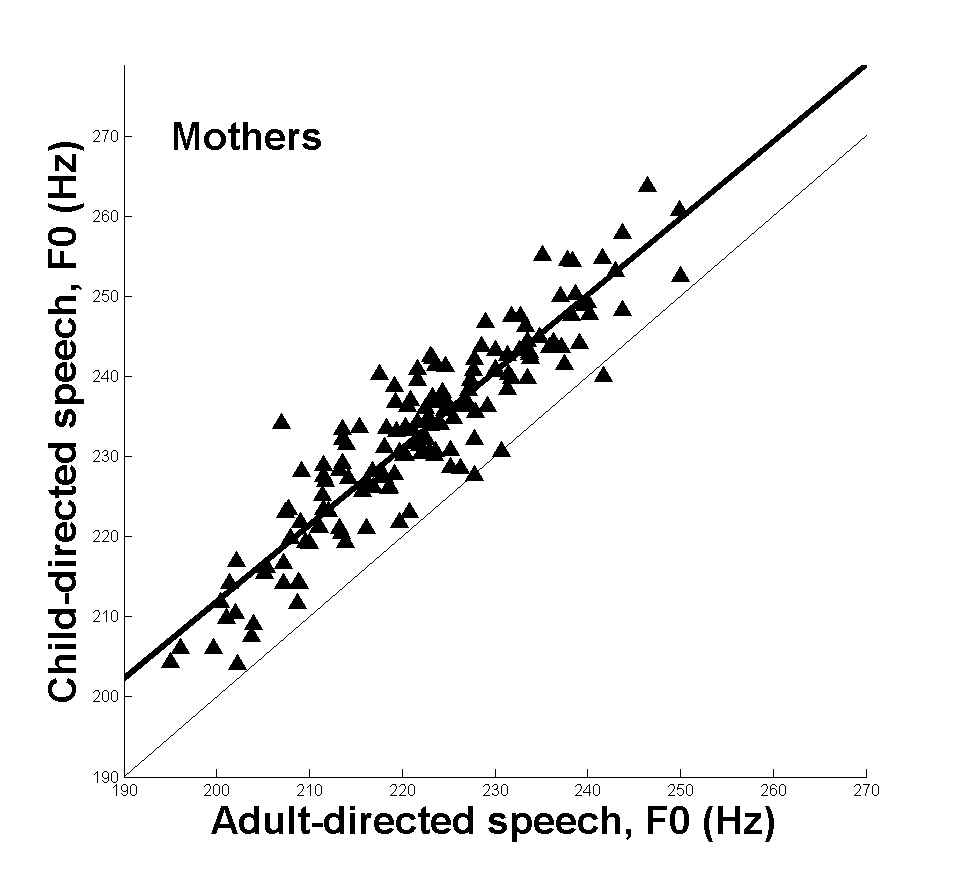
**V. Results**

In accord with our expectation, mean *f*0 values for mothers and fathers were consistent with known values for adult women and men (*M*mothers=227.5 Hz, *SD*mothers=54.2 Hz; *M*fathers=148.5 Hz, *SD*fathers=40.4 Hz).

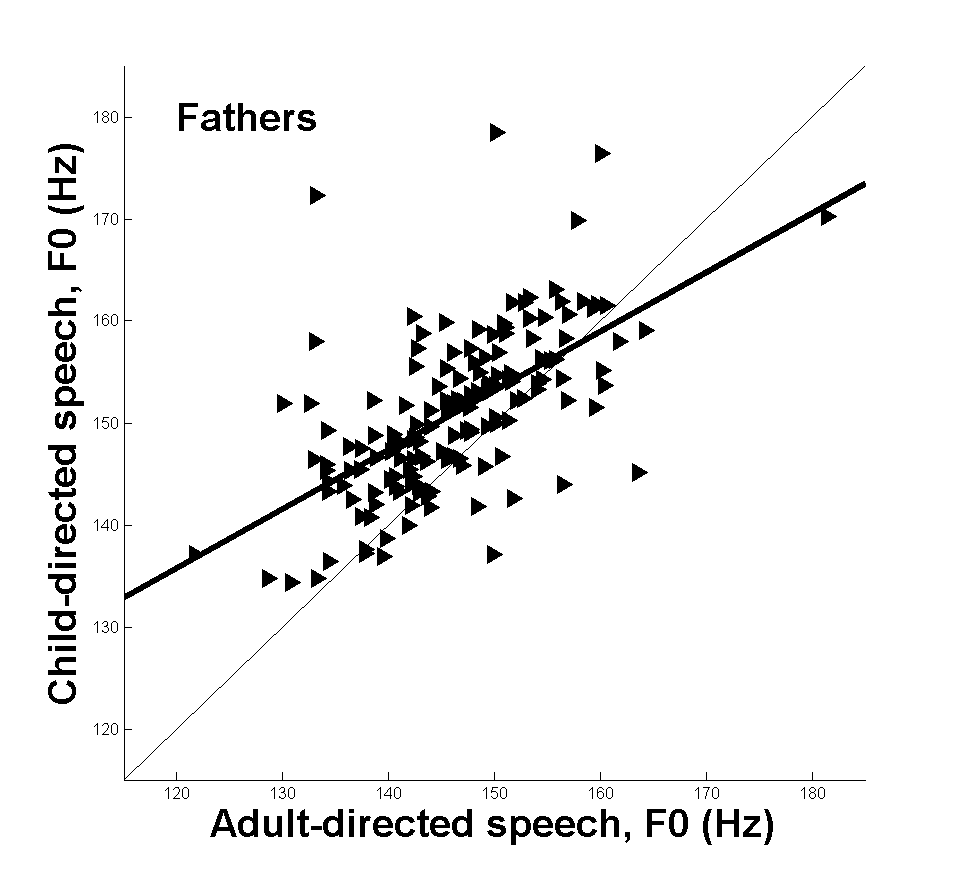
During periods of ADS, mothers’ mean *f*0 was 222.1 Hz (*SD*=53.6 Hz) and during CDS it was 233.0 Hz (*SD*=54.7 Hz). The difference between mothers’ ADS and CDS was significant (*t*(151)=27.89, *p*<10-60).

During periods of ADS, fathers’ mean *f*0 was 146.1 Hz (*SD*=39.4 Hz) and during CDS it was 150.9 Hz (*SD*=41.3 Hz). The difference between fathers’ ADS and CDS was significant (*t*(151)=8.07, *p*<10-12).

In Figures 1 and 2, the lighter line is a linear bisector of the equally scaled square figure. The heavier line is the least squares fitted linear regression for the distribution shown in each figure. The fit of the line was significant for both mothers (*R*2=.844, *p*<10-61) and fathers (*R*2=.373, *p*<10-16). Both mothers and fathers used higher mean *f*0 values in the CDS condition than in the ADS condition, although the relationship was stronger for mothers than for fathers.



*Figure 1*. The fundamental frequency of mothers’ adult-directed speech by child-directed speech is shown. The bisector is shown by the light line, and the least-squares linear regression is shown by the heavy line.



*Figure 2*. The fundamental frequency of fathers’ adult-directed speech by child-directed speech is shown. The bisector is shown by the light line, and the least-squares linear regression is shown by the heavy line.

**VI. Conclusion and Future Directions**

In this paper we showed that a very large database of naturally-collected audio can be processed and analyzed for features known to important for human communication. Here we analyzed hundreds of daylong recordings collected from the auditory perspective of a preschool child in his or her normal family routine. Thousands of hours of audio were collected *in situ*, diarized with automatic speech processing techniques from the LENA Foundation, then further processed by our algorithms to extract a speech feature, fundamental frequency, of specific talkers in the context of the diarization coding.

This work has two main goals. First, this is a proof of concept that a very large database of wild-type auditory data can be successfully captured and processed in a way that is useful to a wide variety of researchers. Researchers in computer science, algorithms, speech science, automatic speech recognition, speech and language disorders, engineering, bioacoustics and other fields may benefit from techniques such as those described here. Theoretical implications include applying this technology and approach to better understand fields from database management to the implementation of language in human communication systems. Practical implications of this work include better understanding of early human communication, improving algorithms and processing techniques for automatic speech processing and automatic speech recognition, and identifying communication characteristics of children who may be at risk of developmental delay or disorder.

Second, this work addresses the question of how fathers and mothers control their speech in different communicative contexts, namely when talking with adults or when talking with children. The fundamental frequency shift described here has been similarly described by others, but it has not been demonstrated with a very high number of observations or in highly naturalistic environments as shown here. Another question of interest here that has not been examined thoroughly in the literature is the speech behavior that fathers show in the presence of adults and children. Here again we show that fathers’ speech patterns are similar to mothers in gross respect—that is, on average fathers use higher *f*0 in CDS as compared to ADS—but the pattern is not identical to mothers. A detailed description of the difference is beyond the scope of this report but may reveal important differences between mothers and fathers.

Having demonstrated the fundamental utility of a very large database speech corpus and a fully explicated example, we expect this research program to have several fruitful avenues of future research. First, there is a need to refine and extend existing approaches to data collection, analysis, and processing. Researchers have reported LENA system performance, and we expect this work to continue. Nevertheless, there remain certain proprietary aspects of the system that are inaccessible to researchers. Further, the LENA system may not be possible or appropriate to use in some applications, such as for children with sensory or other disorders. To date, the LENA system has no fully functional alternatives. To address this, we are working toward developing an alternative system without proprietary restrictions. This work also includes improving algorithms in the pre- and post-processing stages of raw data analysis. Due to the large volume of data to be processed, improved efficiency and reliability are needed.

Second, this technology and approach have great potential to impact at-risk populations including children with developmental disorders and children and families from low socio-economic or other disadvantaged backgrounds. In one project, we are looking at the effect of mild-to-moderate hearing loss on the speech development of preschoolers. We are using the automatic methods here to assess speech production characteristics and compare between preschoolers with and without hearing loss.

Third, this technology can lead to better understanding of typical development. As wearable biotechnology rapidly grows and changes, researchers have a dramatically different ability to observe and document typical development, not only in the domain of communication and language but also in domains such as motor control or sociobiological characteristics, to name a just two. It is currently not well understood how observable patterns in various domains interact or may be related. For example, the work reported here suggests that fathers may use different speech characteristics than mothers in the speech they engage in with their children. Exploring these differences in a variety of contexts will help researchers better understand the role of fathers.

Fourth, despite the advantages, data collection and analysis remains a challenging task. To reduce the burden and positively leverage the results of many researchers working in this field, there are efforts to archive and document audio data, associated metadata, and processing tools (such as the pitch determination algorithm extractor used in this work). The accessible online repository HomeBank (<http://homebank.talkbank.org/>) [cite VanDam et al 2016 *Sem Speech Lang*] makes a wide variety of data available to researchers to explore new possibly applications, improve the technology, and contribute additional (raw) data.

The work reported here is an early demonstration of new, exciting technology and its application to a practical question of interest to researchers in speech and allied fields. The methods and procedures hold great promise to advance both the theoretical underpinnings and the practical application of this emerging science.