Chapter 15

# Metrical Parameters and Lexical Dependency: Acquiring L2 Stress

John Archibald
*University of Calgary*

## 1. INTRODUCTION

In this chapter, I discuss the results of several empirical studies on the acquisition of metrical phonology. The data have been collected from monolingual and bilingual children, as well as from monolingual adult and bilingual adult language learners. I use these data to make more general statements about (a) which native language (L1) representations transfer into the second language (L2) grammar, (b) a principle of the Learning Theory known as Lexical Dependency and its relationship to interlanguage change, and (c) the lexical parameterization hypothesis.

In order to paint a complete picture of the acquisition process, we need to consider more than just principles and parameters, but also the learning theory necessary to account for the changing of the interim grammar. This has been convincingly argued for by Dresher (1992), Dresher and Kaye (1990), Lightfoot (1991), Newson (1990), and Saleemi (1992). In this chapter, I present an analysis of the acquisition of stress that draws on both universal principles and parameters, and a particular learning theory. Within this framework we can model the changing grammar of a language learner.

I have conducted studies that have investigated the acquisition of English metrical parameters by speakers of Spanish (Archibald, 1993b), Polish (Archibald, 1992), and Hungarian (Archibald, 1993a). I assume the parameters[1] proposed by Dresher and Kaye (1990) shown in (1):

(1) P1: The word-tree is strong on the [Left/Right].
P2: Feet are [Binary/Unbounded].
P3: Feet are built from the [Left/Right].
P4: Feet are strong on the [Left/Right].
P5: Feet are quantity-sensitive (QS) [Yes/No].
P6: Feet are QS to the [Rhyme/Nucleus].
P8A: There is an extrametrical syllable [No/Yes].
P8: It is extrametrical on the [Left/Right].

The chart in (2) shows the settings for the parameters given in (1) for Spanish, Polish, Hungarian, and English and from it one can discern the differences among them and infer the likely places of transfer:

(2)

| | *Spanish* | Polish | *Hungarian* | English |
|---|---|---|---|---|
| P1 (word tree) | right | right | left | right |
| P2 (foot type) | binary | binary | binary | binary |
| P3 (built from) | left | left | left | left |
| P4 (strong on) | right | right | left | right |
| P5 (QI/QS) | QS | QI | QS | QS |
| P6 (sensitive to) | rhyme | NA | nucleus | rhyme |
| P8 (extrametrical) | yes | no | no | yes |
| P8A (extramet. on) | right | NA | NA | right |

In Polish, P5 and P8 are major sources of transfer errors, whereas in Hungarian P1, P3, P6, and P8 are major sources. Spanish stress assignment is very much like English stress assignment, but when we look at the Spanish participants learning English we see evidence for the principle of Lexical Dependency, and for the transfer of certain L1 structures that influence stress assignment.

In order to understand Lexical Dependency, let us turn to a discussion of the L2 lexicon. There have been some studies concerned with the L2 lexicon (Carroll, 1992; Meara & Ingle, 1986), but there has been little done in terms of the nature of the phonological representation of the lexical entry. I would like to consider the nature of the phonological representation in the L2 lexical entry, and use these data to propose a general model of a learning path in second language acquisition (SLA) that is driven by the L2 lexicon.

## 2. GRAMMATICAL CHANGE

In this section I discuss some issues that I feel are relevant to a consideration of how grammars change over time. Following Lightfooot (1991), I feel that there are certain similarities between how the grammars of individuals change and how languages change. Therefore, certain aspects of diachronic

linguistics can illuminate the study of the change in an individual's grammar. I, therefore, include language acquisition in the term *language change.*

One of the questions that needs to be addressed regarding developmental stages is the pattern of change. Is it gradual or discontinuous? In other fields of language change such as historical linguistics, the assumption has been that language change is, for the most part, gradual. The dominant metaphors of these changing systems come from evolutionary theory. Some paleontologists view evolutionary change as gradual, whereas others argue that the system remains static for long periods of time and then there are catastrophes that change the system dramatically. For gradualists, discontinuities were assumed to result from gaps in the fossil (or linguistic) record. This second model of change is known as the *punctuated equilibrium* model (Gould & Eldredge, 1977).

In terms of developmental stages, then, I would argue that we can envisage two extreme paths,[2] shown in Fig. 15.1. For the Neogrammarians, sound change was across-the-board and gradual, whereas other kinds of change could be lexically based and abrupt. Sorace (1992) argued that we see both patterns in SLA in that semantically-based aspects of auxiliary selection are acquired gradually, whereas syntactically based aspects are acquired abruptly. Later, we see evidence that if we adopt a lexically based model of SLA, we should expect the second pattern. I argue this to be the case.

As in evolutionary studies, the fact that the systems can go long periods of time without changing (what is usually termed *fossilization*) should be of interest. In the words of punctuated equilibrium theorists, "stasis is data." In SLA we see stasis. Our mechanism of change has to take into account these periods of stasis. I propose that the second model describes the situation better.

### 2.1. What Changes?

In developing a theory of language change in SLA, we also need to address the question of what changes. As in Archibald (1993a), I assume that it is the

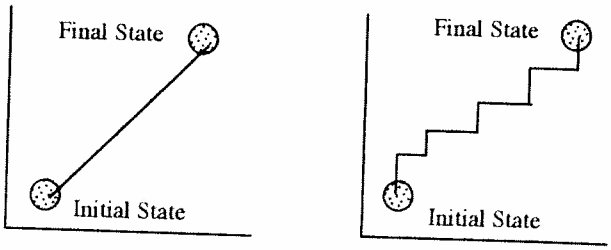



FIG. 15.1. Two paths of interlanguage development.

parameters that are being reset. A question that needs to be addressed is this: Once a parameter resets, does it affect all relevant lexical items immediately? Or does the change move through the lexicon in a way similar to the model of Lexical Diffusion in historical change (Wang, 1977)?

At first it might appear that it would be undesirable to say that the grammar is capable of representing some lexical items as having parameter setting *x*, while other lexical items of the same class have parameter setting *y*, but I would argue that as we find this situation in monolingual speakers, it should not be ruled out as a possibility for bilingual speakers. For example, monolingual English speakers have well-defined subvocabularies that operate according to different principles (e.g., Latinate vs. Non-Latinate vocabulary; nouns vs. verbs). The same claim could be made about bilingual systems: They are autonomous but modular, and the lexical entries have addresses that allow cross-linguistic activation (see Carroll, 1992). We will see that the principle of Lexical Dependency is operative in the L1. This principle of the learning theory also affects the L2. The diagram in Fig. 15.2 will show what I have in mind.

Carroll (1992) argued that in cognate vocabulary, we see L2 input activating an L1 entry. This supports the notion of linked addresses. In addition, White (1992), in discussing verb raising, explicitly argued for the L2 learner's ability to maintain both settings of a parameter until the system stabilizes (see Archibald, 1994, for a further discussion of this issue).

## 3. LEXICAL PARAMETERIZATION

Borer (1984) argued that parameterization is essentially lexical. This notion was extended by Wexler and Manzini (1987), who proposed the Lexical Parameterization hypothesis that argued there were two kinds of parame-

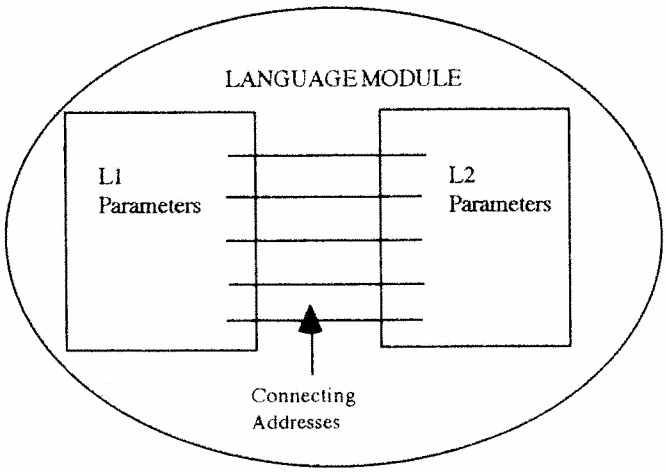



FIG. 15.2. A model of a bilingual lexicon.

ters: lexical and nonlexical. Some parameters could be said to apply to grammars as a whole (e.g., word order parameters, or bounding nodes), whereas other parameters were relevant to each lexical item (e.g., governing category). They stated that "values of a parameter are associated not with particular languages, but with particular lexical items in a language" (p. 55).

But, immediately, certain learnability problems arise from this position. How can we maintain learnability and still say that parameters need to be set for each lexical item? Safir (1987) noted that "restricting interlanguage variation to the lexicon is a very appealing step. . . . Under a properly conceived account of the lexicon, it might not be necessary to assume language learning of any other sort" (p. 78). But he also points out that the Lexical Parameterization Hypothesis (in conjunction with the Subset Principle) solves the problem of overgeneralization, but leads to a problem of undergeneralization. How does the child generalize if parameters are set for each lexical item?

I think that one way of solving this problem is via the Lexical Dependency Hypothesis proposed by Newson (1990). Newson proposed that parameter setting may well proceed through the lexicon (like Lexical Diffusion). The Lexical Dependency Hypothesis then assumes that parametric values are part of the lexical entry. Newson (1990) proposed that we should "suppose that there is a mechanism which generalizes learned information concerning a particular lexical item to all other relevant lexical items. . . . Let us refer to this learning mechanism as a 'Lexical Dependency' as the setting of the lexical parameters of certain items are, under these assumptions, dependent on those of others" (p. 179). This would be one way to reduce the learnability problem. The learner would not actually have to set the parameters for every lexical item. Rather, once certain entries had been set, that information could be generalized to other "relevant" lexical items. I put the word "relevant" in scare-quotes as it seems that in order for this to be an improvement, we need to be explicit as to what we mean by "relevant." Otherwise, we fall into the same traps as unformalized theories of induction (Mathews, 1989). We need to be explicit as to what representations they refer to.

I find this an appealing notion in that it recognizes that some aspects of language acquisition involve induction whereas other aspects involve deduction. It also seems fitting that the domain of inductive learning is the lexicon, where obviously arbitrary representations must be constructed.

As we saw before, something like Lexical Dependency appears to be operating in monolingual speakers who construct separate subvocabularies. Initially, they may assume that all lexical items are represented in the same fashion. Then, on the basis of positive (or perhaps indirect negative) evidence, they would change the representation of relevant lexical items (e.g., all verbs). If Lexical Dependency operates within a language (as Newson,

1990; Wexler & Manzini, 1987, argued) then I will assume that it is the learner's assumption for learning an L2 as well.

## 4. DEFINING RELEVANCE

Assuming the basic mental architecture as set out in Jackendoff (1987), I propose that relevance must be defined in terms of the characteristics of the lexical representations. A relevant item can only be identified in terms of the representation. Therefore, relevant could not mean everything that was [± slimy] if we were not assuming that the feature [slimy] was part of the lexical representation. So relevance must be defined in terms of such things as grammatical category, subcategorization, phonological representation, gender, and so forth. If, then, the learner's current mental representation includes something about syllable weight, or vowel tension, or gender, then that feature could be used as a basis of inductive generalization, that is, copy all of these features onto all other [+feminine] words or [+Noun] words. By allowing for this kind of copying (with reference to structure), we could account for individual variation. For example, an L2 participant who was not representing vowel tension might not set up the same word classes as a native speaker (e.g., may treat *aróma* the same as *cínema*) because vowel tension is not considered a relevant item (i.e., is not part of the representation). Another learner may behave differently as a result of having a slightly different representation, maybe one that included vowel tension, but not syllable weight. We might therefore, see the individual profiles shown in (3):

| (3) | Profile #1 | Profile #2 | Profile #3 |
|---|---|---|---|
| | aróma | ároma | ároma |
| | agénda | agénda | ágenda |
| | cínema | cínema | cínema |

As other aspects of the representation change, so does the notion of relevant. For example, if the learner begins by not distinguishing nouns from verbs, or masculine from feminine, and then later changes that representation (on the basis of positive evidence), as shown in Fig. 15.3, then suddenly there is a new class of relevant items. At Time 1 the learner could generalize to other nouns (or perhaps to other nonverbs). At Time 2 the learner could generalize to all other nouns, or all other [+feminine] nouns. This, in effect, is a retreat from overgeneralization. At Time 1 the learner's rules would be overgeneral, at Time 2 the domain of application becomes more restricted. Therefore, not directly relevant cues could have an effect on this learning path.

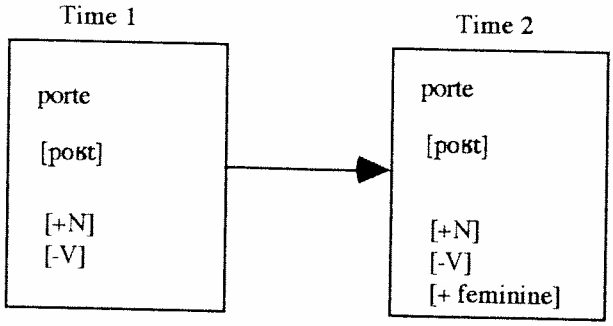



FIG. 15.3. Representational change, time, and relevance.

Note, though, that all of the grammars that could be generated by these principles and parameters would be natural languages. The learner is not generating unconstrained hypotheses. The hypotheses are constrained by the mental representations. Thus, we constrain inductive generalization. The learner can generalize but the generalizations are based on representations, not on processes as in previous views of inductive hypothesis testing.

Changes in lexical representation, then, start with a word and spread through the lexicon. This assumes that the learner is able to perceive a mismatch between input and output (i.e., that learning is error driven).

### 4.1. The Domain of Relevance

One problem that has to be sorted out is how many features *relevant* can refer to. The learner would assume that the features could be changed in all relevant:

- words
- nouns
- nouns with closed penults
- animate nouns with final long vowels

Most General Assumption
|
|
Most Specific Assumption

If we do not constrain these choices then we run the risk of generating unnatural languages. For example, a learner would have the option of treating all animate nouns with final long vowels as quantity-sensitive but all other words as quantity-insensitive, and no natural language has this option available. The question, then, is how the learner determines the domain of relevance.

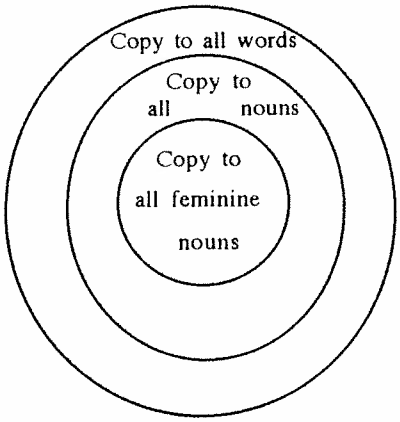



FIG. 15.4. Relevance and the subset principle.

From the above examples, it emerges that the choices are nested in a subset relation, as shown by way of example in Fig. 15.4. When copying features onto other relevant items, the learner could copy to all other lexical items, or only to items that matched in a certain number of features. We would assume that the subset value would be the unmarked choice. Then there would be positive evidence for the learner to reset. This awaits empirical confirmation from a longitudinal study.

## 5. WHERE ARE THE PARAMETERS?

I would like to consider two ways of refining the Lexical Parameterization Hypothesis in terms of the location of parameters in the grammar. Wexler and Manzini (1987) proposed that there were two types of parameters: lexical and nonlexical. Nonlexical parameters are properties of grammars as a whole whereas lexical parameters reside in the lexicon. I refer to this as the Two-Level Version. The problem with the Two-Level Version is how the learner is able to determine which parameters are lexical and which parameters are nonlexical. It is possible that this is stipulated as one of the defining characteristics of the parameter, but this does not seem viable, as a property that is lexical in one language may be nonlexical in another (stress, for example).

Another possible solution for the Two-Level Version is for the learner to assume that all parameters are nonlexical unless there is positive evidence to the contrary. There is something counterintuitive, though, about placing all parameters of variation outside the lexicon as a default value.

I think, though, that there is another interpretation consistent with these facts. What I refer to as the One-Level Version assumes that all

parameters are in the lexicon, and some of them get set as a result of a Universal Grammar (UG) property. So, for example, if a particular grammatical category is a case assigner, then all entries that contain a representation that refers to that grammatical category would get that property via the Lexical Dependency Principle.

If the learner assumes that all parameters are nonlexical and waits for positive evidence, the question remains as to how the learner would act when having to reset. Would the learner (a) reset the nonlexical parameter's value, or (b) reset the lexical–nonlexical parameter? The One-Level Version avoids this problem and is therefore preferable.

## 6. TRIGGERS

Another question that needs to be asked in relation to the lexical representation is *What evidence acts as a trigger to change the representation?* We need to be explicit as to our theory of triggers. Lightfoot (1991) addressed these issues under the heading of Degree-0 Learnability. Although I cannot see any direct implications of Degree-0 learnability for phonology, the nature of the trigger is relevant to the notion of Blame Assignment. This is a problem of matching the input and the mental representation (i.e., the learning theory). In any complex system it is a problem for the system to know what aspect of the representation to change when an error in output has been identified. Imagine that the system has 50 parameters, and the system notices that it is generating some incorrect output—how does the learner know which parameter to change? Clark (1992) proposed an interesting answer to this question. In his view of grammars competing via natural selection, all the learner has to do is change one parameter and natural selection will take care of the rest. Although potentially this is a solution to the Blame Assignment problem, I do not feel that it is a viable one due to problems in defining the Fitness Evaluation Metric. At least in the area of stress, it is meaningless to say that a grammar that had the extrametricality parameter setting wrong would be a better fit than the grammar that had the tree-construction parameter and quantity-sensitivity parameter wrong. I choose to explore the notion proposed by Dresher and Kaye (1990), and Dresher (1992) that triggers must be appropriate to particular parameters, and in this way, account for the problem of Blame Assignment.

The following are some possible relations that might guide the notion of appropriate cues. As stated before, we assume that learning is error driven. Thus, in order for the learner to make a change to the grammar, there must be a perceived mismatch between the input and the (potential) output. I would argue that the kind of mismatch would determine the

kind of change the learner would make in the grammar. I propose relations of the following type: If you find *x* where you were expecting *y*, change parameter *z*. See (4) for specific examples[3]:

| (4) Find (data) | Expecting (grammar) | Change |
|---|---|---|
| primary stress | secondary stress | word tree dominance |
| secondary stress | unstressed | unbounded → bounded |
| unstressed | secondary stress | bounded → unbounded |
| primary stress | unstressed | dominance (w → s) |
| unstressed | primary stress | dominance (s → w) |
| stress at edge | no stress at edge | direction |
| no stress at edge | stress at edge | extrametricality |
| irregular rhythm | regular rhythm | QI → QS |
| regular rhythm | irregular rhythm | Qs → QI |

Expectations, too, are governed by representations. This is a kind of deterministic learning procedure in that the actions available to the learner are dictated by the nature of the trigger. This is, in effect, arguing that indirect negative evidence may have a role to play in parameter resetting (see Carroll & Swain, 1993; Lasnik, 1990; Saleemi, 1992). The parameter will be reset on the basis of appropriate cues, and then the change will spread through the lexicon.

I now show how Lexical Dependency is relevant to L1 acquisition, and then demonstrate how it operates in SLA.

## 7. LEXICAL DEPENDENCY AND CHILD LANGUAGE

Imagine a child setting up a representational system that would account for English stress. One of the things that would have to be noted is that Nouns and Verbs behave differently when it comes to stress assignment. Children begin by assigning stress to a word regardless of grammatical category. In Archibald (1995), I argue that children are extracting the stressed syllables from adult input and mapping these syllables onto a universal trochaic template. Combined with a process of prosodic circumscription, this can account for the stress pattern and syllabic truncation of child forms. Such an analysis is consistent with data on the acquisition of Dutch stress (Fikkert, 1994), Spanish stress (Hochberg, 1988), French stress (Allen & Hawkins, 1980), and English stress (Archibald, 1995). In all of these languages, a trochaic stress pattern is clearly dominant in the child

forms, regardless of grammatical category. The data shown in (5) (from Montés Giraldo, 1976) are representative of Spanish L1 stress:

(5)

| *Adult Form* | *Child Form* |
|---|---|
| cáscara | kakála |
| estómago | tomágo |
| lámpara | pampála |
| fósforo | popúlo |
| hipopótamo | popotámo |

Although there is disagreement as to how to best explain the stress pattern of the adult forms of this type (Harris, 1991; Roca, 1990), it is clear that the child begins by treating the lexical items with a marked (or exceptional) stress pattern as if they were unmarked (or regular).

Fikkert (1994) presented the data shown in (6) that indicate the trochaic pattern in acquiring Dutch stress:

(6)

| Dutch | English Gloss | Adult Form | Child Form |
|---|---|---|---|
| kameel | camel | /kamél/ | [kíkɔmɔ́i] |
| konijn | rabbit | /konɛ́in/ | [tɔ́tɔtɛ́in] |
| kameel | camel[4] | /kamél/ | [kánəmèu] |
| ballonnen | balloons | /balɔ́nən/ | [báulowlɔ́mə] |

The data shown in (7) are from a bilingual Spanish/English child (see Archibald, 1995) and also support this trochaic bias preference:

(7) Samantha's Data (2;6)

| English | | Spanish | |
|---|---|---|---|
| Target | Form | Target[5] | Form |
| puppies | pɑ́pis | rána | yána |
| purple | pípəl | brúja | byúxa |
| yellow | yɛ́yow | manzánas | mansánas |
| Babar | bábá: | chiquíta | kikíkə |
| crocodile | kakodáyo | cára | káya |
| Cinderella | sɪnəwɛ́yə | váso | báso |
| crying | fwayɪ́ŋ | cása | yáka |
| Snow White | sów wáyt | nuéva | yéma |
| apples | ǽpows | naríz | neníy |
| outside | áwsayd | níños | níños |
| Alicia | alíysyə | | |
| Maria | miynía | | |
| Joshie | jɑ́šiy | | |
| Gary | gáriy | | |

| | |
|---|---|
| Oda | ódə |
| Maya | máyə |
| horsie | hɔ́siy |

Although English and Spanish stress systems are similar (see Archibald, 1993b), it is clear in both languages that the child is assigning penultimate stress to all forms.

It appears, then, that children begin by assigning stress to all words in the same way.[6] Later, when the child noted a mismatch between their production and perception of certain forms, they would have to consult their mental representation for a word to see what might be responsible. One of the things that the child would find there would be the grammatical category. So they might decide that Nouns and Verbs behave differently with respect to stress. Forms such as [$_N$ récord] [$_V$ recórd] might be particularly useful. The learner would have to determine what was causing the variation in stress assignment. Grammatical category would be a possible choice. Then Lexical Dependency would change the representation for all relevant items, that is, all other verbs. Dietrich (1990) argued that both for children in L1 acquisition and for adults in L2 acquisition, nouns are learned before verbs. This is supported by the production data presented here (for L1 acquisition).

Carroll's (1989) article on the acquisition of gender is also relevant here. The English speakers learning French assumed that their L2 had the same set of features as their L1 did (where L1 lacked gender). The L2 learner's initial assumption is that the L2 lexical item has all the same features as the L1 entry. Carroll argued that gender was available in UG for primary acquisition but was not triggered in secondary acquisition. This is different from parameter resetting in that the representation is lacking a feature, as shown in (8):

| (8) English | French |
|---|---|
| [+N] | [+N] |
| [−V] | [−V] |
| | [+feminine] |

As opposed to changing a feature, as shown in (9):

| (9) English | French |
|---|---|
| [+quantity-sensitive] | [−quantity-sensitive] |

This argues that L2 learners can reset lexical parameters but not trigger new structure. Archibald (1997) provides further support of this claim when looking at speakers of tone and pitch accent languages acquiring English stress.

## 8. DATA FROM NATIVE SPEAKERS

In order to make the claim that adult, nonnative speakers can refer to aspects of the lexical entry in determining relevant structures, I wanted some evidence that adult, native speakers can do the same thing. I conducted a small test (10 participants) in which the participants were presented with a list of 10 words (some of which were real but likely unknown English words, and some of which were nonce forms). They were asked to read the list out loud and the sessions were recorded. Later the stress placement was transcribed. Five of the participants were told that the words on the list were nouns,[7] and five were told that the words on the list were verbs.[8] I was interested in seeing if native speakers would treat the words significantly differently depending on whether they were treating them as nouns or verbs. I would interpret a significant difference as an indication of a native speaker's ability to consult a lexical entry and generalize feature values of known items to unknown items. Relevance is described with reference to lexical representation.

The results were arrived at in the following way. Stress placement for each item was determined. Syllable positions were encoded as follows:

Final syllable: 1
Penultimate syllable: 2
Antepenultimate syllable: 3

General tendencies of stress placement could be determined by calculating the mean score for each item in each list. That is to say, for the word *retand* in the noun list one participant assigned final stress (score = 1) and four participants assigned penultimate stress (score = 2). The mean would be $1 + 2 + 2 + 2 + 2 = 9/5 = 1.8$. On the verb list, however, all five participants assigned final stress to *retand*. The mean score would be $1 + 1 + 1 + 1 + 1 = 5/5 = 1$. This reveals a difference in the treatment of nouns and verbs for native speakers when dealing with novel forms. The patterns produced are indicated in Table 15.1. The means are shown in Table 15.2. An ANOVA revealed that there was a significant difference between the two patterns. The results were as follows: $F = 3.465$, $p = .03$. Clearly, the monolingual native speakers were treating novel nouns and verbs differently with respect to stress placement. The general pattern can also be seen in Fig. 15.5. Note that the Noun stress was always closer to the beginning of the word than the Verb stress. I would argue that all of this shows that native speakers can consult such features as grammatical category when faced with novel items. They are generalizing on the basis of their representation. I maintain that L2 learners are doing the same thing.

TABLE 15.1
Native Speaker Performance on Nouns and Verbs

| | *Noun* | | | *Verb* | | |
|---|---|---|---|---|---|---|
| | *Final* | *Penult* | *Antepenult* | *Final* | *Penult* | *Antepenult* |
| donegate | 0 | 0 | 5 | 0 | 0 | 5 |
| falloon | 4 | 1 | — | 5 | 0 | — |
| burgee | 0 | 5 | — | 3 | 2 | — |
| nidus | 0 | 5 | — | 0 | 5 | — |
| servale | 3 | 2 | — | 5 | 0 | — |
| mistoban | 0 | 0 | 5 | 1 | 0 | 4 |
| aconvent | 0 | 0 | 5 | 1 | 0 | 4 |
| indumbine | 0 | 0 | 5 | 1 | 4 | 0 |
| retand | 1 | 4 | — | 5 | 0 | — |
| distapse | 0 | 5 | — | 5 | 0 | — |

TABLE 15.2
Mean Native Speaker Scores

| *Word* | *Verb Mean* | *Noun Mean* |
|---|---|---|
| donegate[9] | 3.0 | 3.0 |
| falloon | 1.0 | 1.2 |
| burgee | 1.4 | 2.0 |
| nidus | 2.0 | 2.0 |
| servale | 1.0 | 1.4 |
| mistoban | 2.6 | 3.0 |
| aconvent | 1.8 | 3.0 |
| indumbine | 1.8 | 3.0 |
| retand | 1.0 | 1.8 |
| distapse | 1.0 | 2.0 |

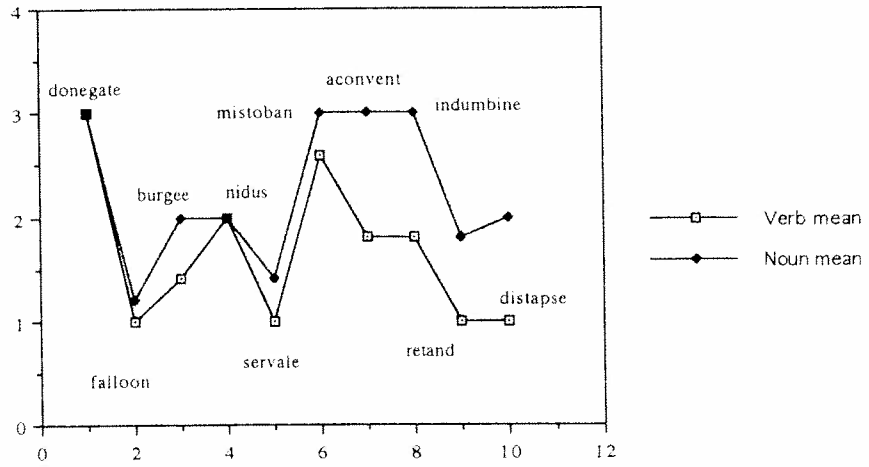



FIG. 15.5. Mean native speaker scores.

## 9. DATA FROM NON-NATIVE SPEAKERS

L2 stress has received very little attention in the literature (James, 1988; Mairs, 1989; Pater, 1991; Youssef & Mazurkewich, chap. 16, this volume). In a number of papers, I have addressed the question of L2 learners acquiring stress. My broad conclusions suggest (a) that adult interlanguages do not violate metrical universals, and (b) that adults are capable of resetting their parameters to the L2 setting. The participants were quite good at putting English stress on the right syllable.[10] Thus, their interlanguages are a combination of UG principles, correct L2 parameter settings (from resetting), and incorrect L1 parameter settings (from transfer). In the studies I have done on the acquisition of English stress by Polish, Spanish, and Hungarian speakers, a certain amount of evidence has emerged suggesting that the participants are operating in accordance with a principle of lexical dependency.

The basic research design used in my earlier studies was to have the participants perform both production and perception tasks related to stress assignment. First they had to read a list of words, and then sentences out loud (see Appendix). Stress placement was transcribed on the key words. Then the participants listened to the same words they produced as they were read out loud on a tape recorder by a native English speaker. The participants had to mark which syllable they perceived stress to be on. In both the production and the perception tasks, transfer of the L1 parameter setting into the L2 grammar was evident.

### 9.1. Spanish Participants

Archibald's (1993b) Spanish study showed that native Spanish speakers learning English transferred their diacritic extrametricality markings. I think that this can be restated in terms of Lexical Dependency. The L2 learner begins with a representation like that of the L1 (including the extrametricality markings). They then transfer this L1 representation to relevant L2 items. Spanish nonverb stress is subject to the following constraint (Harris, 1983):

> Final stress is unmarked in consonant-final words:
> Unmarked: *civíl, mercéd, altár*
> Marked: *móvil, césped, ámbar*

A mechanism that would account for this is final consonant extrametricality. Consider the word "cannibal" in Spanish, *caníbal*. The underlying representation must be [kaníbal], where the final consonant is extrametrical. Otherwise we would get the unmarked stress pattern [kanibál]. When

asked to produce this word in English, [kaníbal] was elicited in informal research sessions with Spanish participants learning English. The lexical marking of extrametricality seems to have been transferred into English. I take this as evidence that the learners can refer to a structural characteristic like [extrametrical consonant] when determining relevance.

Evidence that this is not merely the result of transfer but of some sort of lexical dependency comes to light when we look at noncognates. Spanish speakers had difficulty with a class of words in English that are unsuffixed adjectives ending in a consonant cluster like *robust, overt,* and so forth. Hayes (1980) noted that unsuffixed adjectives receive final stress if they end in a consonant cluster, otherwise they receive penultimate stress, as in (10):

(10) robúst, ovért cómmon, illícit

Hayes' explanation of this is that word-final consonants are extrametrical. Several participants made mistakes on these unsuffixed adjectives with consonant extrametricality. Why should this be the case if Spanish speakers have final consonant extrametricality too? Remember that in Spanish, the unmarked stress pattern for C-final words is final, as in (11):

(11) paréd

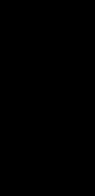
If the Spanish speakers are transferring this we would expect (12):

(12) robúst, ovért

Instead, we found people producing (13):

(13) róbust, óvert

Why? If we assume that they are treating the whole coda as extrametrical then this can be explained. If just the final segment of the rime was extrametrical then we would derive (14):

(14) [robúst]



If the whole coda is extrametrical then we derive (15):

(15) [róbust]
| \/
σ σ
s w
\ /
F

When determining the representation of English words, then, the Spanish speakers are generalizing from their L1 setting of final consonant extrametricality to an assumption that the L2 has final coda extrametricality. This is another example of a representationally based form of inductive generalization.

## 9.2. Polish Participants

Archibald (1992) showed how Polish participants learning English treated nouns and verbs differently when assigning stress. In a class of words represented by *horizon* (nouns with penultimate stress due to a tense vowel in the penult), the most common error made by Polish participants on all tasks was to stress the initial syllable (i.e., *hórizon*). However, in a class of words represented by *astonish* (verbs with penultimate stress due to a lack of a heavy syllable in the penult), the most common error pattern was to stress the final syllable (i.e., *astonísh*). I argued that the learners had determined that English nouns have final rhymes that are extrametrical (if the final vowel is lax) whereas English verbs do not. In terms of lexical dependency, this suggests that the learners are able to consult the notion of grammatical category when assigning stress; they were treating nouns and verbs differently.

Similarly, there is evidence that the participants were treating all verbs as a coherent class. The behavior of the Polish participants was consistent with a process that could be phrased as *if it's a verb, stress the final syllable.* For many of the items tested (i.e., those with a heavy, final syllable) this would yield the correct result, as shown in (16):

(16) •maintáin, appéar, eráse, decíde, achíeve
•collápse, eléct, obsérve, adápt, convínce

But for other items (i.e., those with light, final syllables) this would lead to the wrong form, as shown in (17):

(17) •astonísh, edít, cancél, considér, interprét

I take this as further evidence that the nonnative speakers are copying the representation of stress assignment to other relevant items in the lexi-

con. Again in this case the relevant class is determined by grammatical category.

We also see evidence from the Polish participants that they can use orthographic clues to determine relevance. There was strong evidence that the Polish participants were treating words ending in an *-e* in the spelling differently. The relevant words were *erase, decide, achieve, collapse, observe, convince, hurricane, baritone, antelope,* and *candidate.* Regardless of whether it resulted in the correct form, these items tended to receive final stress. As previously noted, this overlaps with the process of stressing the final syllable in verbs, and often produces the correct stress pattern. Where it doesn't work is in the class of words with secondary stresses: *húrricàne, báritòne, ántelòpe,* and *cándidàte.* For these words, the participants tended to produce final stress. I would argue that they are taking the final *-e* in the spelling to be a reliable cue for the final vowel to be tense, and therefore receive stress. This results in forms like *hurricáne, baritóne, antelópe,* and *candidáte.* It could conceivably be argued that they have not acquired the rhythm rule that retracts the stress to the first syllable on words like this, but there is one piece of evidence that makes me question that analysis. The word *matador* was also tested in this class, and it resulted in very few errors. It is the only word in this class (i.e., with a secondary stress) that does not end in an *-e* in the spelling. On this word, too, initial stress was very common, arguing that stress retraction was not problematic.

The error patterns produced by the Polish participants when dealing with English stress demonstrate that they are able to copy the representation of stress assignment parameters to other members of a relevant class.

### 9.3. Hungarian Participants

One of the characteristics of Hungarian stress assignment is that it is quantity-sensitive to the nucleus rather than to the rhyme (as in English). Although primary stress will always appear on the initial syllable, a branching nucleus (CVV) will attract a secondary stress whereas a branching rhyme (CVC) will not. I would argue that this type of quantity-sensitivity may well have an effect on the participants' placement of primary stresses in English. This is an empirical question that can be addressed by looking at the word classes with tense vowels compared to the word classes with closed syllables. If the Hungarian participants were referring to their L1 representations to make generalizations about the L2 stress assignment, we would expect them to be more accurate on words in (18i) than on the words in (18ii):

(18) (i) maintáin, appéar, eráse, decíde, achíeve (stress on CVV)
(ii) collápse, eléct, obsérve, adápt, convínce (stress on CVC)

This was, in fact, the case. The Hungarian participants had much less difficulty[11] on the final syllable when it had a long vowel in it (i) than when it was closed by a consonant (ii). I would argue that the Hungarian participants are able to base their generalizations about L2 stress on their L1 lexical representations of phonological structure. Particularly, they appear to refer to their L1 domain of quantity-sensitivity when assigning English primary stress.

Further evidence that they are generalizing comes from the behavior witnessed on the class of words that includes *hurricane, baritone, antelope, candidate,* and *matador.* If the Hungarian participants were solely transferring their L1 parameter settings, we would expect them to get these words right, as they have initial stress (like Hungarian words). However, many of the participants produced these words with final stress. I would suggest that this follows from the explanation proposed previously where the participants tend to place stress on a syllable with a branching nucleus. It is conceivable that again *matador* is exceptional in this class in that it is the only form that does not have a diphthong in the final syllable. Perhaps the Hungarian participants are then treating it as if it were in a class of words like *cinema.*

The data from the Hungarian participants, then, suggest that L2 learners are able to generalize their L1 parameter settings to relevant lexical items in the L2.

## 10. CONCLUSION

In this chapter, I have argued that the principle of Lexical Dependency is of explanatory value when describing the grammars of monolingual and bilingual children, and monolingual adults. It is therefore not surprising that it appears to be relevant to the grammars being constructed by adult L2 learners, as well.

Interlanguage grammars can be seen as a combination of (incorrect) L1 parameter settings, and the effects of UG manifested in parameters that have been reset to the (correct) L2 value. In the domain of metrical phonology we find an area rich in theoretical models that prove to be of considerable explanatory power when looking at the phonology of interlanguage gramamars.

But, in order to account for how interlanguage grammars change, we need more than just UG. We also need a learning theory. I have argued here that L2 learners' grammars are influenced by a learning theory that contains a principle of Lexical Dependency, coupled with what I have called the One-Level version of the lexical parameterization hypothesis. This suggests that when we look at L2 learners' grammars, we will see

representations that can be explained within a principles and parameters model, and patterns of change that are consistent with a punctuated equilibrium model.

Our goal is to build a theory that accounts for both representation and process in an L2 grammar, and I hope that this chapter has contributed one step toward this objective.

## ACKNOWLEDGMENTS

I thank Wayne O'Neil for his constructive feedback while preparing this chapter. The work of Elan Dresher has obviously triggered much of my own research program. The simplifications and wild extrapolations are mine own.

## APPENDIX

*Class 1*

1. The thing I love about coffee is the aroma.
2. In the summer I like to visit Manitoba.
3. The town asked for a big loan to build an arena.
4. I have never met anyone from Minnesota.
5. You can see the sun a bit above the horizon.

*Class 2*

1. We can't talk about that, it's not on the agenda.
2. The exam committee couldn't reach a consensus.
3. I'm thirty years old and I still have my appendix. 4. When it gets hot I like to sit on the veranda.
5. He didn't read the book, he just read a synopsis.

*Class 3*

1. On Saturdays I like to go to the cinema.
2. When I was in school I learned to throw the javelin.
3. It has a strong taste but I really like venison.
4. She lives in the United States of America.
5. I was trying to fix the doors on the cabinet.

*Class 4*

1. I find that position much too tiring to maintain.
2. I don't think she's as old as she might appear.
3. You can record over the songs you want to erase.
4. The committee will support whatever you decide.
5. I was amazed by what you were able to achieve.

*Class 5*

1. I really didn't think that the building would collapse.
2. The delegates were still not sure who they should elect.
3. You can't take part in the class but you're allowed to observe.
4. When I came to Canada, it was hard to adapt.
5. Don't talk to me, Bob's the person you have to convince.

*Class 6*

1. Roberta is not very easy to astonish.
2. This new manuscript is quite difficult to edit.
3. I can't come on Friday, I guess I'll have to cancel.
4. They made Tony an offer he's going to consider.
5. Some of the results were difficult to interpret.

*Class 7*

1. Edmonton was devastated by a hurricane.
2. In the opera company, Bob's the best baritone.
3. When we all went to the zoo we saw an antelope.
4. They asked me but I don't want to be the candidate.
5. My brother always wanted to be a matador.

## ENDNOTES

1. I have not included some of their parameters that are not directly relevant to the issues discussed here.
2. Robert Murray (personal communication) has pointed out that we see both patterns. Gradualism may be the norm unless there is high contact between two languages, which would result in punctuated equilibrium. The similarities between high contact and bilingualism are obvious.
3. Much more explicit claims are made in Dresher (1992). My proposals are merely illustrative of the nature of appropriate cues. Dresher is much more explicit as to the interactions of parameter settings, and the robustness of cues. The interested reader should see the original.
4. The two examples of *camel* are taken from different children.
5. The accents show the stress placement, and are not found in Spanish orthography.
6. In all of the above data, though, it must be noted that the child's forms consist almost entirely of nouns. Verbs are lacking in the data. Ultimately, we would need data from a later stage in the child's acquisition to see how verbs were stressed.
7. The target words, in this case, were preceded by the word *the*.
8. The target words, in this case, were preceded by the word *to*.
9. About 70% correct for the production task, and about 85% correct for the perception task. The tasks are described subsequently.
10. See Archibald (1993a) for a full analysis.
11. Though the stress did not change on this word, the vowel quality of the final vowel did. A common production was "the [dɑnəgət]" but "to [dɑnəgeɪt]."

## REFERENCES


Allen, G. D., & Hawkins, S. (1980). Phonological rhythm: Definition and development. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology: Vol. 1. Production* (pp. 227–256). New York: Academic Press.

Archibald, J. (1992). Transfer of L1 parameter settings: Some empirical evidence from Polish metrics. *Canadian Journal of Linguistics, 37,* 301–339.

Archibald, J. (1993a). *Language learnability and L2 phonology: The acquisition of metrical parameters.* Dordrecht: Kluwer.

Archibald, J. (1993b). The learnability of English metrical parameters by adult speakers of Spanish. *International Review of Applied Linguistics, XXXI,* 129–142.

Archibald, J. (1994). A formal model of learning L2 prosodic phonology. *Second Language Research, 10,* 215–240.

Archibald, J. (1995). The acquisition of stress. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 81–110). Mahwah, NJ: Lawrence Erlbaum Associates.

Archibald, J. (1997). The acquisition of English stress by speakers of non-accentual languages: Lexical storage versus computation of stress. *Linguistics, 35,* 1.

Borer, H. (1984). *Parametric syntax.* Dordrecht: Foris.

Carroll, S. (1989). Second language acquisition and the computational paradigm. *Language Learning, 39,* 535–594.

Carroll, S. (1992).On cognates. *Second Language Research, 8,* 93–119.

Carroll, S., & Swain, M. (1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition, 15,* 357–386.

Clark, R. (1992). The selection of syntactic knowledge. *Language Acquisition, 2,* 83–150.

Dietrich, R. (1990). Nouns and verbs in the learner's lexicon. In H. W. Dechert (Ed.), *Current trends in European second language acquisition research* (pp. 13–22). Clevedon, UK: Multilingual Matters.

Dresher, B. E. (1992). *Acquiring stress systems.* Paper presented at the DIMACS symposium, MIT.

Dresher, B. E., & Kaye, J. (1990). A computational learning model for metrical phonology. *Cognition, 34,* 137–195.

Fikkert, P. (1994). *The acquisition of prosodic structure.* Dordrecht, Holland: Institute of Generative Linguistics.

Gould, S. J., & Eldredge, N. (1977). Punctuated equilibria: The tempo and mode of evolution reconsidered. *Paleobiology, 3,* 115–151.

Harris, J. (1983). *Syllable structure and stress in Spanish: A nonlinear analysis.* Cambridge, MA: MIT Press.

Harris, J. (1991). With respect to metrical constituents in Spanish. In H. Campos & F. Martínez-Gil (Eds.), *Current studies in Spanish linguistics* (pp. 447–473). Washington, DC: Georgetown University Press.

Hayes, B. (1980). *A metrical theory of stress rules.* Unpublished doctoral dissertation, MIT.

Hochberg, J. (1988). Learning Spanish stress: Developmental and theoretical perspectives. *Language, 64,* 683–706.

Jackendoff, R. (1987). *Consciousness and the computational mind.* Cambridge, MA: MIT Press.

James, A. (1988). *The acquisition of a second language phonology.* Tübingen: Gunter Narr Verlag.

Lasnik, H. (1990). On certain substitutes for negative data. In H. Lasnik (Ed.), *Essays on restrictiveness and learnability* (pp. 184–197). Dordrecht: Kluwer.

Lightfoot, D. (1991). *How to set parameters.* Cambridge, MA: MIT Press.

Mairs, J. L. (1989). Stress assignment in interlanguage phonology: An analysis of the stress system of Spanish speakers learning English. In S. Gass & J. Schachter (Eds.), *Linguistic*

*perspectives on second language acquisition* (pp. 260–284). Cambridge, UK: Cambridge University Press.

Mathews, R. J. (1989). The plausibility of rationalism. In R. J. Mathews & W. Demopoulos (Eds.), *Learnability and linguistic theory* (pp. 51–76). Dordrecht: Kluwer.

Meara, P., & Ingle, S. (1986). The formal representation of words in an L2 speaker's lexicon. *Second Language Research, 2,* 160–171.

Montes Giraldo, J. (1976). El sistema, la norma y el aprendizaje de la lengua [The system, norm, and learning of language]. *Boletín del instituto Caro y Cuervo, 31,* 14–40.

Newson, M. (1990). Dependencies in the lexical setting of parameters: A solution to the undergeneralisation problem. In I. M. Roca (Ed.), *Logical issues in language acquisition* (pp. 177–198). Dordrecht: Foris.

Pater, J. (1993). Theory and methodology in the study of metrical parameter (re)setting. In J. Mathews & L. White (Eds.), *McGill Working Papers in Linguistics, 9,* 211–244.

Roca, I. (1990). Diachrony and synchrony in word stress. *Journal of Linguistics, 26,* 133–164.

Safir, K. (1987). Comments on Wexler & Manzini. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 77–90). Dordrecht: D. Reidel.

Saleemi, A. (1992). *Universal Grammar and language learnability.* Cambridge, UK: Cambridge University Press.

Sorace, A. (1992, August). *Acceptability hierarchies, gradual versus continuous change and permanent indeterminacy in non-native grammars.* Paper presented at Language Acquisition Research Symposium '92.

Wang, W. S.-Y. (1977). *The lexicon in phonological change.* The Hague, The Netherlands: Mouton.

Wexler, K., & Manzini, R. (1987). Parameters and learnability in binding theory. In T. Roeper & E. Williams (Eds.), *Parameter setting* (pp. 41–76). Dordrecht: D. Reidel.

White, L. (1992). On triggering data in L2 acquisition: A reply to Schwartz and Gubala-Ryzak. *Second Language Research, 8,* 120–137.

# The Generative Study of Second Language Acquisition

Edited by

Suzanne Flynn
*Massachusetts Institute of Technology*

Gita Martohardjono
*City University of New York*

Wayne O'Neil
*Massachusetts Institute of Technology*

LEA
1998 LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey London