How Nature Meets Nurture: Universal Grammar and Statistical Learning

Jeffrey Lidz and Annie Gagliardi

Presented by: Alandi Bates

2 Theoretical Frameworks

- Input-driven: learning is essentially a form of memory either essentially statistical or biased in extra-linguistic conceptual cognition
- Knowledge-driven: learning as inference where the learner is endowed with a system of knowledge, typically called Universal Grammar (UG) that specifies a possible space of grammars.

2 Theoretical Frameworks

- Input-driven: these approaches share the perspective that generalizing beyond the input is driven by properties of the learner that are not fundamentally linguistic in nature.
- Knowledge-driven: domain specific representations provide the foundation for generalizing beyond experience and experience provides the information that is relevant to the identification of those representations.

Where these 2 perspectives differ

- How linguistic representations develop
- The role that input plays in learning

How linguistic representations develop

- Input-driven
  - Abstract linguistic representations are arrived at by a process of generalization across specific cases
  - Abstract representations are therefore the output of a learning process that goes from specific to general

- Knowledge-driven
  - Abstract linguistic representations are part of the inherent linguistic capacities of the learner
  - They aren't arrived at by generalization
  - They use the input to identify the abstract representations apparent in the surface form of the exposure language

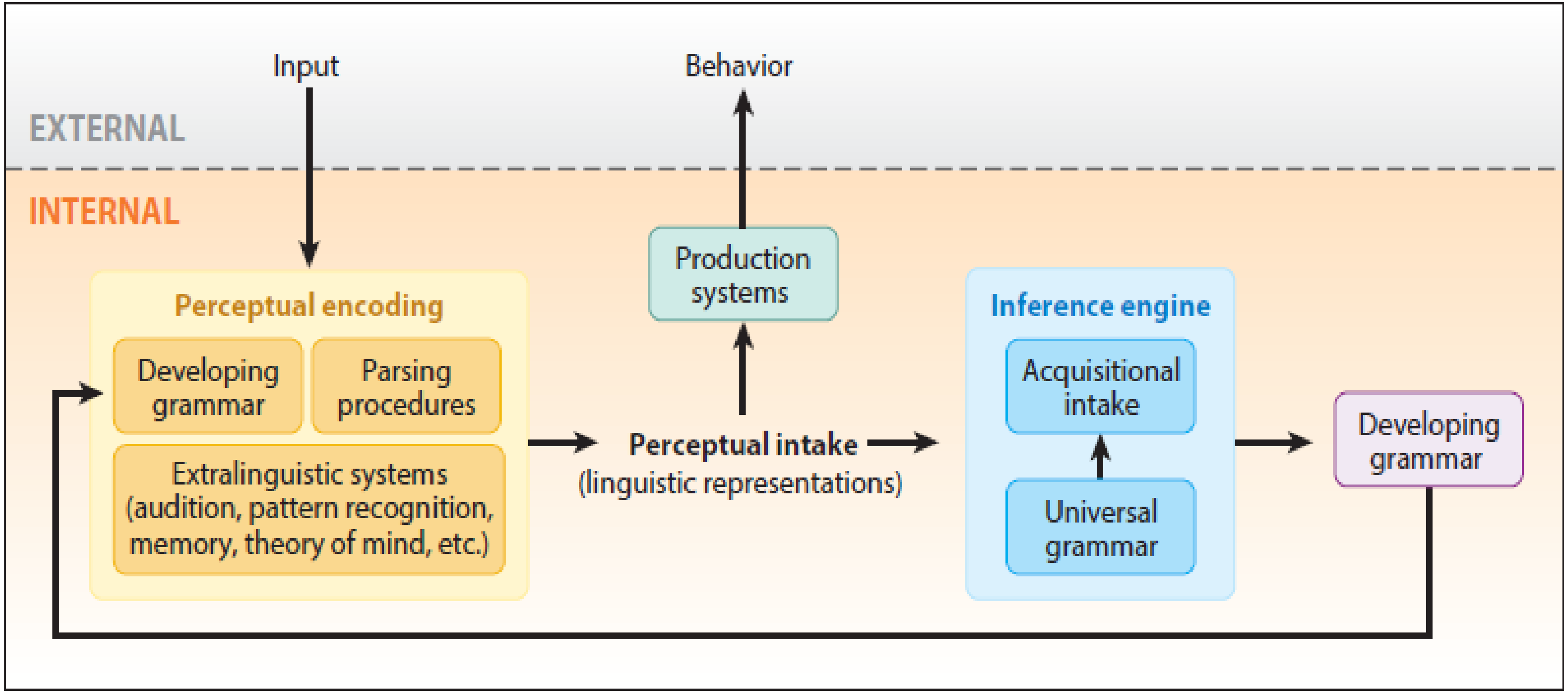
The role that input plays in learning

- Input-driven
  - What is learned is a recapitulation of the inputs to the learner
  - Acquired representations are a compressed memory representation of the regularities found in the input
  - New inputs evoke past experience via some metric of similarity
  - New sentences are possible to the extent that they are similar to past experiences
- Knowledge-driven
  - The learner searches the input for cues to aid in choosing an appropriate abstract representation
  - Which they themselves are innate and define the cues that learners use to identify them
  - Once identified what the learner carries forward is a representation that may bear no obvious relation to the input that triggered it
  - This representation allows for specific generalizations that may not be at all similar to the experiences of the learner

Where these 2 perspectives do NOT differ

- Input plays a significant role in language acquisition.
- Acquisition of a particular grammar is guided to a large extent by experience
- However, where we do expect to see differences is in the relation between the features of experience and the acquired linguistic knowledge

Lidz and Gagliardi (2014) model of Language Acquisition

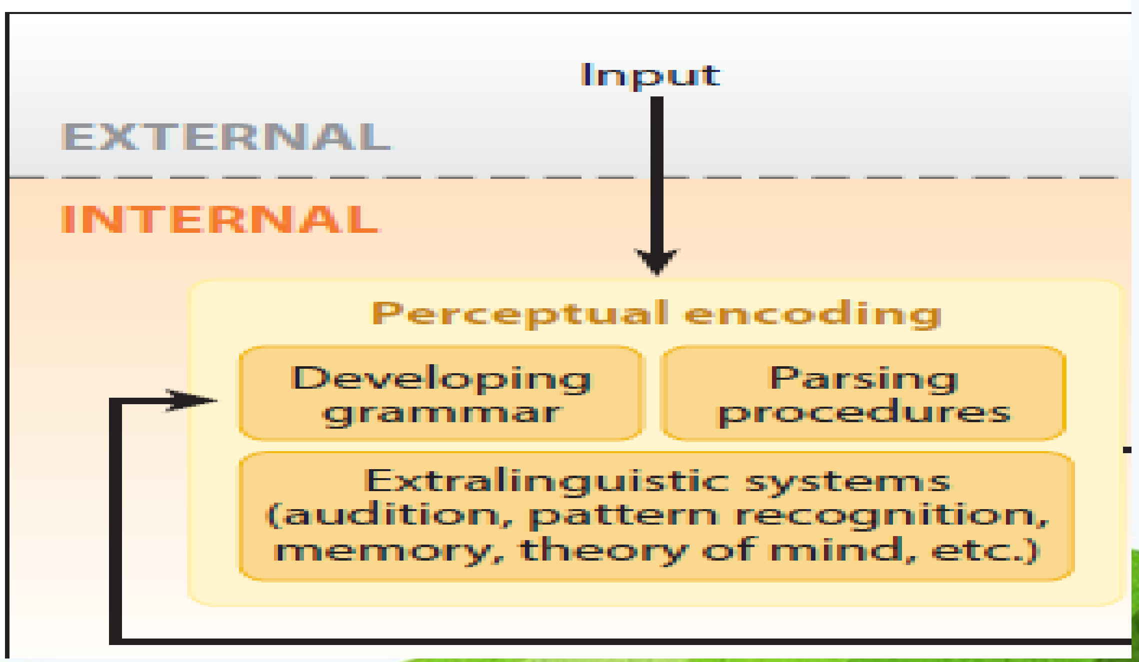


Lidz and Gagliardi (2014) model of Language Acquisition

- Highlights the link between INPUT and OUTCOMES while maintaining essential insights of the knowledge-driven perspective.
- Separates language acquisition into 3 components:
  - The intake component
  - The UG component
  - The inference component

Lidz and Gagliardi (2014) model of Language Acquisition

- Highlights the link between INPUT and OUTCOMES while maintaining essential insights of the knowledge-driven perspective.
- Separates language acquisition into 3 components:
  - The intake component
    - Identifies features of the information to which learners are sensitive



Lidz and Gagliardi (2014) model of Language Acquisition

- Intake component
  - Input feeds into the Perceptual encoding mechanism which is comprised of the Developing Grammar, Extra-linguistic systems that are constrained by audition, pattern recognition, memory, theory of mind, and others, and Linguistic systems such as Parsing Procedures
  - These components of the Perceptual encoding mechanism output the Perceptual intake
  - Which then feeds into the Inference engine and Production systems that support comprehension and behavioral responses

Lidz and Gagliardi (2014) model of Language Acquisition

- *Highlights the link between INPUT and OUTCOMES while maintaining essential insights of the knowledge-driven perspective.*
- *Separates language acquisition into 3 components:*
  - The UG component
    - Identifies the class of representations that shape the nature of human grammatical systems

Lidz and Gagliardi (2014) model of Language Acquisition

- UG component
  - Determines the hypothesis spaces (possible grammars) the learner entertains and the constraints on the interpretation of the abstract representations that make up the grammars
  - Generates the predictions about what the learner should expect to find in the environment from every possible grammar defined by UG
    - This is what is fed into the Acquisitional intake

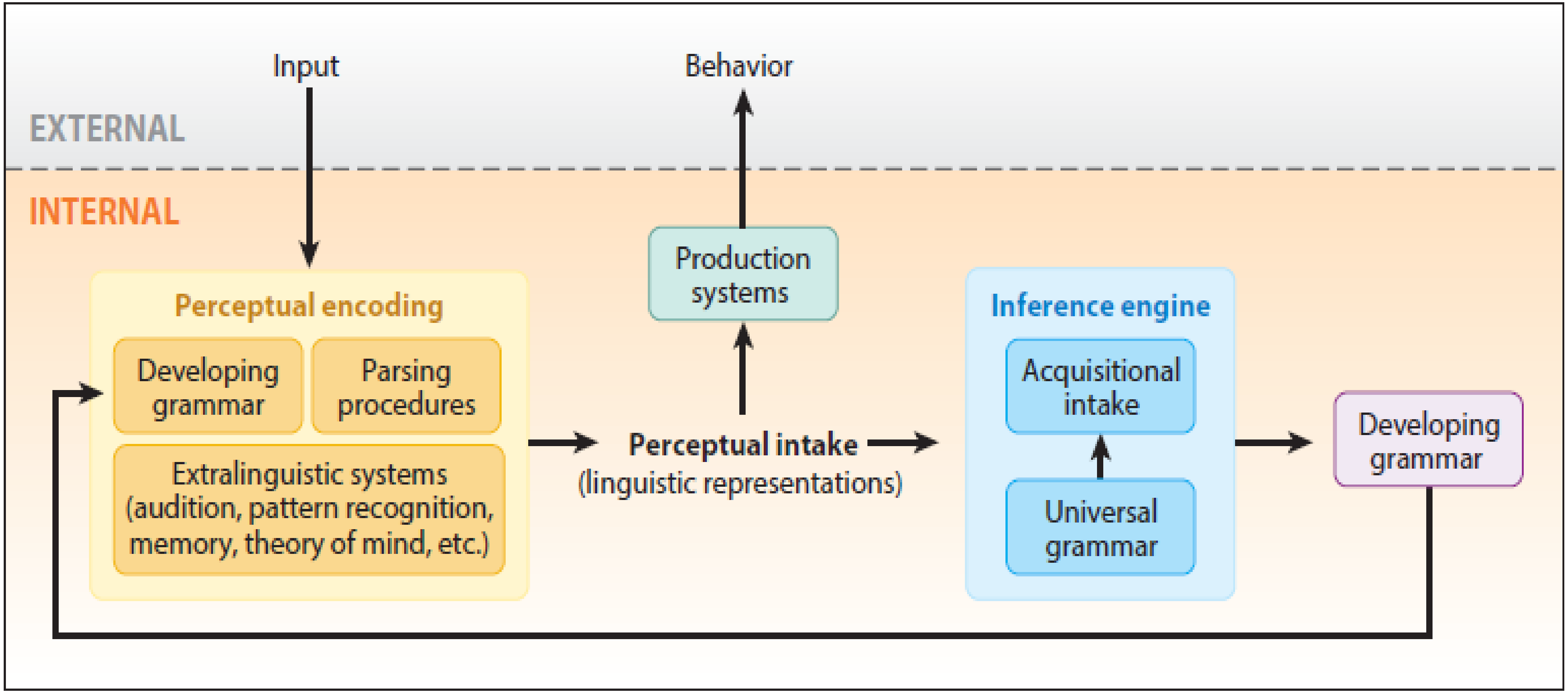
Lidz and Gagliardi (2014) model of Language Acquisition

- Highlights the link between INPUT and OUTCOMES while maintaining essential insights of the knowledge-driven perspective.
- Separates language acquisition into 3 components:
  - The inference component
    - Identifies how learners combine the intake with UG in selecting a particular grammar

Lidz and Gagliardi (2014) model of Language Acquisition

- Inference component
  - Although UG is a part of the Inference engine it only supplies a set of possible representations and constraints on how grammars can be constructed. It doesn't make the inference.
  - Inference takes places in the Acquisitional intake that compares the perceptual intake to the predictions generated by all of the possible grammars defined by UG
  - The selected representation is inferred from the combination of these two mechanisms and is then added into the Developing grammar
  - Once this has been updated the new Developing grammar is used to encode subsequent input that makes up further perceptual intake

Lidz and Gagliardi (2014) model of Language Acquisition



Evidence for all 3 components

1. Although learners are sensitive to statistical features of their language environment they don't always match objective measures of informativity (Perceptual intake mechanism)
2. UG alone doesn't solve the language acquisition problem, but is just one very important component that must be matched with an inference mechanism (Inference component)
3. Some of the statistical sensitivities of the learner are defined by UG which allows the learner to have knowledge of information outside of what is available in the input/intake (UG)
    1. Shows that UG adds the space of possible grammars to be considered and defines the data that enables the inference engine to link perceptual intake to the target grammar under construction

Statistical Sensitivities

- Hart & Risley (1995): High SES 3 year olds had an average vocabulary of 1000 words whereas Low SES children of the same age had an average vocabulary of 500 words. That although SES predicted both the parental input and child vocabulary growth, measures of input were stronger predictors of child outcome
- Hoff (2003): Found diversity of word types and MLU of the parents accounted for the variability in the rate of vocabulary growth in children
- Rowe (2002): Found it's not just quantity of speech that aids vocabulary growth but that input has different effects on language development as a function of what the child already knows
  - 2year olds: quantity of speech
  - 3 year olds: diversity of vocabulary used by parents
  - 4 year olds: complexity of narratives and decontextualized speech

Statistical Sensitivities

- These same findings are evident in syntax as well
- Huttenlocher et al. (2002): Found that variation in caregiver syntactic complexity explains the complexity of 5 year olds' speech
- Huttenlocher et al. (2010): found that higher diversity of caregiver speech predicted several properties of their child's speech: lexical diversity, diversity within constituents, and diversity in clause type
- Although this establishes the relationship between the data and the speech produced, effects of input seem to express themselves in the frequency with which children use complex structures and not the initial acquisition of those structures
  - Suggesting that the effect of input has more to do with the mechanisms by which knowledge of grammar are deployed and not the knowledge itself.

Statistical Sensitivities

- Legate and Yang (2007): link the variability of the use of root infinitives to the evidence available about the verbal morphology of the language being acquired. They propose UG with the intake predict the acquisition of tense
  - Spanish: 60% support grammar with tense; 2 years
  - French: 39% support grammar with tense; 2 years 8 months
  - English: 6% support grammar with tense; 3 years 5 months

Statistical Sensitivities

- These data are compatible with both perspectives
- Where they differ is in how the learner projects beyond their experience to structures not experienced yet

Statistical Sensitivities: Artificial Language

- Saffran et al. (1996): 8 month olds can segment words according to TPs in a toy language of three syllable words
- Johnson & Tyler (2010): similar to Saffran but with words of varied syllable number; infants unable to perform segmentation in this more realistic setup
- Gomez & Gerken (1999): infants sensitive to TPs between words at the sentence level
- Gomez (2002): infants can learn dependencies of nonadjacent words if there aren't any adjacent dependencies to track
- Gomez & Maye (2005): variability of space between elements encourages better learning
- Lany & Gomez (2008): nonadjacent dependencies can be learned but only once adjacent dependencies have already been acquired

Statistical Sensitivities: Artificial Language

- Takahashi (2009): created a corpus of utterances that had a statistical signature of constituent structure but that did not express possibility of movement rules
- Both 18 month olds and adults could distinguish between sentences with moved constituents and moved non-constituents
  - Without clues about movement in the data how were they able to recognize that movement was possible
  - It is important to distinguish between statistical sensitivity and the consequences of said sensitivity
    - It is only through the consequences of sentences outside of the data that we can draw conclusions about the role that input plays in acquisition
    - When learners generalize to structures that are absent or dissimilar from the input then the role of UG and its contributions can be proposed while still acknowledging the role of statistical sensitivity

Distinguishing INPUT from INTAKE

- Gagliardi & Lidz (2014): found children acquiring Tsez are sensitive to certain noun-internal features out of proportion to their predictive reliability in the input unlike the Bayesian model that relied more heavily on the more predictive features of the semantic cues
  - The preference for using phonological information over semantic information likely reflects perceptual intake in the initial stages of noun class learning
  - The Acquisitional intake may place a greater weight on phonological features for informing and forming noun classes
    - Perhaps an innate bias provided by UG to treat formal categories as formally and not semantically conditioned

Universal Grammar not sufficient to drive learning

- Understood since Chomsky (1965) that UG isn't the only and sufficient component for the acquisition of language since UG is abstract and needs a mapping between these abstractions and the concrete nature of the grammar being acquired
- Han et al. (2007): Although UG provides the constraints on naturally acquirable language statistical information is needed in order to identify the expression of these in the surface form of the language
  - Quantificational sentences in Korean

Universal Grammar not sufficient to drive learning

- Han et al. (2007): Interpretation of these sentences can go 2 ways
- (4a) Khwuki monste-ka motun khwuki-lul mek-ci ani ha-yess-ta
  - cookie monster-NOM every cookie-ACC eat-CI NEG do PST-DECL
  - ‘Cookie Monster didn’t eat every cookie.’ (long negation)
- (4b) Khwuki monste-ka motun khwuki-lul an mek-ess-ta
  - cookie monster-NOM every cookie-ACC NEG eat-PST-DECL
  - ‘Cookie Monster didn’t eat every cookie.’ (short negation)

Universal Grammar not sufficient to drive learning

- Han et al. (2007): Interpretation of these sentences can go 2 ways
- If the universal takes scope over negation this would mean that Cookie Monster ate none of the cookies
- If negation takes scope over the universal this would mean that Cookie Monster ate some but fewer than all the cookies
  - Tested 4 year olds and adults and found that within each group interpretation was split between both possible interpretations suggesting that each speaker controls only one grammar with the variability following from which grammar they were implementing

Universal Grammar not sufficient to drive learning

- Han et al. (2007): Interpretation of these sentences can go 2 ways
- Proposed that this is due to the sparseness of the data available to accurately select the correct grammar
- In the absence of this evidence learners picked one grammar at random from the set of structures licensed by UG
- This highlights the need for a separate mechanism aside from UG that allows this to occur
  - UG specifies the options but does not determine the grammar

The dissimilarity of Intake stats and acquired knowledge

- Poverty of the stimulus problem (Induction problem) in Kannada
  - Ditransitives in this language have flexible word order and the optionality of a benefactive affix on the verb. Theses 2 features together allow for 4 possible ditransitive combinations
  - (5a) hari rashmi-ge pustaka-vannu kalis-id-a
    - hari rashmi-DAT book-ACC send-PST-3SM
    - 'Hari sent the book to Rashmi.'
  - (5b) hari rashmi-ge pustaka-vannu kalis-i-koTT-a
    - hari rashmi-DAT book-ACC send-PP-BEN.PST-3SM
    - 'Hari sent the book to Rashmi.'
  - (5c) hari pustaka-vannu rashmi-ge kalis-id-a
    - hari book-ACC rashmi-DAT send-PST-3SM
    - 'Hari sent the book to Rashmi.'
  - (5d) hari pustaka-vannu rashmi-ge kalis-i-koTT-a
    - hari book-ACC rashmi-DAT send-PP-BEN.PST-3SM
    - 'Hari sent the book to Rashmi.'

The dissimilarity of Intake stats and acquired knowledge

- Poverty of the stimulus problem (Induction problem) in Kannada
  - Ditransitives in this language have flexible word order and the optionality of a benefactive affix on the verb. Theses 2 features together allow for 4 possible ditransitive combinations
- With asymmetries in binding across constructions
  - Benefactive construction the dative can bind into the accusative independent of word order but in nonbenefactive constructions binding can only occur when the dative appears first

The dissimilarity of Intake stats and acquired knowledge

- (6a) Q-DATx ACCx V+BEN
  - Rashmi pratiyobba hudugan-ige avan-a kudure-yannu tan-du-koTT-aLu
    - Rashmi every boy-DAT 3SM-GEN horse-ACC return-PPL-BEN.PST-3SF
    - 'Rashmi returned every boy his horse.'
- (6b) ACCx Q-DATx V+BEN
  - Rashmi avan-a kudure-yannu pratiyobba hudugan-ige tan-du-koTT-aLu
    - Rashmi 3SM-GEN horse-ACC every boy-DAT return-PPL-BEN.PST-3SF
    - 'Rashmi returned his horse to every boy.'
- (6c) Q-DATx ACCx V
  - Rashmi pratiyobba hudugan-ige avan-a kudure-yannu tan-d-aLu
    - Rashmi every boy-DAT 3SM-GEN horse-ACC return-PST-3SF
    - 'Rashmi returned every boy his horse.'
- (6d) *ACCx Q-DATx V
  - *Rashmi avan-a kudure-yannu pratiyobba hudugan-ige tan-d-aLu
    - Rashmi 3sm-GEN horse-ACC every boy-DAT return-PST-3SF
    - 'Rashmi returned his horse to every boy.'

The dissimilarity of Intake stats and acquired knowledge

- Poverty of the stimulus problem (Induction problem) in Kannada
  - Ditransitives in this language have flexible word order and the optionality of a benefactive affix on the verb. Theses 2 features together allow for 4 possible ditransitive combinations
- With asymmetries in binding across constructions
  - Benefactive construction the accusative can bind into the dative only when the accusative appears first but in nonbenefactive constructions binding can occur independent of word order

The dissimilarity of Intake stats and acquired knowledge

- (7a) *DATx Q-ACCx BEN
  - *sampaadaka adar-a lekhan-ige pratiyondu lekhana-vannu kaLis-i-koTT-a
    - editor it-GEN author-DAT every article-ACC send-PP-BEN.PST-3SM
    - 'The editor sent its author every article.'
- (7b) Q-ACCx DATx BEN
  - sampaadaka pratiyondu lekhana-vannu adar-a lekhan-ige kaLis-i-koTT-a
    - editor every article-ACC it-GEN author-DAT send-PP-BEN.PST-3SM
    - 'The editor sent every article to its author.'
- (7c) DATx Q-ACCx unaffixed
  - sampaadaka adar-a lekhan-ige pratiyondu lekhana-vannu kaLis-id-a
    - editor it-GEN author-DAT every article-ACC send-PST-3SM
    - 'The editor sent its author every article.'
- (7d) Q-ACCx DATx unaffixed
  - sampaadaka pratiyondu lekhana-vannu adar-a lekhan-ige kaLis-id-a
    - editor every article-ACC it-GEN author-DAT send-PST-3SM
    - 'The editor sent every article to its author.'

The dissimilarity of Intake stats and acquired knowledge

- There is an interaction between morphology, word order, and grammatical function of the quantifier
  - When Benefactive present dative argument is syntactically prominent for binding
  - When Benefactive is absent accusative is syntactically prominent for binding
  - These patterns are not reflected in the input as per Viau & Lidz (2011). Few examples that do occur are not sufficient for drawing conclusions about the binding properties of this language
    - Must involve projections from other facts that are more accessible

2 Contributions of UG to learning

- Viau & Lidz (2011): Identified 2 contributions that UG affords acquisition of binding facts
  - First contributed by Harley (2002) is that it defines the space of possible languages that links the syntax of possession with the syntax of ditransitives
  - Second is UG allows the syntax to define the Acquisitional intake from which the statistical inference can proceed

2 Contributions of UG to learning

- Viau & Lidz (2011): Identified 2 contributions that UG affords to the acquisition of binding facts
  - First contributed by Harley (2002) is that it defines the space of possible languages that links the syntax of possession with the syntax of ditransitives
    - UG enables a complex set of facts to follow from a single representational parameter concerning the syntax of possession relations
      - If the language exhibits possession structure where the possessed his higher than the possessor that structure can be recruited in ditransitives
        - Treating the goal argument as a possessor thereby making it syntactically prominent

2 Contributions of UG to learning

- Viau & Lidz (2011): Identified 2 contributions that UG affords to the acquisition of binding facts
  - Second is UG allows the syntax to define the Acquisitional intake from which the statistical inference can proceed
  - UG defines the information that children use to determine whether a given ditransitive uses a goal-prominent or theme-prominent syntax
    - Surface form varies cross-linguistically for ditransitives
      - English: 2 forms distinguishable by word order
      - Kannada: 2 forms distinguishable by an affix on the verb and not word order
      - Spanish: 2 forms distinguishable by clitic doubling of the dative argument and not word order

2 Contributions of UG to learning

- Viau & Lidz (2011): Identified 2 contributions that UG affords to the acquisition of binding facts
  - The matching of strings to the underlying structure can be accomplished by tracking NPs in the surface form that occur as the dative argument
  - The dative in a goal-prominent argument is restricted to being a possessor of the theme so the NPs that appear in this surface form are themselves restricted
  - Tend to be animate
  - Perceptual intake: needs representations of morphological and word order variability as well as grammatical functions of each argument
  - Acquisitional intake: a representation of the proportion of animate to inanimate datives for each of the morphological and word order variants
    - Learner encounters a construction statistically biased to animate goals this skews the distribution to support an inference that it involves goal-prominent syntax

Summarizing UG contributions

- 2 Contributions of UG to language acquisition
  - Explains specific ways children project beyond their experience (input encountered)
  - Defines the Acquisitional intake which enables statistical-distributional evidence informative to acquiring the correct grammar
    - Kannada ditransitives :
      - The first contribution explains why children have knowledge about binding patterns across novel sentence types
      - The second contribution explains how the observations can be linked to those binding patterns

Conclusions

- In some cases it is the learner's expectations about how languages are structured that define the information in the environment that drives learning
- Theory of UG not equivalent to a theory of language acquisition
- The model integrates the statistical sensitivities of the learner in the formation of the perceptual and acquisitional intake with UG and shows the differences between input and intake
- Aims to unify the input-driven and knowledge-driven perspectives into a single workable framework

Summary Points provided by Lidz and Gagliardi (2014)

- 1. Effects on language acquisition of statistical–distributional patterns in the input are expected under any theory of language acquisition, even those with a large contribution of innate knowledge.
- 2. The hypothesized existence of UG does not remove the need for a theory of learning that explains how experience contributes to language acquisition.
- 3. We deconstruct the language acquisition device into three parts: intake mechanisms, UG, and inference mechanisms.
- 4. UG explains the dissimilarity between experience and acquired knowledge. It also allows learners to draw inferences from statistical–distributional evidence to the grammatical representations responsible for producing that evidence.
- 5. Intake mechanisms explain how learners filter their input to identify critical information to learn from. Input may be filtered by information processing mechanisms, prior knowledge, or expectations associated with particular grammatical hypotheses.
- 6. Inference mechanisms connect UG with the intake to determine the appropriate mapping from abstract representations to surface form.