



Published in final edited form as:

*J Child Lang.* 2013 June ; 40(3): 672–686. doi:10.1017/S0305000912000141.

## What counts as effective input for word learning?\*

**LAURA A. SHNEIDMAN, MICHELLE E. ARROYO, SUSAN C. LEVINE, and SUSAN GOLDIN-MEADOW**

University of Chicago

### Abstract

The talk children hear from their primary caregivers predicts the size of their vocabularies. But children who spend time with multiple individuals also hear talk that others direct to them, as well as talk not directed to them at all. We investigated the effect of linguistic input on vocabulary acquisition in children who routinely spent time with one vs. multiple individuals. For all children, the number of words primary caregivers directed to them at age 2;6 predicted vocabulary size at age 3;6. For children who spent time with multiple individuals, child-directed words from ALL household members also predicted later vocabulary and accounted for more variance in vocabulary than words from primary caregivers alone. Interestingly, overheard words added no predictive value to the model. These findings suggest that speech directed to children is important for early word learning, even in households where a sizable proportion of input comes from overheard speech.

---

Many studies have demonstrated the important link between child-directed speech from a primary caregiver and the child's lexical, semantic and syntactic development (e.g. Barnes, Gutfreund, Satterly & Wells, 1983; Gleitman, Newport & Gleitman, 1984; Hart & Risley, 1995; Hoff, 2003; Huttenlocher, Haight, Bryk, Seltzer & Lyons, 1991; Huttenlocher, Vasilyeva, Cymerman & Levine, 2002; Newport, Gleitman & Gleitman, 1977; Rowe, 2008). However, to our knowledge, no study has considered the impact that input from other naturalistic sources has on children's later language development. Many children live in households where multiple adults and other children are present for large portions of the day. In such households, young children are likely to hear speech directed to them from older siblings and other household members. Moreover, these children are likely to overhear speech not directly addressed to them at all. Little is known about the frequency of speech from these input sources in multi-party households, nor about the impact that speech input from different sources has on children's later language outcomes.

The goal of this article is to begin to address these issues by considering lexical input and vocabulary development in children growing up in households where multiple speakers are regularly present. We chose to focus on vocabulary because learning new words may be particularly sensitive to speech directly addressed to the child. Mutual engagement between caregivers and children (more likely when children are directly addressed) may plausibly enhance children's attention to cues that make a speaker's referential intent easier to discern (e.g. Tomasello, 1995), and could be key for learning new words. Indeed, prior research has found that children who participate in more episodes of joint engagement with caregivers

---

\*This research was supported by P01HD40605 to Goldin-Meadow and Levine.

© Cambridge University Press 2012

Address for correspondence: Laura Shneidman, Center for Early Childhood Research, University of Chicago, 5848 S. University Ave., Chicago, IL, 60637. lauras@uchicago.edu.

have larger vocabularies than children who participate in fewer of these episodes (Carpenter, Nagell & Tomasello, 1998; Tomasello & Farrar, 1986; Tomasello & Todd, 1983).

Experimental evidence suggests that children have the ability to learn new words from overhearing speech (e.g. Akhtar, 2005; Akhtar, Jipson & Callanan, 2001; Floor & Akhtar, 2006; Shneidman, Buresh, Shimpi, Knight-Schwartz & Woodward, 2009). At age 1;6, children can learn a novel object label from an overheard utterance when the learning task is sufficiently simplified (i.e. when children are tested immediately following a labeling demonstration; Floor & Akhtar, 2006). By age 2;6, children show a robust ability to learn from overheard interactions in an experimental paradigm (Akhtar, 2005). However, in these studies, children's attention was relatively restricted, which may have facilitated learning. Children may be less likely to learn words from natural, overheard interactions than from experimentally engineered interactions simply because, in the natural situations, their focus of attention is less constrained. Indeed, in an experimental paradigm that was designed to be more naturalistic (children were given a distracting toy to play with during the experimental session), two-year-old children were only able to learn a novel label from overheard speech when the speaker used a high-pitched register that mimicked child-directed speech (Shimpi & Akhtar, 2011). Naturally occurring overheard speech may therefore be less effective in promoting word learning than experimentally engineered overheard speech.

This article is organized as follows. In the first part, we characterize lexical input for children growing up in households where multiple speakers are regularly present. We ask how early input for these children compares to input for children growing up in households where most of the child's time is spent with a single caregiver. Prior research shows that primary caregivers talk significantly less, and are less responsive to children's communicative verbalizations, when a sibling is present than when a child is alone with the caregiver (Huttenlocher, Vasilyeva, Waterfall, Vevea & Hedges, 2007; Jones & Adamson, 1987; Oshima-Takane & Robbins, 2003; Wellen, 1985). Children growing up in households where they spend much of their time with multiple people might therefore receive less directed input from their primary caregiver than children growing up in households where they spend most of their time alone with a single caregiver. Alternatively, because children in multi-party households are surrounded by individuals who also have the potential to interact with them, these children could hear more total directed input than children who spend most of their time with a single caregiver. In addition, children in multi-party households likely have a greater potential to be exposed to overheard input precisely because there are more multi-party conversations in these households. In the first part of our article, we quantify children's input from these sources.

In the second part, we consider the impact that directed and overheard input has on lexical acquisition for children growing up in these different types of households. Specifically, we describe the relation between naturally occurring input at age 2;6 and children's receptive vocabulary at age 3;6. We ask which measures of input – directed input from the primary caregiver, directed input from all household members, or all input (directed and overheard input) – most strongly relate to children's later vocabulary. We chose to consider input at age 2;6 because prior work has shown that children demonstrate a robust ability to learn from overheard speech at this age in an experimental paradigm (e.g. Akhtar, 2005). We chose to assess children's vocabulary at age 3;6 because previous work has shown that early parental input has an impact on child vocabulary comprehension one year later (e.g. Huttenlocher *et al.*, 1991; Rowe, 2008).

## METHOD

### Participants

Thirty monolingual English-speaking families containing a target child were selected from a sample of sixty-five families participating in a longitudinal study of language development in the greater Chicago area. The children were videotaped in their homes for 90 minutes every four months, from age 1;2 to 3;6. To select the children for this study, we considered all sixty-five children from the larger sample and made note of children's social partners across home visits. We chose children who represented two ends of a spectrum: children who tended to spend their days with more than one individual (the multi-speaker group), and children who tended to spend their days with a single individual (the single-speaker group).

Fifteen households (5 with boys and 10 with girls) were selected for the MULTI-SPEAKER group, and fifteen (7 with boys and 8 with girls) for the SINGLE-SPEAKER group. On average, target children in the multi-speaker families had more than one individual around them in 83% ( $SD=14\%$ ) of the eight home visits, compared to 30% ( $SD=18\%$ ) in the single-speaker families ( $t(28)=9.11, p<0.001$ ). Across the eight visits, the average number of individuals around the children in the multi-speaker group was 2.4 people ( $SD=0.49$ ) compared to 1.4 ( $SD=0.29$ ) for children in the single-speaker group ( $t(28)=6.88, p<0.001$ ). All of the children in the multi-speaker group had more than one individual around them during the target visit at age 2;6 (average number of individuals present=2.8 people; range 2 to 5 people); none of the children in the single-speaker group had multiple speakers around them during the target visit.

Ten children from multi-speaker families had siblings (3 children had 1 older sibling, 4 children had 2 older siblings, 1 child had 3 older siblings, 1 child had 4 older siblings, and 1 child had 5 older siblings). Two children had their mother and father present with them throughout the day, two children lived with multiple extended family members, and one child lived with his mother and her roommates. Some of the target children from the single-speaker families also lived with siblings (5 children had 1 older sibling) and/or other family members; however, these family members had work or school obligations and were typically not with the child during the day, which resulted in the primary caregiver being alone with the target child for extended periods of time.

The primary caregivers in the multi-speaker families had diverse educational backgrounds: four had graduated from high school, three had some college experience, three had graduated from college, and five had an advanced degree. There was also diversity in the educational level of the primary caregiver in the single-speaker households: one had graduated from high school, three had some college experience, seven had graduated from college, and four had an advanced degree. The multi-speaker and the single-speaker families were roughly equal in distributions by ethnicity and income level (see Table 1).

### Procedure

Children from the target families were videotaped at age 2;6 (range: 2;6 to 2;7) in their homes for 90 minutes by an experimenter who kept a camera on the target child. We provided no instructions to the families as to who should be present at these visits, other than to encourage the families to interact as they normally would if the experimenter were not present. When the child was aged 3;6 (range: 3;5 to 3;7), the experimenter administered the Peabody Picture Vocabulary Test (PPVT) to the child (PPVT III; Dunn & Dunn, 1997).

The speech utterances that the children produced were transcribed from the videotapes, along with the speech utterances produced by all others within earshot of the child (defined as any intelligible utterance audible on the videotape). These utterances were categorized as

either being directed to the child or overheard by the child. Utterances that were directed to a group of individuals that included the child were coded as directed utterances. All utterances were classified as coming from the primary caregiver, coming from a child speaker (under thirteen years), or coming from an adult speaker other than the primary caregiver. Utterances in phone conversations, talking to oneself, and talking to pets were not transcribed: (1) because they did not generally direct the child's attention to objects or events in the immediate surroundings and thus were less likely to be useful for word learning; and (2) because in these situations no human interlocutor was present and previous work has suggested that attention to both conversational partners is important for children to learn novel words from overheard speech (Shneidman *et al.*, 2009). These occurrences were infrequent.

### Input language measures

Using the videotape from the 2;6 visit, we measured the number of word tokens and word types in the target child's input calculated in three different ways: (1) speech directed to the child from the primary caregiver; (2) speech directed to the child from any other household member; and (3) overheard speech. We defined each family's primary caregiver by parental report except in cases where parents defined themselves as dual primary caregivers (2 families). In these cases, we chose the primary caregiver based on who directed the most speech utterances to the child at the 2;6 visit. Fourteen mothers and one father were primary caregivers in the single-speaker group. In the multi-speaker group, all primary caregivers were mothers.

## RESULTS AND DISCUSSION

### Child language at 2;6 and 3;6

Using the videotape taken at 2;6, we calculated the number of word tokens and word types that the target child produced. There were no reliable differences between the single-speaker and multi-speaker households in the number of word tokens children produced at 2;6 (single-speaker:  $M=1479$ ,  $SD=1175$ ; multi-speaker:  $M=1534$ ,  $SD=567$ ,  $t(28)=0.16$ , n.s.), nor in the number of word types children produced at 2;6 (single-speaker:  $M=208$ ,  $SD=113$ ; multi-speaker:  $M=222$ ,  $SD=67$ ,  $t(28)=0.40$ , n.s.). There were also no differences in the PPVT scores of children from single-speaker and multi-speaker households at 3;6 (single-speaker:  $M=107$ ,  $SD=17$ ; multi-speaker:  $M=103$ ,  $SD=19$ ,  $t(28)=0.70$ , n.s.).

Thus, on average, the groups did not differ in either the size of their productive vocabularies at 2;6 or their receptive vocabularies at 3;6. We turn next to a comparison of the linguistic environments to which children growing up in the multi-speaker vs. single-speaker groups were exposed at 2;6.

### Linguistic environment in single-speaker and multi-speaker households

**The child's sources of input**—All of the utterances that the children in the single-speaker group heard during the target session were directed to them from their primary caregiver. In contrast, only 69% ( $SD=15\%$ ) of the utterances that children in the multi-speaker group heard came from child-directed speech. Of these utterances, 76% ( $SD=22\%$ ) came from the primary caregiver, 17% ( $SD=24\%$ ) came from other adults, and 7% ( $SD=11\%$ ) came from children under thirteen years.

The remaining 31% ( $SD=15\%$ ) of total utterances that children in the multi-speaker households heard came from overheard speech. Of these utterances, 51% ( $SD=15\%$ ) came from the primary caregiver, 33% ( $SD=27\%$ ) came from other adults, and 16% ( $SD=20\%$ ) came from other children. Controlling for amount of talk, the word types that children heard

in directed speech input were as likely to appear on the PPVT outcome measure as the word types heard in overheard speech input ( $t(14)=0.02$ ,  $p=0.80$ ).

**Speech directed to the child**—There were no differences between the single-speaker and multi-speaker groups in the total number of words (tokens) or the number of different words (types) directly addressed to the child by the primary caregiver; see the black bars in Figure 1. Children in the single-speaker group heard, on average, 3606 ( $SD=1712$ ) tokens and 444 ( $SD=136$ ) types in directed input from their primary caregiver during the 90-minute taping sessions, and children in the multi-speaker group heard on average 3094 ( $SD=1463$ ) tokens and 399 ( $SD=89$ ) types in directed speech from their primary caregiver during the 90-minute taping sessions; these means did not differ significantly (Tokens:  $t(28)=0.65$ , n.s.; Types:  $t(28)=1.07$ , n.s.). In addition, the number of tokens ( $M=3606$ ) and types ( $M=444$ ) that children in single-speaker households heard from their primary caregivers did not differ significantly from the number of tokens ( $M=4116$ ,  $SD=1465$ ) or types ( $M=460$ ,  $SD=90$ ) that all household members, including the primary caregiver, directed to the child in multi-speaker households (Tokens:  $t(28)=0.87$ , n.s.; Types:  $t(28)=0.40$ , n.s.); compare the black bars on the left side of each graph to the black and gray bars on the right side in Figure 1.

**Overheard speech**—When overheard speech is added to the mix, the total number of word tokens ( $M=6286$ ,  $SD=1837$ ) and word types ( $M=626$ ,  $SD=121$ ) children in multi-speaker households received was significantly higher than the number that children in single-speaker households received (Tokens:  $t(28)=4.13$ ,  $p<0.001$ ; Types:  $t(28)=3.90$ ,  $p<0.001$ ); compare the black bars on the left side of each graph to the black, gray and white bars on the right side in Figure 1. Thus, the children in our study who spent many hours with a variety of speakers heard more input overall (equal amounts of speech directed to them, and more overheard speech) than children who spent most of their waking hours with a single caregiver.

### Input from the primary caregiver and vocabulary development

Previous work exploring the relationship between language input and vocabulary development has considered only speech that the primary caregiver directly addresses to the child. We ask next whether the relationship between input and vocabulary development is altered when considering children raised in households with multiple sources of input.

To examine the impact of input from the primary caregiver on vocabulary, we entered input tokens from the primary caregiver at 2;6, and group (single-speaker vs. multi-speaker) into a regression model predicting children's 3;6 PPVT score. We found a significant effect of directed input from the primary caregiver at 2;6 on children's 3;6 vocabulary score ( $\beta=0.54$ ,  $p=0.023$ ). This effect indicates that every standard deviation change in word tokens from the primary caregiver at 2;6 was positively associated with a 0.54 standard deviation difference in children's PPVT scores at 3;6. Importantly, we found no effect of group (single-speaker vs. multi-speaker) on 3;6 PPVT score ( $\beta=0.02$ , n.s.), and we found no interaction between group (single-speaker vs. multi-speaker) and input from the primary caregiver ( $\beta=0.02$ , n.s.). Directed speech from the primary caregiver was thus significantly related to later vocabulary for children in both single- and multi-speaker households. This model accounted for 27.5% of total variance in children's 3;6 vocabulary score.

We found the same pattern of results when we used word TYPES as our measure of input. Types addressed to the child by the primary caregiver at 2;6 significantly related to 3;6 PPVT score ( $\beta=0.57$ ,  $p=0.007$ ). There were no group or interaction effects. Since word tokens and word types are highly correlated in our sample ( $r=0.94$ ,  $p<0.001$ ), it is impossible to determine which measure better accounts for the relation between input and vocabulary.

## Input from all sources and vocabulary development

Given that children growing up in the multi-speaker households received directed input from individuals other than the primary caregiver, as well as input from overheard speech, our next goal was to consider how these various sources of input related to later vocabulary in these children. Recall that for the multi-speaker group, we calculated three measures of word tokens: (1) the number of word tokens directed to the child from the primary caregiver; (2) the number of word tokens directed to the child from other household members; and (3) the number of overheard word tokens, defined as any token not directed to the child that was audible to the child. Summed, these measures represent children's total linguistic input. We used regression models to examine the relation between word tokens in input at 2;6 and child PPVT score at 3;6. Results are displayed in Table 2.

Model 1 shows the results of a simple regression with the number of tokens directed to the child by the primary caregiver as the sole predictor of 3;6 PPVT scores. This measure did not reliably predict PPVT ( $F=2.82$ ,  $p=0.18$ ).

In Model 2, we included the number of tokens directed to the child from other household members as an additional input predictor. In contrast to the previous model, this model did significantly predict children's vocabulary at 3;6 ( $F=3.90$ ,  $p=0.049$ ). In this model, directed tokens from the primary caregiver was a significant predictor of vocabulary score ( $\beta=0.60$ ,  $p=0.035$ ) and directed tokens from others approached significance ( $b=0.50$ ,  $p=0.061$ ). Model 2 accounted for 39% of the total variance in 3;6 PPVT score, compared to 18% for Model 1. The difference in the  $r$ -squared values between Model 1 and Model 2 approached significance ( $F=4.28$ ,  $p=0.06$ ).

In Model 3, we added tokens overheard by the child to the model. This measure did not predict vocabulary score ( $F=2.38$ ,  $p=0.13$ ), and Model 3 accounted for no more variance in PPVT scores than Model 2 (i.e. 39% of the total variance).

We found the same pattern of results when we used word TYPES as our measure of input (see Table 3). Note that the predictors in Table 2 are additive, i.e. adding 'Direct tokens from primary caregiver' to 'Direct tokens from others' results in the total number of direct tokens the child hears. In contrast, the predictors in Table 3, which are based on types rather than tokens, are not additive simply because the same word type often occurred in more than one category (e.g. the word *dog* might be directed to the child and also overheard by the child). Summing these categories would therefore overestimate the number of different words the child hears. As a result, the predictors in Table 3 are cumulative (i.e. 'Direct types from primary caregiver alone' is a subset of 'All directed types') and thus differ from the predictors in Table 2. Types directed to the child by the primary caregiver did not significantly relate to children's PPVT scores ( $F=3.62$ ,  $p=0.08$ ), whereas types directed from all household members did ( $F=8.02$ ,  $p=0.01$ ). All types (directed and overheard) failed to predict children's subsequent PPVT score ( $F=2.28$ ,  $p=0.16$ ).

To summarize, Model 2 best described the relation between input and later vocabulary acquisition in the multi-speaker households for both word tokens and word types – Model 2 (which used all speech directly addressed to the child) was significant, whereas neither Model 1 (which used only speech directed to the child by the primary caretaker) nor Model 3 (which used all speech the child heard, even overheard speech), was. Our findings underscore the fact that not all input has equal potential for supporting word learning. We found that speech directed to the child uniquely predicts children's PPVT score. If we consider only speech overheard by the child in a model that predicts PPVT score, we find no relation between input and vocabulary ( $\beta=-0.06$ ,  $p=0.80$ , for overheard tokens;  $\beta=-0.06$ ,

$p=0.82$  for overheard types). There appears to be something special about speech directed to the child for informing lexical development.

## DISCUSSION

Most studies exploring the relation between early language input and child vocabulary acquisition have focused exclusively on the impact that directed input from the child's primary caregiver has on child vocabulary. But many children live in households where they routinely interact with a number of other people in addition to their primary caregiver. We found that children growing up in multi-speaker households heard more total tokens and types (directed and overheard) than children growing up in households where most of their time was spent with a single caregiver. However, we found no significant differences in the amount of speech that was DIRECTED to the child across household types. These results suggest that children growing up in the multi-speaker households do not experience either a deficit or surplus in child-directed speech. Only when we include overheard speech do we find that children in multi-speaker families hear more words (and more different types of words) than children in single-speaker families.

Importantly, our data demonstrate that, for children who spend time with multiple speakers, directed input from the primary caregiver alone may not be the best measure of these children's effective input. We accounted for more variance in children's PPVT score when we considered directed speech from ALL household members, and not just speech from the primary caregiver. This finding has important methodological implications for research in language development. In households with multiple speakers, it is important to consider child-directed speech from all household members as potential sources of input.

Our study replicates previous work that has found a relation between children's input in DIRECTED SPEECH and later vocabulary (e.g. Barnes *et al.*, 1983; Gleitman *et al.*, 1984; Hart & Risley, 1995; Hoff, 2003; Huttenlocher *et al.*, 1991; Huttenlocher *et al.*, 2002; Newport *et al.*, 1977; Rowe, 2008). However, we have extended these previous studies by demonstrating that the sources of the directed-speech are irrelevant – specifically, it did not matter whether the speech to the child came from primary caregivers or from other adults and older siblings. Moreover, neither overheard input nor the sum of all input (directed and overheard) significantly predicted subsequent vocabulary. These findings suggest that children aged 2;6 (or at least children aged 2;6 who are also directly addressed in conversation) do not readily make use of overheard input when learning words in naturalistic situations. We speculate that overheard speech may have little impact on word learning precisely because this type of input is not likely to occur in situations of mutual engagement between speakers and children, engagement that has been found to facilitate language acquisition (e.g. Carpenter *et al.*, 1998).

Previously, researchers have reported that children aged 2;6 are able to learn words from overheard speech in the experimental laboratory (e.g. Akhtar *et al.*, 2001). We did not replicate these findings in a naturalistic setting. Naturally occurring overheard speech may differ from experimentally engineered overheard speech (and from child-directed speech) in ways that could make it a difficult source of input for word learning. For example, speakers might make reference less frequently to visually present objects, or they might use more complicated syntax, when not talking directly to children than when addressing them directly. In addition, overheard speech could be quieter than directed speech, or could refer less frequently than directed speech to situations that are interesting or salient to young children. In our study, we included ALL speech directed to an interlocutor other than the child in our measure of overheard speech, no matter how accessible that input was to the child. Future research is needed to disambiguate the exact circumstances under which learning

from overheard speech does occur, and whether learning from overheard input changes over the course of development.

In the current study, we considered only the relation between the number of words children hear and their later vocabulary. Directed speech might be particularly useful for learning vocabulary because, as noted earlier, word learning is likely to occur in situations of mutual engagement where the speaker's referential intentions are made clear to the child (e.g. Tomasello, 1995). However, directed speech in the context of mutual engagement may be less critical for other aspects of children's language development, for example, syntax learning. One-year-old children can learn to distinguish grammatical from ungrammatical constructions after exposure to an artificial speech stream that has only predictive dependencies as cues to word order (Saffran, Hauser, Seibel, Kapfhamer, Tsao & Cushman, 2008). If the ability to abstract patterns is all that is needed to learn syntactic rules, then overheard input could be just as useful for fostering syntactic development as directed speech is.

Further, while the current research only considered quantitative properties of input in single-speaker and multi-speaker families, there are also likely to be qualitative differences that occur in input as a function of the number of interlocutors present in an interaction. There is evidence, for example, that the content of mother speech to children changes when other siblings are present; mothers use more metalinguistic language when alone with children, and more social language in triadic interaction (Oshima-Takane & Robbins, 2003). Future research should explore what effect these types of qualitative factors have on children's subsequent acquisition. Indeed, some theorists have hypothesized that the qualitative properties of overheard speech could make it MORE useful than directed speech for certain aspects of language learning precisely because overheard speech is not tailored to the level of the child. For example Blum-Kulka and Snow (2002) suggest that overheard speech could be key to learning narrative structure and humor, aspects of language to which children might otherwise not be exposed.

It is also important to note that children living in cultures where their primary linguistic input is overhead speech seem to experience no acquisition delays (see review in Lieven, 1994). However, the contexts in which overheard speech is used in these cultures may differ from the contexts in which it is used in the United States, perhaps making overheard speech more useful for word learning in other cultures. For example, de Leon (1998) notes that caregivers in the Tzotzil Mayan culture have almost constant physical contact with young children and, as a result, the children are consistently positioned in a manner that places them in coordinated attentional focus with their caregivers. Children in this physical arrangement may be more likely to attend to objects that caregivers label, even when those labels are not addressed to them.

In addition, on the occasions when children growing up in cultures where they are rarely directly addressed do NOT share the speaker's vantage point, they may be able to make use of attentional strategies that could help them learn from overhead speech – strategies that may be less prevalent in children growing up in a culture where direct engagement is common. The ability to actively attend to one's environment is an important skill for learning vocabulary in non-dyadic interactions, particularly if speakers do not refer to objects that are in the child's attentional focus. There is, indeed, evidence that children growing up in cultures where they receive the majority of their input from overheard speech are more skilled at this kind of third party observation than children who regularly experience directed interaction (Chavajay & Rogoff, 1999).



Finally, structural properties found in overheard speech in these communities might also be able to facilitate learning. For example, Brown (1998) has argued that the dialogic repetition overheard by children growing up in a Tzeltal Mayan village (e.g. repetition of the same verbs across conversational turns) facilitates early acquisition of verb roots (which are the first words learned by Tzeltal children) by highlighting those roots. Future work will need to consider the relation between directed vs. overheard input and language learning in communities where overheard speech represents a large proportion of early language input.

In conclusion, our study highlights the need to carefully consider what constitutes linguistic input for a child. For children in our study who routinely heard talk from many speakers, the size of their receptive vocabulary, as measured by the PPVT, was related to talk directed to them by ALL speakers, not just their primary caregiver. Interestingly, however, overhead speech did NOT add predictive value to this relationship. Thus, not all speech that children hear is equally relevant for word learning, at least in a culture where children are routinely addressed directly. Our data do not speak to children growing up in cultures where they are rarely addressed directly, and where they may, by necessity, make more effective use of overheard speech. Our findings suggest that language acquisition researchers must consider the sources of linguistic input children experience, characteristics of this input, the child's cultural history of exposure to directed and overheard speech, as well as the processing skills children bring to bear on the input. Together, these factors are likely to determine what counts as effective input for word learning.

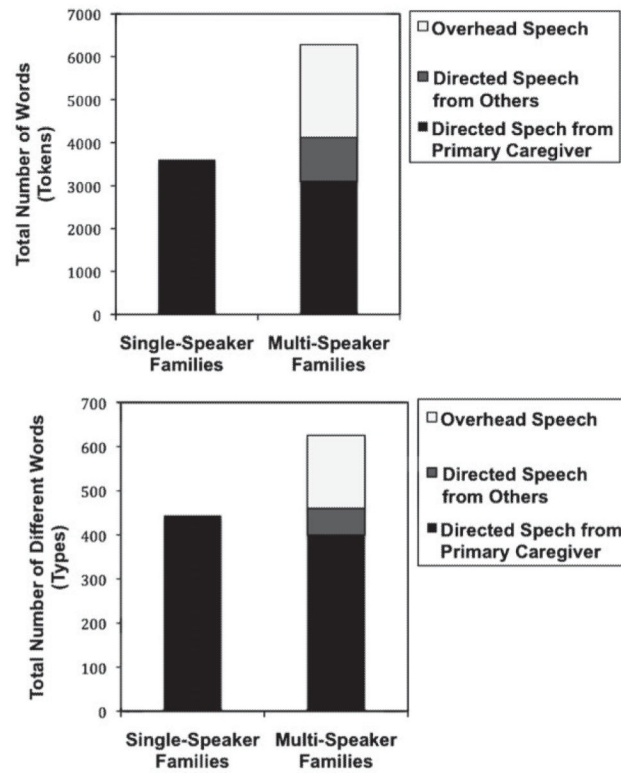
## Acknowledgments

We thank K. Schonwald, B. Trofater and J. Voigt for their administrative and technical help, K. Brasky, E. Croft, K. Duboc, B. Free, J. Griffin, S. Gripshover, C. Meanwell, E. Mellum, M. Nikolas, J. Oberholtzer, L. Rissman, M. Ryan, B. Seibel, K. Utich and J. Wallman for help in data collection and transcription. We are grateful to the parents and children who participated in the study. Portions of this research were presented at the annual meeting of the Boston University Conference on Language Development, November 2006. This article is dedicated to the memory of Michelle E. Arroyo.

## REFERENCES

- Akhtar N. The robustness of learning through overhearing. *Developmental Science*. 2005; 8:199–209. [PubMed: 15720377]
- Akhtar N, Jipson J, Callanan MA. Learning words through overhearing. *Child Development*. 2001; 72:416–30. [PubMed: 11333075]
- Barnes S, Gutfreund M, Satterly D, Wells G. Characteristics of adult speech which predict children's language development. *Journal of Child Language*. 1983; 10:65–84. [PubMed: 6841502]
- Blum-Kulka, S.; Snow, CE., editors. *Talking to adults: The contribution of multiparty discourse to language acquisition*. Lawrence Erlbaum Associates Publishers; Mahwah, NJ: 2002.
- Brown P. Children's first verbs in Tzeltal: Evidence for an early verb category. *Linguistics*. 1998; 36(4):713–53.
- Carpenter M, Nagell K, Tomasello M. Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*. 1998; 63(4):176.
- Chavajay P, Rogoff B. Cultural variation in management of attention by children and their caregivers. *Developmental Psychology*. 1999; 35:1079–90. [PubMed: 10442876]
- de Leon, Lourdes. The emergent participant: Interactive patterns in the socialization of Tzotzil (Mayan) infants. *Journal of Linguistic Anthropology*. 1998; 8:131–61.
- Dunn, LM.; Dunn, LM. *Peabody Picture Vocabulary Test*. 3rd edn. American Guidance Service; Circle Pines, MN: 1997.
- Floor P, Akhtar N. Can 18-month-old infants learn words by listening in on conversations? *Infancy*. 2006; 9:327–39.

- Gleitman LR, Newport EL, Gleitman H. The current status of the motherese hypothesis. *Journal of Child Language*. 1984; 1:43–79. [PubMed: 6699113]
- Hart, B.; Risley, T. *Meaningful differences in the everyday experience of young American children*. Brooks; Baltimore: 1995.
- Hoff E. The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*. 2003; 74:1368–78. [PubMed: 14552403]
- Huttenlocher J, Haight W, Bryk A, Seltzer M, Lyons T. Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*. 1991; 27:236–48.
- Huttenlocher J, Vasilyeva M, Cymerman E, Levine S. Language input and child syntax. *Cognitive Psychology*. 2002; 45:337–74. [PubMed: 12480478]
- Huttenlocher J, Vasilyeva M, Waterfall HR, Vevea JL, Hedges LV. The varieties of speech to young children. *Developmental Psychology*. 2007; 43:1062–83. [PubMed: 17723036]
- Jones CP, Adamson LB. Language use in mother–child and mother–child–sibling interactions. *Child Development*. 1987; 58:356–66.
- Lieven, EVM. Crosslinguistic and crosscultural aspects of language addressed to children. In: Gallaway, C.; Richards, BJ., editors. *Input and interaction in language acquisition*. Cambridge University Press; Cambridge: 1994. p. 56-73.
- Newport, EL.; Gleitman, H.; Gleitman, LR. Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In: Snow, CE.; Ferguson, CA., editors. *Talking to children*. Cambridge University Press; New York: 1977. p. 109-149.
- Oshima-Takane Y, Robbins M. Linguistic environment of second born children. *First Language*. 2003; 23:21–40.
- Rowe ML. Child-directed speech: Relation to socioeconomic status, knowledge of child development, and child vocabulary skill. *Journal of Child Language*. 2008; 35:185–205. [PubMed: 18300434]
- Saffran J, Hauser M, Seibel R, Kapfhamer J, Tsao F, Cushman F. Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition*. 2008; 107(2):479–500. [PubMed: 18082676]
- Shimpi, P.; Akhtar, N. Learning actions and words from third-party interactions; Unpublished paper presented at the biennial meeting of the Society for Research in Child Development; 2011;
- Shneidman LA, Shimpi PM, Sootsman-Buresh J, Knight-Schwartz J, Woodward AL. Social experience, social attention and word learning in an overhearing paradigm. *Language Learning and Development*. 2009; 5(4):266–81.
- Tomasello, M. Joint attention as social cognition. In: Moore, C.; Dunham, P.; Philip, J., editors. *Joint attention: Its origins and role in development*. Lawrence Erlbaum Associates; Hillsdale, NJ: 1995. p. 103-130.
- Tomasello M, Farrar MJ. Joint attention and early language. *Child Development*. 1986; 57:1454–63. [PubMed: 3802971]
- Tomasello M, Todd J. Joint attention and lexical acquisition style. *First Language*. 1983; 4:197–211.
- Wellen CJ. Effects of older siblings on the language young children hear and produce. *Journal of Speech & Hearing Disorders*. 1985; 50:84–99. [PubMed: 3974217]



**Fig. 1.** The total number of words (tokens, top graph) and total number of different words (types, bottom graph) that children in single-speaker and multi-speaker families heard in speech directed to them by their primary caregiver, in speech directed to them by others in the household, and in speech that they overheard at age 2;6.

**TABLE 1**  
**The distribution of single-speaker and multi-speaker families by ethnicity and income**

Child group	Yearly household income	Family ethnicity			Total <sup>c</sup>
		African American	Caucasian	Other <sup>b</sup>	
Single-speaker families	Below \$50,000	2	2	2	6
	Above \$50,000	0	6	1	7
Multi-speaker families	Below \$50,000	3	1	1	5
	Above \$50,000	1	7	1	9
Total <sup>d</sup>		6 (22%)	16 (59%)	5 (19%)	27 <sup>a</sup>

<sup>a</sup>One single-speaker family and one multi-speaker family with incomes below \$50,000 did not report ethnicity. These families were excluded from this table.

<sup>b</sup>The 'Other' category included Asian and Hispanic families, along with one family of mixed ethnicity.

<sup>c</sup>The relative proportion of single-speaker and multi-speaker families in low- versus high-income groups was roughly equal: the majority of single-speaker (54%) and multi-speaker (64%) families had high income.

<sup>d</sup>The relative proportion of single-speaker and multi-speaker families within each ethnicity was roughly equal: the majority of the single-speaker (62%) and multi-speaker (57%) families were Caucasian.

**TABLE 2**  
**Regression models using input measures at 2;6 (word tokens) to predict children's**  
**receptive vocabulary skills at 3;6 (PPVT) in the multi-speaker group (n=15)**

Predictors	PPVT 3;6 $\beta$ (standardized)		
	Model 1	Model 2	Model 3
Direct tokens from primary caregiver	0.42	0.60*	0.61*
Direct tokens from others		0.50+	0.50+
Overheard tokens			0.002
R <sup>2</sup> statistic	0.18	0.39	0.39

NOTE:

\*  $p < 0.05$ , +  $p < 0.07$ .

**TABLE 3**

Regression models using input measures at 2;6 (word types) to predict children's receptive vocabulary skills at 3;6 (PPVT) in the multi-speaker group (n=15)

Predictors	PPVT 3;6 $\beta$ (standardized)		
	Model 1	Model 2	Model 3
Direct types from primary caregiver alone	0.47		
All directed types		0.62*	
All types in input (directed and overheard)			0.39
R <sup>2</sup> statistic	0.22	0.38	0.15

NOTE:

\* $p < 0.05$ .