International Journal of Speech-Language Pathology

ISSN: 1754-9507 (Print) 1754-9515 (Online) Journal homepage: https://www.tandfonline.com/loi/iasl20

The development of a speech sound screening test for European French-speaking monolingual and bilingual children

Margaret Kehoe, Nathalie Niederberger & Anne-Laure Bouchut

To cite this article: Margaret Kehoe, Nathalie Niederberger & Anne-Laure Bouchut (2020): The development of a speech sound screening test for European French-speaking monolingual and bilingual children, International Journal of Speech-Language Pathology, DOI: 10.1080/17549507.2020.1750699

To link to this article: https://doi.org/10.1080/17549507.2020.1750699

International Journal of Speech-Language Pathology, 2020; Early Online: 1–10

Taylor & Francis
Taylor & Francis Group
Check for updates

The development of a speech sound screening test for European French-speaking monolingual and bilingual children

MARGARET KEHOE, NATHALIE NIEDERBERGER & ANNE-LAURE BOUCHUT

Faculté de psychologie et des sciences de l'éducation, Université de Genève, 28, bd du Pont-d'Arve, Geneva, Genève

Abstract

Purpose: This study presents normative data on a screening test, referred to as the *Dépistage Rapide Articulation et Phonologie (DRAP)*, designed to detect European French-speaking children who are at risk for having speech sound disorders.
Method: The test consists of 20 words which contain late-acquired phonological features such as medial and final /r/, consonant clusters, /s/C sequences, and alveolar and post-alveolar fricatives. The test was given to 196 children, monolingual and bilingual, living in Geneva and San Francisco, and ranging in age from 2.11 through to 6.11. Our analyses examined the influence of bilingualism, context (Geneva or San Francisco), gender and age on the test scores and also looked at the influence of these factors on phonological features in the test.
Result: There were no strong effects of bilingualism, context, and gender on the test results but strong effects of age. Validity and internal consistency of the test were in the acceptable range. A focus on phonological features indicated that children had difficulty with final clusters and post-alveolar fricatives.
Conclusion: The study provides encouraging results for the use of this test as a screening measure with French-speaking children.

Keywords: *Speech sound development; screening test; bilingualism*

Introduction

Until recently, clinicians have faced many challenges when working with French-speaking children who have speech sound disorders. First, there has been a dearth of information on speech sound development, and second, there have been few assessment tools available to identify speech sound disorders. The last decade has seen an augmentation of studies on French speech sound development (Brosseau-Lapré & Rvachew, 2014; MacLeod, Sutton, Trudeau, & Thordardottir, 2011) and on the creation of tools to identify speech sound disorders in this population (Bérubé, Bernhardt, & Stemberger, 2015; MacLeod, Sutton, Sylvestre, Thordardottir, & Trudeau, 2014; Martinez Perez, Masson, & James, 2015; Rvachew et al., 2013). However, our understanding of French speech-sound development remains limited in comparison to other languages such as English and Spanish. Furthermore, most of the recently developed tests have been normed on Canadian French children. There are few studies on the speech sound development of European French children.

The current research is a normative study of a speech sound screening test for European French-speaking children. This test is called the *Dépistage Rapide Articulation et Phonologie (DRAP)*. The test developed out of the need by the second author to have a screening measure for testing the speech of European French-speaking children, aged three to seven years, attending school in the United States. Children were also tested in a Francophone context (Geneva, Switzerland) to determine whether their results would be similar to those obtained by French-speaking children abroad. In the *Introduction*, we provide an overview of what we know about speech sound development in French, a review of assessment procedures for French-speaking children and a justification as to why a new screening test is warranted.

Correspondence: Margaret Kehoe, Faculté de psychologie et des sciences de l'éducation, Université de Genève, 28, bd du Pont-d'Arve, Geneva 1205, Switzerland. Email: Margaret.Winkler-Kehoe@unige.ch

ISSN 1754-9507 print/ISSN 1754-9515 online
Published by Taylor & Francis
DOI: 10.1080/17549507.2020.1750699

We also examine the influence of bilingualism on speech sound development.

Speech sound and syllable structure development in French

In 2007, Rose and Wauquier-Gravelines published a chapter on French speech acquisition in the *International Guide to Speech Acquisition* (McLeod, 2007). One of their main observations was the lack of empirical data which existed on speech sound development in French. Since the publication of this chapter, authors have responded to this lack of information by conducting studies on consonant and syllable structure development in French.

MacLeod et al. (2011) conducted a normative study of consonant acquisition in 156 typically developing Québec French-speaking children, aged 1.8–4.5. They found that consonant acquisition could be described in terms of three periods: an early period in which children, aged less than 3.0, have acquired (i.e. 75% of children accurately produce the sound in three word positions) four consonants /t, m, n, z/; a middle period in which children, aged 3.0–4.5 have acquired an additional twelve consonants /p, b, d, k, g, ɲ, f, v, ʁ, l, w, ɥ/; and a late period in which children, aged 4.5 and older, are still in the process of acquiring four consonants /s, ʃ, ʒ, j/. Prior to this, Aicart-De Falco and Vion (1987) also observed that the greatest numbers of errors in French consonant acquisition concerned the quadrilateral formed by /s, z, ʃ, ʒ/. They tested a group of European French-speaking children, aged three to six years, and documented a high proportion of voicing and place of articulation (PoA) errors with the four consonants /s, z, ʃ, ʒ/.

In the area of syllable structure, attention has been given to the production of codas and clusters. Hilaire-Debove and Kehoe (2004) found that word-final codas were present a high percentage of the time in the productions of one- and two-syllable words by French-speaking children aged 1.8–2.8 (approximately 88% and 80% in one- and two-syllable words respectively). Similarly, Brosseau-Lapré and Rvachew (2014) found that word-final codas were preserved in the speech of phonologically delayed French-speaking pre-school children. In contrast, Rvachew et al. (2013) observed that word-internal codas were often subject to deletion even in children aged 6 to 7 years, particularly in multisyllabic words such as *garderobe* /gaʁdəʁob/ "wardrobe" and *tournevis* /tuʁnəvis/ "screwdriver".

As for clusters, MacLeod et al. (2011) found varied age of acquisition depending upon the segmental make-up of the cluster: /bl/ and /fl/ clusters were acquired first whereas /tr/ clusters were acquired later. Other researchers examining cluster development in French have observed that word-initial clusters are acquired before -final clusters (Demuth & Kehoe, 2006) and that consonant +/l/(C/l/) clusters are acquired earlier than consonant +/r/(C/r/) clusters (Kehoe, Hilaire-Debove, Demuth, & Lleó, 2008). Little attention has been given to the acquisition of /s/ C clusters in French since their presence is restricted in comparison to other languages, being only of the type /s/ + stop clusters.

In sum, French-speaking children have difficulties with alveolar and post-alveolar fricatives (e.g. /s, z, ʃ, ʒ/) in comparison to other consonants (Aicart-De Falco & Vion, 1987). They appear to acquire word-final codas easily but may have difficulties with word-internal codas (Rvachew et al., 2013). They make more errors on word-final in comparison to -initial clusters and C /r/ in comparison to C /l/ clusters (Demuth & Kehoe, 2006).

Tests of French speech sound development

Apart from the increasing number of studies on French speech sound development, recent years have seen the construction of different tests to assess the speech sound and syllable structure development of French-speaking children. Some of these are screening tests, which aim to identify children who are at risk for having a speech sound disorder (MacLeod et al., 2014; Rvachew et al., 2013); others are comprehensive tests aimed at identifying speech sound disorders which have a phonological or speech motor component (Bérubé et al., 2015; Martinez Perez et al., 2015). There also exist tests on the market which contain subtests targeting phonology and articulation such as the Nouvelles Epreuves pour l'Examen du Langage (N-EEL, Chevrie-Müller & Plaza, 2001) and EXALANG 3–6 years (Helloin & Thibault, 2006). Since the focus of this study is on developing a screening test, we describe two screening tests available for French-speaking children.

Rvachew et al. (2013) provide normative data on a screening test for Canadian French-speaking children – the *Test de Dépistage Francophone de Phonologie* (TDFP). The test consists of 30 words selected from a larger phonological test, *Test Francophone de Phonologie* (TFP; Paul & Rvachew, 2009). An important characteristic of the test design was that it approximates the syllable structure characteristics of spoken French by having a high proportion of multisyllabic words: 20% are one-syllable; 50% are two-syllable; and 30% are three- and four-syllable words. Rvachew et al. (2013) administered the test to 61 children, aged 6–7 years, 25 of which spoke a language other than French at home. There were no significant differences between age groups, nor were there any significant effects of language background on the children's speech. Children, aged 6–7 years, articulated 65-70% (21/30) of words correctly suggesting to Rvachew et al. (2013) that the word set was difficult enough to avoid ceiling effects.

MacLeod et al. (2014) report the results of a screening test for Canadian French-speaking children. Their test is based on words taken from the

Casse-Tête d'Evaluation de la Phonologie (CTEP, Auger, 1994). The authors chose 40 words from the CTEP which contain all the consonants of French in initial, medial, and final position and which are easily identifiable by two-year-old children. They collected normative data on 243 French-speaking monolingual children from Québec, aged between 1.8 and 4.5 years, who were divided into six age-groups of approximately 6 months. Results indicated a significant effect of age: percent consonants correct (PCC) increased from 56% for the age-range 1.8–1.11 to 90% for the age-range 4.0–4.5. There was no significant effect of gender on the test results.

One might wonder why the construction of an additional screening test is necessary given the existence of the *TDFP* and *CTEP (adapted)*. First, the *DRAP* contains 20 words in comparison to the *TDFP* and *CTEP* (adapted) which contain 30 and 40 words respectively. It can be administered in less than 5 minutes. Second, the 20 words have been carefully selected to target sounds and sounds structures which pose difficulty for children at the later stages of phonological acquisition. They include medial /r/ (*fourchette, escargot, arbre*), final /r/ (*dinosaure*, chaussure), initial clusters (*clé, fleche, plume, crayon, grenouille, trois*), final clusters (*table, ventre, arbre*), C/l/ clusters (*clé, fleche, plume, table*), C/r/ clusters (crayon, grenouille, trois, ventre, arbre), /s/C sequences (*ski, masque, escargot*), /s,z/ (*six, soleil, chaise, chaussure, dinosaure*) and /ʃ,ʒ/ (e.g. *chaise, chaussure, fleche, jaune, rouge*) (see Supplementary Appendix 1 for an English translation of words in the *DRAP*). Thus, the test does not sample all the sounds and syllable structures of French in different word positions but only those purported to be difficult for children at the latter stages of acquisition. Finally, it contains words familiar to European French-speaking children. The use of the *TDFP* and *CTEP* (adapted) would require modification for European French speaking children since certain words would be unfamiliar (e.g. *TDFP*: *beigne* "donut", *glissade* "slide"; *CTEP* (adapted): *beigne* "donut", *chandail* "shirt", *persil* "parsley", *yogourt* "yoghurt").

Influence of bilingualism on speech sound development

As several recent surveys indicate, speech-language therapists see an increasing number of children who speak more than one language. A great deal of attention has been given to the dangers of under- or over-diagnosing bilingual children with speech and language disorders. This danger is pronounced in the area of lexical and morpho-syntactic development where differences exist between typically developing monolingual and bilingual children (Hoff et al., 2012). In the area of speech sound development, findings are equivocal as to whether bilingual children present with greater numbers of speech sound errors than monolingual children (Hambly, Wren, McLeod, & Roulstone, 2013).

In the current study, we examine whether we observe increased numbers of errors in bilinguals in comparison to monolinguals. The children in Geneva are simultaneous bilinguals. They have started acquiring French from birth or early on while attending pre-school. In contrast, the children in San Francisco have not all been exposed to French from an early age. Many have only been exposed to French since starting school at the ages of 3–5 years. Their speech may be similar to child second language learners, containing greater numbers of errors due to L1 transfer than children who have been exposed to French from early on. In fact, studies on the speech sound development of children who are exposed to their second language after three years indicate that these children are not strongly disadvantaged in their speech sound production. Holm, Dodd, Stow, and Pert (1999) report findings on the phonological development of 35 Mirpuri-/Punjabi-/Urdu-speaking children who acquired English on entering school at about 4 years of age. The youngest age-group of children (aged 4.8–5.6) had PCC scores in English of 86% which were not very different from the PCC in their home languages of 84%, despite the fact that they had only been exposed to English for approximately one year. Similarly, Morrow, Goldstein, Gilhool, and Paradis (2014) documented high consonant accuracy in consecutive bilinguals who were not exposed to English until 3.3, some later than 5.0. At the first wave of testing, approximately 9 months following exposure to English, the average PCC was 89%. These results attest to good performance by young bilinguals on tests of speech sound production, even after limited exposure to the target language.

Current study

The aim of the study is to collect normative data on the *DRAP*. First, we investigate the influence of bilingualism, context, gender and age on the screening results. Second, we provide preliminary information on the test's validity by comparing the results on the *DRAP* with a more comprehensive assessment of phonology and articulation (i.e. subtests of the N-EEL). We also measure sample-specific internal consistency and conduct an item analysis. Third, we examine children's performance on phonological features in the *DRAP* and describe the errors patterns obtained on these features with the aim of revealing which features are more prone to errors than others.

Method

Participants

Participants included 101 French-speaking children, aged 2.11–6.10, tested in Geneva and 95 French-speaking children, aged 3.0–6.11, tested in San Francisco. The original sample tested in Geneva was

115 children but 14 children were excluded due to being too young ($n = 3$), incomplete test results ($n = 6$), unintelligible speech ($n = 2$), and insufficient information provided in the parental questionnaire ($n = 3$). The original sample tested in San Francisco was 97 children: two participants were excluded due to incomplete test results. Based on the parent questionnaire, all children had normal hearing, were in good health and were developing normally. We do not have individual socioeconomic status information on the children. The children in Geneva attended preschools and schools in middle-class areas whereas the children in San Francisco attended a private "Lycée" and were from middle to upper class families.

In the Genevan sample, bilingual status was determined based on a parent questionnaire in which the parents indicated whether their child spoke another language at least 30% of the time in addition to French. They were required to indicate which language the child spoke at home and with whom, and at what age the child had acquired French. Of the 63 Genevan bilinguals, 11 were trilingual speaking two languages at home. In the San Franciscan sample, all children were bilingual or trilingual. Supplementary Appendix 2 presents information on the number of participants according to bilingual status, gender, and age group, or in the case of the San Franciscan data, whether French was spoken by one parent, two parents or no parent. The age-groups were as follows: 3 (2.11–3.11), 4 (4.0–4.11), 5 (5.0–5.11), and 6 (6.0–6.11).

Stimuli

The *DRAP* was developed by the second author, Nathalie Niederberger. It consists of 20 words familiar to children, aged 3 to 6 years: 18 of the 20 words can be found in parent-report inventories such as the *l'Inventaire Français du Développement Communicatif* (IFDC) (16-30 months) (Kern & Gayraud, 2010) and the *Developpement du langage de production en français* (Bassano, Labrell, Champaud, Lemétayer, & Bonnet, 2005). The stimuli were carefully selected to target late-acquired sounds and syllable structures. Specific phonemes such as /ʃ/ were targeted in different word positions (e.g. /ʃosyʁ/, /ʃɛz/, /fuʁʃɛt/, /flɛʃ/). The words and images were selected after pilot testing, which made it possible to eliminate or modify items that were too easy (e.g. *maison* "house" and *feuille* "leaf") or originally ambiguous (e.g. *plume* – a drawing of a bird was added; ski – a snow-covered mountain was added; *six* - was represented by six dots, as on a dice). The list of 20 words and the sound and syllable structures they target are shown in Supplementary Appendix 3. The stimuli were represented as coloured drawings on plastic sheets with a maximum of six pictures per page.

Procedure

Children were tested individually in a quiet room either at the pre-school or school while interacting with one or two native French-speaking examiners. The children were required to name the picture following the question "Qu'est-ce que c'est?" What is that? or "Comment ça s'appelle?" What is that called? The procedure took less than 5 minutes to complete. The children's speech was recorded with a digital audio recorder (Marantz Tascam DR-2d) and a hand-held unidirectional condenser microphone in Geneva and Easi-Speak Pro digital hand-held recorder in San Francisco.

Validity

In order to determine whether children's results on the screening test were correlated with a more comprehensive phonological measure, we tested 11 children from the Genevan sample on the "phonology and articulation" and "phonology and memory" subtests of the N-EEL (Chevrie-Müller & Plaza, 2001), which is a French standardised battery of language measures. The 11 children (6 bilinguals, 7 male) were aged from 5.1 to 6.10 and were chosen randomly from the larger group. This group was made up of older children; hence, younger children would need to be tested in the future to ensure that validity results hold for younger children as well. The "phonology and articulation" subtest is made of 3 parts: 1a. frequent monosyllabic words (number of items is 22); 1b. frequent monosyllabic words containing clusters or fricatives (14 items); 1c. frequent multisyllabic words (25 items). The child receives two points if they correctly produce the word, one point if they make a mild phonological error (e.g. devoicing error) and 0 points if they make other errors. The "phonology and memory" subtest test has two parts. In "empan", the child needs to repeat 9 low frequency multisyllabic words; In "phonologie", the child needs to do the same but this time the words contain phonological structures such as clusters or codas (e.g. *cosmopolitisme*), The word is scored as correct or incorrect.

Data analyses

Using Phon, a software program specifically designed for the analysis of phonological data (Rose & MacWhinney, 2014), each child's wave file was segmented, and stimulus words were identified and transcribed. Two French-speaking graduate students, who had experience in phonetic transcription, performed the analyses. They transcribed each child's productions in broad phonetic transcription. The transcribed data were transferred to Excel and coded according to the phonological criteria under consideration.

Table I. Means and standard deviations of scores on the *Dépistage Rapide Articulation et Phonologie (DRAP)* according to bilingual status and context.

	Whole Word		PCC[a]		Vowel Errors		
	Mean	SD	Mean	SD	Mean	SD	n
Monolingual	16.50	3.33	94.31	6.10	0.45	0.65	38
Bilingual Geneva	15.63	3.92	91.97	8.87	0.75	1.15	63
Bilingual San Fran.	15.93	3.61	92.74	7.37	0.48	0.73	95

[a]Percent consonants correct.

Reliability

Twenty-one participants (11 in Geneva; 10 in San Francisco) were re-transcribed by a second transcriber using the Blind Transcription function of the Phon program. Point-to-point agreement in terms of consonant (90% in Geneva; 94% in San Francisco) and vowel transcription (98% in Geneva; 97% in San Francisco) was good.

Coding and statistical analyses

Data were analysed using mixed effects logistic regression. This allowed us to model production accuracy based on binomial data as well as examine the effects of predictor variables on the dependent variables, while taking into account random effects related to participant and stimulus item. The analyses were performed using R statistical software (R Development Core Team, 2015) and the lme4 package (Bates, Mächler, Bolker, & Walker, 2015) for mixed effects models. Comparisons were made using likelihood ratio tests (LRT) which yield a chi-square statistic. To determine differences between groups, we employed Tukey multiple comparisons.

In each model, bilingual status (monolingual vs. bilingual), context (Geneva vs. San Francisco), gender (male vs. female), and age (3, 4, 5, or 6 years) were entered as fixed effects. The random part of the model included random intercepts for participants and items. Random slopes on fixed effects were initially included but subsequently removed due to lack of convergence. The model was fitted using maximum likelihood estimation.

The analyses were conducted on three dependent variables: whole word score, PCC, and vowel errors. The first measure, whole word score is an efficient clinical yes-no measure: each individual word production was coded as either correct (1) or incorrect (0). Words were considered incorrect when they contained consonant or vowel errors. The second measure, PCC, is widely used in research and clinical practice. The response variable in our model for PCC was a proportion score: number of consonants correct/number of total consonants. For example, *escargot* /ɛskaʁgo/ produced as [ɛkago] was coded as 2/4 for PCC. Finally, we included a measure which considered the number of vowel errors children made. Each individual word production was coded as containing a vowel error or not. Examples of vowel errors include [ʃyʃyʁ] for *chaussure* /ʃosyʁ/ and [flaʃ] for *fleche* /flɛʃ/.

Table II. Means and standard deviations of scores on the *Dépistage Rapide Articulation et Phonologie (DRAP)* according to age.

	Whole Word		PCC[a]		Vowel Errors		
	Mean	SD	Mean	SD	Mean	SD	
Age 3	12.38	4.89	85.17	11.64	1.03	1.21	29
Age 4	15.35	3.35	91.50	7.35	0.63	0.72	43
Age 5	16.39	2.96	94.03	5.43	0.50	0.89	76
Age 6	17.92	2.17	96.61	3.71	0.31	0.62	48
All	15.94	3.66	92.80	7.68	0.56	0.88	196

[a]Percent consonants correct.

Result

Normative results on whole word scores, PCC and vowel errors

First, we examined the influence of bilingualism, context, gender and age on the screening results. Table I presents the means and standard deviations for the three main dependent variables of the *DRAP* according to whether the children were monolingual or bilingual, and whether they lived in Geneva or San Francisco. The average whole word score was 16 (out of a total of 20), the PCC score was approximately 93% and the number of vowel errors per child was less than 1.0 meaning that many children did not make vowel errors. Table II presents the means and standard deviations for the three main measures of the *DRAP* according to age. Whole word scores ranged from 12.0 to 18.0, PCC scores ranged from 85 to 97%, and the number of vowel errors ranged from 1.0 to less than 1.0 (i.e. 0.31) when comparing the youngest and oldest groups of children.

We ran mixed models entering all predictor variables as fixed effects and using random intercepts for participants and items. One predictor was found to be significant in the model based on whole word scores: age ($\chi^2(3) = 46.59$, $p < 0.001$). Older children obtained higher scores than younger children. Tukey multiple comparisons with Bonferroni corrections indicated that there were significant age effects between three- and five-year-olds ($z = 4.75$, $p < 0.001$), three and six-year-olds ($z = 7.00$, $p < 0.001$), four and six-year-olds ($z = 4.52$, $p < 0.001$) and five- and six-year-olds ($z = 3.34$, $p = 0.004$). Importantly, there were no differences between three- and four-, and four- and five-year olds.

Two predictors were found to be significant in the model based on PCC: bilingual status ($\beta = 0.55$, $\chi^2(1) = 5.29$, $p = 0.02$) and age ($\chi^2(3) = 44.17$, $p <$ 0.001). Greater consonant precision was associated

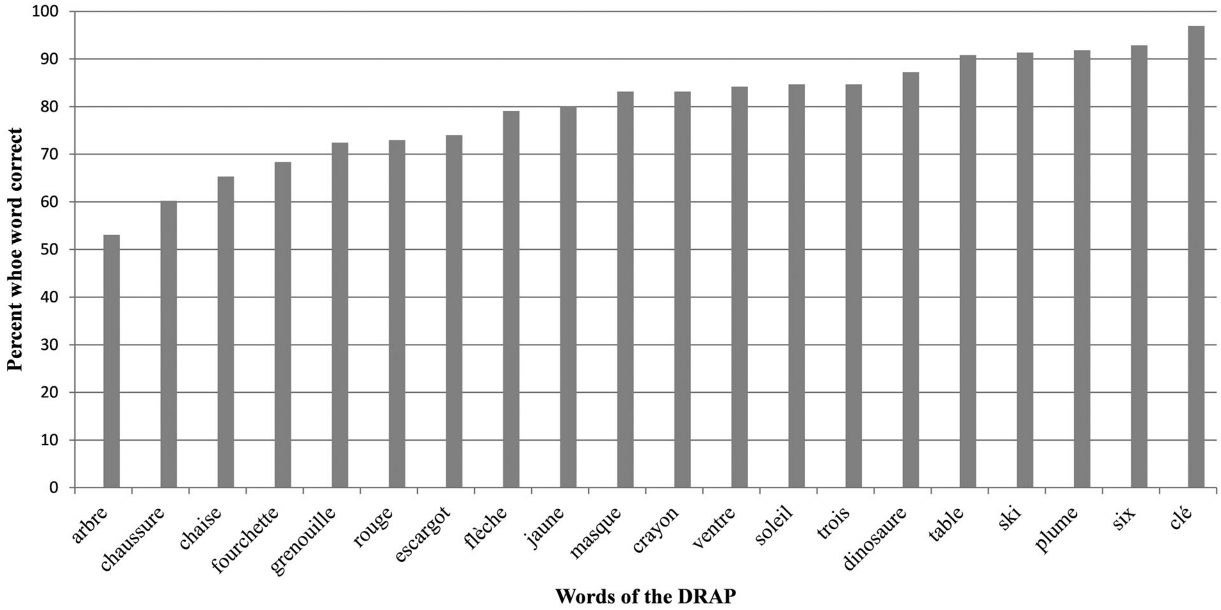

Figure 1. Percentage of children obtaining a correct whole word score on individual words of the *Dépistage Rapide Articulation et Phonologie (DRAP)*.

with being monolingual and being older. Tukey multiple comparisons with Bonferroni corrections indicated that there were significant age effects between three- and five-year-olds ($z = 4.69$, $p < 0.001$), three and six-year-olds ($z = 6.82$, $p < 0.001$), four and six-year-olds ($z = 4.58$, $p < 0.001$) and five- and six-year-olds ($z = 3.16$, $p = 0.008$). Once again, there were no differences between three- and four-, and four- and five-year olds.

One predictor was found to be significant in the model based on number of vowel errors: age ($\chi^2(3) = 12.50$, $p = 0.006$). Older children made fewer vowel errors than younger children. Tukey multiple comparisons with Bonferroni corrections indicated that three-year-olds made more vowel errors than six-year-olds ($z = -3.41$, $p = 0.003$). Supplementary Appendix 4 summarises the significant age effects for the three measures based on the Tukey multiple comparisons.

Validity, internal consistency and item analysis

Second, we examined validity and reliability aspects of the *DRAP*. We measured Pearson Product Moment correlations between scores on the *DRAP* and scores on the N-EEL subtests "phonology and articulation" and "phonology and memory" for the subgroup of 11 children. We obtained significant positive correlations between results on the N-EEL subtest "phonology and articulation" (monosyllabic and multisyllabic words combined) and the whole word measure (r(9) = 0.71, $p = 0.01$) and between results on the N-EEL and PCC scores (r(9) = 0.80, $p = 0.003$). There were no significant correlations between scores on the *DRAP* and scores on the subtest "phonology and memory" either for "empan" or "phonologie".

We examined the reliability of the *DRAP* using Cronbach's alpha, which is a measure of internal consistency or how closely related the test items are as a group. The internal consistency based on the entire sample ($n = 196$) was = 0.82, which is within the acceptable range (Tavakol & Dennick, 2011).

We also conducted an item analysis in which we calculated the percentage of children obtaining a correct whole word score on individual items of the *DRAP*. These results are presented in Figure 1. As can be seen, percentages ranged from 53% for the word arbre through to 97% for the word clé. Words which proved the most difficult for children were words with post-alveolar fricatives (*chaussure, chaise, fourchette, rouge*), medial /r/ codas (*arbre, fourchette*) and C/r/ clusters (*arbre, grenouille*). Words which posed few difficulties were short (*clé, six, ski*), contained C/l/ clusters (*clé, plume, table*) and alveolar fricatives (*six, ski*). Importantly, there were no words on the *DRAP* in which all children scored at 0 or 100% suggesting that there were no words that were too easy or too difficult.

Findings on phonological features of the DRAP

Third, we examined children's production accuracy on phonological features sampled in the *DRAP*. Table III presents the means and standard deviations for medial /r/ codas, final /r/ codas, initial clusters, final clusters, C/l/ clusters, C/r/ clusters, /s/C sequences, alveolar fricatives, and post-alveolar fricatives for the four different age groups of children as well as for the entire group of children. As can be seen, there were few developmental effects for final /r/ and C/l/

clusters. Children as young as three years obtained scores in excess of 90% for these features. Children experienced the most difficulty with final clusters and post-alveolar fricatives. At age 6, children still obtained scores below 90% (final clusters: 83.92%; post-alveolar fricatives: 88.54%) for these features. At age 3 years, children also experienced difficulty with medial /r/ and C/r/ clusters (i.e. scores below 70%); however, by age 6 years, scores were around 90% or greater.

The major error pattern and the most frequent minor error patterns for the different phonological features are presented in Supplementary Appendix 5. The main error pattern of medial and final /r/ was deletion of /r/. Similarly, the predominant pattern for C/l/ and C/r/ clusters was deletion of the liquid element. In the case of /s/C sequences, it was deletion of /s/. Final clusters were more prone to deletion of the liquid element than initial clusters (i.e. 71 vs. 46%).

Table III. Means and standard deviations of percentage scores on the *Dépistage Rapide Articulation et Phonologie (DRAP)* according to phonological features and age.

	Medial /r/		Final /r/		
	Mean	SD	Mean	SD	n
Age 3	64.37	35.56	100.00	0	29
Age 4	76.74	34.53	97.67	10.65	43
Age 5	84.65	24.00	98.03	12.76	76
Age 6	91.67	22.28	100.00	0	48
All	81.63	29.28	98.72	9.39	196

	Initial Clusters		Final Clusters		
	Mean	SD	Mean	SD	n
Age 3	81.61	18.01	71.26	26.31	29
Age 4	88.37	15.66	87.60	19.28	43
Age 5	93.20	13.38	83.33	24.04	76
Age 6	97.57	8.41	88.19	18.82	48
All	91.50	14.57	83.67	22.77	196

	C/l/		C/r/		/s/C		
	Mean	SD	Mean	SD	Mean	SD	
Age 3	92.24	13.53	67.59	26.41	72.41	39.91	29
Age 4	95.35	13.64	82.33	22.77	86.82	26.37	43
Age 5	96.38	11.33	84.74	21.38	91.67	19.72	76
Age 6	97.92	6.98	90.42	15.97	97.92	10.67	48
All	95.92	11.43	83.06	22.34	89.29	24.89	196

	Alveolar Fricatives		Postalveolar Fricatives		
	Mean	SD	Mean	SD	n
Age 3	83.33	23.99	59.77	37.67	29
Age 4	90.31	15.53	75.19	31.36	43
Age 5	91.45	15.03	86.62	22.11	76
Age 6	95.14	10.29	88.54	19.19	48
All	91.27	16.13	80.61	28.16	196

Realisation of a cluster with substitution of one of the members of the cluster was a frequent minor error pattern for C/l/, C/r/ and /s/C sequences. Interestingly, substitutions in which /l/ or /r/ became glides were only seen in the San Franciscan data (e.g. [fwɛʃ] for *flèche* /flɛʃ/; [kwɛ'jɔ̃] for *crayon* /kʁɛ'jɔ̃/).

In terms of error patterns for alveolar and post-alveolar fricatives, the most common pattern was a PoA error in which alveolar fricatives became post-alveolar (e.g. [ʃɛʒ] for *chaise* /ʃɛz/) and the reverse, post-alveolar fricatives became alveolar (e.g. [flɛs] for *flèche* /flɛʃ/). However, some of the PoA substitutions also included distortions, in which it was difficult to transcribe the exact sound (an acoustic quality between /s/ and /ʃ/ or /z/ and /ʒ/). The other main error patterns were voicing and manner substitutions. The voicing errors for alveolar and post-alveolar fricatives were only present in the Genevan data (e.g. [ʃɛs] for *chaise* /ʃɛz/; [ʁuʃ] for *rouge* /ʁuʒ/).

Discussion

The current study presents normative findings on a speech sound screening test called the *DRAP*, which was administered to monolingual and bilingual children in Geneva, and bilingual French-speaking children in San Francisco. Our analyses show effects of age on all measures and a small effect of bilingualism on one measure. Gender and context did not influence any test measure, although the latter influenced the types of errors children made.

Normative results on the DRAP

The *DRAP* is a screening test designed to identify children who are at risk of having a speech sound disorder. These children should obtain poorer results on the *DRAP* than the ones obtained by the norming population presented in this study (see Niederberger, Kehoe & Bouchut (submitted), for findings on the *DRAP* with a clinical population). Taking into consideration the means and standard deviations presented in Table II, normative cut-offs can be derived based on 1, 1.5, and 2 standard deviations below the mean for whole word and PCC scores (see Table IV). A whole-word score which is 1 standard deviation below the mean would identify 20% of the sample as requiring additional testing, a value which is higher than the incidence of speech sound disorders in the general population (approximately 6 to 7% according to Broomfield & Dodd, 2004). Thus, we could also

Table IV. Recommended cut-off values for identifying children, who require additional phonological testing, based on 1, 1.5, and 2 standard deviations below the mean.

	Whole Word Score				PCC[a] Score			
	Mean	1 SD	1.5 SD	2 SD	Mean	1 SD	1.5 SD	2 SD
3	12	8	5	3	85	74	69	63
4	15	12	10	9	92	84	80	77
5	16	13	12	11	94	89	86	83
6	18	16	15	14	97	93	91	89

[a]Percent consonants correct.

consider the use of cut-offs such as 1.5 or 2 standard deviations below the mean. Such criteria would identify 13% and 6% respectively as being at risk. In the case of vowels, more than three vowel errors at age 3 and more than two vowel errors at ages 4–6 would represent results outside the normative data and, thus, be a cause of concern. Given that the *DRAP* only contains 20 items, a more conservative cut-off (i.e. 1 standard deviation) might be warranted in order to avoid missing children who have speech-sound disorders.

Our statistical analyses indicated that bilingualism had a significant effect on PCC scores but no effect on whole word scores and vowel errors. Given that the percent difference in PCC scores was small (i.e. 2%) between monolingual and bilingual children, we do not advocate the use of separate norms for monolingual and bilingual children. Our results are consistent with those of Rvachew et al. (2013), who did not find many differences between monolinguals and bilinguals on the *TDFP*. As Hambly et al.'s (2013) review article suggests, there is no clear evidence that bilingualism leads to a general pattern of slower or faster acquisition of speech sounds. Certain studies, however, have reported accelerated or delayed development of speech sound acquisition in bilinguals. These studies have focussed on language structures or constellations in which cross-linguistic interaction was predicted (Fabiano-Smith & Goldstein, 2010; Keffala, Barlow, & Rose, 2018). It is possible that if we had focussed on certain target structures (e.g. codas) and bilingual populations (e.g. French-Portuguese), larger differences between monolinguals and bilinguals may have been observed. Our analyses did not show any influence of context on phonological production accuracy. That is, bilinguals acquiring French in Geneva, where French is spoken in the environment, had similar results to bilinguals acquiring French in San Francisco, where English is spoken in the environment and French is spoken at school. Gender also did not prove significant in any of our analyses meaning that no additional norms are needed for boys versus girls. Similarly, MacLeod et al. (2014) reported no significant gender effect on the *CTEP* (adapted). In general, gender does not have a strong influence on phonological acquisition, although there are some studies which find slightly better performance for girls than boys in speech sound mastery (Smit, Hand, Freilinger, Bernthal, & Bird, 1990).

The main statistically significant effect in our analyses was age. Developmental effects were present on the three measures: whole word scores, PCC, and number of vowel errors. MacLeod et al. (2014) reported significant age effects on their screening measure *CTEP* (adapted) for the age range 1.10 through to 4.5, when divided into six five-month intervals. Differences were not present amongst all age levels but were present between the youngest and oldest age ranges. Rvachew et al. (2013) did not obtain significant age effects but they included only two age groups of children, that is, 6- and 7-years. Our study included a wider age range of children and due to the small numbers of participants at certain age ranges, we divided the sample into yearly rather than five- or six-month intervals. In the case of whole word scores and PCC, there were several significant age contrasts. Overall, our test appears to be sensitive to age differences, which we interpret as a positive feature.

Our test was not sensitive to all age differences, however, rendering the task of generating precise normative data difficult. The cut-off criteria presented in Table IV should be considered approximate given that not all ages differed significantly from each other on all measures of the *DRAP*. Three- and four-, and four- and five-year-olds, for example, did not differ from each other on whole word scores and PCCs, meaning that more generous cut-off criteria would be needed for these groups. Further testing of the *DRAP* with a larger normative sample would be necessary in order to obtain more accurate cut-off values per age.

Validity, internal consistency and item analysis

We documented positive significant correlations (coefficients from 0.71 to 0.80) between scores on the *DRAP* and scores on a more thorough test of speech sound development (i.e. subtest "phonology and articulation" of N-EEL), indicating that the *DRAP* was estimating what it purports to test. We did not document any correlations between the *DRAP* and a test of phonological memory, consistent with the fact that the test was not designed to test this aspect of phonological knowledge. Other analyses indicated that the test had acceptable internal consistency and that there were no items in the test that were too easy or difficult.

Phonological features of the DRAP

The *DRAP* was designed to test children on late-acquired phonological features so we also examined children's production accuracy on these individual features. Our study showed that children realized /r/ in word-final position, although they frequently deleted it in word-internal position (e.g. [ɛska'go] for *escargot* /ɛskaʁ'go/ but [dino'zɔʁ] for *dinosaure* /dino'zɔʁ/). Other researchers have noted the high preservation rates of word-final consonants (Hilaire-Debove & Kehoe, 2004) and the vulnerability of word-internal codas to deletion in French (Brosseau-Lapré & Rvachew, 2014; Rose, 2000; Rvachew et al., 2013). In terms of clinical implications, our results suggest that omission of final /r/ in the words of the *DRAP* should be a cause of concern in children aged three-years or more, whereas omission of /r/ in word-internal position occurs through to five years or later.

Our findings support studies by Demuth and Kehoe (2006) and Demuth and McCullough (2009) which show superior performance for word-initial as compared to -final clusters in French. Children obtained percent scores of 10% higher for word-initial in comparison to -final clusters at almost all age ranges. Our study also showed that children produced C/l/ clusters correctly whereas they often made errors with C/r/ clusters; /s/C sequences were situated in between the two extremes of difficulty. In terms of clinical implications, our results suggest that errors with C/l/ clusters in the words of the *DRAP* at the age of three years may already be a cause of concern but errors with /s/C sequences and C/r/ clusters may continue through to four and five years.

Our results confirmed those of Aicart-De Falco and Vion (1987) which noted that the greatest numbers of consonantal errors in French-speaking children, aged 3 to 6 years, concern the alveolar and post-alveolar fricatives. The most prominent error patterns were substitutions of PoA, in which alveolars became post-alveolars, and the reverse. In addition, Aicart-De Falco and Vion (1987) reported voicing errors, in which voiced fricatives became voiceless. Like Aicart-De Falco and Vion (1987), we found that PoA errors were the most frequent (approximately 60-76% of errors patterns) followed by voicing errors (15–23%).

MacLeod et al. (2011) found /ʃ, ʒ/ to be acquired (i.e. produced accurately by 75% of children) by 3.6–4.0, which is consistent with the findings of this study, in which 4-year-old children produced target post-alveolar fricatives with 76% accuracy. Rvachew et al. (2013) found /ʃ/ to be mastered (i.e. produced accurately by 90% of children) by children aged 7 years, but not by children aged 6 years. These findings are also not dissimilar to ours in that we found that post-alveolar fricatives were not yet produced with 90% accuracy at age 6. In terms of clinical implications, our results suggest that errors on target post-alveolar fricatives in the words of the *DRAP* occur through to the age of six years and cannot be used as a clinical marker. In contrast, errors with alveolar fricatives should not be present after 5 years.

In our analyses of error patterns, we observed small differences between the two language contexts. In the San Francisco data, there were occasions in which children substituted glides for liquids. Liquid gliding is a common phonological process whereby English-speaking children simplify clusters (Smit, 1993). The use of this process by the San Franciscan children appears to be an example of cross-linguistic transfer (of error patterns) from English onto French since such forms were not attested in the Genevan data. The Genevan children made voicing errors in which voiced alveolar and post-alveolar fricatives became voiceless. Such errors have been reported for French-speaking children in France (Aicart-De Falco & Vion, 1987), although their prevalence varies between regions (Houdebine, 1985). They were not present in the San Franciscan data.

Conclusion

This study reports data on a quick-to-administer screening test to detect speech sound disorders in French-speaking children, aged approximately 3–6 years. Our preliminary results suggest that the test is sensitive to age differences but not to context or gender. Differences between monolinguals and bilinguals were present for one phonological measure, PCC, but the magnitude of the difference was small (i.e. 2%). A detailed analysis of the phonological features of words in the test reveals that certain features were more prone to errors than others, furnishing information that may prove useful when the test is employed in a clinical setting. We are currently collecting data on the *DRAP* with a clinical population in order to determine whether it effectively screens out children with speech sound disorders.

Acknowledgements

We would like to thank Kopika Kannathasan, Audrey Burkhardt, Constance Terrail and Tanya Bella Bancaleiro for their help in testing the children and Chloe Girardier for her help in testing the children and in transcribing the data. In addition, we would like to thank the personnel and teachers at «EVE Espèces de vie enfantine du secteur université» and in the Genevan public-school system for their collaboration in the recruitment of children.

Declaration of interest

No potential conflict of interest was reported by the author(s).

Supplementary material

Supplemental data for this article can be accessed at https://doi.org/10.1080/17549507.2020.1750699.

References

Aicart-De Falco, S., & Vion, M. (1987). La mise en place du système phonologique du français chez les enfants entre 3 et 6 ans: une étude de la production. *Cahiers de Psychologie Cognitive*, *7*, 247–266.

Auger, D. (1994). *Casse-Tête d'évaluation de la phonologie*. Montréal: Published by author.

Bassano, D., Labrell, F., Champaud, C., Lemétayer, F., & Bonnet, P. (2005). Le DLPF: Un nouvel outil pour l'évaluation du développement du langage de production en français. *Enfance*, *57*, 171–208. doi:10.3917/enf.572.0171

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 48. doi: doi:10.18637/jss.v067.i01

Bérubé, D., Bernhardt, B.M., & Stemberger, J.P. (2015). A test of French phonology: Construction and use. *Canadian Journal of Speech-Language Pathology and Audiology*, *39*, 62–101.

Broomfield, J., & Dodd, B. (2004). The nature of referred subtypes of primary speech disability. *Child Language Teaching and Therapy*, *20*, 135–151. doi: doi:10.1191/0265659004ct267oa

Brosseau-Lapré, F., & Rvachew, S. (2014). Cross-linguistic comparison of speech errors produced by English- and French-speaking preschool-age children with developmental phonological disorders. *International Journal of Speech-Language Pathology*, *16*, 98–108. doi: doi:10.3109/17549507.2013.794863

Chevrie-Müller, C., & Plaza, M. (2001). Nouvelles Epreuves pour l'Examen du Langage (N-EEL). London: Pearson.

Demuth, K., & Kehoe, M. (2006). The acquisition of word-final clusters in French. *Journal of Catalan Linguistics*, *5*, 59–81. doi: doi:10.5565/rev/catjl.79

Demuth, K., & McCullough, E. (2009). The longitudinal development of clusters in French. *Journal of Child Language*, *36*, 425–448. doi: doi:10.1017/S0305000908008994

Fabiano-Smith, L., & Goldstein, B. (2010). Phonological acquisition in bilingual Spanish-English speaking children. *Journal of Speech, Language, and Hearing Research*, *53*, 160–119. doi: doi:10.1044/1092-4388(2009/07-0064

Hambly, H., Wren, Y., McLeod, S., & Roulstone, S. (2013). The influence of bilingualism on speech production: A systematic review. *International Journal of Language and Communication Disorders*, *48*, 1–24. doi: doi:10.1111/j.1460-6984.2012.00178.x

Helloin, M.-C., & Thibault, M.-P. (2006). EXALANG 3-6 ans. Campbell: Happy Neuron Pro.

Hilaire-Debove, G., & Kehoe, M. (2004). Acquisition des consonnes finales (codas) chez les enfants francophones: des universaux aux spécificités de la langue maternelle. Actes de XXVèmes Journées d'Etude sur la Parole, FES: Fez, Maroc.

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*, 1–27. doi: doi:10.1017/S0305000910000759

Holm, A., Dodd, B., Stow, C., & Pert, S. (1999). Identification and differential diagnosis of phonological disorder in bilingual children. *Language Testing*, *16*, 271–292. doi: doi:10.1191/026553299674008527

Houdebine, A.-M. (1985). *La phonologie de l'enfant français de six ans: variétés régionales*. Hamburg: Helmut Buske.

Keffala, B., Barlow, J., & Rose, S. (2018). Interaction in Spanish-English bilingual's acquisition of syllable structure. *International Journal of Bilingualism*, *22*, 16–37. doi: doi:10.1177/1367006916644687

Kehoe, M., Hilaire-Debove, G., Demuth, K., & Lleó, C. (2008). The structure of branching onsets and rising diphthongs: Evidence from the acquisition of French and Spanish. *Language Acquisition*, *15*, 5–57. doi: doi:10.1080/10489220701774229

Kern, S., & Gayraud, F. (2010). *L'inventaire français du développement communicatif*. Grenoble: Editions La Cigale.

MacLeod, A.A., Sutton, A., Sylvestre, A., Thordardottir, E., & Trudeau, N. (2014). Outil de dépistage des troubles du développement des sons de la parole: Bases théoriques et données préliminaires. *Canadian Journal of Speech-Language Pathology and Audiology*, *38*, 40–57.

MacLeod, A.N., Sutton, A., Trudeau, N., & Thordardottir, E. (2011). The acquisition of consonants in Québécois French: A cross-sectional study of pre-school aged children. *International Journal of Speech-Language Pathology*, *13*, 93–109. doi: doi:10.3109/17549507.2011.487543

Martinez Perez, T., Masson, P., & James, E.G. (2015). Construction et normalisation d'épreuves aidant au diagnostic de la dyspraxie verbale en Belgique francophone. *Rééducation Orthophonique*, *263*, 149–164.

McLeod, S. (Ed.) (2007). *International guide to speech acquisition*. Clifton Park, NY: Thomson Delmar Learning.

Morrow, A., Goldstein, B.A., Gilhool, A., & Paradis, J. (2014). Phonological skills in English language learners. *Language, Speech, and Hearing Services in Schools*, *45*, 26–39. doi: doi:10.1044/2013_LSHSS-13-0009

Niederberger, N., Kehoe, M. & Bouchut, A-L. (submitted). DRAP : un test de Dépistage Rapide de l'Articulation et de la Parole pour les enfants de 3 à 6 ans.

Paul, M., & Rvachew, S. (2009). *Test Francophone de Phonologie (Unpublished Test)*. McGill University, Montreal, Quebec.

R Development Core Team. (2015). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Rose, Y., & MacWhinney, B. (2014). The PhonBank initiative. In J. Durand J, U. Gut, & G. Kristoffersen (Eds.), *The Oxford handbook of corpus phonology* (pp. 380–401). Oxford: Oxford University Press.

Rose, Y. & Wauquier-Gravelines, (2007). French speech acquisition. In S. Mc Leod. (Ed.), *International guide to speech acquisition* (pp. 364–384). Clifton Park, NY: Thomson Delmar Learning.

Rose, Y. (2000). *Headedness and prosodic licensing in the L1 acquisition of phonology*. Unpublished doctoral dissertation, McGill University.

Rvachew, S., Marquis, A., Brosseau-Lapré, F., Paul, M., Royle, P., & Gonnerman, L.M. (2013). Speech articulation performance of francophone children in the early school years: Norming of the Test de Dépistage Francophone de Phonologie. *Clinical Linguistics & Phonetics*, *27*, 950–968. doi: doi:10.3109/02699206.2013.830149

Smit, A. (1993). Phonologic error distributions in the Iowa-Nebraska articulation norms project: Word-initial consonant clusters. *Journal of Speech and Hearing Research*, *36*, 931–947. doi: doi:10.1044/jshr.3605.931

Smit, A., Hand, L., Freilinger, J., Bernthal, J., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders*, *55*, 779–798. doi: doi:10.1044/jshd.5504.779

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. doi: doi:10.5116/ijme.4dfb.8dfd