There are more valence alternations than the ditransitive

Hao Sun & Jean-Pierre Koenig

University at Buffalo

1 Introduction

Recent research has witnessed a growing interest in determining how much usage frequency explains language patterns and language processing (Bybee, 2006; Pierrehumbert, 2001; Jurafsky, 2002). This paper focuses on the degree to which the relative frequency of verbs in syntactic frames (or valence alternations) may explain the acquisition of these frames (or these alternations). Valence alternations can be roughly defined as a set of verbs alternatively taking two or more sets of similar syntactic dependents with similar truth conditions. The English ditransitive alternation, one of the most widely discussed valence alternations, is illustrated in the sentences in (1).

(1) a. John gave a book to Mary. (prepositional object frame)
   b. John gave Mary a book. (double object frame)

(1a) illustrates the prepositional object (PO) frame, where an NP and a PP follow the verb, whereas (1b) illustrates the double object (DO) frame where the verb takes two NP complements. It has been argued that the existence of a verb that occurs particularly frequently in a syntactic frame or an alternation can facilitate the learning of that frame or alternation (Goldberg, Casenhiser & Sethuraman, 2004). In other words, statistical skewing may facilitate language acquisition (hereafter, the statistical skewing hypothesis).

In this paper, we investigate further an underlying assumption of the statistical skewing hypothesis, namely that the lexical distribution of most syntactic frames (or valence alternations) shows the same kind of skewing as the ditransitive alternation. More specifically, we try to answer the following three questions.

(i) Is the statistical skewing characteristic of the ditransitive alternation representative of all valence alternations (from now on, except when explicitly noted, we do not distinguish
between individual syntactic frames and valence alternations, as the distinction is not relevant for our purposes)?

(ii) Is statistical skewing a plausible learning theory for all valence alternations?

(iii) What is the similarity in the distributions of verb frequencies between child speech, child-directed speech, and adult-directed speech?

To answer these questions, we performed a large-scale corpus study that investigated the frequency distributions of verbs that appear in five valence alternations. We collected data from naturalistic child-directed speech and adult-directed speech from two corpora, the British National Corpus (BNC) and the Child speech Data Exchange System (CHILDES, MacWhinney, 2000). We then expanded and quantified researchers’ intuitions about statistical skewing using information-theoretic and statistical measures to compare different alternations.

2 Previous studies

Traditional accounts of language development rely on abstractions over linguistic categories (Chomsky, 1980; Pinker, 1989). In contrast, usage-based or exemplar-based theories argue for the relevance of more concrete and fine-grained linguistic representations on grammatical knowledge. Robust frequency effects have been found in phonology, morphology, language comprehension, production and acquisition that attest of the relevance of concrete and fine-grained representations. To take but one example, frequent words are more likely to resist analogical leveling in historical changes and easier to undergo grammaticalization than low frequency words (Bybee, 2006) and are accessed faster from long-term memory than infrequent words (Forster & Chambers, 1973; Garney et al., 1997); frequent words are also faster to undergo phonological reductions (Bybee & Scheibman, 1999). (See Bybee, 2006 and Jurafsky, 2003 for summaries.)

More importantly for our paper, the learning of the ditransitive has been argued to be influenced by the relative frequency of occurrence of particular verbs in the ditransitive (Goldberg et al., 2004; Casenhiser & Goldberg, 2005; Ellis & Larsen-Freeman, 2009). Goldberg and her colleagues (Goldberg et al., 2004) performed a corpus study of the ditransitive construction in both child speech (CS) and child-directed speech (CDS). They concluded that give accounts for the lion’s share of the token frequency of all the verbs that appear in this
construction and showed experimentally that such skewed frequency distribution facilitates the learning of valence alternations (see Goldberg et al., 2004; Casenhiser & Goldberg, 2005).

But is the ditransitive alternation representative of the fifty or so valence alternations Levin (1993) reports? This is an important question as psycholinguistic work on valence alternations almost disproportionately study the ditransitive and implicitly assume it is representative of all English valence alternations. Language production studies on valence alternations, especially structural priming studies, have a long tradition of focusing on the ditransitive alternation (Bock, 1986; Pickering & Branigan, 1998, among many others). The same is true of cross-linguistic or bilingual structural priming studies (Pappert et al., 2011; Hartsuiker, Pickering & Veltkamp, 2004; Huttenlocher, Vasilyeva, & Shimpi, 2004) or eye-tracking studies (Arai, Van Gompel, & Scheepers, 2007; Thothathiri & Snedeker, 2008). Psycholinguists explicitly or implicitly seem to assume that valence alternations all behave the same and that what is true of the ditransitive is true of all valence alternations.

3 Experiment setup

3.1 Target valence alternations

We investigated the verb frequency distributions from a total of five valence alternations: the conative, the causative-inchoative, the spray-load, the active-passive, and the ditransitive alternations.

(2) Conative Alternation

(a) The child hit the toy. (transitive frame)

(b) The child hit at the toy. (conative frame)

(3) Causative-inchoative Alternation

(a) The boy broke the window. (causative frame)

(b) The window broke. (inchoative frame)

(4) Spray-load alternation

(a) The man sprayed the paint on the wall. (locative frame)
(b) The man sprayed the wall with the paint.  

(5) Active-passive alternation

(a) The little girl ate the apple.  

(b) The apple was eaten by the little girl.  

We chose these alternations both because there is extensive work on each alternation and because these alternations differ both in the kind of differences in meaning or information structure that exist between the two alternative frames and the number of verbs that participate in these alternations.

3.2 Corpora

We collected verb frequency distributional data from two corpora. We collected child speech and child-directed speech from the Child speech Data Exchange System (CHILDES) (MacWhinney 2000). The CHILDES corpus is the product of the joint work of the language development community and consists of transcribed conversations. The other corpus is the spoken section of the British National Corpus (BNC, Leech, 1992), from which we collected adult-directed speech. We recombined most of the English data from the CHILDES corpus which sum up to 4 million words to facilitate data collection.

3.3 Methodology

We combined automatic searching with manual filtering and participating verbs were extracted according to Levin’s verb lists (Levin, 1993). We first searched the BNC and the CHILDES corpora with part of speech search patterns. We then manually filtered the data that matched the search patterns.

3.4 Preliminary results

The general trends in verb frequency distributions are represented in Figure 1.
**Figure 1:** The frequency distribution of verbs that appear in each alternation in child-directed speech and adult-directed speech. The X-axis denotes participating verbs ranked by frequency. The Y-axis represents the probability of each verb occurring in the alternation. Verbs are ranked by decreasing frequency.

As Figure 1 shows, the frequency distribution of verbs in each alternation exhibits the long-tail distribution predicted by Zipf’s law (Zipf, 1931) according to which the frequency of a word is inversely proportional to its rank. For better visualization of the data, all verb frequencies are normalized to the range 0 to 1. The ditransitive stands out in Figure 1 as a single peak appears at the left of the graph (corresponding to the verb *give*). In contrast, other alternations display much
flatter frequency curves, although the data we collected for some alternations is very sparse, in particular for the spray/load alternation.

**Table 1:** Basic descriptive statistics of the type and token frequency of verbs that appear in the five valence alternations. Numbers on the left of the slash correspond to Child Directed Speech (CDS); numbers on the right of the slash correspond to Adult Directed Speech (ADS)

<table>
<thead>
<tr>
<th>Alternation</th>
<th>Verb type inventory size</th>
<th>Token Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ditransitive</td>
<td>17/34</td>
<td>6289/12428</td>
</tr>
<tr>
<td>Spray-load</td>
<td>5/19</td>
<td>175/224</td>
</tr>
<tr>
<td>Causative-inchoative</td>
<td>33/39</td>
<td>8159/8151</td>
</tr>
<tr>
<td>Conative</td>
<td>10/10</td>
<td>1158/3373</td>
</tr>
<tr>
<td>Active-Passive</td>
<td>45/77</td>
<td>112046/179705</td>
</tr>
</tbody>
</table>

The type and token frequencies of verbs participating in each alternation are listed in Table 1. As can be seen, verb type inventory size varies a lot across alternation and corpora, and this has ramifications for our skewing and similarity measures. As will be discussed in the following sections, verb type inventory size has a large effect on measures of statistical skewing and cross-corpora comparisons of verb frequency distributions.

### 3.5 Age of acquisition

A prerequisite for the validity of the statistical skewing hypothesis is that the most frequent verbs that occur in an alternation should be acquired very early; otherwise, their frequent occurrence could not help acquisition. This section provides an indirect and partial test of this age-of-acquisition (prerequisite) prediction made by the statistical skewing hypothesis.
Table 2: Verbs participating in five alternations that are among the first 100 verbs acquired by children (Fenson, et al., 1994). Verbs within each alternation are ordered in increasing age of acquisition. Numbers in parenthesis indicate the frequency rank of a verb in an alternation.

<table>
<thead>
<tr>
<th>Alternation</th>
<th>Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ditransitive</td>
<td>give (1), bring (4), show (2)</td>
</tr>
<tr>
<td>locative</td>
<td>put (1)</td>
</tr>
<tr>
<td>conative</td>
<td>knock (1), bite (2)</td>
</tr>
<tr>
<td>causative</td>
<td>open (1), break (3), close (2), drop (7)</td>
</tr>
<tr>
<td>active-passive</td>
<td>break (1), close (8), finish (6), tear (9)</td>
</tr>
</tbody>
</table>

The data on the age of acquisition of verbs is based on Fenson et al., 1994, where English speaking children were tested on the comprehension and production of various words. The most frequent 5 verbs that participate in each alternation that are included in the list of the first 100 verbs acquired by 2-year-old toddlers (Fenson et al., 1994). Table 2 shows that the most frequent verbs in the ditransitive, the conative, and the causative-inchoative alternations are all acquired very early, namely before children are 2-years old, as predicted by the statistical skewing hypothesis.

4 Experiment 1: statistical skewing

In Experiment 1, we collected verb frequency distribution in the whole CHILDES data set as well as the spoken BNC data set.

4.1 Methodology

Two analyses were conducted on verb frequency distributions, one on each syntactic frame or argument-structure construction and one on each alternation. For the second analysis, the frequency of occurrence of a verb in an alternation was defined as the minimal frequency of occurrence of that verb in either frame.
4.2 Results and observations


<table>
<thead>
<tr>
<th>Child-directed speech</th>
<th>Adult-directed speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frame 1</td>
<td>Frame 2</td>
</tr>
<tr>
<td><strong>ditran</strong></td>
<td>give (DO)</td>
</tr>
<tr>
<td><strong>spray</strong></td>
<td>mark (loc)</td>
</tr>
<tr>
<td><strong>caus</strong></td>
<td>turn (act)</td>
</tr>
<tr>
<td><strong>con</strong></td>
<td>cut (tran)</td>
</tr>
<tr>
<td><strong>pass</strong></td>
<td>do (act)</td>
</tr>
</tbody>
</table>

Table 3 lists the most frequent verbs that appear in the different frames in the child-directed speech and the BNC. We excluded change and knock from further analysis, because change and knock have very special meanings in child speech. Both the DO frame and the PO frame in our two corpora have the same verb give as the most frequent verb. In child-directed speech, give alone accounts for 68.47% and 84.48% of the verbs that appear in the dative frame and the ditransitive frame respectively. In contrast, there is more variation in the most frequent verbs that appear in other valence alternations, as in no other alternation is the most frequent verb in both frames the same.

It is also worth noting that resultative passives, in contrast to process passives are most frequent in child-directed speech and child speech, as is illustrated by the sentences in (6). This resultative passive bias is not present in adult-directed speech: Resultative passives are common in adult-directed speech but they are not the most frequent kind of passives.
(6) resultative passives that appear in child-directed speech
(a) They’re called cement mixers.
(b) Her school was closed.

The predominance of resultative passives in child-directed speech and child speech suggests that the acquisition of passives might initially involve learning resultative passives and then extending the passive frame to cover process passives at later stages of language development.

4.3 Measuring statistical skewing

We now present the results of two measures of statistical skewing, one that uses descriptive statistics and the other information entropy. We also report on the performance of two plausible probability measures of statistical skewing, skewing and kurtosis, and show that these two measures are not appropriate for capturing the intuition behind the statistical skewing hypothesis.

Statistical skewing can be defined, intuitively, as the minimal number of verb types needed to reach a specific probability threshold. We thus first ranked verbs that participate in a specific frame in a descending order of frequency and then used quartiles to capture this intuition. Here, quartiles are defined as the number of verbs required to reach 25%, 50% and 75% of the total probability mass.

Table 4: Three quartile measures of statistical skewing for verbs that alternate in child-directed speech and in adult-directed speech

<table>
<thead>
<tr>
<th>quartiles</th>
<th>Child-directed speech</th>
<th>Adult-directed speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>ditran</td>
<td>0.25 0.5 0.75</td>
<td>0.25 0.5 0.75</td>
</tr>
<tr>
<td></td>
<td>1 1 1</td>
<td>1 1 3</td>
</tr>
<tr>
<td>spray</td>
<td>1 2 3</td>
<td>2 4 7</td>
</tr>
<tr>
<td>caus</td>
<td>1 3 6</td>
<td>2 5 12</td>
</tr>
<tr>
<td>conative</td>
<td>1 3 4</td>
<td>1 2 4</td>
</tr>
<tr>
<td>passive</td>
<td>2 5 11</td>
<td>3 7 15</td>
</tr>
</tbody>
</table>
As shown in Table 4, the ditransitive alternation requires fewer verbs to reach target thresholds in both child-directed speech and in adult-directed speech than other alternations and the causative/inchoative and the active/passive alternations need the most verbs to reach the threshold. Furthermore, alternations seem to be more skewed in child-directed speech than in adult-directed speech, something of relevance to the statistical skewing hypothesis (since it claims that the more skewed an alternation is, the easier it is to learn).

We also use information entropy to measure statistical skewing following previous treatments of statistical skewing (Gomez, 2002; Ellis & Larsen-Freeman, 2009) and reduce the effect of verb inventory size by computing entropy ratios against baselines. Entropy, an information theory term, measures the average information content in a series of events by measuring their uncertainty (Shannon, 1948). Intuitively, a series of predictable events contain less valuable information than a series of unpredictable events. As shown in Equation (1), entropy is defined as the average of information content.

$$H(X) = - \sum_{i} P(x_i) \cdot \log P(x_i)$$  \hspace{1cm} (1)

Here, we set the base of the logarithm to 2 for sake of simplicity, in which case entropy can be interpreted as the average number of bits needed to encode a sequence of verbs that participate in an alternation. It is worth noting that entropy is sensitive to the number of values of the variable under study: the entropy of a discrete variable $x$ increases with the number of possible values of $x$. This is worrisome for us when measuring skewing of verb frequency distributions, as the number of verbs participating in an alternation varies by alternation and corpus. To eschew this issue, we provide baselines for each alternation from two corpora and measure to what extent the entropy of an alternation deviates from its baseline entropy by computing the ratio between a target entropy and a random baseline entropy. We used as random baseline entropy the average entropy of 100 random frequency distributions that involved the same number of verb types as the alternation under consideration.
**Figure 2:** Entropy of the frequency distributions of verbs that alternate in child-directed speech and adult-directed speech: (a) entropy over actual frequency distributions (b) entropy ratios, computed as target entropies divided by corresponding baseline entropies.

Figure 2 shows the result of entropy analyses of our five target alternations with Figure 2(b) teasing apart the statistical skewing effect from the verb type size effect. If a ratio in Figure 2(b) is below 1, the corresponding entropy is lower than its baseline. As Figure 2 shows, there is considerable variation in entropy across alternations and corpora. The ditransitive has the lowest entropy and entropy ratio among the five alternations, which suggests that the ditransitive is very skewed (regardless of verb inventory size). The active-passive alternation in Figure 2(a) has the highest entropy, which is due to its large participating verb type size, as shown by the fact that its entropy ratio is close to the entropy ratios of the conative, the spray-load and the causative-inchoative in Figure 2(b). In Figure 2(a), the entropy of child-directed speech is consistently lower than the entropy of adult-directed speech, but Figure 2(b) presents a more mixed result, suggesting that the skewing effect in child-directed speech is partly due to its small verb inventory size. Overall, the results of our entropy analysis agree with the more intuitive descriptive statistics analysis we reported in Table 4.

5 **Experiment 2: cross-corpora similarities**

In a second experiment, we tested the similarity of child speech to child-directed speech and adult-directed speech. This is important, as if the statistical skewing hypothesis is correct, children must pick up on regularities in the frequency distribution they are exposed to (in child
directed speech). We therefore expect the frequency distribution of verbs in children’s productions to more closely resemble the frequency distribution of verbs in child directed speech than the frequency distribution of verbs in adult directed speech. Additionally, measures of cross-corpora similarities are of relevance to at least two classes of theories. First, child-directed speech is argued to be qualitatively different from adult-directed speech in terms of its lexicon, syntactic structures and sentence intonations. If true, we would expect child speech and child-directed speech verb frequency distributions to be more similar than child speech and adult directed speech verb frequency distributions. Second, language acquisition theories differ on the extent to which children imitate their language input.

5.1 Data and methodology

Data collection procedure was similar to Experiment 1. To measure similarity of verb frequency distributions, we used symmetric Kullback-Leibler divergence (KL-divergence) as our measure, another information-theoretic measure widely used in computational linguistics to assess cross-corpora distribution similarities, or more precisely distances.

$$KL(P||Q) = - \sum_{i} P(x_i) \cdot \log \left( \frac{P(x_i)}{Q(x_i)} \right)$$

Empirically, KL-divergence measures the distance between two probability distributions $P$ and $Q$, as shown in Equation (2). If two probability distributions $P$ and $Q$ are exactly the same, the KL-divergence between $P$ and $Q$ equals 0. The original KL-divergence is asymmetric; in other words, the distance from $P$ to $Q$ is not the same as the distance from $Q$ to $P$. Here, we define a symmetric version of KL-divergence shown in Equation (3), where $P$ and $Q$ are interchangeable. The lower the symmetric KL-divergence, the more similar two frequency distributions are.

$$symmetricKL(P, Q) = KL(P||Q) + KL(Q||P)$$

Similar to Experiment 1, the baseline condition was set to the average (symmetric) KL-divergence computed from 100 random frequency distributions. Because KL-divergence is also sensitive to verb inventory size, we divided the target symmetric KL-divergences by our baseline symmetric KL-divergences.
5.2 Results

The results from the computation of the symmetric KL-divergence of verb frequency distributions over the three pairs of corpora we analyzed are represented in Figure 3.

**Figure 3:** Symmetric KL-divergences between child speech and child-directed speech (CS-CDS), between child speech and adult-directed speech (CS-ADS) and between child-directed speech and adult-directed speech (CDS-ADS): (a) KL-divergence for (target) verb frequency distributions, (b) KL-divergence ratios, computed as target KL-divergence divided by the corresponding baseline KL-divergence.

Figure 3, overall, represents measures of distance between verb frequency distributions across corpora. Ratios below 1 in Figure 3(b) correspond to cases where the KL-divergence is lower than the random baseline divergence. At least three salient trends can be observed in Figure 3. First, the ditransitive has very consistent frequency distribution patterns across corpora and the divergence measure, around 0.5, is well below 1 in all the three comparisons. In comparison, the causative-inchoative and the active-passive have much larger symmetric KL-divergence. Second, the conative and the spray-load alternations suffer from data sparsity. Sparsity is probably the reason that child speech and adult-directed speech do not share any alternating verbs in the conative alternation. Third, there is variation among alternations across the three corpora. The largest KL-divergence appears in the active-passive alternation in the CS-.
ADS comparison, around 3.5, which supports our argument that passives in CS and in ADS are qualitatively different. More generally, the observation that child-directed speech deviates more from adult-directed speech than from child directed speech is in accordance with previous corpus studies (Buttery & Korhonen, 2005; Buttery, 2006). In cases where there is enough data, CS-CDS shows consistently lower divergence scores than CS-ADS, suggesting that children’s language productions relatively closely mirror their language input in terms of verb frequency distribution.

6. Conclusions and future research

In this paper, we conducted a large-scale corpus study of five valence alternations in English. We investigated verb frequency data on child-directed speech and adult-directed speech and compared the frequency distributions of participating verbs in the ditransitive alternation to four other valence alternations. Our results suggest that the ditransitive does not seem to be representative of all valence alternations. None of the other four valence alternations are skewed like the ditransitive alternation is in both child-directed speech and in adult-directed speech. Many psycholinguistic studies implicitly or explicitly assume that valence alternations are homogeneous, generalizing from properties of the ditransitive alternation to all English alternations. Our results suggest that such extrapolation might not always be warranted. The difference between the verb frequency distribution of the ditransitive alternation and other alternations further suggests that the statistical skewing hypothesis is more plausible as an explanation of the ease of learning for the ditransitive alternation than for other alternations. Finally, our study suggests that learning the adult-like active-passive alternation might follow a different path than learning other alternations, given the differences between the active-passive alternation in child directed speech and adult direct speech. As the most frequent verbs in the active-passive alternation used by children (and in child directed speech) are resultative passives for the most part, process passives (that form the bread and butter of adult passives) are acquired later.

The results of our second study where we used symmetric KL-divergence to compare verb frequency distributions across corpora show that the frequency distribution of verb uses roughly
reflects the frequency distributions of their input in that child speech deviates less overall from child-directed speech than from adult-directed speech.

There remain two important questions for future research. The first question concerns the representativeness of the valence alternations we investigated, as we studied only about 10% of the alternations listed in Levin (1993). Other, less common alternations, might involve different verb frequency distributions than the ones we discussed here or so few verbs that generalizing from the association of a very frequent verb and an alternation might not even be feasible. The second question pertains to the role that the various skewing patterns we found in our corpus study might play in valence alternation learning, in particular whether the reduced skewing we found in all but the ditransitive construction might still facilitate learning. We are currently devising an artificial grammar learning experiment to answer this question.

References


Appendix A

A summary of CHILDES corpus data used in this corpus study is presented in Table A1.

Table A1: Child-directed speech and child language data used in this study

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Year</th>
<th>Age of children</th>
<th>Data Source</th>
<th>Year</th>
<th>Age of children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bates</td>
<td>1988</td>
<td>1;8 and 2;4</td>
<td>Haggerty</td>
<td>1929</td>
<td>2;6</td>
</tr>
<tr>
<td>Berstein</td>
<td>1982</td>
<td>1;1–1;11</td>
<td>Hall</td>
<td>1979</td>
<td>4;6-5;0</td>
</tr>
<tr>
<td>Bloom</td>
<td>1970,</td>
<td>1;9–3;2</td>
<td>Higginson</td>
<td>1983-1984</td>
<td>0;11-2;11</td>
</tr>
<tr>
<td></td>
<td>1974,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1975</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bloom</td>
<td>1973</td>
<td>1;4.21–2;10</td>
<td>McCune</td>
<td>1970s</td>
<td>0;9-3;0</td>
</tr>
<tr>
<td>Bohonan</td>
<td>1976</td>
<td>2;8 and 3;0</td>
<td>Morisset</td>
<td>1990</td>
<td>2;6</td>
</tr>
<tr>
<td>Brent</td>
<td>2001</td>
<td>0;6-1;0</td>
<td>New England</td>
<td>1989</td>
<td>1;2-5</td>
</tr>
<tr>
<td>Brown</td>
<td>1973</td>
<td>1;6-5;1</td>
<td>Post</td>
<td>1986</td>
<td>1;7.5-2;8.7</td>
</tr>
<tr>
<td>Clark</td>
<td>1976</td>
<td>2;2–3;2</td>
<td>Rollins</td>
<td>2003,</td>
<td>0;6-1;0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2006,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2011</td>
<td></td>
</tr>
<tr>
<td>Demetra</td>
<td>1989</td>
<td>2;0–3;11</td>
<td>Sachs</td>
<td>1969-1973</td>
<td>1;1 -5;1.</td>
</tr>
<tr>
<td>Demetra</td>
<td>1989</td>
<td>2;1-2;9</td>
<td>Soderstrom</td>
<td>1998-2003</td>
<td>0;6-1;0</td>
</tr>
<tr>
<td>Evans</td>
<td>NA</td>
<td>First graders</td>
<td>Suppes</td>
<td>1972-1973</td>
<td>1;11-3;11</td>
</tr>
<tr>
<td>Hayes</td>
<td>NA</td>
<td>1;6–4;0</td>
<td>Tardif</td>
<td>1991-1992</td>
<td>2;0</td>
</tr>
<tr>
<td>Feldman</td>
<td>2001</td>
<td>1;2-2;3</td>
<td>Valian</td>
<td>1991</td>
<td>1;9.20–2;8.24</td>
</tr>
<tr>
<td>Garvey</td>
<td>1973</td>
<td>2;10–5;7</td>
<td>VanHouten</td>
<td>1986</td>
<td>2;0-3;0</td>
</tr>
<tr>
<td>Gathercole</td>
<td>1980</td>
<td>2;9–6;6</td>
<td>VanKleeck</td>
<td>NA</td>
<td>3;0</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------</td>
<td>---------</td>
<td>--------</td>
<td>------</td>
<td>-----------------</td>
</tr>
</tbody>
</table>