Rapid Word Learning in Trilingual Children: One Spurt or Three?

Aaron Albin and Lisa Gershkoff-Stowe

1. Introduction

1.1. The vocabulary spurt

Longitudinal studies that chart the nature and use of children's earliest words consistently note a common pattern of development. Children start out slowly adding one word at a time but then show a sudden acceleration in the rate of learning towards the end of the second year (Bates et al., 1994; Bates, Bretherton, & Snyder, 1988; Bloom, 1973; Dromi, 1987; Gershkoff-Stowe & Smith, 1997; Mervis & Bertrand, 1994). To illustrate, a child with a typical trajectory might produce approximately 10 words at 12 months, 50 words at 18 months, and 500 words at 24 months (Fenson et al., 1994). This translates to a rate of 0.83 words/month from 0-12 months, 6.67 words/month from 12-18 months, and 75 words per month from 18-24 months. In this hypothetical example, the acceleration in rate is exponential, since the change from 0.83 to 6.67 is an 800% increase, and the change from 6.67 to 75 is a 1125% increase. This 'explosion' in the size of a child's productive lexicon is commonly referred to as the vocabulary spurt.

Why the vocabulary spurt occurs, and what metric should be used to identify it, are issues that have been extensively debated in the literature. Indeed, whether the phenomenon exists at all is still a matter of controversy. One problem is that not all children experience a spurt. For example, Goldfield & Reznick (1990) used maternal diaries to track the lexical development of 18 children, beginning when they were 14 months of age until each child reached the 75-word mark. They found considerable variation in individual growth curves, with some children showing a steady, linear increase in new word acquisitions and others exhibiting the sharp increase characteristic of the spurt. A second problem is that a lexical spurt may not be limited to a particular developmental period. For instance, in a recent study, Dandurand & Shultz (2011) used densely-recorded observations to detect statistically significant shifts in linearity in the vocabulary data of 20 children. Parents kept daily records of their children's vocabulary acquisitions beginning at 15 months of age until each child reached the 90-word mark. Using a technique based on Functional Data Analysis, Dandurand & Shultz found that most children experienced multiple spurts at different ages and of varying degrees of intensity. Finally, a third problem associated with studying the vocabulary spurt concerns the methods researchers use to detect nonlinearities in children's word learning. Bloom (2000) has argued that changes in vocabulary growth reflect a gradual continuous increase in the number of words acquired rather than an abrupt change in rate. Furthermore, the criteria used for detecting a vocabulary spurt, according to Bloom, lack objectivity and precision. For example, many researchers require children to attain some arbitrary number of words (e.g., 10 new object words in a 3-week period) rather than to assess the rate of change itself. A similar argument has been made by Ganger & Brent (2004) who questioned whether a 'spurt' refers to a nonlinearity in the overall size of the child's lexicon or in the rate of word learning.

* Aaron Albin, Pennsylvania State University - University Park. Lisa Gershkoff-Stowe, Indiana University - Bloomington. This project benefited from many helpful comments from the audiences at previous presentations of the material, both at the Indiana University Second Language Studies Colloquium (2/20/2015) as well as the 13th Generative Approaches to Second Language Acquisition conference (3/4/2015). This research was supported in part by the Indiana University Pervasive Technology Institute and the Office of the Vice President for Information Technology.

 Proceedings of the 13th Generative Approaches to Second Language Acquisition Conference (GASLA 2015), ed. David Stringer et al., 14-30. Somerville, MA: Cascadilla Proceedings Project.

1.2. Exogenous vs. endogenous accounts

Broadly speaking, two explanations of the vocabulary spurt can be identified - an 'endogenous' and 'exogenous' account. The endogenous account maintains that the spurt is a consequence of changes in how the infant brain develops and learns. Examples of factors that have been cited as possible causes of the spurt include a greater efficiency at processing words for retrieval (Dapretto & Bjork, 2000; Gershkoff-Stowe, 2002), improved word segmentation (Newman et al. 2006; Plunkett, 1993; Walley, 1993), a more refined conceptual categorization of objects (Gopnik & Meltzoff, 1987), and changes in hemispheric specialization (Mills, Coffey-Corina, & Neville, 1993). Indeed, it has been argued that these and other factors may be at work simultaneously, though most likely occurring in an unsynchronized fashion (Mayor & Plunkett, 2010). At a broader level, the underlying assumption tying all such endogenous accounts together is that, at a certain time in development, one or more qualitative and discontinuous changes occur, leading to the spurt.

This position may be contrasted with the *exogenous account*, which maintains that the spurt is merely a byproduct of the natural variation in the time it takes to learn words. For example, words that are less frequent in the child's input may require more time before they are learned. McMurray (2007) uses computational simulations to demonstrate how this factor alone can create a spurt-like pattern in the growth of a child's lexicon. The underlying assumption behind the exogenous account is that rate of word learning is input-driven, such that the spurt happens at a certain point as input is accumulated.

Distinguishing the two accounts is generally a challenge. The core problem is that, with monolingual children, an increase in age is generally accompanied by a concomitant increase in input, thus the two explanations are intertwined. However, the two accounts are more easily distinguishable for simultaneous multilingual children. Such children frequently have an imbalance in the amount of input they receive in each of their languages. Moreover, the ratio of input between the different languages can fluctuate over time (e.g., if a parent is gone on a business trip, or if a relative visits from out of town). Consequently, for simultaneous multilinguals, an increase in age does not necessarily correspond to an increase in input.

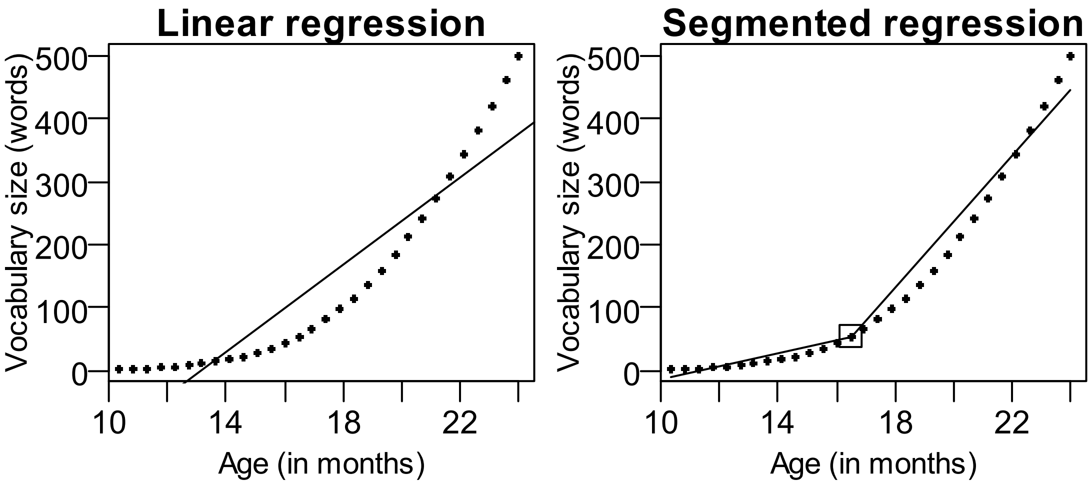
The present study contributes to the debate regarding the mechanisms underlying the vocabulary spurt by analyzing the longitudinal word learning trajectory for each language in two simultaneous trilingual children. In particular, two research questions are addressed. First, *is a spurt observed in all three languages*? Second, for any languages for which a spurt is observed, *which better predicts its timing: accumulated input or child's age*? If a spurt occurs when the child hits a certain threshold of accumulated input for a given language, this would suggest that input is the better predictor, thus supporting a strictly exogenous account. Alternatively, if the timing of the spurt is synchronized across languages (such that it occurs for all languages within some fixed age window), this would suggest that the child's age is the better predictor, thus supporting the endogenous account.

The remainder of this paper describes the study conducted to address these two research questions. Section 1.3 describes how the notion of a vocabulary spurt will be operationalized for the purposes of the present study. Section 2 documents the details on the two children as well as how their input and output was measured. Section 3 presents the empirical results, and Section 4 concludes the paper with a discussion of the theoretical implications.

1.3. Operationalizing the vocabulary spurt

In light of the problems associated with measures of the vocabulary spurt raised in Section 1.1 (e.g., Bloom, 2000; Dandurand & Shultz, 2011; Ganger & Brent, 2004), the notion of a spurt will be operationalized as follows. Consider the hypothetical data in Figure 1.

Figure 1. Two models fit to hypothetical data on vocabulary growth



In both panels, the *x*-axis represents a child's age, and the *y*-axis represents the size of that child's productive vocabulary. The left panel contains a linear regression, whereas the right panel contains a segmented (or 'piecewise') regression. The crucial difference between the two is that, with linear regression, a single line is fit to the data, whereas with a segmented regression, two lines are fit. For the segmented regression, both lines are constrained so that they must pass through the same point in the middle (which in the present context represents the spurt). For both models, it is possible to evaluate how well the model fits the raw data - that is, how accurately the regression line(s) reflect the original individual data points. The approach adopted here is to calculate the difference between each data point and its corresponding location along the regression line, then collapse positive and negative divergences by taking the absolute value. For N original data points, this produces N *absolute residuals*, where lower values indicate a better fit of the model to the data. Regardless of whether the model represents a linear regression or a segmented regression, this approach generates the same number of absolute residuals, each of which is tied to one of the original data points. Given this structure, the absolute residuals for the two models can be compared in a paired *t*-test to determine whether one model systematically fits the data better than the other. By virtue of the fact that the segmented regression involves fitting two lines rather than one, it is automatically guaranteed to have a better fit. However, in certain cases the improvement in fit may be negligible and therefore fail to reach significance in the *t*-test. Alternatively, in the cases where the two-line model fits the data *significantly* better than the one-line model, it may be declared that there is a nonlinearity in the trajectory. For the hypothetical data in Figure 1, this is indeed the case - the absolute residuals are significantly smaller for the two-line model compared to the one-line model. This fact is readily visible in the figure: whereas in the left panel the regression lines are quite far removed from the raw data points in all but a few locations, in the right panel the regression lines track the raw data rather closely. This criterion will be used in the present study to operationalize whether a spurt occurs. If the two-line model fits the data significantly better than the two-line model (for data points spanning some stretch of time), the growth of the child's productive vocabulary will be treated as exhibiting enough of a nonlinearity to be deemed a spurt.[1]

In cases where a spurt thus defined can be discerned, a second, related issue is precisely *when* it occurs. In the segmented regression (two-line model), this translates to the location of the inflection point between the two lines. Throughout the rest of the present paper, this location is determined by applying the algorithm described in Albin & Rankinen (2014). The algorithm loops through every data point and calculates what the resulting absolute residuals would be if that point were the spurt location

[1] At a large enough time scale (e.g., birth to 5 years), the trajectory will almost certainly be nonlinear, hence the 2-line model would be nearly guaranteed to fit better than a 1-line model. This method is optimized for analyzing smaller time scales (such as the 14-month range in Figure 1), at which such a conclusion is far from guaranteed.

(i.e., inflection point). The distribution of absolute residuals for any given candidate spurt location can be summarized by calculating its median (i.e., central tendency). If one candidate spurt location results in a systematically better fit relative to some other candidate, the median would be lower (suggesting the distribution of absolute residuals is shifted toward lower values). By treating every data point as a candidate spurt location and calculating medians in this way, it is possible to determine which candidate inflection point results in the overall best fit to the raw data. In the right panel of Figure 1, the point surrounded by a box (where the two lines meet) was determined to result in the best-fitting model.

In this way, the operationalization of vocabulary spurt is broken into two separate problems (whether it exists, and if so, where it occurs), and each of these problems is given its own quantitative solution.

2. Method

This section overviews the method applied to explore the above research questions. Section 2.1 provides demographic background details on the two children and how frequently they visited the lab. Section 2.2 reports on how estimates of the children's linguistic input was obtained. Section 2.3 describes how the video recordings of interactions in the lab were used to estimate the size of each child's lexicon in his/her three languages.

2.1. Details on the two children

The present study investigates the lexical development of two simultaneous trilingual children acquiring English, Japanese, and either Mandarin or Cantonese. The children were observed over the course of approximately one year through regular visits to the laboratory. Both children were born and reared in the United States, hence the input outside the home was mostly English. The other two languages for each child (Japanese and Mandarin/Cantonese) were the native languages of the parents. While the four parents differed in the extent of their caretaking roles, all were passionate about their child learning their native language. As such, each parent mainly used his/her native language when speaking to the child at home. As a result, the English input the children received from their parents at home was relatively limited in quantity and passive in nature.[2] Moreover, the two parents for a given child did not know each other's languages. Consequently, the overall result is that each child experienced a relatively well-separated division of input, split among [Ambient language = English], [Parent A = Home language 1], [Parent B = Home language 2].

More specific details regarding the two children in the present study are provided in the table below.

Table 1. Details on the two children

Child	Gender	L1 of mother	L1 of father	Start age	End age	Study duration	Number of observations
C	Female	Japanese	Cantonese	12.4 months	24.8 months	12.4 months	14 sessions
K	Male	Mandarin	Japanese	15.6 months	29.2 months	13.6 months	20 sessions

As indicated in the first column ("Child"), the two children will be referred to by the pseudonyms "Child C" and "Child K". The next three columns provide the child's gender and the native language ("L1") of the two parents. The following two columns contain the age of the child at the beginning and

[2] During the lab visits, the parents did frequently address their child in English for the benefit of the other interlocutors in the room. However, since only a small fraction of the child's overall linguistic input occurred during lab visits, this fact is not expected to unduly skew the results.

end of the study.[3] Note that the observation period for both children started well before, and ended well after, the typical spurt age (cf. the 18-20 month range mentioned above in Section 1). Note also that the period of observation began and ended slightly earlier for Child C. The "Study duration" column reflects the end age minus the start age, and the final column ("Number of observations") reports the number of times the child visited the lab. Child C was studied for a shorter overall duration and had fewer total observations. The visits for Child C were also more spaced out, averaging 12.4/14 = 0.89 months between visits as opposed to 13.6/20 = 0.68 months for Child K. For both children, however, the median interval between visits was 21 days (i.e., 3 weeks).

2.2. Data on each child's linguistic input

At each visit to the lab, the parents were asked to indicate what percent of the child's input since the last visit to the lab was in each of the three languages. The exact wording of the instrument was as follows:

Please indicate the following information about your child's language exposure. You are free to leave out any information. We are interested in knowing what percentage of the time you estimate your child hears each language and in what context. For example, on an average day your child is exposed to English 30% of the time (by you, your spouse, and caregivers at his/her daycare), another language 40% of the time by you, and a third language 30% from your spouse. Include as many speakers as needed for each language.

Pooling across both children and the two home languages (Japanese and Mandarin/Cantonese), the specific sources of input that were reported in this task are summarized in Table 2.

Table 2. Reported sources of input in English and the two home languages

Language	**Outside home**	**At home**
English	daycare, play group, parents' workplaces, music/gym/yoga class, YMCA	friends, house guests, babysitter
Home languages	parent, daycare, play group, Saturday school	parent, friends, house guests, family members, roommate, roommate's friends/family, DVDs

From this larger set, the reported sources of input were highly variable from week to week. For example, Child K's maternal grandmother came to visit from Taiwan (thus temporarily increasing input in Mandarin), and Child C's mother went away to a conference (thus temporarily decreasing input in Japanese). As mentioned in the Introduction, such fluctuations in the balance of input between the three languages help create a dissociation between biological age and linguistic input, thus making it possible to tease apart the endogenous and exogenous accounts. As such, the variability in input from week to week is essential for the purposes of the present study (rather than merely a source of noise).

2.3. Data on each child's lexical output

During each visit to the lab, in addition to completing the input estimate task, parents also engaged their child in structured play activities. For example, a set of farm toys was always in the lab, creating a consistent environment from week to week. Children's books were also read during many of the lab visits - either English ones in the lab or ones brought from home in the child's home languages. The

[3] Age was calculated in terms of intervals between the child's birthday each month. For example, if the child was born on February 14th, then on March 14th the child was counted as 1 month old. Decimals reflect fractions across such intervals, e.g., 15 days into a 30-day period between the birthday in two adjacent months would be 0.5.

duration of these play sessions varied from week to week between the two children but typically ranged from 30 minutes to 1 hour.

Both parents were present for most of the lab sessions, and each used their L1 in most of their child-directed speech. Parents were encouraged to interact with their child as naturally as possible during each session. In the few cases when a parent was absent, a graduate student lab assistant who spoke the same L1 as the missing parent filled in for him or her. There was also always at least one native English speaker in the lab at every session who spoke English exclusively - typically the second author and sometimes also a lab assistant. This structuring of the interlocutors ensured that the child had an equal opportunity to use all three languages during every session.

During each session, the child's interactions with his/her parent(s) and the other interlocutors in the room were video recorded by an additional lab assistant (who did not interact directly with the child). A team of undergraduate and graduate student transcribers then went through these video recordings and transcribed all child utterances and child-directed speech contained therein. Students who were native speakers of the various languages in question were recruited to transcribe the corresponding portions of the recordings. The transcriptions thus created followed CHAT conventions, the standard conventions for the CHILDES corpus exchange platform, modified as needed to facilitate trilingual data. The transcription process was extremely labor intensive, taking three years to complete, beginning in 2011 and ending in 2014. The aggregate length of the transcripts is over 90,000 lines (approximately 32,000 for Child C and 58,000 for Child K).

Each main tier line in the transcript was tagged for language via three-letter codes (`eng`/`jpn`/`zho`/`yue` for English/Japanese/Mandarin/Cantonese). Cases where it was ambiguous which language was being spoken were explicitly marked as such with the language code `amb`. Such cases were generally quite rare since each child's three languages are typologically unrelated (and therefore there is only minimal shared cognate vocabulary between the three.) An example of a main tier line is `*MOT: [- jpn] mama koko itai.`, where `MOT` indicates the mother is speaking, `jpn` indicates the language being used is Japanese, and `mama koko itai` is the utterance itself. For all such non-English lines of the transcript, English translations were added in a secondary tier, e.g., `%eng: Mama hurts here` for the above example.

Each main tier line in the transcript was analyzed for morphological structure using the MOR ("morphology") function in the CLAN software. For each of the four languages in question (English, Japanese, Mandarin, and Cantonese), all lines tagged as that language were analyzed using a corresponding morphology grammar specific to that language.[4] The output of the morphology analysis for a given main tier line was added to the transcript as an additional secondary tier line (beginning with `%mor:`). Using MOR, the three words in the above example were analyzed as `n|mama=mother` (the noun *mama* "mother"), `n:deic:dem|koko=here` (the deictic demonstrative noun *koko* "here"), and `adj|ita-PRES=aching` (the present-tense form of the adjective *ita-* "aching"). Note that the last word in this example has been lemmatized, i.e., abstracted away from its specific form and associated with a higher-level unit (all conjugations of the adjective *ita-*). In a similar way, the English words *played* and *playing* were identified as forms of *play*.

For the purposes of quantifying each child's lexical output, the FREQ ("frequency") command in CLAN was used to analyze all main tier lines marked `*CHI:` (produced by the child) except those marked as `amb` (ambiguous). With the `%mor:` tiers added to the transcript as described above, FREQ provides counts of not only word *tokens* but also word *types* (using the lemmatization information just discussed). Since the goal of the present study is to chart changes in the children's *overall lexicon size*, the latter is of interest, since it only counts all unique words the child produces in a given session (regardless of how many times any given word happened to be said). For this reason, each child's lexical output was quantified as word types rather than word tokens.

---

[4] The English morphology grammar is included with the base distribution of CLAN. The morphology grammars for Japanese, Mandarin, and Cantonese are publicly distributed at http://childes.psy.cmu.edu/morgrams/. The lexicon files forming the backbone of these grammars were modified as needed to accommodate the specific words used in the present dataset (e.g., idiosyncratic words used within one specific family or low-frequency words missing from the original lexicon files).

Since the present study tracks growth in overall lexicon size over time, what matters is how many *new* word types the child produces in a given transcript (relative to the previous one). For example, if the child produces *dog* at sessions 2 and 4 but not at session 3, *dog* may be counted as a "new" word at session 2 (and session 3 may be treated as an accidental gap).[5] To extract this information from the transcripts, rather than running the FREQ command on each individual transcript for a given session, it was run on *the collective set of all transcripts up to that session*. For example, at session 5, rather than running FREQ on the fifth transcript, it was run on all text from the first five transcripts combined. This analysis therefore shows how the *corpus of everything the child has ever produced* evolves over time, with changes from session to session reflecting only the new word-types added to the corpus.

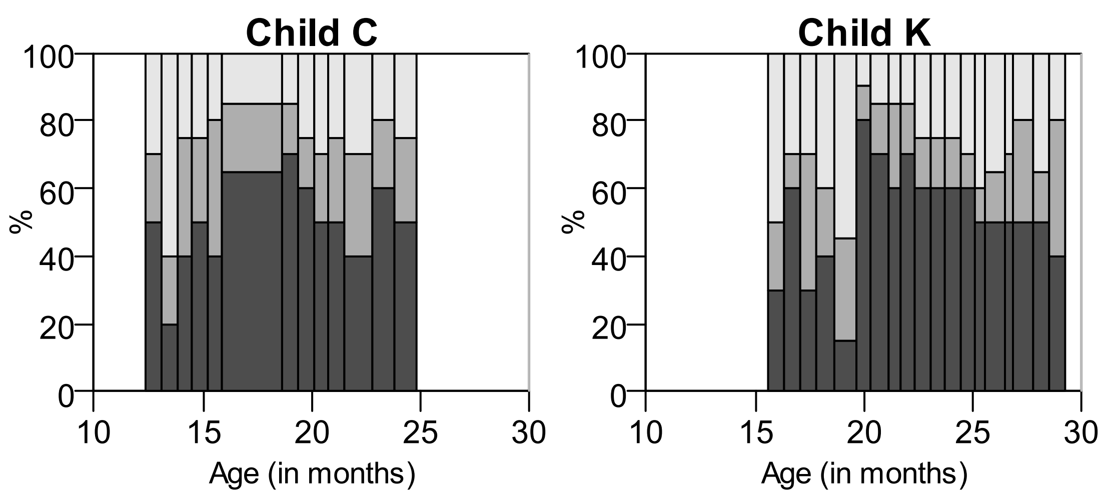
3. Results

The following three sections report the empirical results of the present study. Section 3.1 analyzes the data collected from parents' self-reports on the children's linguistic input. Section 3.2 explores the extent to which this information on *linguistic input* can predict the location of spurts in each child's three languages. Section 3.3 examines whether information on the children's *age* at each lab visit can be used to predict the timing of the spurts.

3.1. Input

The raw data on each child's linguistic input are provided in the columns labeled "Input (proportions)" in the Appendix. The following plots illustrate what these data look like in their original form as proportions (e.g., 20% language A, 40% language B, and so on). Each thin vertical line dividing two adjacent stacks of bars corresponds to a visit to the lab. The relative heights of the bars correspond to the distribution of input during a given range of time (i.e., from one lab session to the next). The darkest grey bars on the bottom correspond to English, the intermediate grey bars in the middle correspond to Japanese, and the lightest grey bars on the top correspond to Mandarin/Cantonese. The sharp increase in English input for Child K around 20 months occurred when he started attending full-time daycare.

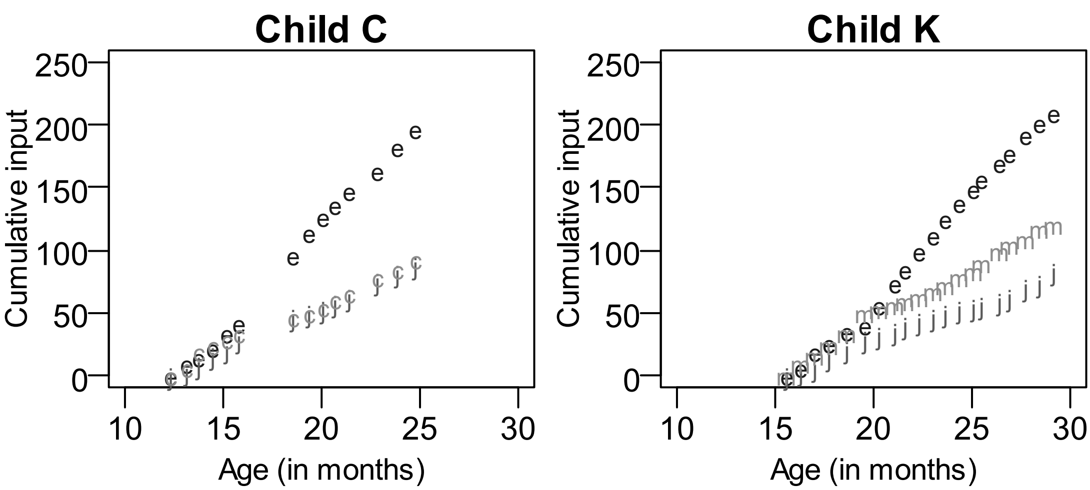
Figure 2. Raw data on linguistic input (Proportion between adjacent sessions)



[5] Of course, counting a word as "known" whenever the child first produces it is an overly simplistic criterion, since a child may produce a word without fully comprehending its meaning (and since a word's meaning itself can often be quite complex). However, this is one of the only ways word knowledge can be operationalized given the available data for the present study (transcripts of recordings spaced roughly three weeks apart).

These data can be translated into a more useful format for the purposes of analysis. If, for example, Child K's parents reported that he received 50% English, 25% Japanese, and 25% Mandarin input over the course of the 20 days between two sessions, this could be recoded as 10 days English input, 5 days Japanese input, and 5 days Mandarin input. This transformation simply takes the percentages of input between sessions and multiplies it with the number of days between sessions (e.g., 50% English * 20 days = 10 days of English input). By expressing the output of this transformation as a cumulative sum (e.g., rather than {1, 2, 3}, taking {1, 1+2, 1+2+3}), the data can be transformed into units that represent the *days of accumulated linguistic input since the child's first visit to the lab*. The data thus transformed are listed in the columns labeled "Input (cumulative days)" in the Appendix. In this format, the accrual of input over time looks like the following. Languages are indicated by the first letter of their names: [e]nglish, [j]apanese, [m]andarin, and [c]antonese.

Figure 3. Cumulative linguistic input (in days)

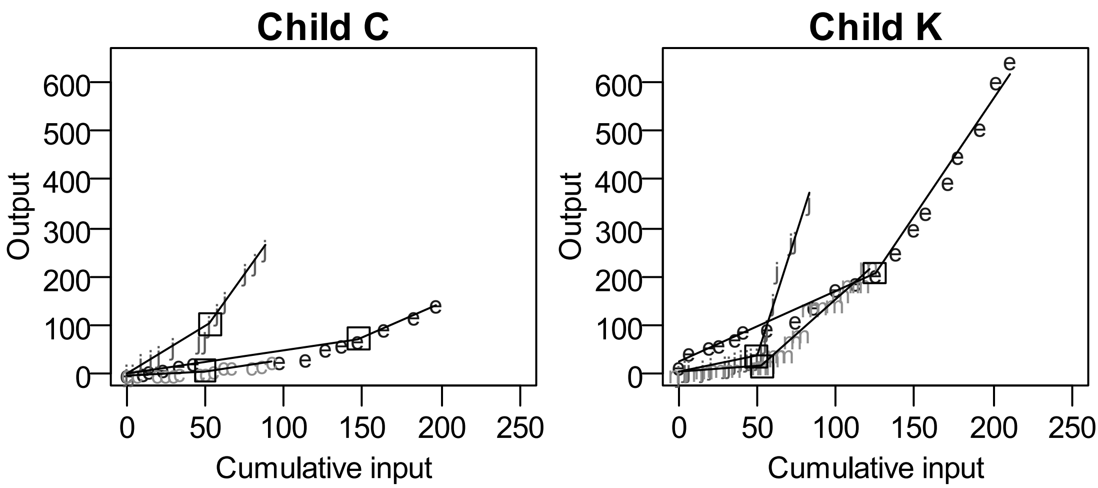


This representation makes it clear that both children accumulated substantially more English input than either of their home languages. Considering that English is the ambient language for both children, this finding is not surprising. Child C accumulated a balanced amount of input in her two home languages (Japanese and Cantonese). In contrast, Child K accumulated slightly more input in Mandarin than in Japanese.

3.2. Output predicted by input

The following plot shows what it looks like if this measure of cumulative linguistic input is used as a predictor for the child's lexical *output*. The *x*-axis ("Cumulative input") represents the same data as the *y*-axis of the previous plots (Figure 3). The fact that English extends further to the right than the two home languages reflects the skewing towards English input just discussed. The *y*-axis ("Output") represents the cumulative count of word types at each time point (as established in Section 2.3). The data on the children's lexical output are provided in the columns labeled "Output (cumulative types)" in the Appendix.

Figure 4. Output predicted by input

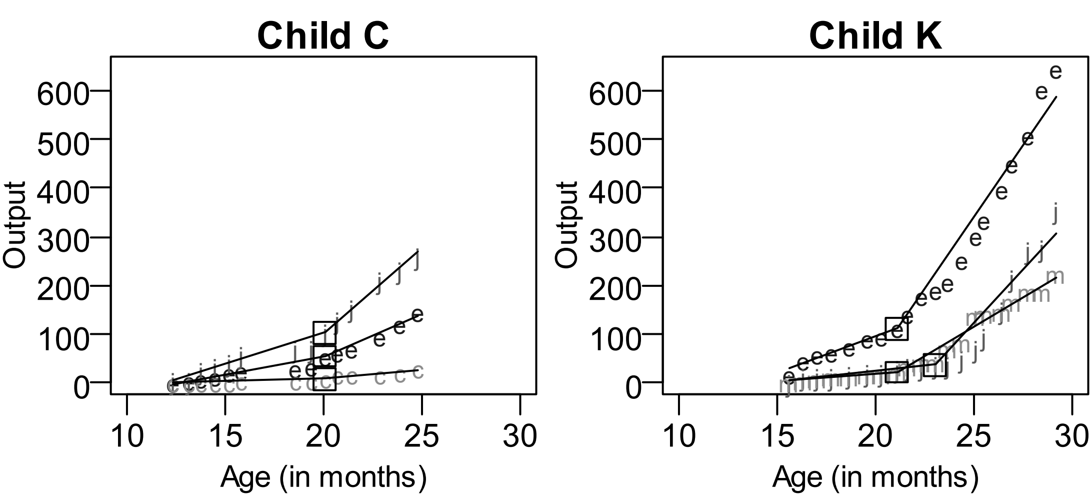


The lines in Figure 4 reflect the result of applying the algorithm described in Section 1.3. More specifically, for each language, regression lines were fit to the left and right of each of Child C's 14 data points, as well as each of Child K's 20 data points Since two points are needed to fit a regression, the leftmost and rightmost points (i.e., the first and last lab visits) were excluded, thus leaving 14-2=12 candidate spurt locations for Child C and 20-2=18 candidates for Child K. The spurt points creating the best-fitting models (with the lowest median absolute residual) are indicated in Figure 4 with a box (as was done in Figure 1). This analysis makes it clear that spurts occur for both children's home languages at around 50 days of accumulated input. The spurt for English occurs much later, at 125-150 days. This dissociation in timing between the three languages is somewhat surprising under a strictly exogenous account predicting spurts to be tied to the input.

3.3. Output predicted by age

The following plots show the same data except the *x*-axis has been swapped out for the child's age at each session.

Figure 5. Output predicted by age



Represented in this way, the timing of the spurts is much better synchronized. For Child C, the best-fitting spurt point occurs at session 9 (20.1 months) for all three languages. For Child K, it occurs at session 9 (21.1 months) for English and Mandarin and session 11 (23 months) for Japanese.

Details on the changes in the slope of the line before vs. after the inflection point appear in the following two tables. All changes are over 200% (double), and one even exceeds 1000% (i.e., accelerating to over 10 times the original rate)

Table 3. Change in slope for Child C

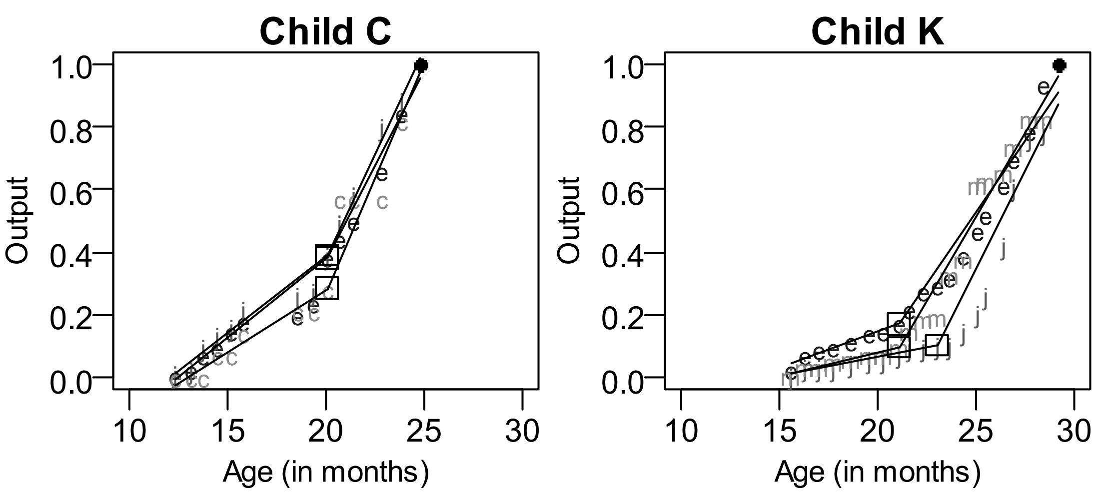
Language	Words per month Before	Words per month After	Percent change
English	90	218	242%
Japanese	159	437	275%
Cantonese	14	51	364%

Table 4. Change in slope for Child K

Language	Words per month Before	Words per month After	Percent change
English	203	798	394%
Japanese	58	597	1021%
Mandarin	47	325	695%

The patterns in Figure 5 are somewhat obscured by the fact that the three trajectories rise to different heights (indicating differing sizes of the children's productive lexicons in their three languages). This information can be normalized out if the *y*-axis is re-scaled in terms of percentages relative to the last session. For example, by Session 11, Child K had demonstrated knowledge of 37 word types in Japanese, which, relative to his 355 word types at Session 20, is 37/355 or 10.4%. If the data are normalized in this way, as in the following plots, the similarity in the trajectory for the three languages becomes even clearer.

Figure 6. Output predicted by age (Percentages relative to last session)



At chance, the probability of all three spurt points occurring at the same session for Child C (among her 12 sessions in the middle, excluding the first and last ones) would be $1/12^3$, or p=0.0006. In a similar vein, by calculating all logically possible permutations, it can also be demonstrated that the probability of all three spurt points falling within two sessions of each other for Child K is p=0.05. Thus, for both children, it is possible to reject the null hypothesis that the synchronicity in spurt locations is due to chance alone.

As discussed in Section 1.3, by comparing the goodness-of-fit of a linear regression vs. a segmented regression (i.e., a one-line vs. a two-line model), it is possible to determine whether the observed nonlinearity in the trajectory for each combination of language and child is significant. The two-line model achieves a significantly better fit in 5 out of 6 cases (with the only exception being Child C's Cantonese data). The fact that, as a whole, a two-line model fits significantly better than a single-line model suggests that there is indeed some sort of a nonlinearity (a 'spurt') occurring in the data.

4. Discussion

Having surveyed all of the empirical results, it is now possible to revisit the two research questions asked at the outset of the present study. This is the focus of Sections 4.1 and 4.2 below, which revisit the first and second research questions, respectively. Section 4.3 explores the broader theoretical implications of these findings. Section 4.4 discusses the present study's limitations and directions for future research. Section 4.5 concludes the paper with a brief summary of the main findings.

4.1. Research Question 1 (A spurt found for all languages?)

The first research question was, "Is a spurt observed in all three languages?" As established in Section 1.3, the relative fit of the one- and two-line models can be used as a criterion for whether the word learning trajectory shows a notable nonlinearity. In five out of six cases, the two-line model (segmented regression) did indeed fit the data significantly better than the one-line model (linear regression). Thus, in the majority of cases, the two children under examination in the present study did indeed show spurts for each separate language.

The one exception is Child C's Cantonese data, for which the two-line model did not fit significantly better than the one-line model. Consequently, under this definition, no strong evidence for a spurt could be found. This is likely tied to the relatively small size of her vocabulary for that language, having only demonstrated productive knowledge of 28 words by the last visit to the lab. With a sample that small, there is simply not enough room for a robust nonlinearity to manifest itself. Thus, the apparent spurt for Cantonese in the left panel of Figure 6 may be illusory, as it is based on relative rates of change in a small sample (Bloom, 2000). Note, for example, that at the time of the spurt (20.1 months), Child C had produced 103 unique words in Japanese, 55 in English, and only 8 for Cantonese. These data suggest Child C was a relatively unbalanced trilingual (at least for the duration of the study); hence, her Cantonese data cannot be compared alongside that of her other two languages on equal grounds.

For comparison, at the time of Child K's spurts (21.1 months for English/Mandarin and 23 months for Japanese), he had produced 141 unique words in English, 37 in Japanese, and 33 in Mandarin. This finding is interesting considering that, in the child language literature, a spurt is generally assumed to occur close to the 50 word mark (Bates, Bretherton, & Snyder, 1988). Child K's Japanese crosses the 50-word mark at 24.4 months, two sessions after the spurt for that language. Likewise, his Mandarin does so at 23 months, three sessions after the spurt for that language.

Thus, with both children taken collectively, as many as half of the spurts (3 out of 6) occurred prior to the often cited 50-word threshold. On the one hand, huge individual differences are found in the size and rate of vocabulary growth in both monolingual and bilingual children (D'Odorico et al. 2001; De Houwer, 2009; Goldfield & Reznick, 1990; Pearson, Fernandez, & Oller, 1993). Indeed, both Child C and Child K fall within the broad range of data reported by Fenson et al. (1994) on typically developing, monolingual children. On the other hand, it is also possible that the 50-word criterion, originally developed in a monolingual context, may be ill-suited for identifying a word spurt in trilinguals. If, hypothetically, a monolingual and trilingual child both show spurts at the same age, then the fact that a trilingual child's learning efforts are split between multiple languages makes it likely that the spurt will occur at a smaller overall vocabulary size for any given language. Support for this idea can be found in the study by Pearson & Fernandez (1994), which examined individual patterns of vocabulary growth in 18 Spanish-English bilingual children between 8 and 30 months of age. Using the criterion of 20 new words reported in a month as their definition of a vocabulary spurt, they found that none of the children showed a spurt in both languages at once. Rather, children appeared to concentrate growth in one language at a time. As the authors noted, "to have a spurt in each language would require learning 40 new words a month including at least 20 words a month in the second language as well" (p. 639).

4.2. Research Question 2 (Predicting spurt timing: input or age?)

The second research question asked, "For any observed spurt, which better predicts its timing: accumulated input or child's age?" In the present study, there was no threshold of accumulated input at which spurts occurred (Figure 4). Despite accumulating more input overall in English (Figure 3), this did not lead to an earlier spurt for that language in either child. In comparison, age was found to be much better at predicting the timing of the spurt (Figures 5-6). Child C's Japanese spurt was synchronized with her English spurt, and the spurts for Child K's three languages were also all relatively close to each other (at an order of magnitude tighter than the spread in Figure 4). These results suggest that sharper rates of vocabulary growth are found for all three languages after a certain *age range*, not a certain input threshold.

Finally, note how the age range for spurts in the present study (20-23 months) is comparable to that of monolingual children reported in the literature (Benedict, 1979; Bloom, 1973; Nelson, 1973), especially given the individual variation noted above. Assuming that the *overall* amount of linguistic input for the two children (pooling across all three languages) is roughly comparable to that of a monolingual child, since the input is split among three languages, they necessarily received much less input *per language* than a monolingual. As such, if spurts were driven entirely by input, there would most likely be a substantial delay for trilinguals relative to monolinguals. The present study, however, found this not to be the case. Instead, the results are compatible with an endogenous account in which one or more developmental changes occurring at the age range in question, regardless of the number of languages spoken by the child.

4.3. Implications

The findings from this study make two important contributions to the child language literature. First, given that relatively few empirical studies have been conducted on trilingual child language acquisition, the present study provides new data on the pace and timing of lexical growth in two simultaneous trilinguals during the second year of life. Specifically, both children's growth in each language was found to follow a nonlinear pattern, reflecting a shift from slow to fast new word acquisition. This result is consistent with patterns of lexical development in monolingual and bilingual children (D'Odorico et al. 2001; De Houwer, 2009; Pearson, Fernandez, & Oller, 1993). However, differences were also found between the two children in terms of the precise timing (i.e., age of onset) and number of accumulated words at the point of nonlinearity, as might be expected given the large intra-individual differences reported in the literature. Finally, and perhaps most importantly, although slight variations in the onset and number of words produced at the time of spurt were found to occur within the same child across the three languages, these were not tied to the amount of linguistic input received.

This last point highlights the second main contribution of the present study: using the same child to make comparisons across languages. This unique design controls for individual differences while varying the nature of the learning environment (and therefore the amount of accumulated input) for each language. Accordingly, this made it possible to test different theoretical predictions about the underlying source of the word spurt. Recall from Section 1.2 that two broad lines of interpretation, referred to here as the "endogenous" and "exogenous" approaches, have attempted to account for the mechanisms driving rapid vocabulary growth. The present study tested the predictions of these two accounts by analyzing the time course of vocabulary development in two trilingual children. Analyses revealed that age is a better predictor of spurt timing than accumulated input.

The finding is compatible with an endogenous account predicting qualitative changes in cognitive, neurological, and/or linguistic abilities near the end of the second year. In contrast, the finding that age is a more reliable predictor of the spurt is hard to explain under an exogenous approach that takes spurts to be strictly driven by linguistic input. Of course, this does not imply that factors such as amount of input or word frequency are irrelevant to vocabulary acquisition. Indeed, there is ample evidence to suggest that word learning crucially depends on opportunities to hear words in a social context (e.g., Hart & Risley, 1995). The present study simply demonstrates that these factors *by themselves* are not sufficient to explain the attested patterns. Rather, the shift from slow to fast rates of lexical growth may be a confluence of multiple unsynchronized changes in domain-specific knowledge and skills.

4.4. Limitations and future directions

The present study is not without its limitations. The data presented above are based on only two children, both of whom are being raised by academic parents who are particularly interested in their child's language development. Thus, it is unclear how representative the participants in the present study are to other children growing up in trilingual households. Another methodological concern is that the data collected during regular visits to the laboratory are, at best, approximations of the children's lexical development. Additional data from the MacArthur-Bates Communicative Inventory (Fenson et al., 2007) as well as diary data kept by parents at home, though not yet fully analyzed, will provide a more reliable assessment of the children's vocabulary growth. Finally, as with most case study reports, parents' descriptions of the amount of input children received in each language were anecdotal. As such, it is necessary to proceed with caution when interpreting the present results.

With respect to the children's linguistic input, the present study focused exclusively on the *quantity* of the input. Additional analyses with respect to the *quality* of input might also prove insightful. For example, the child's language choices may have been impacted by whether parents tended to use explicit strategies (such as instructing the child to say a word or correcting the child) or implicit strategies (such as repeating, translating, or simply "moving on") (Quay, 2001).

Given the vastness of the current dataset, there are a number of other theoretically important avenues for future investigation. For instance, an analysis of when translation equivalents (e.g., words meaning *dog*) first emerge in the three languages relative to each other could reveal an interesting hierarchy, perhaps tied to the relationship between the three languages in that child's input or output. Additionally, by splitting up the analysis into parts of speech, it may be possible to determine the presence of a part-of-speech bias (e.g., toward nouns or verbs) in each of the child's three languages. Other language variables, such as mean length of utterance (MLU), type/token ratio, and frequency of word usage, hold promise for a better understanding of multilingual development as well as elucidating fundamental processes in the development of language such as the vocabulary spurt.

The present study can thus be viewed as a first step in a much larger project. The 90,000-line transcript on which the current analyses are based is a valuable resource for understanding the complicated dynamics behind simultaneous trilingual acquisition. Clearly, the analyses above have only begun to tap its immense potential.

4.5. Conclusion

The goal of this paper was to test the claim that the marked acceleration in lexical growth observed in most children towards the end of the second year arises out of the statistical properties of word distributions in the child's language environment (McMurray, 2007). A novel approach based on statistical model comparison was used to identify if and when a shift in rate of vocabulary growth occurred in two simultaneous trilingual children. The results showed that, given a sufficient critical mass of words (Marchman & Bates, 1994), both children achieved a lexical spurt in all three languages, independent of their exposure patterns - contrary to the exogenous account of the vocabulary spurt. Thus, to address the question posed in the present paper's title ("One spurt or three?"), the answer appears to be that trilingual children exhibit *one simultaneous spurt in all three languages*.

Appendix

Cells in bold and underline indicate spurt location (as determined by the algorithm from Section 1.3).

Language codes: eng = English, jpn = Japanese, yue = Cantonese, zho = Mandarin

Full data for Child C

Session number	Age (months)	Input (proportions)			Input (cumulative days)			Output (cumulative types)		
		eng	jpn	yue	eng	jpn	yue	eng	jpn	yue
1	12.4	N/A	N/A	N/A	0	0	0	1	4	0
2	13.1	0.5	0.2	0.3	10.5	4.2	6.3	4	7	0
3	13.8	0.2	0.2	0.6	14.7	8.4	18.9	10	25	0
4	14.5	0.4	0.35	0.25	23.1	15.75	24.15	14	35	2
5	15.2	0.5	0.25	0.25	33.6	21	29.4	21	43	2
6	15.9	0.4	0.4	0.2	42	29.4	33.6	26	57	4
7	18.6	0.65	0.2	0.15	96.6	46.2	46.2	29	70	6
8	19.4	0.7	0.15	0.15	113.4	49.8	49.8	34	73	6
9	20.1	0.6	0.15	0.25	126	52.95	55.05	**<u>55</u>**	**<u>103</u>**	**<u>8</u>**
10	20.8	0.5	0.2	0.3	136.5	57.15	61.35	64	129	16
11	21.5	0.5	0.25	0.25	147	62.4	66.6	72	153	16
12	22.8	0.4	0.3	0.3	163.8	75	79.2	95	211	16
13	23.8	0.6	0.2	0.2	182.4	81.2	85.4	122	232	23
14	24.8	0.5	0.25	0.25	196.4	88.2	92.4	144	263	28

Full data for Child K

Session number	Age (months)	Input (proportions)			Input (cumulative days)			Output (cumulative types)		
		eng	jpn	zho	eng	jpn	zho	eng	jpn	zho
1	15.6	N/A	N/A	N/A	0	0	0	19	3	2
2	16.3	0.3	0.2	0.5	6.3	4.2	10.5	47	13	7
3	17	0.6	0.1	0.3	18.9	6.3	16.8	58	13	8
4	17.8	0.3	0.4	0.3	25.8	15.5	23.7	63	13	12
5	18.6	0.4	0.2	0.4	36.2	20.7	34.1	77	17	13
6	19.6	0.15	0.3	0.55	40.7	29.7	50.6	93	20	16
7	20.3	0.8	0.1	0.1	56.7	31.7	52.6	97	22	16
8	21.1	0.7	0.15	0.15	74.2	35.45	56.35	112	28	22
9	21.7	0.6	0.25	0.15	85	39.95	59.05	**141**	34	**33**
10	22.4	0.7	0.15	0.15	99.7	43.1	62.2	179	34	42
11	23	0.6	0.15	0.25	112.3	46.25	67.45	191	**37**	44
12	23.7	0.6	0.15	0.25	124.9	49.4	72.7	207	37	73
13	24.4	0.6	0.15	0.25	137.5	52.55	77.95	251	52	85
14	25.1	0.6	0.1	0.3	149.5	54.55	83.95	303	73	138
15	25.6	0.5	0.1	0.4	157	56.05	89.95	335	94	141
16	26.5	0.5	0.15	0.35	170.5	60.1	99.4	396	150	147
17	26.9	0.5	0.2	0.3	177.5	62.9	103.6	451	215	165
18	27.8	0.5	0.3	0.2	191.5	71.3	109.2	510	272	186
19	28.5	0.5	0.15	0.35	202	74.45	116.55	604	279	186
20	29.2	0.4	0.4	0.2	210.4	82.85	120.75	646	355	225

References

Albin, Aaron, & Rankinen, Wil (2014). A "pivot" model for extracting formant measurements based on vowel trajectory dynamics. *Journal of the Acoustical Society of America*, *136*(4), 2082. https://github.com/usagi5886/pivot

Bates, Elizabeth; Bretherton, Inge; & Snyder, Lynn Sebestyen (1988). *From First Words to Grammar*. New York, New York: Cambridge University Press.

Bates, Elizabeth; Marchman, Virginia; Thal, Donna; Fenson, Larry; Dale, Philip; Reznick, J. Steven; Reilly, Judy; & Hartung, Jeff (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, *21*(1), 85–123.

Benedict, Helen (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, *6*(2), 183–200.

Bloom, Lois (1973). *One Word at a Time: The Use of Single Word Utterances Before Syntax*. The Hague, The Netherlands: Mouton.

Bloom, Paul (2000). *How Children Learn the Meanings of Words*. Cambridge, Massachusetts: MIT Press.

D'Odorico, Laura; Carubbi, Stefania; Salerni, Nicoletta; & Calvo, Vincenzo (2001). Vocabulary development in Italian children: A longitudinal evaluation of quantitative and qualitative aspects. *Journal of Child Language*, *28*(2), 351-372.

Dandurand, Frédéric, & Shultz, Thomas R. (2011). A fresh look at vocabulary spurts. In Carlson, Laura; Hölscher, Christoph; & Shipley, Thomas F. (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 1134-1139). Austin, Texas: Cognitive Science Society.

Dapretto, Mirella, & Bjork, Elizabeth L. (2000). The development of word retrieval abilities in the second year and its relation to early vocabulary growth. *Child Development*, *71*(3), 635-648.

De Houwer, Annick (2009). *Bilingual First Language Acquisition*. Bristol: Multilingual Matters.

Dromi, Esther (1987). *Early Lexical Development*. Cambridge, UK: Cambridge University Press.

Fenson, Larry; Marchman, Virginia A.; Thal, Donna J.; Dale, Philip S.; Reznick, J. Steven; & Bates, Elizabeth (2007). *The MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual* (2nd ed.). Baltimore: Paul H. Brookes Publishing Company.

Fenson, Larry; Dale, Philip S.; Reznick, J. Steven; Bates, Elizabeth; Thal, Donna J.; & Pethick, Stephen J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development 59*(5), 1-173.

Ganger, Jennifer, & Brent, Michael R. (2004). Reexamining the vocabulary spurt. *Developmental Psychology*, *40*(4), 621-632.

Gershkoff-Stowe, Lisa (2002). Object naming, vocabulary growth, and the development of word retrieval abilities. *Journal of Memory and Language*, *46*(4), 665-687.

Gershkoff-Stowe, Lisa, & Smith, Linda B. (1997). A curvilinear trend in naming errors as a function of early vocabulary growth. *Cognitive Psychology*, *34*(1), 37–71.

Goldfield, Beverly A., & Reznick, J. Steven (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*(1), 171–183.

Gopnik, Alison, & Meltzoff, Andrew (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development*, *58*(6), 1523–1531.

Hart, Betty, & Risley, Todd. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore: Paul H. Brookes Publishing.

Marchman, Virginia, & Bates, Elizabeth (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, *21*(2), 339–366.

Mayor, Julien, & Plunkett, Kim. (2010). Vocabulary spurt: Are infants full of Zipf? In Ohlsson, Stellan, & Catrambone, Richard (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 836-841). Austin, TX: Cognitive Science Society.

McMurray, Bob. (2007). Defusing the childhood vocabulary explosion. *Science*, *317*(5838), 631. http://www2.psychology.uiowa.edu/Faculty/Mcmurray/recent/science/

Mervis, Carolyn B., & Bertrand, Jacquelyn (1994). Acquisition of the novel name-nameless category (N3C) principle. *Child Development*, *65*(6), 1646–1662.

Mills, Debra L.; Coffey-Corina, Sharon; & Neville, Helen J. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology*, *13*(3), 397-445.

Nelson, Katherine (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, *38*(1-2), 1-135.

Newman, Rochelle; Ratner, Nan Bernstein; Jusczyk, Ann Marie; Jusczyk, Peter W.; & Dow, Kathy Ayala (2006). Infants' early ability to segment the conversational speech signal predicts later language development: A retrospective analysis. *Developmental Psychology*, *42*(4), 643-655.

Pearson, Barbara Zurer; Fernández, Sylvia C.; & Oller, D. Kimbrough. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning*, *43*(1), 93-120.

Pearson, Barbara Zurer, & Fernández, Sylvia C. (1994). Patterns of interaction in the lexical development in two languages of bilingual infants. *Language Learning 44*(4), 617-653.

Plunkett, Kim (1993). Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language*, *20*(1), 43-60.

Quay, Susan (2001). Managing linguistic boundaries in early trilingual development. In Cenoz, Jasone, & Genesee, Fred (Eds.), *Trends in Bilingual Acquisition* (pp. 149-199). Amsterdam: John Benjamins.

Walley, Amanda C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review*, *13*(3), 286–350.

Proceedings of the 13th Generative Approaches to Second Language Acquisition Conference (GASLA 2015)

edited by David Stringer, Jordan Garrett, Becky Halloran, and Sabrina Mossman

Cascadilla Proceedings Project   Somerville, MA   2016

Copyright information

ISBN 978-1-57473-470-6 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Albin, Aaron and Lisa Gershkoff-Stowe. 2016. Rapid Word Learning in Trilingual Children: One Spurt or Three? In *Proceedings of the 13th Generative Approaches to Second Language Acquisition Conference (GASLA 2015)*, ed. David Stringer et al., 14-30. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3272.