Predicting a Child’s Trajectory of Lexical Acquisition

Nicole Beckage, Michael Mozer, Eliana Colunga
University of Colorado Boulder
Lexical Acquisition Trajectories

• Can we predict the words a child will learn next based on information about the child?

• Can we predict what words a child will learn next given the words a child already knows?

• Track the learning of a specific set of lexical items for individual children.
  • What are the general trends of word learning in individual children?
  • How does the current lexical items inform future lexical acquisition?
MacArthur-Bates CDI

- **Parent report** the individual words that their child can **produce**.
- **Fixed set** of nearly 700 well studied words
- **Normative** age of acquisition norms.
Normative CDI data

<table>
<thead>
<tr>
<th></th>
<th>month 16</th>
<th>month 17</th>
<th>…</th>
<th>month 29</th>
</tr>
</thead>
<tbody>
<tr>
<td>airplane</td>
<td>38.5</td>
<td>39.4</td>
<td>…</td>
<td>95.0</td>
</tr>
<tr>
<td>light</td>
<td>35.9</td>
<td>30.3</td>
<td>…</td>
<td>90.0</td>
</tr>
<tr>
<td>zoo</td>
<td>9.0</td>
<td>9.1</td>
<td>…</td>
<td>66.7</td>
</tr>
</tbody>
</table>

- For each words, what **percentage of children at a give age** are reported to produce it.

- **Cross sectional data.** Over 1000 CDIs, with at least 70 children of each age included in norms.

- Further separated to allow for **female** and **male** specific age of acquisition norms.
In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

**Abstract**

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.

**Introduction**

In this paper, we explore whether there are regularities in the growth of a child’s vocabulary that allow the trajectory of an individual’s learning to be predicted. How does a child’s current vocabulary inform and relate to their vocabulary in the future? We know that deficits in a child’s early lexicon is a predictor of future language skills. One source of information that can be used to model vocabulary acquisition is the Communicative Development Inventory (CDI).

Despite their shortcomings, the CDI norms may be useful for characterizing an individual child’s lexical growth. In this work, we develop models that attempt to predict when a specific word will be learned by a particular child. The models are based on two qualitatively different sources of information: a representation describing the child (age, sex, and quantifiers of vocabulary size) and a representation describing the specific words a child is ready and able to learn, early learning, and this understanding should inform cognitive theories of development. On a practical level, these models may support the development of interventions to boost language acquisition.

**Keywords:** Language acquisition; word learning; lexical acquisition.
Predictive modeling

• Using *regularized logistic regression*, can we predict the words a child is likely to learn next?
• Regression model allows us to investigate what *types of information* are relevant for predicting future lexical growth.

• Train **individual models** for each word (e.g. 680 models).
• **Cross-validation** to determine regularization parameter and model coefficients.
Predictive modeling

• Can we predict the individual lexical items that a child is likely to learn?

• Three types of information:
  • normative age of acquisition trends for specific words
  • information specific to the child
  • lexical items known by the child

<table>
<thead>
<tr>
<th></th>
<th>month 16</th>
<th>month 17</th>
<th>...</th>
<th>month 29</th>
</tr>
</thead>
<tbody>
<tr>
<td>airplane</td>
<td>38.5</td>
<td>39.4</td>
<td>...</td>
<td>95.0</td>
</tr>
<tr>
<td>light</td>
<td>35.9</td>
<td>30.3</td>
<td>...</td>
<td>90.0</td>
</tr>
<tr>
<td>zoo</td>
<td>9.0</td>
<td>9.1</td>
<td>...</td>
<td>66.7</td>
</tr>
</tbody>
</table>

Probability of learning word \( i \) is the (smoothed) rate of change based on the norms.
Predictive modeling

• Can we predict the individual lexical items that a child is likely to learn?

• Three types of information:
  • normative age of acquisition
  • information specific to the child
  • lexical items known by the child

<table>
<thead>
<tr>
<th></th>
<th>age</th>
<th>sex</th>
<th>...</th>
<th>voc. sz</th>
<th>dog</th>
<th>house</th>
<th>...</th>
<th>zoo</th>
</tr>
</thead>
<tbody>
<tr>
<td>kid A</td>
<td>16.2</td>
<td>F</td>
<td>...</td>
<td>32</td>
<td>0</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>17.1</td>
<td>F</td>
<td>...</td>
<td>49</td>
<td>1</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>18.9</td>
<td>F</td>
<td>...</td>
<td>132</td>
<td>1</td>
<td>0</td>
<td>...</td>
<td>1</td>
</tr>
<tr>
<td>kid B</td>
<td>19.3</td>
<td>M</td>
<td>...</td>
<td>257</td>
<td>1</td>
<td>0</td>
<td>...</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>20.5</td>
<td>M</td>
<td>...</td>
<td>345</td>
<td>1</td>
<td>1</td>
<td>...</td>
<td>0</td>
</tr>
</tbody>
</table>

Probability of learning word $i$ is a function of child or lexical features
Normative model

• If all children learn somewhat the same then the normative model should do really well at predicting individual word learning.

• Two types of normative models.
  • One based on male and female children together (benefit of larger data set).
  • One based on the sex of the individual child.
Child and word feature models

- Individual learner information is useful in capturing language acquisition.
  - The sex of the child as well as the CDI percentile have been shown to relate to language learning ability.
  - Does that help with the predicting acquisition of individual words?
- Individual lexical items may tell us about future lexical learning.
  - The current vocabulary likely captures information about the language environment.
  - And interests of the specific child.
Results & discussion

<table>
<thead>
<tr>
<th></th>
<th>llk train</th>
<th>llk test</th>
<th>AUC</th>
<th>% best fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>norms</td>
<td>-123076</td>
<td>-30849</td>
<td>.588</td>
<td>1.39</td>
</tr>
<tr>
<td>norms(m/f)</td>
<td>-123813</td>
<td>-31135</td>
<td>.563</td>
<td>1.39</td>
</tr>
<tr>
<td>child</td>
<td>-84698</td>
<td>-22774</td>
<td>.812</td>
<td>67.02</td>
</tr>
<tr>
<td>word</td>
<td>-65703</td>
<td>-24059</td>
<td>.801</td>
<td>30.20</td>
</tr>
</tbody>
</table>

- Both child features and word features improve over the norms.
- The child feature model overwhelmingly and significantly outperforms all other models.
Improving the models

• If child features and lexical features improves over normative acquisition, does including both features further increase performance?
  • Only true if features contain different and independent types of information.

• It seems likely that the lexical items should be useful.
  • However, the representation is sparse—vector of 0s and 1s indicating if a word is produced by the child.
  • Does reducing the density of this representation increase performance?
Results for extended models

<table>
<thead>
<tr>
<th></th>
<th>total llk</th>
<th>AUC</th>
<th>pos. param</th>
<th># params</th>
</tr>
</thead>
<tbody>
<tr>
<td>child</td>
<td>-22774</td>
<td>.812</td>
<td>7</td>
<td>4.70</td>
</tr>
<tr>
<td>word</td>
<td>-24059</td>
<td>.801</td>
<td>650</td>
<td>31.20</td>
</tr>
<tr>
<td>redux-word</td>
<td>-22654</td>
<td>.810</td>
<td>19</td>
<td>8.39</td>
</tr>
<tr>
<td>redux-w+child</td>
<td>-22848</td>
<td>.803</td>
<td>25</td>
<td>9.72</td>
</tr>
<tr>
<td>ensemble</td>
<td><strong>-22263</strong></td>
<td><strong>.816</strong></td>
<td>26</td>
<td>13</td>
</tr>
</tbody>
</table>

- PCA reduction on **lexical** vector results in **increased** performance (comparable to child model).
- Including all features in one logistic regression **decreases** fit (generalization, sparsity issue).
- Averaging probability estimates from child model and word model separately, **significantly** increases fit.
Significance

• The way we represent **lexical knowledge** affects fit!

• Building a model based on **child** features alone and **word** features alone is better than a joint model.
  • The information from child and word features is partially **independent**.
  • **Both child** and **word** features **contribute** to predicting word learning.

• We can use these types of **model evaluation** metrics to predict future lexical acquisition.

• AND also as a method to **understand** what features are important/salient to young learners.
Feature analysis

- **Percentile** is the predictor that appears in the most (87%) models.
  - Other work shows that early talkers and late talkers learn differently, both in the rate of learning and the words they learn.

- **Vocabulary size** is also common (75%).

- Even though percentile is a function of vocabulary and age, the three often are in all three models (38%).

- Many word models have a small (<8) words that are really useful in prediction (that’s part of the reason the reduction helped). *Can we understand lexical connectivity by exploring this more?*
Future directions

• Extend lexical prediction to other models.
  • Neural network models.
  • Graph theoretic growth models.

• Capture the relationship between learning specific words.
  • Is the production of certain words predictive of learning other words? Can this help us understand relevant features for early lexical learning?

• What is the best representation of the vocabulary knowledge that accurately predicts learning of individual lexical items? For individual children?
Why we care...

• If we know what words a child is ready to learn we may be able to facilitate learning.

• If we understand the role of the current vocabulary on future vocabulary growth, we can hopefully catch and possibly correct individual lexical trajectories.

• We can use this type of modeling approach to capture lexical relations that can inform our understanding of the relevant features of word learning.
Thanks!

To Eliana Colunga and the DACS lab for the beautiful longitudinal data and feedback on the work.

Email me at nicole.beckage@colorado.edu with questions and comments!