Evaluating rate and accuracy of real word vs. non-word diadochokinetic productions from childhood to early adulthood in Hebrew speakers

Michal Icht a, Boaz M. Ben-David b,c,d,e,*

a Department of Communication Disorders, Ariel University, Israel
b Baruch Ivcher School of Psychology, Interdisciplinary Center (IDC) Herzliya, Israel
c Department of Speech-Language Pathology, University of Toronto, Toronto, ON, Canada
d Rehabilitation Sciences Institute, University of Toronto, ON, Canada
e Toronto Rehabilitation Institute, University Health Networks (UHN), ON, Canada

ARTICLE INFO

Keywords:
Oral-Diadochokinesis
Oral-DDK Rate
Oral-DDK accuracy
Preschool children
Early adulthood
Non-words

ABSTRACT

Background: Oral-Diadochokinesis (oral-DDK) tasks measure how quickly and accurately one can repeat a series of target sounds. Thus, they are a popular tool for evaluating oral-motor skills for individuals with various speech disorders. Typically, oral-DDK tasks involve rapid repetition of non-words. For several populations (e.g., young children, older adults), it has been suggested that repetitions of real words may be more suitable, commonly resulting in faster rates. Yet, the literature is either silent or inconsistent regarding this real-word repetition advantage for other age groups, from preschoolers to young adults. It is not clear whether performance accuracy is affected as well. Specifically, for Hebrew speakers, this data is missing.

Aims: The goal of this study was to compare rate and accuracy for non-word and real-word repetition, in four groups of Hebrew-speaking individuals; preschoolers (5 years old), younger elementary school children (7 years old), adolescents (15 years old) and young adults (25 years old). Secondary goals were to provide a developmental pattern for oral-DDK rates for Hebrew speakers, and to compare it to the English norms.

Methods & Procedures: All participants (n=150) had typical speech and language development. They were asked to repeat “pataka” (non-word) and “bodeket” (Hebrew real word) as quickly and accurately as possible for 10 sec. Production rates (syllables per second) and accuracy (on a 5-point scale) were measured.

Outcomes & Results: As expected, oral-DDK rates gradually increased with age, with similar rates for both real- and non-words. Accuracy scores were higher for real- than non-word repetition, across all age groups. For the group of school-age children, the Hebrew rates differed from the English ones.

Conclusions & Implications: A real-word repetition advantage was documented only for repetition accuracy, but not for rate. These findings can be explained as each stimulus involves different demands on an individual’s neuro-motor and linguistic processing abilities. Further research using real- and non-word tasks should be conducted with clinical populations to assess whether both procedures could assist in differential diagnosis between various speech disorders. Finally,
the large differences between children of different ages, as well as the apparent rate differences between Hebrew and English, highlight the need to create age- and language-sensitive norms.

1. Introduction

Oral-Diadochokinesis (Oral-DDK) tasks are common diagnostic tools, widely used by speech-language pathologists (SLPs) in clinical contexts as well as for research purposes. As these tasks measure how quickly and accurately one can repeat a given sequence of sounds, they are important in oral-motor assessment (Bernthal et al., 2008; Kent et al., 1987; Williams & Stackhouse, 2000). This task is especially important in pediatric assessment, as early detection of speech-language pathologies allows for early interventions (American Speech-Language-Hearing Association, 2016).

One of the most common oral-DDK stimuli is the tri-syllabic non-word “pataka” (Sequential Motor Rates; Fletcher, 1972; Potter, 2005). Yet, it has been suggested that using a non-word may hamper the performance of several populations (e.g., young children, older adults), providing an underestimation of their oral-motor abilities. Possibly, for these groups, repeating a real word, such as the English word "buttercup", may provide better and more precise evaluation. Such real-word advantage has been previously documented for Hebrew speaking school-age children (9-11 years old; Icht & Ben-David, 2015) and older adults (Ben-David & Icht, 2017). However, other studies failed to find this real-word advantage (e.g., Potter, 2005).

The goal of the current study was to compare the rate and accuracy for non-word and real-word stimuli in four groups of Hebrew speaking individuals, preschoolers (5 years old), younger elementary school children (7 years old), adolescents (15 years old) and young adults (25 years old). This may provide knowledge on the existence of the real-word advantage in oral-DDK in these age groups, or its lack thereof, and offer normative datasets for performance among Hebrew speakers, currently missing in clinical use.

1.1. Using oral-DDK tasks for assessment of children and adolescents

Traditional oral-DDK tasks assess the speed with which an individual can move the articulators (mainly, lips and tongue) in a non-linguistic context. They are designed to reflect oral-motor skills, for example, the ability to access a new (or less familiar) motor sounds, they are important in oral-motor assessment (Bernthal et al., 2008; Kent et al., 1987; Williams & Stackhouse, 2000). This task is especially important in pediatric assessment, as early detection of speech-language pathologies allows for early interventions (American Speech-Language-Hearing Association, 2016).

Oral-DDK tasks are widely used in the assessment of children across the world, forming an essential part of the diagnostic battery in many developmental speech and language disorders (e.g., stuttering and dysarthria). For children with typical development, there is a gradual increase in oral-DDK rates through the childhood years, across different languages (e.g., English: Canning & Rose, 1974; Fletcher, 1972; Mahler, 2012; Brazilian-Portuguese: Modolo et al., 2011; Thai: Prathanee et al., 2003; Hebrew: Icht & Ben-David, 2015). Taking English speakers as an example, the non-word (“pataka”) oral-DDK rate for preschool children is about 3.5 syl/sec. The production rate increases to 4.5 syl/sec, by the age of 6 years (Robbins & Klee, 1987), and adult-like rates (6 syl/sec) are achieved by the age of 9 to 10 years (Canning & Rose, 1974) or by the age of 15 years (Fletcher, 1972).

Table 1 provides a brief summary of oral-DDK performance rates in preschool English-speaking children, as collected from the literature. A quick glance in Table 1 reveals that there is a high degree of variability among these studies of children’s oral-DDK performance (in terms of Ns, measures, and scores, as some were converted from the original reporting), as well as a great degree of variability in the results they obtained. For example, non-word repetition rates for 5-year-olds range from 2.7 to 5.5 syl/ sec. In a review of the oral-DDK literature, Icht and Ben-David (2018b) noted that the high variance in performance rates documented across

<table>
<thead>
<tr>
<th>Study/Age group:</th>
<th>Real word (/pattycake/or /buttercup/)</th>
<th>Non-word (/pataka/)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canning and Rose, 1974</td>
<td>4.34 (.18)</td>
<td>4.06 (.28)</td>
</tr>
<tr>
<td>Williams and Stackhouse, 2000</td>
<td>3.96 (.61)</td>
<td>4.05 (.48)</td>
</tr>
<tr>
<td>Yaruss and Logan, 2002*</td>
<td>4.16 (.75)</td>
<td>3.88 (.66)</td>
</tr>
<tr>
<td>Potter, 2005</td>
<td>3.71 (.49)</td>
<td>2.7 (.33)</td>
</tr>
<tr>
<td>Mahler, 2012</td>
<td>4.52 (.66)</td>
<td>5.04 (.82)</td>
</tr>
<tr>
<td></td>
<td>4.95 (.72)</td>
<td>5.52 (.96)</td>
</tr>
</tbody>
</table>

* The averages represent a mix of non-word and real-word repetition. According to the authors "When a child was either unwilling or unable to comply with this request [repeating “puh-tuh-kuh”], he was asked to produce the word “pattycake” instead" (p.70).
different age groups, may impair the ability to detect atypical performance, for example, performance that varies from the norm by 2.5 standard deviations (SDs). The authors suggested that these deviations across studies may stem from differences in the tasks’ instructions, in the modeling protocol, and in the target stimuli (for further discussion, see: Icht & Ben-David, 2015).

Finally, gender was, at times, found to affect oral-motor skills in children. For example, Smith and Zelaznik (2004) found that boys (until age 5 years) show a slower maturational course of speech motor development. However, in most of the studies which examined oral-DDK rates, no gender-effects were found among children as well as adults (e.g., Fletcher, 1972; Icht & Ben-David, 2014, 2015; Modolo et al., 2011; Robb et al., 1985; Topbas, 2010; Zamani et al., 2017).

Reviewing the pertinent literature reveals a relative lack of data on oral-DDK rates for children in languages other than English. This may be important as oral-DDK rates were found to vary substantially across languages, even in non-word repetition (Icht & Ben-David, 2014). For example, English oral-DDK rates for 9-year-olds were found to be 3.89 syl/sec (Fletcher, 1972), whereas for Brazilian-Portuguese speakers, they reached 5.10 syl/sec (Modolo et al., 2011; see Table 3 in Icht & Ben-David, 2014). Such a wide range necessitates language-specific normative data to provide an accurate diagnosis. The current study aims to provide such normative data for Hebrew speakers in four age groups, from preschoolers to young adults.

1.2. Real vs. non-word repetition

Several researchers questioned the compatibility of using non-word repetition for the assessment of young children (Canning & Rose, 1974; Cohen et al., 1998). For example, Yaruss and Logan (2002) suggested that some preschool age children have difficulties complying with the instructions for non-word repetition, and some even fail to complete the task. Icht and Ben-David (2014) suggested that the abstract non-familiar nature of nonsense syllable sequences might be challenging for young children, as they may be preoccupied by trying to decipher the linguistic meaning of the “funny” words. In order to make this task simpler and more accessible for children, the use of real words has been suggested.

In contrast, other researchers suggested avoiding real-word repetition in oral-DDK tasks, as it involves different demands on an individual’s neuro-motor and linguistic processing than non-word repetition (Williams & Stackhouse, 1998). The inclusion of real words adds a linguistic dimension to the task (Wilcox et al., 1996) accessing the ability to access a stored motor program through linguistic cues, while non-word repetition assesses a non-linguistic (“pure”) motor ability. Attempting to reconcile these approaches, it has been recommended to use a dual-protocol, testing both real and non-words in the oral-DDK battery (Canning & Rose, 1974; Robbins & Klee, 1987) to allow a wider and more accurate evaluation.

Most of the studies that examined children’s performance with real-word stimuli found faster oral-DDK rates relative to non-word stimuli (Canning & Rose, 1974; Icht & Ben-David, 2015; Robbins & Klee, 1987; Yaruss & Logan, 2002). Williams and Stackhouse (1998; 2000) suggested that this real-word advantage becomes clearer as the child reaches school age. A similar pattern, namely a real-word advantage, has been also found for older adults (in Hebrew; Ben-David & Icht, 2017). In this study, older adults (aged 66–95 years) repeated real words faster than non-words by 13.5 percent. Contrary to these findings, Potter (2005) found a reversed pattern, with higher rates for non-words than for the real word “pattycake” with preschool children. The results of these studies support using a dual-protocol (real- and non-word repetition) to improve differential diagnosis of normal and pathological performance in this task.

1.3. Oral-DDK accuracy

The majority of above-mentioned studies examined repetition rate, whereas repetition accuracy has been mostly neglected. Since many individuals with speech disorders exhibit imprecise articulation (Bernthal et al., 2008), it has been suggested that the accuracy of oral-DDK productions may provide an important supplement to traditional measures of rate (Yaruss & Logan, 2002). Indeed, children with typical development were found to produce errors during standard DDK tasks (Canning & Rose, 1974; Henry, 1990; Williams & Stackhouse, 2000), similar to errors that typically occur in conversational speech. Analyzing the frequency and types of these errors may provide SLPs with valuable information about children’s oral-motor development. The impact these errors might have on oral-DDK rate, may also carry clinical importance. As noted by Williams and Stackhouse (2000), production accuracy is important for rehearsing new vocabulary and when spelling complex words. Accuracy of production may be also related to sequencing abilities (particularly in poly-syllabic words), which may be impaired in children with persisting speech difficulties (e.g., residual speech errors; Stackhouse, 1992).

In pediatric assessment, production accuracy is significant as it may be more closely related to oral-motor maturation than rate. For example, in a group of children with typical development, aged 3- to 5-year-old, Williams and Stackhouse (2000) reported a general increase in accuracy with age across different stimuli (real word, non-word and syllable sequences). That is, overall performance of the 3-year-olds was significantly poorer than that of the 4- and 5-year olds (see also Canning & Rose, 1974; Fletcher, 1972; Henry, 1990).

The scoring procedure for repetition accuracy, unlike repetition rate, is less clear. Studies that report accuracy use different scoring measures, not always fully described. The variety of methods used across studies make it difficult to know what accuracy measure to use applying the norms to clinical practice. Williams and Stackhouse (2000) scored the accuracy of both one repetition (first attempt) and five repetitions of the target sequence and compared them to an adult model. Accuracy of first repetition of the target increased with age, with the 3-year-olds performing worse than the other two age groups. Likewise, accuracy of five repetitions of the target also increased with age. It is noteworthy that for the 3 and 4-year-olds, real word repetition was significantly better than non-word repetition; yet, this difference was not evident for the 5-year old children. A different accuracy scoring procedure was used by Yaruss and Logan (2002). They scored the number and types of articulation errors produced during all DDK trials and found no correlations between age and measures of DDK accuracy (i.e., the frequency of articulation errors or speech disfluencies).
In the current study, we suggest a relatively-simple alternative scoring method for oral-DDK accuracy that aims to encompass the different approaches employed by other studies in the pertinent literature. The proposed scoring method is based on counting articulation errors and converting the number of errors to an accuracy score. This results in a 5-point Likert rating scale based on a numerically and qualitatively defined classification, to be detailed in the Method section. The suggested scale takes into account that both articulation rate and accuracy increase with maturation. Such a procedure echoes the standard method for identifying speech errors in other tasks (e.g., a standard picture-naming task) routinely employed in the evaluation of children’s speech abilities (Yaruss & Logan, 2002). As posited by Milloy (1985), this approach reflects both consistency and accuracy.

1.4. The present study

The goal of the study was to examine oral-motor development for Hebrew speakers, comparing non-word and real-word oral-DDK rates and accuracy for four groups of Hebrew speaking individuals, preschoolers, younger elementary school children, adolescents, and young adults. These comparisons are important given the inconsistent findings in the literature on a real-word advantage, and the lack of data on accuracy in real- vs. non-word repetition. As no performance norms for oral-DDK rates and accuracy are available for Hebrew speaking preschool children and adolescents, we aimed to provide such preliminary baseline data, essential for accurate clinical diagnosis and evaluation. We also aimed to examine the maturation pattern of oral-DDK rates, by comparing the preliminary norms obtained in the current study with previously reported rates of Hebrew-speaking school-age children. Finally, we compared the Hebrew preliminary norms to English speakers’ data, as collected from the literature, to examine possible language-specific differences.

2. Method

The study received prior approval from the ethics board of Ariel University. Written informed consent was obtained from all participants or participants’ parents and all participants assented orally.

2.1. Participants

The current investigation focused on four groups of Hebrew speaking individuals with typical speech (articulation) abilities - - preschoolers (5 years old), younger elementary school children (7 years old), adolescents (15 years old), and young adults (25 years old). Inclusion criteria were as follows: (a) native Hebrew speakers, (b) no history of speech or hearing difficulties and/or neurological disease, and (c) no abnormal oral structure/ function. Participants were recruited by publishing a Call for Participation in various parents’ groups via social media (e.g., Facebook) and on campus (Ariel University).

One hundred and sixty-two potential participants (or the parents of the younger participants) who responded were invited to an individual screening session. A research assistant (trained SLP student) provided an explanation of the research and its goals and made sure these candidates met the inclusion criteria by an interview. The aforementioned inclusion criteria were reported by the parents of the younger participants, or by self-report for the young adults by filling a written questionnaire (taken from Icht & Ben-David, 2018a). Twelve of the candidates were dismissed after the screening phase (six were bilinguals, and six had non-age-appropriate inter-dental lisp errors; Bat-El, 2009). The remaining one-hundred and fifty individuals participated in the experiment; demographic data is presented in Table 2.

The sample included individuals from several medium sized cities in the center of Israel, with a comparable socioeconomic status (high) and cultural background. These factors were measured by reporting family income, parental education, and occupational status. Typical speech (articulation) abilities were verified by two research assistants using a common articulation and naming test, as will be described next.

2.2. Assessments

Articulation test. To ensure typical phonological and articulatory development and abilities, all participants performed a common Hebrew articulation and naming test (Rosin & Yakir, 2000). This unstandardized test comprises 32 pictures of familiar objects, mono- to four-syllable words. All consonants appear in initial-, medial-, and final-positions. All vowels are also represented in full. Participants were required to name the pictures. In case a participant failed to name the picture (less than 5 percent), the research assistant named it and asked the participant to repeat the word.

Table 2

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>n</th>
<th>Gender (m, f)</th>
<th>Age (Mean, SD)</th>
<th>Age Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschoolers</td>
<td>35</td>
<td>13, 22</td>
<td>5:00 (0:01)</td>
<td>4:10-5:03</td>
</tr>
<tr>
<td>Younger elementary</td>
<td>35</td>
<td>15, 20</td>
<td>6:11 (0:02)</td>
<td>6:09-7:03</td>
</tr>
<tr>
<td>Adolescents</td>
<td>40</td>
<td>20, 20</td>
<td>15:07 (0:05)</td>
<td>15:01-16:05</td>
</tr>
<tr>
<td>Young adults</td>
<td>40</td>
<td>20, 20</td>
<td>24:06 (1:10)</td>
<td>21-29</td>
</tr>
</tbody>
</table>

M. Icht and B.M. Ben-David
coded the participants’ responses, to ensure inter-rater reliability). In the few cases when the two did not reach an agreement (transcription disagreement that occurred between the two listeners; less than 5%), they consulted with the first author (an experienced SLP) to reach a decision. Eight children from the 5-year-old group and two children from the 7-year-old group were identified with inter-dental lisping errors (age appropriate in Hebrew; Bat-El, 2009).

**Oral-DDK tasks.** Participants performed two oral-DDK tasks, one using a non-word and one using a real word, in a counterbalanced order. For non-word, we used the tri-syllabic sequence “pataka”, a widely used sequence, for which normative data for school age children and for adults is available in Hebrew (Icht & Ben-David, 2014; 2015). For real word, we used the word “bodeket” (the female rendition of the noun examiner or the verb inspecting in Hebrew). This stimulus has been offered and validated in previous studies (Ben-David & Icht, 2017; Icht & Ben-David, 2015). To ensure the familiarity of this word for the younger age groups, prior to administering the real-word task, the 5- and 7-year-olds were shown a colored drawing of a female doctor examining a child using a statoscope. The research assistant briefly described the picture, repeating the target word twice.

Oral-DDK repetitions were audio-recorded for each participant, using a high-quality digital recorder (Olympus Europa SE & Co. KG; VN-8500PC). It was held about 20 cm from the participant’s mouth and slightly tilted toward the speaker (following Icht & Ben-David, 2018a).

### 2.3. Procedure

The participants were tested individually in a quiet room at their home (or on campus, for the group of young adults). Before arrival, the participants or their parents, were informed on the study and its goals, filled out the demographics questionnaire, and signed a consent form. At the beginning of the experimental session, the articulation test was administered. Then, each participant received a short explanation regarding the oral-DDK task. First, the research assistants demonstrated the task (real- or non-word repetition) once to each participant, and the participant was asked to practice the task once. The younger participants were shown an appropriate picture to ensure that the real word was familiar, and that they had a lexical representation for this item. Then, the participant was requested to repeat the tri-syllabic stimulus, as quickly and accurately as possible, stating: “Now, please repeat this word [or ‘funny word’] once again, as quickly as possible, without making a mistake, until I will signal you to stop”, for 10 sec, using a count-by-time method (see, Icht & Ben-David, 2014). To encourage the participation of the younger participants, the research assistants used a variety of non-verbal motivating techniques, such as nodding and gesturing. All participants successfully completed both tasks. The order of tasks was counterbalanced across participants. The whole session lasted about 10 minutes.

### 2.4. Scoring procedure and data analysis

**Oral-DDK rates.** Oral-DDK rate (syl/sec) was calculated by multiplying the total number of tri-syllabic sequences (“pataka” or “bodeket”) produced by each participant in 10 sec, by 0.3. In case that a sequence was only partially completed by the time 10 sec elapsed, it was excluded. The syllable count was verified using PRAAT software (Boersma & Weenink, 2017), by listening to the recordings and by visual examination of the output by the two research assistants, following the procedure described in Icht and Ben-David (2018b).

**Oral-DDK accuracy.** The suggested accuracy scoring method follows Yaruss and Logan (2002) and uses phonetic transcriptions of all the consonants produced when repeating real or non-words in a given time (e.g., 10 sec in the current study). The suggested scale considers that with maturation performance improves: the number of repetitions increases, and the number (and percentage) of errors decreases. It is based on counting the total number of errors produced and converting the number to a score on a 5-point Likert scale. Scores range from 0, denoting that all productions in the participant’s sample were accurate and non-consistent (or an inability to compete the task) to 4, denoting accurate and consistent productions of all targets in the sample. Scores of 1, 2 and 3 were provided by counting the number of errors. A score of 3 indicates good accuracy and consistency level, with three or fewer articulation errors. We suggest that one, two or three errors do not clinically differ, representing fewer than 10% of the mean value of typical performance for the youngest age group tested (preschoolers; Mean of 37 syllables) and fewer than 5% of the mean value of typical performance for the oldest age group tested (young adults; Mean of 71 syllables). On the other extreme, a score of 1 indicates poor accuracy and consistency levels, with at least seven articulation errors. This number was chosen, as it represents about 20% errors for the youngest age group tested (preschoolers) and about 10% errors for the oldest age group tested (young adults). A score of 2 indicates Medium accuracy and consistency levels with number of errors ranging between these two boundaries (four, five or six errors). As aforementioned, as age increases, speech rate and accuracy increase. Hence, we suggest that 10% errors represent good accuracy for preschoolers (receiving the score of 1), but poor accuracy for young adults (score of 3).

Six types of articulation errors are coded: (1) insertions or (2) deletions of sounds (e.g., production of incomplete sequences), (3) changes in voicing or (4) placement, (5) exchanges between sounds, and (6) perseveration of sounds. The occurrence of such errors may also point to production inconsistencies, that is, variations of the target sequences along the entire trial. Following Crary (1993), an error was coded when all target syllables were present but in wrong order. As recommended by Yaruss and Logan (2002), errors did not include systematic errors (phonological processes), but rather non-systematic errors that occasionally occur in spontaneous speech (note, individuals with such phonological processes were not included in the current sample). Misarticulation of vowels was not considered as an error.

To define production accuracy, two research assistants, trained SLP students, separately listened to the recordings and phonetically transcribed the consonants produced by the children during their oral-DDK productions (for a similar procedure, see Yaruss & Logan, 2002). The number of errors involving the aforementioned deviations from the target consonants (/p, t, k/ or /b, d, k/) was counted,
enabling them to assign each participant an accuracy score (for a comparable intelligibility scale, see Icht, 2019). For a full description of scoring criteria and some examples, refer to Table 3a. If a disagreement between the research assistants was identified (less than 5 percent of the recordings), the first author (an experienced SLP) reached a decision.

2.5. Statistical analyses

Repeated-measures mixed-model $2 \times 4 \times 2$ ANOVAs were conducted separately for production accuracy and for rate (as dependent variables), with stimulus type (real word vs. non-word) as a within-participant variable, and age group (preschool, younger elementary school, adolescents, and young adult groups) and gender as between-participants variables. For accuracy, pairwise planned comparisons followed this ANOVA, to directly compare the accuracy scores, averaged across real word and non-word repetitions, between all age groups. For production rate, the ANOVA was followed by a Pearson’s correlation analysis between non-word and real word.

To detect possible speed-accuracy tradeoff, a univariate ANOVA of repetition rate (dependent variable) was conducted, with age group (x4: preschool, younger elementary school, adolescents, and young adult groups) and accuracy perfection (x2: flawless vs. flawed production) as between participants variables. Finally, additional data taken from previous studies was compared to the current study, in a series of independent sample t-tests. Partial eta squared ($\eta_p^2$) was used as an effect size measure in all statistically significant tests.

3. Results

Table 4 presents oral-DDK accuracy scores (means and SDs) and rates (syl/sec) for real- and non-word repetition, across the different age groups. A quick glance at Table 4 reveals higher accuracy scores for real word repetition across all age groups, with no effect for age. Repetition rates, however, apparently did not differ by the types of stimuli, real- and non-words, and appeared to be higher as age increases.

3.1. Production accuracy

As a first step we analyzed production accuracy, using a repeated-measures mixed-model ANOVA with stimulus type (real-, non-word) as a within-participant variable, and age group and gender as between-participants variables. Analysis showed that across groups, accuracy was higher for real word than non-word repetition, $F(1, 142) = 30.21, p < .01, \eta_p^2 = .18$, with mean accuracy scores of

| Table 3a |
| Types and examples of articulation errors (based on: Yaruss & Logan, 2002, Table 1) |
| Types of articulation error | Examples |
| Insertions of sounds - an additional consonant or syllable is inserted | “pta-ta-ka”, “bok-de-ket” |
| Deletions of sounds - a consonant or syllable is omitted | “pa-ta”, “bo-de-et” |
| Changes in voicing - a voiceless consonant is produced instead of a voiced consonant, or vice versa | “ba-ta-ka”, “bo-te-ket” |
| Changes in placement - a consonant is produced with an incorrect place of articulation | “pa-ta-ta”, “do-de-ket” |
| Exchanges between sounds - two syllables within the same sequence are exchanged | “pa-ka-ta” |
| Perseveration of sounds - a produced syllable is repeated | “pa-pa-ka” |

| Table 3b |
| Oral-DDK accuracy scoring |
| Accuracy score | Description |
| 0 | All productions in the participant’s sample were inaccurate and non-consistent, or an inability to complete the task |
| 1 | Poor accuracy and consistency, articulation errors $\geq 7$ |
| 2 | Medium accuracy and consistency, $4 \leq$ articulation errors $\leq 6$ |
| 3 | Good accuracy and consistency, articulation errors $\leq 3$ |
| 4 | Accurate and consistent productions of all targets in the sample, flawless performance |

Table 4

Oral-DDK accuracy scores (Means, and SDs in parenthesis) and rates (syl/sec) for real- and non-word repetition across the different age groups

| Age group | n | Oral-DDK accuracy (from 0 to 4) | Oral-DDK rates (syl/sec) |
| Preschoolers (5 years old) | 35 | 3.28 (.72) | 3.66 (.53) |
| Elementary school (7 years old) | 35 | 3.35 (.71) | 3.61 (.52) |
| Adolescents (15 years old) | 40 | 3.20 (.70) | 3.70 (.51) |
| Young adults (25 years old) | 40 | 3.48 (.70) | 3.87 (.51) |
Table 5
Planned comparisons of oral-DDK accuracy scores and rates between the different age groups. The asterisk (*) represents a significant difference at \( p < .05 \).

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Oral-DDK accuracy</th>
<th></th>
<th></th>
<th>Oral-DDK rates</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate (SE)</td>
<td>Estimate (SE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preschoolers - Younger elementary school</td>
<td>-.01 (.11)</td>
<td>.96</td>
<td>-.46 (.16)</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>Preschoolers - Adolescents</td>
<td>.02 (.11)</td>
<td>.82</td>
<td>-.81 (.15)</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>Younger elementary school - Adolescents</td>
<td>.03 (.10)</td>
<td>.78</td>
<td>-.35 (.15)</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>Young adults - Preschoolers *</td>
<td>-.20 (.11)</td>
<td>.06</td>
<td>-.32 (.15)</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>Young adults - Younger elementary school *</td>
<td>-.19 (.10)</td>
<td>.06</td>
<td>-.26 (.15)</td>
<td>&lt;.01</td>
<td></td>
</tr>
<tr>
<td>Young adults - Adolescents *</td>
<td>-.23 (.10)</td>
<td>.02</td>
<td>-.51 (.15)</td>
<td>&lt;.01</td>
<td></td>
</tr>
</tbody>
</table>

3.71 vs. 3.28, respectively. We did not find a main effect for age group, \( F(3, 142) = 2.13, p = .10 \), and the two effects did not interact, \( F(3, 142) = .51, p = .68 \). Indeed, as Table 4 shows, the advantage in accuracy for real words over non-words was consistent across all age groups (about .4 on a 0-4 scale). We also found a main effect for gender, with higher accuracy scores for females, 3.60 vs. 3.44, \( p = .68 \). Indeed, as Table 4 shows, the advantage in accuracy for real words over non-words was consistent across all age groups (about .4 on a 0-4 scale). We also found a main effect for gender, with higher accuracy scores for females, 3.60 vs. 3.44, \( p = .68 \). However, gender did not interact with the stimulus type, \( F(1, 142) = 1.43, p = .23 \), or with age group, \( F(3, 142) = 1.55, p = .21 \).

Pairwise planned comparisons followed the ANOVA, to directly compare the accuracy scores, averaged across real word and non-word repetitions, between all age groups. The analyses revealed that the group of young adults was more accurate across all age groups (at least marginally significant differences). These groups of preschoolers, younger elementary school children, and adolescents did not significantly differ from each other (see Table 5 for detailed data).

3.2. Oral-DDK rate

Next, we analyzed production rates, using a repeated-measures mixed-model ANOVA with stimulus type (real-, non-word) as within-participant variable, and age group and gender as between-participants variables. We did not find a significant main effect for the stimulus type, \( F(1, 142) = .03, p = .88 \), nor for gender, \( F(1, 142) = .29, p = .60 \), and the two did not significantly interact, \( F(1, 142) = .16, p = .69 \).

As expected, we found a main effect for age group, \( F(3, 142) = 238.15, p < .01 \), \( \eta^2_p = .83 \), indicating higher repetition rates as age increased. Pairwise planned comparisons followed this ANOVA, to directly compare repetition rates between the different age groups. These analyses revealed that oral-DDK rates differed significantly between all age groups, as detailed in Table 5. However, age group did not interact with the other factors, stimulus type, \( F(3, 142) = .37, p = .77 \), or gender, \( F(3, 142) = 1.86, p = .14 \). Indeed, in a follow-up analysis conducted in each age group separately, we did not find a significant difference between real word and non-word oral-DDK rates in either group, \( t < .8, p > .45 \) for all.

Interestingly, non-word repetition rates and real word repetition rates significantly correlated across age groups (using Pearson’s correlation coefficient), \( r(150) = .90, p < .01 \), supporting the criterion validity of the real word task. In a separate post-hoc analysis, this correlation was found to be significant in three of the four age groups: preschoolers: \( r(35) = .52, p < .01 \); younger elementary school: \( r(35) = .25, p = .14 \); adolescents: \( r(40) = .57, p < .01 \); young adults: \( r(40) = .61, p < .01 \).

3.3. Possible speed-accuracy tradeoff

A careful analysis of the accuracy data revealed that the great majority of participants (98 percent) scored 3 or 4 on the real word task (none scored 0, only a single participant, from the preschool group, scored 1; two scored 2: one preschooler and one adolescent). Although accuracy scores were lower for the non-word task, a similar trend was noted, and about 87 percent of participants scored 3 or 4 (a single participant, from the preschool group, scored 1; 19 participants scored 2: five preschoolers, five young elementary schoolers, seven adolescents and two young adults). Consequently, we created a new accuracy perfection variable, flawless (accuracy score = 4) vs. flawed (accuracy score < 4) production, reflecting a higher percentage of flawless real- than non-word production, across age groups (74% vs. 47.3%, respectively; \( p < .01 \), related sample McNemar test). Note, such separation between participants producing error-free iterations versus those producing more errors was already suggested by Yaruss and Logan (2002).

In order to detect possible speed-accuracy tradeoff, we conducted a univariate ANOVA with age group and accuracy perfection (flawless vs. flawed production) as between participants variables, and repetition rate as a dependent variable. Non-word accuracy perfection was not found to affect non-word repetition rate, \( F(1, 142) = .08, p = .78 \), and it also did not interact with age group, \( F(3, 142) = .88, p = .45 \). Similarly, real word accuracy perfection did not affect real word repetition rate, \( F(1, 142) = .06, p = .81 \), and it did not interact with age group, \( F(3, 142) = .96, p = .42 \). In sum, no speed-accuracy tradeoffs were found across age groups and for both types of stimuli.

3.4. A developmental course for Hebrew speakers

A secondary goal of this study was to provide a developmental course of repetition rates for Hebrew speakers, from preschool to adulthood. To that end, we conducted another analysis of oral-DDK rates, averaged across real- and non-word repetitions, adding previously published rates for elementary school age children (60 participants aged 9-11 years; Icht & Ben-David, 2015) that used the
same protocol. A series of independent sample t-tests revealed that this additional group of elementary school age children (9-11-years-old) significantly differed in mean repetition rates from all the groups tested in the current study, \( p < .01 \) for all comparisons: \( t(93) = 7.57, t(93) = 4.55, t(98) = 11.57 \) and \( t(98) = 13.84 \) for a comparison with preschool, young elementary school, adolescent and young adult groups, respectively. A visual depiction of the developmental course of repetition rates for Hebrew speaking children and adolescents is provided in Figure 1.

3.5. A cross-language comparison: Hebrew and English oral-DDK rates

Table 6 presents average non-word repetition rates for Hebrew speakers as obtained in our lab, in comparison to normative data obtained for English speakers from comparable age groups. Table 6 also presents sets of two-sample unpaired t-tests, directly comparing norms between the two languages. No significant differences were found for the groups of preschoolers, adolescents, and young adults. However, for the two groups of school-age children (7- and 9-11 years old), repetition rates significantly differed (see Table 6). Clearly, differences in several participant- and experimental-related factors (Ns, age, study design) may lead to variability between studies; significant effects, or their lack-there-of, should be taken with caution.

4. Discussion

The goal of the current study was to provide a wide picture on oral-DDK performance maturation for Hebrew speakers. We examined Hebrew speakers with typical development of four age groups, preschoolers (5 years old), younger elementary school children (7 years old), adolescents (15 years old), and young adults (25 years old), and compared their performance between a non-word and a real word oral-DDK task, in terms of rate and accuracy. Results showed that accuracy scores were higher for real word repetition relative to non-words across all age groups. Oral-DDK rates increased with age and were similar for real- and for non-words. Comparing our data with previously obtained data for 9-11-year-old Hebrew speakers indicated a clear maturation pattern in repetition rates for individuals studied. Finally, it appears that Hebrew preliminary rate norms do not necessarily match English norms, especially for elementary school children.

4.1. Oral-DDK accuracy

Assessing the accuracy of the oral-DDK productions may be a meaningful measure, as speech accuracy is a fundamental aspect of speech production commonly used in clinical evaluation (Yaruss & Logan, 2002). Since children with typical development, as well as individuals of clinical populations (e.g., adults with Parkinson’s disease; Ackermann, Hertrich, & Hehr, 1995), produce errors during DDK tasks, an accuracy measure may be an important addition to the traditional measure of oral-DDK rate. We suggest using a relatively simple scoring procedure, based on identifying and counting articulation errors, a practice routinely performed by SLPs in other clinical tasks, such as articulation test or analysis of a conversational speech sample (Yaruss & Logan, 2002). Since this procedure is familiar to SLPs, it can be successfully applied in daily clinical practice. Following the recommendations of Yaruss and Logan (2002), we defined specific scoring criteria (on a 5-point scale) derived from the most common production errors in terms of articulation accuracy and consistency (see Table 3b).
Using the suggested scoring system, real word repetitions were more accurate than non-word repetitions across all age groups, both in average score and in the percentage of flawless productions. These results suggest that the non-word task may be more difficult than the real word task. On the one hand, as more errors occurred with non-word repetition, analyzing their type may provide important information regarding specific articulation errors. On the other hand, the real word task may better reflect conversational speech. These results highlight the potential importance of a dual-protocol, assessing both real- and non-word repetition.

This pattern of results, higher accuracy scores for real- than non-words, may serve as a benchmark for typical performance. Deviations from this performance pattern (e.g., similarly low accuracy scores for both real- and non-words, or a mixed pattern with higher accuracy in non-word than real word task) may hint at possible oral-motor or speech difficulties. Hence, a dual protocol may be of special importance in clinical evaluations of individuals with speech and language disorders, as a diagnostic measure and even as a prognostic indicator. Clearly, further studies focusing on special populations are called for to confirm this possibility.

Interestingly, in the current study no significant speed-accuracy trade-offs were documented. Specifically, the rate of flawless repetitions was not affected by age, even though the increased number of repetitions with age also increased the chance for an articulation error. These findings highlight the improved performance with maturation and stress the importance of adding the accuracy score as a separate measure of the repetition rate, providing clinicians with useful information on speaking abilities.

The apparent separation between these measures, raises the possibility of identifying subgroups of participants of different functional profiles, for example, participants producing fewer errors in high rates, vs. those producing more errors with slower rates. Further research will be necessary to investigate this possibility and to examine the speech characteristics of each subgroup.

Finally, we were not surprised to find the group of young adults to be significantly more accurate than all younger age-groups (which did not differ in terms of repetition accuracy). As speech production abilities continue to develop well into adolescence, some aspects of the speech mechanism are not fully matured until late adolescence (Cheng et al., 2007). This seems to be the case for oral-DDK accuracy as well, for both types of stimuli.

4.2. Oral-DDK rates

Comparing performance rates (syl/sec) between the different age groups revealed the expected gradual increase in rate with aging, as the slowest group was the preschool children (with a mean rate of 3.78 syl/sec), followed by the group of school age children (4.24 syl/sec), that was followed by the adolescents (6.58 syl/sec). The group of young adults showed the fastest rates (7.09 syl/sec). This trend echoes previous reports in the literature regarding the gradual increase in oral-DDK rates as age increases and the speech production system matures. Clearly, as children develop, they gradually acquire new and advanced oral-motor and speech abilities. The gradual maturation of these capabilities is reflected by the increase in speed of performance (Fletcher, 1972; Kent & Former, 1987).

The present study provides oral-DDK rate datasets for Hebrew speaking children and adolescents with typical development and typical speech abilities, currently missing, that can be reliably used by clinicians. Deviations from these age-expected norms may hint at atypical performance that merits further evaluation. Note, the current oral-DDK rates for young adults are in line with those obtained in a previous study of Hebrew speakers (age range: 20-45 years; Icht & Ben-David, 2014), strengthening the external validity of the present study and its findings.

Interestingly, no real word advantage was found for repetition rates in all the age groups tested in the current study. This finding is in contrast with some of the literature (for preschool children in English, Williams & Stackhouse, 1998; in Persian, Zamani, Rezai, & Garmatani, 2017), but is in line with others (e.g., Potter, 2005). As mentioned by Arthur (2017), several developmental factors influence an individual’s non-word repetition rates, including vocabulary, auditory processing, articulatory output processes, and literacy development. In future studies, researchers may wish to further test this, as possibly related to vocabulary as well (see, Edwards, Beckman, & Munson, 2004).

<table>
<thead>
<tr>
<th>Age group</th>
<th>Hebrew</th>
<th>English</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschoolers (5 years old)</td>
<td>3.78 (.81)</td>
<td>4.14 (1.05)‡</td>
<td>t(53) = 0.48, p = .63</td>
</tr>
<tr>
<td>Younger elementary school (7 years old)</td>
<td>4.22 (.79)</td>
<td>5.16 (.57)‡</td>
<td>t(53) = 2.21, p = .03</td>
</tr>
<tr>
<td>School-age children (9-11 years old)</td>
<td>4.55 (1.03)‡</td>
<td>4.23 (.23)‡</td>
<td>t(106) = 1.97, p = .05</td>
</tr>
<tr>
<td>Adolescents (15 years old)</td>
<td>6.54 (.78)</td>
<td>6.25 (.21)‡</td>
<td>t(43) = 1.5, p = .14</td>
</tr>
<tr>
<td>Young adults (25 years old)</td>
<td>7.13 (.78)</td>
<td>6.55 (.94)‡</td>
<td>t(62) = 0.39, p = .39</td>
</tr>
</tbody>
</table>

Norms taken from:
1 Robbins and Klee, 1987
2 Fletcher, 1972
3 Icht & Ben-David, 2015
4 Robb et al., 1985
5 Topbas, 2010.
4.3. Clinical implications

Comparing Hebrew and English norms. As Table 6 shows, average non-word repetition rates obtained in Hebrew and in English were overall similar. However, for elementary school children aged 7-11 years, significant differences in performance were found. These findings suggest that oral-DDK rates appear to be impacted by differences in language (or culture), thus one universal oral-DDK norm (e.g., for USA English speakers) cannot be a-priori used across languages and cultures (for a discussion, see Icht & Ben-David, 2014). Therefore, it is important to set language-specific norms that can meet the needs of the clinical practice.

The developmental course of oral-DDK rates for Hebrew speakers. In order to provide a developmental pattern for oral-DDK rates in Hebrew speaking children and adolescents, we added to the current results, the findings of a previous study conducted among 9-11-year-olds (Icht & Ben-David, 2015). These results are presented in Figure 1 and give a clear picture of a gradual increase in oral-DDK rates along the childhood and adolescence years. This pattern echoes the literature indicating that the time course of development for speech motor coordination is protracted (Smith & Zelaznik, 2004).

A real word advantage in repetition rates was only found for the additional group of elementary school age children, 9-11 years old. Recall that cross-language differences in norms were also obtained in this age group, whereas for the groups of adolescents and young adults, we noticed no language- or stimulus-type-effects. Clinically, this suggests the need for taking extra-caution in testing elementary school children, where there is a special need to use language-sensitive norms, as well as to consider using both real- and non-word repetition tasks.

We were not surprised to observe unique performance in elementary school age children, since differences between the maturation pattern of various motor tasks are noted in this age-range (Walsh & Smith, 2002). It has been suggested that during development, varying trade-offs between several motor aspects of speech production (e.g., speech movement amplitude, velocity of articulatory movement) occur, driven by the goal of increasing speech rate (Smith & Zelaznik, 2004; Walsh & Smith, 2002). Specifically, the 7 to 12 years old period was found to be characterized by a unique pattern in the development of coordinative synergies for speech production. In this period, children show no evidence for increased stability in oral-motor coordination for speech (Smith & Zelaznik, 2004). A similar effect specific for this age-group was identified in other speech sub-systems, for example, oral-motor reflexes (jaw-closing and jaw-stretching systems), where 7-8 years seems to be a transitional period, with large responses relative to 4-6 years and to adults (Smith et al., 1991; Wood & Smith, 1992). Finally, in the current study, only the group of 7-year-olds did not show the correlation between real- and non-word repetition rates obtained in all other tested age groups. Taken together, it is reasonable that specifically in this age range of elementary school years, the maturation pattern of the non-word oral-DDK task differs from that of the real word task, supporting a possible clinical benefit of using both tasks.

4.4. Limitations and future recommendations

A limitation of this study is that the real word “bodeket” does not share the exact same phonological structure and features with the non-word “pataka”. This may have an impact on performance. However, the differences are minute and similar in scope to those found with the commonly used English real words for oral-DDK testing (“buttercup”, “pattycake”), and this word was found to be appropriate in previous studies (Ben-David & Icht, 2017; Icht & Ben-David, 2015). In addition, the data regarding school age children (9-11 years-old) was collected in a different study, by different research assistants. Although the assessment protocol was identical, this by itself may have an impact on performance, and warrants replication in a future study.

Since the current study focused on testing children and young adults with typical development, further evaluation of the dual protocol with special populations (e.g., children with Childhood Apraxia of Speech) can provide significant clinical information on this test. Focusing on special populations, it may be interesting to compare the type of articulation errors or speech disfluencies between real- and non-word repetitions. Future studies are needed to further validate the current accuracy scoring procedure, as this is the first study to apply it. Evaluating physiological factors (e.g., measuring jaw opening) or adding an acoustical analysis of the participants’ productions can also be clinically important. Given social restrictions due to COVID-19 pandemic, we also recommend adapting the oral-DDK task to telehealth (remote assessment, see Ben-David et al., 2020) in an effort to offer access to necessary health care.

In sum, we suggest that further research should be conducted with clinical populations (children as well as adolescents), assessing both rate and accuracy using a dual assessment procedure, which includes repetition of both types of stimuli, non- and real words. Possibly, such a protocol may be significant in the differential diagnosis of various speech disorders, providing the clinician a wide profile of oral-motor abilities, and allowing a comparison of neuromotor skills (non-word repetition) with linguistic skills (real word repetition) assessing both articulation errors and speed. This approach can provide diagnostically relevant data.

Author Statement

Both authors contributed to the research equally on all aspects (methodology, formal analysis, data curation, writing and project administration).

Declarations

Ethics approval and consent to participate. This study was approved by the Institutional Ethics Committee, Ariel University. Potential participants or their parents received written and oral information on the study, after which their written consent was obtained. All participants gave their oral consent.
Source of Funding. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no competing interests.

References


Journal of Psycholinguistic Research, 46, 897–904.