bringing down the average accuracy for this type. It remains a possibility that some controls rephrased this stimulus, deriving a legitimate structure: *He flew through, on to the outside*. Given the design flaw in juxtaposing *on* and *to* (albeit a legitimate combination in the target language), a more detailed examination of L2 knowledge of flow-trajectory must be left for future work, in which the *to*-PP might be replaced with, for example, an *into*-PP. The ternary combinations of Type D again proved more difficult for lower-level learners, although accuracy

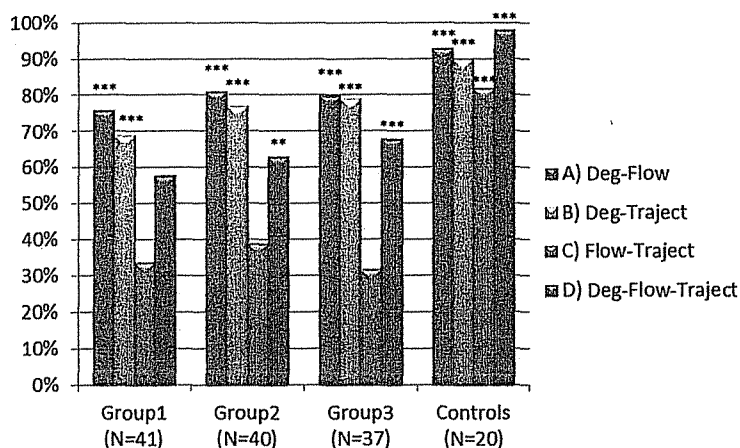


Fig. 5.3 GJ task accuracy scores by proficiency level, with significance above chance (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$)

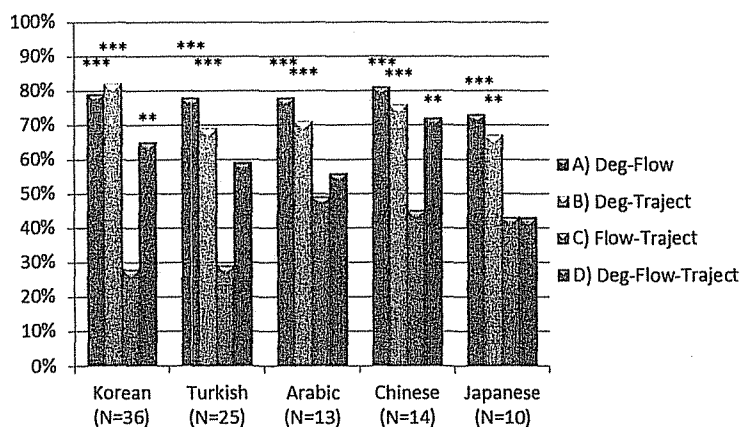


Fig. 5.4 GJ task accuracy scores by L1, with significance above chance (*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$)

generally improved with proficiency. Group 1 results were not significantly above chance, while Groups 2 and 3 showed increasingly significant rates of accuracy. This pattern conforms to our understanding of these combinations as involving a higher processing load. Native-speaker responses were unequivocal at 98 % accuracy.

To summarize the results, just as in the previous experiment, the learners revealed impressive rates of accuracy for the binary combinations of Types A and B, consistently underperformed on the flawed Type C, and showed increased accuracy with proficiency on the ternary combinations of Type D. Again, there was no L1 effect.

In order to uncover any task effects, a third ANOVA was conducted with task and stimulus type as within-subject factors and proficiency group and L1 as between-subject factors. A main effect of task was found, $F(1, 119) = 8.632$, $p = .004$, which may be understood on closer analysis of the performances of learners grouped by proficiency level and by language background. The proficiency groups displayed slightly different patterns of responses by task. Group 1 and the control group showed no task effect. Group 2 performed differently on Type A in task 1 and task 2: $t(41) = 2.142$, $p = .035$; however, the effect size was small ($\eta^2 = .037$), and there were no differences for B, C, and D. Group 3 performed differently on Type D ($t(36) = 2.774$, $p = .006$); again, the effect size was small ($\eta^2 = .061$), and there were no differences for A, B, and C. Despite the general task effect, this more detailed analysis in terms of proficiency levels reveals that the similarities in performance are more striking than the differences.

An analysis by L1 group furnishes a similar understanding. Of the 24 possible pairwise comparisons (6 language groups \times 4 stimulus types), only three produced a significant difference: Arabic speakers only on Type D, Japanese speakers only on Type D, and Turkish speakers only on Type B. In conclusion, while the ANOVA did reveal a main effect of task, the results of the two experiments remain highly comparable.

Recall that the main purpose of the Aladdin experiments was to test for knowledge of the universal hierarchy in the toughest test case scenario: when the L1 has no multiple modifiers. An additional goal was to test whether the acquisition of lexical semantics was not only *necessary* but also *sufficient* for the acquisition of the syntax of P-modifiers. The implications of these results for pedagogy are clear in these two regards: even when the L1 has no instances of multiple P-modifiers, the hierarchy is naturally manifested in adult L2 acquisition, with rates of accuracy reliably and significantly above chance. Once the lexical semantics of individual modifiers was in place, this was sufficient for accuracy on the hierarchy, even at the lowest levels of proficiency, with no instruction on particular combinations necessary. The implication for high-advanced levels of instruction is that the syntactic complexity of modifier hierarchies actually comes for free; teaching materials should focus on lexical semantics rather than word order in this domain.

5.4 L2 Acquisition of Adjective Order

As mentioned earlier, L2 acquisition of adjective ordering restrictions (AOR) remains to be investigated, and their commonplace inclusion in ESL curricula is unprincipled: they appear at introductory, intermediate, or advanced levels of textbooks, and materials are not recycled. This is an uncharted area of L2 knowledge. However, recent syntactic research has expanded our understanding of language universals in this domain: the same syntactic hierarchy is found in all languages that allow direct hierarchical modification, with some systematic variation (e.g., Cinque 2010; Laenzlinger 2005; Scott 2002; Shlonsky 2004; Sproat and Shih 1991).

The second part of the project on the L2 acquisition of modifiers draws on such work and seeks to uncover what role, if any, this hierarchy plays in the acquisition of English as a second language (Stringer et al. In prep).

Although the data from these large-scale experiments are still under analysis, the interim findings are directly relevant to the question of how generative research can inform classroom pedagogy, and the following work-in-progress report extends the previous discussion by showing how generative research may shed light not only on language universals but also on the issue of L1 transfer. That the two Aladdin experiments did not reveal L1 transfer was not surprising, as the five L1s chosen for analysis were alike in the relevant respect: none of them had multiple modifiers. The syntax of adjectives, however, admits much more crosslinguistic variation than that of P-modifiers, and in approaching the design of these experiments, an attempt was made to test specifically for L1 influence. After a dismissive approach to L1 transfer in generative L2 research of the 1970s and 1980s (e.g., Dulay and Burt 1974; Krashen 1981), most researchers have come to believe that L1 influence has a major role to play in acquisition of syntax (Schwartz and Sprouse 1996), phonology (Strange and Schafer 2008), morphology (Montrul 2000), and the lexicon (Stringer 2010). However, the fact remains that in several subdomains, transfer is not apparent or at best has effects at later stages of development (see Hawkins 2001, for discussion). If L2 learners of English follow very different paths of development in their understanding of English adjective order, in ways predictable on the basis of their L1 grammars, this would support an approach to pedagogy which takes the L1 into account, ideally with teaching materials specific to particular L1 groups of learners. If, however, there is no L1 influence in this domain, this would in turn support the use of the same teaching materials irrespective of the language background of learners.

The L1 groups selected for investigation included Arabic-, Korean-, and Chinese-speaking learners of English, as these differ from English and from each other in interesting ways. Arabic has strict AOR in the relevant conditions, but post-nominally, and in the mirror order, as shown in the following example (Fassi-Fehri 1999: 107):

12. l-kitaab-u l-?axdar-u š-šaḡiir-u Arabic
the-book-NOM the-green-NOM the-little-NOM
“The little green book.”

In contrast, Chinese and Korean usually mark adjectives with an “adjective marker,” homophonous with a relative clause marker and glossed here as REL, in which case ordering restrictions do not apply.³ These variations are based on a Chinese example provided by Sproat and Shih (1991: 565–566).

³The assumption here is that Korean adjectives are, in fact, relative clauses. The link between relative clauses and attributive adjectives is well-known, albeit complex and controversial (compare: *the boy who is tall* and *the tall boy*; *the train which is moving fast* and *the fast-moving train*). Crosslinguistically, it appears that adjectives marked with relative markers, just like relative clauses themselves, are not subject to ordering restrictions (compare: *the {great new/*new great} café*; *the café {which is great and which is new/which is new and which is great}*). For a review of analyses linking relatives and attributive adjectives, see Alexiadou et al. (2007).

13. hǎo-de yuán-de pánzi/yuán-de hǎo-de pánzi Chinese
good-REL round-REL plate/round-REL good-REL plate
“nice round plate”
14. metji-n dung-eun jeopsi/dung-eun metji-n jeopsi Korean
nice-REL round-REL plate/round-REL nice-REL plate
“nice round plate”

Thus far, Chinese appears to pattern like Korean; however, there is a crucial difference. While the relativizer is obligatory in Korean, it is optional for most monosyllabic adjectives in Chinese. When it is omitted, AOR robustly reappear (as if by magic), as shown below.

15. hǎo yuán pánzi/*yuán hǎo pánzi Chinese
good round plate/round good plate
“nice round plate”






An additional condition in Chinese is that direct adjectival modification is limited to two elements, one *nonabsolute* (i.e., gradable) adjective and one *absolute* (i.e., ungradable) adjective (Sproat and Shih 1991: 588–591). This restriction has also been observed in other languages, such as Italian (Cinque 1994: 95–96, fn.15), and will be relevant to interpretation of the experimental data discussed below.

The predictions formulated on the basis of L1 differences were as follows. First, Arabic learners might have initial mirror order but then reset the relevant parameter, understood here in terms of “snowball movement” (Shlonsky 2004). Second, on the assumption that attributive adjectives are hosted in dedicated functional projections above NP (Cinque 2010; Laenzlinger 2005; Scott 2002), Koreans might have difficulties with the instantiation of new functional categories and be subject to prolonged confusion. Third, there should be facilitation for Chinese learners in contrast to Korean learners regarding nonabsolute–absolute combinations.


In this preliminary L2 study of the relevant syntax and semantics, investigation was restricted to modification of object nominals (rather than event nominals) and to 14 of the proposed universal types, categorized in a simplified version of Laenzlinger’s (2005) variant of the hierarchy: [evaluative opinion > [scalar physical property size > length > height > speed > depth > width > [measure temperature > wetness > age > [non-scalar physical property shape > color > nationality > material]]]]. At a higher level of categorization, nonabsolute adjectives were understood to precede absolute (i.e., non-scalar physical property) adjectives, and this distinction was also coded in the experimental design.

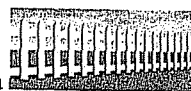
An experiment was administered to 204 ESL learners from 14 different L1 backgrounds, across 5 levels of proficiency, as well as 20 native controls, to examine knowledge of universals and possible L1 transfer effects. This was a binary preference task, with recorded oral delivery of 47 stimuli controlling for prosody and the answer sheets incorporating an original rebus design, such that images replaced the objects described. As instruction on all of the high-frequency adjectives was not possible prior to testing, images were used instead of gaps in written stimuli, to facilitate lexical retrieval. The idea of the rebus technique was taken not from adult word puzzles but from children’s books. Rebus designs are quite common in early


literacy texts, often being incorporated into poems and songs with which children are already familiar. Examples such as the following are easily constructed.⁴

16. Twinkle, twinkle, little , how  wonder what you are
 above the  so high, like a  in the sky.

This technique was adapted for the purposes of the experiment in order to furnish examples with combinations of two, three, and four adjectives, as in (17–19) below. (Note that the images in the actual experiment were in color.) Participants heard recordings of the full sentences (controlled for prosody) read once with one order of adjectives and once with the alternative order and then had 4 seconds to complete the forced preference task by circling either A or B.

17. Brian is talking on a  phone.
 (A) *thin great* (B) *great thin*

18. The house has a  fence.
 (A) *long white wooden* (B) *wooden white long*

19. Daniel likes these  flowers.
 (A) *plastic pink small nice* (B) *nice small pink plastic*

Test materials included 14 binary combinations of nonabsolute–absolute (e.g., opinion–material, *dangerous stone steps*), 8 binary combinations of nonabsolute–nonabsolute (e.g., size–age, *big old car*), 2 binary combinations of absolute–absolute (e.g., color–material, *pink plastic umbrella*), 8 further combinations of nonabsolute–nonabsolute specifically targeting scalar physical properties (e.g., length–height, *long high wall*), and 4 combinations each of 3 adjectives (e.g., *big old stone tower*) and 4 adjectives (e.g., *beautiful long white wooden chair*).

While the complete analysis of the results with appropriate statistical analysis is not yet complete, as testing of controls is still in progress, initial results from the learner groups have already furnished striking patterns, confounding expectations, and pointing toward unexpected implications for classroom pedagogy. The relevant

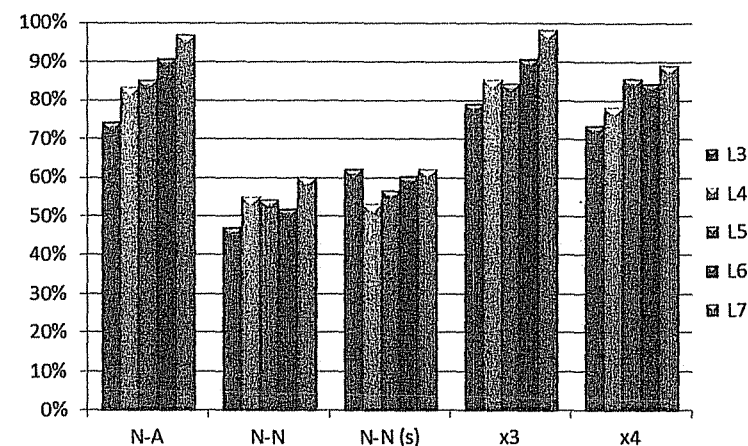


Fig. 5.5 Experiment III: Preference task, results by proficiency group (levels 3–7). Percentage accuracy on adjective combinations: nonabsolute–absolute (N-A), nonabsolute–nonabsolute (N-N), nonabsolute–nonabsolute scalar (N-N (s)), strings of 3 adjectives (x3), and strings of 4 adjectives (x4)

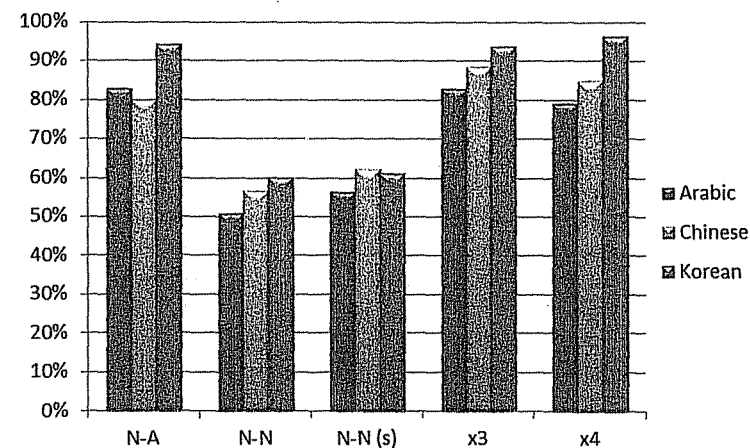


Fig. 5.6 Experiment III: Preference task, results by L1 group. Percentage accuracy on adjective combinations: nonabsolute–absolute (N-A), nonabsolute–nonabsolute (N-N), nonabsolute–nonabsolute scalar (N-N (s)), strings of 3 adjectives (x3), and strings of 4 adjectives (x4)

interim results are provided in Fig. 5.5, which shows accuracy scores in terms of the 5 levels of proficiency, and Fig. 5.6, which contrasts the accuracy scores for Arabic ($N=119$), Chinese ($N=23$), and Korean ($N=21$) learners.

A repeated measure ANOVA reveals three sets of findings. First, let us consider the results in terms of proficiency level. All proficiency levels, from level 3 (TOEFL

⁴The example in (16) is a popular English children's song: "Twinkle, twinkle, little star / How I wonder what you are / Up above the world so high / Like a diamond in the sky."

PBT <400) to level 7 (TOEFL PBT >500), display robust knowledge of the ordering of nonabsolute and absolute adjectives, with accuracy improving as the proficiency level increases (74 %, 83 %, 85 %, 91 %, 97 %). This was an unexpected finding: although this distinction is relevant for languages such as Chinese as discussed earlier, Korean and Arabic learners of English have no specific evidence either in the L1 or the L2 that this distinction is more important than any other in the adjectival hierarchy. In stark contrast, the performance of participants is at or just above chance for combinations of two nonabsolute adjectives, even at higher levels of proficiency. The differences between accuracy on nonabsolute–absolute combinations and nonabsolute–nonabsolute combinations (of both subtypes) are significant for all comparisons ($p < .001$).

Second, we may revisit the hypotheses regarding L1 effects. All three groups revealed knowledge of the nonabsolute–absolute distinction (Arabic, 83 %; Chinese, 80 %; Korean, 94 %), and none performed above chance on combinations of two nonabsolute adjectives. Arabic learners did not show any evidence of a reliance on the mirror order, either at lower or higher stages of proficiency. In addition, the projected difference between L1 Korean and L1 Chinese never materialized. Rather than being at a disadvantage due to the lack of any evidence for AOR in the L1, the Korean learners actually outperformed the Chinese learners on the nonabsolute–absolute distinction. However, when proficiency was factored into the analysis, there was no significant difference between the two language groups. Thus the hypothesis that L1 transfer should lead to different paths of development, and thus different types of teaching materials in idealized learning environments, was roundly falsified.

Third, performance on ternary and quaternary combinations was significantly above chance at all levels and improved over the proficiency range; though given the poor results on adjectives matched for “absoluteness,” we must question whether these high scores reflect knowledge of more complex manifestations of the hierarchy or simply a successful test strategy based on knowledge of absoluteness (one alternative always began with a nonabsolute adjective, and one with an absolute adjective).

The two main conclusions from this early stage of analysis with potential for direct application to classroom pedagogy are as follows. First, AOR do not come for free, beyond the nonabsolute–absolute distinction. Unlike the hierarchy of P-modifiers, which appears to manifest itself naturally in the course of acquisition, the adjectival hierarchy seems to be much more complicated for learners to acquire in cases where elements are matched for absoluteness. While the nonabsolute–absolute distinction is plausibly part of Universal Grammar, teaching materials must be developed to advance students’ knowledge of other combinatorial possibilities within the hierarchy. Second, there appears to be a complete absence of crosslinguistic influence. Despite intriguing differences in L1 grammars, it appears that the knowledge L2 learners bring to the acquisition process is common to learners from markedly different language backgrounds. They all show a remarkably robust understanding of the nonabsolute–absolute distinction yet reveal identical patterns of persistent confusion with regard to

other binary combinations. An important practical implication is that teaching materials need not be tailored to the L1s of the learners.

5.5 Conclusion

Generative approaches to second language research have the potential to furnish valuable insights for classroom pedagogy, even if pedagogy is not the primary focus of such research. The purpose of this chapter was to highlight the implications of formal research on the syntax of modifiers for syllabus design and materials development. The pedagogical implications drawn from the two sets of studies briefly outlined here involve two areas of obvious relevance to language instruction: the nature of language universals and the potential role of L1 transfer. The first series of experiments examined the nature of universals in the syntax of P-modifiers and concluded that this hierarchy is naturally manifested, even when there is no evidence for it in the L1, as soon as the lexical semantics of individual modifiers has been acquired. Thus, as this aspect of grammar is incorporated into advanced materials, explicit instruction on the hierarchy is not necessary: the acquisition of syntax in this case is driven by the acquisition of vocabulary.

A work-in-progress report was also provided on a large-scale experiment targeting the syntax of attributive adjectives. Two pedagogical implications are readily apparent from the initial analysis of these results. First, the totality of the hierarchy does not appear to be gifted to L2 learners as part of the universals providing linguistic scaffolding to the L2 acquisition process. It is theoretically significant that knowledge of the nonabsolute–absolute division in the hierarchy is impressively robust, even though this distinction does not exist in all L1s and is not more readily apparent than any other distinction in the hierarchy of English adjectives. However, participants’ performance on combinations of adjectives matched for absoluteness remained at chance right up through the highest levels tested, not only for combinations of scalar physical properties, which was somewhat expected (e.g., length–depth, *long thin pencil*; depth–width, *deep wide river*), but also across more fundamental divisions in the hierarchy (e.g., opinion–age, *great new haircut*; size–age, *big old car*). Targeted instruction on these combinations is therefore necessary, although arguably through enhanced input rather than through the memorization of metalinguistic rules. To date, adjective order has been invariably taught in the form of a single, rule-oriented chapter in a grammar book, with no recycling. Perhaps the promotion of implicit learning might be more effective, by exposing learners to relevant input through systematic inclusion of binary combinations of adjectives in course materials across the syllabus. The second implication emerging from this study is that contrary to initial hypotheses, there appears to be no evidence whatsoever of L1 transfer in this domain. Thus, despite significant crosslinguistic variation in adjectival syntax, the design of teaching materials can proceed on the assumption that all learners follow the same path of development regardless of language background.

There is currently an apparent chasm between UG-inspired studies of L2 acquisition and classroom-oriented L2 research. Researchers in each tradition tend to frequent different conferences and write for different audiences. There has been a somewhat dismissive attitude toward pedagogy in generative circles, perhaps arising in part from the historical need to gain independence from schools of education and engage more fully with other disciplines such as linguistics and psychology, while in research on L2 pedagogy, there has been a growing conviction that generative approaches are irrelevant to the classroom. However, such perceptions are unfortunate, and they clearly damage the potential for researchers to engage in interdisciplinary work in applied linguistics. In this chapter, I have argued that formal research whose primary goals are not pedagogical in nature can nevertheless have interesting, practical, and direct implications for language instruction in the classroom, a conclusion brought home by all the contributions to this volume. It is to be hoped that such studies encourage awareness of the need to reconnect formal L2 research with language teaching and foster interdisciplinary understanding within a more unified field of second language studies.

Acknowledgements This research was conducted with the collaboration of colleagues in the Department of Second Language Studies as well as teachers and students in the Intensive English Program (IEP) at Indiana University. Many thanks especially to Kathleen Bardovi-Harlig, Doreen Ewert, Michelle Fleener, Marlin Howard, and Heidi Vellenga. Thanks also to Stephanie Dickinson of the Indiana Statistical Consulting Center (ISCC) for her invaluable help with data analysis. The experimental work on modifiers was coauthored with graduate students who, in their own careers, are building bridges between formal L2 research and classroom pedagogy: Beatrix Burghardt, Jung-Eun Choi, Khanyisile Dlamini, Cleyera Martin, Hyun-Kyoung Seo, and Yi-Ting Wang.

5.6 Appendix I: The Aladdin Slides

Example Slides

1. Here is Aladdin. Here is the wizard. Here is a very beautiful lamp.
2. Aladdin and the wizard are going to the cave.
3. Aladdin takes the magic lamp from the wizard.

Stimulus Slides

- | | |
|---|------------------------|
| 4. He flies right up out of the cave. | [DEG [TRAJECT]] |
| 5. He flies on through to the outside. | [FLOW [TRAJECT]] |
| 6. He flies straight on over the camels. | [DEG [FLOW]] |
| 7. He flies right on up into the clouds. | [DEG [FLOW [TRAJECT]]] |
| 8. He goes crash into the birds. | ONOMATOPOEIA |
| 9. The lamp falls right back down onto a tree. | [DEG [FLOW [TRAJECT]]] |
| 10. The lamp falls on down to the ground. | [FLOW [TRAJECT]] |
| 11. Aladdin flies right down in front of a waterfall. | [DEG [TRAJECT]] |
| 12. He flies whoosh over a lake. | ONOMATOPOEIA |

- | | |
|---|------------------|
| 13. Aladdin flies straight on under a rock. | [DEG [FLOW]] |
| 14. Aladdin flies right on across the desert. | [DEG [FLOW]] |
| 15. He flies straight through into the city. | [DEG [TRAJECT]] |
| 16. Oh no! The lamp is not in his bag! | FILLER |
| 17. Aladdin flies straight back across the desert. | [DEG [FLOW]] |
| 18. He flies right back under the rock. | [DEG [FLOW]] |
| 19. He flies back over to the waterfall. | [FLOW [TRAJECT]] |
| 20. He flies straight down behind the tree. | [DEG [TRAJECT]] |
| 21. Aladdin flies right out from behind the tree. | [DEG [TRAJECT]] |
| 22. The wizard falls splash into the lake. | ONOMATOPOEIA |
| 23. Aladdin comes straight out from behind the waterfall. | [DEG [TRAJECT]] |
| 24. He flies straight back across to the rock. | [DEG [FLOW]] |
| 25. He flies right back into the desert. | [DEG [FLOW]] |
| 26. Aladdin touches the lamp. The genie appears! | FILLER |

References

- Alexiadou, A., L. Haegeman, and M. Stavrou. 2007. *Noun Phrase in the generative perspective*. Berlin: Mouton de Gruyter.
- Azar, B.S., and S.A. Hagen. 2006a. *Basic English grammar*, 3rd ed. White Plains: Pearson.
- Azar, B.S., and S.A. Hagen. 2006b. *Basic English grammar*, 3rd ed.: Workbook. White Plains: Pearson.
- Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)* 57(1): 289–300.
- Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.
- Cinque, G. 1994. On the evidence for partial N-Movement in the Romance DP. In *Paths towards Universal Grammar: Studies in honor of Richard S. Kayne*, ed. G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, and R. Zanuttini, 85–110. Washington, DC: Georgetown University Press.
- Cinque, G. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. Oxford: Oxford University Press.
- Cinque, G. 2010. *The syntax of adjectives*. Cambridge, MA: MIT Press.
- Dulay, H., and M. Burt. 1974. Natural sequences in child language acquisition. *Language Learning* 24: 37–53.
- Emonds, J.E. 1976. *A transformational approach to English syntax*. New York: Academic.
- Fassi Fehri, A. 1999. Arabic modifying adjectives and DP structures. *Studia Linguistica* 53: 105–154.
- Gass, S.M. (with J. Behney and L. Plonsky). 2013. *Second language acquisition: An introductory course*, 4th ed. New York: Routledge.
- Gradman, H. 1971. The limitations of contrastive analysis predictions. *Proceedings of The Pacific Conference on Contrastive Linguistics and Language Universals (PCCLLU): University of Hawai'i Working Papers in Linguistics* 3/4, 11–16. Honolulu, HI: Department of Linguistics, University of Hawai'i.
- Grant, L. 2010. *Well said*, 3rd ed. Boston: Heinle.
- Hawkins, R. 2001. *Second language syntax: A generative introduction*. Oxford: Blackwell.

- Krashen, S.D. 1981. *Principles and practice in second language acquisition*. London: Prentice-Hall.
- Lado, R. 1957. *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
- Lado, R., and C.C. Fries. 1954. *English pronunciation: Exercises in sound segments, intonation and rhythm*. Ann Arbor: University of Michigan Press.
- Laenzlinger, C. 2005. French adjective ordering: Perspectives on DP-internal movement types. *Lingua* 115: 645–689.
- Maurer, J. 2000. *Focus on grammar: An advanced course for reference and practice*, 2nd ed. White Plains: Longman.
- Montrul, S. 2000. Transitivity alternations in second language acquisition: Toward a modular view of transfer. *Studies in Second Language Acquisition* 22: 229–274.
- Schachter, J. 1974. An error in error analysis. *Language Learning* 24: 205–214.
- Schwartz, B.D., and R.A. Sprouse. 1996. L2 cognitive states and the full transfer/full access model. *Second Language Research* 12: 40–72.
- Scott, G.-J. 2002. Stacked adjectival modification and the structure of nominal phrases. In *Functional structure in DP and IP: The cartography of syntactic structures*, vol. 1, ed. G. Cinque, 91–120. Oxford: Oxford University Press.
- Shlonsky, U. 2004. The form of Semitic noun phrases. *Lingua* 114: 1465–1526.
- Sproat, R., and C. Shih. 1991. The cross-linguistic distribution of adjective ordering restrictions. In *Interdisciplinary approaches to language*, ed. C. Georgopoulos and R. Ishihara, 565–593. Dordrecht: Kluwer.
- Strange, W., and V.L. Schafer. 2008. Speech perception in second language learners. In *Phonology and second language acquisition*, ed. J.G. Hansen Edwards and M.L. Zampini, 153–191. Philadelphia: John Benjamins.
- Stringer, D. 2005. *Paths in first language acquisition: Motion through space in English, French and Japanese*. Doctoral dissertation, University of Durham.
- Stringer, D. 2010. The gloss trap. In *Linguistic relativity in SLA: Thinking for speaking*, ed. Z.-H. Han and T. Cadierno, 102–124. Clevedon: Multilingual Matters.
- Stringer, D., B. Burghardt, H.-K. Seo, and Y.-T. Wang. 2011. Straight on through to Universal Grammar: Spatial modifiers in second language acquisition. *Second Language Research* 27: 289–311.
- Stringer, D., J.-E. Choi, K. Dlamini, and C. Martin. In prep. *Adjective ordering restrictions in English as a second language*. Unpublished manuscript, Indiana University.
- Thewlis, S.H. 2007. *Grammar dimensions 3: Form, meaning, and use*, 4th ed. Boston: Heinle & Heinle.

Chapter 6

The Syntax-Discourse Interface and the Interface Between Generative Theory and Pedagogical Approaches to SLA

Elena Valenzuela and Bede McCormack

6.1 Introduction

For almost 25 years, generative research on second-language acquisition has examined second-language learners' understanding of various linguistic properties such as island effects (White 1989), subadjacency constraints (Schachter 1989), case and tense (Lardiere 1998), IP (Haznedar 2001) and interface properties (White 2009). Since generative SLA research such as this typically looks at the acquisition of some property of grammar that is not explicitly taught in the classroom, little work has been done that investigates how classroom language teachers might accelerate acquisition of these structures in an instructed L2 setting.

In an attempt to address this gap between theory and practice, the current study examines the acquisition of UG-constrained properties related to the syntax/pragmatics and discourse/pragmatics interface. In particular, we examine patterns of topic-comment knowledge among two groups of learners: L1 English speakers and L1 Spanish speakers, each learning the other's language. Since topic-comment structures are primarily found in spoken language rather than written, they are less commonly taught in L2 classrooms, with the exception of Spanish clitics and clitic placement which feature regularly in SSL lessons.

By taking a balanced look at the difficulties these two groups of learners face in topicalising object nouns in their respective target languages, we hope to show

E. Valenzuela (✉)

Modern Languages and Literatures, University of Ottawa, Ottawa, ON, Canada
e-mail: evalenzu@uottawa.ca

B. McCormack

Department of Education and Language Acquisition, LaGuardia Community College,
City University of New York, Room B234RR, 31-10 Thomson Avenue,
Long Island City, NY 11101, USA
e-mail: bmccormack@lagcc.cuny.edu