Order of L2 Acquisition of Prosodic Prominence Patterns: Evidence from L1 Spanish/L2 English Speech

Emily Nava and Maria Luisa Zubizarreta
University of Southern California

0. Introduction

Prosodic transfer is a rich area of research for second language acquisition (SLA), encompassing multiple aspects of a speaker's production (Archibald 1997, 2006, Goad & White 2006, Gut 2003, Barry 2007, *inter alia*). In the research reported here we are particularly interested in the L2 acquisition of main sentence prominence or nuclear stress (NS), regarding how the L1 (Spanish) and the L2 (English) differ in important respects. There a few (published) studies on this aspect of SLA prosody, a notable exception being Archibald op.cit. who describes the phrasal stress of one Hungarian adult and one Polish adult learning English, finding evidence of L1 phrasal stress transfer in the speech of both L2 learners. The present research project goes beyond description, aiming to characterize the acquisition process of Spanish native speakers as they move towards native-like prosodic production of English NS placement.

This paper makes two contributions: 1) It argues that the distinct timing of the L2 acquisition of different types of prosodic patterns follow from the presence versus absence of competing L1algorithms, providing evidence on the one hand for the existence of prosodic transfer, and on the other hand, for a modular view of NS placement. 2) It presents preliminary evidence from L2 speech that the type of phrasal prominence pattern found in a given language is intricately related to the language's underlying rhythmic organization: stress-timed rhythm in the case of English and syllable-timed rhythm in the case of Spanish. More precisely, the results suggest that in order to acquire the Germanic NS algorithm, the L2 learner must have moved from a syllable-timed to a stress-timed rhythm. If this rhythm-NS connection is established for L2 speech (a connection that is entirely untaught and unconscious), it would provide strong evidence for the contention that the "language instinct" is still alive in adult SLA (Schwartz 1998).

Section 1 details how English and Spanish differ both with regards to main phrasal prominence and rhythmic organization. Section 2 describes the rhythmic differences between syllable and stress-timed languages. Section 3 presents an experiment that speaks to the timing of the L2 acquisition of NS placement. Section 4 presents a second experiment, addressing the issue of the connection between the L2 acquisition of NS and rhythm. Section 5 is a brief conclusion.

1. Main prominence at the phrasal level: Nuclear stress

In this section, we lay out the view that main phrasal prominence is determined by a grammatically-encapsulated algorithm, namely a Nuclear Stress Rule (or NSR). Furthermore, we show that different NS algorithms are at work in English and Spanish (Nava & Zubizarreta to appear).

---

† Acknowledgments: The research reported here was in part supported by NSF grant #0444088 (PI: ML Zubizarreta), a USC Provost Grant "Advancing Scholarship in the Humanities and Social Sciences (PI: ML Zubizarreta), and by a USC Undergraduate Research Grant. We thank our undergraduate research assistants, Monica Bennett, Ashley Flor, Greg Madan, and Chris Van Booven for their help with data collection and analysis. We thank H.Nam and L. Goldstein for help with voicing calculations.

1.1. Nuclear stress in English and Spanish

In every sentence there is one word in particular that receives main prominence relative to its syntagmatic cohort. Such rhythmic prominence is known as Nuclear Stress (NS). In many languages, NS serves the important function of identifying the focus domain of a sentence (Chomsky 1971, Jackendoff 1972, Ladd 1996, Zubizarreta 1998, among others):[1]

(1) That part of the sentence that is *interpreted* as focused must contain the rhythmically most prominent word.

Chomsky & Halle 1968 argued that rhythmical prominence at the phrasal level is determined on the basis of the Nuclear Stress Rule or NSR. [2] The proposed rule assigns prominence right-most in wide-focused contexts. The NSR correctly accounts for transitive and ditransitive cases like those in (2). (The location of NS is indicated with underlines.) The prosodic patterns generated by the NSR are often referred to as the "unmarked stress pattern".[3]

(2) a. Barbara eats <u>pasta</u>.
b. Barbara drew pictures on the <u>covers</u>.

However, it has been noted that Chomsky & Halles's NSR fails to account for cases of non-sentence final NS placement in English. One such case is the SV intransitives illustrated in (3) (Schmerling 1976, Selkirk 1984, Gussenhoven 1984, Sasse 1988). Another one is the case of reduced relatives, illustrated in (4) (Bolinger 1972):

(3) a. A <u>window</u> closed.
b. The <u>butter</u> is melting.
(4) a. This is too high a <u>price</u> to pay.
b. We have <u>books</u> for sale.

As noted by the above-mentioned authors, while NS frequently falls on the N(oun) in the above-mentioned structures, it may also fall on the verb:

(5) The butter is <u>melting</u>.
(6) We have books for <u>sale</u>.

Such variability in NS placement has given rise to much debate and has prompted some researchers (e.g. Bolinger op.cit. and Sasse op.cit.) to question the assumption that there exists a grammatically-encapsulated rule that generates NS placement. Bolinger 1972 argued that the relevant notion that guides placement of NS depends on relative semantic weight and (un)predictability of the predicate. Sasse 1987 argued that placement of NS depends on whether the statement is categorical or thetic.

Our own position is that NS in wide-focus contexts (so-called "unmarked" NS) is generated by grammatically-encapsulated algorithms. In some cases, such as intransitive SVs, the grammar generates two NS patterns; which of the two patterns is selected by a speaker depends on pragmatic or discourse factors. Important evidence for such a view is provided by the cross-linguistic differences

---

[1] Thus, *John broke his <u>leg</u>* (with main prominence on the object) can be interpreted as having wide focus, where the entire sentence is under assertion. It is therefore a natural answer to the question *What happened*?. On the other hand, *<u>John</u> broke his leg* cannot be so interpreted. That intonation pattern imposes narrow focus on the subject and can only function as an answer to the question *Who broke his leg*?.

[2] Such a view is known as the "Stress first" theory. An alternative view, known as the "Pitch first" theory, is put forth in Selkirk 1984, 1995 and Gussenhoven 1984. In that view, a grammatical algorithm determines the distribution of pitch accents (which is in turn sensitive to the focus structure of the sentence) and the last pitch-accented word is interpreted as rhythmically most prominent. See Ladd 1996 for an overview.

[3] Prosodic patterns with only a narrow focus interpretation, such as *<u>John</u> broke his leg* or *<u>Barbara</u> eats pasta*, are referred to as "marked stress patterns". They are generated by a late rule that shifts the unmarked NS onto the subject in order to meet the requirement imposed by the interpretative principle in (1).

between Germanic and Romance, which cannot be reduced to pragmatic or discourse factors. (For a more detailed discussion, see Nava & Zubizarreta in press.)

In Spanish, NS falls invariably sentence-finally in wide-focused contexts, in cases where there is word order variability (as in (7)), as well as in cases where there is no word order variability (as in (8)). A pragmatic-discourse based account cannot fully account for such cross-linguistic differences. See also Ladd 1996, who makes the same point.

(7) a. Se cerró una ventana. (VS) b. Una ventana se cerró. (SV)
Closed a window A window closed

(8) Tenemos libros para vender.
(We) have books for sale

We will adopt here the NS algorithm proposed in Zubizarreta 1998 and Zubizarreta & Vergnaud 2005, where it is argued that the NSR applies to *a metrically-interpreted syntactic structure* and consists of two parts: 1) one which is sensitive to selectional relations between metrical sister nodes and gives primacy to the argument (S-NSR), and 2) one which applies when the S-NSR fails to apply (C-NSR). The S-NSR is the specific rule (given its narrow context of application) and the C-NSR is the general rule.

(9) Given two sister nodes A and B, if A is a head that selects B, then B is rhythmically Strong (S-NSR) Otherwise, the right-most node is rhythmically Strong. (C-NSR).

While both the S-NSR and the C-NSR are operative in English (or more generally in Germanic), only the C-NSR is operative in Spanish (or more generally in Romance). From here on, we will refer to the S-NSR (characteristic of Germanic) as the specific NSR and to the C-NSR (active in both Germanic and Romance) as the general NSR. In the next section, we explain the source of this cross-linguistic difference and illustrate how the revised NSR works.

1.2. The source of cross-linguistic differences: the metrical (in)visibility of functional nodes

As was mentioned earlier, the NSR applies to a binary metrical structure, which is built on the basis of a syntactic structure. The cross-linguistic difference between the two sets of languages with respect to the application of the specific NSR is due to the metrical status of functional categories. Functional categories in Germanic *may* be metrically invisible, which entails that the metrical structure in Germanic may be a proper *subset* of the syntactic structure. However, in Romance (in Spanish in particular) all functional categories are systematically metrically visible; therefore the metrical structure *is* to a great extent equivalent to the syntactic structure in this language type, pace traces (which are always metrically invisible). The metrical (in)visibility of functional categories is due to the following fundamental distinction between the two sets of languages:

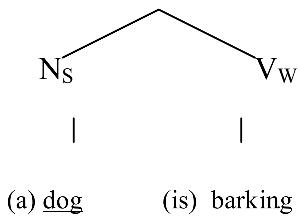
(10) In Germanic, but not in Romance, function words are generally lexically unstressed or reduced.

(11) Only unstressed or reduced function words may give rise to metrically invisible functional categories.

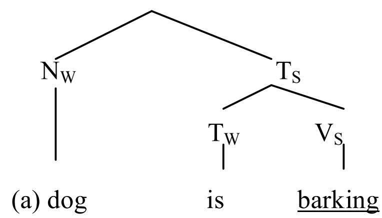
We illustrate the consequence of (10) and (11) with respect to the application of the NSR given in (9).

In English, the functional category T(ense) may or may not be analyzed as metrically invisible. If T is metrically invisible, the specific NSR will apply and designate N as rhythmically strong; see (12). If T is metrically visible, the specific NSR fails to apply; the general NSR applies instead, designating V as rhythmically strong; see (13).

(12)

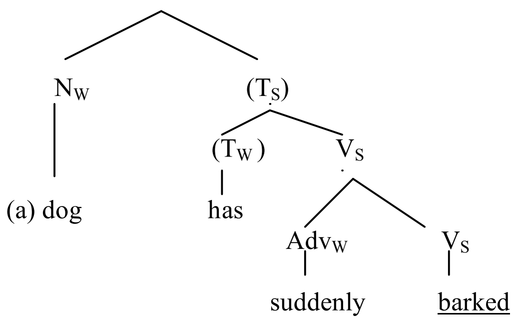


(13)



When an Adverb intervenes between the subject and the verb, the application of the specific NSR is bled (i.e. there is no metrical sisterhood between N and V). Therefore, the general NSR applies across the board, designating the last constituent as rhythmically most prominent; see (14).

(14)



The NSR stated in (9) is assumed to apply to compounds as well, an assumption that is warranted by the fact that stress placement in compounds is also sensitive to the argument/adjunct distinction (see Selkirk 1984, Cinque 1993). In the case of OV-compounds, such as (15), the specific NSR applies, assigning NS to the metrical sister node of the head of the compound:

(15) a. Mary's a pasta-eater.
b. The kids went bird-watching.

The pitch accent (PA) associated with the constituent designated as carrying the NS by the NSR is known as the nuclear pitch accent (nuclear PA) and is perceived as the most rhythmically most prominent one in the sentence (Pierrehumbert 1980, Beckman and Pierrehumbert 1986, Ladd 1980, 1996).

In section 2, we will see that the existence or inexistence of unstressed/reduced function words in a given language is intimately related to the low-level rhythmic properties of that language. But before turning to the rhythmic characteristics of Germanic and Romance, we will discuss another prosodic difference between English and Spanish, namely Anaphoric deaccenting.

1.3. Anaphoric deaccenting

A(naphoric) deaccenting refers to the process of deleting a pitch accent associated with a previously mentioned constituent (Ladd 1980, 1996), which in turn triggers a shift in prominence placement to the metrical sister node. Some examples taken from our experimental data are given in (11); italics indicate that a word has undergone A-deaccenting. In these statements, the entire sentence is focused or under assertion, as indicated by the *why*-question to which these statements function as an answer. A subpart of the focused constituent having been mentioned in the context question, A-deaccenting applies. This triggers a shift in NS placement, namely a shift of NS onto the verb in (11a) and onto the direct object in (11b).[4]

(11) a. Because I collect *stamps*. [Why are you buying that old stamp?]
b. Because I am drawing pictures *on the covers*. [Why are these pictures missing their covers?]

A-Deaccenting is independent from the NSR algorithm; it applies to the output of the NSR. This speaks to the modular nature of main prominence assignment in wide focus contexts in English: there is a part of phrasal prominence that is accomplished by the NSR, given in (9), and another part is due to a latter rule, namely NS-shift, which is a side-effect of A-Deaccenting (see note 4).

Some Romance languages, such as French (Ronat 1982, Zubizarreta 1998) and Brazilian Portuguese (Moraes 1998), have A-Deaccenting. However Spanish is among those that do not (Ladd 1996, Cruttenden 1997, Zubizarreta 1998). If we compare the English example in (11a) with its Spanish counterpart, the previously mentioned constituent still receives main sentence stress:[5]

(12) Porque colecciono sellos. [¿Por qué compras ese sello tan viejo?]

In section 3, we present an experiment that speaks to the issue of transfer and the timing of L2 acquisition of English NS and A-deaccenting by Spanish native speakers. We turn next to a discussion of the cross-linguistic rhythmic differences between stress-timed languages (like English) and syllable-timed languages (like Spanish).

2. Cross-linguistic rhythmic classification

As outlined in section 1, NS is computed on the metrical tree, which may be a subset of the syntactic tree, in particular when it contains a functional node that is interpreted as metrically invisible. Because functional categories may be metrically invisible in English but not in Spanish, we end up with distinct prosodic patterns in certain structures in these two languages. This brings us to a discussion of further differences between English and Spanish, namely with respect to low-level rhythm. We begin with a discussion of the cross-linguistic rhythmic classification and a review of background research on rhythm in L2 speech.

2.1. Distinguishing rhythm across languages

There is a long, debate-filled tradition behind the categorization of languages based on rhythmic properties. The earliest studies and subsequent proposals culminated in the use of the term "syllable-timed" to refer to those languages with isochronous syllable-durations, and "stress-timed" to those with isochronous inter-stress intervals (Pike 1945, Abercrombie 1967, among others). However, numerous subsequent studies have failed to provide acoustic evidence for strict measures of isochrony. Following Dauer's 1983 widely-accepted analysis, languages are now considered to be organized along a continuum, more or less syllable-timed or more or less stress-timed, based on language-specific phonotactic properties: vowel reduction, syllable structure inventory, and physical correlates of word-level stress. English (stress-timed) and Spanish (syllable-timed) are sharply distinguished by the properties given below:

(13) a. English has vowel reduction but Spanish does not.

[4] The shift in NS placement following A-deaccenting follows from the fact that in the languages under discussion, a non-pitch accented word is always less prominent than a word that is pitch-accented.
[5] However, it is not the case that Spanish has no other recourse in contexts of givenness than to repeat previously mentioned information. For example, in the case of the object, for which there is a pronominal clitic form, the known or previously-mentioned information may be cliticize onto the verb, in which case the V is the phrase-final element and receives main prominence. Cf. *Porque los colecciono*.

b. The syllable structure inventory of English is more varied than that of Spanish: open syllables comprise 68% of total syllable types in Spanish, as opposed to 44% in English; 60% of syllables are of the type CV in Spanish, whereas only 34% in English.
c. The difference in duration of vowels in stressed syllables compared to that of unstressed syllables is much greater in English than in Spanish: vowels in stressed syllables are 50% longer than unstressed syllables in English; in Spanish there is only a 10% difference.

Attempts to provide evidence for rhythmic classification on the basis of consonantal and vocalic interval variability have included a number of different measurement techniques, which have all yielded significant cross-linguistic rhythm distinctions across typologically different languages; cf. Ramus et al. 1999, Grabe & Low 2000, Gibbon & Gut 2001. More recent research in this area by Dellwo et al. 2007 suggests that it is more accurate to think of this cross-linguistic distinction in terms of voiceless-to-voiced ratios in speech, rather than in terms of vowel and consonant interval durations. We return to this last point in section 3.

The above-mentioned rhythmic classification has been extended to L2 speech as well. White & Mattys 2007 report on a number of studies that look at values of L2 speech (resulting from a variety of measurement techniques), all of which report values that are intermediate between the values of the L1 and the target L2. Most relevant to the current study, Carter 2005 found that the L2 English speech of Spanish/English bilinguals had nPVI -V values (normalized Pairwise Variability Index of vowels) between the low nPVI-V value for Spanish and the high nPVI-V value for English.

Gut 2003 looked at L1 speakers of English and Romance languages learning German as their L2. The study investigated L1 rhythmic influence and transfer by focusing on the difference in vowel reduction across speaker populations. Vowels reduce or delete in unaccented syllables in even more contexts in English than in German. The Romance languages included in the study (French, Italian and Romanian) do not have vowel reduction. Gut found L1 rhythmic influence in the L2 population speech: L1English speakers reduced vowels in more contexts in their L2 German as compared with the L1 German control group, and Romance native populations did not show enough reduction, showing evidence of L1 transfer via their low level of vowel reduction.

The latter study is of particular import to the current project. If *the existence of unstressed (or reduced) vowels, in particular in function words, is a hallmark of stress-timed languages*, then we would not expect to find unstressed function words in a syllable-timed language like Spanish, and the existence of unstressed function words in English would be one of its stress-timed characteristics. In section 4, we address this expectation with experimental data, and provide evidence from two separate experiments that a relationship exists between the acquisition of stress-timed rhythm and Germanic NS. In the next section, we present an experiment that speaks to the timing of L2 acquisition of English NS and A-deaccenting.

3. Ordering of Acquisition: Germanic NS and A-Deaccenting

The differences between English and Spanish with regards to phrasal prominence realization as well as with regards to low-level rhythmic organization have been outlined in the previous sections. Two separate experiments were designed to see how L2 production might differ from that of ENC production in both the phrasal and rhythmic realms. We begin with the question of phrasal prominence and transfer.

It has been shown in section 1 that both the S-NSR and the C-NSR are operative in English, while only the C-NSR is operative in Spanish. Given the existence of prosodic transfer at the rhythmic level (as mentioned above), it is reasonable to find prosodic transfer in the realm of phrasal stress as well. We therefore put forth the hypothesis in (14) and its related predictions in (15).

(14) *Phrasal Prominence Transfer (PPT) Hypothesis*
L1Spanish speakers of L2English, in particular non-high proficiency speakers, will transfer the NSR from their native language.

(15) a. L1 Spanish speakers will place NS on the verb rather than on the subject in intransitive SV structures in English.

b. L1 Spanish speakers will place NS on the verb rather than on the object in the English compound OV structures.

The hypothesis in (14) implies that the acquisition of the Germanic NSR by Spanish natives requires that the learners restructure their native NS algorithm. However, it is not necessary for speakers to restructure any algorithm in order to acquire A-Deaccenting because Spanish does not have a counterpart to this algorithm. Within a grammar-competition model (Yang 2002), transfer can be understood as a competition between L1 and L2 grammatical algorithms. In the case at hand, the L2 learner is confronted with two competing NSR algorithms: the C-NSR (which is active both in the L1 and the L2) and the S-NSR (which is active only in the L2 grammar). On the other hand, in the case of A-Deaccenting, there is no competing L1 algorithm because Spanish lacks this algorithm all together. For this reason, we expect the following timing of acquisition:

(16) Native speakers of Spanish will acquire A-deaccenting before acquiring Germanic NSR.

The following predictions result from the hypothesis in (16):

(17) a. Speakers with Germanic NS in their speech will also have A-deaccenting.
b. Speakers with A-deaccenting in their speech may or may not have Germanic NS.

In section 3.1 we discuss an experiment designed to test the above hypotheses.

3.1. Experiment 1: Question and Answer (Q&A) protocol

Participants consisted of a control group of 30, adult English Native speakers (ENC), and an L2 group of 45 adult L1Spanish/L2 English speakers. A Cloze test (Oshita 1997) was used as an independent measure of proficiency; the ENC speakers also completed the Cloze test for the purpose of comparison. ENC scores ranged from 70-75 (average 73). Twenty-six L2ers were classified as high proficiency learners (range of 66-73, average 70), and 19 L2ers were classified as intermediate (range of 58-65, average 63). Age at testing for the high proficiency L2ers ranged from 19 to 55 (average 34); age of exposure to English ranged from 3-50 (average 12). Age at testing of the intermediates ranged from 23-52 (average 33); age of exposure to English ranged from 4 to 50 (average 21). [6]

Experiment 1 was designed to elicit NS production at the phrasal level; it consisted of a scripted Q&A dialogue between the experimenter and the participant, allowing us to control for the appropriate connection between prosody and discourse factors. A variety of syntactic structures were paired with different information structure contexts: the wide-focus contexts probe for the neutral, grammar-generated patterns and contexts with previously mentioned information probe for A-deaccenting. Two lists of test items were constructed using a Latin square design. Each list had 45 target stimuli, and an equal number of fillers; the fillers balanced out both the force and form of the sentences in the dialogue.

The protocol used for testing was quite rich; here we present that subset of the data relevant to our present concerns. In particular, in the case of intransitive SV structures, we will consider here only the set of unaccusative verbs that consistently gave rise to the Germanic prosodic pattern in wide focus contexts in the case of the native English speakers. The unaccusative verbs used were *come, enter, arrive, appear, escape, vanish, broke, close, open, die*. (See Nava and Zubizarreta to appear for a more detailed discussion of the SV intransitive data.)

(18) a. SV (12 tokens), SAdvV (4 tokens), SVAdv (4 tokens)
b. OV compounds (4 tokens)
c. SVO (4 tokens); SVOPP (4 tokens)
d. SVO with previously mentioned O (4 tokens); SVOPP with previously mentioned PP (4 tokens)

[6] L2ers were tested in Los Angeles except for 9 participants tested in their native country, Paraguay. While the Paraguayan participants had spent at most two months in the US, they had attended a bilingual American-Paraguayan school from early childhood.

Participants' responses were recorded and analyzed using PitchWorks software program. Data were coded for the presence vs. absence of PA and for the location of the nuclear PA by two native speakers of English.

The results of NS production for unaccusative verbs in SV, SadvV, and SVAdv structures are given in Table 1 below; repesentative examples are given in (19) and (20).

Table 1. Unaccusative structures

Unaccusative	Prosodic Pattern [**S** v]	Prosodic Pattern [S Adv **V**]	Prosodic Pattern [S V **Adv**] [7]
ENC	97%	91%	80%
L2	23%	91%	94%

(19) SV unaccusative, wide focus
Q: What was that crashing sound?
a. A glass broke. (ENC) b. A glass broke. (L2)

(20) SAdvV / SVAdv, wide focus
Q: What happened?
a. A glass suddenly broke. / A glass broke suddenly. (ENC)
b. A glass suddenly broke. / A glass broke suddenly. (L2)

The wide focus SV structures is where we see the greatest difference between the ENC and L2 populations; the ENC group places NS on the subject 97% of the time, while the L2 group placed NS on the subject only 13% of the time, a statistically-significant difference (χ^2 = 137.45, p = .000). This difference in performance can be attributed to the ENC group employing the specific NSR (proper to Germanic), whereas the majority of L2ers used the general NSR instead, due to the activation of the competing L1grammar.

It is important to note that a significant difference is not observed in the case of the SAdvV or the SVAdv contexts. This is due to the fact that in these syntactic structures, general NSR applies across the board in both languages. In the case of the SAdvV structure, there was sentence-final NS placement at a rate of 91% for both populations. In the case of SVAdv structure, the majority of NS placement was sentence-final for both groups as well; the 14% difference between the L1 and L2 groups was not statistically significant (χ^2 = .59, p = .443).

As mentioned in section 1, the specific NSR is also operative in the case of transitive OV compounds; see example in (20). The data given in Table 2 expose a significant difference between the control and the test groups, with the ENCs producing NS on the argument 96% of the time, and L2ers doing so 43% of the time (χ^2 = 37.54, p = .000). The L2 speakers' preference for sentence-final NS again provides evidence for transfer of native NS algorithm.

Table 2. Transitive compound structures

Transitive Compound	Prosodic Pattern [S [**O** v]]
ENC	96%
L2	43%

(21) Q: Did Barbara like the Italian restaurant?
a. Oh yes. She's a pasta-eater. (ENC) b. Oh yes. She's a pasta-eater. (L2)

To conclude, the significant contrast in the SV intransitive and in the OV compound data between the ENC and the L2 group lends support to the PPT Hypothesis in (14).

We now explore the second hypothesis, put forth in (16), by looking first at cases of broad focus transitive and ditransitive contexts, then turning to transitive and ditransitive structures with previously

[7] NS placement fell on the verb in the remaining cases, with the Adv deaccented.

mentioned material. The results are summarized in Table 3 and test items that exemplify each of the four cases are given in (22)-(25).

Table 3. Wide focus transitives with & without previously mentioned (+/-PM)

Transitives	Pattern [S [v **O**]] (-PM)	Pattern [S [**V** o]] (+PM)	Pattern [S [V o **PP**]] (-PM)	Pattern [S [V **O** pp]] (+PM)
ENC	80%	82%	60%	80%
L2	99%	30%	86%	22%

(22) SVO transitive, object not previously mentioned:
Q: Do you have a hobby? A: Yes, I collect stamps.

(23) SVOPP (with no previously mentioned material within the VP):
Q: Are you finished with the coloring books? A: No, I'm drawing pictures on the covers.

(24) SVO transitive, object previously mentioned:
Q. Why are you buying that old stamp?
a. Because I collect *stamps*. (ENC) b. Because I collect stamps. (L2)

(25) SVOPP, PP previously mentioned:
Q. Why are these notebooks missing their covers?
a. Because I'm drawing pictures *on the covers*. (ENC)
b. Because I'm drawing pictures on the covers. (L2)

In cases where the answer did not contain previously mentioned material within the VP, the data speaks to the operation of the NSR, with NS being produced on the direct object in the case of the transitives and on the PP in the case of the ditransitives for both groups; see examples in (22) and (23). While the L2ers especially favored this output (see columns 2 and 4 in Table 3), the ENCs diverged from the predicted results most notably in the case of the ditransitives. This difference is mostly due to one token in particular where ENCs consistently placed NS on the object and which we attribute to pragmatic factors.[8]

We turn next to cases with previously mentioned material; see examples in (24) and (25). The data is summarized on columns 3 and 5 of Table 3, which reveals the expected differences between the ENC and the L2ers. Recall that Spanish (unlike English) does not deaccent previously mentioned material; i.e. it lacks A-deaccenting. In the case of the transitives, the ENC group deaccented the previously mentioned object 82% of the time, placing NS on the V in those cases. The L2ers deaccented at a significantly lower rate, placing NS on the object only 23% of the time (χ^2 = 29.36, p = .000). The ditransitive cases also revealed differences in prosodic patterns between the two groups. ENCs deaccented the previously-mentioned prepositional phrase, placing NS on the object 80% of the time, whereas L2ers did so in only 14% of the cases. This difference was also statistically significant (χ^2 = 36.24, p = .000).

Thus far, it has been established that some L2ers show prosodic patterns that differ substantially from the ENC groups, specifically where the S-NSR applies (intransitives, OV transitive compounds), as well as in the case of the A-Deaccenting. We now further scrutinize the behavior of those speakers who have acquired either Germanic NS or A-Deaccenting or both in order to address the timing of acquisition question put forth in (16). It was considered that a learner had acquired an L2 algorithm if (s)he produced native-like prosodic patterns in two-thirds of the relevant test items (i.e. above chance level). Table 4 crosstabulates the number of participants that have acquired the relevant prosodic patterns (namely, NS and A-deacc) according to that criterion.

[8] The context question for this token was 'Why are all the cars slowing down?', and the answer, 'Because there is ice on the road.' Participants' tendency to place NS on the object in that case may be attributed to the noteworthiness or unexpectedness of the presence of ice. See Nava & Zubizarreta in press for a detailed discussion of the data and the relevance of the pragmatic notions "unexpectedness" and "noteworthiness".

Table 4. Proficiency levels: L2 prosodic proficiency and Cloze test proficiency

NS	A-deacc	High prof.	Interm. Prof.
+	+	9	0
–	+	7	3
–	–	10	16
+	–	0	0

Only 9 high proficiency speakers have both Germanic NS and A-Deaccenting in their speech, 7 high and 3 intermediate have A-Deaccenting but do not have Germanic NS, and finally 10 high and 16 intermediate have neither Germanic NS nor A-Deacc in their speech. *Crucially, we do not find the speakers with Germanic NS in their speech but without A-Deaccenting.* Hence the predictions made by the hypothesis in (15) are borne out: A-Deaccenting is learned before Germanic NS, which supports the idea that it is easier to acquire an L2 rule that is not in competition with a rule in the L1 grammar.

A further layer of proficiency is peeled back when we look at the rate of copula reduction in the speech of both populations, picking up the discussion of metrical invisibility of functional categories (unstressed function words) mentioned in sections 1 and 2. The experiment was not designed with the question of copula reduction in mind, however contexts containing copula (22 in total) were mined to further unearth evidence of the relationship between metrical invisibility and the Germanic NS pattern. If a relationship is indeed revealed, it is expected that L2ers with Germanic NS will have copula reduction in their speech as well.

(26) a. L2 learners with Germanic NS in their speech will produce unstressed copulas at a comparable rate to that of native speakers.
b. L2 learners who have not acquired the Germanic NSR will have a lower rate of copula reduction in their speech as compared to that of native speakers.

The number of unstressed copulas was tabulated for both the control and the test groups, breaking down the L2 group by prosodic proficiency. The ENC produced 98% of reduced copulas. The ++L2 group produced 92%, the - + L2 group produced 71%, and the - - L2 group produced 67% of reduced copula. These results validate the predictions in (26). Those L2ers who have acquired Germanic NS produce unstressed copulas close to the same rate as that of the ENC (within the 90 percentile range). Conversely, speakers without an unstressed copula rate similar to that of ENC have not acquired Germanic NS. This correlation speaks to the further prerequisite for the acquisition of the Germanic NSR: speakers must acquire fundamental phonotactic characteristics (for instance, vowel reduction as in the case of unstressed copulas) that underwrite English-like rhythm. We expand on this last point in the following section.

4. The Nuclear Stress-Rhythm connection

Recall that metrical invisibility of functional categories in English is crucially cued by reduced function words in these languages. Furthermore, reduced function words are one particular instantiation of the more general characteristic of stress-timed languages. On the other hand, the systematic metrical visibility of functional categories in Spanish is cued by the non-reduced nature of function words in this language, which stems from the general phonotactics characteristic of syllable-timed languages. Under this view, functional categories and the prosodic status of the words that realize them constitute the lynchpin of the connection between nuclear stress pattern and the low level rhythm of a language. This leads us to postulate the following hypothesis:

(27) *The Nuclear Stress-rhythm connection*:
Non-sentence final prominence in wide-focus contexts is found in stress-timed languages only.

Hypothesis (27) makes an important prediction with respect to the English speech of Spanish natives:

(28) Speakers that have acquired the Germanic NS rule have approximated English rhythm. Speakers that have not acquired the Germanic NS rule have not approximated English rhythm.

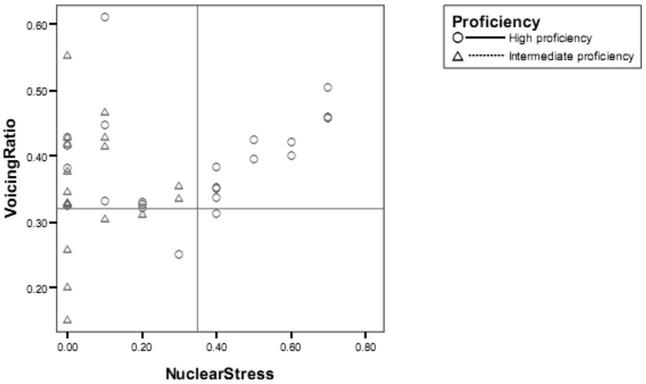
In order to test (28), a second experiment was designed to gauge the low-level rhythm of the same subjects that participated in the first experiment.

4.1. Experiment 2: Voicing ratio

ENC participants were asked to read a passage known as "The North Wind and the Sun", after the Q&A experiment.[9] L2 participants were asked to read the passage in English first and then the passage in Spanish ("El viento norte y el sol").[10] In addition, 20 monolingual Spanish speakers were recorded in Spain for the sake of comparison[11]. Following Dellwo et al. 2007, we made use of voicing parameters to classify languages according to rhythm. The software vocDetect (as modified by Hosung Nam at Haskins Laboratory) was used to find the voiced and voiceless sequences in the speech stream.[12] A function was developed in MatLab to calculate the voiceless to voiced ratio (Goldstein & Nava). When run through MatLab, a value was given for the ratio of voiceless intervals over voiced intervals for each separate sentence of the passage. An average was then taken across sentences to obtain a value for each participant; the average across participant values yielded group values.

The voicing ratio (or VR) for each group are as follows (on a scale of 0 to 1): .51 for the ENC, .37 fro L2 English, .28 for L1 Spanish, and .17 for Monolingual Spanish. These results show a successful replication of the previous findings regarding L1 rhythmic classification, as well as research that finds L2 rhythm values between those of the target language and the L1 (see 2.1).[13] Using a one-way analysis of variance (ANOVA) a highly significant difference was found to exist among the voicing ratio values of the ENC, the L1Spanish of the L2ers, and the Spanish of the Monolingual Spanish speakers ($F(2,79) = 50.94$; $p < .0000$). A one-way ANOVA analysis also revealed a significant difference between the ENC and L2 English group: ($F(3,59) = 5.83$; $p = .0015$).

We turn next to the hypothesis in (25) and its related predictions in (26). To this end, we plotted the L2 English participants' data, with VR values on the y-axis and NS values on the x-axis as shown in the figure below. The lines mark the midpoint of the L2ers' data range for both value sets.



The learners to the right of the vertical line are those that have produced 50% or more of Germanic NS patterns. Learners that are above the horizontal line are those that may be considered to have move towards an English type rhythm. The expectation is that learners that are found to the right

[9] The North Wind and the Sun (and its counterpart in other languages) is a passage often used in phonetic studies because it is a phonotactically-well balanced passage (for the language in question).

[10] 45 L1Spanish/L2English participants comprise the L2 group, however the voicing ratio values given here reflect the values of only 41 participants due to one missing file and three corrupted sound files.

[11] Many thanks to Ben Parrell for recording this group.

[12] http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html

[13] The difference between the L1Spanish and the monolingual Spanish groups seem to indicate that the rhythm of the L2 also has an effect on the rhythm of the L1. We will return to this issue in future research.

of the vertical line will also be above the horizontal line. The expectation is fulfilled, with the exception of one learner located just below the horizontal line.[14] On the other hand, if we look at the upper left quadrant of the above plot, we see that there are learners that have English-like VR but no or very little Germanic NS in their speech. This suggests that acquisition of English rhythm is a prerequisite for the acquisition of Germanic NS, but not vice-versa: acquisition of Germanic NS is not a prerequisite for the acquisition of VR. Indeed, L2 rhythm seems to be acquired before the Germanic NS rule by Spanish natives. Of course, these observations are incomplete without concomitant regression analyses. In Nava & Zubizarreta in preparation we report such analyses and a more detailed discussion of the data. In that work, we also show that, as expected, regression analyses on VR values and percentage of A-deaccenting in the L2 speech of Spanish natives reveal that there is no connection between rhythm and A-deaccenting.

5. Conclusion

The current project provides evidence of phrasal stress transfer in the speech of L1Spanish/L2English learners. Said transfer is attributed to an additional algorithm for NS placement that is operative in English but not in Spanish, namely the specific NSR. It was also shown that L2 learners acquire A-Deaccenting before Germanic NS, in support of the modular view of NS placement.

Furthermore, we put forth the hypothesis that in the case of Spanish natives learning English, the acquisition of the low-level L2 rhythm is a prerequisite to acquire the Germanic NS rule. We presented data that lends preliminary support for this hypothesis, which in turn lends support to the contention that the innate language instinct is still active in second language acquisition, although the relevance of the age factor in the L2 acquisition of prosody still needs to be investigated (see note 14).

Finally, we note that the research reported here constitutes one more example that SLA provides excellent testing grounds to examine the mental representation of native grammars through the lens of the L2 interlanguage. If the rhythm-NS connection is correct, as suggested by the evidence based on L2 speech, it provides evidence in favor of a theory of NS that can establish such a connection in the first place.

References

Archibald, John. 1997. The acquisition of L2 phrasal stress. In M. Young-Scholten and S.J. Hannah, eds. *Focus on Phonological Acquisition*. John Benjamins.

- 2006. Phonology of L2 acquisition In R. Wiese, ed. *The Encyclopedia of Language & Linguistics,* 2nd Edition. Elsevier.

Barry, William. 2007. Rhythm as an L2 problem: How prosodic is it? In Jurgen Trouvain and Ulrike Gut (eds.), *NonNative Prosody. Phonetic Description and Teaching Practice*. Berlin, New York: Mouton de Gruyter 2007.

Bolinger, Dwight. 1972. Accent is Predictable (if You are a Mind-Reader). *Language* 48, 633-644.

Carter, Paul. 2005. Quantifying rhythmic differences between Spanish, English, and Hispanic English. In R. S. Gess, & E. J. Rubin (Eds.), *Theoretical and experimental approaches to romance linguistics: Selected papers from the 34th linguistic symposium on romance languages* (Current issues in linguistic theory 272) Amsterdam, Philadelphia: John Benjamins , 63–75.

Chomsky, Noam and Morris Halle. 1968. *The sound pattern of English*. New York: Harper Row.

Cruttenden, Alan. 1997. *Intonation*. Cambridge University Press.

Dauer, Rebecca. 1983. Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11, 51-62.

Dellwo,Volka, Fourcin, Adrian and Evelyn Abberton. 2007. Rhythmical classification based on voice parameters. *International Conference of Phonetic Sciences* (ICPhS).

Goad, Heather and Lydia White. 2006. Ultimate attainment in interlanguage grammars: a prosodic approach. *Second Language Research* 22, 3 243-268.

Gibbon, Dafydd. and Ulrike Gut. 2001. Measuring Speech Rhythm. In: *Proceedings of Eurospeech*, Aalborg, 91-94.

Gut, Ulrike. 2003. Prosody in second language speech production: the role of the native language. *Fremdsprachen*

[14] The L2 participants in the upper right quadrant had a Cloze test score that ranged from 67 to 75 (average 72) and age of exposure that ranged from 3 to 15 (average 9), suggesting that age of exposure might very well be a significant factor. We intend to investigate this issue further in future work.

Lehren und Lernen 32, 133-152.
Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents*. Dordrecht: Foris.
-2004. *The phonology of tone and intonation*. Cambridge, UK; New York: Cambridge University Press.
Hualde, Jose Ignacio. 2007. Stress removal and stress addition in Spanish, *Journal of Portuguese Linguistics* 6, 58-89.
Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
Flege, James. 1995. Second language speech learning theory, findings and problems. In Strange, W. (ed.), *Speech perception and linguistic experience: Issues in cross-linguistic research*. Timonium: York Press, 233 - 277.
Ladd, Robert. 1980. *The Structure of Intonational Meaning: Evidence from English*. Indiana University Press, Bloomington.
- 1996. *Intonational Phonology*. Cambridge University Press.
Low, Ee Ling, Grabe, Esther, and Francis Nolan. 2000. Quantitative characterizations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, 43, 377-401.
Nava, Emily and Maria Luisa Zubizarreta. In press. Deconstructing the Nuclear Stress Algorithm: Evidence from Second Language Speech. In *The Sound Patterns of Syntax*, N. Erteschik-Shir & L. Rochman, eds. OUP.
- In preparation. *The Nuclear Stress-Rhythm Connection. Evidence from Second Language Speech.*
Ramus, Franck, Marina Nespor, and Jacques Mehler. 1999. Correlates of linguistic rhythm in the speech signal. *Cognition* 73, 265-292.
Ronat, Mitsou. 1982. Logical Form and discourse islands. *Journal of Linguistic Research* 2, 33-48.
Reinhart, Tanya. 2006. *Interface strategies: optimal and costly computation*. Cambridge, Mass.: MIT Press.
Sasse, Hans-Jurgen. 1987. The thetic/categorical distinction revisited. *Linguistics* 25, 511-580.
Schwartz, Bonnie. 1998. The Second Language Instinct. *Lingua* 106, 133-160.
Schmerling, Susan. 1976. *Aspects of English sentence stress*. Austin: University of Texas Press.
Selkirk, Elizabeth. 1984. *Phonology and Syntax. The Relation between Sound and Structure*. MIT Press.
-1995. Sentence Prosody: Intonation, Stress, and Phrasing. In J.A. Goldsmith (ed.), *The Handbook of Phonological Theory*, Blackwell, Cambridge, 550 –569.
Sosa, Juan Manuel. 1999. *La entonación del español*. Madrid: Ediciones Cátedra.
Trouvain, Jurgen and Ulrike Gut. 2007. *Non-Native Prosody. Phonetic Description and Teaching Practice*, Mouton de Gruyter, Berlin/New York.
White, Laurence and Sven L. Mattys. 2007. Calibrating rhythm: First language and second language studies. *Journal of Phonetics* 35: 501-522.
Yang, Charles. 2002. *Knowledge and Learning in Natural Language*. Oxford University Press.
Zubizarreta, Maria Luisa. 1998. *Prosody, focus, and word order*. Cambridge, Mass.: MIT Press.
Zubizarreta, Maria Luisa and Jean Roger Vergnaud. 2005. Phrasal Stress, Focus, and Syntax. In M. Everaert and H. van Riemsdijk, eds., *The Syntax Companion*. Blackwell.

Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)

edited by Jean Crawford, Koichi Otaki, and Masahiko Takahashi

Cascadilla Proceedings Project   Somerville, MA   2009

Copyright information

ISBN 978-1-57473-436-2 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Nava, Emily and Maria Luisa Zubizarreta. 2009. Order of L2 Acquisition of Prosodic Prominence Patterns: Evidence from L1 Spanish/L2 English Speech. In *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America (GALANA 2008)*, ed. Jean Crawford, Koichi Otaki, and Masahiko Takahashi, 175-187. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #2318.