Body-part categories of early-learned verbs: Different granularities at different points in development.

Josita Maouene
Indiana University
Brain and psychological Sciences
jcmaoune@indiana.edu

Shohei Hidaka
Indiana University
Brain and psychological Sciences
shhidaka@indiana.edu

Linda B. Smith
Indiana University
Brain and psychological Sciences
smith4@indiana.edu

***Abstract* - This paper builds on our previous finding that early verbs are strongly related to body parts. One evidence for this relation is the strong word associations among adults between common verbs and body parts. Although many common verbs are related to body parts, the prior evidence suggests that some verbs are strongly related to highly specific body regions (e.g., fingers) and others to larger or more diffuse regions (e.g., hand and arm). Here we ask whether this granularity or specificity in associations is related to age of acquisition. We examine the structure of adult associations of common verbs to body parts as a function of age of acquisition for a 101 verbs normatively acquired between 16 to 30 months. And we propose a new analysis to look at the development of granularity over a short time period: 16 months and for a small number of verbs: 101. We generated verb clusters based on body parts features, and analysed how these body-parts-based clusters account for variance of age of acquisition (AoA) of verbs. By applying this analysis from the 50 earliest learned verbs to the 50 latest learned ones, we found several clusters relevant to AoA in different granularity of body parts. The results fit with growing behavioural and neuro-imaging results on the role of the body – and sensory-motor interactions in the world – in verb processing.**

Index Terms – verb acquisition, body-part clusters, granularity, time of acquisition

I. Introduction

Young children learn verbs in the here and now of their own activity. They move, jump, skip or hop with their legs, they hold, cut, slice with their hands. These different actions have different granularities: move is more general than jump and jump is more general than hop. If bodily actions are deeply related to meaning –and the processes through which children acquire verb meaning – then one might expect the acquisition of verbs to be related to specifics of bodily action. If the acquisition of verbs is related to the specificity of the body part doing the action and if highly localized verbs with respect to the body part doing the action are acquired early, we ask: what is the most basic level of granularity (in terms of associated body-parts) for verbs –at the level of fingers, for example, or hand region? And (2) does that basic level change with development?

*II. Embodiment and language

There is growing interest in how sensori-motor interactions are integral to the acquisition of knowledge and to the development of cognitive processes that bear on that knowledge [1], [2]. In language, growing evidence suggest that the sensori-motor nature of processing does not concern just auditory processing, or phonological processing or motor planning and execution of speech, but language as a whole: language might reflect a general sensori-motor ability in humans [3], [4].

III. Body parts and Verbs

Recent behavioural and neuro-imaging studies with adults suggest strong links between bodily action and the understanding of common verbs. Studies have shown, for example, that performing an opposing action (e.g., pulling with one's hands when the verb is *push)* interferes with word recognition [5]. Further in functional MRI studies, presentation of a verb has been shown to activate the specific motor areas involved in producing the action labelled by that verb [6]. The connections between verbs and bodily action may also be relevant to how children learn and represent verb meanings. There has been little study of the role of embodied actions in children's verb acquisitions, although there are many observations that point to the value of this approach. For example, L. Bloom [7] noted that children learn and use words because they are relevant to their own goals, desires and actions. From this perspective, early verb meanings might be embedded in the child's physical actions in the world, rather than merely in the relations among words. Consistent with this idea, a few studies suggest that early verb use is often in relation to the child's perspective [8], [9]. For example, Huttenlocher [10] found that children both comprehended and produced verbs more when they referred to their own actions rather than when they were about the actions of others.

This work was supported by the NICHHD, grant R01 HD0 28675 to the third author.

III. GRANULARITY IN CATEGORIZATION

Categories in any domain may be defined broadly or narrowly, and a key psychological questioning in any domain is the psychologically relevant granularity. For example, in the 70's Rosch and colleagues [11] proposed that objects categories have three psychologically relevant clustering levels: a very detailed subordinate level (e.g., race car), a middle, basic, level granularity (e.g., car), and a sparse superordinate level (e.g., vehicle). Frequency of use, early acquisition, speed of recognition and reaction time experiments all indicated the basic level of nouns to be psychologically *basic*. These three levels also organize nouns into a hierarchy of nested categories, enabling narrow and broader inferences and thus a system of categories at different levels of granularity are important to understanding nouns.

The levels of granularity are not obvious for verbs unlike for nouns. There have been diverse attempts to seek psychological granularity of verbs in the context of objects named events in adults and children [12], [13], [14]. But most studies in developmental psycholinguistics have considered that verbs have a different system, a more horizontal one [15]. In the adult literature, we did find some hint at a hierarchical organization of verbs in Wordnet [16] with the concepts of hypernym and entailment: the verb Y is a hypernym of the verb X if the activity X is a (kind of) Y (travel is an hypernym of movement), the verb Y is entailed by X if by doing X you must be doing Y (to sleep is entailed by to snore). But these lack a psychologically relevant basis. We therefore propose that verbs also have potentially multiple levels of psychological partitions. Our analysis on two different properties of verbs –body parts and age of acquisition– is designed to reveal both within and between levels of verbs we call different granularities.

IV. RATIONALE

We hypothesize that body parts are psychologically relevant for the basic level of verbs the same way that Rosch established that shape is relevant for the basic level of nouns. If this holds true, we should find that groups of verbs organized by body parts should be correlated to other properties of verbs. In the present study, we consider the correlation between body parts and AoA of early-learned verbs. If body parts past the correlation test with AoA, we can infer that body parts are a good structural property of the basic level of verbs.

V. THE CORPUS OF VERB-BODY PART ASSOCIATIONS

Maouene, Hidaka & Smith [17] collected the body-part associations of 101 English verbs nouns typically learned by children by 30 months of age.

Method

The potential relevance of body parts to individual verbs was measured by providing adults with a list of verbs and asking them to supply *one* body part associated with each verb. The participants were not told the reasoning behind the task and they were *not* asked for the body part associated with action; instead participants were free to supply whatever body part popped into their heads for whatever reason. The rationale for the association task is this: If early verbs are associated with bodily actions done by particular body parts --and this is shared knowledge by mature speakers of the language – then adults should (1) systematically associate specific body parts with specific verbs and (2) they should agree with each other. These judgments comprised the data set for the computational analyses.

Subjects

The participants were 50 college undergraduate students at Indiana University; all were native speakers of English.

Stimuli

The verbs studied were 101 action terms on the MacArthur Communicative Developmental Inventory [18]. This inventory was developed from a normative study of over 1000 16 to 30 month-olds children learning English and is widely used to measure individual children's productive vocabulary. The 101 verbs in the inventory are the most frequent verbs in the productive vocabulary of children learning English during this developmental period. There is also normative data on the proportion of children producing each verb at monthly intervals from 16 to 30 months, allowing us to determine the order of acquisition of verbs in this corpus. The earliest verb on this list (*go*) is known by 50% of the children at 19 months and the latest verb (*wish*) is not known by 50 % of the children until after 30 months.

Procedure

The participants were given a randomly ordered list of the 101 verbs and asked to supply the one body part that first came to mind when they thought of the verb. There were no constraints on the body part terms that participants could offer; they were free to supply any body part, at any level of scale - e.g., fingernails, fingers, hands, arms, whole body.

Analyses

From these judgments, we created a body-part vector for each verb. This vector represents the number of adult judgers who listed each body part as associated with the verb. Nested body parts (e.g., lip, mouth, head) were treated separately. For example, the "meaning" vector for *bite* has these values within it: 29 mouth, 19 teeth, 1 head, and 1 lip as these are the numbers of the 50 adults who gave each one body part when given the word *bite*. In total, overall, 61 unique body parts were offered. We then used a correspondence analysis (CA) in order to reduce the

dimensionality of the data. In our data, most of the 61 body parts dimensions were sparse but a small portion of the body parts (~ 15 different body parts accounted for 84% of the total number of associations) was heavily used. From the CA we kept the four dimensions that had the largest variances, because the other higher dimensions (dimension having the fifth to the eighth largest variance) had almost the same structure as the first four ones. These four dimensions account for 42.5 and appear to correspond roughly to ear-verbs = dimension 1, mouth-verbs = dimension 2, eye-and-brain verbs = dimension3, arm-and-leg-verbs = dimension 4. This data compression strongly suggests that the corpus of verbs as a whole is systematically related to an organized set of body parts. Table 1 below gives the four dimensions, the correlations, the body parts and some verbs concerned by these correlations.

TABLE 1: THE FOUR DIMENSIONS ACCORDING TO THE CA.

Dim	Corr.	Body parts	Verbs
1	0.9427	ears	hear, listen
		mouth, lips,	bite, blow, drink, eat, feed, kiss,
2	0.8576	tongue	lick, say, sing,
3	0.7627	eyes, brain	cry, find, hide, look, read,
		arm, hand,	bring, build, buy, catch, clap,
4	0.7209	finger	clean, cook, draw,
		tongue, leg,	chase, cry, dance, find, go,
−4	0.7209	feet , toes	hurry, jump, kick, lick,

Next we submitted the largest group of verbs (the hand-leg verbs) to a correspondence analysis. Here the groups of verbs seem organized by finer grained regions of the body: finger, knuckle, hand, hand and arm. This finer grained substructure is shown in two dimensions in the figure below. It is this suggestion of nested categories of finer and coarser associations to body parts that motivated the present analysis.

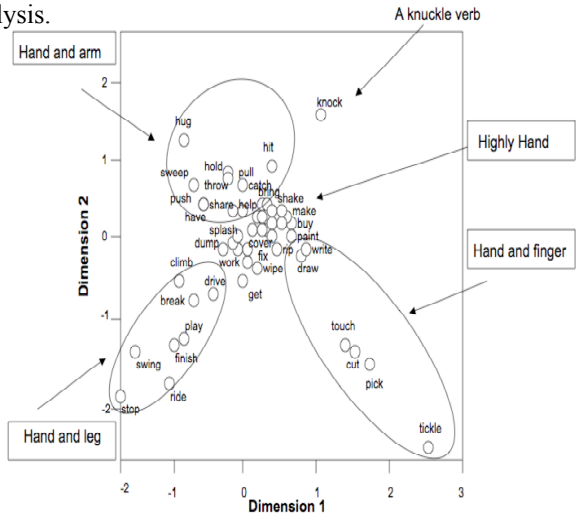

Fig.1: Feature space of arm verbs and their subsets described by the correspondence analysis.

VI. THE NEW ANALYSES

The two questions we seek to answer are (1) what is the most basic level of granularity (in terms of associated body-parts) for verbs –at the level of fingers, for example, or hand region? And (2) does that basic level change with development? In this analysis, we define the "basic level", the level at which clusters structured by body parts explain most significantly the differences in AoA of verbs.

Here, we draw an outline of the analysis. At first, we selected a group of verbs (for example, 50 verbs) ranging from 16 to 30 month of age in the MCDI lists and we sorted these verbs in 52 particular age ranges called Time Window. (Fig. 3). Each TW was submitted to a hierarchical clustering in order to characterize the verbs acquired at a particular age range based on body parts. This method enabled us to look through the developmental changes by shifting the TW one by one. Second, by applying a clustering method to the verbs within each TW, we obtained multiple levels of clustering of verbs based on body parts features. If some verbs in each body-based cluster have different AoAs, a statistical difference among clusters should show. Thus we measured the between- and within-cluster variance ratio of each level of clusters as a degree of differentiation in verb acquisition. We called this measure Acquisition Differentiation Index (ADI). Higher ADIs indicate that the verbs in the same cluster are acquired at similar periods. Applying this procedure to all possible combinations (different TWs and different levels of clusters), we drew an ADI map coordinated by the TWs and the levels of clusters. The map is literally a description of a developmental trajectory of how verbs' body part structure changes over time.

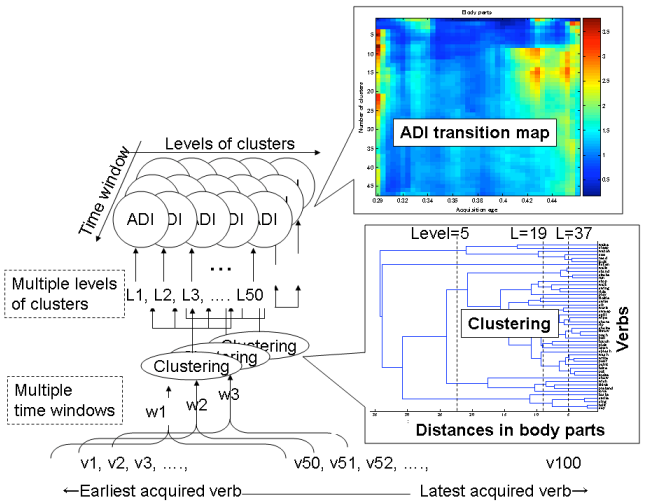

Fig. 2: The procedure of analysis.

A. Clustering

For the hierarchical clustering analysis of a group of verbs, we used the first four dimensions with the larger variance in the 61 body parts features (See the previous section for the detail). The idea is that two verbs that share a short distance (i.e. relevant to similar body parts) are clustered together. For example, *kick* and *jump* sharing *leg*

or fee tend to have a short distance and be in the same cluster whereas *bite and jump* share a long distance (*jump* is much less associated to *mouth* than *bite*) and are each in a different cluster.

B. Age differentiation Index and Correlations with Body parts

In order to evaluate the degree of differentiation of AoA in each set of verb clusters, we defined the *Age Differentiation Index* (ADI). Specifically, we calculated how much variance in acquisition age is accounted by a given set of clusters using a ratio of between-cluster variance to within-cluster variance. We used the rates of children who acquire verbs (0% to 100%) over 15 months in the MCDI list (from 16 month of age to 30 month of age). The formal definition of the ADI is given as an F-value as follows:

$$\mathrm{ADI} = BW^{-1}$$

where $W = \sum_{m}^{M}\sum_{c}^{N}\sum_{i}^{N_c}\frac{\left(R_{cim} - \overline{R}_{cm}\right)^2}{M\left(N_c - 1\right)}$ is the within-cluster variance and $B = \sum_{m}^{M}\sum_{i}^{N}\frac{(\overline{R}_{cm} - \overline{R}_{m})^2}{M(N_c - 1)}$ is the between cluster variance where R_{cim} is *i*-th verb's acquisition rate in *m*−th month belonging to *c*-th cluster, $\overline{R}_{cm}$ and $\overline{R}_{m}$ are average acquisition rates within cluster *c* in all verbs, and M , N , and N_c are number of months (M=15), number of verbs (N=50), number of verbs within cluster *c* (Nc depends on a analysis condition) respectively. If the categories given by the clustering are independent to the acquisition rates $M\left(N_c - 1\right)$ of verbs, the ADI would follows the F distribution with and $M\left(N - N_c\right)$ degree of freedom. We thus used this statistical test as criterion of significant ADI (specifically F>2.5 is our criterion). If the ages of acquisition in the two clusters are different, the variance ratio (ADI) of the two clusters would show a significant difference.

C. Time windows

We applied the procedure of clustering and variance ratios described above to different groups of verbs learned at different periods. Specifically, the 101 verbs were sorted by the average acquisition rates from 16 month to 30 month of age. Note that the acquisition order and the mean acquisition rates over 15 months of the 101 verbs is highly correlated (R=0.977). Since there is no a priori basis to deciding which developmental windows is best for comparing the specificity of body part associations, we looked at different time windows but for the clarity of the presentation, we only present the results with the time window 50 (TW=50). So this means we have 52 age groups from the earliest 50 to the latest 50 verbs. The earliest age group includes the first acquired verb (i.e., having the highest acquisition rates) up to fiftieth acquired one, and the second age group includes the second acquired verb up to the fifty-first acquired verb, and so on. In particular, if we had TW=101, we would have only one age group including all verbs. Thus the size of the time window (TW) is a parameter indicating the granularity of the time domain of verbs.

Next, we examined the distribution of the clusters by their *Peaks*. A peak is defined as the local maximum of cells which all have a significant higher ADI than a criterion. In our case, ADI=2.5 ($p<0.05$ criterion) was used as the threshold. We found that the number of peaks from TW=35 to TW=55 are stable around seven to nine peaks (Figure 3). So we chose the time window 50 resolution with eight peaks. The figure below shows the distribution of the number of peaks for each size of the time window.

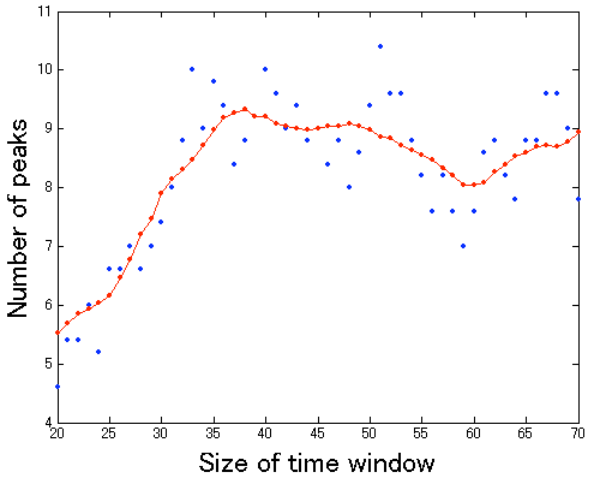

Fig.3 : The blue dots show average numbers of peaks for each size of time window. The red line shows the moving average of the average numbers of peaks.

VII. RESULTS

The result of the ADI map (Fig. 4) shows eight localized islands or Peaks. So the test of significantly positive correlations between body parts and AoA is passed.

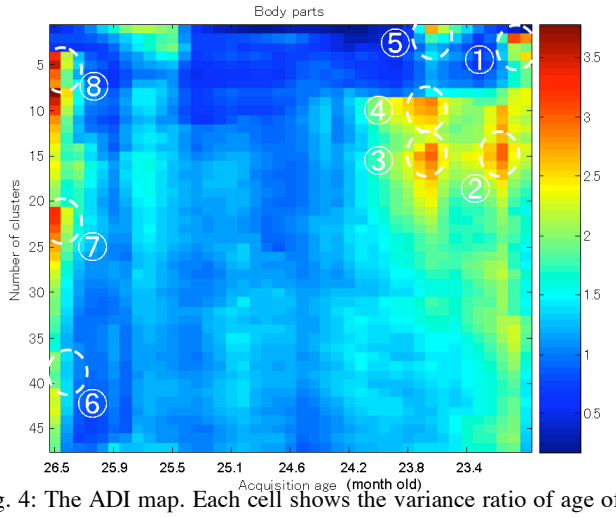

Fig. 4: The ADI map. Each cell shows the variance ratio of age of acquisition of verbs in light of within- and between-body-based clusters. TW50. 8 peaks of maximum correlation show in red.

The further analyses answer our two questions and offer two potentially important results about verb categorization: First, the most basic level of granularity in terms of the associated body parts for verbs is not very fine coarse, nor very general, it is regional. Three clusters have a significant correlation at the first mean age of acquisition (23.4 month), one cluster at the level of hand (and not at the level of the finger or the upper body), one at the level of the eye (and not at the iris or at the face level) and one at the level of the mouth (and not at the tongue or at the lower-face level). Second, this basic level develops: in the 50 verbs learned in the midst of this span, one cluster including ear-verbs (and not lobes or head sides) gets added, i.e. has a significantly different AoA than the other verb clusters. In the 50 verbs learned at the end of this span, one cluster gets added at the basic level: the eye-brain verbs (and not at the left hemisphere or the upper-body level)(Peak 8). This basic level of verb clustering follows a partition of the body that is neither too general nor too specific. This result is reminiscent of what Rosh said on the specific status of the basic level for objects. Below, examples of clusters of verbs for each peak verbs. Any clusters having at least two verbs are shown with their examples. The within clusters are separated by slashes. The peaks in bold correspond to the changes in the basic level.

TABLE 2: SAMPLE OF VERBS FOR EACH PEAK OF THE TW50.

Peak	Mean age	No. of clusters	Examples
1	**22.5**	**3**	**sing, blow, bite, drink/** wipe, catch, splash, carry/ **find, hide, watch, look**
2	22.7	15	**chose, drive/** look, read/ **find, watch/** break, play/ **ride, stop/** drop, wipe/ **blow, bite/** kick, dance
3	23.3	15	**have, knock,/** look, read/ **find, watch/** break, play/ **ride, stop/** draw, give/ **kick, dance**
4	23.5	10	**ride, stop/** spill, have/ **kick, dance/** talk, sing/ **find, watch**
5	**25.9**	**2**	**listen, hear/** finish, pour, taste, pick, lick, shake
6	26.8	37	**pour, build/** chase, hurry/ **paint, give/** cut, touch/ **make, put/** shake, share/ **say, talk/** listen, hear
7	26.8	19	**wish, hate/** pick, cut/ **rip, dump/** pretend, think/ **pour, build/** smile, talk/ **hurry, stand/** listen, hear
8	**26.8**	**5**	**wish, hate/** pour, pick/ **read, talk/** wake, find/ **listen, hear**

.

A third potentially interested result for the extension of the verbs within the basic level is that the body parts relevant to AoA look more *differentiated* in verbs learned later. In other words, children seem to learn earlier verbs relevant to large partitions of the body (eyes, mouth, and hands), and they seem to learn later verbs relevant to more differentiated partitions of body parts (eyes, mouth, hands, brain, ears, and their subsets). The results in the left end of the ADI map (Fig. 4) indicate that children learn verbs at different levels of clustering in parallel. Figure 5 shows the clustering tree of the 50 latest verbs that corresponds with the peaks in the ADI map. Looking at the specific items in these clusters (Table 2), brain verbs (e.g., *wish* and *hate*) or arm verbs (e.g., pour and pick) in the 50 latest learned verbs (Peak 8 in Table 2) have more subsets of brain verbs (e.g., *wish*, *hate*/ *pretend*, *think*) and arm verbs (e.g., *pour*, *build*/ *pick*, *cut*) (Peak 7). It indicates that children learn these hand and brain verbs at a different period *but also* that they learn "finger" verbs and "arm" verbs which are subsets of hand-general, and or "emotional" and "cognitive' verbs (or brain-general) verbs at a different period too. This result suggests that children develop different partitions of the verb space by varying the degree of differentiation in body parts: hands verbs and brain verbs have a different AoA, but also arm verbs and hand verbs (or emotional verbs and cognitive verbs) have a different AoA. In sum, multiple levels in the latest TW suggest that the structure of the meaning of the verbs acquired later is hierarchical or is at least more complicated than a single level structure.

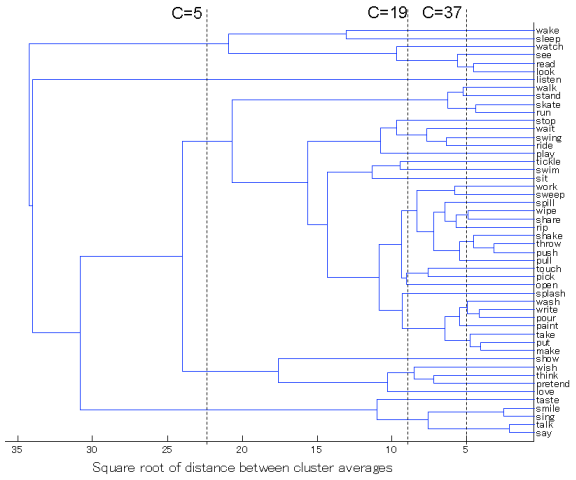

Fig. 5: the latest 50 learned verbs clustering based on body parts features.

VIII. DISCUSSION

These new analyses provide four different results. AoA and verbs clustered by body parts are correlated. The basic level of verbs clustered by the adult associations to body parts is a middle level of generality. Verbs relevant to the middle partitioning by body part clusters tend to be learned earlier and get complemented over time. In the latest period, different partitions of the verb space by varying the degree of differentiation in body parts. These results suggest that like in noun or event categorization, time (AOA) and part structure (body parts) is fundamental to verb categorization. Further, differentiation (middle to fine-graded verbs), a major landmark of development, is observable as well as, complexity, another landmark of development. The next step will be to test these results in a series of experimental studies with children.

REFERENCES

[1] W. Barsalou, "Perceptual symbol systems," *Behavioral and Brain Sciences*, 1999, vol. 22, pp. 577-660.

[2] M. Wilson, "Six views of embodied cognition", *Psychonomic Bulletin & Review, 9*, 2002, pp. 625-636.

[3] A.E. Hernandez, and P. Li, "Age of acquisition: its neural and computational mechanisms", *Psychological Bulletin*, 2007, vol. 133, no. 4), pp. 638-650.

[4] Bates, E., L. Begnini, I. Bretherton, L. Camaioni, & V. Volterra *The emergence of symbols: Cognition and communication in infancy*, New York, Academic Press, 1979.

[5] A.M. Glenberg, T. Gutierrez, T., J.R. Levin, S. Japuntich, & M.P. Kaschak, "Activity and imagined activity can enhance young children's reading comprehension," *J. of Educational Psychology*, vol.96, 2004, pp. 424-436.

[6] O. Hauk, L. Johnsrude, & F. Pulvermuller, "Somatotopic representation of action words in human motor and premotor cortex, " *Neuron*, Jan 22, vol. 41, no 2, 2004, pp. 301-307.

[7] L. Bloom, *Language development from two to three,* Cambridge, Cambridge University Press, 1992.

[8] M.D.S. Braine, "Children's first words", *Annual Review of Psychology*, vol. 52, 1976, pp. 369-396.

[9] M. Tomasello, *First verbs: a case study of early grammatical development*, 1992, Cambridge University Press.

[10] J. Huttenlocher, P. Smiley, and R. Charney, "Emergence of action categories in the child. Evidence from verb meaning." *Psychological Review*, vol. 90, no 1, 1983, pp. 72-93.

[11] E. Rosch, C. B. Mervis, W.D. Gray, D. M. Johnson, and P. Boyes-Braem, " Basic objects in natural categories," *Cognitive Psychology,* vol. 8, no 3, pp. 1976, 382-439.

[12] R.G. Barker, and H. F. Wright, *Midwest and its children: The psychological ecology of an American town*, Evanston, Illinois, Row and Compagny, 1954.

[13] M. W. Morris, and G.L. Murphy, "Converging operations on a basic level in event taxonomies," *Memory and Cognition*, vol. 18, no 4, 1990, pp. 407-418.

[14] A. Rifkin, "Evidence for a basic level in event taxonomies," *Memory and Cognition*, vol.13, no 6, 1985.

[15] R. Michnick Golinkoff, K. Hirsh-Pasek, C.B. Mervid & W.B. Frawley, "Lexical Principles Can Be Extended to the Acquisition of Verbs," In *Beyond Names for things, Young children acquisition of verbs*, M. Tomasello & W.E Merriman Eds. Hillsdale, New Jersey, Erlbaum, 1995, pp. 185-221.

[16] C. Fellbaum, *Wordnet: an electronic lexical database*, Cambridge Mass, MIT Press.

[17] J. Maouene, S. Hidaka, L.B. Smith, "Body parts and early-learned verbs", *Cognitive Science, in press.,*

[18] L. Fenson, P.S. Dale, J.S. Reznick, and E. Bates, "Variability in early communicative development," *Monographs of the Society for Research in Child Development, vol. 59*, no 5, 1994, pp. v-173.