Basic Functionality

**readcsv**- takes in a csv file and only keeps the rows pertinent to child data (discards empty, teacher, or lab rows). May also discard unnecessary columns.

**childLoc**(id, time) – given a child’s id and a time, returns the location of that child as an (x, y) pair

since Ubisense and LENA data are discrete and not continuous, the child’s location will be returned as the most recent last known location given by the data

**childDist**(id, id, time) – calculates distance between two children at a certain time

**isTalking**(id, time) – returns a Boolean value for whether or not the child is talking at the given time

I plan to use the Child\_Voc\_Duration for this. I need clarification though on what it means. For example, at clock time 8:27:59 AM for Child 2, the Child\_Voc\_Duration column has a value of 2. So, can we assume that Child 2 was talking from 8:27:57 to 8:27:59? i.e., the LENA stops recording after a segment of speech is completed.

Advanced Functionality

With the basic methods defined above, we can do a number of things. Some possibilities include

* Counting the number of interactions between two children over a given time interval, where an interaction occurs when childDist <= 1.5 and **isTalking** is true for both children at least some point during the interval. An interaction ends when **isTalking** becomes false for both children for a certain length of time or the children move out of a 1.5m range from each other.
* Returning a list of all the children that a given children talks to over a given time interval
  + List could be ordered from greatest number of interactions to least

Open to other suggestions as well