Input Effects on the Acquisition of Finiteness

Matthew Rispoli and Pamela Hadley

1. Introduction

Understanding input effects is crucial for delineating the roles of Universal Grammar and learning in grammatical development. Unlike the input effects seen in vocabulary development (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Hart & Risley, 1995), effects on grammatical development have been more difficult to identify (Newport, Gleitman, & Gleitman, 1977; Hoff-Ginsberg, 1998; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; see also Valian, 1995 for review). One weakness of earlier studies has been their generality. For example, Newport et al. (1977) tried to uncover relationships between ten aspects of maternal speech, some structural and some interactional, and four measures of structure in child output, as well as the ubiquitous measure of mean length of utterance (MLU). Unfortunately, a methodological problem arises when a large number of correlations are tested; the *multiplicity problem* (Benjamini & Hochberg, 1995). As Valian (1995) has pointed out, studies in the mold of Newport et al. (1977) have turned up "no robust findings" (p. 512). In contrast, Hadley, Rispoli, Fitzgerald, and Bahnsen (2011) focused on a single input variable, input informativeness for tense, based on the theoretical work of Legate and Yang (2007). Hadley et al. found that proportionately richer signaling of English tense in parental input predicted individual differences in the rate of children's acquisition of finiteness between 21 and 30 months of age.

Another weakness of earlier studies has been a lack of control for differences in children's developmental levels when parent input is sampled. Taking Newport et al. (1977) as an example, age and level of syntactic developmental varied in the children studied. The effects of age and prior knowledge had to be controlled statistically. In contrast, Hadley et al. (2011) identified an optimal point in time to explore input effects based on Rispoli, Hadley, and Holt's (2009) quantitative model of growth in finiteness. Because the vast majority of children developing typically do not show emergence of finiteness morphemes at 21 months, but growth in the finiteness system is evident shortly thereafter, Hadley et al. measured input when children were 21 months old. Additionally, child vocabulary at 21 months was used to control for individual differences in development before measureable growth in finiteness began.

Although the results of Hadley et al. (2011) supported Legate and Yang's (2007) Variational Learning (VL) hypothesis, VL is incompatible with one empirical finding of the Rispoli et al. (2009) finiteness growth model. VL predicts variation at the onset of learning. Rispoli et al. (2009) found that finiteness productivity for the group was not significantly different from zero at the centering point (i.e., 21 months) nor were there significant individual differences among children at 21 months of age. Thus, a no-intercept model was the best fit to the data. In other words, children acquiring the finiteness system of English do not vary at the onset of learning as VL predicts. Rather, productivity of finiteness marking grows gradually from zero from during the third year of life. Additionally, the VL model does not attempt to explain differences in the learning of individual finiteness morphemes, but this almost certainly exists and must be part of the process of acquiring finiteness in English. On the other hand, the Gradual Morphosyntactic Learning hypothesis (GML; Rispoli & Hadley, 2011; Rispoli, Hadley, & Holt, 2012) is compatible with a no-intercept model for the acquisition of finiteness. GML also directly addresses how the learning of individual morphemes fits into the larger picture of acquiring

* Matthew Rispoli, University of Illinois. mrispoli@illinois.edu. Pamela Hadley, University of Illinois. We would like to acknowledge the National Science Foundation for funding this research (NSF BCS 08-22513).

 Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012), ed. Chia-Ying Chu et al., 121-127. Somerville, MA: Cascadilla Proceedings Project.

the system of marking finiteness in English and has generated specific hypotheses concerning the facilitating and non-facilitating aspects of parents' grammatical input.

During the third year of life, children's productivity for copula grows more rapidly than for any other morpheme in the tense / agreement suite (Hadley & Short, 2005; Rispoli et al., 2009), giving rise to the following sequence:

Copula > verb-*s*, Aux Do, verb-*ed* > Aux BE

Part of this sequence is related to sheer input token frequency. Copula BE is the most frequent morpheme class in parent input (Hadley et al., 2011; Fitzgerald, Hadley, & Rispoli, 2013). Copula BE is three times more frequent than auxiliary BE and six times more frequent than verb-*s*. The position of verb-*s* in the sequence makes it plain that token frequency explains only why copula should be first in the sequence. Rispoli et al. (2012) provided an analysis of the finiteness system based on positional and functional processing (Bock & Levelt, 1994), which was congruent with the rest of the observed sequence of emergence.

Rispoli et al. (2012) went further than an explanation of the sequence in which finiteness markers increase in productivity. They presented evidence that early productivity of copula *is* facilitates the productivity of other markers of the [3rd person singular subject, present tense] feature bundle (verb –*s*, auxiliary *does*). They called this mechanism *cross-morpheme facilitation* (see also Leonard, Camarata, Brown & Camarata, 2004). Because cross-morpheme facilitation is at play, input to the most frequent part of the finiteness system, namely copula *is* can indirectly affect the growth of a less frequent part of the system, namely verb-*s*.

In the present study, we investigate whether *diversity* in parents' input for the most frequent morpheme in the finiteness system, copula *is*, is related to children's acquisition of the verb-*s* inflection. We operationalize diversity as a *type-token ratio* (TTR). When transitional probabilities in the input contexts surrounding a morpheme are high, it is difficult to isolate the morpheme (Thompson & Newport, 2007). We hypothesize that low input diversity in neighboring morphemes encourages rote learning and hinders morphemic analysis, whereas high input diversity facilitates analysis. We test our hypotheses with a new investigation of the acquisition of finiteness in the third year of life, starting with parent input at 21 months and ending with an assessment of children's acquisition of verb-*s* and verb past at 36 months. Our predictions are as follow:

1. The type / token ratio (TTRs) for copula *is* in parent input at 21 months will be positively correlated with children's copula *is* productivity at 30 months.
2. Because of cross-morpheme facilitation, parents' copula *is* input TTRs from 21 months will have a significant relationship with children's accuracy on verb-*s* at 36 months, but NOT with verb past.

2. Method

Our participants were 42 children (21 girls), developing typically in general communication, expressive vocabulary, and phonology. All children passed the communication screening of the *Ages and Stages Questionnaire* (Bricker & Squires, 1999) at 21 and 24 months, had expressive vocabularies at or above the 10th percentile at 24 months on the *MacArthur-Bates Communicative Development Inventory: Words and Sentences* (CDI; Fenson et al., 2007), and passed phonology screenings to ensure that final /s, z, t, d/ were in their phonological repertoires at 30 and 36 months (Bleile, 1995; Rice & Wexler, 2001). All families were monolingual Standard American English speakers.

The data used in the current study consisted of the CDI total vocabulary size at 21 months, child sex, a 30-min parent input sample at 21 months, a 1-hr child language sample at 30 months, and verb inflection probe scores from the *Test of Early Grammatical Impairment* (TEGI; Rice & Wexler, 2001) at 36 months. Child CDI vocabulary size at 21 months was a control variable. Sex was included as a control variable because girls have larger vocabularies than boys do at 21 months of age. From the 21-month input sample, we calculated our input variable, the copula *is* input TTR. Copula *is* type was based on morphemes surrounding *is* or *'s*, and operationalized in the following manner:

(a) type for contracted *'s* was defined as the word to which *'s* was contracted (e.g., it*'s* what*'s*)
(b) type for *is* initial was defined as the morpheme immediately following (e.g., *is* that...,*is* the...)
(c) type for *is* final was defined as the morpheme preceding (e.g., there it *is,* you know what that is)
(d) type for full (non-contracted) *is* between words was the combination of preceding and succeeding words (e.g., this *is* a…, what *is* that?).

In accordance with our hypotheses, we had intermediate and long-term dependent variables. Our intermediate dependent variable was 30-month child copula *is* productivity score. This productivity measure was a modification of the tense agreement productivity (TAP) score first introduced in Hadley and Short (2005). The TAP scoring procedure is in effect a type-based, high frequency filter applied to children's spontaneous productions in a 1-hr language sample. It samples across the range of finiteness morphemes (for more information on computation and validity of the TAP score, see Rispoli et al., 2009; Hadley et al., in press). For this study, we tallied only sufficiently different productions of copula *is*. The high-frequency filter in the TAP criteria does not allow copula *is* when contracted to a pronoun (*he's, what's*) to contribute to the score, because such forms are high frequency in the input. Additionally, the type basis of the TAP score criteria views *is* with the same subject as the same type (e.g. *is that hot*? *that is hot*). Only one token per type can contribute to the productivity score.

Our long-term dependent variables were the 36-month TEGI scores for verb-*s* and verb past. These probes consist of standard prompts to elicit inflected verb forms. The probes are structured in the following manner: (a) For verb-*s*, a picture of a person with activity or profession traits is shown to the child and the examiner says, "*Here is a dancer. Tell me what a dancer does.*" (b) For verb past, two pictures are shown, the first with a person in the midst of an activity and the second with the activity finished. The examiner says to the child. "*Here the girl is skating. Now she's all done. Tell me what she did*". There are ten probes in the verb-*s* series and 18 in the past series with ten regular and eight irregular verb targets.

3. Results

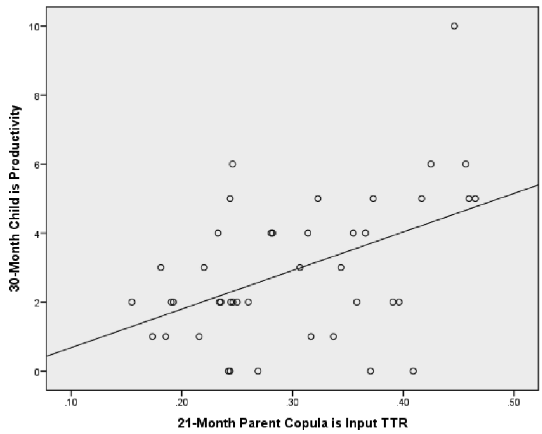
In our 21-month 30-min input samples, parents supplied between 32 and 126 tokens of copula *is* ($M = 74$, $SD = 25$), ensuring the robustness of the TTR calculation. The parents' copula *is* input TTRs ranged from .15 to .47 ($M = .30$, $SD = .09$). There was ample variation in the dependent and the vocabulary control variable as well (see Table 1).

Table 1: Descriptive Statistics for Parent Input, Child Outcome, and Control Variables

Variable Type	Variable Name	*M*	*SD*	Range
Independent:	Parent Input Copula *is* TTR (21 mo.)	.30	.09	.15-.47
Dependent:	Copula *is* productivity (30 mo.)	2.9	2.1	0-10
	TEGI Verb-*s* (36 mo.)	58.3%	33.9%	0-100%
	TEGI Verb Past (36 mo.)	51.9%	29.5%	0-93%
Control Variable	Child CDI total vocabulary (21 mo.)	207	144	37-622

The parents' copula *is* input TTR at 21 months was significantly related to the children's copula *is* productivity at 30 months, $r = .468$, $p = .002$. When we view a scatterplot of these data, we see that one child had a copula *is* productivity of 10 at 30 months, and that the next highest productivity score was six. To guard against the influence of this possible outlier, the data point was removed and the correlation was rerun with 41 children. The correlation remained significant, $r = .404$, $p < .01$.

Figure 1: 21-month Input TTR and 30-month Child Productivity



Recall that the TEGI verb probe scores were our long-term dependent variables. Two hierarchical multiple regressions were conducted using verb-*s* and past tense TEGI scores as dependent variables. The control measures, CDI vocabulary at 21 months, and child sex were entered first. The copula *is* input TTR was allowed entry in stepwise fashion (*F to enter*, $p = .05$). This means that the TTR could only enter the regression if it explained unique variance in the dependent variable over and above the variance explained by the control variables. The regression for verb-*s* was significant, $R^2 = .388$, $p = .001$. Semi-partial correlations revealed that 21-month CDI accounted for 6% of the unique variance, which was not significant $t = 1.873$, $p = .07$. Sex did not account for a significant portion of unique variance, 3%, $t = 1.557$, $p = .129$. On the other hand, the copula *is* input *TTR* accounted for approximately 20% of the unique variance in verb–*s*, $t = 3.320$, $p = .002$. In addition, approximately 9% of the variance was shared by the copula *is* TTR and control variables (see Figure 2). The regression for verb past turned out quite differently. The overall regression was significant, $R^2 = .348$, $p < .001$. CDI vocabulary accounted for approximately 19% of the unique variance, $t = 3.824$, $p = .001$. The child's sex accounted for approximately 15% of the variance, $t = 2.923$, $p = .006$. The copula *is* input TTR did not account for a significant portion of unique variance, t = .613, $p > .05$ (see Figure 3).

Figure 2: Variance Accounted for in 36-month Verb-s TEGI Probe Scores

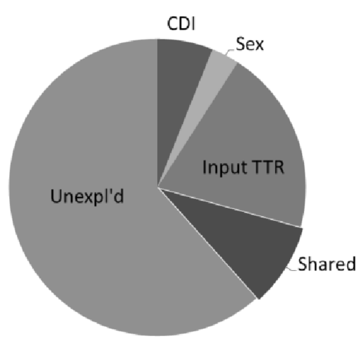
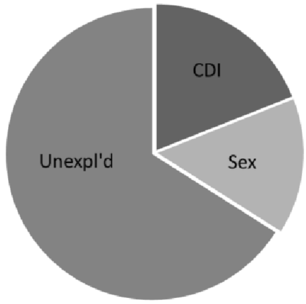


Figure 3: Variance Accounted for in 36-month Verb Past TEGI Probe Scores



4. Discussion

The parent copula *is* input TTR from 21 months was related to child copula *is* productivity at 30 months, nine months later. Furthermore, the copula *is* input TTR accounted for a significant, unique portion of variation in the accuracy of verb-*s* at 36 months, a full 15 months later. But the relationship between copula *is* input and verb inflection was specific to the verb affix that had the same feature specification as copula *is*, namely, [3rd person singular subject, present tense], supporting the hypothesis of cross-morpheme facilitation. Critically, the copula *is* input TTR was unrelated to verb past TEGI scores.

These findings are notable for several reasons. They are relevant to the general question of how children use input to learn individual morphemes. The diversity of the input should not be overlooked, as it may well facilitate acquisition. In this regard, copula *is* may be an example of a more general principle. If a morpheme is repeatedly heard by the child in only one lexical context, there is no reason for the learner to isolate it as a morpheme. It would more likely be learned as a rote portion of a larger lexical unit. With diversity in neighboring morphemes, the child should be better able to isolate the morpheme and pinpoint its grammatical feature context (Pinker, 1984; Slobin, 1973). Although these findings provide initial evidence for the plausibility of the hypothesis, future research is needed to dissociate input diversity from input frequency, something we have not done in the current study.

The findings also support the plausibility of a facilitative learning process. Adapting a term from Leonard et al. (2004), GML refers to this developmental form of generalization as *cross-morpheme facilitation* (Rispoli & Hadley, 2011; Rispoli et al., 2012). Facilitation occurs at the level of grammatical features. In particular, copula *is* bears the features [3rd person singular subject, present tense]. These are the same features born by verb-*s***.** The present research suggests that this mechanism is fairly constrained to this feature specification, and does not extend to past verb inflection. The hypothesis would suggest that facilitation should also affect the acquisition of auxiliary *does* and auxiliary *is*. However, investigating these relationships was beyond the scope of the current study. Additionally, the hypothesis proposed in Rispoli et al. (2012), indicates that positional and functional processing in sentence production (Bock & Levelt, 1994) limit the extent of facilitation. Future research will be needed to determine the extent of facilitation.

Our findings cast doubt on the validity of the claim that the acquisition of tense and agreement in the third year of life proceeds in a piecemeal fashion (Pine, Conti-Ramsden, Joseph, Lieven, & Serriatrice, 2008). Because the input for copula *is* predicts verb-*s* acquisition, Usage-Based accounts of the acquisition of tense and agreement must concede that "abstraction" is underway during the third year of life. Not only is copula *is* a different morpheme than verb-*s*, the two occupy different surface syntactic positions. Therefore, their relationship must be very abstract indeed. Only generative approaches to grammatical development would posit such abstract knowledge so early in the third year of life (Wexler, 2011).

There are multiple limitations to this study that must be kept in mind as future research proceeds. Our study characterized input based on grammatical learning mechanisms hypothesized in GML. It will be important to investigate whether these properties of parent input are related to socioeconomic status (SES) in light of the SES-related differences that have been reported (e.g., Chevrot, Nardy &

Barbu, 2011; Hoff-Ginsberg, 1998; Huttenlocher, Vasilyeva, Waterfall, Vevea, & Hedges, 2007; Rowe, 2008). However, the relationship between SES and the data-providing features of input for morphosyntactic learning are poorly understood. For example, the relationship between SES and lexical diversity in parent input is well documented, as is the relationship between parent lexical diversity and children's subsequent vocabulary growth (Huttenlocher et al., 2010). It is possible that relationships between SES and children's grammatical acquisition may be mediated by the effect of lexical diversity on morpheme learning. If that were so, SES *per se* is not responsible for slowing grammatical acquisition. Parental education and/or income do not alter the rate of grammatical development directly. They do so through SES-related input differences. To understand the nature of learning mechanisms and individual differences in children's rate of learning, researchers must measure input directly without depending on SES as a surrogate for input quantity and quality measures (Miller, 2012). Future research will need to determine if the grammatical input properties characterized in this study are associated with SES and determine whether the relationships found hold after controlling for SES and differences in parents' general lexical diversity.

In conclusion, this research provides new evidence that specific properties of grammatical input affect the rate of the acquisition of finiteness in the third year of life. Input effects stand side by side with biological factors endogenous to the child (Hadley & Holt, 2006; Hadley et al., 2011; Rice, Gayan, Smith, 2009), and in no way replace these effects. An important future goal of research in this field is to understand the ways in which biological, environmental and developmental factors join in determining an individual child's rate of grammatical development. This goal is not only important from a theoretical point of view, but also from an applied point of view. It is our hope that this research contributes to that goal.

References

Benjamini, Yoav, & Hochberg, Yosef. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B, 57*(1), 289 – 300.

Bleile, Ken. (1995). *Manual of articulation and phonological disorders*. San Diego, CA: Singular Publishing Group, Inc.

Bock, Kathryn, & Levelt, Willem. (1994). Language production: Grammatical encoding. In M. Gersenbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). San Diego, CA: Academic Press.

Bricker, Diane, & Squires, Jane. (1999). *Ages and Stages Questionnaires: A Parent Completed Child-Monitoring System*. Baltimore, MD: Brookes.

Chevrot, Jean-Pierre, Nardy, Aurelie, & Barbu, Stephanie. (2011). Developmental dynamics of SES related differences in children's production of obligatory and variable phonological alternations. *Language Sciences, 33,* 180-191.

Fenson, Larry, Marchman, Virginia, Thal, Donna, Dale, Philip, Reznick, J. Steven, & Bates, Elizabeth. (2007). *MacArthur-Bates Communicative Development Inventories*. Baltimore, MD: Brookes.

Fitzgerald, Colleen, Hadley, Pamela, & Rispoli, Matthew. (2013). Are some parents' interaction styles associated with richer grammatical input? *American Journal of Speech-Language Pathology, 22,* 476-488.

Hadley, Pamela, & Holt, Janet. (2006). Individual differences in the onset of tense marking: A growth curve analysis. *Journal of Speech, Language, and Hearing Research, 49,* 984-1000.

Hadley, Pamela, Rispoli, Matthew, Fitzgerald, Colleen, & Bahnsen, Alison. (2011). Predictors of morphosyntactic growth in typically developing toddlers: Contributions of parent input and child sex. *Journal of Speech, Language, and Hearing Research, 54*, 549-566.

Hadley, Pamela, Rispoli, Matthew, Holt, Janet, Fitzgerald, Colleen, & Bahnsen, Alison. (in press). The Growth of Finiteness in the Third Year of Life: Replication and Predictive Validity. *Journal of Speech, Language, and Hearing Research*.

Hadley, Pamela. & Short, Heather. (2005). The onset of tense marking in children at-risk for SLI. *Journal of Speech, Language and Hearing Research. 48*, 1344-1362.

Hart, Betty & Risley, Todd. (1995). *Meaningful differences in the everyday experiences of young American children.* Baltimore, MD: Paul H. Brookes Publishing Co.

Hoff-Ginsberg, Erika. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics, 19*, 603-629.

Huttenlocher, Janellen, Haight, Wendy, Bryk, Anthony, Seltzer, Michael, & Lyons, Thomas. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*, 236-248.

Huttenlocher, Janellen, Vasilyeva, Marina, Cymerman, Elina, & Levine, Susan. (2002). Language input and child syntax. *Cognitive Psychology, 42*, 337-374.

Huttenlocher, Janellen, Vasilyeva, Marina, Waterfall, Heidi, Vevea, Jack, & Hedges, Larry. (2007). Varieties of caregiver speech. *Developmental Psychology, 43*, 1062–1083.

Huttenlocher, Janellen, Waterfall, Heidi, Vasilyeva, Marina, Vevea, Jack, & Hedges, Larry. (2010). Sources of variability in children's language growth. *Cognitive Psychology, 61*, 343–365.

Legate, Julie, & Yang, Charles. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition, 14*, 315-344.

Leonard, Laurence, Camarata, Stephen, Brown, Barbara, & Camarata, Mary. (2004). Tense and agreement in the speech of children with specific language impairment: Patterns of generalization through intervention. *Journal of Speech, Language, and Hearing Research, 47*, 1363–1379.

Miller, Karen. (2012). Not all children agree: Acquisition of agreement when the input is variable. *Language Learning and Development, 8*, 255-277.

Newport, Elissa, Gleitman, Lila, & Gleitman, Henry. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In Catherine Snow & Charles. Ferguson (Eds.) *Talking to children: Language input and interaction* (pp. 109-150). New York, NY: Cambridge University Press.

Pine, Julian, Conti-Ramsden, Gina, Joseph, Kate, Lieven, Elena, & Serriatrice, Ludovica. (2008). Tense over time: testing the Agreement/Tense Omission Model as an account of the pattern of tense-marking provision in early child English. *Journal of Child Language, 35*, 55-75.

Pinker, Steven. (1984). *Language Learnability and Language Development.* Cambridge, MA: Harvard University Press.

Rice, Mabel, Smith, Shelley, & Gayan, Javier. (2009). Convergent genetic linkage and associations to language, speech and reading measures in families of probands with specific language impairment. *Journal of Neurodevelopmental Disorders, 1*, 264–282.

Rice, Mabel, & Wexler, Kenneth. (2001). *Test of Early Grammatical Impairment*. San Antonio, TX: The Psychological Corporation.

Rispoli, Matthew, & Hadley, Pamela. (2011). Toward a theory of gradual morphosyntactic learning. In Inbal Arnon & Eve Clark (Eds.), *Experience, variation and generalization: Learning a first language* (pp. 15-33). Amsterdam, the Netherlands: John Benjamins.

Rispoli, Matthew, Hadley, Pamela, & Holt, Janet. (2009). The growth of tense and agreement. *Journal of Speech, Language & Hearing Research, 52,* 930-944.

Rispoli, Matthew, Hadley, Pamela, & Holt, Janet. (2012). Sequence and system in the acquisition of tense and agreement. *Journal of Speech, Language and Hearing Research, 55*, 1007-1021.

Rowe, Meredith. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language, 35*, 185-203.

Slobin, Dan. (1973).Cognitive prerequisites for the development of grammar. In Charles Ferguson & Dan Slobin (Eds.), *Studies in child language development*. (pp. 175-208). New York: Holt, Rinehart & Winston.

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

Valian, Virginia. (1995). Input and Language Acquisition. In William Ritchie & Tej Bhatia (Eds.).*Handbook of Child Language Acquisition* (pp 497-530). San Diego, CA: Academic Press.

Wexler, Kenneth. (2011). Grammatical computation in the Optional Infinitive Stage. In Jill de Villiers & Thomas Roeper (Eds.) *Handbook of Generative Approaches to Language Acquisition* (pp 53- 118). Dordrecht: Springer DOI 10.1007/978-94-007-1688-9_3

Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)

edited by Chia-Ying Chu,
Caitlin E. Coughlin, Beatriz Lopez Prego,
Utako Minai, and Annie Tremblay

Cascadilla Proceedings Project   Somerville, MA   2014

Copyright information

ISBN 978-1-57473-463-8 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, 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:

Rispoli, Matthew and Pamela Hadley. 2014. Input Effects on the Acquisition of Finiteness. In *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)*, ed. Chia-Ying Chu et al., 121-127. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3073.