Processing Determinism[*]

William O'Grady

I propose that the course of development in first and second language acquisition is shaped by two types of processing pressures—internal efficiency-related factors relevant to easing the burden on working memory, and external input-related factors such as frequency of occurrence. In an attempt to document the role of internal factors, I consider a series of case studies involving contrasts that are rarely instantiated in the input, yet show early mastery. I conclude with some general remarks about the nature of development and the possibility that it unfolds more uniformly than suggested by most previous work.

1. Introduction

Proficiency in a language, we all agree, is the product of development. But this only raises the deeper question, on which there is no agreement, of how and why language develops in the way that it does.

The standard way to approach this question is to examine particular features of language with a view to identifying the forces that shape their emergence. This is also the strategy that I adopt here. I begin, in the next section, by briefly outlining three approaches to the study of linguistic development, including the processing-based approach that I favor. Sections 3 and 4 focus on a series of case studies involving first and second language acquisition that appear to confirm the importance of processing pressures to the architecture of development. Section 5 considers the possibility that development, so shaped, is uniform, contrary to first appearances. Section 6 presents a brief conclusion.

The overarching framework within which I will lay out my proposal is emergentist in character. That is, I assume that the properties of language, and the manner in which those properties are acquired, reflect the interaction of nonlinguistic forces and propensities. On this view, what we think of as development is not the product of a specialized language acquisition device; rather its course is defined less directly, by events and influences that I will attempt to outline here, at least in a preliminary way.

---

2. Development

For the purposes of discussion, I will treat development as the series of changes that bring a learner's language in closer alignment with the language to which he or she is exposed. Signs of development pervade the language acquisition process. The syllables produced by children acquiring English are initially open, then permit voiceless obstruents in coda position, and finally allow voiced obstruents in that position (Ingram 1976): [dɑ] > [dɑk] > [dɑg] 'dog'. Sentential negation is initially expressed just by *no*, then by a preverbal negative, and finally by a negative that co-occurs with an auxiliary verb (Déprez & Pierce 1993): *no > I no(t) go > I won't go*. The aspectual suffix *-ing* is acquired before the definite article *the* (Brown 1973:274): *-ing > the*. And so on.

Two types of developmental progression are often distinguished in the literature. The first, involving what are sometimes called 'developmental sequences,' is manifested in the step-wise emergence of a single phenomenon, such as the ability to articulate syllables or to express negation, as just noted. The second type of progression, commonly referred to as 'developmental order,' is seen in the emergence of two or more phenomena relative to each other. The early mastery of *-ing* compared to *the*, also alluded to above, is a classic example. I focus here on the study of developmental sequences, which (I believe) offer the more promising opportunity to uncover the forces that underlie the emergence of language. However, I will briefly consider the problem of developmental order in section 5.

2.1 Theories of development

There is very substantial disagreement on how and why language unfolds in the particular way that it does. Several quite different approaches can be discerned.

A number of scholars have staked out positions that place a maximum emphasis on the learner's sensitivity to language use by others. Often labeled the 'usage-based theory of learning' (Tomasello 2003, MacWhinney 2004, Ambridge & Lieven 2015), this approach holds that learning arises through 'language use itself, by means of social skills like joint attention, and by means of powerful generalization mechanisms' that are highly sensitive to experience (Behrens 2009:383). Soderstrom et al. (2009:409) go so far as to call this perspective 'the new paradigm of language acquisition.'

At the other extreme lies classic work in Universal Grammar (Chomsky 1981, Crain & Thornton 1998, Roeper 2007, Snyder & Lillo-Martin 2011), which holds that the human language faculty is built around an inborn system of linguistic categories and principles. This approach attributes significant aspects of development to the interaction of experience with innately given constraints and parametric options (e.g., Baker 2001). Although some early work in this area raised the possibility that parameter setting might require very minimal exposure to input, more recent proposals leave room for gradual changes that may require

large amounts of experience (e.g., Yang 2004, Chomsky 2008:7, Sharwood Smith & Truscott 2014).[1]

A third approach, which I favor, assigns a significantly smaller role to the input than is common in usage-based theories, and no role at all to Universal Grammar. Instead, the developmental profile of language is taken to be largely the product of processing pressures.

(1)   Processing Determinism
       The course of development is shaped by processing pressures.
       (less costly > more costly)

Ideas along these lines have been put forward in various forms over the years (e.g., O'Grady 2005; Pienemann 2005, this issue; VanPatten 2012; Filipović & Hawkins 2013; among others.) Space does not permit a survey of this work, or a comparison among the various approaches. Suffice it to say that the differences appear to lie largely in two considerations: the type of factors used to calculate processing cost, and the type of data used to test the validity of those calculations. I outline my position on these two matters next.

2.2 An emergentist approach to processing determinism

In the version of processing determinism that I propose, two quite different types of factors create the pressures that shape the course of development. On the one hand, there are external factors that arise from the environment, including the composition and quantity of the input to which learners are exposed. Perhaps the best documented of these factors is frequency of occurrence: all other things being equal, a more frequent form is acquired before a less frequent counterpart (Jurafsky 2003, Menn 2000, Ferreira 2003); for a general review, see Jaeger & Tily (2011) and Ambridge et al. (to appear).[2] There is an apparent straightforward explanation for this correlation: the more frequently a form is heard and used, the stronger and more accessible the corresponding processing routine becomes (e.g., Paradis 2004:28).

On the other hand, there are internal pressures that stem from the burden that particular operations place on working memory, quite independently of frequency of occurrence. Such effects are not always easy to discern, however, because of a confound involving rate of occurrence—low-cost processing routines (and the associated linguistic patterns) are likely to be used more frequently. In order to

---

[1] A further possibility, put forward by Wexler (1990), among others, holds that the availability of parameters is controlled by a maturational timetable, so that relevant input, even if ample, may initially have no impact.

[2] The effects of frequency can be modulated in various ways. For instance, in the Competition Model (e.g., MacWhinney 1987), the key factor is the reliability of particular cues rather than their simple rate of occurrence. N. Ellis (2006) discusses various other ways in which frequency effects are modified and mitigated by a range of factors.

ensure that processing effects are not just frequency effects, it is necessary to investigate linguistic phenomena whose emergence either runs counter to input frequency or occurs in the absence of rich input. I return to this matter in the next two sections.

In contrast to much of the contemporary literature on development, I focus my discussion on comprehension rather than production, for two reasons. First, as noted above, in order to neutralize the effect of experience, it is vital to consider phenomena that are rarely instantiated in the input. An insistence on production data in such cases would actually subvert the study of development.

Second, development is possible in the absence of production, as the study of mutism reveals (Lenneberg 1962, Hickok 2009). In contrast, language cannot be acquired without the ability to comprehend at least part of the input to which one is exposed, and comprehension very routinely precedes production (Hsi 2014). An understanding of how the ability to comprehend emerges is therefore absolutely central to the investigation of development.

I outline the essentials of my approach in the next section with the help of brief case study of English reflexive pronouns, drawing on the much more detailed exposition offered in O'Grady (2015). Section 4 presents three additional case studies that further illustrate the advantages of a processing-based theory of development.

3. Pronoun interpretation

It is impossible to propose a theory of development without also taking a position on the type of system that it produces. Traditionally, that system is assumed to include a grammar that determines whether sentences are well-formed, independent of the processing routines that are used in the course of speech and comprehension. I reject this idea in favor of the view that processing routines do the work traditionally assigned to grammatical rules (e.g., O'Grady 2005, 2015) and must therefore be the locus for the study of development. The example of English reflexive pronouns helps illustrate this point.

It has long been understood that prototypical reflexive pronouns in languages such as English seek out a 'co-argument antecedent' (Jespersen 1933:111). Thus, the reflexive pronoun in (2), the direct object of *scratch*, can only refer to Mickey, the subject of the same verb. It cannot refer to Donald or to Donald's friend, let alone to someone not mentioned in the sentence.

(2)   Donald's friend said that [**Mickey** scratched **himself**].
                          ↑_____|

Elsewhere (e.g., O'Grady 2005, 2015), I have proposed that this interpretive constraint is the product of processing pressure rather than a principle of Universal Grammar, as commonly proposed. The key idea is simply that

processing cost is minimized if the reflexive pronoun is interpreted immediately and locally—taking its co-argument *Mickey* as antecedent.

A different processing routine underlies the interpretation of plain pronouns such as *him* and *her*. Because plain pronouns have a crucial role in linking sentences to situations, the search for their antecedent focuses on referents that are prominent in the discourse, without regard for locality (Song & Fisher 2007, Foraker & McElree 2007, among many others).

(3)   *Mickey scratched **him***

As frequently observed (e.g., Reinhart 2006:181ff, Sekerina et al. 2004, and Reuland 2011:127), a search through discourse for a pronoun's antecedent is likely to incur processing costs beyond those required for the immediate and local interpretation of a reflexive pronoun.

Turning now to development, we find an extremely suggestive fact: whereas children seem to master the interpretation of reflexive pronouns by age three, some show a tendency to misinterpret *him* to mean *himself* in the sort of situation depicted in Figure 1 (Chien & Wexler 1990, Conroy et al. 2009, Love et al. 2009, Van Rij et al. 2010, Bergmann et al. 2012).

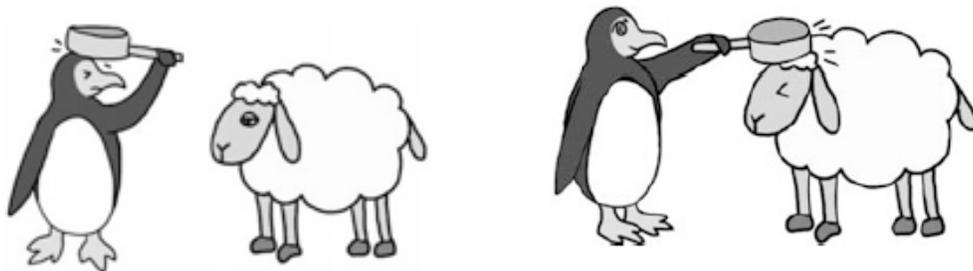Look, a penguin and a sheep are on the sidewalk.

FIGURE 1. Sample picture from Van Rij et al. (2010: 749)

Some children respond 'yes' when the sentence *The penguin is hitting him* is matched with the picture on the left—a sign that they interpret *him* to mean *himself*. Importantly, there is no comparable tendency to err when *The penguin is hitting himself* is matched with the picture on the right, which children correctly reject.

Reflexive pronouns and plain pronouns thus manifest different developmental profiles, one characterized by early mastery and the other by the potential for early confusion. Why should this be?

Table 1. Developmental sequences for reflexive and plain pronouns

Reflexive Pronouns	Plain Pronouns
1. Correct interpretation	1. Possible misinterpretation as a reflexive
	2. Correct interpretation

Why should this be?

Frequency of occurrence offers no explanation for the observed asymmetry: despite their susceptibility to misinterpretation, plain pronouns occur far more commonly in caregiver input than do their reflexive counterparts, as illustrated by the results of a search that I conducted of speech to Adam, Eve and Sarah in the CHILDES data base.

Table 2. Instances of pronoun types in maternal speech to Adam, Eve & Sarah

	Reflexive Pronouns	Plain pronouns
3rd masc sg	15 (*himself*)	292 (*him*)
3rd fem sg	1 (*herself*)	1079 (*her*)
3rd pl	1 (*themselves*)	465 (*them*)
Total	17	1836

There is however a straightforward processing explanation for children's difficulty with plain pronouns, even after exposure to ample input. As noted above, the interpretation of plain pronouns calls for wider-ranging and potentially more costly processing activity than is required in the case of reflexive pronouns. Early errors on plain pronouns reflect a propensity to avoid those additional demands, consistent with the frequent observation that children have more difficulty than adults weighing and integrating information from discourse and the ambient situation (e.g., Clackson et al. 2011:140 and the references cited there). Put simply, children interpret plain pronouns as if they were reflexive pronouns in order to minimize processing cost.[3]

The emergence of constraints on pronoun interpretation is an instructive example for two reasons. First, it illustrates how processing pressure can shape development: the routine for interpreting reflexive pronouns is acquired first because it is less costly. Second, the developmental lag associated with plain pronouns demonstrates that internal processing pressures can trump frequency of occurrence: reflexive pronouns occur far less frequently than plain pronouns, but are mastered first.

---

[3] Interestingly, we appear not to find comparable errors in production (Bergmann et al. 2012)—for a good reason: because the referent is already known to the speaker, processing cost is not reduced by encoding it with the wrong type of pronoun, regardless of the length of the referential dependency.

4. Additional effects of internal processing pressures

Another way to probe the factors relevant to development is to examine cases of language acquisition for which experience offers little or no relevant input, leaving the mechanisms of development to operate essentially on their own, guided primarily by internal processing pressures. I will consider three apparent examples of this situation, one from first language acquisition and the other two from second language acquisition.

4.1 Scope preferences in first language learners of English and Korean

A signature feature of natural language is its capacity to express negation and quantification, as well as their interaction with each other ('scope'). The sentence in (4) offers a simple example.

(4)   Donald didn't send all the letters.

The dominant interpretation of this sentence is that Donald sent only some of the letters (the *not > all* reading, in which the negative is said to have wide scope), not that every letter remained unsent (the *all > not* reading, in which quantifier has wide scope). Interestingly, the opposite is true in Korean, where the preferred interpretation is that none of the letters were sent. (Subj = subject marker; DO = direct object marker; Pst = past tense; Decl = declarative)

(5)   Donald-ka motun phyenci-lul an ponay-ss-ta.
        Donald-Subj all letter -DO not send-Pst-Decl
        'Donald sent none of the letters.' (All the letters were unsent.)

The contrast between the two languages is partly due to the fact that the negative precedes the direct object in English (an SVO language), but follows it in Korean (an SOV language).

(6)a.   English:
        Donald did**n't** send **all** the letters.
                Neg   Quantifier
    b. Korean:
        Donald **all** the books **not** send.
            Quantifier   Neg

Because the negative precedes the quantifier in English, the processor encounters it first and registers its presence in the sentence.

(7)   The processor encounters the negative:
        Donald **didn't** send
                Neg

Upon subsequently encountering the quantified NP, the processor has two options, neither of which has an obvious cost advantage over the other.

On the one hand, the processor can allows the negative to scope over the just-encountered quantifier (the *not > all* interpretation). For the sake of illustration, assume a situation in which there are three letters, just two of which were sent.[4]

(8)a.   The negative influences the interpretation of the quantifier (*not > all*):
        Donald did**n't** send all the letters
                ✉ ✉ | ✉

On the other hand, the processor could just as easily interpret the quantified NP independently of the negative, yielding the reading in which what was not sent corresponds to the full set of letters.

(8)b.   The negative doesn't influence the interpretation of the quantifier (*all > not*):
        Donald did**n't** send all the letters
                ✉ ✉ ✉

Interestingly, English-speaking children allow both interpretations (Musolino, Crain, & Thornton, 2000; Musolino & Lidz, 2006). The strong preference for the *not > all* interpretation that is characteristic of adult speakers emerges much later.[5]

Now let us consider the Korean version of the *not-all* pattern exemplified in (6b). The two key steps in the interpretation of this pattern can be schematically depicted as follows.

(9)a.   The processor encounters the quantified NP:
        Donald all the letters ....
                ✉ ✉ ✉

    b. The processor encounters the negative:
        Donald all the letters **not** send
                ✉ ✉ ✉

---

[4] This is of course just one of several scenarios compatible with wide scope for the negative.

[5] This later development may well be driven by experience. One possibility, suggested by Musolino and Lidz (2006, p. 834), assigns a role to input: Over a period of years, English speakers note that the *all > not* interpretation is usually expressed by patterns such as *Donald didn't send any letters* or *Donald sent no letters*, inviting the inference that *Donald didn't send all the letters* is reserved for the *not > all* interpretation.

As illustrated here, the quantified NP is encountered and interpreted (step a) well before the appearance of the negated verb. This leaves just one easy option at the point at which the processor comes upon the negative: the *all > not* interpretation, as shown in (9b). Derivation of the *not > all* option requires revision of that initial interpretation, as depicted below.

(10)   Donald all the letters not send
         ✉ ✉ ✉   →   ✉ ✉ | ✉

Given the additional cost associated with this interpretation, it is not surprising that Korean speakers strongly prefer the *all > not* reading and that this preference is manifested even in very young learners (Han, Lidz, & Musolino, 2007; O'Grady, 2013).

Table 3 summarizes the developmental profile for scopal preferences in English and Korean.

Table 3. The developmental sequences for scope in L1 English and L1 Korean

English	Korean
1. both *not > all* & *all > not* accepted	1. strong preference for *all > not*
2. strong preference for *not > all*	2. no change

It is unlikely that the initial step in either developmental progression is shaped by the input (although the second step in English may well be; see note 5). My search of the CHILDES corpora for Adam, Eve and Sarah uncovered just 8 sentences (all from speech to Adam) in which a negated verb takes a universally quantified NP as direct object. The CHILDES database currently contains just a small sample for Korean (12 bi- one-hour sessions, all collected over a three-month period, beginning when the child was two years old). That sample offers no instances of negated sentences containing *motun* 'all/every,' and a search of the Sejong corpus of spoken and written Korean (http://www.sejong.or.kr), consisting of more than ten million words, revealed just 23 examples of this pattern in adult-to-adult communication.

Crucially, however, processing pressures offer a straightforward explanation for the course of development in English and Korean. The reason that children learning English initially allow both the *not > all* and the *all > not* interpretations is that the two options are comparable in processing cost. And the reason that children learning Korean initially allow only the *all > not* interpretation is that it has a lower processing cost than its *not > all* counterpart, which requires an extra step in Korean. In both languages, the facts follow largely from internal considerations of interpretive economy rather than from experience.

4.2 Scope preferences in L2 learners of English and Korean

The scope interpretation facts for second language learners reveal a further striking contrast. Whereas Korean-speaking learners of English initially transfer the dominant *all > not* interpretation of their language to English, English-speaking learners of Korean do not transfer the preferred *not > all* interpretation of English to Korean (O'Grady 2013).

Table 4. The developmental sequences for scope in L2 English and L2 Korean

Korean-speaking learners of English as a second language	English-speaking learners of Korean as a second language
1. *all > not* (transfer)	1. *all > not* (no transfer)
2. *not > all*	2. no change

Once again, it is unlikely that input could explain this asymmetry. Not only are the relevant sentences rare in both languages, instruction appears not to be helpful either. As I have reported elsewhere (O'Grady 2013), American learners of Korean receive no instruction on scopal interpretation (yet correctly avoid transfer), whereas Korean learners of English receive relevant instruction (but nonetheless succumb to transfer).

Processing considerations provide a plausible explanation for the facts, if we make a simple assumption: learners have a propensity to transfer the entrenched routines of their first language to a second language, but only if that does not increase processing cost in the second language (O'Grady 2010, 2013; O'Grady et al. 2011; see also Filipović & Hawkins 2013). The two outcomes reported in table 4 now follow.

First, as depicted in (11), Korean learners of English transfer the processing routine for the *all > not* interpretation from their native language to English, because this familiar option incurs no more processing cost in English than the *not > all* interpretation (which is why it is also attested in children learning English as a first language).

(11)   Donald did**n't** send all the letters

⊠ ⊠ ⊠

Lee (2009) documents the subsequent gradual emergence of the *not > all* interpretation in advanced second language learners, presumably in response to experience.

Second, English-speaking L2 learners are reluctant to transfer the favored *not > all* reading of their native language to Korean, because that interpretation is more costly than the *all > not* alternative in Korean. As observed in our discussion of (10) above, repeated here as (12), the *not > all* option in Korean can be derived only by revising the earlier *all > not* interpretation.

(12)   Donald all the books **not** send

   ✉ ✉ ✉   →   ✉ ✉ | ✉

Processing considerations militate against such an operation, leaving English-speaking learners with the correct interpretation for the Korean pattern, even in the absence of relevant instruction or input.

4.3 The interpretation of null subjects by second language learners

   It is well known that languages such as Japanese and Mandarin Chinese permit null subjects whose interpretation is dependent upon a prominent element in the discourse. Less well known is a curious difference between the two languages with respect to the interpretation of the null pronoun (*pro*) in examples such as the following. (Cl = classifier; Gen = genitive; Hon = honorific)

(13)   Japanese
   San-nin-no keisatu-ga Sato-san-no ie-ni   kita.
   three-Cl-Gen police-Subj Sato-Hon-Gen house-to came
   'Three police officers came to Sato's house.'

   pro Yamada-san-no   ie-ni-mo   kita.
       Yamada-Hon-Gen house-to-also came
   '[They] also came to Yamada's house.'

(14)   Mandarin
   Sān-ge   jingchá lái-le   Zuoténg jiā.
   three-Cl police come-Asp Zouteng house
   'Three police officers came to Ms Sato's house.'

   pro ye   láu-le   Shāntián jiā.
       also come-Asp Shantian house
   '[They] also came to Shantian's house.'

As noted by Monou & Kawahara (2013), Mandarin allows the *pro* subject in the second sentence to refer to the three police officers mentioned in the first sentence (the 'definite' interpretation), roughly as the English pronoun *they* does. In contrast, Japanese allows both this interpretation and an interpretation in which *pro* picks out any three policemen (the 'indefinite' interpretation). Thus whereas the Japanese sentence can be used to describe the situation depicted in either figure 2 or figure 3, the Mandarin sentence is appropriate only for the first case, where there is just a single set of policemen.
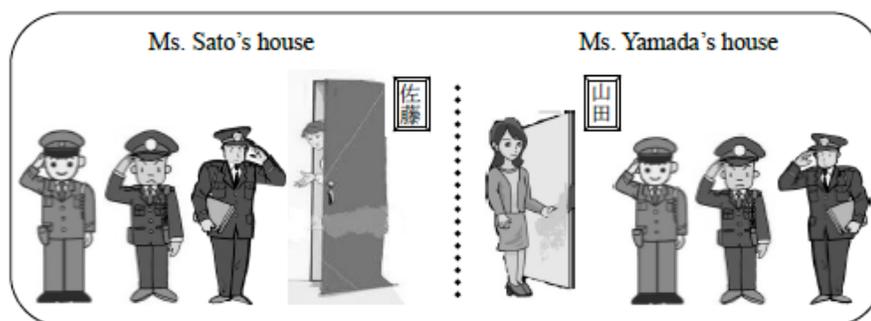
Figure 2. The same three policemen come to Ms Sato's house and to Ms. Yamada's house (Monou & Kawahara 2013)
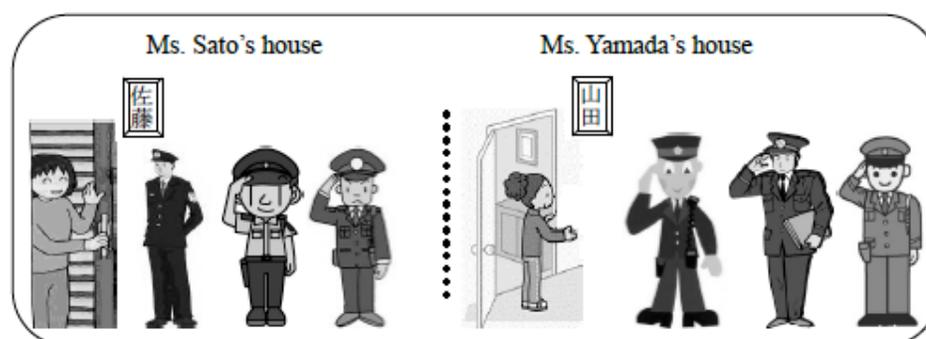

Figure 3. One set of three policemen come to Ms Sato's house, and another set of three policemen come to Ms. Yamada's house (Monou & Kawahara 2013)

The interpretation of *pro* depicted in figure 2 involves reference to a previously established set of policemen, whereas its interpretation in the case of figure 3 requires the construction of a new set of policeman. The latter interpretation is almost surely more costly, consistent with the well-established principle that the introduction of new referents creates additional processing burden (Crain & Steedman 1985, Kurtzman & MacDonald 1993:257-78). Not surprisingly, as Monou & Kawahara report, the indefinite reading is typologically marked as well: any language that allows it also allows the definite reading, but not vice versa. This is the usual situation in cases of processing-cost asymmetries (Hawkins 2004).

Monou & Kawahara investigated the interpretation of Mandarin *pro* by 22 adult Japanese speakers (age 20;6 to 22;1) who were taking an introductory-level course in Mandarin at a university in Japan. At the time of the Monou & Kawahara's experiment, the participants had been studying Mandarin for an average of just 11 months.

A truth value judgment test was conducted to assess the learners' interpretation of the null pronoun in contexts supporting the definite and

indefinite reading (five instances of each type). Whereas the L2 learners accepted the definite interpretation in Mandarin at a rate of 86.4%, they allowed the indefinite interpretation just 33.5% of the time. Moreover, 10 of the 22 participants never accepted the indefinite reading, and just 7 accepted it in more than two of the five test items. (The participants were also tested to make sure that they allowed the indefinite interpretation in Japanese, their native language; all did.)

In sum, when Japanese speakers learn Mandarin as a second language, their course of development seems to be characterized by a strong preference for the definite interpretation of *pro* over its indefinite counterpart.

(15)   definite interpretation > indefinite interpretation

Why should development proceed in this way, especially when both interpretatios are allowed in the first language? Monou & Kawahara establish that students receive no relevant instruction, and patterns in which a null pronoun has a numeral-bearing antecedent are surely too rare to permit an inductive generalization in the first year of study.

None of this matters if, as I suggest, internal processing pressures are responsible for the developmental asymmetry. The high processing cost associated with the indefinite interpretation makes it an unattractive option for second language learners, even if it is permitted in their first langauge, thereby yielding the preference summarized in (15).

4.4 A tentative picture of development

The key idea that I have been outlining is simply that development is best understood in terms of processing cost, which in turn is determined by two sets of factors.

On the one hand, there are internal processing pressures of various sorts:

i.   Locality and immediacy favor co-reference involving a co-argument antecedent rather than an antecedent in a different clause, or even a different sentence. This preference underlies children's early success in the interpretation of reflexive pronouns in English.

ii. Incrementality favors scopal interpretations that can be constructed in a cumulative linear manner, without the need to revise previously assigned interpretations. This factor underlies the early preference for the *all > not* reading in Korean, while at that same time explaining why there is no comparable initial preference in English, with its different word order. It also accounts for the reluctance of English speakers to allow the *not > all* interpretation in Korean despite its predominance in English.

iii. Referent stability favors the identification and maintenance of a particular referent (or set of referents) throughout the discourse. This preference is responsible for the reluctance of Japanese learners to accept the indefinite interpretation of null pronouns in Mandarin even though this interpretation is allowed in their native language.

On the other hand, external factors involving the form and use of the ambient language also contribute to processing cost—and hence to the course of development. The most obvious of these, of course, is frequency of occurrence, which has a major effect on the strength and accessiblity of processing routines, as observed in section 2. But a number of other external factors also seem to come into play, at least three of which merit mention.

i. Perceptibility ('acoustic salience') is known to affect the speed and success of processing (Dick et al. 2001). Not surprisingly, given its relevance to processing, this factor is also implicated in shaping the course of development, often in interaction with frequency factors. As Hsieh et al. (1999) observe, for instance, plural -s is not only more frequent overall than the later-acquired verbal -s in caregiver speech, it is more likely to occur in sentence-final position, where its perceptibility is further enhanced by lengthening; see also Song et al. (2009) and Sundara et al. (2011).

ii. Morphophonological regularity relates to processing cost in two quite different ways. On the one hand, the more learners encounter verbs that take an irregular past tense in a language such as English, the less they encounter verbs that form their past tense with the suffix -ed. This has the effect of limiting the opportunities to strengthen the regular routine. On the other hand, the existence of subtypes of irregularity (internal change, as in *run-ran*; final-consonant modification, as in *send–sent*; partial suppletion as in *think–thought*; full suppletion as in *go–went*) adds to the processor's burden by creating the need for a series of special lexical routines, each of which requires strengthening. The consequences are reflected both in processing (frequent irregulars are retrieved faster) and in development (infrequent regulars are less likely to be acquired by each successive generation of children); see Pinker (1999:125 & 200) for general discussion.[6]

iii. Semantic complexity, as measured by the number of semantic features encoded by a particular morpheme, appears also to be relevant. Complexity of this sort is found in the English verbal suffix -s, which encodes third person, singular agreement, and present tense, compared to the single feature encoded

---

[6] Filipović & Hawkins (2013) object to the attempt to equate ease of processing with ease of learning, noting that irregulars—once acquired and entrenched—are often easy to process. This is correct; however, *at the point at which they are being learned,* irregulars run against the grain of a more general processing routine, with the consequences for processing cost and development that I have noted.

by the earlier acquired plural suffix (e.g., Brown 1973:370-71, Goldschneider & deKeyser 2001:24-25). Other examples of semantic complexity include the grammaticalized preposition *a* in Spanish, which is used for direct objects whose referent is both specific and human (*ver **a** Juan* 'see Juan'); the 1st person pronoun *kāua* in Hawaiian, which is both dual and inclusive (it refers to the speaker plus the addressee, but no one else); and the ergative case marker *ne* in Hindi, which marks subjects, but only if the verb is transitive and in the past tense. All three phenomena are known to be problematic for language learners (e.g., Montrul & Bowles 2009 for Spanish, Housman et al. 2011 for Hawaiian, Montrul et al. 2012 for Hindi), arguably because of the processing cost associated with the activation of multiple semantic contrasts with each use of the morpheme.

In sum, two very different sets of processing pressures, one internal and the other external, contribute to development in the theory that I propose. Often the two interact, but they are not the same thing. As we have seen, there are cases where internal processing pressures alone seem to forge the course of development. And, of course, it is easy to think of cases where external factors are the primary determinant of development. (One obvious example involves lexical learning, whose precise course appears to reflect experience quite faithfully; see, e.g., Huttenlocher et al. 1991 and Hoff & Naigles 2002). As we will see next, a distinction between the two types of pressures is also crucial to estimating the extent to which development is uniform.

5. Uniformity in development

It is commonly believed that development manifests a high degree of variation: no two learners proceed at the same rate, make the same errors, or attain proficiency in exactly the same way (de Bot et al. 2007, van Dijk et al. 2011, Verspoor & van Dijk 2011). Against this view, I propose that development is in fact highly uniform, in the following sense.

(12) The Uniformity Thesis
      Development is uniform, where processing cost is relevant.

Various issues and challenges arise at once (see, e.g., R. Ellis 2015), some of which I consider next.

5.1 The data bearing on development

One very major issue involves the type of data that is relevant to the assessment of uniformity. In a great deal of influential research, this data comes primarily from learners' spontaneous speech (e.g., Brown 1973 and Pienemann &

Kessler 2011, to take two very different examples). This practice is problematic for at least two reasons.

First, as noted in section 2.2, a focus on production is ill suited to the investigation of patterns that occur infrequently—many of which are critical to understanding the role of internal processing pressure in shaping development. Reflexive pronouns and scope are cases in point: they are quite rare in spontaneous speech, yet provide crucial insights into the course and nature of development.

Second, a focus on speech production essentially guarantees a level of variation likely to mask developmental regularity. Even mature monolingual speakers of a language differ in their ability to express themselves in extemporaneous speech, and there can be fluctuations in performance literally from moment to moment. Variation and instability are no doubt magnified in learners, whose processing routines are less entrenched than those of mature speakers and hence even more susceptible to the factors that make speech difficult in the first place. Conclusions based on production data need to be tempered by the insights afforded by controlled experiments, especially those involving comprehension.

5.2 Variation in rate of development and in ultimate attainment

Another issue calling for comment involves the existence of variation that is unrelated to processing cost. One example of this involves rate of development, which undoubtedly varies from individual to individual, as first documented in Brown's (1973) study of Adam, Eve and Sarah. As Brown observed, Eve's rate of development far exceeded that of the other two children, despite similarities in the order in which various morphemes were acquired.

This sort of variation is irrelevant to the Uniformity Thesis. The cost of maintaining a vocabulary of, say, 300 words, or the cost of implementing a processing routine for subject relative clauses, is presumably the same regardless of the age of acquisition. Differences in rate of acquisition, whatever their source, therefore do not constitute a challenge to the uniformity claim.

A second type of variation stems from differences across individuals in cerebral plasticity (e.g., Herschensohn 2009), which underlies the ability to form new processing routines, and in working memory capacity (e.g., Just & Carpenter 1992), which supports the use of those routines in speech and comprehension. Both factors have played a prominent role in theorizing about variation in second language acquisition, but it is vital to distinguish in this regard between attainment and path. There is undoubtedly variation in the ultimate product of development, especially in the case of second language acquisition, and this variation may well turn out to reflect difference in cerebral plasticity and/or working memory. However, there need not be variation in the *course* of development, which is the sole concern of the Uniformity Thesis. I return to this matter in section 5.5.

5.3 Differences that reflect variation in the input

Another sort of variation, ubiquitous in language acquisition, can be traced to differences in the input. The most obvious and clear-cut examples involve learner populations in distinct dialect areas. Children growing up in American almost surely learn the words *truck*, *mail*, and *parking lot* before *lorry*, *post*, and *car park*, while the reverse is true for children being raised in England. (Similar examples can easily be constructed for second language learners.)

There are comparable differences in morphosyntactic development. It would not be surprising to find that children in England use *shall* before their counterparts in America. Nor would we be surprised to discover that American children become familiar with the *don't need* pattern (*They don't need to go*) before the alternative *need not* construction (*They needn't go*), whereas the reverse happens in England.

Developmental variation that is triggered by differences in input does not undermine the Uniformity Thesis. To the contrary, it confirms it. If in fact input-related factors affect processing cost as I propose (section 2.2), differences in experience *must* affect the course of development. Processing determinism allows no alternative. Children in America should uniformly learn *truck* before *lorry*, and children in England should uniformly learn *lorry* before *truck*.

5.4 Variation in developmental order

As noted at section 2, there are two types of developmental progression—developmental sequences, which involve the emergence of a single form-meaning mapping, and developmental order, which involves the emergence of a set of (typically unrelated) mappings—such as the fourteen morphemes investigated in Brown's (1973) pioneering work.[7]

A great deal of effort has been devoted to the study of developmental order, especially in the context of second language acquisition. Although some early work reported significant uniformity in the order in which Brown's fourteen morphemes are mastered (Dulay & Burt 1973, 1974; Bailey et al. 1974; Goldschneider & deKeyser 2001), an ambitious study by Murakami (2013) suggests otherwise. Drawing on data from 140,000 essays produced by 46,700 L2 learners of English from ten different L1 backgrounds, Murakami uncovered 'large intra- and inter-learner variation in morpheme development' (p. 236). I believe that this result will be replicated in other languages, including the many (such as Korean) that have dozens and even hundreds of grammatical morphemes rather than the handful found in English.

---

[7] Present progressive, *in*, *on*, plural, past irregular, possessive, uncontractible copula, articles, past regular, third person regular, third person irregular, uncontractible auxiliary, contractible copula, contractible auxiliary

Crucially, though, there are reasons to be cautious about the conclusion that this sort of variation undermines the Uniformity Thesis. For one thing, the observed differences come largely from the study of production (see above). Moreover, a good deal of the variation can be traced to differences in the composition and quantity of the input (e.g., Murakami 2013: 241), and is therefore expected for the reasons noted in the previous section.[8]

A further sort of variation in L2 developmental order reflects the influence of the learner's first language. As Luk & Shirai (2009) observe, for instance, speakers of Japanese, Korean and Chinese have an advantage over speakers of Spanish in the mastery of the English pre-nominal genitive (e.g., *Jerry's book*), for which their languages, but not Spanish, have an approximate counterpart. On the other hand, speakers of Spanish manifest an advantage in the mastery of articles and the plural marker (which their language has, but Japanese, Chinese and Korean don't). Murakami (2013) documents further L1 effects drawing on a much larger data base.

My view of transfer (O'Grady 2010, 2013) is that it too is the product of processing pressure: learners seek to facilitate processing in their L2 by calling upon the already entrenched low-cost routines of their L1. A key piece of evidence for this view was discussed in sections 4.2 and 4.3, where it was noted that transfer is impeded when an L1 processing routine carries a higher cost in the second language than in the first.

There should of course be uniformity in the cost-related consequences of particular instances of transfer. For example, speakers of English should always be repelled by the cost of the *not > all* interpretation in Korean. But this is not what is at stake in the variation observed by Luk & Shirai and by Murakami. The variation they report reflects differences in the processing routines available in individual first languages, some of which have (for example) pre-nominal genitives and some of which don't. If transfer is motivated by processing considerations, differences of this type should be (uniformly) manifested in different populations of second language learners. There is currently no reason to think that this is not the case.

5.5 The critical test of uniformity

How then can the Uniformity Thesis be put to the test? The answer is simple: it can be tested most directly by investigating phenomena that show the fewest

---

[8] A further problem, largely ignored in the literature, is the difficulty of calculating processing cost when one is dealing with multiple variables and multiple unrelated morphemes. Is the determiner *the* more or less semantically complex than past tense *-ed*? Is there a difference in the acoustic salience of the preposition *in* and the suffix *-ing*? Is the morphophonemic alternation that inserts [ə] after a stem-final coronal continuant in the third person singular (hence *push[əs]*, but *put[s]*) more or less costly than the insertion of [ə] after a stem-final coronal stop in the past tense suffix (hence *push[t]*, but *putt[əd]*)? The answer is not only that we do not know, but also that no serious attempt has been made to find out.

confounds and extraneous influences, namely phenomena whose properties are largely shaped by internal processing pressures.

The four phenomena examined in sections 3 and 4 are a case in point, as their key properties and their course of development appear to reflect internal processing pressures rather than the influence of experience or other external factors. The predictions of the Uniformity Thesis, summarized in table 5, are therefore straightforward. In keeping with other work on processing-based approaches to markedness and development (e.g., Hawkins 2007), these predictions are formulated in terms of what should *not* happen: if X is easier to process than Y, X should not emerge later than Y.

Table 5   Developmental predictions derived from internal factors

Phenomenon	Predictions
English pronouns	The interpretation of reflexive pronouns should not be mastered later than the interpretation of plain pronouns.
scope in Korean as an L1	The *all > not* interpretation should not emerge later than the *not > all* interpretation.
scope in Korean as an L2	The *all > not* interpretation should not emerge later than the *not > all* interpretation.
the interpretation of null pronouns in Chinese as an L2	The indefinite interpretation of *pro* should not initially be favored over the definite interpretation.

The available literature offers significant support for these predictions. There is good evidence that English-speaking children sometimes interpret plain pronouns as if they were reflexives (e.g., Conroy et al. 2009); there is no evidence of a systematic error in the reverse direction, except possibly in the case of Specific Language Impairment (e.g., Joanisse & Seidenberg 2003). First language learners of Korean in O'Grady et al.'s (2011) study show an early preference for *all > not* scope; although some also allow the *not > all* interpretation, none show a preference for this reading. English-speaking L2 learners of Korean show a uniform preference for the *all > not* interpretation in O'Grady's (2013) experiment; none manifest the opposite preference. No beginning Japanese-speaking learners of Chinese as an L2 in Monou & Kawahara's study preferred the indefinite interpretation of *pro*. These findings point to a fundamental similarity in the course of development when it is shaped by internal processing pressures. [9]

---

[9] Two other predictions, both correct, can be derived with the help of minimal additional assumptions. First, all other things being equal, English-speaking children should initially exhibit no preference for either the *not > all* or *all > not* interpretation in their native language; see section 4.1. Second, assuming that L2 learners transfer the stronger routine from their L1 to the L2

6. Conclusion

The investigation of linguistic development requires us to confront the classic challenge of cognitive science: why do mental systems such as language have the particular properties that they do, and how do those properties emerge in the course of acquisition? There is, I believe, a single answer to the two questions, which is that processing pressures both shape language and determine the developmental course that leads to its acquisition—the thesis of processing determinism. All other things being equal, the pieces of language, ranging from vocabulary to grammatical patterns, emerge in a way that reflects their processing cost. (Put another way, low processing cost is a powerful attractor state.)

As I have tried to show, a variety of factors, some external and some internal, interact to determine processing cost. On the one hand, the mechanisms relevant to processing are highly sensitive to features of the input, including frequency of occurrence, perceptibility, morphophonological regularity, and the like. On the other hand, because processing is subtended by limited working memory resources, learners manifest an instinctive preference for efficiency—with consequences that play a major role in shaping development, as we have seen.

The facts of development are no less complex than the facts of language, of course, and the explanatory challenges are enormous. Processing determinism offers a way to approach those challenges as part of an integrated theory of how language works—how it is used, how it is acquired, how it changes over time, and so on. Further research in this area will no doubt shed light on whether this sort of approach is on the right track.

References

Ambridge, Ben & Elena Lieven. 2015. A (neo-)constructivist account of child language acquisition. In B. MacWhinney & W. O'Grady (eds.), *Handbook of language emergence*. Boston: Wiley. In press.

Ambridge, Ben, Evan Kidd, Caroline Rowland, & Anna Theakston. to appear. The ubiquity of frequency effects in first language. *Journal of Child Language*.

Bailey, N., Madden, C., & Krashen, S.D. 1974. Is there a "natural sequence" in adult second language learning? *Language Learning* 24, 235-243.

Baker, Mark. 2001. *The atoms of language: The mind's hidden rules of grammar*. New York: Basic Books.

---

if the processing cost is the same in the second language (O'Grady 2013), Korean-speaking L2 learners of English should initially prefer the *all > not* interpretation in English; see section 4.2.

Bates, Elizabeth & Brian MacWhinney. 1987. Competition, variation, and language learning. In B. MacWhinney (ed.), *Mechanisms of language acquisition*, 157-93. Mahwah, NJ: Erlbaum.

Bates, Elizabeth & Brian MacWhinney. 1988. What is functionalism? *Papers and Reports on Child Language Development* 27: 137–52.

Behrens, Heike. 2009. Usage-based and emergentist approaches to language acquisition. Linguistics 47, 383-411.

Bergmann, Christina, Markus Paulus & Paula Fikkert. 2012. Preschoolers' comprehension of pronouns and reflexives: The impact of the task. *Journal of Child Language* 39, 777-803.

Brown, Roger. 1973. A first language: The early stages. Cambridge, MA: Harvard University Press.

Chien, Yu-Chin & Kenneth Wexler. 1990. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1, 225–295.

Chomksy, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.

Chomsky, Noam. 2008. The biolinguistic program: Where does it stand today? Available at http://clt.blcu.edu.cn/resource/[[1].Chomsky_2008_bio-todayGood.gd.pdf

Clackson, Kaili, Claudia Felser, & Harald Clahsen. 2011. Children's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language* 65, 128-44.

Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz, & Colin Phillips. 2009. Equal treatment for all antecedents: How children succeed with Principle B. Linguistic Inquiry 40, 446-86.

Crain, Stephen & Rosalind Thornton. 1998. *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics*. Cambridge, MA: MIT Press.

Crain, Stephen & Mark Steedman. 1985. On not being led up the garden path: The use of context by the psychological parser. In D. Dowty, L. Kartunnen, & A. Zwicky (eds.), *Natural language parsing*, 320-58. Cambridge, UK: Cambridge University Press.

de Bot, K., Lowie, W., & Verspoor, M. 2007. A Dynamic Systems Theory approach to second language acquisition. Bilingualism: Language and Cognition, 10(1), 7–21

Déprez, Vivianne & Amy Pierce. 1993. Negation and functional projections in early grammar. *Linguistic Inquiry* 24, 25-67.

Dick, Frederic, Elizabeth Bates, Beverly Wulfeck, Jennifer Utman, Nina Dronkers, & Morton Gernsbacher. 2001. Language deficits, localization and grammar: Evidence for a distributive model of language breakdown in aphasics and normals. *Psychological Review* 108, 759–788.

Dulay, H.C., & Burt, M.K. (1973). Should we teach children syntax? *Language Learning*, 23, 245-258.

Dulay, H.C., & Burt, M.K. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24, 37-53.

Ellis, Nick. 2006. Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing,

blocking, and perceptual learning. *Applied Linguistics* 27, 164 –194.

Ellis, R. 2015. Researching acquisition sequences: Idealisation and deidealisation SLA. *Language Learning* 65.

Ferreira, Fernanda. 2003. The misinterpretation of noncanonical sentences. *Cognitive Psychology* 47, 164-203.

Filipović, Luna & John Hawking. 2013. Multiple factors in second language acquisition: The CASP model. *Linguistics* 51, 145-76.

Foraker, Stephani & Brian McElree. 2007. The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language* 56, 357-83.

Goldschneider, Jennifer Robert Dekeyser. 2001. Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. Language Learning 55. 27-77.

Han, Chung-hye, Jeffrey Lidz, & Julien Musolino. 2007. V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry* 38, 1-48.

Hawkins, John. 2004. *Efficiency and complexity in grammars*. Oxford, UK: Oxford University Press. Jaeger, T. Florian & Harry Tily. 2011. On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science* 2, 323-35.

Hawkins, John. 2007. Acquisition of relative clauses in relation to language universals. *Studies in Second Language Acquisition* 29, 337-44.

Herschenson, Julia. 2009. Fundamental and gradient differences in language development. Studies in Second Language Acquisition 31, 259-89.

Hickok, G. 2009. Understanding language without ability to speak. *Talking Brains: News and Views on the Neural Organization of Language*. http://www.talkingbrains.org/2009/03/understanding-language-without-ability.html

Housman, Alohalani, Kaulana Dameg, Māhealani Kobashigawa, & James Brown. 2011. Report on the Hawaiian oral language assessment (H-OLA) development project. *Second Language Studies* 29, 1-59.

Hsieh, Li, Laurence Leonard & Lori Swanson. 1999. Some differences between English plural noun inflections and third singular verb inflections in the input: The contribution of frequency, sentence position, and duration. Journal of Child Language 26, 531–543

Ingram, David. 1976. Phonological disability in children. London: Edward Arnold.

Jaeger, T. Florian & Harry Tily. 2011. On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science* 2, 323-35.

Jespersen, Otto. 1933. *Essentials of English grammar*. London: Allen & Unwin.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review 99, 122–149.

Joanisse, Marc & Mark Seidenberg. 2002. Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain and Language* 86, 40-56.

Jurafsky, Dan. 2003. Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hayy & S. Jannedy (eds.), *Probabilistic Linguistics*, 39-95. Cambridge, MA: MIT Press.

Kurtzman, Howard & Maryellen MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition* 48, 243-79.

Lee, Sunyoung. 2009. Interpreting scope ambiguity in first and second language processing: Universal quantifiers and negation. Ph.D. dissertation, University of Hawaii at Manoa.

Lenneberg, E. 1962. Understanding language without ability to speak: a case report. *Journal of Abnormal and Clinical Psychology 65*, 419-425.

Lidz, Jeffrey & Julien Musolino. 2002. Children's command of quantification. *Cognition* 84, 113-54.

Lieven, Elena. 2010. Input and first language acquisition. Lingua 120, 2546-56.

Love, Tracy, Matthew Walenski, & David Swinney. 2009. Slowed speech input has a differential impact on on-line and off-line processing in children's comprehension of pronouns. *Journal of Psycholinguistic Research* 38, 285-304.

Luk, Zoe & Yasuhiro Shirai. 2009. Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -*s*, articles, and possessive *'s*. *Language Learning* 59, 721-754.

MacWhinney, Brain. 1987. The Competition Model. In B. MacWhinney (ed.), *Mechanisms of language acquisition*, 249-308. Mahwah, NJ: Erlbaum.

MacWhinney, Brian. 2004. A multiple process solution to the logical problem of language acquisition. Journal of Child Language 31, 883-914.

Menn, Lise. 2000. Let's face a simple question: Why is canonical form simple? *Brain and Language* 71, 157-59.

Monou, Tomoko & Shigeto Kawahara. 2013. The emergence of the unmarked in L2 acquisition: Interpreting null subjects. Proceedings of the Fourteenth Tokyo Conference on Psycholinguistics. Available at http://user.keio.ac.jp/~kawahara/pdf/MonouKawahara.pdf

Montrul, Silvina & Melissa Bowles. 2009. Back to basics: Differential object marking under incomplete acquisition in Spanish heritage speakers. Bilingualism: Language and Cognition 12 , 363-83.

Montrul, Silvina, Rakesh Bhatt & Archna Bhatia. 2012. Erosion of case and agreement in Hindi heritage speakers. *Linguistic Approaches to Bilingualism* 2, 141–176.

Murakami, Akira. 2013. Individual variation and the role of L1 in the L2 development of English grammatical morphemes: Insights from learner corpora. PhD dissertation. Department of Theoretical and Applied Linguistics, University of Cambridge.

Musolino, Julien, Stephen Crain, & Rosalind Thornton. 2000. Navigating negative quantificational space. *Linguistics* 38, 1-32.

Musolino, Julien & Jeffrey Lidz. 2006. Why children aren't universally successful with quantification. *Linguistics* 44, 817-852.

O'Grady, William. 2005. *Syntactic carpentry: An emergentist approach to syntax*. Mahah, NJ: Erlbaum.

O'Grady, William. 2011. Relative clauses: Processing and acquisition. In E. Kidd (ed.), *The acquisition of relative clauses: Processing, typology and function*, 13-38. Amsterdam: John Benjamins.

O'Grady, William. 2013. The illusion of language acquisition. *Linguistic Approaches to Bilingualism* 3, 253-85.

O'Grady, William. 2015. Anaphora and the case for emergentism. In B. MacWhinney & W. O'Grady (eds.), Handbook of language emergence. Boston: Wiley. In press.

O'Grady, William, Hye-Young Kwak, Miseon Lee, & On-Soon Lee. 2011. An emergentist perspective on partial language acquisition. *Studies in Second Language Acquisition* 33, 323-45

Paradis, Michel. 2004. *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.

Pienemann, Manfred. 2005. An introduction to Processability Theory. In M. Pienemann (ed.), *Cross-linguistic aspects of Processability Theory*, 1-60. Amsterdam: John Benjamins.

Pinker, Steven. 1999. Words and rules. New York: Basic Books.

Reinhart, Tanya. 2006. *Interface strategies: Reference-set computation*. Cambridge, MA: MIT Press.

Reuland, Eric. 2011. Anaphora and language design. Cambridge, MA: MIT Press.

Roeper, Tom. 2007. *The prism of grammar: How child language illuminates humanism*. Cambridge, MA: MIT Press.

Sekerina, Irina, Karin Stromswold, & Arild Hestvik. 2004. How do adults and children process referentially ambiguous pronouns? *Journal of Child Language* 31, 123–152.

Sharwood Smith, Michael & John Truscott. 2014. *The multilingual mind: A modular processing perspective*. Cambridge: Cambridge University Press.

Shi, Rushen. 2014. Functional morphemes and early language acquisition. *Child Development Perspectives* 8, 6 – 11.

Snyder, William & Diane Lillo-Martin. 2011. Principles and parameters theory and language acquisition. In P. Hogan (ed.), *The Cambridge encyclopedia of language sciences*, 670-73. Cambridge, UK: Cambridge University Press.

Soderstrom, Melanie, Erin Conwell, Naomi Feldman & James Morgan. 2009. The learners as statistician: Three principles of computational success in language acquisition. *Developmental Science* 12, 409-11.

Song, Hyun-joo & Cynthia Fisher. 2007. Discourse prominence effects on 2.5-year-old children's interpretation of pronouns. *Lingua* 117, 1959-87.

Song, J.Y., Sundara, M., Demuth, K., 2009. Phonological constraints on children's production of English third person singular -s. *Journal of Speech, Language, and Hearing Research* 52, 623-642.

Sundara, M., Demuth, K., Kuhl, P., 2011. Sentence-position effects on children's perception and production of English third person singular -*s*. *Journal of Speech, Language, and Hearing Research* 54, 55-71.

Thompson, Susan & Elissa Newport. 2007. Statistical learning of syntax: The role of transitional probability. *Language Learning and Development* 3, 1-42.

Tomasello, Michael. 2003. *Constructing a language: A usage-based theory of language*. Cambridge, MA: Harvard University Press.

van Dijk, M., Verspoor, M., & Lowie, W. (2011). Variability and DST. In M. Verspoor, K. deBot, & W. Lowie (Eds.), A dynamic approach to second language development: Methods and techniques, 55–84. Amsterdam: John Benjamins.

Van Geert, Paul & Marjolijn Verspoor. 2015. Dynamic systems and language development. In B. MacWhinney & W. O'Grady (eds.), Handbook of language emergence. Boston: Wiley. In press.

Van Rij, Jacolien, Hedderik van Rijn, & Petra Hendriks. 2010. Cognitive architectures and language acquisition: A case study in pronoun comprehension. *Journal of Child Language* 37, 731-66.

VanPatten, Bill. 2012. Input processing. In S. Gass & A. Mackey (eds.), *The handbook of second language acquisition*, 268-81. New York: Routledge.

Verspoor, M., & van Dijk, M. (2011). Visualizing interactions between variables. In M. H. Verspoor, K. de Bot, & W. Lowie (Eds.), A dynamic approach to second language development: Methods and techniques, pp. 85–98. Amsterdam: John Benjamins.

Wexler, Kenneth. 1990. Innateness and maturation in linguistic development. Developmental Psychobiology 23, 645-60.

Yang, Charles. 2004. Universal Grammar, statistics or both? *Trends in Cognitive Science* 8, 451-56.