New insights from daylong audio transcripts of children's language environments

Jessica L. Montag (jmontag@illinois.edu)
Department of Psychology, 603 E Daniel Street
Champaign, IL 61820

Abstract

Recent technological advances and research trends have enabled the collection and analysis of multi-hour or daylong recordings of children's auditory environment. While this technology has allowed researchers to sample language experience from multiple contexts across the day, challenges remain with respect to how these audio recordings can or should be coded and analyzed. Daylong audio samples have the potential to transform our understanding of the language input that children encounter, but new analysis techniques may be necessary to take advantage of these new opportunities. The present work explores the linguistic content of the transcripts of three daylong recordings with the goal of understanding the content of these recordings in order to develop new ways to analyze and gain insight from these recordings.

Keywords: language development, daylong audio, corpus

Introduction

There is a great deal of evidence suggesting that aspects of the language environment, broadly defined, contribute to language outcomes. Many features of the language environment have been shown to predict language outcomes, including the amount of language (Weisleder & Fernald, 2013), variability or diversity of that language, (Huttenlocher et al., 2010) and multiple social factors (Hirsh-Pasek et al., 2015). These findings have generated a great deal of interest in investigations of naturalistic language environments, with the goal of understanding which aspects of the language environment are associated with which language outcomes. This knowledge may help us understand both healthy development and remediation for children whose language skills lags behind their peers.

Investigations that aim to link language environments with language outcomes are not trivial endeavors, and researchers encounter both methodological and theoretical obstacles. Methodological challenges include the practicalities of collecting samples, especially representative samples, of children's language environments, and then subsequently transcribing, coding or otherwise transforming the audio data into a form that is useful for research.

Other challenges are theoretical. In the field of language development, there is some consensus with respect to how we measure language outcomes. Vocabulary inventories and behavioral measures such as looking time or eye gaze are often used to assess linguistic knowledge. While these measures are imperfect, there is nonetheless some agreement, as evidenced by the large number of studies that use these methodologies, that they indicate something meaningful about children's knowledge. However, there is substantially *less* agreement with respect to what the appropriate measures of the environment might be that predict these language outcomes. There is uncertainty regarding which *constructs* serve as theoretically relevant predictors of language outcomes, and how these constructs should be *operationalized* (Montag, Jones & Smith, 2018). For example, linguistic diversity is associated with positive language outcomes (Huttenlocher et al, 2010; Rowe, 2012), but it can be measured in many ways: number of unique words in a transcript, number of unique words relative to a measure of total transcript size, number of unique words in each unit of some amount of time, and the temporal spacing of words over time and context. Understanding how to operationalize a measure such as linguistic diversity is not obvious. Different methods of operationalization may have different implications for the theoretical links between language input and outcomes, and the learning processes that underlie language development. Our theories shape the selection and operationalization of the constructs we use to predict language outcomes, and in turn the selection and operationalization of constructs affect our theories.

New advances in data collection and data analysis (Bergelson et al., 2019; Gilkerson & Richards, 2008; Le Franc et al., 2018), which in the near future could include machine learning techniques to aid more rapid speech-to-text transcription, may solve some aspects of the methodological challenges associated with collecting and transcribing or coding natural language environments. However, without appropriate theoretically motivated plans for data analysis, we risk having the data but lacking the inferential methods for using that data to answer our research questions. A central proposition in the present work is that new methods that allow us to collect and analyze longer recordings do not merely provide *larger* datasets, but *different* datasets and we need both analytic tools and well-defined theory to make sense of this new data.

The present analyses explore the content of three fully transcribed daylong audio recordings. The goal of this work is two-fold. The first goal is to better understand: *what kinds of data are in daylong audio transcripts?* To this end, analyses look at the number of words and unique words, and how they are distributed over time in the home language environment. The second goal is to uncover ways to *operationalize variables of interest* that emerge from the exploration of the content of the daylong recordings.

The goal of this work is not to draw conclusions regarding the structure of early language input on the basis of three transcripts, but rather explore hypothetical dimensions that might vary as a means of starting to understand how to analyze daylong recordings in a sensible, theoretically motivated way. To date, no analysis has explored the content of three fully transcribed days of child-available speech so the findings and lessons learned here may provide insights into how larger datasets may be best analyzed.

Methods

The three daylong audio recordings included in the present analyses were collected using the LENA system (Language Environment Analysis; LENAFoundation). LENA devices are small audio recorders that children wear in the pocket of a custom piece of clothing that record long intervals of the auditory environment in a minimally intrusive manner.

Audio Recordings

One recording (Child 1) was of a 12-month-old child and was recorded and transcribed by VanDam (2018). The other two recordings (Child 2A and 2B) are of the same child approximately one month apart at 10 and 11 months and were recorded by Fausey and Mendoza (2018) and transcribed by the author's lab. The VanDam transcript and the Fausey audio recordings were all retrieved from the HomeBank online repository (VanDam et al., 2016). Both target children were girls and lived in English-speaking homes. All recordings took place in the child's home.

Child 1's recording includes a full day of a child's auditory environment. Child 2A's recording includes nearly a full day of audio, but with a 4-hour interval where the child left the home, so audio was not recorded. Child 2B's recording only contains audio from the morning and evening, because the child attended daycare during the day.

Table 1: Recordings included in present analyses

File	Age	Awake Hours Recorded*	Total Words
Child 1	1 year, 7 days	9.5	27,471
Child 2A	10 mo., 9 days	5.2	26,435
Child 2B	11 mo., 7 days	2.9	15,027

*Awake hours are approximate, given challenges with judging when children fell asleep during naps and bedtime.

Transcribing Procedure

For information about the transcription of Child 1's audio, see VanDam (2018). Child 2A and 2B were transcribed in CHAT format (MacWinney, 2000) using ELAN software (Lausberg & Sloetjes, 2009). When transcribing using ELAN, utterances are diarized (partitioned by speaker) enabling subsequent analyses by speaker. Utterances are also "segmented" meaning that the beginning and endings of utterances are identified with timestamps that can be used to analyze exactly when the utterance was produced. The audio was transcribed by one research assistant and then checked for accuracy by a different research assistant. Full transcripts will be made available at the HomeBank online repository upon publication of the final manuscript.

Analysis

All analysis code was written in Python. All analyzed speech is child-available speech (speech that was captured by the audio recorded regardless of whether it was address to the child or another individual) though subsequent analyses could analyze only child- or adult-directed portions of the transcript. Words were not lemmatized, and contractions and other word shortenings were left intact. For example, "don't" and "cuz" were *not* recoded as "do not" and "because." The present analyses aimed to modify the content of the speech as little as possible, though subsequent analyses may make different choices which would be equally valid given the goals of the research endeavor.

Results

The results presented here aim to describe and visualize the language that appears in children's language environments. First, analyses describe the presence of language throughout the day and the lexical diversity of that language. Then, analyses of individual words capture the exact words that appear in day-to-day speech, and the distribution of individual words over time.

Total Words over Time

Figure 1 illustrates the occurrence of words over the course of the day. Child 1 was recorded all day and the three notable gaps in figure correspond to naps. Child 2A was recorded nearly all day, with a gap in the morning corresponding to a nap and a larger gap in the afternoon when the child left the home. Child 2B was recorded in the morning before daycare and in the evening upon returning home.

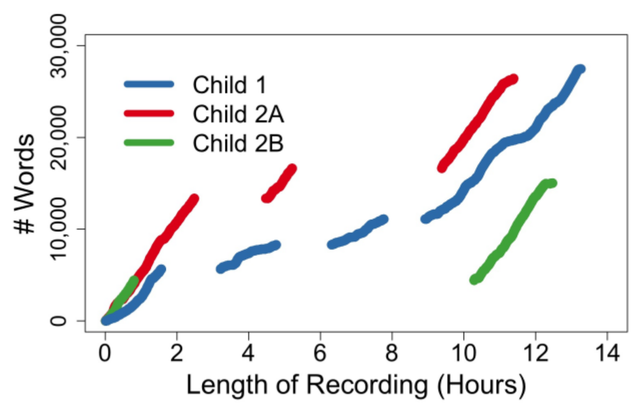

Figure 1: Total count of child-available words over time

It would be ideal to have full recordings for all children, but the reality is that researchers are typically prohibited from recording outside of the home given concerns about speakers who have not consented to being recorded. Many young children indeed spend a good deal of time outside of the home, either at daycare or other activities (outside play, running errands, visiting family and friends) so rather than exclude these families or these days from analyses, it is still interesting to understand what home language environments are like on days where children spend a substantial amount of time outside the home. Future research may provide insight into how home language environments may or not be different from daycare or other environments outside the home in which children experience language input.

The present work analyzes too small a sample to draw solid conclusions, but two additional features of the daylong input are worth noting. First, there is a remarkable amount of speech in these recordings, and few periods of silence when children are awake. Much of this may be overheard speech, and future work may reveal how typical or atypical this profile of language input might be. Second, Child 1's recording was more than three times the length of Child 2Bs, yet they heard less than twice as many words. It would be interesting if a substantial portion of a child's language input came from early morning and evening events, which would have implications for the temporal dynamics of language experience and how to model language environments from incomplete data. These are exactly the questions that daylong audio recording methodologies might answer.

Operationalization of Lexical Diversity

Another way to visualize the amount of language captured in the audio recordings is in Figure 2, in which the total number of words in the recording is plotted along the x-axis and the total number of unique words in plotted along the y-axis. The lexical diversity of the three samples is approximately equal (the three lines are mostly overlapping) though there may be more moment-to-moment variability in Child 1 as shown by the more variable slope. For example, the increase in slope at the end of the day corresponds to an extended book reading activity before bedtime where many words and many new words were uttered. Figure 2 also shows how unique word types accumulate as the total number of words increases. The three curves illustrate the classic function described by Heaps' (1978) and Herdan's (1960) laws; as a sample size increases, the rate of encountering new unique words decreases. These laws provide a challenge for operationalizing lexical diversity. Larger samples necessarily have less lexical diversity, so type-token ratios are unsuitable when there is variability in sample size (for additional discussion, see Montag et al., 2018; Richards, 1987).

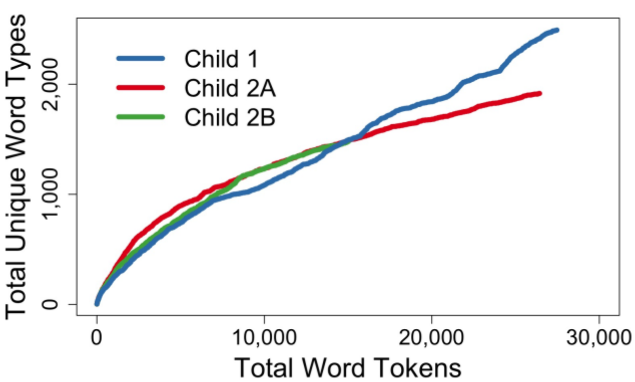

Figure 2: Accumulation of unique words given total words

Given that lexical diversity must be computed over a given sample size, and samples vary across families for a variety of reasons (true differences in amount to speech, differences in amount of speech that was captured by the audio recorder) type and token counts or ratios are insufficient for operationalizing lexical diversity in a way that allows samples to be compared to each other. Figure 3 offers an alternative way to operationalize lexical diversity. This figure shows the expected number of unique words in samples of different sizes, drawn from the full transcript.

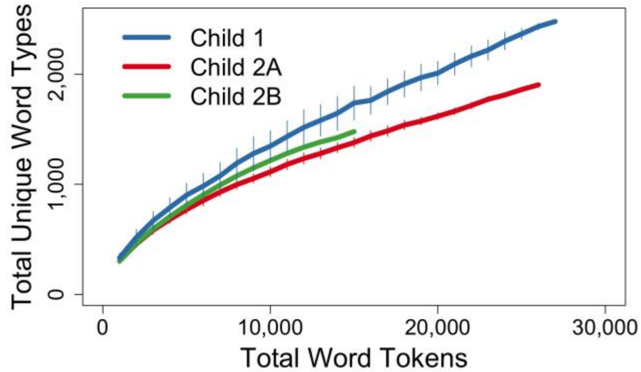

Figure 3: Simulated number of unique word types at different sample sizes

Samples of different sizes (in Figure 3, samples that increase in increments of 1,000 words) are sampled from the full transcript and the total unique words in each sample is calculated. Crucially, samples are selected contiguously rather than randomly because variability of context itself is associated with greater lexical diversity (Montag et al., 2018). For example, different words occur during bedtime, playtime, or mealtime. Contiguous samples reduce the contribution of diversity of contexts itself (more different conversational contexts sampled) as a contributor to lexical diversity in larger sample sizes. The sampling procedure is then repeated 100 times. Error bars represent standard deviations of unique word counts across those 100 samples. Lexical diversity should be calculated *relative* to some sample size, and this method allows for unique word counts to be estimated for a sample given properties of the whole transcript such that individual corpora or transcripts of different lengths can be compared to each other.

A goal of this work is to better understand how variables of interest, like lexical diversity, can be operationalized and used to predict various language outcome measures. To this end, analyses like those in Figure 3 yield a measure of lexical diversity that is as independent of sample size as may be possible—the total number of unique words in a sample of 1,000 or 5,000 or 10,000 words—which could be used as a predictor in models of word learning or language outcomes.

Operationalization of Words over Time

Daylong recordings highlight the fact that language is not distributed evenly over time. Some intervals have large amounts of speech and others contain less speech. Defining variability in language over time has both methodological and theoretical applications.

Methodologically, many investigations of child language environments record small time intervals, often 60-90 minutes though studies vary wildly in the number of time intervals selected over days, months, or years and generalize the findings to estimate a child's language experience (Hart & Risley, 1995; Huttenlocher et al., 2010; Rowe, 2012). There are certainly practical and sensible reasons to record smaller unit of time and extrapolate, but an important question is how representative that smaller sample might be of broader language experience. Children's language input may or may not be generally consistent over time, i.e., how much consistency is there in the amount of language encountered in different hours throughout the day. Further that consistency itself may vary across families such that for some children the amount of language they hear may be fairly consistent across the day such that different hours contain approximately equal amounts of speech, which for other children some hour might contain large amounts of speech while others contain very little.

Figure 4 provides one operationalization of the distribution of words over time. These histograms show the number of words in 1,000 randomly selected 1-hour intervals in the transcripts of Child 1 and Child 2A. Start times were randomly selected and the number of words contained in the subsequent hour of recording time was recorded. These intervals include naps (intervals where we know the child did not experience language) but do not include time when the child left the home and was not recorded (intervals where we know nothing about what language was present).

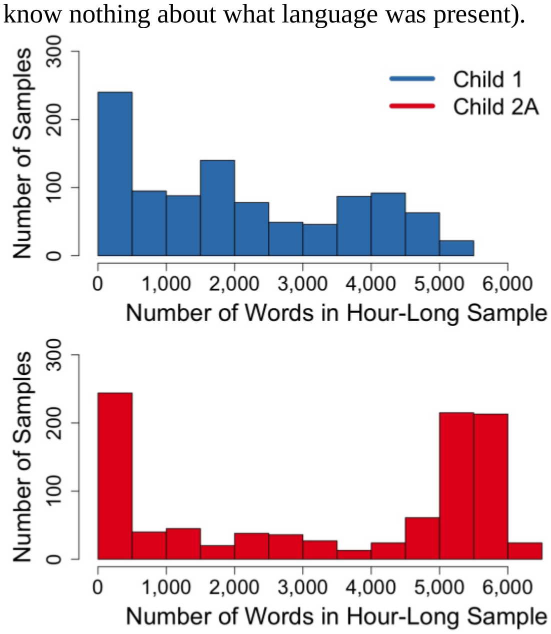

Figure 4: Number of total words in 1,000 randomly selected hours from Child 1 and Child 2A's transcript

Features of these histograms illustrate important details of the temporal dynamics of words in time, as well as consequences of the audio recording methodology. First, and most obviously, the distributions of the two histograms are different. The shape of the distributions captures the fact that Child 1's recording comprised a whole day without gaps (periods of silence correspond to naps) while Child 2A's recording contained a true gap in the middle where the child left the home, so a smaller proportion of Child 2A's transcript was true silence. The histograms reflect features of the hour-to-hour variability in each transcript during the day, including differences in speaking rates of caregivers, though also quirks of what events were and were not recorded in the two samples. Histograms such as these could be one way to operationalize how consistent speech might be throughout the day. For example, there is moderate variability in the number of words that Child 1 heard hour-to-hour. No hour was particularly verbose or sparse (ignoring naps). In Child 2A's environment, single hours either contained large amounts of speech or no speech, with few intermediate hours. Naps and outside excursions make these particular histograms somewhat challenging to interpret, but analyses like these could in principle be used to describe the hour-to-hour consistency of speech in a child's environment.

Second, with the exception of the high number of samples with fewer than 500 words of speech (these samples include naps) the distribution of Child 1 is mostly flat and Child 2A is highly skewed. There does not seem to be a "representative hour" as would be suggested by a normal distribution. The flat or skewed distributions of word counts in randomly selected hours suggests that recoding (or selectively transcribing from a longer recording) a single hour of a child's language input and extrapolating to a full day may present challenges because researchers may not know a-priori the shape distribution that the sample is drawn from, or how representative that hour might be. These findings are consistent with other work that questions the generalizability of a single hour of densely coded input (Bergelson et al., 2018; Mendoza & Fausey, 2019) or variability in language input across different contexts (Snoderstrom & Wittebolle, 2013). Given realities associated with the ability (or rather, lack of ability) to record full days of audio, or completely transcribe and code full days of audio, understanding the dynamics of words over time may be necessary for us to most effectively use imperfect data to uncover truths about language distributions.

Theoretically, understanding how language is distributed over time may have implications that derive from theories of learning. For example, the distribution of learning trials in time is known to affect learning outcomes (Carvalho & Goldstone, 2015; Estes, 1955). The distribution of words (or unique words, complex sentences, a child's own words, or any other relevant feature) in time may be an individual difference variable, in addition to linguistic quantity or diversity that may be of interest when linking language experience to language outcomes.

Item Analyses

In addition to the amount and distribution of overall language, daylong audio recordings also allow researchers to understand dynamics and variability of individual words.

Table 2 shows the 20 most frequent words in the three transcripts analyzed here, as well as all 120,000 words of child speech addressed to (or available to) children under the age of 24 months from the CHILDES corpus (MacWhinney, 2000). CHILDES data was accessed from childesdb (Sanchez et al., 2019) using the childesr package in R (Braginsky, Sanchez & Yurovsky, 2019).

Table 2: List of the 20 most frequent words in all three transcripts and about 120,000 words of child-directed speech from the CHILDES corpus

	Child 1		Child 2A		Child 2B		CHILDES	
1	you	1104	you	1339	you	745	you	6667
2	i	766	i	714	i	452	the	2843
3	the	741	the	650	the	413	that	2313
4	and	497	to	574	it	300	it	2248
5	to	471	a	457	to	276	a	2220
6	a	467	your	427	your	264	oh	1764
7	it	438	it	407	a	239	i	1711
8	that	394	that	330	that	232	ah	1629
9	your	346	yeah	310	is	209	to	1386
10	go	293	and	304	yeah	208	and	1356
11	no	248	okay	286	and	164	what	1295
12	in	231	oh	284	oh	160	yeah	1264
13	is	230	what	275	on	155	can	1230
14	oh	221	do	274	in	147	see	1207
15	have	221	we	268	okay	138	is	1201
16	ok	212	is	262	what	134	do	1175
17	here	208	on	259	no	129	here	1166
18	we	204	can	252	we	127	look	1153
19	she	203	want	221	can	127	on	1094
20	do	201	her	221	are	125	go	1063

There are clear similarities and differences across the corpora. For example, "I" is the second more frequent word in all three daylong transcripts, but it is the seventh most frequent word in CHILDES. The word "I" is often the most frequent word in naturalistic adult spoken corpora, so the higher frequency if "I" may reflect the higher incidence of adult-to-adult overheard speech the daylong transcripts, or some other asymmetry in the contexts sampled across the corpora. Likewise, the verbs "see" and "look" are relatively more frequent in CHILDES, which again could indicate a greater proportion of goal-directed caregiver-child interactions in CHILDES relative to the daylong recordings. Alternately, variability in individual word occurrences across corpora may be purely coincidental and not worth further inquiry; we do not yet know.

Differences between the daylong transcripts and CHILDES may suggest that contexts, activities, or families sampled in each corpus may systematically vary such that overall word frequency estimates are meaningfully different. In a trivial sense, this is surely true. For example, "bath" and "diaper" occur less frequently in CHILDES than would be expected given their frequencies in each of the three daylong transcripts. One possible explanation is that bathing or toilet contexts may be under-represented in CHILDES given that smaller units of time (like play-time or snack-time) were often sampled rather than full days. There may be other relevant differences in word choice that arise as a consequence of how audio was sampled. Future analyses may begin to provide answers. The lexical inventories of the daylong transcripts versus CHILDES further highlight the claim that daylong recordings provide not only more but different information about language environments, in this case sampling from contexts that otherwise may be absent from other assessments of language environments.

Beyond the frequencies with which different words appear, words also vary in how they are distributed over time. Some words may be approximately evenly distributed across a day while others may be "bursty" such that the word appears many times in a single context, and rarely or never in other contexts. In addition to the distribution of talk throughout the day, distributions of individual words, and the words with which they tend to co-occur may have implications for language learning (e.g., Elman, 1990; Landauer & Dumais, 1997). The distribution of words over time—whether they are consistently distributed or bursty, may be an important item-level difference to consider when predicting word learning, specifically why children normatively learn certain words before others. Word burstiness may be an important factor to consider, given the fact that (as previously discussed) the learning literature suggests that how examples are distributed through time have implications for learning (Carvalho & Goldstone, 2015; Estes, 1955).

There is considerable variability in how consistently words are distributed over time. Figure 5 shows the appearance of seven different words in the daylong transcripts of two children. Each vertical line corresponds to a one-minute period in which one of the seven target words appeared.

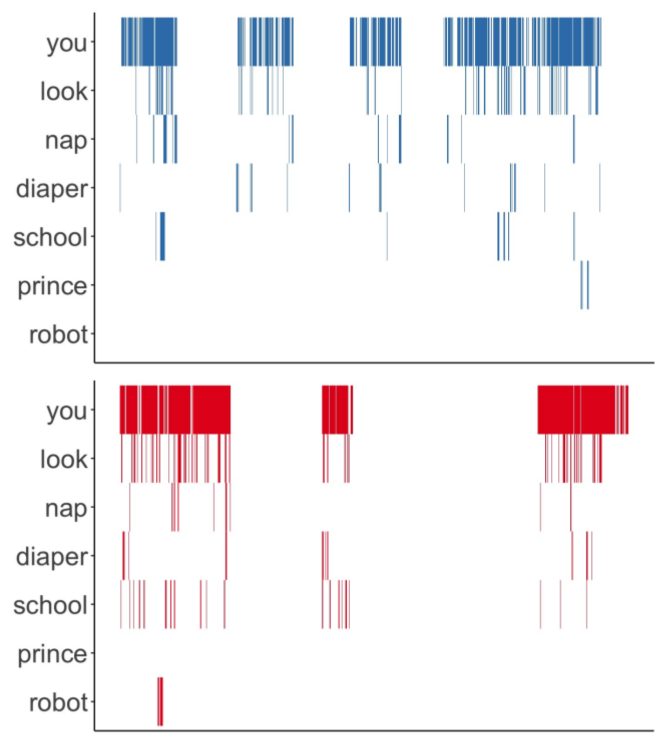

Figure 5: Occurrences of single words during each one-minute interval in each daylong transcript (top = Child 1, bottom = Child 2A)

The word “you” is the most frequent word in both transcripts. The gaps where ‘you’ was not uttered correspond to naps and Child 2A’s excursion outside the home. Other high-frequency words like “look” occur throughout the day. “Nap” and “diaper” occur less frequently, and in bursts rather than distributed equally in time. It is possible that these bursts correspond to times of day in which the word is relevant to the activities taking place. For example, is it clear for Child 1 (who took three naps) that the word “nap” often precedes or follows real naps. Likewise, the word “diaper” may coincide with diaper changes—note that both families utter the word “diaper” immediately after waking in the morning. Next, “school” may be an example of a word that has different temporal distributions in the two families. Though “school” is overall more frequent in Child 1’s transcript, its mentions appear many times in a single context and never or rarely in other contexts. In Child 2A’s transcript, school appears more consistently throughout the day. Finally, some very contextually bound (bursty) words appeared during book reading. Child 1 happened to hear a book about a prince and Child 2A happened to hear a book about a robot. Instances of these words are temporally restricted to the book reading event and appear nowhere else. If differences in the temporal distribution of words were a consistent difference in the children’s experience, it would be interesting if either profile, more or less bursty, were associated with better learning.

The goal of the present work is not to describe the temporal distributions of certain words, but to explore the dimensions along with distributions may vary. In addition to overall frequency, words vary in how bursty or temporally clustered they may be. This burstiness dimension may be important for understanding both variability across items in their age of acquisition, or even individual variability in the acquisition of specific words. It is a measure of a child’s language environment whose contribution the field has to date been largely unable to measure and is an example of a possible theoretical contribution of daylong audio recordings.

The present analyses that suggest that rather than simply provide more information about a child’s language environment, daylong recordings may provide different information. This different information may allow us to generate and test new predictions and theories regarding the data input with which children learn language and the learning processes that underlie language learning.

Discussion

The present analyses explore the information contained in daylong transcripts of children’s auditory input. The goal is to better understand how these recordings can be used in new ways to extend our understandings of the content of early language environments and how language environments contribute to language learning. We are only beginning to understand the content and structure of language environments at a developmental scale, and work like this may contribute to the development of methodologies and theories for understanding language learning.

The present analyses should be considered a methodological proof-of-concept, rather than concrete descriptions of language environments. It is hard to draw strong conclusions about natural language environments from three transcripts. However, these analyses identify features that vary across children or across words, and how those features may be operationalized with the eventual goal exploring links between language input and language out comes. New constructs and operationalization schemes may be developed when evaluating language input in new ways.

Finally, this work is not meant to suggest that some investigations are preferable to others when describing language environments, merely that analyses of daylong recordings provide different information. Investigations that focus on specific activities: mealtime, playtime, daycare, etc. are crucially important because context-by-context variability is a key feature of children’s language environments. It is the *conjunction* of multiple different investigations that will be most informative of the properties of language input that children use to learn language.

Acknowledgements

This work was supported by NIH Grant R03 HD096157, NSF grant BCS-1749594 and a James S. McDonnell Foundation Scholar Award to JLM.

References

Bergelson, E., Amatuni, A., Dailey, S., Koorathota, S., & Tor, S. (2019). Day by day, hour by hour: Naturalistic language input to infants. *Developmental Science, 22*, e12715.

Bergelson, E., Casillas, M., Soderstrom, M., Seidl, A., Warlaumont, A. S., & Amatuni, A. (2019). What do north American babies hear? A large-scale cross-corpus analysis. *Developmental Science*, *22*, e12724.

Braginsky, M., Sanchez, A. & Yurovsky, D. (2019). childesr: Accessing the 'CHILDES' Database. R package version 0.1.2. https://CRAN.R-project.org/package=childesr

Carvalho, P. F., & Goldstone, R. L. (2015). The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. *PB&R, 22*, 281–288.

Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*, 179-211.

Estes, W. K. (1955). Statistical theory of spontaneous recovery and regression. *Psych. Review, 62*, 145-154.

Fausey, C. M. & Mendoza, J. K. (2018). FauseyTrio HomeBank Corpus. doi: 10.21415/T5JM4R

Gilkerson, J., & Richards, J. A. (2008). *The LENA natural language study*. Boulder, CO: LENA Foundation.

Hart, B., & Risley, T. R. (1995). Meaningful differences in the everyday experiences of young American children. Baltimore: Brookes.

Heaps, H. S. (1978). Information retrieval. Computational and theoretical aspects. New York: Academic Press.

Herdan, G. (1960). Type-token mathematics. Vol. 4. The Hague: Mouton.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., . . . Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science*, *26*, 1071-1083.

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. Cognitive Psychology, 61, 343–365.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*, 211–240.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, *41*, 841-849.

Le Franc, A., Riebling, E., Karadayi, J., Wang, Y., Scaff, C., Metze, F., & Cristia, A. (2018). The ACLEW DiViMe: An Easy-to-use Diarization Tool. In *Interspeech* (pp. 1383-1387). Hyderabad: India.

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. 3rd Edition. Mahwah, NJ: Lawrence Erlbaum Associates.

Mendoza, J. K., & Fausey, C. M. (2019, Nov. 26). Everyday Music in Infancy. https://doi.org/10.31234/osf.io/sqatb

Richards, B. (1987). Type/token ratios: What do they really tell us? *Journal of Child Language*, *14*, 201–209.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child directed speech in vocabulary development. *Child Dev.*, *83*, 1762–1774.

Sanchez, A., Meylan, S. C., Braginsky, M., MacDonald, K. E., Yurovsky, D., & Frank, M. C. (2019). childes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior Research Methods*, *51*, 1928-1941.

Soderstrom, M., & Wittebolle, K. (2013). When do caregivers talk? The influences of activity and time of day on caregiver speech and child vocalizations in two childcare environments. *PloS one*, *8*, e80646.

Montag, J. L., Jones, M. N., & Smith, L. B. (2018). Quantity and diversity: Simulating early word learning environments. *Cognitive Science*, *42*, 375-412.

VanDam, M. (2018). VanDam Public Daylong HomeBank Corpus. doi:10.21415/T5388S

VanDam M, Warlaumont AS, Bergelson E, Cristia A, Soderstrom M, De Palma P., MacWhinney B. HomeBank (2016) An online repository of daylong child-centered audio recordings. *Seminars in Speech & Lan.*, *37*, 128-141.

Weisleder, A., & Fernald, A. (2013). Talking to children matters early language experience strengthens processing and builds vocabulary. *Psych. Science*, *24*, 2143–2152.