Empirical and Theoretical Approaches to Language Acquisition

Empirical and Theoretical Approaches to Language Acquisition:

A Generative Perspective

Edited by

Merle Weicker, Rabea Lemmer,
Andrea Listanti and Angela Grimm

Cambridge
Scholars
Publishing

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Introduction

Merle Weicker, Rabea Lemmer, Andrea Listanti and Angela Grimm

1. About this volume

This edited collection contains 15 papers originally presented at the 15th *Generative Approaches to Language Acquisition* (GALA 15) conference, which took place at Goethe University Frankfurt am Main, Germany, in September 2022. The volume provides a well-selected sample of recent research in the field of language acquisition. The papers in this volume present original research that underwent a peer-review process before publication. The empirical studies in this volume cover various types of learners, ranging from first to second language acquisition over to impaired language acquisition. The research presented provides crucial insights into the foundational questions of language acquisition research: How can we explain that children acquire their language(s), given the complexity of the language system? Are their acquisition paths uniform? Or are there individual differences? Which challenges do children with language or hearing impairment face?

The volume contains research on a variety of acquisition phenomena, covering different areas of the language system. The contributions focus on morpho-syntax, phonology, and semantics in a rich sample of languages like German, English, Spanish, Italian, European and Brazilian Portuguese, Greek, and Romanian. The methodologies employed encompass corpus analyses as wells as experimental on-line and off-line paradigms for the investigation of language processing, language production, and language comprehension.

Despite their diversity, all contributions share a generative perspective on language acquisition. The contributions highlight the importance of theoretically guided experimental work and the fruitful connection between empirical language acquisition research and theoretical linguistics. Language acquisition research in this vein is crucial for our understanding of human cognition. Furthermore, it can offer novel insights into theoretical

linguistics. The volume also reflects the progress made in language acquisition research regarding its methodological standards. By bringing together diverse topics within the same framework, we hope to inspire researchers to continue using modern experimental paradigms to test predictions about typical and atypical language acquisition driven by linguistic theory.

2. Overview of chapters

The book is organized according to learner profiles. The first part ("Monolingual Language Acquisition") contains in-depth studies investigating monolingual acquisition at the word, phrase, and sentence level. The second part ("Bilingual Language Acquisition") comprises studies on different types of bilinguals and second language learners. The third part ("Impaired Language Acquisition") is devoted to studies on children with developmental language disorder and children with hearing impairment. Across these sections, language production and language comprehension are examined using a variety of methodologies.

Part I opens with the chapter *Interpreting Variable Input* by Ana Teresa Pérez-Leroux. By reviewing different case studies, the chapter addresses the intriguing question how children manage grammatical learning under conditions of scarce and opaque grammatical input, sociolinguistically variable grammatical phenomena, and phenomena undergoing diachronic change. Pérez-Leroux argues for contrast inferences as a key acquisition mechanism, which allows children to incorporate noisy input into the developing abstract grammar.

Chapter 2 (*Children Do Not Distinguish 'The Green Leaves and The Yellow Leaves' from 'The Green and Yellow Leaves'*) by Adina Camelia Bleotu and Tom Roeper explores the comprehension of two Romanian coordination structures involving nouns and adjectives. Based on the results of a picture-selection task, Bleotu and Roeper conclude that the syntax-semantics mapping of these structures poses a challenge for 4-year-olds; unlike adults, they do not associate different structures with different interpretations.

Chapters 3 to 6 are devoted to the production of syntactic structures. Chapter 3 (*The 'Growing Trees' Approach and the Development of the Left Periphery in Brazilian Portuguese*) by Miguel Meira and Elaine Grolla investigates developmental steps in the acquisition of Brazilian Portuguese syntax. By looking at longitudinal data produced by a child between ages 1 to 5, the authors provide evidence for the cartographic hypothesis of a

"growing" syntactic tree explaining the order of emergence of syntactic structure.

Maria Lobo, Joana Batalha, Antónia Estrela and Bruna Bragança follow this line of research for the written modality in Chapter 4 (*Syntactic Complexity in Early Written Development*). Lobo et al. elicited written narratives through a sequence of images. Their results indicate that the diversity of connectives that children resort to and the increase in the proportion of subordination processes provide a valuable measure to assess syntactic complexity in early writing. The authors identify three profiles of early writers, based on syntactic complexity measures combined with their mastery of punctuation norms.

The interface to semantics and pragmatics is the topic of Chapter 5 (*Late Acquired Lexical and Syntactic Markers in Italian Rhetorical Questions*) by Maria F. Ferin and Tanja Kupisch. The authors investigate possible acquisition orders of optional morpho-syntactic markers used in Italian rhetorical questions. Four groups of children between six and nine years and a control group of adults were tested with an elicited production task. The results point to optional pragmatic marking as a very late acquired phenomenon, as all the age groups show that the acquisition of this interface structure is still ongoing.

Syntactic parameter setting is discussed in Chapter 6 (*Applying a Constraint Demotion Algorithm to Syntactic Parameter-Setting*) by Kangzheng Gao. Taking a computational approach, the author validates the "Constraint-based Learner" algorithm, which successfully models the learners' transition from a subset to a superset grammar. The results of the test offer solutions for a computational representation of optimality within the framework of Optimality Theory.

The last two chapters in Part I shift the focus to phonology and its role for lexical access. Chapter 7 (*The Role of Phonological Features in Lexical Processing by German-Speaking School Age Children*) by Bénédicte Grandon and Esther Ruigendijk adds a processing perspective. In two studies, the authors examined to what extent the saliency of segments facilitates lexical decision and processing in children and adults. They conclude that larger distinctions between competing words support lexical processing.

Ioanna Kappa (Chapter 8: *On the Reduction of Non-Strident Voiced Fricative Plateau Clusters in Child Greek*) discusses phonological acquisition in Greek. The aim of Kappa's study was to identify factors driving the reduction patterns in plateau clusters of voiced fricatives. Among factors like place of articulation, sonority, positional faithfulness,

and contiguity, only the first factor mattered and reflected featural markedness.

Part II is devoted to children and adolescents who acquire more than one language. The first two chapters in Part II focus on simultaneous and sequential bilingual school age children. Chapter 9 (*Bilingual Children's Explicit Knowledge and its Relationship to Priming Within- and Across-Languages*) by Ioli Baroncini and Jacopo Torregrossa examines the extent to which VSO word-order is accepted in Italian by Greek-Italian bilingual children, considering that this word-order is possible in Greek and dispreferred in Italian. Using an acceptability-judgement task they show that the children accepted VSO sentences in Italian to a low extent suggesting that explicit metalinguistic knowledge may affect priming effects at an implicit, subconscious level.

Chapter 10 (*Does Branching Directionality Influence Bilingual Children's Acquisition of Recursion? A Case Study of English Recursive Adjectives and Possessives*) by Usha Lakshmanan, Deborah Foucault and Tom Roeper presents a study on the comprehension and production of English indirect recursion with a sample of children acquiring typologically diverse languages. First findings suggest that the acquisition of recursive structures may be influenced not only by the branching direction at the local phrasal level, but also by the branching direction at a more global level of the two languages being acquired.

Two chapters that examine older second language learners complement Part II. Magdalena Wojtecka, Roland Nenno and Stefanie Haberzettl (Chapter 11: *Are Adolescent L2-learners with L1 Persian Sensitive to the Strong Relationship between Verb Placement and Finiteness in German Main Clauses?*) investigate the acquisition of finiteness and verb placement by analyzing spontaneous speech data of 11 adolescent L2-learners of German. Results show that while adolescent L2-learners almost never used incorrectly inflected verbs, they encountered greater difficulties with verb placement in German main clauses. This finding indicates a significant difference compared to adult L2-learners of German.

Elisa Di Domenico (Chapter 12: *Over-Use of Overt Subject Pronouns in L2 Italian: The L1 Counts!*) used an elicited production task to investigate the use of overt subject pronouns in L2 Italian by near-native speakers whose L1 was either a pro-drop (Greek) or a non-pro-drop language (English). Contrary to the findings of previous studies, results show that the L1—in addition to bilingualism *per se*—affects the overuse of subject pronouns in the L2.

Part III completes the volume by taking into account children with language impairment. Chapter 13 (*The Acquisition of Object Clitics in Children with a Phonetic-Phonological Disorder: Evidence from Comprehension and Production*) by Elena Pagliarini, Maria Milan and Fabrizio Arosio investigates developmental language disorder. The authors aim to tap into the origin of the difficulties in clitic production observed in children with developmental language disorder. More precisely, they examined whether the deficits in Italian-speaking children are caused by morphosyntactic or by prosodic limitations. The authors report that children diagnosed with a grammatical deficit exhibited greater difficulties in clitic comprehension and production than those with phonological impairments. The authors argue for a distinction of subtypes of developmental language disorder in order to tailor intervention.

The volume closes with two chapters addressing hearing impairment: In Chapter 14 (*Verbal Agreement Inflection in Children with Hearing Loss – The Role of Hearing Device, Perceptual Prominence and Syllable Structure*) by Martina Penke and Laura Gerkens, treatment-specific effects of hearing aids and cochlear implants in 20 German pre-school children are examined. Their analyses of experimental as well as spontaneous speech data focuses on the acquisition of high- and lower-pitched consonants. Results show no evidence for treatment-specific effects. The data rather indicate problems with specific agreement markers. The authors suggest that perceptual factors related to the acoustic properties of critical consonants and their syllabic position affect the production of verbal agreement markers in children with hearing loss, independent of their hearing device.

Chapter 15 (*A'- and A-movement in Portuguese Cochlear Implanted Children: The Effects of Length of Language Exposure*) by Mara Moita reports data on the acquisition of syntactic structures with A'- and A-movement. The analysis of comprehension data from 47 Portuguese children with hearing impairment and cochlear implants shows that these children do not exhibit a general syntactic deficit. Rather, the data suggests a delay in the acquisition of *wh*-questions and relative clauses depending on the length of language exposure and the age of cochlear implant activation.

PART I:

MONOLINGUAL LANGUAGE ACQUISITION

CHAPTER 1

INTERPRETING VARIABLE INPUT[1]

ANA T. PÉREZ-LEROUX

Abstract

When faced with variable or noisy input, what do children do? Children manage grammatical learning under conditions of scarce and opaque grammatical input, sociolinguistically variable grammatical phenomena, and phenomena undergoing diachronic change. Minimalist approaches, by reducing grammar variation to differences in the encyclopedic and functional inventories, reframe grammar learning as lexical learning. Learning what words mean (also known as solving the mapping problem) depends on two core inferential mechanisms: syntactic bootstrapping, and contrast. Learners attend to probabilistic distributional information and use it to support inferences about the meanings of forms, and the identification of contrast sets to which contrast inferences can apply. The developmental and linguistic literature both contain multiple proposals about contrast inferences of various types. These can be unified into a single mechanism, Gricean in nature, from which listeners derive meanings (or relations between forms) by reasoning over what speakers have said and what they could have said instead. Contrast inferences can extend beyond word learning and scalar implicatures and serve as foundation of learning across the various levels of language knowledge. Upon identifying formal differences learners can follow a multilevel hierarchical array of mapping decisions where contrasts are first interpreted in terms of grammaticality, and subsequently in terms of truth-functional values, perspectives on situations, pragmatic enrichments, or sociolinguistic attributions. By committing learners to interpret variation at one level or another, contrast inferences represent a key engine for children's development of abstract

---

grammars and their mastery of the rich variation existing in their language communities.

1. Introduction

When faced with variable or noisy input, what do children do?

This question is timely. Johnson and White (2020) argue for placing variation at the core of the study of language acquisition:

> […] language is by its very nature variable, and much of this variability is informative, as it is (probabilistically) governed by a variety of factors— including linguistic context, social or cultural context, the relationship between speaker and addressee, a language user's geographic origin, and a language user's gender identity. […] *we anticipate a paradigm shift in the way many language researchers conceptualize the challenge of early acquisition.* (Johnson and White 2020, 1)

Developmental sociolinguistics as a project situates language acquisition research within the context of socio-cognitive development. To undertake this important goal requires we reconsider some implicit assumptions in the field, generative and otherwise, which have hindered research on the acquisition of variable phenomena. I will start by reviewing some basic assumptions about learning and variable input, and then consider work illustrating how children learn under various input conditions: scarce and opaque grammatical input, variable grammatical phenomena, and aspects of grammar undergoing diachronic change. The minimalist assumption that the lexicon is the locus of parametric variation commits us to approach all grammar learning as lexical learning (Chomsky 2001). This requires us to revisit two core learning mechanisms: one that relates words to sentences (the Syntactic Bootstrapping Hypothesis) and another that relates words to other words (Principle of Contrast). I conclude by proposing that some general form of contrast inferences is central to all language learning, and can serve to integrate the sociolinguistic and the grammatical dimensions of language acquisition.

2. A bias towards categorical phenomena in acquisition

We can start by admitting we struggle with variability. I will support this claim with three anecdotes. Responding to a paper about the interactions between grammatical development and language change (Cournane and Pérez-Leroux 2020), an anonymous reviewer incredulously asked "Do you mean to say that children are exposed to input from different stages of the

evolution of the language?" Another time, an audience member criticized the results of an elicited production study: "You cannot interpret these data because your adults only succeed 60% of the time." An anonymous reviewer wrote about a report of individual response patterns: "The authors could not find group level effects, so they had to resort to individual analyses." To these remarks one could respond "Yes: language communities are likely to contain speakers whose speech represent different stages of diachronic evolution. Yes, we can study phenomena where there is no single dominant adult response. And, no, there is nothing wrong with individual analyses. Language lives in the mind, and minds often hold contents and representations that are similar but different."

Some will be tempted to blame Chomsky for our current mistrust of variability. After all, he did propose the study of an ideal speaker-hearer, living in a perfectly homogeneous linguistic community (Chomsky 1965, 3). A disconcerting choice of words, but the goal was not to endorse ideal norms, but to work with a smaller version of a language, in order to get started building some syntactic formalizations. Few semanticists complain that Montague (2002) limited the scope of semantics when he wrote a model for a fragment of English. Acquisitionists' mistrust of variability likely has a closer source. Brown (1973), the field's founding father, proposed the very useful 90% success criterion in obligatory contexts as acquisition criteria. Everyone followed suit, proceeding to concentrate on phenomena where adults perform at 100%, and mostly avoiding the other stuff. Hammers are great tools, but they draw our attention to nails. Pinker's views about the basic properties of the language learning were more problematic. Pinker declared uncontroversial that learning involved no memory for sentence form: "First I assume that the child has no memory for the input other than the current sentence-plus-inferred meaning and whatever information about past inputs is already encoded into the grammar at that point." (Pinker 1984, 31). Today we know that children have particularly strong syntactic priming effects, which essentially is implicit memory for sentence form (Bencini and Valian 2008). Under the view that learners do not store sentence structure in memory there is no room for incorporating statistical information about structures into the learning process. These biases limit the type of phenomena that we are willing to investigate.

Recent conceptual shifts have opened the door to the study of variability in acquisition. The basic goal of generative approaches remains the same: to understand how grammar, the mental symbolic systems that underlie human language, develops in the minds of children. Grammars exist within a delimited space of formal possibilities. Current generative grammar differs crucially from the classic Principles and Parameters model in that

parameters of syntactic variation are more narrowly defined; so is the scope of learning. According to the Uniformity Principle, languages are seen as fundamentally uniform, with variation restricted to the lexical and functional inventory of a grammar, their visible properties, and distributions (Chomsky 2001). Put broadly, what is learned is not syntax but the units that serve as building blocks of syntax. Thus grammar learning is lexical learning, which is different from saying that all grammar is lexical (Ambridge, Pine, and Lieven 2014).

If so, what is the role of Universal Grammar (UG) in grammar learning? A theory of UG is not a theory of acquisition (Lidz and Gagliardi 2015). UG structures deductions about sentences that fall outside of experience, defining both the representational boundaries of language (Moro's 2016 impossible grammars; Longobardi's 2018 schemata) and the evidence learners use to infer particular properties. Many key questions remain: how the learner uses available evidence, and fills in the details of grammar, and crucial to our discussion here, what happens if the target of learning is not categorical, but probabilistic behavior?

Early generative efforts to study variability addressed learner-internal variability, not target language variability, which was considered problematic in the context of parameter setting.[2] Probabilistic behavior was modelled as the retention of different parameters within the same grammar (Roeper's 1999 Universal Bilingualism; Yang's 2002 variational learning). At that point, syntactic learning was characterized as the selection of target parameters and deselection of alternatives.[3] Yang (2016) is one of the few works to directly address language internal variability in syntactic learning. Yang returns to the classic formulation of learning as rule formation, and proposes concentrating on the input distribution of regular patterns and exceptions. He assumes that children readily infer rules from detection of similarities, but only adopt a rule as productive given a certain balance between rule-compliant cases and exceptions. He offers the Tolerance principle, a precise formalization of this balance, which effectively accounts for cases of overgeneralization in various domains and has been extended beyond the initial corpus-based studies (Li et al. 2021; Henke 2023) to promising initial behavioral tests (Li and Schuler 2023; Schuler, Yang, and Newport 2016). As pointed out by Lidz and Gagliardi, the statistical effects of input are not incompatible with theories that emphasize innate linguistic

---

[2] Much work in the generative framework centered on the optionality of certain functional elements (Paradis and Genesee 1997; Rizzi 2008; Pérez-Leroux, Pirvulescu, and Roberge 2018, among many others).
[3] See also the proposal in Fodor (1998), where syntactic parameters are defined as bits of structure, or treelets.

knowledge. These authors propose three required components for an acquisition model: i) the perceptual encoding of input, ii) the deductive system that defines both the space of possible representations and what counts as relevant data, and iii) the inference engine that connects the intake to a specific grammar (Lidz and Gagliardi 2015, 5).

Current interest in using probabilistic information in grammar acquisition owes much to Lila Gleitman's work on verb learning (Lidz 2020). Gleitman's work demonstrated that limited exposure to ambiguous contexts, supported by syntactic information (i.e., the contexts in which a word appears), is sufficient for learners to learn verb meanings. The proposed inferential mechanism, known as syntactic bootstrapping, is essentially probabilistic. According to Chomsky, her results are central to the whole generative enterprise, because they demonstrate that learning requires "prior constraints, C. S. Pierce's "abductive principle" that puts a limit upon admissible hypotheses" (Chomsky 2020, xv). For Gleitman, the deductive and the inferential are both essential components of learning:

> The learner who can observe everything can drown in the data. Two kinds of capacity and inclination rescue the learning device. The first is a general learning procedure that can extract, combine, and coordinate multiple probabilistic cues at several levels of linguistic analysis (in the spirit of many machine-learning and constraint-satisfaction proposals […]. However, for such a probabilistic multiple-cue learning process to work at all it requires unlearned principles concerning how language realizes conceptual structures; and similarly unlearned principles for how these mappings can be discovered from their variable and complex encoding in speech within and across languages […]. (Gleitman et al. 2005, 28)

Syntactic bootstrapping shows how learners can use probabilistic distributional information to map out form and meaning. Many question remain about how children deal with variability in the input. When is variability noise, and when is it useful information? When faced with some variant, how do children know that it is part of the system to be learned (competence) or the result of performance? In sections 3–5 I explore what happens when evidence for some aspect of morphosyntax to be learned is i) scarce or opaque; ii) when it reflects sociolinguistic variability?; and iii) when variability reflects ongoing diachronic changes in the target language?

3. Faced with scarce or opaque input

The deductive component of language shines brightest in poor input conditions, where the child is forced to enrich sparse data with

generalizations and deductions. I will review two such cases, the acquisition of recursive nominal embedding, and the problem of opaque clues about the finer points of Korean clausal structure.

3.1 Recursion across languages

Recursion, the ability to generate phrases self-embedded in categories of the same type, has been described as a language universal (Hauser, Chomsky, and Fitch 2002). That does not mean that all languages have the same types of recursive structures, or that all forms of embedding can be used recursively, as pointed out by Roeper and Snyder (2004) for German possessive -*s*, which is not used iteratively (*Marias Haus* 'Maria's house', but not **Marias Vaters Haus* 'Maria's father's house'). This suggests we must separate learning a simple modification rule (i.e., one level of embedding) from learning that a rule is recursive (two or more levels of embedding). Roeper and Snyder also observed that in corpus data, parentals use recursive embedding infrequently. For nominal modification, the input for recursion is extremely poor, as shown by an analysis of the entire corpus for two Spanish children Juan (over 21,000 words) and Maria (over 80,000 words) (Pérez-Leroux 2022). Parents use simple modification infrequently (less than 150 in each corpus), and recursive modification was vanishingly rare: three cases were found in each corpus.

How is recursion learned, if it is so rare? There are two possibilities: direct exposure to an instance of recursive embedding, or inference from the simple embedding case. Widmer et al. (2017) studied synchronic and diachronic variation in forms of embedding to examine if recursion could become extinct from a language. They found that while new forms of recursion emerge and disappear rapidly in diachronic time, all languages retained some form of recursion. Li et al. (2021) had proposed that productivity of the simple rule is a condition for recursion, and can be learned from the distributional properties of input at first level of embedding. Specifically, they proposed that in a configuration N1-linker-N2, evidence of the structural substitutability of N1 and N2 is sufficient for establishing the productivity of the embedding rule, and that such evidence is distributionally available to the child.

In a series of studies, we compared production in five languages that varied in branching directionality and homogeneity, and diversity of forms used in recursive nominal embedding (German, English, French, Spanish, and Japanese). We elicited nominal recursion in children ages 4-6 to investigate whether the complexity of the inventory of embedding forms had an impact on the age of acquisition of recursive modification. For

instance, Japanese has two nominal embedding forms, shown in (1): the particle *no*, and relative clauses (Pérez-Leroux et al. 2023a). Both are left-branching and recursive. In contrast, German has a diverse set of embedding strategies, some right-branching others left-branching. This yields a broad range of configurations for level 2 modification, partially illustrated in (2). In Pérez-Leroux et al. (2022, 65–66) we found four different morphosyntactic forms for possessive embedding, including genitive case, possessive -*s*, relative clauses, and various prepositions. The most common were *von* ('from') (as in (2) and *mit* ('with').

(1) a. Japanese Recursive NO
         Taro   no   otoosan   no   Kuruma-o
         Taro   GEN   father   GEN   car-ACC
         'Taro's father's car'
     b. Japanese Recursive Relative Clauses
         Orenji   no   tana   no   ue-ni   aru
         orange   GEN   shelf   GEN   on-in   be
         koppu   no   naka-ni   haiteru   haburashi
         cup   GEN   inside-in   be   toothbrush
         'the toothbrush that is inside of the cup that is on the orange shelf'

(2) German Recursive Dative Possession (Pérez-Leroux et al. 2022, 69)
     der   Ballon   von   dem   Affen
     the.NOM   balloon   of   the.DAT   monkey
     vom   Clown
     of.DAT   clown
     'The clown's monkey's balloon'

    We found no evidence that inventory complexity impacts children's use of recursive nominal embedding (Pérez-Leroux et al. 2023c). Having to learn more embedding forms, and that some forms might not be recursive, does not lead to age differences across languages: children can use recursive nominals earlier but become productive after the age of five. When faced with diverse forms, German children overgeneralize the less restrictive/more productive *von + dative* at the expense of other embedding strategies. They use Saxon possessive -*s* frequently but do not overgeneralize it.

    Language type did not seem to confer advantages or disadvantages in children's emerging productivity with recursive structures. This suggests that lower frequencies of input for recursion is not a problem, because

children learn recursion from the more common simple embedding case. The distributional patterns of simple embedding are learned early, but generalized to recursive structures only as children's capacities for sentence planning mature (McDaniel, McKee, and Garrett 2010). This explains why we observe individual variability but crosslinguistic comparability in the developmental timeline. When uncertain as to whether a marker allows recursion children use a different structure but do not fail to learn to produce recursive descriptions. Diachronically specific forms of recursion are unstable because transmission may fail due to the scarcity of direct evidence. Nonetheless, as Widmer and colleagues demonstrate, recursion *per se* remains robust because other forms become generalized to recursive contexts.

3.2 Korean scope

Low frequency is not the only input challenge child learners face. Some types of structures are ambiguous on the surface, such as the underlying structural position of Korean verbs. Descriptions of Korean disagree as to whether the sentence final verb is in low or high position (Han, Musolino, and Lidz 2016). There are no clear distributional cues to settle the matter. Quantifier interpretation is used as evidence, since Korean has "frozen scope", where quantifiers are unambiguously interpreted in their hierarchical position. Objects always raise out of VP; the negation+V complex may or may not raise; if it raises, negation has scope over the object. Thus, the interpretation of negation relative to a quantified direct object serves to diagnose whether the verb remained in V (low position), as in (3b), or has raised to T (high position), as in (3c).

(3)   a.   Subject Object-QP Neg V
             Khwuki monste-ka   motun   khwuki-lul
             cookiemonster-NOM   every   cookie-ACC
             an   mek-ess-ta.
             NEG   eat-PST-DECL
             'Cookie monster didn't eat every cookie.'
       b.   Verb + Neg remains low, Quantified object has scope
             over negation ('all did not') Every > Neg
             'None of the cookies were eaten.'
       c.   Verb to T, Negation has scope over the quantified
             object ('not all did') Neg > Every
             'It is not the case that CM ate every cookie, he might
             have had some.'

Han, Musolino, and Lidz (2016) tested subject and object scope interpretation with respect to negation. The interpretation of subject quantifiers did not vary, always scoping over negation. Object quantifiers were interpreted variably, across individuals. Speakers had one or the other scope pattern, and their intuitions were both consistent and persistent over time. Descriptions of Korean contradict each other because there are two types of speakers, hiding behind the surface ambiguity of Korean sentences. Crucially, these authors also found that children were as categorical as adults, but their scope patterns did not correlated to their parents'.

Resilient properties of language develop in the absence of robust experience. Recursion is rare but children develop it at comparable pace in languages with diverse and complex inventories as in languages with simple inventory. Korean scope shows that individuals can infer categorical systems from input that is variable at the community level. It seems that children succeed even when the experience is insufficient by either extending generalizations from one domain to another (from simple to recursive embedding), or by randomly selecting one categorical option when input is ambiguous. This suggests that under sparse or opaque input conditions children's behavior is more categorical than their experience, not less.

4. Faced with variable input

4.1 Children can approximate variable input

How do children learn linguistic variability that is socially-conditioned (i.e., dependent on speaker (age, gender, class) or context (politeness, power imbalance, etc.))? The sociolinguistic literature suggests that for many properties of language, children readily match adult patterns of sociolinguistic variability. Young children reproduce gender differences in speech early (Fung, Schertz, and Johnson 2021), and associate variation to speaker gender in artificial language settings (Samara et al. 2017). In a study of the Buckie dialect of Northern England, Smith, Durham and Richards (2013) note that 2- to 3-year-olds match caregivers' rates and stylistic distribution of local diphthong variant. Nardy, Chevrot, and Barbu (2014) found that kindergarteners from families who speak the standard variety adapt to regional phonetic variability in French, depending on their interaction with children from the community. Miller and Cardenas (2020) found that children match adults' patterns of final -*s* lenition in Dominican Spanish along the lines of gender and formality.

Nonetheless, early mastery of socially-variable morphosyntax is not guaranteed. Smith and Durham (2019) found different behavior for two variable phenomena in Buckie English: present tense person agreement and regional versus standard forms of negation. Younger children learn the two negation forms sequentially: they start with the standard *not* and only later use the local form *na*. Situational context constrains variant selection in older children and their caregivers, but not in younger children. In contrast, for person agreement, children acquire the formal split early (NPs vs. pronouns), and match caregiver variable patterns of selection across lexical verb types and clause types (interrogatives vs. declaratives). In the case of variability with Spanish differential object marking (DOM), Callen and Miller (2022) found that children share many of the same features with their caregivers, but younger children were insensitive to the association of DOM and specificity. In some situations, researchers have observed both simplification of a system and maintenance of alternating variants, even within the same phenomenon. Exposed to the unpredictable but stable variability of the two-gender system of Fering, a North Frisian dialect, most of the children can replicate adult patterns of variability; while other children, particularly those with less exposure to the Fering variety, simplify the system to one grammatical gender (Hendricks, Miller, and Jackson 2018).

Shin and Miller (2021) propose a developmental sequence for (sociolinguistically) variable forms, where contextual overlap has to be learned. They hypothesize that one form emerges first, followed by complementary distribution when the second form is acquired. Eventually, the overlap between forms in variable contexts increases to adult levels. They also note that learning mechanisms discussed in the developmental literature—including regularization, the Mutual Exclusivity Principle, and the Principle of Contrast—must also be involved in explaining the developmental trajectory of socially-conditioned variable phenomena (Shin and Miller 2021, 23).

4.2 Children can be delayed due to phonetic variability

Phonetic context and phonetic variability impact children's performance with grammatical markers. Miller and Schmitt (2012) demonstrate how degree of phonetic variability across regional varieties can impact age of acquisition of a morpheme. They tested the socially-stratified patterns of final -*s* lenition in children learning three varieties of Spanish. Independently from socioeconomic status, children learning the varieties with more lenition are over two years behind in plural production, and have

poorer comprehension of plurals. Similar results have been found for English (Davies, Rattanasone, and Demuth 2017).

Our group has recently undertaken a study of the impact of vowel variability in the acquisition of Spanish gender in bilinguals, a feature encoded in most nouns in word-final unstressed low and mid vowels. This phonetic context is sensitive to contact-induced neutralization via influence from the reduced inventory of unstressed vowels in English (Pérez-Leroux et al. 2023b). We observed individual differences in phonetics, with some children demonstrating distinct vocalic spaces for word-final -*e, -a, -o*, while others produce centralized, overlapping vowels. We also observed some evidence of grammatical reorganization. Overall accuracy was high, as the majority of the children were at ceiling for gender. However, one fifth of the children gave no evidence of having grammatical gender. The individual patterns were diverse: some had a feminine bias, where they correctly used more feminine words overall (compared to the other children) and also had errors with masculine. We also found children who created a novel genderless determiner *da*, who used masculine default, and who had inconsistent-patterns, using both genders variably for the same nouns, as in Fering. Crucially, while we observed a range of individual phonetic and grammatical performance, we found no association between vowel centralization and loss of the gender distinction.

Phonetic variability in morphosyntax leads to acquisition delays in some cases but not others. Variability itself is not always problematic; only sometimes. For some aspects of language, young children match caregivers' patterns of variability, leading to the conclusion that children attain sociolinguistic competence early. In other domains children learn some variants first, and the alternations later. Unfortunately, the emergent literature on learning variation fails to separate the types of conditioning factors (linguistic context, speaker-based, social/discourse context, and even the absence of conditioning in unpredictable variation). The developmental path of learning variation might be determined not only by the frequency of competing forms and the strength and type of association (distributional vs. semantic), but also by the type of conditioning factor.

5. Phenomena undergoing variation and change

Interesting questions arise from studying phenomena undergoing diachronic change. Is variation related to in-progress language change treated by children as noise? Are children sensitive to age-related variation, as other sociolinguistic phenomena? Generative syntacticians favor the notion that children drive language change (van Gelderen 2011). Cournane (2014)

argued that regular patterns of cyclic change (such as modal verbs, which start as lexical verbs, become root modals, shift to epistemic interpretations, and eventually disappear and fall into disuse) result from specific child learning biases. Cournane (2014) found preliminary evidence for children advancing the grammaticalization cycle of some English modal verbs. Cournane and Pérez-Leroux (2020) found that at age five children advance the epistemic interpretation of *must*, relative to adults. Subsequent work confirms children in other languages are biased for epistemic overgeneralization of root modals, as predicted (Cournane and Veselinović 2022)**.**

Cycles are a special case of language change. In most cases, language change consists of competition between forms. Labov (2001) proposed that adolescents are the main vectors of diachronic changes, choosing some variants over others. According to Labov, children first learn sociolinguistic variation. Upon reaching adolescence, they choose to advance some phenomena change, a phenomena called incrementation. There is evidence in the sociolinguistic literature that adolescents have higher rates of use of some innovative forms (Labov 2001, 2007; Tagliamonte and D'Arcy 2009; Holmes-Elliott 2016; Denis et al. 2019; among others).

In recent work, Hall (2020) and Hall and Maddeaux (2020) compare the relative contribution of pre-adolescents and children to one case of ongoing phonetic change. Rates of fronted realization of /u/ are increasing across North America, and particularly in Canada. The variant is favored in post coronal contexts (i.e., in words such as *tooth* or *soon* versus *food* or *boot*). Both pre-adolescents and children use fronted /u/ at higher rates than adults, but this is more pronounced in younger children, who overextend /u/-fronting to non-coronal contexts. Another case of diachronic change is the increase of possessive -*s*, relative to *of*. The alternation is governed by animacy, relation type, possessor length, and phonological ending. Hall and Pérez-Leroux (2022) found that children have the same lexical-semantic associations as adults, but choose *of* at higher rates, with a peak for genitive preference in ages 4 to 6, followed by retrenchment to adult baseline for the middle group of children, and a second increase as children approach adolescence. These two studies support Hall's (2020) proposal that linguistic change has two distinct sources: overgeneralization and incrementation, processes that have different cognitive bases.

Before mastering the features that regulate an alternation, younger children address (diachrony-related) variation by overgeneralizing one form. This is a common acquisition scenario, as shown by Spanish DOM. Adolescents have already learned the inventory of forms. Thus their overselection of certain variants must have a different basis, potentially

related to identity construction. Sankoff (2018) shows that variant selection can continue to change during the lifespan. Aging adults in her study demonstrated developmental trajectories for different phenomena. In some cases, older adults maintained the same pattern of variation they had in their youth. For others, the same older adults kept pace with the evolving choices of younger speakers. Sankoff also observed retrograde change, when older adults reversed overall trends of quantitative change, choosing instead to sound like the elders of their youth.

We are reminded of Kroch's (1989) idea that language change can be modelled (as in population biology) in terms of grammar selection. Speakers can select their linguistic affiliation, favoring some variants over others. Nonetheless, however skilled children may be at detecting and learning variation, they do not have immediate access to all the features that condition variation, and are therefore prone to overgeneralization. One may observe that semantic features—which are learned by less-direct means— are often involved when children shift the balance between forms.

6. Enriching the toolkit: Contrast inferences

So, how do children learn variable properties of language? Which properties of the input enter the grammar and which are discarded as noise? The input contains both variation (changes or differences in condition) as well as variability (inconsistency). The child does not know which is which: different phonetic realizations may (or may not) mean different markers; different words or morphemes may (or may not) reflect different meanings, registers or sociolects. Some specific capacities must be involved, including the abilities to:

(i)   Immediately detect formal differences;
(ii)   Track distribution of forms, assess their predictability, and use this information to segment speech into parts and sort out how these parts are combined, and
(iii) Use these distributions to help map contrasts into grammar and meaning.

Learners use distributional information to make decisions about what contrasts enter the grammar; about the scope of use of grammatical forms (rule vs. exceptions; idiomatic vs. systematic) (Yang 2016; Radulescu, Wijnen, and Avrutin 2019), and about which level to represent a given contrast. Any formal alternation can be interpreted as relevant to

grammatical possibilities, semantic differences, contextual/pragmatic differences, or as differences between types of speakers.

Accepting that grammar learning is lexical learning entails we can use the results from the study of vocabulary development to understand grammatical learning. Children do not learn words gradually, across situations. Instead, they are thought to make immediate inferences in word learning, subject to confirmation in subsequent encounters (Trueswell et al. 2013). To make these inferences, learners pay attention to the context where new words are used, giving primacy to sentence context (syntactic bootstrapping). Simultaneously, they contrast new words to known words. Historical formulations of how contrast supports early word learning include Markman and Wachtel's (1988) Mutual Exclusivity, a lexical assumption that new nouns are interpreted by learners to refer to mutually exclusive object categories, and Clark's (1987) Principle of Contrast, a pragmatic-based assumption that differences in form imply differences in meanings. These lexical principles are extensions of the classic Gricean insight that speakers derive meanings by reasoning over what speakers have said and what they could have said instead (Grice 1975).

Proposals based on contrast inferences are ubiquitous in language analysis. In morphology, we have the blocking principle, which posits that irregular forms block the application of a rule (Aronoff 1976). In semantics, there is the Principle of Maximize Presupposition (competition of forms according to their presuppositions; Heim 1991), and scalar implicatures in pragmatics (attribution of non-literal meanings on the basis of comparing alternatives) (see Schlenker 2012, who proposes that the former is a subcase of the latter). Contrast inferences have been implemented as principles of grammatical learning, as in Wexler and Culicover's (1980) Uniqueness Principle (one deep structure leads to one surface structure), or Yang's (2016) Tolerance Principle (about quantitative competition between rules). All these principles are predicated on implicit comparisons between forms.

Contrast inferences might be relevant to language learning across language domains beyond the lexicon. However, research on children's pragmatic abilities also indicates that children might have difficulties with scalar implicatures, for instance failing to extract "not all" inferences about statements such as *some of the animals are sleeping* (e.g. Papafragou and Musolino 2003; Chierchia et al. 2001; Noveck 2001). These failures have received various explanations, as resulting from pragmatic deficits, insufficient knowledge of the scales of lexical comparison, or limitations in executive function that limits the ability to hold in attention the relevant alternatives.

The presence of general difficulties with scalar comparisons presents a challenge for the view that contrast inferences play a central role in

grammatical learning. Therefore, these results deserve close inspection. Some recent studies suggest that children can derive some implicatures. Barner, Brooks, and Bale (2011) show that children fail with lexicalized quantifier scales (*all*/*some*), but succeed with sentences containing contextual alternatives (*only the cat and the dog are sleeping*). Thus, some inferences are accessible to children, but other inferences present specific challenges. Foppolo et al. (2021) further find success with scalar implicatures correlates to Theory of Mind development, which indicates these implicatures have high executive function demands. Skordos and Papafragou (2016) note that presence of explicit lexical alternatives help children generate scalar implicatures. Their study controlled whether participants heard *all* before judging *some*. Lexical support improved performance, but only when the stronger scalar element was perceived to be relevant. In contexts where relevant alternatives are established by using another quantifier, such as *none*, explicit mention of the alternative was not necessary. While we need to learn more about what moderates children's limitations and successes with implicatures (Bill et al. 2021), it seems clear that children are able to augment sentence interpretation by comparing potential alternative sentences. Importantly, this approach has been extended to the study of grammatical markers, as shown by Tieu et al. (2020) on the acquisition of plural semantics.

Contrast inferences encompass the various implicit assumptions that listeners and learners make about the expressive choices of speakers. No parallel has been drawn between sentence-level implicatures and lexical learning principles, despite their shared Gricean foundations. If children can compare statements to refine their interpretation, and compare words to decide their meanings, can we not apply it to the overall process of building a grammar? We should. I propose that contrast inferences apply at different levels in the organization of language. Because these levels can incorporate extralinguistic dimension, the approach can be extended to the question of learning from variable input. Any detectable distinction of forms can be interpreted along a hierarchy of levels:[4]

(4)   Hierarchy of mapping levels
        Level of contrast   Output
    a. Lexical   Distinct lexical entries
        cat/dog
        Morphosyntax   Grammaticality intuitions
        *I sing/*Me sing*

---

[4] I leave aside the phonological level, which does not directly map onto meaning and can be learned distributionally.