Mothers' Work Status and 17-month-olds' Productive Vocabulary

Abstract

Literature examining the effects of mothers' work status on infant language development is mixed, with little focus on varying work-schedules and early vocabulary. We use naturalistic data to analyze the productive vocabulary of 44 17-month-olds in relation to mothers' work status (Full-time, Part-time, Stay-at-home) at 6 and 18 months. Infants who experienced a combination of care from mothers and other caretakers had larger productive vocabularies than infants in solely full-time maternal or solely other-caretaker care. Our results draw from naturalistic data to suggest that this care combination may be particularly beneficial for early lexical development.

Keywords: language development, maternal work status, mother-child relations

Mothers' Work Status and 17-month-olds' Productive Vocabulary

North America and Europe have seen a marked increase in women's part-time employment over the past 20 years. In 2017, 68% of British women with children under age 2 worked, compared with <50% a decade earlier (Office of National Statistics, 2017). For the first time since such records have been kept, more mothers are working part-time than are staying home: in 2017, 40% of women with children under 2 worked part-time, while 32% stayed at home (Office of National Statistics, 2017). Trends are similar in the USA: 62% of mothers with children under 2 worked in 2016, of which 26% were part-time (Bureau of Labor Statistics, 2017).

How these changes affect language development is unclear. A large-scale study by the National Institute of Child Health and Human Development (NICHD, 2000) found no difference between the linguistic abilities of infants spending more hours in childcare (i.e. infants whose mothers worked longer hours) and those receiving exclusively maternal care. In fact, infants attending childcare had better linguistic skills than those in full-time maternal care when the childcare was of high *quality* (incorporating a range of environmental measures). This may be due to maternal input: Booth and colleagues (2002) found that while non-working mothers spent more time with their infants, full-time-working mothers spent more hours in high quality interactions (measured by e.g. responsiveness and affection). Language outcomes at 1;5 did not differ, suggesting that working mothers offset time spent away with linguistically-supportive interactions at home.

Timing of mothers' return to work may also influence outcomes. Han, Waldfogel and Brooks-Gunn (2001) found that children whose mothers returned to work in the nine months after birth (rather than thereafter) had lower vocabulary scores 3 years later. While Han et al.

found no effect of work status after nine months of age, Gregg, Washbrook, Proppoer and Burgess (2005) found that full-time, but not part-time, maternal employment before 18 months had a small negative effect on childhood cognitive outcomes. Similarly, Waldfogel, Han and Brooks-Gunn (2002) analyzed a large-scale sample to show that >21 hours/week of maternal employment in children's first postnatal year may affect vocabulary outcomes at age 3-4 years.

Taken together, previous work suggests there may be an advantage for care from other caretakers in the first two years, but maternal care appears to be particularly important early on. These results generally rely on off-line measures taken during large-n epidemiological studies of diverse populations with many interacting variables (e.g. Gregg et al., 2005; NICHD, 2000). We take a different approach: within a smaller, more homogenous sample (n=44), we directly measure infants' language productions through home recordings and parental vocabulary reports. This provides a focused view of maternal work status (full-time, part-time, stay-at-home) between 6 and 18 months and language development. This analysis is exploratory, but based on previous literature, we expect to see a lexical advantage for infants whose mothers did not work at 6 months compared with infants whose mothers worked full-time. As few studies have considered part-time work, we have no strong *a priori* predictions about how this work status will relate to infants' earliest stages of word production.

Methods

The present analysis includes data taken from a larger longitudinal study on early noun-learning from 6-18mo. The full study is described in more detail elsewhere (e.g. Bergelson & Aslin, 2017; Bergelson et al., in press; see Supplementary Materials). Given our goal of examining spontaneous word production across children (which is relatively sparse before age

two), the present analysis focuses solely on productive vocabulary at 1;5 – the endpoint of home recordings in the larger study – alongside demographic information taken at 0;6 and 1;6.

Participants

Forty-four infants completed the larger study, including one pair of dizygotic twins. Infants (21 females) were raised in largely white, middle-class households in New York state. Thirty-three of the mothers had ≥BA degree, forty-two infants were Caucasian, two were mixed-race. All infants were full-term with no reported speech or hearing problems. See Table 1 for age information and Supplementary Materials for detailed sample demographics. The present study was conducted according to guidelines laid down in the Declaration of Helsinki, with written informed consent obtained from a parent or guardian for each child before any assessment or data collection. All procedures involving human subjects in this study were approved by the IRB at both the University of Rochester (where the data were initially collected) and Duke University (where they continue to be analyzed).

PLEASE INSERT TABLE 1 ABOUT HERE

Parental report data

We derived children's total productive vocabulary based on the MacArthur-Bates Communicative Development Inventory (CDI, Words and Gestures, Fenson et al., 1994) completed at 1;5 (the same age as the final home recordings analyzed here). From demographics questionnaires completed at 0;6 and 1;6, we analyze our key variable of maternal work status, alongside child gender, and two measures of maternal education: a 6-point scale from <high school to advanced degree (see details in Supplementary Materials), and mothers' vocabulary

score from NIH toolbox Picture Vocabulary Test (PVT; this test has been validated for ages 3-85, Heaton et al 2014).

We first considered mothers' work status at our two time-points as full time (FT), part-time (PT), or stay-at-home (Home). At 18mo., FT mothers worked 30+hrs. per week (M=38, Range=30-45, SD=4.3; 6m: M=38, Range=30-45, SD=4.4), PT mothers worked 6-24hrs. (M=18.5, SD=6; 6m: M=15.6, Range=6-24, SD=7.9), and Home mothers worked 0hrs. Mothers who freelanced (hours not specified; n=1) or were part-time students (n=1) were classed as PT. One mother who did not work due to a disability stayed at home with her infant; she was classed as Home.

We next assessed whether maternal work status changed from 6-18mo., creating three 'care-schedule' categories. Infants with Home mothers at both 0;6 and 1;6 are classed as Home-only; infants with FT mothers at both time-points are Care-only, and infants whose care shifted to or from PT between time-points are classed Mixed-care: they experienced both maternal and other-caretaker care over 6-18mo. (see Table 2). To avoid making assumptions about infants' care based on their mother's work status alone – e.g. infants with Home mothers may still have attended daycare – we verified caregiving details with mothers at the time of analysis; these follow-ups largely confirmed our initial categorization (see Supplementary Materials).

PLEASE INSERT TABLE 2 ABOUT HERE

Video and audio recordings

Infants were recorded at home for one hour at 1;5 wearing two small Looxcie video cameras attached to a hat or headband. If infants seemed likely to remove the cameras, mothers were also outfitted with a Looxcie camera on a headband. Additionally, a camcorder (Panasonic

HC-V100/Sony HDR-CX240) was set up in the home. Trained research assistants annotated the video recordings from a merged video combining the head-camera and camcorder feeds using Datavyu (Datavyu Team, 2014).

Audio recordings were collected on a separate day from videos. Infants wore a waistcoat containing a LENA recorder (LENA Research Foundation) which captured language input for a full day. Three hours were selected by extracting the top three non-overlapping 'talk' hours in each recording, calculated by averaging the LENA algorithm's output for child and adult vocalization count across the day. Rarely, the recording revealed that the child was asleep or not wearing the recorder more than ten contiguous minutes of the top-three hours; when this happened, additional time was coded from other parts of the recording. These data were annotated in CLAN (MacWhinney, 2000), as described below.

Data annotation and aggregation

Given the larger study's goal of examining noun input and learning, and the preponderance of nouns in infants' early vocabulary (Fenson et al., 1994), trained researchers annotated concrete nouns (e.g. *dog* or *computer* but not *happiness*) directed towards or said by the child, along with several properties of interest to the larger study (Bergelson & Aslin, 2017). Here we consider only *child* noun productions, combining data from audio and video recordings. Annotators were not aware that subsequent analyses would examine maternal work status.

Child productions were checked and transcribed by the first author (a trained phonetician). A second phonetician, blind to the original transcriptions and the purposes of this study, re-transcribed 10% of infant productions (n=213). Transcribers' agreement was 75.6% (Cohen's kappa=.754), excluding voicing contrast (e.g. /kæ/ vs. /gæ/ for *cat*) and gradient

differences in vowel height/backness (e.g. /kæ/ vs. /kɛ/, cf. Macken & Barton, 1980; Rvachew, Mattock, Polka & Ménard, 2006). Disagreements were resolved through discussion. Each word's lemma (i.e. its 'basic' form) was then added (e.g. an infant's production of *doggy* and *dog* would be counted as two tokens of one lemma, DOG). Finally, type and token counts (i.e. number of different nouns and total noun instances per child) were computed.

Results

All analyses were conducted in R save the Bayesian ANOVAs which used JASP; all data and analyses are publicly accessible at https://github.com/cathelaing/Laing-Bergelson-Infancy. We analyzed three measures of infants' vocabulary at 1;5 in relation to care-schedule: (1) total reported productive vocabulary (CDI) (2) noun types and (3) noun tokens produced in audio and video home records. Five infants' (2 females) CDI data was not available at 1;5. Reported vocabulary ranged from 1-372 words; M=56.05, SD=72.2. Audio and video data were recorded for all 44 infants.

Shapiro-Wilkes tests indicated all three measures (CDI, types, and tokens) were not normally distributed (all $p<.05$). They were thus log-transformed (after adding one, since five infants (3 males) produced zero words in the recorded data). The transformed data were normally distributed, and contained no outliers (Shapiro-Wilkes, all ps$>.05$; all observations within 3SD of the mean for each measure). All analyses were performed using log-transformed data.

As expected, Pearson product-moment correlations on the log-transformed data revealed significant correlations between our three child vocabulary measures (Table 3).

PLEASE INSERT TABLE 3 ABOUT HERE

We next examined the effect of care-schedule, alongside two variables known to relate to vocabulary: maternal education (PVT score and education level) and sex (Fenson et al., 1994). First, we conducted ANOVAs with only care-schedule as a predictor, and found that for each vocabulary measure, care-schedule accounted for significant variance (see Table 4).

Next, we conducted model comparisons between null models that included sex and maternal education (separately for sex+PVT score and sex+educational-level, since these are both proxies for maternal education), and models that included these variables along with care-schedule. For all three vocabulary measures, the model with care-schedule in addition to the other variables provided a better fit than the model with sex and PVT score alone. For types and tokens, but not CDI, the model with care-schedule and the other variables provided a better fit than the model with sex and maternal education alone (see Table 4). We also ran Bayesian ANOVAs in JASP (JASP Team, 2018) on the log-transformed data with our two variables alongside care-schedule. Bayesian analysis quantifies support for the alternative hypothesis, H_1, compared with the null hypothesis, H_0 (see Kass & Raftery, 1995), represented by a Bayes factor, BF_{10}. E.g., a BF_{10} of 7.67 (Table 4) suggests a model with caretype (H_1) is 7.67 times more likely than the null model. BF_{10} were calculated using nested model comparisons against the null model, the default analysis in JASP. For each variable we ran three models: DV~caretype; DV~caretype+sex+education+sex*education; and DV~caretype+sex+PVTscore+sex*PVTscore (see Table 4). In each case, JASP compared these models against a model without the effect of caretype; see output on GitHub: https://github.com/cathelaing/Laing-Bergelson-Infancy.

PLEASE INSERT TABLE 4 ABOUT HERE

Care-Schedule Comparison

Given that ANOVA results showed a significant role for care-schedule over our three vocabulary measures, we conducted follow-up independent-samples t-tests to compare number of words produced across groups (see Table 5). Mixed-care infants had significantly higher word production than both Home-only and Care-only infants for both noun-types and noun-tokens (see Fig. 1). Home-only and Care-only infants were statistically indistinguishable on both measures.

PLEASE INSERT TABLE 5 ABOUT HERE

PLEASE INSERT FIGURE 1 ABOUT HERE

Discussion

We find that care-schedule from 6-18mo. was significantly linked with infants' early vocabulary development across vocabulary measures: reported vocabulary (CDI), and naturalistic observation of volubility (noun-tokens) and lexical diversity (noun-types). Overall, infants who experienced both maternal and other-caretaker care had significantly higher productive vocabulary than infants who were cared for full-time either by their mothers *or* by other caretakers.

While our analysis was exploratory, the results converge with and expand the wider literature. Like Booth et al. (2002), we found no difference between infants who experienced only maternal care and those in full-time care by other caretakers in the first 18 months. Extending previous research, we show that mothers' who work part-time in the first 18 months have infants with larger vocabularies; this suggests that part-time work may positively relate to early language development.

Most existing studies in this area use offline measures and epidemiological approaches (Gregg et al., 2005; NICHD, 2000) that often rely on parental report. While we found that care-

schedule accounted for variance in all three vocabulary measures both directly and when considered in models in addition to sex and maternal vocabulary score (PVT), when maternal education level was included instead of PVT, care-schedule no longer improved model fit for CDI data but continued to improve fit for the home-recording vocabulary measures. This may be due to sample size (we were lacking CDI but not home recordings for 5 infants), or to factors inherent across care-schedule groups. On one hand, parental report accuracy may differ depending on how much time the caregivers spend with infants. On the other hand, both CDI measures correlated highly and significantly with noun-types and -tokens in the home recordings. Given that the CDI scores of mixed-care infants were >2.5x those in the other two groups, we find it most likely that with a larger sample, this measure too would render results as robust as those from home recordings.

Thus, the inclusion of naturalistic data in the present study presents a valuable new perspective, providing a direct measure of infants' vocabulary that can be examined in relation to further social and demographic variables. The current work, though modest in scope, is a first step in tying naturalistic observational data of language production to maternal work status in infancy.

Previous large-n research has been critical for unpacking the interlocking effects of SES, race, age, and child-care quality in addition to work status. We approached this question from the opposite perspective: we examined a small population with limited demographic variability. While this limits our ability to generalize, it did permit us to conduct in-depth family-level analysis of maternal work status with other variables essentially held constant. Encouragingly, these results are largely in line with previous large-n work. Consistent with Han et al. (2001), our results suggest that maternal work at 6mo. is tied to lower vocabulary at 17mo. Moreover, our

findings expand on their results to show that this applies to FT, but not PT work. Consistent with Waldfogel and colleagues (2002), we too see that infants who were in full-time maternal care until 18mo. had smaller vocabularies than those with a combination of maternal and other-caretaker care.

This study is further limited in two ways. First, though relevant in previous work (NICHD, 2000), data concerning care type and quality were not collected; we cannot speculate on how care-type affected early linguistic experience. Second, given that 17 months is essentially the onset of substantive word production, we cannot anticipate whether the effects we observe regarding maternal work-status will have longer-term effects; follow-up at later ages, along with more in-depth analysis across lexical classes, would allow us to test the extent of these effects (cf. Walker et al., 1994).

Our results suggest that maternal care may be linked to even very early linguistic abilities. More effective maternal engagement alongside the wider speaker variability conferred by other caretakers may be important in establishing early linguistic skills (cf. Rost & McMurray, 2010). Additionally, PT mothers may spend more of their non-work time engaging in linguistically-supportive interactions with their infants (cf. Booth et al., 2002) than both FT and Home mothers. In turn, mixed-care infants may receive more high-quality interactions with their mothers than both Home-only and Care-only infants, as in many cases PT mothers are infants' sole caretakers for substantial parts of the week. While care 'quality' is notoriously hard to operationalize, solo care generally allows more 1:1 time, and thereby more supportive early interactions (Ramírez- Esparza, García- Sierra & Kuhl, 2014). These suggestions are not mutually exclusive; our results are an initial suggestion that part-time care may provide a 'sweet spot' between more one-to-one interactions at home, alongside a diverse input in early

experiences. Infants experiencing both maternal and other-caretaker care may have, linguistically-speaking, the best of both worlds, seen in observably more voluble and variable word production by 17 months.

Works Cited

Bergelson E., Amatuni, A., Dailey, S., Koorathota, S., Tor, S. (accepted pending minor revisions). Day by Day, Hour by Hour: Naturalistic Language Input to Infants. *Developmental Science*.

Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-mo-olds *Proceedings of the National Academy of Sciences*, *114*(49), 12916-12921.

Booth, C., Clarke-Stewart, K. A., Vandell Lowe, D., McCartney, K., & Tresch Owen, M. (2002). Child-care usage and mother- infant "quality time." *Journal of Marriage and Family*, *64*, 16–26. https://doi.org/10.1111/j.1741-3737.2002.00016.x

Bureau of Labor Statistics (2017). *Employment Characteristics of Families - 2016*. Retrieved from https://www.bls.gov/news.release/pdf/famee.pdf, 11/24/17.

Datavyu Team (2014). *Datavyu: A Video Coding Tool*. Databrary Project, New York University. http://datavyu.org.

Fenson, L., Dale, P.S., Reznick, J.S., Bates, E., Thal, D.J., Pethick, S.J., Tomasello, M., Mervis, C.B. and Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, i-185. DOI: 10.2307/1166093

Gregg, P., Washbrook, E., Propper, C., & Burgess, S. (2005). The effects of a mother's return to work decision on child development in the UK. *The Economic Journal*, *115*(501), 48–80. https://doi.org/10.1111/j.0013-0133.2005.00972.x

Han, A. W., Waldfogel, J., & Brooks-gunn, J. (2001). The effects of early maternal employment on later cognitive and behavioral outcomes. *Journal of Marriage and Family*, *63*(2), 336–354.

Heaton, R., Akshoomoff, N., Tulsky, D., Mungas, D., Weintraub, S., Dikmen, S., Beaumont, J, Casaletto, K., Conway, K., Slotkin, J. & Gershon, R. (2014). Reliability and validity of composite scores from the NIH Toolbox Cognition Battery in adults, *Journal of the International Neuropsychological Society*, *20*(6), 588-598. 10.1017/S1355617714000241

Higgins, C., Duxbury, L., & Johnson, K. L. (2000). Part-time work for women: Does it really help balance work and family? *Human Resource Management*, *39*(1), 17–32.

https://doi.org/10.1002/(SICI)1099-050X(200021)39:1<17::AID-HRM3>3.0.CO;2-Y

JASP Team (2018). JASP(Version 0.8.6)[Computer software]. Available online at https://jasp-stats.org.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American statistical association*, *90*(430), 773-795.

LENA Research Foundation (2016). https://www.lena.org/.

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

NICHD. (2000). The relation of child care to cognitive and language development. *Child Development*, *71*(4), 960–980.

Gershon, R., Wagster, M., Hendrie, H., Fox, N., Cook, K. & Nowinski, C. (2013). NIH toolbox for assessment of neurological and behavioral function. *Neurology*, *80*(11 Supplement 3), S2-S6.

Office for National Statistics (2017). *Families and the labour market, England, LFS time series datasets* [data file]. Retrieved from https://goo.gl/Aj3SBA, 11/24/17.

Ramírez- Esparza, N., García- Sierra, A., & Kuhl, P. K. (2014). Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. *Developmental Science*, *17*(6), 880-891. https://doi.org/10.1111/desc.12172

Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, *15*(6), 608-635. DOI: 10.1111/j.1532-7078.2010.00033.x

Waldfogel, J., Han, W.-J., & Brooks-Gunn, J. (2002). The effects of early maternal employment on child cognitive development. *Demography*, *39*(2), 369–392.

Table 1. *Infants' age (mean and SD) at data collection for each data type.*

Data type	**Mean(SD), in days**
CDI	525.44(7.7)
Audio Home-Recording	521.71(3.26)
Video Home-Recording	519.16(3.03)
Demographics 6m	184.43(3.64)
Demographics 18m	549.23(3.74)

Table 2. *Care-schedule groups, as defined by mothers' work status at 6 and 18 months. Number of females shown in parentheses.*

Care-schedule	**Work status 0;6**	**Work status 1;6**	**n Infants**	**Group total**
Home-only	Home	Home	15(7)	15(7)
Mixed-care	Home	PT	4(1)	12(8)
Mixed-care	Home	FT	1(1)	
Mixed-care	PT	PT	5(4)	
Mixed-care	PT	FT	2(2)	
Care-only	FT	FT	17(6)	17(6)

Table 3. *Reported productive vocabulary from CDI reports (full inventory + nouns only) and types and tokens produced in home recordings.*

Variables	N infants	1	2	3	4
1. CDI (full)	39				
2. CDI (nouns)		.97***			
3. Noun-types	44	.71***	.75***		
4. Noun-tokens	44	.64***	.68***	.92***	

*p < .05; **p < .01; ***p < .001

Table 4. *Effect of care-schedule, maternal education, and child gender on productive vocabulary (CDI, noun-types, noun-tokens). Columns 2-5 show descriptive statistics (M(SD)) for our full sample (Col. 2) and three care-schedule groups (Col. 3-5). Column 6 shows ANOVA results of model comparisons between a model with our main effect + predictors (care-schedule + sex+PVT or education) and a null model with predictors only (PVT or education + sex). Column 7 shows BF_{10} for each model. All tests were conducted on log-transformed data; M(SD) is presented untransformed for interpretive clarity.*

		M	**$M_{Home-Only}$**	**$M_{Mixed-Care}$**	**$M_{Care-Only}$**	**ANOVA**	**BF_{10}**
CDI	caretype	56.05(72.2)	21.42(14.2)	109.73(106.1)	44.4(46.3)	F(2,35)=4.04, *p*=.03*	2.41
	caretype+ sex+PVT					F(2,35)=4.25, *p*=.02*	3.03
	caretype+ sex+ed					F(2,31)=2.65, *p*=.21	1.5
Noun-Types	caretype	9.3(14.3)	6.5(7.5)	17.9(19.5)	5.8(12.7)	F(2,41)=6.37, *p*<.01**	10.67
	caretype+ sex+PVT					F(2,41)=4.14, *p*=.02*	2.74
	caretype+ sex+ed					F(2,37)=5.28, *p*<.01**	8.1
Noun-Tokens	caretype	41.8(59.9)	29.3(36.3)	89.5(81.5)	19.2(39.2)	F(2,41)=8.48, *p*<.001***	38.64
	caretype+ sex+PVT					F(2,41)=6.25, *p*<.01**	10.9
	caretype+ sex+ed					F(2,37)=9.08, *p*<.001***	40.93

*p < .05; **p < .01; ***p < .001

Table 5. *T-test results for group comparisons, including effect sizes determined by Cohen's d and BF_{10}. Asterisks represent significance at α=.05. (N.B.: statistics conducted over log-transformed data)*

	Home vs. Mixed	**Mixed vs. Care**	**Home vs. Care**
CDI	*t*(16.7)=-2.64, *p*=.02*, *d*=*-1.24*, BF_{10}=4.14	*t*(17.27)=-1.49, *p*=.15, *d*=-.62, BF_{10}=.88	*t*(24.72)=-1.56, *p*=.13, *d*=*.59*, BF_{10}=.85
Types	*t*(24.29)=-2.56, *p*=.02*, *d*=*-.98,* BF_{10}=3.36	*t*(25.02)=-3.59, *p*=.001***, *d*=-1.33, BF_{10}=22.12	*t*(29.66)=1.01, *p*=.32, *d*=-.36, BF_{10}=.5
Tokens	*t*(24.94)=-2.84, *p*=.009**, *d*=-1.06, BF_{10}=4.78	*t*(26.86)=-4.46, *p*<.001***, *d*=-1.6, BF_{10}=101.5	*t*(29.56)=1.41, *p*=.17, *d*=-.5, BF_{10}=.71

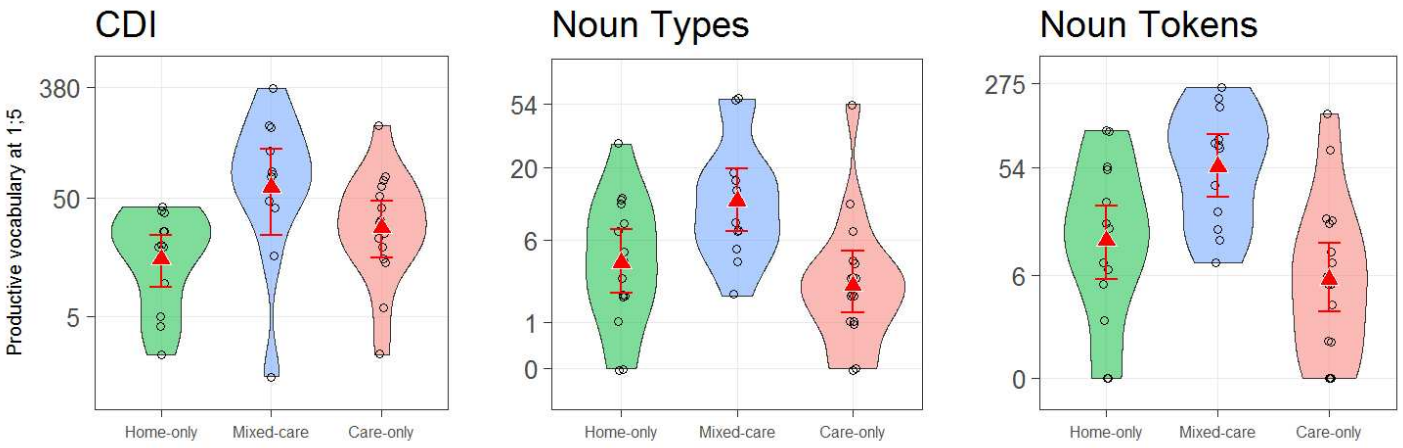

Figure 1. *Infant productive vocabulary at 1;5 plotted as a function of care-schedule for CDI, noun-types and noun-tokens. The y-axis uses a natural log scale. Circles represent individual infants' data, jittered slightly to show overlapping data-points. Triangles represent group means and 95% bootstrapped non-parametric confidence intervals. Non-overlapping confidence intervals correspond to significant condition differences (see Table 5).*